



Paper to be presented at the DRUID Academy Conference 2016 in Bordeaux,
France on January 13-15, 2016

Varieties of the Concept of National Innovation Systems: Towards Complementarity

Marija Rakas
Aalborg University
Business and Management
rakas@business.aau.dk

Daniel S. Hain
University of Aalborg
Department of Business and Management
dsh@business.aau.dk

Abstract

The main idea of the paper is to map the advances in the studies on National Innovation Systems (NISs) or National Systems of Innovation (NSIs) since its emergence to its more recent developments. The first aim of the paper is to come to a more comprehensive understanding of the initial versions of the NIS analytical framework and its development over time in the context of a variety in conceptualisations and operationalization of the concept that exist in the literature on NISs. The second aim of the paper is to provide an insight into whether (or not) the contributions in the NIS branch of literature have led over time to a more comprehensive understanding and a coherent framework suitable for conducting a systematic analysis of NISs, and to what extent these developments might be attributed to different reasons such as availability of data, methodological tools and/or gaps in theoretical understandings given the complex and multidimensional nature of the phenomena under study. A greater comprehensiveness in understanding and systematic operationalization of the NIS concept also has a potential to contribute to the development of comparative studies of different NISs. In order to map the advances of the studies on NISs or NSIs, originating from the core contributors Christopher Freeman, Bengt-\r{A}ke Lundvall, and Richard Nelson to recent developments, we manually categorized the most relevant publications in this branch of literature along the dimensions (i.) conceptual and analytical focus, (ii.) theoretical aspects, and (iii.) research methodology applied. We further relate the observed developments to the organization of knowledge production in terms of the relational structure among the contributions based on a bibliometric network analysis. In addition, we deployed techniques from the field of natural language processing to enrich the study with the thematic information on the context of the contributions.

Varieties of the Concept of National Innovation Systems: Towards Complementarity *

Marija Rakas and Daniel S. Hain

Aalborg University, Department of Business and Management, IKE / DRUID, Denmark

Abstract: The main idea of the paper is to map the advances in the studies on National Innovation Systems (NISs) or National Systems of Innovation (NSIs) since its emergence to its more recent developments. The first aim of the paper is to come to a more comprehensive understanding of the initial versions of the NIS analytical framework and its development over time in the context of a variety in conceptualisations and operationalization of the concept that exist in the literature on NISs. The second aim of the paper is to provide an insight into whether (or not) the contributions in the NIS branch of literature have led over time to a more comprehensive understanding and a coherent framework suitable for conducting a systematic analysis of NISs, and to what extent these developments might be attributed to different reasons such as availability of data, methodological tools and/or gaps in theoretical understandings given the complex and multidimensional nature of the phenomena under study. A greater comprehensiveness in understanding and systematic operationalization of the NIS concept also has a potential to contribute to the development of comparative studies of different NISs.

In order to map the advances of the studies on NISs or NSIs, originating from the core contributors Christopher Freeman, Bengt-Åke Lundvall, and Richard Nelson to recent developments, we manually categorized the most relevant publications in this branch of literature along the dimensions (i.) conceptual and analytical focus, (ii.) theoretical aspects, and (iii.) research methodology applied. We further relate the observed developments to the organization of knowledge production in terms of the relational structure among the contributions based on a bibliometric network analysis. In addition, we deployed techniques from the field of natural language processing to enrich the study with the thematic information on the context of the contributions.

*Paper prepared for the 2016 DRUID Academy, University of Bordeaux

1 Introduction

The NIS concept has diffused rapidly since its emergence in the mid-1980s among both researchers and policy makers. However, researchers and policy makers attach very different meanings to the NIS concept. A variation in the main focus and approach in studying and conceptualizing National Innovation Systems (NISs) also exists among the originators of the concept, between the one developed by Lundvall (1992) and the one originating from Nelson (1993) and Freeman (1987). In the following years several different conceptual frameworks have been developed using the term “systems of innovation”. The coexisting conceptual frameworks include regional (Cooke, 2001; Malmberg and Maskell, 2002), sectoral (Breschi and Malerba, 1997; Malerba, 2002, 2005), technological (Bergek et al., 2008; Carlsson and Stankiewicz, 1991; Hekkert et al., 2007), business (Whitley, 2000) and social systems of innovation and production (Amable, 2000).

Despite the coexistence of the constructs in the literature there are still difficulties in understanding what each construct means, and how they are related to each other. The conceptual frameworks have been developed in different contexts, the theoretical underpinnings are not always explicitly stated or are drawing upon different disciplinary fields, and the encompassing concepts are not always well developed or are not defined in the same manner (Edquist and Hommen, 2008). Therefore, there seem to be no consensus in terms of conceptualization of the term. Related, it is not clear how research methodologies used for empirical studies have shaped the development of the NIS concept over time. While low consensus in respect to the conceptual framework elements does not necessarily represent a problem, especially when the research field is still evolving, it might limit the systemic advances of knowledge (Pfeffer, 1993).

The main aim of the paper is to come to a more comprehensive understanding of the nuances and varieties of conceptualisations of NIS as an analytical framework, and their development over time. The second aim of the paper is to provide an

insight to what extent the contributions in the NIS literature have over time led to a more coherent and comprehensive framework for conducting a systematic analysis of NISs, and provided empirical evidence for the theoretical claims made. Such contributions can to some extent be made by studying different aspects on NISs in isolation. Yet, in the spirit of the originators, a comprehensive analysis of the framework optimally includes the joint effects of – and interaction between – organizations, institutions, and the structure in which they relate to each others. Consequently, a development towards more inclusive analysis in respect to its different building blocks within the NIS framework could reflect the growing sophistication of the research community towards its operationalization. In contrast, an ongoing trend to study different aspects in isolation might be attributed to reasons such as availability of data, methodological tools and/or gaps in theoretical understandings given the complex and multidimensional nature of the phenomena under study.

Further reasons for how a theoretical and analytical framework develops can be found in the underlying structure of researcher communities working on it. First, the behavior of individuals, reflected in their preference towards a certain theoretical and analytical focus, is strongly affected by the way how they relate to and interact with larger collectives (Burt, 1992; Granovetter, 1973; Merton, 1957; Simmel, 1955). Second, the way how specialized research communities interact between each others might manifest in an emerging common consensus on theoretical understanding, and an exchange methodological practice, enabling them to manage the increasing complexity of jointly operationalizing different theoretical and analytical perspectives (Jones et al., 2008; Singh, 2005; Wagner and Leydesdorff, 2005)

Consequently, studying the organization of knowledge production within an scientific field or branch – as reflected by the network structure of research communities – provides valuable explanation for its development. In this paper, we analyse the structure of the NIS literature and its contributors revealed by related academic publications on the topic. Therefore, we use a set of different methods to gain insights on the development of how the framework is conceptualized and operationalized,

and why that is the case. First, to map academic community structure within NIS research, we use different methods such as co-citation analysis and bibliographic coupling. Within these communities we analyze their development over time in terms of produced publications and received citations, and identify main contributions and most influential authors.

To identify specialization pattern among these communities, we in a next step enhance the obtained structural publication information with first indicators on their content. Therefore, we manually categorize publications according to their single or multiple theoretical and analytical focus. Thereby, we are able to investigate if NIS research develops isolated in intellectual tribes, or advances towards integrative operationalization. We complement our mapping of NIS research with an linguistic analysis of publication abstracts by deploying methods from the field of natural language processing. By identifying common topics across publications and communities. The way how science progresses is to a large extent reflected in the use of language, which can indicate general research trends, but also the emergence of common consensus or dispute (Kuhn, 1970). Such topics provide us additional insights on the context of NIS research within and across communities, and its development over time.

The remainder of the paper is structured in the following manner. Section 2 introduces briefly the conceptual framework of NISs as developed by Freeman, Lundvall, and Nelson. We emphasise the main differences and similarities between these original approaches to studying NISs. This section also serves as a basis for the development of the classification scheme that will be used for characterizing the relevant publications in this branch of literature. More specifically, the first dimension of the scheme refers to a conceptual and analytical focus. The scheme will be used for manual classification of the most relevant publications on NISs along the conceptual and analytical focus, theoretical and research methodology dimension in order to identify the nature of the progress in this branch of literature.

Section 3 describes data for mapping the advances in the NIS branch of literature, and introduces the methods used in our analysis, namely co-citation analysis, bibliographic coupling, co-citation analysis and LDA topic modeling. Section 4 presents the analysis of the overall developments in the field and a more detailed analysis of some specific patterns that emerge.

2 Comparison of the Original Versions of the NIS Concept

The initial conceptualizations of the term “National Systems of Innovation” or “National Innovation Systems” in the field of innovation studies is usually attributed to the three founding fathers that are Christopher Freeman, Bengt-Åke Lundvall, and Richard Nelson. Based on former bibliometric research on the development of NIS literature (Fagerberg and Sapprasert, 2011; Teixeira, 2013), and confirmed by the results of this study, the core contributions on the subject are *“Technology, policy and economic performance: lessons from Japan”* by Freeman (1987), *“National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning”* edited by Lundvall (1992), and *“National Innovation Systems: A Comparative Analysis”* edited by Nelson (1993).

The common elements associated with this new branch of literature include a holistic approach to studying innovation, a focus on the interdependencies between the actors, organizations and institutions that influence innovation, and a focus on policy (Fagerberg and Sapprasert, 2011). Beside the common elements, the original versions of the conceptual framework and approaches to studying NISs also differ. Even though, all three major contributions to the NIS literature are usually cited together by the users of this literature (Fagerberg and Sapprasert, 2011) it is not clear to what extent are these differences confronted in the research literature. Edquist and Hommen (2008) also argues that it is fair to say that the researchers tend to adopt Nelson’s, Lundvall’s version or some alternative variation without considering the alternative approaches when studying NISs. While early stages in the development of a novel concept tend to be associated with the frameworks that evolve as they

are empirically applied, it is not evident how the NIS conceptual framework have been refined through its application. Even though the main aim of the paper is to understand how the NIS concept has been used and diffused among the scholars, the useful starting point is to go back in history and compare the three initial versions of the concept, followed by the discussion on different contexts in which each approach has been developed.

Drawing on the earliest versions of the concept as presented in the three NIS core contributions as well as on the originators explanations about the distinctiveness and commonalities between the approaches, the following section will try to relate them. In particular we contrast the approaches in relation to conceptual and definitional issues , the main orientation of studies, the theoretical underpinnings, as well as particular approaches adopted in studying NISs (table x).

According to Lundvall (1992), the NIS concept introduced by Freeman refers to the nation-specific organization of sub-systems, and to the interaction between sub-systems. Yet, Freeman's main focus is the analysis of the organization of R&D and production in firms, the relationships among firms, and the role of the government. The analysis is mainly historical and draws upon innovation theory. On the other side, Nelson analysis focuses on the public-private character of technology and information, and the role of private firms, government and universities in the production of new technology. The distinctiveness of the approach developed by Lundvall in comparison with two other approaches includes the main focus of analysis being an interactive learning. Lundvall concludes that different theoretical approaches bring forward different aspects of the system and that one single approach to studying NIS should not be preferred (Lundvall, 1992).

In the preface of the section on National Systems of Innovation in the volume edited by (Dosi et al., 1988), where all three authors presented their theoretical and/or empirical research on NISs, Nelson emphasized that what is common for his preliminary report that focuses on the United States and Freeman's chapter on Japan is a complex institutional structure that characterize innovation system of the

Table 1: Comparison between the original versions of the NISs concept

	Freeman, C	Nelson, R. R.	Lundvall, B.Å.
Concept definition	<i>"The network of institutions in the public and private sectors which activities and interactions initiate, import, modify and diffuse new technologies may be described as the national system of innovation."</i> Freeman (1987)	<i>"[...] set of institutions whose interactions determine the innovative performance of national firms."</i> Nelson (1993)	<i>"[...] all parts and aspects of the economic structure and the institutional set-up affecting learning as well as searching and exploring - the production system, the marketing system and the system of finance present themselves as a sub-systems in which learning takes place."</i> Lundvall (1992)
Term "National"	NA	national borders	<i>"[...] national system encompasses elements and relationship, either located within or rooted inside the borders of the nation state. "</i> Lundvall (1992)
Term "System"	NA	<i>"[...] the set of institutions whose interactions determine the innovation performance of firms"</i> Nelson (1993)	<i>"[...] a system of innovation is constituted by elements and relationships which interact in the production, diffusion and use of new and economically useful knowledge."</i> Lundvall (1992)
Term "Innovation"	NA	<i>"[...] the process by which firms master and get into practice product design and manufacturing process that are new to them."</i> Nelson (1993)	<i>"[...] on-going process of learning, searching and exploring, which result in new products, new techniques, new forms of organization and new markets."</i> Lundvall (1992)

Table 2: Comparison between the original versions of the NISs concept continued

	Freeman, C	Nelson, R. R.	Lundvall, B.Å.
Analytical framework	NA	Linking institutional arrangements to technological and economic performances.	Interactive learning anchored in the production structure (including “demand conditions” and “supporting industries”) Institutional set-up including “firm strategy” Modes of cooperation and competition
Elements of the system	The role of the Ministry of International Trade and Industry The role of company R&D strategy in relation to imported technology and reverse engineering The role of education and training and related social innovations The conglomerate structure of industry	The institutions and mechanisms supporting technical innovation	The internal organization of firms Inter-firm relationships Role of the public sector Institutional set-up of the financial sector R&D intensity and R&D organization
Main orientation	To develop the idea of “National Systems of Innovation” associated with pervasive changes in technology	To understand and describe differences and similarities between national systems and the extent that these differences explain variation in national economic performances	To provide a theoretical perspective that might be used in case studies and to discuss some important sub-systems
Type of the analysis	Single case study (Japan)	Comparative case study (15 countries)	Conceptual/Theoretical

modern industrialized nations. Additionally, the definition of NIS involves the actors and activities of industrial R&D as well as institutions such as universities and government funds and programs. In contrast, the main focus of analysis in Lundvall's chapter is on user-producer interactions (patterns of interaction between upstream and downstream firms or university-industry interactions) while the NIS concept refers to the networks of technological interactions that tend to be enclosed within the national borders. Additionally, Nelson points that while his and Freeman chapter assume the existence of NISs, Lundvall presents a theory that try to explain this.

In sum, the differences between the approaches can be attributed to narrower or broader definitions of the concept, a main focus of the analysis, and a main theoretical perspective used for studying the NISs. What is regarded as common to all three approaches is a focus on the constituents of systems of innovation, that is, institutions, organizations and interactions.

3 Method

In order to map the structure and development within the branch of NSI research, we use a set of different methods. The following section provides a brief overview over their characteristics and potentials, and our rationales to apply them given the aim of this study.

First, we carry out a bibliometric analysis, where we focus on mapping knowledge and community structure as revealed in by a network-analysis of citation pattern. Our underlying assumption is that the way articles and other contributions refer to each others carries information on their conceptual relatedness. Therefore, we use different techniques to make sense of raw citation information. In a first step, we carry out a simple analysis based on direct citations in order to identify the most important contributions based on their received citations and centrality within the citation network. However, while useful for a first overview on the general relevance of articles, direct citation analysis has shown to not be very accurate in identifying relationships and relatedness between the articles, and communities within the cor-

pus. Therefore in a next step we also do a co-citation as well as a bibliographic coupling analysis.

Our main aim here is to identify existing and emerging community structures within the NIS literature. While up to now only focusing on citation data, we afterwards manually enriched our data with manually selected labels indicating the theoretical and analytical focus, type of research method and purpose of theory in the articles. This manual labeling enriches our analysis with additional information related to the broad qualitative dimensions of the work within the NIS literature and its sub communities, and its development over time. Finally, we deploy methods from natural language processing, namely LDA topic modeling to in a unrestricted and data-driven way add another qualitative layer of information on the context of work within the literature. In the following, we will in more detail describe the data and methods used in our analysis, and elaborate on the rationales behind our choices.

3.1 Data description

For our bibliometric analysis of the NIS literature, we draw from the Thompson Reuter's "Web of Science" (WoS) database. Since the branch of NIS literature can be seen as rather intradisciplinary and heterogeneous, finding an initial corpus of publications proves to be challenging. First intuitive possibility would be to search the database for articles including the terms "National Innovation Systems" or "National Systems of Innovation". This would likely lead to an initial corpus containing very little unrelated studies (false positives), but due to inconsistency in the usage of vocabulary within the community exclude many relevant studies (false negatives). To solve this problem arising from the blurry boundaries of NIS literature, we apply an iterative search strategy. First, we select a first seed article with literature-review characteristics, meaning that it is likely to contain references to a large part of the NIS core literature. In our case, we choose Fagerberg and Sapprasert (2011), a recent review of the emergence and development of NIS literature. WoS offers th function-

ality to search for related articles, meaning such that share common references with the initial one (basically, the same intuition as behind bibliographic coupling, which is explained later). We selected the 5,000 articles that share most common references as our main corpus, and excluded ones that did not receive any citation, leaving 3,705 articles. For further analysis, we extracted citation information our main corpus, expanding our our corpus of 234,045 cited articles. Being more restrictive on this level, we excluded all cited references that receive less than 5 citations from our initial copus, reducing their number to 123,483.

3.2 Organization of knowledge production: Bibliometric methods

After an initial analysis mainly focused in identifying the core contributions within the NIS literature, we proceed with a closer investigation of the structure of knowledge production within the field of NIS literature. While initial data indicates that the field indeed originates from the aforementioned three core contributions, the original theoretical framework(s) as well the focus of research within it has undergone substantial change during its development over the last two decades. The way knowledge is produced - in academia and elsewhere - can to a large extend be explained by its structure of diffusion and relation. For the representation and investigation of such structures, network analysis has proven to be a powerful tool, especially to identify communities focused on the production of particular types of knowledge. Therefore, we use different bibliometric techniques to represent the field of NIS literature as a network based on citation-relationships, namely co-citation analysis and bibliographic coupling. In the following, we briefly explain the intuition and rationales between these two techniques.

3.2.1 Relational mapping: Co-Citation analysis

We co-citation analysis and create a network between these contributions based on their co-citation similarity. A direct citation reflects an author is influenced by the work of another author, but usually does not explicitly indicate the strength or

direction of that influence. Conventionally, it is assumed that each reference makes equal contribution to the citing article. In contrast, the strength of the connection between two articles varies with the frequency in which two articles are cited together by other article. Hence this method is based on the assumption the citation of two articles together by another one contains information on their relatedness. The more often such co-citations occur, the more the corresponding articles can be assumed to share contextual similarities. In other words, these similarities are based on the revealed perception of other researchers in the field, who decide to cite both articles at the same time. While it is generally agreed that co-citations provide more adequate results on the relatedness between articles than simple citation analysis (Boyack and Klavans, 2010), it is not without caveats. Foremost, co-citation analysis represents a “backward looking” way of mapping relationships, since it is based on citation pattern of articles published after the article of interest. Consequently, co-citation analysis tends in a similar way as direct citation analysis to be biased towards pronouncing the relationships among earlier citations, which accumulated more citations over time.

Another advantage of utilizing citation information of the core corpus is that we thereby also include books and book-chapters, which are otherwise covered rather poorly in the mainly on journal articles focused WoS database. Yet, in the field of innovation system research, the most significant contributions consist of books and edited volumes (Fagerberg and Sapprasert, 2011).

3.2.2 Bibliographic coupling

Bibliographic coupling, like Co-citation, is a similarity measure that uses citation analysis to establish a similarity relationship between articles. Two articles are bibliographically coupled if they both cite one or more ore in common, where the “coupling strength” increases with the number of citations to other articles they share. It thus follows the opposite logic that co-citation analysis, in a way that the relationship between articles is not measured by the perception of others, but by their internal

logic of how they relate to the rest of the community. It follows the intuition, that two articles sharing a common pool of references are in some way related in terms of context, methods or theory they use, and to which academic community they relate to.

3.3 Mapping of qualitative characteristics: Contextual methods used

3.3.1 Manual article-classification

Based on the review of similarities and differences between the three original versions of the NIS concept and on the research question posed in this article, the classification scheme has been defined including the three dimensions. The classification scheme will be used for manual classification of a sample of top cited articles in the NIS branch of literature as sourced from the ISI Web of Science database. The aim of the classification was to map the research being conducted in the NIS field in relation to the main focus, purpose of the theory related activities and the type of research methods used by the researchers.

The first dimension used for classifying the identified articles includes the main focus of the research paper being either institutions, organizations, networks, or some combination of the two or all the three categories. The main rationale for classifying the papers based on their main focus was to explore the main approaches taken and elements of the system studied in the NIS literature. A scheme for classifying the articles based on their main focus was developed by authors based on the original version of the NIS conceptualization that put forward different aspects of NISs as their main focus. The three categories may also be regarded as the main building blocks of the NIS approach that include elements of the systems (organizations, institutions) and relationships between them.

The manual classification of the papers under the previously introduced categories was based on reading of the title and abstract of the paper. The example of the research paper focusing mainly on the institutional aspect of NIS includes conceptual paper focusing on the combined formation of NSIs and welfare system as an institu-

tional response to modernization-marginalization polarization (Albuquerque, 2007). The number of papers that could be exclusively classified as having a main focus on institutions was rather limited. Most of the analyzed papers have as their main focus some combination of institutions and organizations, institutions and networks or include all three dimensions. An alternative approach that could be used to identify the main focus of the articles may include the existence and variation in NISs definition by researchers in a sample of selected articles instead of imposing the predefined categories based on the comparison of the original version of the conceptualization of the NIS concept.

The second dimension that has been used for classification of the sample of articles includes the purpose of theory related activities adopted in the articles. The abstracts of the articles were analyzed to determine if the main purpose of the theory related activities belongs to descriptive use of theories, theory building or theory testing. The main rationale for using this dimension was to gain insight into the development and maturity of the NIS branch of literature. There is a general agreement that theory development is a requirement for the development of any field (Kuhn, 1970). Additionally, the maturity of the field may be evaluated based on the proportion of the activities associated with theory building versus theory testing. The early stages of the development of the field are usually related to more extensive work on theory building, while later stages of the development of the field would include more activities connected to theory testing (Handfield and Melnyk, 1998). The abstract of the articles were analyzed to determine if the authors are explicitly stating activities related to theory development or theory testing. The abstracts that contained the statement on the theory validation, extension and/or refinement were grouped together within the category theory testing. The abstract that contained the statement on theory building were grouped within the category theory building. If the purpose of theory related activities was not explicitly stated, the abstract of the article were further analyzed to determine the extent to which the article can be characterized as contributing to building new theory and the extent to which it can be

regarded as theory testing. For example, the explorative research papers were classified under the category theory building. An additional category labeled descriptive use of theory was used for classifying the articles within the sample which abstract contained an explicit statement as being descriptive or had no implicit statement on theory related activities. It is though important to emphasize that the descriptive use of theory may be regarded as an activity related to theory building and should be included in this category, while the articles containing no reference to theory (based on the review of the abstract) implication would better fit within the separate category labeled practitioner or policy oriented studies.

The last dimension that has been used for manual classification of the articles in the sample includes the type of research methodology used. More specifically, the abstracts of the articles were analyzed to determine if the articles fall into the two broadly defined categories named analytical or empirical research methods. Analytical research methods were further categorized as conceptual and formal modeling, while the empirical methods were further categorized as quantitative and qualitative research methodology. The main rationale for using this dimension was to get an insight into the potential systematic pattern in the literature in relation to the research methodology used.

The breadth of the analysis covered by the classification scheme based on the three dimensions can be extended by introducing additional dimensions such as or more fine-grained categories within the each dimension in order to develop more comprehensive understanding of the NIS literature.

3.3.2 Topical mapping: Natural language processing: LDA topic modeling

Our second approach to enrich as such non-contextual documents and their relationships within the empirical analysis, we utilize techniques from the field of natural-language-processing, namely probabilistic topic models. They represent a family of methods that automatically extract thematic information (i.e. topics) from large collections of text documents (i.e. text corpus). They are useful for analyzing

the massive amount of text that would be unfeasible (take too long or be too expensive) to annotate by human reading. Topic modeling algorithms enable content analysis to work on scale beyond reach of human annotation (Blei, 2012). Topic modeling algorithms aim to reliably label large collections of text documents with thematic information (Blei, 2012). These methods try to discover the themes in original texts and how these themes are connected to each other. Topic models are Bayesian models of text document collections. A topic is a distribution of words over a fixed vocabulary. However, even though topics are automatically inferred from text and are in its essence just word distributions over documents, they usually resemble thematic concepts surprisingly well.

We here use the popular Latent Dirichlet Allocation (LDA) topic model to in a data-driven way identify important conversations and themes within the NIS research field. LDA is a generative model, meaning that it takes the end document and tries to find parameters of the process that generated it. LDA is a model of document generation process. According to this model, in order to make a new document, two-step procedure is followed: (1) choose a distribution over topics and (2) for each word in the document, choose a topic from the distribution chosen in first step and draw a word from that topic. When we have a collection of documents, we can use statistical methods to infer the set of topics that generated that collection. The generative model of LDA is presented in figure ??.

4 Results and Discussion

4.1 Direct Citation Descriptives

Table 3 summarizes the most influential contributions to the field of NIS studies, both by citations received by other contributions within our dataset, as well as by their weighted degree within the co-citation network (method to be described later). This results confirm former findings (Fagerberg and Sapprasert, 2011; Teixeira, 2013) that, independent of the metric and corpus used, the books written by Lundvall

(1992), Nelson (1993) indeed represent the by far most significant contributions to the NIS literature, followed by Freeman (1987).

Table 3: Most frequently cited NIS contributions

Direct Citations			Weighted Co-Citations	
Article		Cit.	Article	Cit.
LUNDEVALL B-1992-NATIONAL SYSTEMS INN		124	NELSON R-1993-NATL INNOVATION SYST	5169
NELSON R-1993-NATL INNOVATION SYST		124	LUNDEVALL B-1992-NATIONAL SYSTEMS INN	5095
FREEMAN C-1987-TECHNOLOGY POLICY EC		61	FREEMAN C-1987-TECHNOLOGY POLICY EC	2654
PORTER M-1990-COMPETITIVE ADV NATI		52	PORTER M-1990-COMPETITIVE ADV NATI	2337
FREEMAN C-1995-CAMBRIDGE J ECON		35	FREEMAN C-1995-CAMBRIDGE J ECON	1920
WINTER S G.-1982-EVOLUTIONARY THEORY	31.00		WINTER S G.-1982-EVOLUTIONARY THEORY	1522
PAVITT K-1984-RES POLICY		27	LUNDEVALL B-2002-RES POLICY	1418
DOSI G-1988-TECHNICAL CHANGE EC		25	DOSI G-1988-TECHNICAL CHANGE EC	1258
LUNDEVALL B-1988-TECHNICAL CHANGE EC		25	FREEMAN C-1972-EC IND INNOVATION	1225
LUNDEVALL B-2002-RES POLICY		24	LUNDEVALL B-1988-TECHNICAL CHANGE EC	1194
FREEMAN C-1972-EC IND INNOVATION		23	ETZKOWITZ H-2000-RES POLICY	1156
FREEMAN C-1988-TECHNICAL CHANGE EC		22	PAVITT K-1984-RES POLICY	1145
CARLSSON B-1991-J EVOLUTIONARY EC		20	COHEN WM-1990-ADMIN SCI QUART	1098
COHEN WM-1990-ADMIN SCI QUART		20	FREEMAN C-1988-TECHNICAL CHANGE EC	1065
ETZKOWITZ H-2000-RES POLICY		17	LUNDEVALL B-2007-IND INNOV	1023
GRILICHES Z-1990-J ECON LIT		17	CARLSSON B-1991-J EVOLUTIONARY EC	939
FURMAN JL-2002-RES POLICY		16	FURMAN JL-2002-RES POLICY	914
FREEMAN C-2002-RES POLICY		16	FREEMAN C-2002-RES POLICY	905
LUNDEVALL B-2007-IND INNOV		15	GIBBONS M-1994-NEW PRODUCTION KNOWL	845
DOSI G-1982-RES POLICY		14	NORTH D-1990-I I CHANGE EC PERFOR	840

4.2 LDA topic modeling

Table 4 reports the results of the LDA topic model analysis on the abstracts of our initial corpus of NIS articles. We indeed see well defined linguistic topics emerging by the model, providing us with a good intuition on the context of the corresponding article. For example, topic 18 appears to clearly associate with studies in economic geography, topic 20 with institutional theory, topic 4 with firm level management and entrepreneurship studies, and topic 2 with networks and interaction structures. In a next step, we plan to relate the articles – and on a higher level of aggregation, the communities – with the topics mostly resembling the linguistic structure of the corresponding abstract. That would enable us to do further in-depth studies on the importance of these topics across researchers, communities, and over time.

References

- Albuquerque, E. (2007). Inadequacy of technology and innovation systems at the periphery. *Cambridge Journal of Economics*, 31(5):669–690.
- Amable, B. (2000). Institutional complementarity and diversity of social systems of innovation and production. *Review of International Political Economy*, 7(4):645–687.
- Bergek, A., Jacobsson, S., Carlsson, B., Lindmark, S., and Rickne, A. (2008). Analyzing the functional dynamics of technological innovation systems: A scheme of analysis. *Research Policy*, 37(3):407–429.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4):77–84.
- Boyack, K. W. and Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology*, 61(12):2389–2404.
- Breschi, S. and Malerba, F. (1997). Sectoral innovation systems.
- Burt, R. S. (1992). Structural holes: The social structure of competition. In Nohria, N. and Eccles, R., editors, *Networks and Organizations*, pages 57–91.
- Carlsson, B. and Stankiewicz, R. (1991). In search of useful public policies: key lessons and issues for policy makers. *Journal of evolutionary economics*, 1(2):93–118.
- Cooke, P. (2001). Regional innovation systems, clusters, and the knowledge economy. *Industrial and corporate change*, 10(4):945.
- Dosi, G., Freeman, C., Nelson, R., Silverberg, G., and Soete, L. (1988). *Technical Change and Economic Theory*. Laboratory of Economics and Management (LEM), Sant’Anna School of Advanced Studies, Pisa, Italy.
- Edquist, C. and Hommen, L. (2008). *Small country innovation systems: globalization, change and policy in Asia and Europe*. Edward Elgar Publishing.
- Fagerberg, J. and Sapprasert, K. (2011). National innovation systems: the emergence of a new approach. *Science and public policy*, 38(9):669–679.
- Freeman, C. (1987). *Technology Policy and Economic Performance: Lessons from Japan*. Pinter Publishers, London.
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380.
- Handfield, R. B. and Melnyk, S. A. (1998). The scientific theory-building process: a primer using the case of tqm. *Journal of operations management*, 16(4):321–339.
- Hekkert, M., Suurs, R., Negro, S., Kuhlmann, S., and Smits, R. (2007). Functions of innovation systems: A new approach for analysing technological change. *Technological Forecasting and Social Change*, 74(4):413–432.

- Jones, B., Wuchty, S., and Uzzi, B. (2008). Multi-university research teams: shifting impact, geography, and stratification in science. *science*, 322(5905):1259–1262.
- Kuhn, T. S. (1970). *The Structure of Scientific Revolutions*. University of Chicago Press.
- Lundvall, B.-Å. (1992). *National Systems of Innovation: Toward a theory of innovation and interactive learning*. London: Pinter Publishers.
- Malerba, F. (2002). Sectoral systems of innovation and production. *Research policy*, 31(2):247–264.
- Malerba, F. (2005). Sectoral systems of innovation: a framework for linking innovation to the knowledge base, structure and dynamics of sectors. *Economics of Innovation and New Technology*, 14(1-2):63–82.
- Malmberg, A. and Maskell, P. (2002). The elusive concept of localization economies: towards a knowledge-based theory of spatial clustering. *Environment and planning A*.
- Merton, R. K. (1957). The role-set: Problems in sociological theory. *British journal of Sociology*, pages 106–120.
- Nelson, R. R. (1993). *National innovation systems: a comparative analysis*. Oxford University Press, USA.
- Pfeffer, J. (1993). Barriers to the advance of organizational science: Paradigm development as a dependent variable. *Academy of management review*, 18(4):599–620.
- Simmel, G. (1955). The web of group affiliations. *Conflict and the web of group affiliations*, pages 125–95.
- Singh, J. (2005). Collaborative networks as determinants of knowledge diffusion patterns. *Management science*, 51(5):756–770.
- Teixeira, A. A. (2013). Evolution, roots and influence of the literature on national systems of innovation: a bibliometric account. *Cambridge journal of economics*, page bet022.
- Wagner, C. and Leydesdorff, L. (2005). Network structure, self-organization, and the growth of international collaboration in science. *Research policy*, 34(10):1608–1618.
- Whitley, R. (2000). The institutional structuring of innovation strategies: business systems, firm types and patterns of technical change in different market economies. *Organization Studies*, 21(5):855.

Appendix

Figure 1: The generative model of Latent Dirichlet Allocation

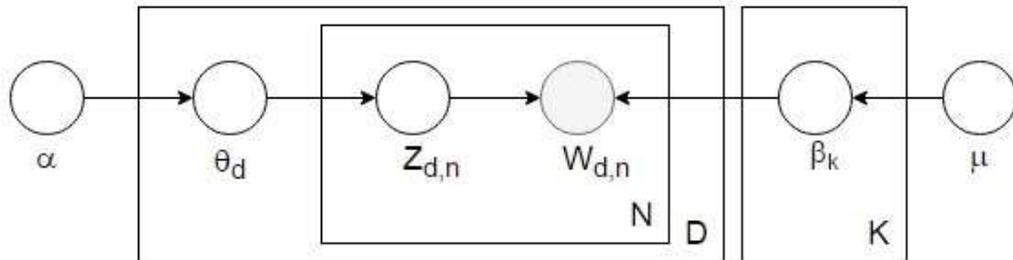


Table 4: Selected LDA topics and associated terms

Topic	Associated terms
1	role, important, universities, collaboration, university, cooperation, linkages, benefits, education, projects, play, interactions, related, less, key, types, collaborative, links, importance, organisations
2	social, networks, network, clusters, actors, cluster, interaction, relationships, relations, case, analysis, networking, community, capital, within, collective, basis, social capital, relevant, way
3	framework, approach, literature, theoretical, processes, concept, perspective, approaches, understanding, empirical, implications, conceptual, insights, analysis, concepts, need, transition, analytical, within, context
4	business, management, strategies, strategic, strategy, capacity, environment, entrepreneurial, entrepreneurship, resources, opportunities, resource, creation, conditions, ability, develop, implications, entrepreneurs, capabilities, critical
5	system, systems, environmental, design, sustainable, problems, challenges, elements, set, health, future, potential, systems innovation, problem, implementation, quality, evaluation, towards, agricultural, many
6	national, level, analysis, european, differences, sectors, levels, countries, patterns, states, indicators, europe, based, united, terms, infrastructure, sector, sectoral, germany, central
7	knowledge, learning, processes, economy, sources, transfer, within, production, interactive, base, flows, forms, knowledgebased, mechanisms, exchange, skills, context, source, creation, knowledge transfer
8	factors, information, diffusion, characteristics, adoption, data, services, practices, based, communication, key, influence, users, ict, identify, important, survey, risk, high, internet
9	countries, international, global, developing, foreign, economies, developed, emerging, biotechnology, industrial, domestic, investment, china, developing countries, trade, competitiveness, national, multinational, world, south
10	firms, innovative, activities, firm, external, manufacturing, small, large, internal, survey, sources, high, size, extent, higher, likely, results, resources, innovation activities, customers
11	science, studies, scientific, analysis, field, years, science technology, literature, researchers, fields, work, academic, attention, issue, last, past, number, time, areas, future
12	model, models, organizational, dynamics, organizations, dynamic, organization, evolution, structure, complex, time, changes, mechanisms, life, agents, existing, point, decision, complexity, group
13	policy, policies, public, government, sector, energy, support, private, development, innovation policy, case, success, programs, initiatives, instruments, makers, policy makers, systemic, governments, efficiency
14	innovation, process, innovations, support, smes, service, open, processes, product, innovation process, enterprises, construction, innovation processes, understanding, types, innovation performance, increasingly, product innovation, innovate, technological innovation
15	results, performance, impact, data, effects, empirical, relationship, effect, findings, positive, patent, activity, influence, significant, evidence, productivity, variables, patents, period, indicate
16	technology, technological, technologies, change, process, capabilities, development, transfer, technical, capability, technological change, technology transfer, technological capabilities, generation, path, technological innovation, diffusion, progress, important, acquisition
17	industry, market, industries, markets, competitive, product, companies, value, production, based, competition, products, advantage, demand, costs, industrial, marketing, chain, supply, competitive advantage
18	regional, local, regions, industrial, region, spatial, geography, within, spillovers, proximity, studies, geographical, production, location, recent, creative, economic geography, agglomeration, urban, importance
19	development, economic, growth, financial, economy, economic growth, capital, economic development, human, relation, evidence, economies, main, contribution, labour, income, recent, increasing, real, employment
20	institutional, theory, institutions, evolutionary, governance, state, economics, political, theories, corporate, change, power, historical, variety, intellectual, issues, transformation, structures, particular, argued