Dynamics in Science-Based Markets:

Two Phases of Development

Zur Erlangung des akademischen Grades eines Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

der Fakultät für Wirtschaftswissenschaften

der Universität Fridericiana Karlsruhe (TH)

genehmigte

Dissertation

von

Dipl.-Math. oec. Inna Haller

Tag der mündlichen Prüfung:	13.05.2009
Referent	Prof. Dr. Werner Rothengatter
Korreferent	Prof. Dr. Wolf Fichtner

Danksagung

Die vorliegende Arbeit entstand während der Zeit meiner Arbeit als wissenschaftliche Mitarbeiterin am Institut für Wirtschaftspolitik und Wirtschaftsforschung (IWW) der Universität Karlsruhe (TH). Daher gilt mein Dank in erster Linie Herrn Prof. Dr. Grupp für die große Unterstützung mit Rat und Tat sowie die Freiheiten, die er mir beim Ausgestalten des Themas gewährt hat. Nach seinem tragischen Tod hat Prof. Dr. Rothengatter die Betreuung dieser Arbeit übernommen und zu einem erfolgreichen Ende geführt. Dafür danke ich ihm sehr herzlich. Herrn Prof. Fichtner danke ich für die Übernahme des Korreferates und zahlreiche Verbesserungsvorschläge.

Mein Dank geht weiterhin an die Mitarbeiter des Instituts für die gute, angenehme Zusammenarbeit und tatkräftige Unterstützung: Eckard Szimba, Dirk Fornahl, Patrick Jochem, Mirja Meyborg, Melanie Funken-Vrohlings. Ein ganz besonderer Dank geht an meinen langjährigen Zimmergenossen am IWW Chung Anh Tran, der mir oftmals in vielen Situationen und Lebenslagen hilfreich zur Seite stand und der demnächst auch sein Promotionswerk zu Ende führt.

Die vorliegende Dissertation wäre nicht zustande gekommen, wenn ich nicht breite Unterstützung seitens meiner Familie gehabt hätte. Danken möchte ich ebenfalls meinem Sohn Dominik, der mir Tag für Tag zeigt, dass es auch andere wichtige Dinge im Leben gibt. Ihm ist diese Arbeit gewidmet.

Karlsruhe, Januar 2010 Inna Haller

Table of Contents

List of Figures

List of Tables

Table of Abbreviations

Akaike Information Criterion	AIC
Amorphous Silicon	a-Si
Augmented Dickey-Fuller	ADF
Cadmium Telluride	cdTe
Citation	C
Co-integrated	CI
Combined Heat and Power	CHP
Concentration Ratio	CR
Coordination Failure Diagnostics	CFD
Copper Indium Diselenide	CIS
Electricity Feed-in Law	StrEG
Error Correction Model	ECR
European Patent Office	EPO
European Research Area	ERA
European Union	EU
Federal Ministry for the Enviroment, Nature Conservation and Nuclear Safety	BMU
German Society for Solar Energy	DGS
Giant Component	GC
Gross Domestic Product	GDP
Gross National Product	GNP
Herfindahl Herschman Index	HHI
Information and Communication Technologies	ICT
Intenational Patent Classification	IPC
International Energy Agency Photovoltaic Power Systems	IEA PVPS
Maastricht Economic Research Institute on Innovation and Technology	MERIT
Mega Watt Peak	MWp
Monocristalline Silicon	m-Si
National Bureau of Economic Research	NBER
National Science Foundation	NSF
Network	NW
Non-Patent Citation	NPC
Non-Patent Citation Share	NPCS
Non-Patent Literature	NPL

Non-Patent Literature Mean	NPLM
Non-Patent Reference	NPR
Organization for Economic Cooperation and Development	OECD
Patent Literature	PL
Photovoltaic	PV
Polycristalline Silicon	p-Si
Public Promotion Catalogue of the Federal Ministry of Education and	FöKat
Renewable Energy Law	EEG
Research and Development	R&D
Revealed Technology Advantage	RTA
Science Citation Index	SCI
Science Intensity	SI
Science-Based	SB
Social network analysis	SNA
Standard International Trade Classification	SITC
Technology in Retrospect and Critical Events in Science	TRACES
United States Patent and Trademark Office	USPTO
Vector Autogressive Model	VAR

1 Introduction

1.1 Motivation

The idea that technological progress plays an important role in the economy of industrialized societies is not new. Its dominating impact was demonstrated by many economists. For instance, Karl Marx saw technological innovations as dynamic driving force of economic and social development. Joseph Schumpeter identified and discussed the great importance of innovation in economies. He saw innovations as a process of "creative destruction" that breaks the old economic structures and creates the new ones. Neoclassical economists attributed economic growth to the increase of capital and labor inputs, and technological progress that was measured as residual factor. Solow (1957) showed that the contribution of technological progress to economic growth is more important than that of capital accumulation and increases in labor. Solow calculated that approximately 87% of U.S. economic growth between 1909 and 1949 was due to technical progress. The Solow neoclassical model was responsible for important advances in economics. However, the technical progress in the neoclassical growth theory is exogenous. It does not explain where technologically driven productivity growth comes from. Starting in the 1980s, the endogenous growth theory has been developed. It views technological progress as a product of economic activity.

However, the renewed importance attributed to technical progress is based on the current discussion about transformation of the modern economy into a knowledge-based economy. Lundvall (1992, 1996) denotes this change as a new phase of economic development that highlights the shortcomings of standard economic theory and policy because knowledge differs substantially from other economic resources. In contrast to conventional factors of production such as capital and labor, knowledge does not decrease in value when it is used. Quite the opposite; knowledge appreciates only through application.

Knowledge is the most important factor of production for long-term growth and job creation. There are a number of empirical facts that confirm this statement:
Overall world manufacturing output grew from $13.9 trillion in 1990 to $19.6 trillion in 2003 after adjusting for inflation. However, the manufacturing output of five high-

technology industries[1] grew faster than this, from $1.5 trillion to $3.5 trillion. Firms and sectors associated with the highest levels of knowledge (technology) display better economic performance, whereas aggregated productivity and employment growth remain modest in most countries (OECD, 1996a). The share of high- and medium-high-technology industries[2] represents about 65% of OECD manufacturing trade. The shift towards knowledge economies has led to knowledge-intensive employment. Much of the employment growth has been in relatively highly skilled knowledge-intensive sectors, with a shift away from the more manual occupations towards the non-manual and more knowledge-intensive occupations (OECD, 2007). Most OECD countries are increasing their investment in the knowledge base[3] (OECD, 2008). This can be taken as evidence of the increased importance attributed to knowledge.

All of these facts reveal a paradigm shift towards a knowledge-based economy. Unfortunately, there is not any clear generally accepted definition of a knowledge-based economy. This problem of definition follows at least partially from the methodological difficulties that arise when attempting to measure the knowledge-based economy. New indicators are required in order to measure innovative performance and other related output of a knowledge-based economy.

The definition of a knowledge-based economy has often been rather rhetorical in nature. Godin (2006) states that the knowledge-based economy presents at least two characteristics: firstly, an increased quantitative and qualitative importance of knowledge. Secondly, application of information and communication technologies (ICT) is considered as an important driver of the knowledge-intensive economy. Application of ICT helps to codify certain types of knowledge and increases its effective usage. All codified knowledge can be easily transmitted over long distances with limited costs.

OECD (1996b, p. 7) provides the following definition of knowledge-based economies:

---

1 Examples of high-technology industries are aircraft, pharmaceuticals, office and computing equipment, communications equipment, and scientific instruments (http://www.nsf.gov/statistics accessed on 04.01.2009)

2 Medium-high-technology includes motor vehicles, electrical equipment and most chemicals (http://stats.oecd.org accessed on 04.01.2009)

3 Investment in knowledge is defined and calculated as the sum of expenditure on R&D, on total higher education (public and private) and on software (OECD, 2008).

> "…economies which are directly based on the production, distribution and use of knowledge and information."

This definition can often be found in scientific documents dealing with the knowledge-based economy. In order to provide this definition with more substance, OECD suggests two further concepts closely linked with the concept of the knowledge-based economy. The first is investment in knowledge, as mentioned above. The second is knowledge-based industries, which, according to Godin (2006), feature the following three characteristics: firstly, a high level of investment in innovation; secondly, intensive use of acquired technology, and finally, a highly educated workforce. The concept of knowledge-based industries focuses not only on the main producers[4] of high-technology goods, but also on intensive users of high technologies who benefit from technological innovations. For this reason, the analysis of knowledge-based industries should include supply and demand driving forces.

The European Union (EU) has also recognized the evidence demonstrating the importance of the qualitative transition in common economic development and has invested in significant opportunities to help this change process. In March 2000, the EU formulated its policies in line with the ambitious objectives of the so-called Lisbon Strategy. The European Council stated that:

> "the shift to a digital, knowledge-based economy, prompted by new goods and services, will be a powerful engine for growth, competitiveness and jobs. In addition, it will be capable of improving citizens' quality of life and the environment." (Presidency Conclusions, Lisbon European Council, 23 and 24 March 2000)

One of the central elements of the EU Lisbon Strategy is the creation of a European Research Area (ERA) that should support cooperation and coordination activities carried out at the European, national and regional level. More teams should be able to form research networks and, in this way, the free movement of people and ideas should be stimulated.

Assuming that a society is to become knowledge-based, it will face challenges on a range of levels. There are some crucial issues can be mentioned in this context: knowledge production, protection, and dealing with spillover effects. Increasing

---

[4] This term refers not only to the firms that produce high-technology goods, but also to the researchers and scientists that offer solutions for diverse technological problems.

complexity of modern technologies leads to higher costs arising by production of knowledge. Knowledge can be copied if the “knowledge producer” does not take the necessary precautions. In a global economy national borders are no longer barriers to knowledge diffusion. Furthermore, knowledge rapidly becomes obsolete.

Globalization pressure on economies and enterprises increases the stringent necessity of obtaining permanent access to new technological achievements. A close relationship between science and technology is one of the most important features of knowledge-based economy. The analysis of innovation processes in science-driven markets turns out to be of great importance for the common economic development and competitiveness of industrialized nations.

Certain institutional arrangements are required for the creation and development of science-based technologies. However, the nature of scientific knowledge frequently displays spontaneous order, i.e. it is the result of activities carried out by many individuals and groups, none of which intend to bring about that particular state of the body of knowledge, either individually or collectively (e.g. Radnitzky (1989)). Therefore it appears to be difficult to forecast precisely the future trends of technological change. Nevertheless, it is helpful to enter into the black box of “science-based” models and to try to understand the dynamics of innovation processes in science-driven markets. The subject of this thesis is the analysis of development pattern of science-based technologies. The most studies dealing with evolution of science-based technologies have rather descriptive nature. The primarily purpose of this thesis is identification of endogenous and exogenous factors that shape the development of such technologies and quantifying the effect of these factors on the technological evolution.

1.2 Overview

This thesis is organized as follows: Chapter 2 discusses different possibilities for the operational distinction of science-based technologies. Additionally, the state of the art concerning the technology cycle of knowledge-intensive technologies is summarized in a stylized model. This model serves as a basis for further investigation. Following the stylized model three different investigation levels of science-based markets can be distinguished: the market level (or product level), the level of technological activities, and the level of scientific efforts. It is not possible to separate these levels from each other completely. There are interdependencies between all levels; e.g. if activities increase or decrease on one level this has an impact on the other levels. However, the usage of different databases makes it difficult or impossible to analyze all three market levels in one common model. For this reason, the empirical analysis of the thesis is carried out in three steps. The main goal of each part is to detect reasons for the occurrence of typical pattern in the market development of science-based technologies.

In order to gain a better understanding of the development pattern of science-based technologies, the solar Photovoltaic (PV) technology is selected for further investigation. The empirical analysis of the thesis is applied to this technological field. In chapter 3 the development of the PV market is described.

Chapter 4 takes a closer look at the solar PV modules quality and its technical improvement over time. The producers of solar cells play a key role as they influence the development and diffusion of PV technology significantly. The central research question therefore relates to the investigation of the technology cycle on product-level. The technological improvement is measured by specific product characteristics of solar modules. However, an innovation has a macro-economic impact only if it spreads quickly and widely (a process known as "diffusion"). In order to investigate economic impact of innovative solar modules, an examination is carried out whether producers of technically advanced products are able to achieve additional market segments in the period of interest.

In chapter 5, the reaction of the solar PV market on political and economic decisions is analyzed. Due to their long-term effect, the impact of such exogenous factors on the technology cycle can vary; depending on the stage of development (see also Schmoch 2007). The reaction of the market together with changes in the environment of the technology also influences the devolution of the technological path. A time series analysis is suitable for studying the dynamic pattern of the technology development and is included in the second part of the empirical investigation

Chapter 6 discusses changes in the networks of scientific community related to the solar PV technology. Especially in the case of science-based technologies the searching procedure for new technical solutions implies a fundamental uncertainty that is involved in innovation processes. This uncertainty affects firm survival essentially. Cooperations with other organizations and researchers may therefore be useful. In this context, cooperation networks represent important channels for the knowledge transfer. Jacobsson et al. (2004) underline that knowledge resources of individual firms can strongly be expanded by the integration into networks. Thus, the network gives access to information and knowledge of other actors. Since the publishing process of a scientific paper in most cases requires an intensive exchange of ideas between authors, co-authorship can be taken as an indicator for the strength of communication links. Relating to this, studying co-author relationships derived from scientific publications in the field of solar PV cells is one method to examine communication taking place in this field. The analysis of communication may offer important insights into patterns of relationships, inter-firm links, and changes in communication over time.

Consequently, the empirical part of section 6 has two main objectives: the first is to analyze whether there are changes in the topology of underlying co-author networks and whether these changes correspond to the occurrence of different phases in the technology cycle. The second is to analyze the cooperation behavior on the level of individual authors and organizations. Who cooperates with whom? Are the preferences regarding organization type by searching cooperation partners?

Studying these three empirical questions – technological improvements of solar cells and its diffusion on the market, the reaction of the market on exogenous factors, and changes in the topology and cooperative behavior of co-authors – helps to find real reasons for the occurrence of typical pattern in the technology cycle of solar PV cells. The thesis concludes with chapter 7.

2 The Two Phases Model of Science-Based Technologies

2.1 Operational Definition of Science-Based Technologies

The subject of this thesis is analysis of science-based (SB) technologies and the development pattern of these technologies. The definition of knowledge-intensive technologies can be directly derived from their strong dependence on new scientific achievements. But how can scientific dependence be measured? Here an operational definition is required. Unfortunately, there is not any clear, generally accepted categorization of a specific technology as science-based. The following section discusses some theories and approaches that can be used for demarcation of science-based technologies.

In general, technology is defined as a set of real, usually tangible tools by which parts of the environment can be transformed (Tronatzki and Fleischer 1990, p.11). The tools and processes in knowledge-based technologies are more complex and compose of numerous different components. Generally speaking, most technologies embody human knowledge directly or indirectly. Some inventions are made by the human mind through synthesis of observation of real world cause and effect principles. In contrast to many medium- resp. low-high-tech technologies, the basic knowledge of science-based technologies cannot be created by casual observations, craft skills or testing and error. The processes of contemporary technologies are often based on scientific[1] knowledge, even if it is not obvious to the user (cp. Tronatzki and Fleischer 1990)[2].

The problem of operational defining science-based technology has a long history. In order to give an impression about the variety of possibilities to define science-based technologies, four definition groups with different aggregation levels are discussed. The first definition originates in the work of Machlup (1962) and gives a wide

[1] “Normal science” means research firmly based upon one or more past scientific achievements, achievements that some particular scientific community acknowledges for a time as supplying the foundation for its further practice (Kuhn 1970, p. 10).

[2] However, Tronatzki and Fleischer (1990, p.13) emphasizes that the definition of technology as “function-performing tool” has also some weaknesses because some innovative practices that cannot be defined as technology in this sense. Innovations in jurisprudence can be taken as example. But such innovations are not the subject of this study. In this analysis only technological innovations and its dynamics are regarded.

definition of “knowledge industries. It includes all occupations that are concerned with producing and handling information rather than goods. According to Machlup (1962) knowledge can be divided into 5 categories[3]:

(1) Practical knowledge: professional, business, workman’s, political, household, and other practical knowledge;
(2) Intellectual knowledge regarded as part of liberal education, humanistic and scientific learning, and general culture;
(3) Pastime knowledge: entertainment and curiosity;
(4) Spiritual or religious knowledge;
(5) Unwanted knowledge accidentally acquired, aimlessly retained;

In this study, Machlup (1962, p. 362) estimated that “knowledge production in 1958 was almost 29 per cent of adjusted GNP.”

Freeman (1982) gives a historical overview about empirical studies that deal with knowledge-intensive technologies. These studies are conducted by European and American economists in three areas of industry: chemicals, plastics, and electronics. According to Freeman, these industrial fields belong to a main cluster of fast growing products in the whole post-war period with growth rates of about 10% per year. The definition of these industries as “research-intensives” is based on consideration of high ratio of professional R&D manpower in relation to total employment or of R&D expenditures to net output. In a similar, way the OECD (1986) defined knowledge-intensity in manufacturing sectors on the basis of R&D intensity. For a given sector, R&D intensity is the ratio of sectoral R&D expenditures to the share of industry output.

One of the most important empirical studies for identification of science-based technology at the sectoral level was made by Pavitt (1984). In this work Pavitt distinguishes between supplier-dominated sectors, scale-intensive sectors, special suppliers, and science-based ones. Based on an analysis of approximately 2000 innovations in British industry within the period of 1945-1979, Pavitt found typical innovation patterns in these broader sectors of industry. Innovations in science-based sectors display a close relationship to basic research and scientific progress. The innovations in these sectors require high investment in research (not only in product development), but offer properties of key technology with a strong diffusion potential in other industrial sectors (see also Martin, 1992). Marsilli (2001) further

[3] Machlup (1962, pp. 21-22).

suggests to split the science-based sectors into two main categories: the “life science-based” (drugs and bioengineering) and the “physical science-based” (computers, electrical telecommunications instruments). However, these studies offer a sectoral classification. Segmentation according to technology fields was not intended.

A high aggregation level of the sectoral approach is its essential weakness. Grupp (1996) writes that technological spillovers make it difficult to set bounds between branches of the economy. Scherer (1982) demonstrates regarding technology flows in the United States that R&D activity is mostly product-oriented and not geared to a branch of industry.

Can a categorization of technologies based on allocation of corresponding products offer an improvement to a more clear demarcation of science-based technologies? An implementation of such approach is suggested by Hatzichronoglou (1997). This method supplements the sectoral approach providing a more appropriate tool for analyzing international trade and improves the aforementioned approach in two ways:

Firstly, an industry may be characterized as strong knowledge-intensive in one country but only slightly knowledge-intensive in another. These differences are a consequence of product heterogeneity within one industry across countries. Secondly, using product lists it is possible to identify the share of high-tech products manufactured by medium-technology sectors.

The product approach supplies a list of products in the high-technology category corresponding to the three-digit SITC Rev. 3 classification of foreign trade. The R&D intensity of product groups is defined here as a ration of R&D expenditures to total sales covering six countries (the United States, Japan, Germany, Italy, Sweden, and the Netherlands). But this approach also has some limitations. First of all, it is a problem of sufficiently disaggregated data. Furthermore, the sectoral approach is more particular. The R&D intensity in each sector is attributed to the principal activity of those firms that establish the sector. In contrast, disaggregation to the product level leads to inconsistence for industries with very complex products. For example, allocation of airplanes complete to the aerospace industry leads to overestimation of R&D intensity in aerospace and underestimation of R&D intensity in electronics.

Another possibility of identifying science-based technology is given through the analysis of patent[4] data or patent application data[5]. Patents are a common source used to analyze and measure inventive activity. An explanation for the assumption that patents can be used as indicators for inventive activity is derived from three conditions that have to be fulfilled for patent granting: novelty, utility and inventiveness of invention. But also indicators based on patent statistics have some advantages and disadvantages. There is a short list of important points for patent analysis that can be also found in OECD (1994, p. 15-17):

First, patents have a close relationship to the output of industrial R&D activity and other inventive and innovative activities. Second, patent data cover a range of technological fields. For this reason, patent data can be used for analyzing the diffusion of key technologies or for generating specialization profiles for countries, regions or companies. Third, patents offer worldwide geographical coverage, as most countries have a patent system and, more important, all of them are represented in large systems like the American and European ones. A further extremely important advantage regarding patents documents is their detailed information content including the year of invention (priority year), technical classification, country of applicant, country of inventor etc., with data going back many years. In recent years, evaluation of this information for economic purposes has expanded rapidly with improved on-line availability of patent data.

Nevertheless, indicators based on patent statistics have a range of shortcomings. First of all, patents are not the only possibility of protecting intellectual property. Not all inventions are patented. Some of them even cannot be patented. There are some technological fields that are excluded from patent protection according to the European Patent Convention (Article 52)[6]. For these cases, other alternatives can be

[4] "A patent is an exclusive right granted by law to applicants / assignees to make use of and exploit their inventions for a limited period of time (generally 20 years from filing). The patent holder has the legal right to exclude others from commercially exploiting his invention for the duration of this period. In return for exclusive rights, the applicant is obliged to disclose the invention to the public in a manner that enables others, skilled in the art, to replicate the invention. The patent system is designed to balance the interests of applicants / assignees (exclusive rights) and the interests of society (disclosure of invention)" (cp. WIPO (2008, p. 10).

[5] It always depends on the underlying research question. If the focus of analysis lies on the invention activity, it is more appropriate to use patent application data. If the research question relates to economic usefulness, the data about granted patents are more adequate.

[6] There are:

(a) discoveries, scientific theories and mathematical methods;

(b) aesthetic creations;

used such as secrecy, rapid launching, and low prices and so on that can supplement or even replace patent protection. In technological areas with short life cycles patent protection can be unattractive because it can take a long time to grant a patent. Second, there are differences in the propensity to patent across firms, sectors and countries, influenced by different national patent systems as well as the patterns of international trade and direct investment. Third, there are differences in the patent values that should be considered by the analysis. The list of methodological problems can be continued. But nevertheless, patents are one of the most used data sources for construction of R&D indicators.

Regarding demarcation of science-based technologies, there are some empirical studies that use patent data. Some of them estimate science-intensity of technologies matching patent and publication data technology at the level of inventors and authors (cp. Coward and Franklin 1989). Three possible types of patent-paper intersections have been investigated:

1. individual name matches between patent inventors and paper authors;
2. institutional name matches between patent assignees and organizations listed as affiliations by authors; and
3. examiner-cited literature references found in patents and base literature papers from the model.

The authors document that author-inventor name matching was the best approach.

The second match procedure was used in order to check the results of the first approach. The third method was the less useful. But this method is used by Narin and Noma (1985) in order to identify science-based technologies. Also Pavitt (1998) writes that patent citation analysis is a less distorting indicator of science-technology linkage. For this reason it is interesting to consider this approach more explicitly.

As already mentioned, one of the three conditions that each invention needs to require in order to get patented is its novelty. By checking the novelty of a patent application the inventors and the patent office examiners prepare a list of citations of published prior art documents. This list can include other patents or “non-patent citations” (NPCs). The patents are more preferable because they describe technical

(c) schemes, rules and methods for performing mental acts, playing games or doing business, and programs for computers;
(d) presentations of information.

features more clear than scientific articles. But occasionally relevant patents are not possible to find. In this case other sources are cited. Based on the assumption that the most "non-patent citations" relates to scientific publication, science-based technologies can be defined as fields with frequent references to scientific publications.

However, this definition includes some critical points that should be taken into account. They are as follows:

– To what extent does the NPCs actually measure the "science-intensity" of patent applications?
– There are two types of references in patent documents with different contextual background.
– There are different citation policies that depend on patent office where the application procedure is in progress.

These critical questions will be discussed below. The first question is about what the NPCs really represent. There are some empirical applications that deal with this problem. Narin et al (1997) investigates the contribution of public science to industrial technology by tracing the rapidly growing citation linkages between US patents and scientific research papers. In this study 430,226 NPRs are investigated on the front pages of 397,660 US patents granted in 1987-1988, and 1993-1994. The result report demonstrates that over 70% of patent citations are linked with public science, authored at academic, governmental, and other public institutions. A more critical study is conducted by Schmoch (1993) who describes science-technology links. The study uses the NPC methodology within two main areas, namely lasers and polyimides for EPO documents. He identifies some methodology problems including the fact that not all patent citations are really linked to science. Secondly, not all NPCs are available in the databases. Some citations can be displayed, but there are limits in the searching procedure. Finally, the number of NPCs in comparison with patent citations is relatively low. For this reason it is more meaningful to investigate samples including patents with at least one citation. Based on these methological problems NPC-analysis in this study does not reveal clear results.

Narin et al. (1997, p. 318) suggest to see the linkage between science and technology as a linear process:

> "The notation that technology springs from a science-base was originally embedded in the "linear model" of innovation: from basis research

> through applied research continuing into technology and resultant economical benefit. This simple linear model has been supplanted by much more complex views of the process, with many feedback loops and major contributions of technology to science, but the origins of research knowledge in basic research still lie at the core of the process".

Other studies (for example Meyer 2000a; Meyer 2000b; Tijssen 2002) stress that it is not adequate to speak about an unidirectional linkage between science and technology reflected in patent citations. One should rather speak about interdependency between science and technology in the context of patent citations. Meyer (2000b, p. 156) characterizes science and technology as "separate and hardly communicating activities", but views "them as dancing partners". This idea can be regarded as the following model that includes two kinds of activities: exploitation and scientific exploration. Exploitation can be seen as technological development, pilot processes and feedback, and exploration to increase the understanding of causal relationships in real world.

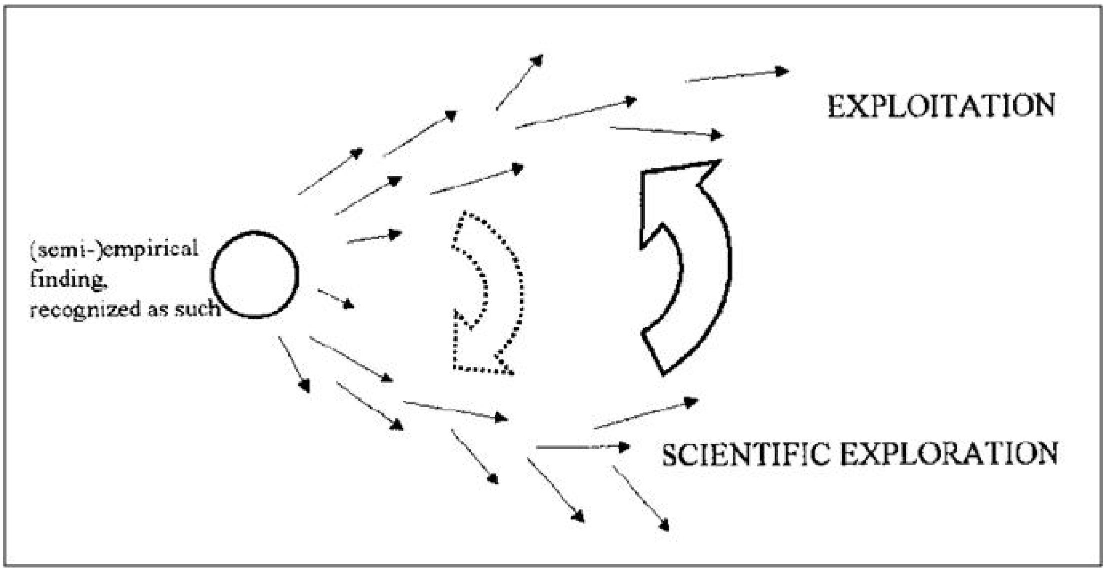


Figure 2-1: The two-branched model[7].

Furthermore, it should be taken into the account that scientific citations only include references to codified knowledge; other knowledge sources like tacit knowledge are not visible using this methodology. But to sum up, various studies confirm that the

[7] Source: Meyer (2000b, p. 157)

NPCs can be taken as good indicators of the knowledge flows between science and technology and therefore useable to identify science-based technologies.

The second question concerns different types of citations. Regarding this, different relevance category codes, also known as relevance indicators can be distinguished. They include: A - Technological background; D - Document cited in application; E - Earlier patent document; L - Document cited for other reasons; O - Non-written disclosure; P - Intermediate document; T - Theory or principle; X - Relevant if taken alone; Y - Relevant if combined with other documents[8]. Consequently, the strength of science linkages varies.

In addition, the references in patent documents can be made by examiner or by inventor resp. applicant. The examiner references are placed on the front page of the patent; in contrast, the applicant references are located in the full text of the patent document. But it is not only a formal difference because from a technical point of view the exact location of examiner references makes it easier to use this information for further analysis. The applicant references are more difficult to extract because they are distributed in the whole document. Comparing examiner citations to applicant citations, different possibilities can occur (Meyer 2000b, p. 158): regularly, examiner citations include all applicant citations, but it can also happen that any applicant citation is taken over by the examiner. But in most cases, examiner citations outnumber the applicant citations. Narin and Noma (1985) document that the share of patents with references on front pages is large enough to identify significant linkages between patented technology and science. In addition to this, these two kinds of references have slightly different backgrounds. The main goal of the examiner is to set up bounds of claims contained in the patent, whereas the applicant rather tends to outline the history, usefulness and development of the patented invention.

The third problem is that there are differences between patent offices regarding patent citation practice. Tussen et al. (2000) report that USPTO[9] patents include much more references to scientific documents as compared to patents filed through the European patent system. It results from the differences in granting procedures between the USPTO and the EPO. Within the USPTO granting process the applicants are obligated to provide an overview of all known relevant documents that

---

[8] http://www.questel.com/en/customersupport/Userdoc/DocPDF/CT_PlusPat.pdf. p. 3
[9] United States Patent and Trademark Office (USPTO)

can be either patents or other written documents. The patent examiner later decides which references are relevant for assessing the claims that have been made.

Issue	EPO	US
"Someone skilled in the art", the "average specialist".	Specialist is well educated, so for him, a less detailed description is sufficient.	Specialist is less educated, so he needs a very detailed description and many references to other documents.
Searches, search reports	It is said, that, in many cases, the EPO report is better due to a broader access to relevant material.	The quality of US researches is limited by their focus on English-language documents. US examiners are under time pressure
Education requirements of patent examiners	Generally higher than those of their US counterparts.	Generally lower than those of their European counterparts.
Claims	Focus on umbrella claims	Many claims
References	No duty of disclosure; Focus on relevant citations	Duty of disclosure: All relevant documents have to be indicated by the applicant party.

Table 2-1: Differences between the European and the US examination practices[10].

References provided by the applicant can be omitted while examiners might add references as well. In contrast to this, there is no obligation for the applicant at the EPO to generate resp. complete the reference list. Table 2-1 summarizes the differences between US and EPO examination practices.

These differences lead to quantitative differences between patent citations generated by the EPO on the one side, and the USPTO on the other Table 2-2 shows summary citation statistics from EP and US patents reported by Bacchiocchio and Montobbio (2004).

[10] Source: Meyer (2000b, p. 163)

	EP Dataset	USPO Dataset
Range of cited patents	1978-1997	1978-1997
Range of citing patent	1979-1998	1979-1998
Potentially cited patents	906,792	1,766,075
Potentially citing patents	984,148	1,734,687
Total citations	959,852[a]	8,080,276[a]
Citations per potentially citing patent	0.98	4.66
Citations per citing patent	1.86	5.59
Cited patents by fields,%[b] (potentially cited patents in parenthesis)		
Chemicals	27.45 (22.1)	17.93 (19.3)
Computers and Communications	10.58 (10.1)	17.60 (12.6)
Drugs and Medical	12.92 (9.5)	10.8 (9)
Electrical and Electronics	12.72 (13)	18 (17.5)
Mechanical	29.89 (35.3)	18.05 (21.2)
Others	6.43 (9.8)	17.62 (20.2)
Cited Patents by country,% (potentially cited patents in parenthesis, %)		
Germany	16.06 (20.1)	5.99 (7.8)
France	6.59 (7.9)	2.34 (3)
Italy	2.73 (3.2)	0.83 (1.2)
United Kingdom	7.57 (6.5)	2.64 (2.9)
Japan	21.82 (18.5)	19.6 (19.9)
United States	31.76 (29.1)	61.09 (54.7)
Sweden and Finland	2.17 (2.5)	0.94 (1.2)
Others	11.29 (12)	6.56 (9.1)
Cited Patents by institutional field,%[c] (potentially cited patents in parenthesis, %)		
not assigned	9.14 (10.6)	14.62 (16.8)
firms	87.46 (86.3)	83.93 (81.5)
non firms	3.40 (3.1)	1.45 (1.6)

a. Cells with the lag $T - t < 1$ have been removed,
b. see the Appendix for the sectoral concordance between EPCD and USPOD,
c. in the EPCD the group called 'firm' includes just companies while in the USPOD this group includes 'non government organization'. The group called 'non firm' in the EPCD includes university and public research centres while in the USPOD dataset is

Table 2-2: Statistics of EP and US patents and citation samples[11].

First of all, there are many more patents and citations at the USPTO than at the EPO in the whole data sample. Consequently, the number of citations per patent between these two patent offices varies significantly. These differences are not unexpected and reflect disparities in institutional processes underlying the citation practices discussed above. Additionally, the institutional, technological and country compositions of citations are displayed in Table 2-2. If the share of total citations of a country is higher than its fraction of total patents (in parenthesis), an above average citation intensity for that country is indicated. The same reasoning occurs for technological fields and institutional types. In the field "Drugs and Medical" in both

[11] Source: Bacchiocchio and Montobbio (2004, p.4).

databases, for example, the average citation intensity is higher than its potential citation intensity. Reverse situation can be regarded in mechanical sectors. It can be explained by the different nature of these technologies and probably by their different average patent scope. However, an unclear picture appears in Chemicals and in Computer & Communications. In the EP dataset, the field “Chemicals” has a relatively higher propensity to cite. In contrast to this, there is a remarkably higher propensity to cite in the field of “Computers & Communications”.

Bacchiocchio and Montobbio (2004) mention three reasons for this bias: One explanation is the difference in patent classification on which technological fields are built up. The matching procedure between the categories of the US National Bureau of Economic Research (NBER) and 30 technological fields is based on European IPC codes may cause some matching may cause some inaccuracies. As a consequence of this, it is possible that technological fields do not have the same size in terms of amount of patents in the different patent offices. Finally, different qualities of patents in the same fields granted by the different patent offices can lead to additional differences. Understanding all these differences between European and US patent practices has to be taken into account by selecting the appropriate data for the analysis. It makes more sense to limit the data collecting process at one patent office in order to identify science-based technology using patent citations.

To sum up, one can say that there is a significant difference between patent and literature citations. The patent citations process has rather interactive social nature including different groups of actors: examiners, applicants, inventors, patent attorneys, etc. Since patent citations have possible legal consequences for grant or refusal of a patent, the selection of patent citations is more careful compared to citations in scientific papers. There are other methodological and technical problems that make it difficult to identify definite dependency of technology from science in terms of high-tech fields. But nevertheless, there are some empirical studies that confirm that patent citation links have mediated nature. For this reason, patent citations can be taken as an indicator for identifying science-based technologies.

2.2 Measuring Concepts of the Non-Patent Citation (NPC) Methodology

In the previous subsection different approaches for identifying science-based technologies have been discussed. This subsection represents the NPC methodology for identifying science-based technologies. But prior to the implementation of the NPC approach, an appropriate measuring concept has to be chosen. There are two possibilities that can be used.

The first possibility relates to the average number of non-patent literature references per patent application in a particular technological field. This method is discussed, for example, by Grupp (1996). This measuring concept can be defined as follows:

i: index for technological field

t: time index

$P_i(t)$: number of patent applications at a particular patent office in technological field *i* in year *t*

$NPL_i(t)$: number of citations to non-patent literature references (NPL) in the patent applications in technological field *i* and in year *t*.

The mean NPL indicator corresponds to the average number of science references per patent application in a particular technological field *i*:

$$NPLM_i(t)=NPL_i(t)/P_i(t) \tag{2-1}$$

If, for example, each patent application in a technological field *i* in the sample has exactly 2 references to non-patent sources, then the indicator value on average is 2. By comparing the NPLM values of investigated technology with the NPLM values of reference unit, the specialization ratio can be calculated. This term can be denoted as the Revealed Technology Advantage (RTA) and is given by:

$$RTA_i^{(1)}(t) = \frac{NPLM_i(t)}{\sum_{i=1}^{N} NPL_i(t) / \sum_{i=1}^{N} P_i(t)} \tag{2-2}$$

A value greater (resp. less) as 1 indicates the above-average (resp. below-) number of non-patent references per patent application in technology *i* compared to the number of non-patent references in patent applications of reference unit. In this context, the RTA's values greater as 1 correspond with the above-average science-intensity of technology *i*.

The second measuring concept is based on the number of patent and non-patent references:

$NPL_i(t)$: number of citations to non-patent literature references (NPL) in the patent applications in technological field *i* and in year *t*.

$PL_i(t)$: number of citations to patent literature references (PL) in the patent applications in technological field *i* and in year *t*.

$C_i(t)$: number of citations in the patent applications in technological field *i* and in year *t*.

Analogously to (2-1) the non-patent citation share (NPCS) can be defined:

$$NPCS_i(t)=NPL_i(t)/(PL_i(t)+NPL_i(t))=NPL_i(t)/C_i(t) \tag{2-3}$$

In a similar way as in (2-2) the specialization ratio for technology *i* is given by:

$$RTA_i^{(2)}(t)=\frac{NPCS_i(t)}{\sum_{i=1}^{N}NPL_i(t)/\sum_{i=1}^{N}C_i(t)} \tag{2-4}$$

The values of both specialization indicators in (2-2) and (2-4) lie between 0 and positive infinitive. It causes some difficulties regarding the interpretation of specialization profiles of the investigated units. In order to simplify the interpretability of specialization ratios, a normalization of specialization ratios is meaningful. There is a range of possibilities for normalization transformation. One of the most common used normalization technique is the Moebius transformation function: $(r^2-1)/(r^2+1)$ with r as specialization ratio as in (2-2) or (2-4). After the normalization procedure the specialization indicator has a variation range limited between -1 and 1. The indicator value is equal to zero when the citation ratio of non-patent literature in the investigated technological field is equal to the citation ratio for patent applications in the reference unit.

Using normalized indicators calculated for different technologies one can compare its science-intensities with the science-intensity of the reference unit. The problem arises by making a statement about significance of differences between specialization profiles. The statistical properties of specialization indicators are still a research question. One exception is a paper published by Schubert and Grupp (2008) that investigates asymptotic normality of one special class of specialization ratios. Here some definitions have to be introduced. The investigated unit is the technological field. This unit has some attributes or objects. The objects here refer to references in patent documents and may take two values: 1 if it is a reference to non-patent literature and 0 otherwise. As reference unit one can take all patent applications independent from technological affiliation at a particular patent office. The specialization indicators have to fulfill some restrictive conditions:

1) The objects have a natural unit. In this case the objects are the citations in patent documents and therefore there is no problem with this assumption.
2) The enumerator and the denominator of the indicator can be interpreted as probabilities, i.e. these numbers have to be less than 1. This condition is uncritical for the specialization ratio defined in (2-4). But it is violated in (2-2). For this reason, the asymptotic normality of indicators like (2-2) cannot be proved by the approach suggested by Schubert and Grupp (2008).
3) The investigated units (here technological fields) have to be independent. This assumption is uncritical because the definition of fields based on the International Patent Classification (IPC). It can be assumed that these technologies do not interfere with each other.
4) The number of objects per investigated units (the number of citations) is non-random. When the number of objects in the reference unit increases, then the number of objects in each of the other units also increases, so that the share of objects for each unit converges to a constant. Both assumptions have a technical nature and do not have any economic background.

2.3 Application of the NPC-Methodology for Identifying Science-Based Technologies

This subsection deals with the application of two measuring concepts defined in (2-2) and (2-4) in order to identify science-intensity of some technological fields. The patent data for this analysis are extracted from the October 2007 edition of PATSTAT, a worldwide patent statistic database that was developed by EPO in 2005. To conduct this analysis, it was necessary to isolate the needed information from the given database. Therefore several assumptions are made.

First of all, it can be assumed that all patents where the patent applicants presume that their invention will be important for the European market are filed at the European Patent Organization (EPO). This constraint also eliminates the problems arising due to different citation practices discussed in 2.1. Hence, the analysis focuses only on patents that are filed at the EPO.

The investigated period covers the period 1989 until 2004. The fact that the EPO was founded in 1978 causes some restrictions in patent analysis (see also Schmoch 2007). Until the end of the 1980s, a steady transition from foreign applications at national European offices to central applications at the EPO can be observed. Currently, the number of direct foreign applications at domestic offices is negligible. The majority of foreign patents are filed through the EPO. As a consequence of this, it is more meaningful to investigate EPO patent applications since about 1990 for investigation of internationally balanced technology trends. In the time between 1978 and 1990, the EPO and the national European offices have to be regarded, and before 1978, only national offices are obtainable. But again, different citation practices have to be taken into account. All these restrictions make it useful to restrict the investigation period to the period of time between 1989 and 2004.

The next point deals with the selection of investigated technologies. Schmoch (2007) lists over 44 technologies which he identified as knowledge-based. These technologies have been taken as the basis of the following investigation. The relevant patent documents are identified by means of the IPC. Additionally, the photovoltaic solar cell technology completes the list. This technology is also the subject of a detailed investigation in the present thesis.

The time index t in equations (2-1) to (2-4) relates to the priority year of the patent application. The priority year is the year of the first application to any patent office in the world in order to safeguard priority claims as given on the EPO document. This year can deviate from the denoted year of actual application at the EPO. The priority year is often taken for patent analysis as the reference date because it represents the year in which the invention is materialized and codified (see also Grupp 1996).

Table 2-3 gives a list of 45 selected technologies[12] with the corresponding IPC reference. The fourth column defined the non-patent literature mean indicator according to equation (2-1) for the whole investigated period 1989-2004.

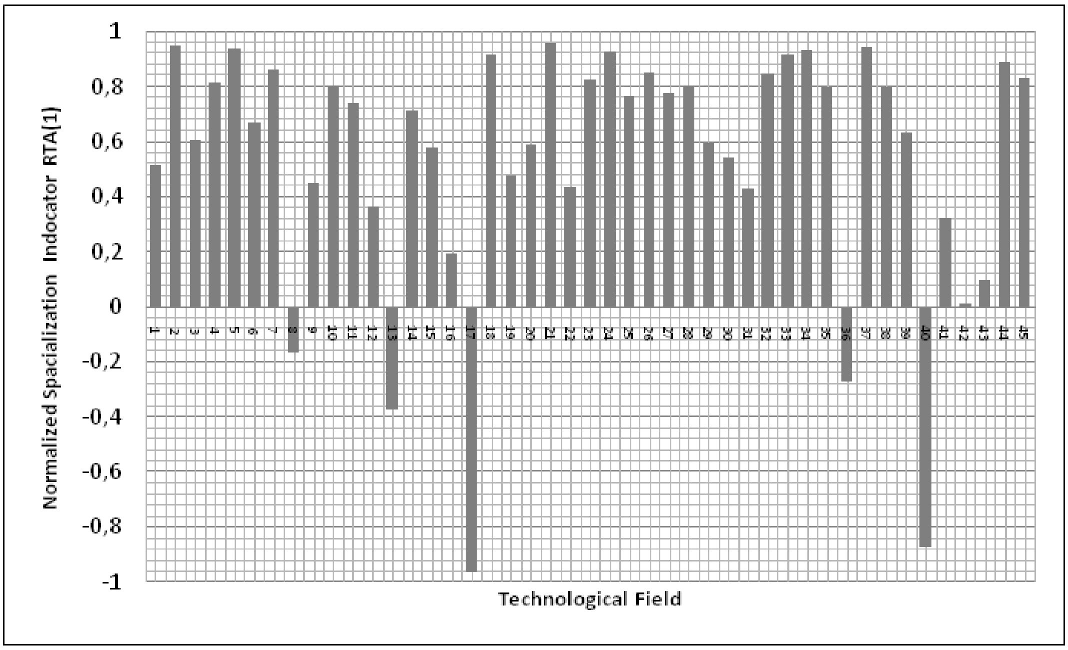


Figure 2-2: RTA$^{(1)}$values of 45 technologies, 1989-2004.

[12] 44 technologies are selected according to Schmoch (2007). The solar PV technology is added to the list.

No.	Field	IPCs	$NPLM_i$	RTA(1)
1	Recovery of poisoned soil by chemical means	B09C 1/08	0.95	0.52
2	Medical preparations containing antigens or antibodies	A61K 39	3.40	0.95
3	Biological treatment of waste water	C02F 3	1.08	0.61
4	Luminescent materials	C09K 11	1.69	0.82
5	Measuring involving enzymes or micro-organisms	C12Q	3.11	0.94
6	Micro-structure technology	B81	1.21	0.67
7	Nano technology	B82	2.00	0.87
8	Gas turbine plants	F02C	0.45	-0.16
9	Simulators for training purposes	G09B 9	0.87	0.46
10	Speech analysis	G10L	1.63	0.81
11	Fuel cells	H01M 8	1.39	0.74
12	Switching for mobile communication	H04Q 7	0.78	0.37
13	Centrifuges with free vortex	B04C	0.36	-0.37
14	Making metallic powder	B22F 9	1.31	0.72
15	Metal working by electric current	B23H	1.04	0.58
16	Plies of pneumatic tyres	B60C 9	0.65	0.19
17	Packaging fragile articles other than bottles	B65B 23	0.08	-0.96
18	Composition of optical fibres	C03C 13/04	2.57	0.92
19	Polymerisation catalysts	C08F 4	0.90	0.48
20	Materials for heat transfer	C09K 5	1.05	0.59
21	Interferons generated by generic engineering	C12N 15/19. /20. /21. /22. /23	3.84	0.96
22	Seismology	G01V 1	0.85	0.44
23	Control of optical properties	G02F 1	1.75	0.83
24	Computer systems according to biological models	G06N 3	2.80	0.93
25	Recording by optical means	G11B 7	1.49	0.77
26	Superconductors	H01B 12	1.90	0.85
27	Integrated circuits	H01L 27	1.53	0.78
28	Batteries	H01M 6	1.63	0.81
29	AD conversion	H03M 1	1.06	0.60
30	Lasers for manufacturing	B23K	0.98	0.54
31	Robotics	B25J	0.85	0.43
32	Immobilised enzymes	C12N 11	1.87	0.85
33	Single-crystal growth from vapour	C30B 23	2.64	0.92
34	Biological biocides	A01N 63	2.95	0.94
35	Diagnosis by magnetic means	A61B 5/055	1.62	0.80
36	Sorting by specific features	B07C 5	0.40	-0.27
37	Peptides	C07K	3.28	0.95
38	CVD coating	C23C 16	1.60	0.80
39	Conjugated artifical filaments	D01F 8	1.13	0.63
40	Alarm locks	E05B 45	0.14	-0.87
41	Exhaust apparatus for purifying exhaust	F01N 3	0.75	0.33
42	Combustion engines. control of fuel injection pumps	F02D 1	0.54	0.01
43	Wind motors	F03D	0.59	0.10
44	Electrically conducting glass	C03C 4/14	2.25	0.89
45	Photovoltaic solar cells	H01L 31/04. /06 or (H01L and solar*)	1.77	0.83
Reference.	All patent applications at the EPO with at least one citation		0.53	

Table 2-3: Applying of the NPC-methodology on the 45 technological fields.

Table 2-4 reports the descriptive statistics of the NPLM indicator for patent applications at the EPO.

Observations	Mean	St. Dev.	Median	Min	Max
45	1.48	0.91	1.31	0.08	3.84

Table 2-4: The descriptive statistics of the NPLM indicator.

In order to identify whether a technology can be defined as science-based or not, a reference data set has to be specified. In this case the values of the NPLM indicator for 45 investigated technologies are compared with the NPLM value for all patent applications at the EPO that meet two conditions:

1. The priority year is between 1989 and 2004.
2. Patent applications have at least one citation to patent or non-patent literature.

The last column in Table 2-3 illustrates the corresponding normalized indicator using the Moebius transformation function suggested in subsection 2.2. Regarding the results in Table 2-3, 40 from 45 technologies can be identified as science-based ones because in 40 cases the NPLM value is greater than the NPLM value for the reference data set (0.53) respectively positive values in $RTA^{(1)}$ (see also Figure 2-2). Table 2-5 confirms differences in means between the NPLMs of the investigated technologies and the NPLMs of patent applications in the reference set ($p<1\%$).

Sample: 1 45
Included observations: 45
Test of Hypothesis: Mean = 0.533792

Sample Mean = 1.482889
Sample Std. Dev. = 0.912904

Method	Value	Probability
t-statistic	6.974158	0.0000

Table 2-5: Equivalence testing of 45 technologies in relation to the NPLM-values.

The weakness of the calculation according to equations (2-1) and (2-2) is that it is not possible to make a statistical statement about the accuracy of the calculated indicator values. Furthermore, a clear differentiation between strong, weak and average science-intensive technologies is also not realizable. Grupp (1996) suggests a pragmatic categorization of strong and weak science-based technologies. Thereby, the splitting based on the observation that the values of the normalized specialization indicator of type (2-2) are not distributed evenly over the whole interval between [-1;1]. According to this strategy, two technologies can be categorized as weak science-intensive technologies, namely “packaging fragile articles other than bottles”(No.17) and “alarm locks” (No.40). Nevertheless, this categorization is somewhat arbitrary because the boundary between strong and weak classes regarding science-intensity cannot be determined statistically.

The approach suggested by Schubert and Grupp (2008) results in several improvements within this situation. According to this approach, one can compare science-intensities of the investigated technologies with the science-intensity of the reference unit using asymptotic normality of $RTA^{(2)}$. The standard deviation of the calculated normalized specialization indicators $RTA^{(2)}$ gives an impression about its exactness.

The calculation of $RTA^{(2)}$ (compare Table 2-6) implies the following results: 30 of the listed technologies have significant higher values of science-intensity compared to the science-intensity of all patent applications at the EPO. Three technologies comprise an average science-intensity, and the science-intensity of 12 technologies is lower than the average. However, the photovoltaic solar cells technology (No. 45 in the list) is identified as technology with high science-intensity regarding both calculations (2-1) and (2-3). These results are important for the ongoing analysis.

There are significant differences between two measuring concepts? Figure 2-3 shows strong monotonous dependency between ranks of 45 technologies after its sorting by score values of $RTA^{(1)}$ and $RTA^{(2)}$. In order to compare the rank results achieved using $RTA^{(1)}$ and $RTA^{(2)}$ calculations, the Pearson rang correlation coefficient is calculated. The Pearson rang correlation coefficient demonstrates a strong significant relationship (r=0.97, t-statistic=31.32) between two approaches (see also Figure 2-3).

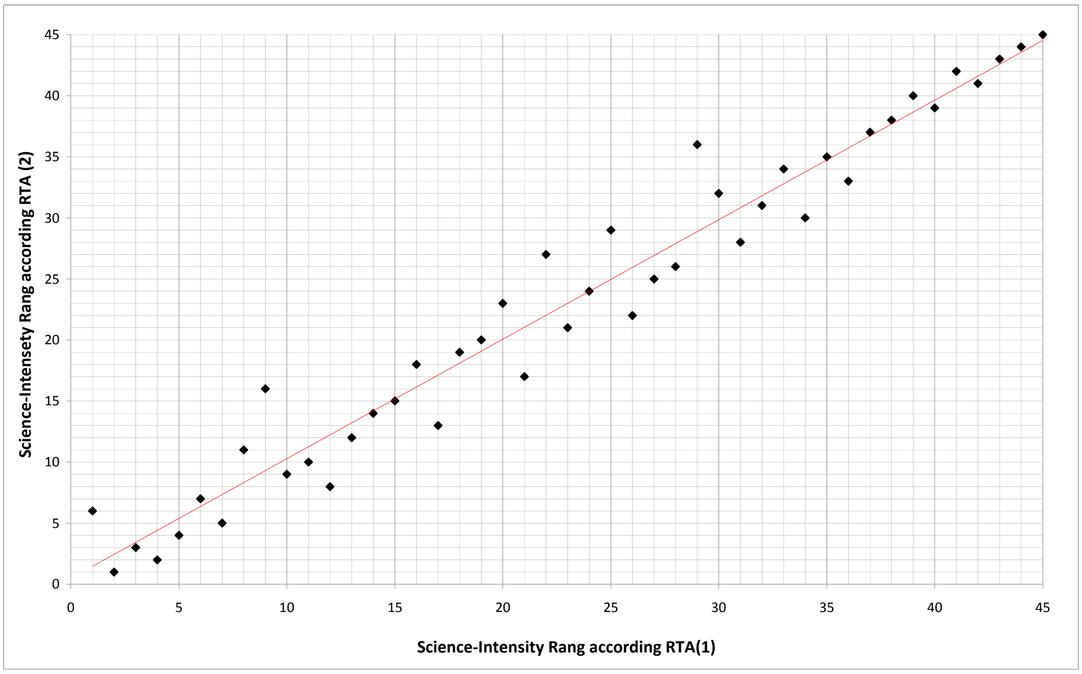


Figure 2-3: Comparison of the science intensity of 45 technologies.

To sum up, the characterization of a technology as science-based and the measurement of the science-intensity degree always depend on exact definition and measuring concept. Unfortunately, a lot of empirical studies do not apply exact measurements of technological dependency on science. The difficulties in demarcation of science-based technologies are surely a reason why explanations of development patterns for science-intensive technologies are still an open research question. By identifying science-based technologies, the researchers frequently go back to their own technological experience and do not try to define this special kind of technologies exactly. Admittedly, it is sometimes difficult to measure technological science-intensity. The exemplary application of this chapter demonstrates how to measure technological science-intensity using patent data.

No.	#NPC	#PC	#C	NPCS	$RTA^{(2)}$	Std. Dev.of $RTA^{(2)}$	Grade of Science-Intensity
1	157	716	873	0.18	-0.02	(0,07)	Average
2	30388	20680	51068	0.60	0.82***	(0,001)	High
3	1593	6258	7851	0.20	0.10***	(0,02)	High
4	2732	6219	8951	0.31	0.47***	(0,013)	High
5	28202	26666	54868	0.51	0.77***	(0,002)	High
6	1437	4859	6296	0.23	0.21***	(0,02)	High
7	460	775	1235	0.37	0.61***	(0,02)	High
8	1092	12569	13661	0.08	-0.68***	(0,02)	Low
9	301	1562	1863	0.16	-0.13***	(0,05)	Low
10	6489	10698	17187	0.38	0.62***	(0,01)	High
11	3865	14326	18191	0.21	0.14***	(0,01)	High
12	9499	39856	49355	0.19	0.04***	(0,01)	High
13	135	1803	1938	0.07	-0.75***	(0,04)	Low
14	742	2142	2884	0.26	0.32***	(0,03)	High
15	739	2533	3272	0.23	0.20***	(0,03)	High
16	621	4762	5383	0.12	-0.44***	(0,03)	Low
17	6	345	351	0.02	-0.98***	(0,01)	Low
18	517	988	1505	0.34	0.55***	(0,02)	High
19	3661	15505	19166	0.19	0.04***	(0,01)	High
20	743	3306	4049	0.18	0.00	(0,03)	Average
21	146	97	243	0.60	0.83***	(0,02)	High
22	542	2476	3018	0.18	-0.03	(0,04)	Average
23	16335	36337	52672	0.31	0.48***	(0,01)	High
24	1720	1104	2824	0.61	0.83***	(0,005)	High
25	9590	32341	41931	0.23	0.21***	(0,01)	High
26	925	1225	2150	0.43	0.69***	(0,01)	High
27	12040	32259	44299	0.27	0.37***	(0,37)	High
28	2157	6597	8754	0.25	0.28***	(0,02)	High
29	1472	4621	6093	0.24	0.26***	(0,02)	High
30	7167	32824	39991	0.18	-0.03***	(0,01)	Low
31	1693	8292	9985	0.17	-0.08***	(0,02)	Low
32	807	1570	2377	0.34	0.54***	(0,02)	High
33	659	642	1301	0.51	0.77***	(0,01)	High
34	2116	2197	4313	0.49	0.75***	(0,01)	High
35	1985	4676	6661	0.30	0.45***	(0,02)	High
36	254	2741	2995	0.08	-0.65***	(0,03)	Low
37	60257	40653	100910	0.60	0.83***	(0,001)	High
38	5398	13918	19316	0.28	0.39***	(0,01)	High
39	393	1503	1896	0.21	0.12***	(0,04)	High
40	14	549	563	0.02	-0.96***	(0,02)	Low
41	2709	18620	21329	0.13	-0.36***	(0,02)	Low
42	87	696	783	0.11	-0.47***	(0,08)	Low
43	367	3371	3738	0.10	-0.56***	(0,03)	Low
44	18	30	48	0.38	0.61***	(0,12)	High
45	1354	3445	4799	0.28	0.40***	(0,02)	High
Reference	**999485**	**4422852**	**5422337**	**0.18**			

Table 2-6: Science intensity of 45 fields according to the $RTA^{(2)}$.

2.4 Double Boom Cycles of Science-Based Technologies: Theory and Empirical Studies

By trying to set up a clear distinction of science-based technologies that have been discussed above, one can see the first step in order to understand the development pattern of these complex technologies. The importance of science-based technologies is undoubted and is discussed in the introduction. However, the question of typical chronological devolution of these complex technologies remains unexplained.

The following section gives an overview about existing theories and empirical studies that deal with the explanation of a typical course of technology cycles. Thereby, a special focus lies in the context of knowledge-based technologies.

There are several suggestions of technology cycles in the literature. Technologies differ by sectors, complexity, and consequently, technology cycles differ as well (cp. Schmoch 2007). Additionally, the aggregation levels with corresponding measuring concepts and temporal extent of the analysis have to be taken into account. The present review of existing technology cycle theories has the following structure. At first, the long-term perspective with regard to the whole industry is discussed. Then the theories regarding meso-level and micro-level of technological development are outlined. Particular attention is given to “science-push” and “market-pull” models that are often applied to describe stimuli of technological development. Concluding, some empirical studies dealing with development of particular science-based technologies are presented. The results are summarized as a descriptive stylized model suggested by Grupp (1998).

The conventional long-wave theory suggested by Kondratiev (1925) is based on the assumption that there is a link between economic development and the rise and fall of technologies. According to the Kondratieff wave theory, capitalist economies peaked and crashed in regular waves that lasted on average 54 years. These waves are centered on emerging industries. Hirooka (2003) gives an overview of studies that recognize the contribution of technological innovations within economic development. In this paper he demonstrates the most probable positioning of Kondratieff-waves as shown in Figure 2-4[13]. The first wave (1789-1825) comprises the period of the first Industrial Revolution; the economic development during this

[13] Kondratieff's study covered the period 1789 to 1926. The later time periods are completed by other scientists.

period can be interpreted as the result of the development of the textile industry. The second cycle was induced by the development of railways and iron production. Between 1900 and 1929 (rise of the third wave), the USA took over the economic leadership, jumping from a developing country to an industrialized country. In this time, the USA introduced a range of important technological innovations, such as in steel production, oil drilling, and the generation of electric power. After World War II started the fourth cycle with various innovations, such as TV, aircraft, petrochemicals, and computers. Figure 2-4 shows that the economic cycle and technology cycle run parallel and have a typical non-linear wave-like pattern.

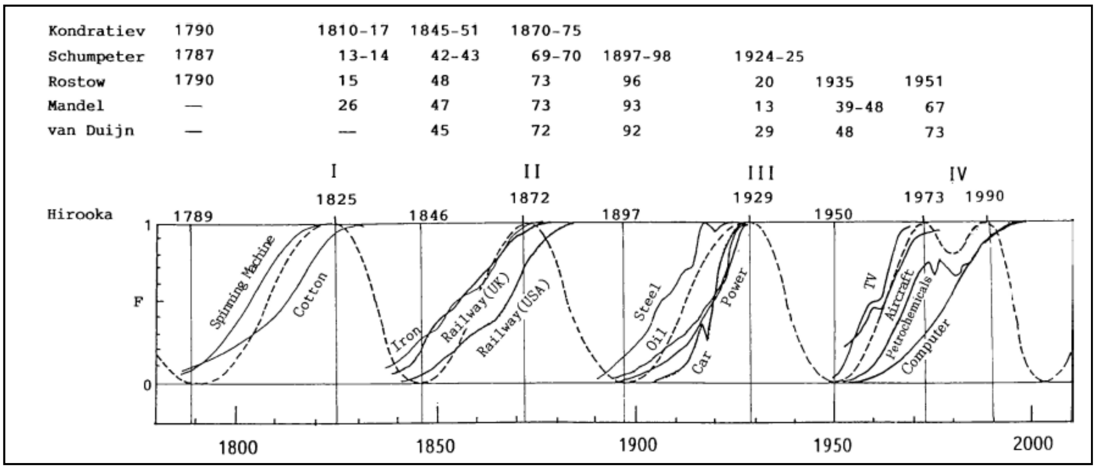


Figure 2-4: Kondratiev's business cycles and the diffusion of innovations[14].

The Kondratiev concept of long waves was taken up by Schumpeter (1939) in his study of classic study of business cycles. According to Schumpeter (1939), an innovation process is the source of fluctuations in business activity. To be more precise, the discontinuous nature of entrepreneurial activity causes instabilities within the capitalist growth process. Various studies have expanded the range of possible cycles, finding longer or shorter cycles in the data. A summarizing of these theories can be found in Freeman (1983).

The majority of economic theorists confirm Schumpeter's emphasis on the role of technological progress in explaining economic growth. However, an investigation of real contribution of particular technologies including the reasons for its development

[14] Diffusion of innovation products —; Kondratiev's business cycles – – –. Source: Hirooka (2003), p. 557

requires a more disaggregated level of analysis. The next level in disaggregation level is meso-level[15] that analyzes the development of particular technologies.
A seminal work regarding meso-level technology is written by Dosi (1982). Dosi (1982, p. 154) defines technological trajectories as

> "a "cylinder" in the multidimensional space defined by technological and economic variables. Thus, a technological trajectory is a cluster of possible technological directions whose out of boundaries are defined by the nature of the paradigm itself".

Dosi's model explains continuous changes and discontinuities in technological innovation that are determinate not only by subsequent implementation stages, and economic and institutional factors but also by fundamental research. Although the model tries to establish a general framework of technological development with by taking into account a range of important factors, Dosi (1982) does not explain typical patterns of technological development in time.

Schmoch (2007) lists a number of publications dealing with cyclic technological developments. Most of them address sales statistics of technology-related products and describe the diffusion of products. In the traditional innovation and diffusion research, one discerns various phases, such as market introduction, market growth, market saturation and decline, sometimes with a strict separation of the innovation and diffusion part. In so doing, one arrived at simple exponential or logistic relations with slow growth in the beginning, a turning point and low growth due to ceiling effects at the end (the so-called "S-curve concept").

The traditional symmetric S-curve concept is further developed by Bass (1969). By subdividing the diffusion process into internal factors (for instance mouth-to-mouth-propaganda, social pressure in society) and external factors (like marketing activities) the diffusion process is modeled in the form of an asymmetric S curve. In the ideal case, even in this model after passing a certain threshold of critical mass, exponential growth is assumed. This model is criticized by the management and organization literature (see Höft, 1991). Even in the case of conventional consumer markets the market mechanisms are much more complex. It is the present state of the art to investigate the mutual relations between growth of product variety and overall market growth in selected sectors of the industry (for instance the US shoe market;

[15] Additionally, one can analyze micro-level of technology that refers to individual products or processes (cp. Schmoch 2007, p.1002).

see Baudisch 2005). First results point out that the dynamics of demand cannot be explained by the postulation of static consumer preferences. Lindqvist (1994) lists other defects of the S-curve concept and thus the actual scientific literature does not any more support this concept anymore.

The dynamics on the time scale of innovation processes in science-based markets has not been explored to a great extent. There are some empirical investigations which deal with the question of market formation in science-based sectors (Schmoch, 2007). The empirical access to study these special markets often uses one or several of the following indicators:

- Measurement of scientific activities by publication statistics (van Raan, 1997)
- Measurement of technological development by patent applications or patent grants, respectively, and
- Measurement of installed or sold (shipped respectively) products to grasp diffusion.

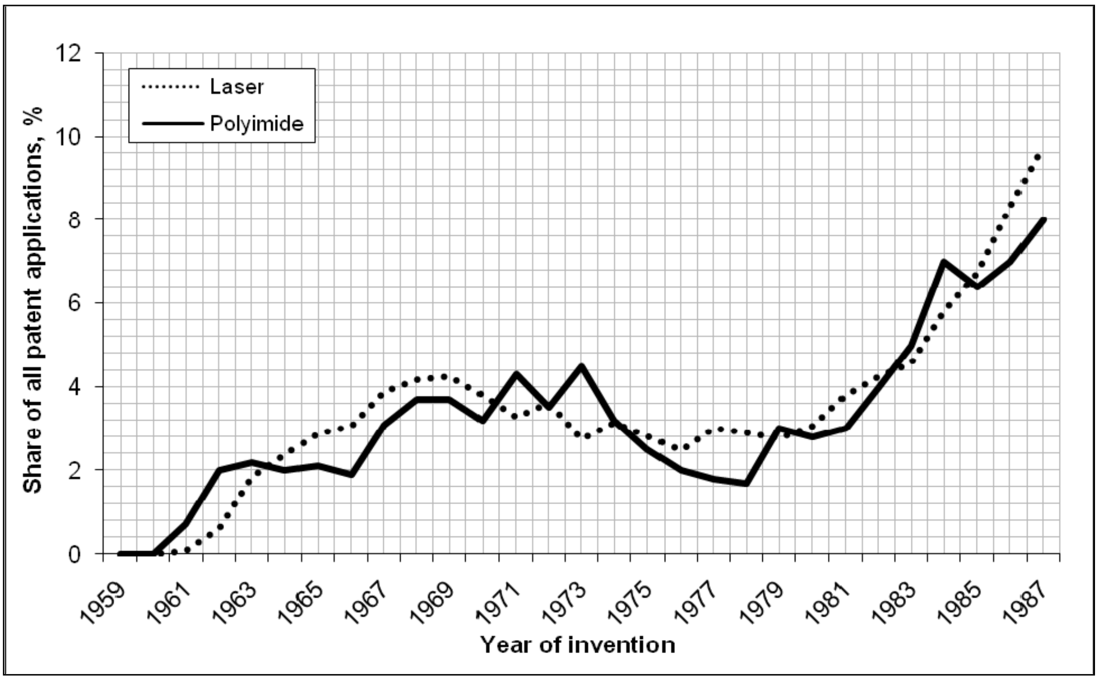


Figure 2-5: Development of patent applications for lasers and polyimides[16].

Grupp and Schmoch (1992) already studied patent applications for polyimides and lasers already at the beginning of the 1990s. Figure 2-5 shows a characteristic of non-linear patterns with two maxima. In the cases of lasers and polymides, the basic

[16] Source: Grupp, Schmoch (1992), p. 278.

scientific theories, the first wave of activities, and the final market-driven growth span a period of fifty years or more (Grupp and Schmoch, 1992). In the case of the laser market, the number of scientific publications was very low in the first years of activity. Yet, with an increasing number of patents also the number of scientific publications grew. From this observation, it is concluded that scientific activities not always precede technological development, but – due to intensive interaction in the scientific community – science and technology are intertwined.

Parallel to the observations by Grupp and Schmoch (1992), Rickerby and Matthews (1991) describe what they called the "technological commercial exploitation curve" for surface engineering (Figure 2-6). Their description is not supported by quantitative data, but is based on qualitative experience of engineers. Striking is the similarity of this qualitative experience with the indicator-based patent curves for lasers and polyimides. In both cases, a characteristic dynamics with two discernable phases of development may be recognized.

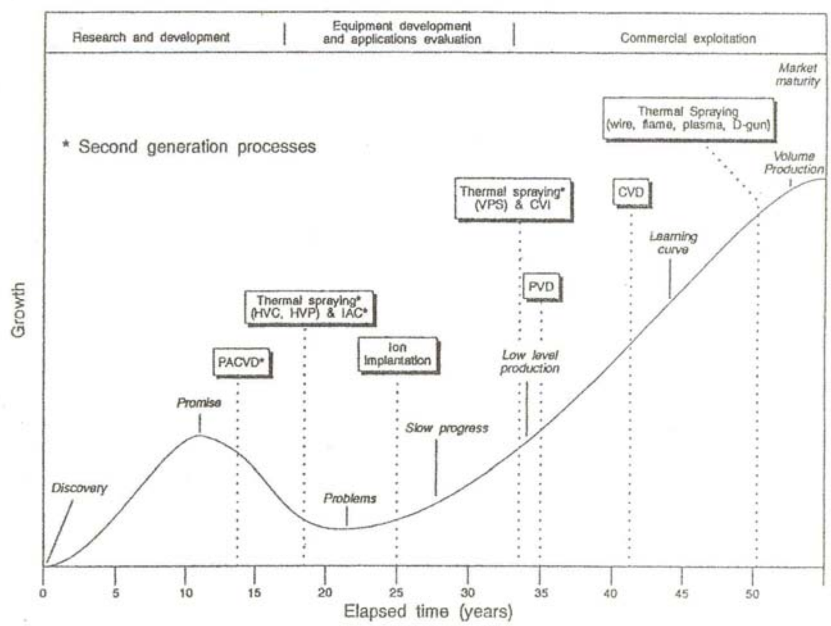


Figure 2-6: Technological and commercial exploitation of surface technologies[17].

The development of some emerging technologies is also illustrated by Gartner consultancy[18], giving a graphical modeling of the maturity, adoption and business application of specific technologies. This hype cycle approach high-lights the progression of an emerging technology from market over enthusiasm through a

[17] Source: Rickerby, Matthews, 1991, p. 347.
[18] www.gartnergroup.com (12.12.2008)

period of disillusionment to an eventual understanding of the technology's relevance and role in a market or domain. Technologies are described in terms of visibility and maturity. The dimension of visibility does not offer any clear differentiation between technology and market development but concludes both kinds of activities (Figure 2-7).

According to Gartner's Hype Cycle graph, handwriting recognition, software as service and location “aware applications have reached the bottom of the trough and are starting to climb into the "slope of enlightenment". In this phase, the majority of consumers, not just the early adopters and technology enthusiasts, start to see the benefits of the technology and become more educated.

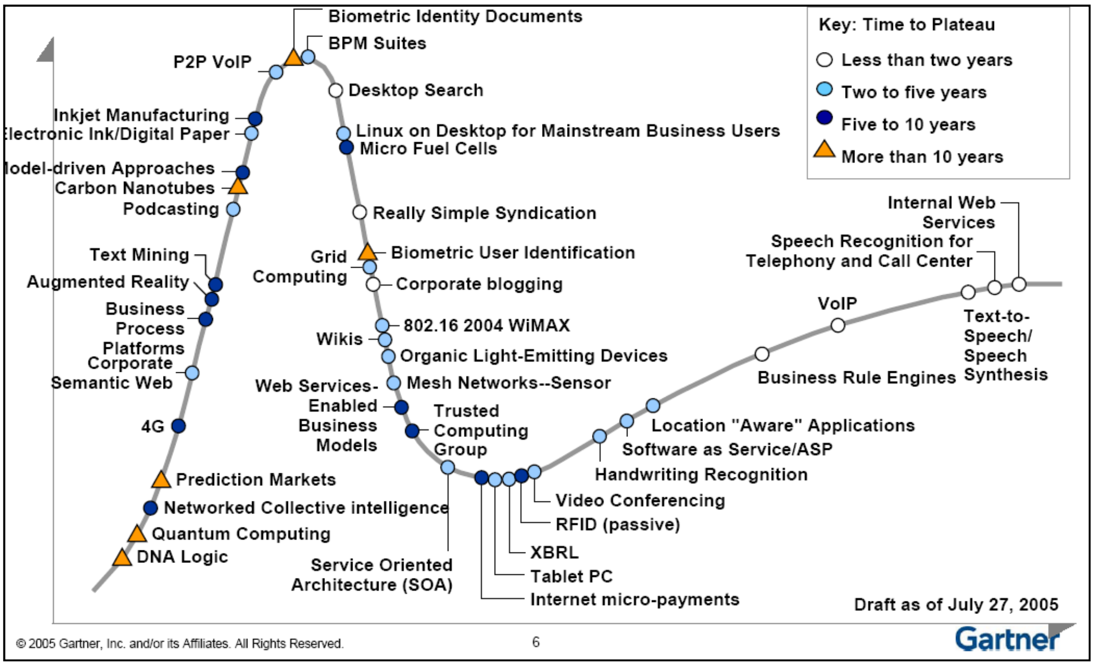


Figure 2-7: Hype Cycle for emerging technologies, 2005[19]

Taken all together, the collected data confirm that there are common patterns in the development path of science-based technologies. This finding is not self-evident. Laser, polyimides, technologies for surface engineering, genetic engineering, pharma-ceuticals, emerging technologies of hype cycle represent quite different technologies in relation to their scientific background, market size, industrial application, etc. The striking similarity regarding the development of these technologies seems to be unexpected. The reason for the “striking correspondence” is the science-intensive nature of both technologies, on the one hand, and potential of

[19] http://www.gartner.com. Accessed on 17.01.2009.

numerous practical applications, on the other hand (see also Grupp and Schmoch, 1992, p. 282). Stokes (1997) pools such technologies together as "Pasteur's quadrant". Although research in this quadrant has potential real-world utility, its investigators never lose sight of the desire to advance scientific understanding.

Two main stages of market formation of “Pasteur’s” technologies can be distinguished. In the first phase "voice of the market" is largely absent and the development goals are oriented towards internal success within scientific communities (Hekkert et al., 2007). The misunderstanding or, better yet, non-interaction between the side of science and the demand side is largely due to intellectual, but also normative differences; questions of safety, standardization, and compatibility are often neglected (Blind, 2004). Although some of these differences may be larger in perception and rhetoric than in reality, they tend to lead to stagnation, and thus, cause the end of the first maximum of activities as observed in the empirical studies. Consequently, the emergence of innovations is seen as "driven by individual genius". If radical innovations had immediately a better price-performance ratio, the substitution on consumer markets would be a simple matter (Geels, 2006). But radical innovations are usually born as "hopeful monstrosities" (Mokyr, 1990), i. e. as interesting and promising ideas with crude performance. Much work is needed to make radical innovations technically and economically viable. Small market niches as "incubation rooms" are also essential to protect their early development (Geels, 2005). Technical feasibility is not the same as product development and introduction on the market.

This stage of development can also be described by the “science-push” model which is explained by Bush (1945). This model was very popular in the 1960s. The “science-push” assumes a linear structure of innovation processes from department to department, starting with scientific discovery und underlines thereby the importance of basic research in the innovation process. According to this model, scientific discoveries are more or less unexpected; at a future date technologists find an application for these inventions, and then, engineers and marketing specialists convert them into products and bring them on the market. The atom bomb can be given as an example of a “science-push” model. Another example: the emergence of biotechnology as a new industry is more a consequence of “science push” efforts because numerous inventions are often undertaken in the absence of clearly defined goals for biotechnology. Nevertheless, one can see that biotechnology has a quite broad spectrum of applications in the consumer and food industry, agriculture, medicine, etc.

Obviously, some scientific discoveries cannot break the deadlock after a first maximum of activities (Scherer, 1986). But it may also happen that further improvements and investigations open a bridge towards demand and consumer preferences after a while. The second stage of market formation begins. This regime shift may give rise to a wave of activities and, indeed, in aforementioned case studies this was always observed. Among the numerous factors that work against the introduction and diffusion of technologies, Kemp et al. (1998) mention technological factors, government policy and regulatory frameworks, cultural and psychological factors, demand and production factors, infrastructure and maintenance, as well as undesirable social and environmental effects on new technologies. Other barriers include high investment costs, “split incentives”, lack of awareness of potentials by customers as well as by policy makers and so on (Philibert 2006). If the scientific and technological potentials of new technology fit with the demand side, market introduction and diffusion may take place. Here, socio-technical alignment is required, where economics, politics, consumer circles, and aspects of quality of life play a role. But the topic of user preferences is underdeveloped in economics; what happens on the demand side remains mostly a black box.

The needs of customers find a better consideration in the “demand pull” model of innovation processes. This model is also known as “market pull” model that assumes the innovation process as a linear sequential process, but at the beginning of the innovation stands “market demand”. The market needs can be clearly articulated by either consumers or an entrepreneur find out a gap in the market. Schmookler (1966) investigates systematically the impact of market demand on technological progress regarding different American industries, such as railway, building, and oil refinery industries over 150 years. His results are:

> “Despite the idea that scientific discoveries and major inventions typically provide the stimulus for inventions, the historical record of important inventions in petroleum refining, paper making, railroading and farming revealed not a single unambiguous instance in which either discoveries or inventions played the role hypothesized. Instead, in hundreds of cases, the stimulus was the recognition of a costly problem to be solved or a potentially profitable opportunity to be seized.” (Schmookler 1966 p. 199)

Which of these models also seems to be more close to reality? Kostoff R. N. (1997) gives an overview of some widely known science-technology evolution case studies which consider the role of science and market needs in innovation processes.

Project Hindsight was an in-depth study which was funded by the Defense Department (DoD (1969)) and identified factors which were important for the effectiveness of research and technology programs in development of approximately twenty weapons systems. The most significant finding was that the results of research were more likely to be applied when the researchers had more attention for the needs of the application engineers. In contrast, the basic research results were less important for the ongoing development of successful weapon systems. This conclusion supports rather the arguments of “demand pull” models.

Another study, namely TRACES (Technology in Retrospect and Critical Events in Science) conducted by Illinois Institute of Technology Research (IITRI 1968) came to opposite conclusions. Within the frame of TRACE key events which had led to five major technological innovations were investigated. The study showed that non-mission research provided the origins from which science and technologies could advance towards innovations. However, a couple of years later, the National Science Foundation (NSF) sponsored Battelle-Columbus Laboratories to conduct a follow-up study to TRACE (Battelle 1973). This study covered 10 civilian innovations which were traced back 35-50 years back. The Battelle report refined the analysis of IITRI 1968, by, for example, identifying “decisive events”, and by stressing socioeconomic and managerial factors (Walsh 1973). There is a short summary of the results (Smith 1987): 15% of decisive events were classified as non-mission research, 45% as mission oriented research, and 39% as development and application work. Concerning factors which affected the rate and direction of innovation, in 87% of decisive events a technical opportunity was important (supporting science push model), while in 69% recognition of a need was important (supporting market pull model). A very good overview of key research papers, reports and studies which compare the contribution of basic and applied research can be found in Wooding (2007).

The linear logic of both models which has been introduced above does not incorporate any feed-back mechanisms and overlook multi-directional information flows between researchers, manufacturers, and marketers. From the early 1970s until the mid-1980s the combination of the previous models became widely accepted. The structure of innovation remained essentially linear and logically sequential. The innovation process is not necessarily continuous and can be divided into a series of functionally distinct but interacting and interdependent stages (Rothwell and Zegveld 1985, p. 50). Numerous feedback loops and persistent interactions complete the model.

Schmoch (2007) assumes that science-based technologies pass through two different development stages: the first stage corresponds rather with the assumption of the "science-push" model, and the second conforms to the "demand pull" model. This assumption is based on empirical investigation of 44 technologies that science dependency is investigated in 2.3. . However, 22 of these fields developed according to a strong double boom pattern, for a further ten fields at least a weakly formed double boom can be considered. Two fields seem to be in the first period of development and only the remaining ten fields indicate no double boom courses, but continual increased scientific development. Still not all science-driven technologies are subject to this wave development, but for quite a large share of these technologies this pattern seems to be appropriated. A standardized reference scheme of the formation of science-based markets (Figure 2-8) summarizes empirical findings and provides the basis for the ongoing analysis.

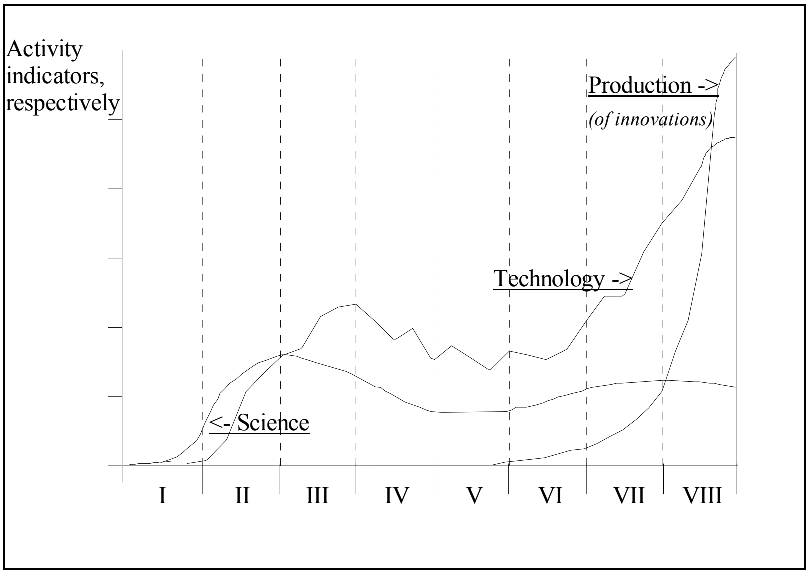


Figure 2-8: The standardized reference scheme of the formation of SB-markets[20].

[20] Three different types of indicators are used: publications (for science), patent applications (for technology) and installed or sold (shipped, respectively) products (for production). Source: Grupp (1998), p.34.

Grupp (1998) distinguishes 8 phases in formation of science-based market:

I: First explorations in the scientific domain.

II: Properly developed science; first technical achievements.

III: Science fully developed; technology still capable of extensions; prototypes.

IV: Difficulties discernible in economic transposition.

V: Temporary stagnation in science and technology; reorientations.

VI: Industrial R&D envisages new possibilities; but still capable of expansion.

VII: First commercial applications; industrial R&D fully developed.

VIII: Penetration of all markets; importance of R&D waning relative to turnover.

The present state in this research area provides us with a lot of studies on either the science push or the demand pull side. Most of these papers are of a qualitative nature. It is the challenge of this thesis to find out which main factors have an impact on the technological development. One can assume that the evolution of technology is influenced by different exogenous factors such as the political, social, economic environment, but also by endogenous factors like technological, institutional conditions and so on. The basic hypothesis is: there are two quite different development phases in the development of science-based markets. These phases differ from each other by sets of determinants influencing its formation. The main aim is to identify these determinants and to measure the impact of these factors in different stages of market development empirically.

3 Case Study – Solar Photovoltaic Market

In order to get a more detailed insight in the nature of technology cycle by science-based technologies the technological field "solar photovoltaic cells" will be analyzed more deeply. This section briefly describes the history of the PV technology. Additionally, the development of the German PV Market is presented. Finally, the analytical framework of the PV case study is discussed.

3.1 History of the PV Market

In 1839, Edmond Becquerel discovered the photovoltaic effect[1] that can be described as a physical process in which a solar cell converts sunlight into electricity. While experimenting with an electrolytic cell made up of two metal electrodes, Becquerel observed that certain materials produce small amounts of electric current when exposed to light. However, he could not give an explanation for the physical nature of this process. It took a couple of decades until Albert Einstein explained this phenomenon in 1904[2] and gave a better understanding of photoelectricity. In that time, very low efficiency of solar cells made it impossible to see sunlight as a power source. E.D. Wilson of Westinghouse Electric's photoelectricity comments the efficiency of solar cells in such way:

> "The photovoltaic cell will not even prove interesting to the practical engineer until the efficiency has been increased at least fifty times" (see Perlin, 2002, p. 20).

For a long time the photovoltaic effect was only a scientific phenomenon with few device applications. It took further 50 years until scientists at Bell Laboratories developed the first crystalline silicon photovoltaic cell that had an efficiency of 4%. This efficiency was soon increased to 6% and then to 11%[3]. The era of power-producing cells has begun.

---

[1] While nineteenth-century scientists called the process photoelectric, by the 1920s it was renamed as the photovoltaic effect (Perlin, 2002, p. 20).

[2] For his theoretical explanation Einstein was awarded a Nobel Prize in 1921.

[3] Zweibel and Hersch (1982, p.7)

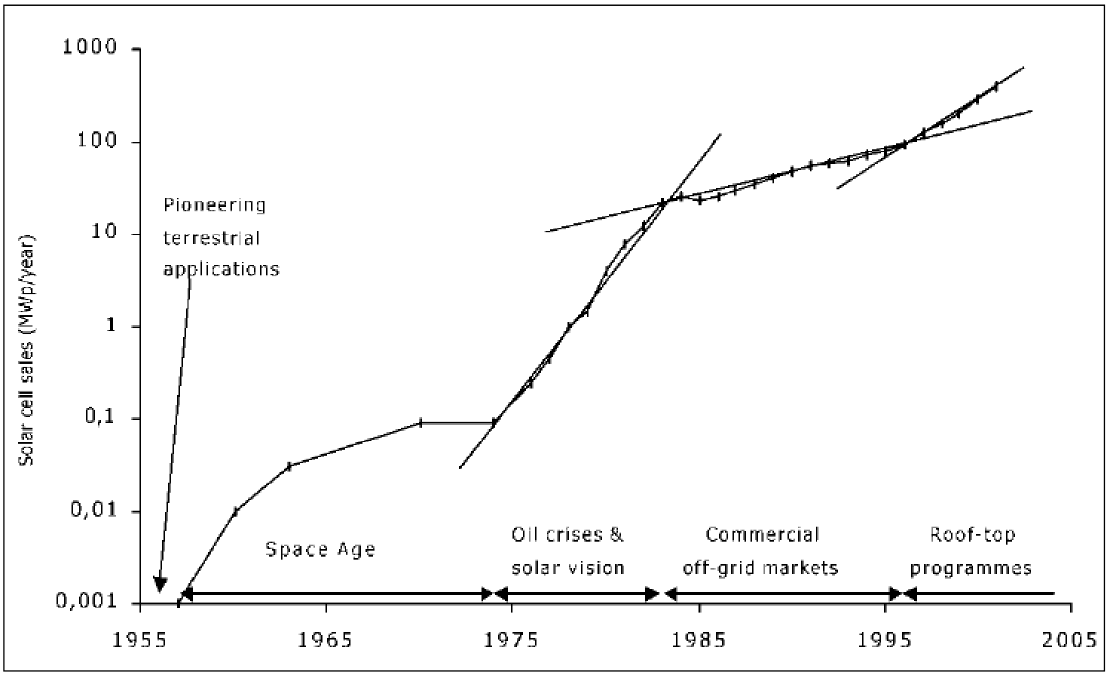


Figure 3-1: The five phases of solar cell diffusion[4].

In general, five periods in the history of the PV technology can be distinguished (see Figure 3-1). In the late 1950ths, there were some pioneering applications in regions being geographically isolated from electric utility lines; anyway, sales remained extremely small. One of the first important applications of solar cells was in space on board of the Vanguard satellite[5]. In the late 1960s, the satellite market was the first significant commercial market with an annual production of about 0.1 MWp[6] per year. Many Soviet and American satellites used solar PV sells as a source of power. However, the terrestrial use was limited due to the high cost of solar cells. Despite a price drop of about 65% between 1956 and 1971, solar cells still cost $100 per watt[7]. The high requirements of aerospace industry in terms of technical accuracy set high standards on reliability of solar cells. If for any reason the cells did not work, not only the mission would be lost, but also millions of dollars spent for equipment.

The great wake-up call came with the oil price crisis in 1973 that stimulated a rapid growth of the PV sector during the 1970s and early 1980s. Increasing production and

---

4 Source: Jacobsson et al. (2004) p. 8.

5 The Vanguard space satellite used a small (less than one-watt) array of cells were applied to power its radio system (see Zweibel and Hersch, 1982, p.8).

6 Megawatt peak (MWp): Unit of power: 1 megawatt = 1,000 kilowatts. The peak watt is the unit used to measure the standardized power output of a photovoltaic cell. It corresponds to one watt of electrical power under standard test conditions.

7 See Perlin (2002, p. 50).

economies of scale brought the price of the PV energy down from 100 $/watt in 1971 to 7.6$/watt in 2007. Also the second oil price crisis resulted in a further rapid increase in solar cell production.

However, steadily declining oil prices during the early 1980s led to a reduction in the funding of photovoltaic R&D. These factors slowed down the growth to approximately 15% per year over the time period 1984 to 1996[8]. A new growth phase started in 1997 which, among others, can be attributed to a range of market formation programs in developed countries.

8 World Photovoltaic Annual Production, 1971-2003. Earth Policy Institute: http://www.earth-policy.org, accessed 17.12.2008.

3.2 Development of the Market for Solar PV Modules in Germany

Germany has a worldwide dominant position regarding installed solar PV capacity, and, in 2005, passed Japan that had the top position in the solar technology sector before. In 2007, it was the largest single market with 1,100 MWp (see Jäger-Waldau 2008, p. 93). This strong position can be explained by the German government support, but also by the efforts of the European Union to increase the share of renewable energies in use within the EU[9]. Therefore, investigation of German PV market is meaningful[10]. The following subsection describes the main milestones in the German PV development in the time period 1975 to 2005.

The market development can be divided into two phases: 1975 until 1990, and the time after 1991[11,12]. The empirical argumentations of this partition will be given in the following chapters. This partition also arises from changes in the promotion policy of German government. The first phase ("science-push") can be characterized by the support of knowledge generation in field of terrestrial solar cells. The promotion policy in the second phase ("demand pull") focuses on stimulation of the demand side. The main goal of the German government in this period was to send positive signals to the market.

In the period 1975 to 1990, according to the Public Promotion Catalogue of the Federal Ministry of Education and Research (FöKat[13]) 30 universities, 18 research institutes, and 50 firms received Federal funding in the technological field of solar cells. In this period a close connection between academic knowledge with industrial knowledge began to arise. There were two Energy Research Programs in that time. The first program "Energy Research and Energy Technologies Program 1977-1980" focused on building-up scientific capacities on universities and extramural research. Within the second program (1980-1990) the first laboratory production was started. Thereby, the financial support of Federal government was concentrated on promotion of research activities in renewable energy sources. There were also some demonstration projects. However, these projects did not have a significant impact on

9 In March 2007 a binding EU-wide target to source 20% of their energy needs from renewable energy sources is signed by European leaders.

10 This geographical limitation is given up by investigation of co-authors networks in chapter 6.

11 The time before 1975 is excluded from the analysis because the most official data does not cover the time before 1975.

12 Jacobsson et al. (2004) suggest also a similar separation.

13 http://foerderportal.bund.de/foekat/jsp/StartAction.do, accessed 17.12.2008.

the market formation in general, but they helped to collect application knowledge by learning more about solar cells.

In general, stimulation of regional industry signalizes a growing need to promote technological development at the national level. In the period from 1975 to 1995, Jacobsson (2004) distinguishes three organization groups promoting renewable energy in Germany. The first type of organizations can be denoted as public organizations, i.e. the German Society for Solar Energy (DGS[14]) is one example. Founded in 1975 with 3,600 members including scientists and industry representatives, its main subject has ever been renewable energy sources. Today, the activities of the DGS include education and promotion of research on renewable energy usage. Besides solar energy sources, it focuses on biomass, small scale biomass heating and CHP[15]. Another example for a public organization in this field is the German Association for the Promotion of Solar Power with actually about 2,500 members. In 1989, this organization proposed the first "cost covering payment" as a market introduction tool for photovoltaic systems.

A second type of promoting organizations is the conventional industry association. The German Solar Industries Association, being one example, was founded in 1978. Initially, it focused on diffusion of technical information on the use of solar technology. But in the 1980s, it also wielded influence on the members of the German Parliament.

A third type of organizations is non-profit organizations at the international level. Eurosolar is an example in this group. It was founded in 1988 as the European Association for Renewable Energies that conducts its work independently from political parties, institutions, commercial enterprises and interest groups. It brings together expertise from the fields of politics, industry, science, and culture to promote the introduction of solar energy[16].

The main goal of these organizations was to generate institutional change including, for example, promoting of technical norms. But more important is the explanatory work amongst policy makers, industry and users which change perceptions of what is possible and desirable in terms of solar energy. This activity helped for configuration of concrete support policies at the different levels from federal to communal.

---

[14] www.dgs.de. accessed 19.12.2008.

[15] Combined Heat and Power.

[16] http://www.eurosolar.de. accessed 19.12.2008.

Summarizing, this phase can be described as preparative phase with main focus on the support of basic research in the PV technology. However, at the end of the 1980s it became clear that cost reductions and competitiveness of renewable technologies could not be achieved only by research activities. Further cost reduction potential was expected from the field of experience through real exploitation of the PV systems. At the same time, the market gave some important inputs concerning consumer needs and expectations. Finally, increasing volume of production and economies of scale led to additional reduction of cost. Consequently, a paradigm shift was required.

This process of change can be disclosed by an increased number of funded organizations. In the second phase ("demand pull") there were[13]: 34 universities, 22 research institutes, 85 firms, and 26 others[17]. The first important milestone in the second period was the 1,000 roofs program (1991-1995) that created an annual market of 2-5.4 MWp per year[18]. Within this program, subsidies of investment costs were granted amounting to 50% for the old West German states, and to 60% for the newly-formed German[19]. Thereby, only investment costs of 27.000 DM/kWp[20] were accepted. Over the course of three years, 2,100 PV systems with the installed capacity of 5.3 MW and corresponding costs of 50 million EUR[21] were supported. All these systems were installed on the roof of private households.

Table 3-1[22] gives an overview of the number of installed PV systems, average costs and capacity installed within this program by the German Federal States. The Federal states got quota for the participation in the 1,000 roofs program. The federal States of Mecklenburg-Western and Saxony-Anhalt did not completely make use of the potential of this program. Consequently, others were allowed to exceed their original limits.

[17] For example, ministries, district or city administrations.
[18] 1,000 PV Roof Program, Germany, 1991-95, Finale Report by BMFT 1996, Bonn, Germany.
[19] This subsidy was additionally raised by 10% (for the newly-formed German states) resp. 20% (for the old West German states) see also Jannsen (2005, p. 213).
[20] It is equals13.800 EUR/kWp.
[21] See Laukamp et al. (2000).
[22] Only the PV Systems with complete information about costs and capacity are listed.

Federal State	Number of installed PV Systems	Systems per 100,000 inhabitants	Average Nominal Power (kWp)	Costs (DM/kWp)
Baden-Wuerttemberg	172	1.67	2.79	23,222
Bavaria	168	1.41	3.00	24,267
Berlin	105	3.03	2.30	23,475
Brandenburg	129	5.08	3.32	23,015
Bremen	67	9.85	2.30	24,998
Hamburg	125	7.32	1.78	25,879
Hessen	147	2,45	2.71	24,805
Lower Saxony	172	2.22	2.26	25,480
Mecklenburg-Western	86	4.70	2.79	23,879
North Rhine-Westphalia	150	0.84	2.17	23,761
Rhineland-Palatinate	156	3.94	2.44	24,061
Saarland	10	0.92	1.84	20,102
Saxony	150	3.28	3.48	24,774
Saxony-Anhalt	90	3.27	3.43	24,089
Schleswig-Holstein	133	4.90	2.31	25,558
Thuringia	137	5.46	2.62	25,414
Germany (total)	1,997	2,45	2.64	24,135

Table 3-1: Data about the PV systems installed within the 1,000 roofs program[23].

The defined limits did not consider the vastly different populations of the German Federal States. The distribution of PV systems per 100,000 inhabitants was highly uneven (see also Berger 2001). North Rhine-Westphalia and Bremen were two Federal States that reached the lowest and the highest installation density, respectively. The top positions were occupied by two city states: Bremen with 9.85 systems per 100,000 inhabitants and Hamburg with 7.32 systems per 100,000 inhabitants. Five federal states (North Rhine-Westphalia, Saarland, Bavaria, Baden-Wuerttemberg, and Mecklenburg-Western Pomerania) had below-average density of systems. This ranking, however, has been changed significantly. Especially Bavaria and Baden-Wuerttemberg have used the 1,000 roofs program as a starting point for the further extension of PV installations.

The 1,000 roofs program was connected with a special scientific measuring program: all PV system operators undertook to protocol monthly yield data and to provide a

[23] Source: Hoffman et al. (1998, p.13), Statistical State Agency Baden-Wuerttemberg.

logbook[24]. In this way, statistically meaningful data were collected in order to study the reasons for faults, failures and poor performance from the installed PV systems. Furthermore, wide experience about integration of PV systems on roofs was collected. Additionally, a social science study was carried out for exploring a social background of PV system users. 1,450 system owners participated in this study and gave additional information on socio-demographic characteristics such as age, regular occupation, and education. The question on the reasons for the installation of PV system was of particular importance (see Hoffmann et al. 1998).

In the period between 1991 and 2000, the Electricity Feed-in Law (StrEG) has become effective. According to the StrEG, grid operators were obligated to connect generators of electricity from renewable sources and to pay premium prices (feed-in tariffs[25]) for supplied electricity. Wind power plants and solar power plants received the highest remuneration with 90% above the mean specific revenues, followed by small hydro, biomass and biogas power plants smaller than 500 kW with 75%[26]. Hydro, biomass and biogas power plants larger than 500 kW, but smaller than 5 MW, received 65% of the mean specific revenues. The law did not cover plants larger than 5 MW (see Bundesgesetzblatt 2000). The public budget was not charged by this initiative, because all accruing costs were allocated to all energy consumers. Nevertheless, the costs for PV systems were relatively high. In addition, declining electricity prices after 1996 led to decreasing tariffs, so that investors in renewable energy sources had reinvestment problems. Therefore, the succeeding Renewable Energy Act (EEG) introduced cost based tariffs. There were other changes including differentiation of tariff rates depending on energy type, size, and site. Also, the range of promoted technologies was extended[27]. Consequently, the calculation of tariffs according to the EEG is more complicated and had to be adapted every two years in order to keep up with technological progress and market development. As a result of the EEG, the share of electricity produced by renewable energy sources has almost doubled from 6.3% in 2000 to 12.0% in 2006 (BMU[28] 2007).

---

[24] Furthermore, about 100 systems were monitored in detail for the time period of two years using remote data acquisition systems at a sampling rate of 5 min (Laukamp et al. 2000).

[25] The premiums in the Electricity Feed-In Law were calculated annually as a percentage of the mean specific revenues for all electricity sold under application of the public electricity grid in the previous year, i.e., the average electricity price for all customers. In this way, the remuneration changed every year. (cp. Bundesgesetzblatt 2000).
http://www.iea.org/textbase/pm/?mode=weo&action=detail&id=31, accessed 18.12.2008.

[26] Remuneration rose to 80% some years later.

[27] The following renewable energy sources are included: hydropower, biomass, landfill gas, sewage treatment plant gas and biogases, as well as wind, solar and geothermal power.

[28] the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety

The next important subsidy program is the federal 100,000 roofs program starting in 1999. A soft loan was granted with the low interest rate of 1.9% over 10 years supporting the installation of PV systems with up to 1 MW. The maximum amount of the loan for private persons was depending on the plant size. For the systems with less than 5kWp capacity, the loan ran up to 13,500 DM/kWp. For larger plants it amounted to 6,750 DM/kWp for the part exceeding the 5 kWp threshold. The loan of 50% was given to companies, regardless of the plant capacity. However, the maximum loan was 6,750 DM/kWp in this case (see also Berger 2001). Private households owned more than 85% of all PV systems and more than 70% of all kW installed within the 100,000 roofs program, while private companies and associations owned the remaining installations (see Rechmuth and Hünnekes (2003)).

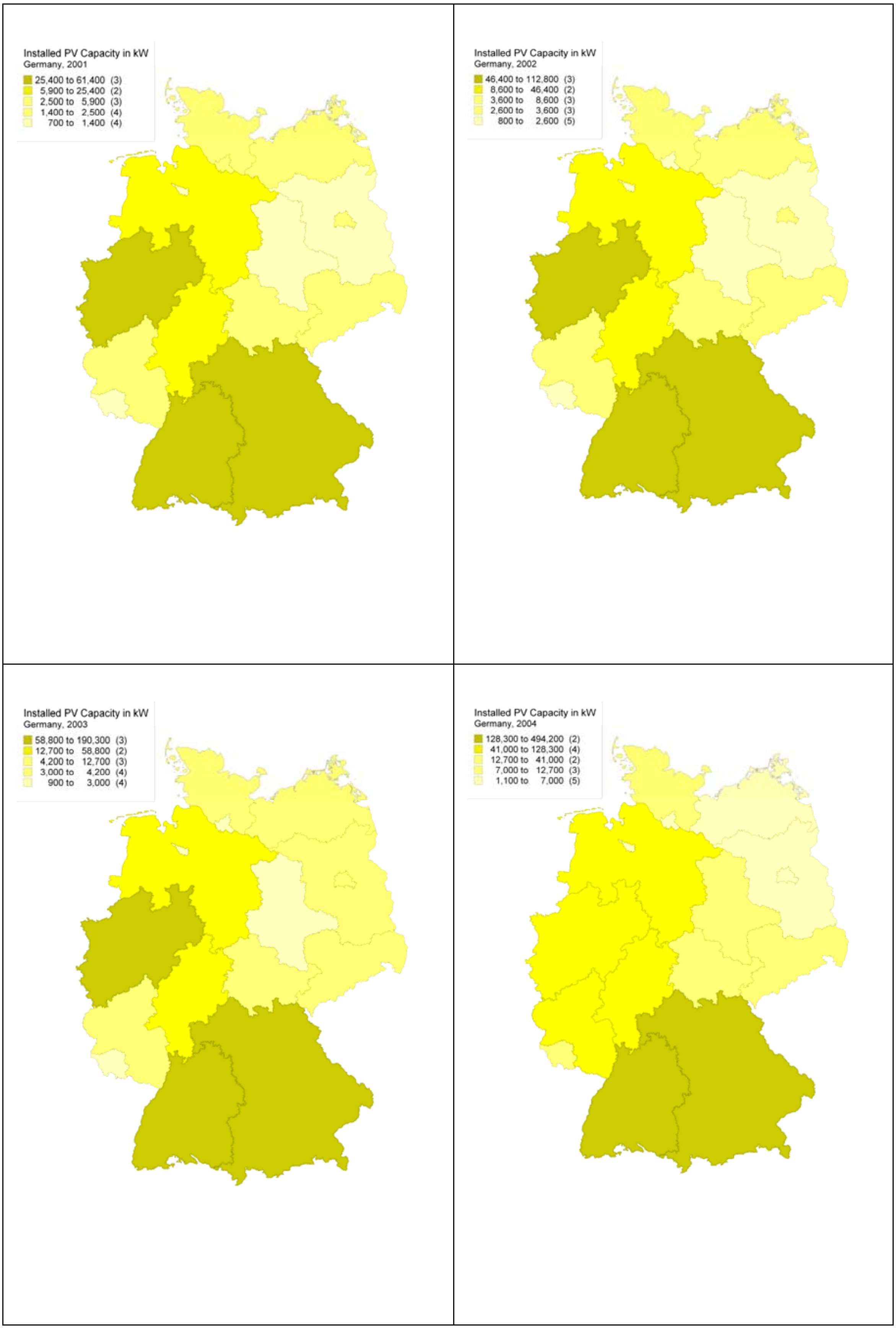
In the period between 1999 and 2003, the main target of the 100,000 roofs program was a total PV installation of 300 MW. Additionally, an intensified support of the national manufactures of PV systems was intended. Production capacities of German PV manufacturers should be used almost completely. Table 3-2 shows the distribution of credits that were contracted out within the 100,000 roofs program. Table 3-2 also reveals the inequality of regional distribution of granted credits. Bremen and Hamburg were leading in terms of density of installed systems PV systems within the 1,000 roofs program. These two city states lost the top positions and performed only relatively weak positions regarding installed capacity within the 100,000 roofs program. In contrast to this, Bavaria had the dominating position, followed by Baden-Wuerttemberg. Lower Saxony reached the 3rd place. Comparing the results of Table 3-1 and Table 3-2, one can recognize that Bavaria, Baden-Wuerttemberg, Lower Saxony have strengthened their positions regarding PV installations significantly and showed a particular affinity to PV technology.

Federal State	average wattage (W) per 1,000 inhabitants	total number of grants per 1,000 inhabitants
Baden-Wuerttemberg	5,517	1.23
Bavaria	8,205	1.74
Berlin	294	0.06
Brandenburg	627	0.12
Bremen	525	0.17
Hamburg	305	0.06
Hessen	1,511	0.4
Lower Saxony	1,668	0.47
Mecklenburg-Western Pomerania	796	0.12
North Rhine-Westphalia	1,587	0.31
Rhineland-Palatinate	1,626	0.41
Saarland	1,291	0.45
Saxony	324	0.09
Saxony-Anhalt	374	0.09
Schleswig-Holstein	794	0.22
Thuringia	754	0.18
Germany (total)	2,772	0.61

Table 3-2: Regional distribution of credits according to the 100,000 roofs program[29].

Such regional distinctions remain regarding the actual data for installed PV capacity. The Photovoltaic Magazine Photon International provides data about installed PV capacity for the time period between 2001 and 2006 (Figure 3-2). This data shows similar regional differences as revealed by the 100,000 roofs program. Again, a strong position of southern regions is obviously. In particular Bavaria is far the leading federal state regarding the PV installation density. The way to this position was certainly paved by the work of local energy societies and a range of additional support programs.

[29] Source: Rechmuth and Hünnekes (2003).


Installed PV Capacity in kW
Germany, 2001
25,400 to 61,400 (3)
5,900 to 25,400 (2)
2,500 to 5,900 (3)
1,400 to 2,500 (4)
700 to 1,400 (4)
Installed PV Capacity in kW
Germany, 2002
46,400 to 112,800 (3)
8,600 to 46,400 (2)
3,600 to 8,600 (3)
2,600 to 3,600 (3)
800 to 2,600 (5)
Installed PV Capacity in kW
Germany, 2003
58,800 to 190,300 (3)
12,700 to 58,800 (2)
4,200 to 12,700 (3)
3,000 to 4,200 (4)
900 to 3,000 (4)
Installed PV Capacity in kW
Germany, 2004
128,300 to 494,200 (2)
41,000 to 128,300 (4)
12,700 to 41,000 (2)
7,000 to 12,700 (3)
1,100 to 7,000 (5)

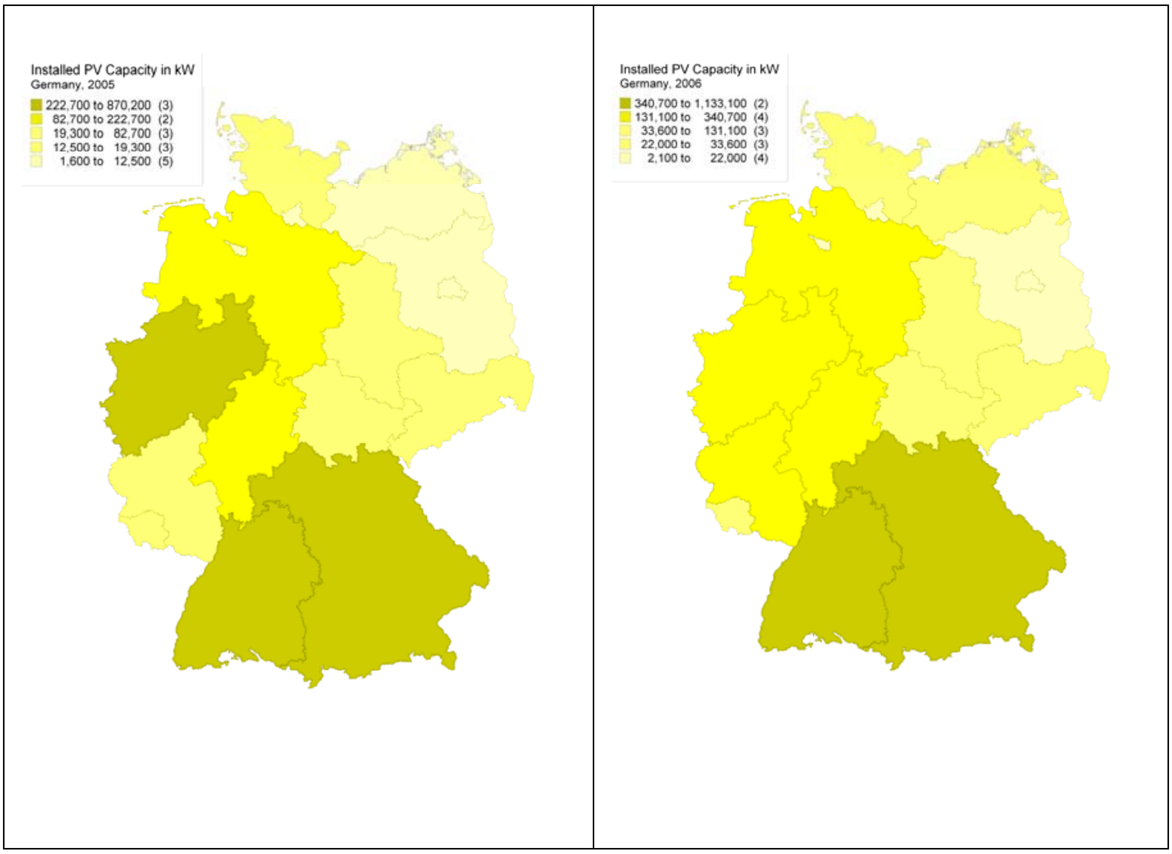


Figure 3-2: Newly Installed PV capacity in Germany (in kW), 2001-2006[30].

There are some reasons for these regional disparities. First, to a great extent the PV electricity comes from small power generators that are dominantly distributed in the build-up areas. The regions with a higher share of the build-up areas offer a better potential for the PV installations.

Second, the income in the southern and western regions is higher than in Eastern Germany. Even with a credit, the starting investments for PV installations are a burden for private households. The third reason is the geographical distribution of the solar energy resource. The electricity produced by the PV systems is largely determined by the amount of incident solar radiation. For this reason some southern areas such as Bavaria and Baden-Wuerttemberg have certain advantages. Places in the north receive in average 25% less daily global radiation on a PV module than the southern regions do.

Finally, photovoltaic as science-based technology issues a range of challenges to universities and research institutions. These organizations open up new growth potentials, push the social acceptance of this technology and affect its diffusion. The

[30] Source: Photon International, own presentation.

spatial distribution of the universities that offer education in the area of Photovoltaics in Germany is also uneven. There are 171 courses of studies related to the Solar Industry in Germany. These courses are allocated to 89 universities in 79 cities (see Franz, 2008). 65 of these institutions are universities of applied sciences. This fact reveals the higher flexibility of such education unit to needs of employment market.

Additionally, the intensity of research activities in these universities differs significantly. In the period of time 1992-2008, most of the research projects (31) with the highest support are listed in Stuttgart (23 Mio. Euro), followed by Darmstadt (12; 8.8 Mio Euro), Konstanz (10; 6.7 Mio Euro) and Munich (10; 6.1 Mio Euro). These high values reveal spatial concentration of the PV research on the South of Germany. It can be assumed that the teaching quality of the universities involved in these research activities has a higher standard. The universities in southern regions are important actors in the national innovation system and have strong impact on the technological development and diffusion of the PV technology.

The 100,000 roofs program ended in 2003. The support of PV systems by soft loans is maintained by the program "Solar Power Generation"[31]. Under this program 43,000 loans representing a total volume of 338.1 MW equivalents to 1,335 Million EUR investments have been granted since 2005. Only in the year 2007, 101.3 MW were supported (see IEA PVPS 2008, p.65).

There are also a range of other investment programs[32] of various Federal States that facilitate the further developing of the PV market in Germany. To sum up, regarding promotion policy a division of German PV market into two phases is meaningful. However, it is still an open question whether these phases are also identifiable regarding the indicators of market development in the stylized model suggested by Grupp (1998). This issue is the research question of the next three chapters in of this thesis.

---

[31] See http://www.kfw-foerderbank.de/EN_Home/Housing_Construction/SolarPower.jsp (accessed 21.12.2008)

[32] For example, PV programmes at schools financed by utilities and governments (e.g. 'SONNEonline' by PreussenElektra, 'Sonne in der Schule' by the Federal Ministry of Economics and Technology and 'Sonne in der Schule' by Bayernwerk

4 Market Analysis in the Field of PV: Measuring Changes of Solar Module Quality over Time

This section deals with a longitudinal analysis of market level measured by development of product characteristics. This approach is based on the theoretical premises that development of science-based technologies is strongly influenced by interaction between demand and supply, at least in the subsequent phases. Market is a place where suppliers and demanders meet to conduct an exchange. Following this concept, the market mechanism provides not only equilibrium prices and equilibrium quantity, but is also a determinant for the direction and extensity of technological progress. This consideration is confirmed by the following citation:

> "From pioneering contributions by Jacob Schmookler (1966) and Richard Nelson (1959), economists have recognized that demand and supply conditions determine not only prices, but also the pace of technological change" (Scherer 1992, p. 1427).

In line with acknowledged importance of innovation, economic research on innovation in the past 40 years has very much focused on the generation side of innovations rather than on the demand side. A reason for this may be the fact that there are more difficulties in analyzing demand in theory and empirical work. However, with the increased interest in user-driven innovation, this is likely to change.

Very early Ironmonger (1972) deals with customer choice where several priorities among wants exist, and theoretically explored choices and wants when technology or quality changes or new commodities are introduced. But in this work, commodities are regarded as homogeneous entities with no differentiation by features offered to customers. However, consumer preferences are not general, but attribute-specific. It is Lancaster (1991)[1] who replaces the price-quantity concept by a "bundle of characteristics" to describe products by rank order. Several other authors extended Lancaster's model towards metric scales in various dimensions[2].

This section uses a metric re-scaling approach, which measures the quality of innovative products using the product's features. This specific approach permits the

[1] See the 1991 book for reference only. His work dates back to 1971.
[2] For a review see Grupp (1998).

systematic measurement of the technical improvement of products and can offer a dynamic comparison between specifications of innovative products and processes. Such measurements show the extent to which a product's "supply" of characteristics matches the "demand" for these characteristics in the market place.

The empirical part of this study has two main objectives: first, to examine whether the efficiency of the market is not disrupted by government activities (e.g. subvention); second, to check whether innovative solar module producers can disproportionately benefit from the growth rates in the PV market.

The remainder of this chapter is structured as follows: after the introduction, section 4.1 deals with the theoretical background of the metric re-scaling approach. In section 4.2, market efficiency is examined, while section 4.3 considers the changes of technical performance of solar modules over time. Paragraph 4.4 represents the evolution of market shares and compares it with the results of the metric rescaling approach in section 4.3. Finally, section 4.5 concludes the chapter with the main findings.

4.1 Theoretic Background: the Metric Re-scaling Approach

A very extensive overview in dealing with the variable "quality" in consumer theory is given by Waldman (2000). Already in classical contributions by Adam Smith and David Ricardo; references to the existence of variability of quality can be found. Adam Smith, for example, mentions the variability in quality of labors, and in his theory of land rent David Ricardo analyzes differences in the productivity of land induced by various grades of land (see Wadman 2000, pp. 7-8). However, the classical consumer theory does not suggest any solution of a problem which is raised when the quality of good is not constant. In the mid 1960s there were several researches who discussed the research question of quality change. Two of them are mentioned in this section: Lancaster (1966, 1971) and Ironmonger (1972).

According to Ironmonger (1972), the consumers desire commodities because commodities satisfy human wants. There is a multiplicity of wants which

> "...are assumed to be so ordered that at a given income and prices the consumer will satiate as many wants as possible, going down the order of priority from the most important to the least" (Ironmonger, 1972, p. 23).

In this context, the goods do not produce the utility directly, but through satisfying some particular separated wants. The major problem in regarding product quality arises, if goods are considered as homogenous units. There are difficulties to compare a product with predecessors if a new commodity has complete new features.

Concerning this matter, Lancaster (1966) presents a "new approach to consumer theory". This approach is based on the assumption that consumers derive the utility not directly from products, but from its properties or characteristics. The properties have a feature to be objective and the same for all consumers. But the utility derived by consumers is subjective and depends on their preference function. In this way, the consumer can be regarded as a producer: consumption is an activity in which goods are input and the collection of characteristics is the output (Lancaster 1966, p. 133). However, the real objective of the consumer is to maximize the utility, and not the amount of characteristics. Thus, the characteristics have to be transformed to utility. Lancaster (1966, p.135) assumes each individual follows an ordinary utility function on characteristics and chooses a situation which maximizes individual's utility.
This approach offers an important advantage over the ordinary consumer theory which completely disregards all intrinsic properties of particular goods. Traditional

theory seems to be inapplicable if goods get a new feature. In this context, there are only the same commodities or different commodities. Following Lancaster, one has to apply a new utility function if a good obtains a new property. This model presumes that the consumer can order a variety of goods in terms of quality. The product is described by a range of characteristics and the consumer has to order each characteristic according to his preferences. Using this transformation, the distance between the first-best and second-best disappears. This disadvantage is improved in the metric re-scaling approach[3].

It is widely accepted that innovation is complex and difficult to measure. The measurement of improvement in product performance poses special difficulties. Most innovation indicators such as R&D expenditures, patents or the number of new or improved products, respectively, do not contain information on a product's performance progress. Here, the established concept developed by Grupp (1994) may be applied.

The metric re-scaling approach permits the systematic measurement of the product's technological improvement and can offer a continued comparison between specifications covering innovative products and processes. In addition, it demonstrates differences in the technical properties of offered products and therefore the variety (Saviotti 1996) or product differentiation of innovations by suppliers as well as intra-firm brand differentiation of the same producer. This approach is based on the measurement of technological characteristics of innovative products and is based on Lancaster's (1991) consumer theory. The following four steps can be distinguished in measurement process of quality chance using metric re-scaling approach (see also Grupp and Maital 2001, Grupp 1998):

1. A selection of the fundamental numerical characteristics of the investigated product, process or service that are relevant for the buyers has to be carried out. This selection is based only on technical features. "Fashion" aspects have to be excluded from the analysis, because it is not possible to map temporary fashion-related characteristics on an ordinary scale.
2. Measurements of selected attributes for the product and its competitors are required.

[3] The metric re-scaling approach is equivalent to the formerly known "technometric approach" (Sahal, 1985), a term which has never become wide acceptance. Nevertheless, under the notion "min-max approach" or "metric re-scaling" it is now more widespread in use, e.g. for the Human Development Index of the United Nations.

3. A transformation of each measured attributes into dimensionless interval [0, 1] is required. Otherwise it is not possible to aggregate different features with different measurement units to one common quality indicator. The “0” in this interval represents the attribute’s lowest value among all competing products, and “1” corresponds to the attribute’s highest value. This transformation for brand *k’* in subset of *k* brands is based on the following metric:

$$K^*(i,j,k',k,t) = \frac{K(i,j,k',k,t) - K(i,j,k_{\min},t_o)}{K(i,j,k_{\max},k,t) - K(i,j,k_{\min},t_0)} \qquad (4\text{-}1)$$

where *K(i,j)* is the specification of the attribute *j* of a product *i*, *t* is the time index and t_0 the initial period. k_{max} and k_{min} denote those brands whose products have the maximal resp. the minimal performance for the corresponding attribute with respect to the total subset.

For some attributes, a higher attribute score (for example weight or price) implies lower product quality. In this case, the inverse formula has to be used:

$$K_{inv}^*(i,j,k',k,t) = \frac{K(i,j,k',k,t) - K(i,j,k_{\max},t_o)}{K(i,j,k_{\min},k,t) - K(i,j,k_{\max},t_0)} = 1 - K^*(i,j,k',k,t) \qquad (4\text{-}2)$$

The index *K** is not dependent anymore on specific physical units. Equation (4-1) resp. (4-2) measure the *j*-th attribute value at time *t* in comparison to the capabilities measured in the initial period t_0. The initial position may be “frozen” and used as a reference point for evaluating the technical improvement over time. This way the technical progress in the product features can be revealed (see also Grupp 1998; Haller and Grupp 2009).

The maximum values of *K** represent the highest, and the minimum values of *K** correspond to the lowest technological standards. At the beginning of the analysis, the maximum value of *K** is 1. This value can be achieved only for that innovation[4] which has the best performance value regarding characteristic *j* among all competing products. The value of *K*=0.5* can be interpreted in the following way: regarding the characteristic *j,* the product *i* has the same distance to the best product like to the

[4] It is possible that more than one product share the top position with respect to one technical characteristics.

worst product. The maximum value of 1 can be exceeded in the later periods as a consequence of technological improvement. In general, the values of K^* do not have any upper limit.

4. Finally, from the micro-level a metric profile may be aggregated[5] on the level of all j specifications per product i if functional characteristics or (revealed) preferences F (Saviotti and Metcalfe, 1984) are defined:

$$K^*(i,k',t) = \frac{\sum_j K^*(i,j,k',k,t) \cdot F(i,j)}{\sum_j F(i,j)} \tag{4-3}$$

The preferences can be derived from utility functions, by introspective or market observation, from expert knowledge or by use of hedonic prices.

How can the metric re-scaling approach be interpreted? Future product characteristics are unknown, i.e. the market of innovation can be regarded as a discovery process. According to the metric re-scaling approach, the peculiarity of attributes at t_0 is fixed and can be taken as the yardstick for measurement at future dates. Since neither the ideal performance values of actual characteristics, nor the performance values at the end of the observation is known, the yardstick is taken from the heterogeneity of the present sample at t_0, not from idealized asymptotic values. This corresponds to evolutionary thinking, showing the evolution of product features that evolved from history, where ever it arrived at nowadays, and not a future ideal equilibrium.

There are some assumptions in this approach which have to be discussed further. (see also Grupp 1998; Haller and Grupp 2009). The first critical point relates to the selection of relevant characteristics. Already Lancaster (1971) discusses this limitation. At the standpoint of the market, the interaction between R&D, marketing personal, and consumers identifies missing product features. Controlling competition realizes these market niches and bridges revealed gaps. But also for the analysis, the selection of product characteristics has to be considered carefully. There is a necessary, but not sufficient condition for the process of attribute selection: if two products have no differences in any investigated characteristics but in the price, then one important attribute is missed in the analysis. Furthermore, the most important

[5] It is clear that the last aggregation step leads to loss of information. It is not possible to identify the source of technical progress on this aggregate level because technological improvement can concern either only one attribute, or a combination of attributes.

characteristics can be identified using interviews with experts. Grupp (1998) reports that already a handful of interviews with experts provide a good overview on the most of the market-relevant product attributes.

Another point of criticism is that the establishment of the maximum and the minimum performance values that can be interpreted as more or less arbitrary. However, such approach is a standard practice in the literature. The quality or technological level of the following products can be measured in comparison with the product quality of the market leader at the point of time t_0. In the chronological consideration, one puts the technological market know-how is oriented on the level of the leading company at the beginning of the analysis. The following products are compared with technological standards of market at the time $t=t_0$.

A critical point is also to the fact that innovative process features are not directly covered in this concept. In cases where the producer marks these due to marketing or law (for instance environmental, nutrition, or health relevant details such as "non-bleached", "sugar-free"), one can code this information by a binary variable. Otherwise, features such as "warranty time" may be considered as a catch-all variable for process innovation.

Adittional critical point is the establishment of the preference profile in equation (4-3). Considering each attribute separately, consumers would never select an inefficient product which has some poor characteristics and any superior characteristics compared to other technical alternatives. However, if the product has a tuple of characteristics, the following situation occurs relatively often: a product gains an advantage regarding some attributes, but at the same time has lower performance values in aspect of other attributes. In this case, the consumer is coerced to choose a product with the best performance values for more relevant attributes and to accept poorer performance of less important attributes. This leads to an aggregation problem which can be only solved if manufactures and consumer have an agreement about their preferences. However, in reality consumers and producers have a non-uniform rating in relation to cost-benefit ratio of the products. Nevertheless, empirical studies show that a relatively consistent evaluation of important attributes and their weighting can be achieved. The example of comparative consumer tests in Germany reveals that the deviation in preferential weighting of product characteristics is relatively low and lies in any case under ten per cent (Grupp 1998, p. 138). Grupp (1998) also reports that inaccuracies in preferential profiles have less consequence on the results of the analysis due to the incomplete list of characteristics.

In conclusion, the main problem of the data re-scaling approach is a problem of data acquisition, i. e. the collection of detailed information about technical characteristics of the products. Unfortunately, there are not any standard references which can be taken for this analysis. Theoretically, information sources such as producer and consumer data, or expert judgements are available as potential data bases. In practice, however, the problem of data collection is exacerbated by the fact that different references tend to different statements. This problem arises, for example, because customers and producers have different needs, meanings, terminology, expectations, or interpretations. Furthermore, the measurement of technical properties can be influenced by environmental factors such as light, temperature, or air moisture. The quality of a product can also vary as a consequence of manufacturing circumstances. The poor data quality is a well-known problem and has consequences for interpretation of analysis results. But it is really difficult and costly to improve such data basis significantly.

4.2 Investigation of Market-Efficiency

The subsection 4.2 deals with the investigation of market efficiency for the PV market in Germany. The analysis of market efficiency is important because only an efficient market has an important property handling with market information correctly. Knowing how to handle product information effectively implies correctly setting of prices quickly and demand accommodation. Under these conditions, inefficient, technical underdevelopment products would be displaced from the market and lose in competition against the technological progressive modules provided that technological better products have an appropriate cost-performance ratio.

The basic principle of free market economy is that markets have particular affinity to autonomous adaptation. Market responses to price signals; exogenous regulation is not necessary and not desirable. Otherwise, market biases can arise which give their owners unique position of power and control over customers and competitors. In normal market conditions, market actors react automatically as rational actors and follow their individual terms of reference. The market mechanism makes sure that non competitive actors are pushed out of the market. In the case of the PV market, one can often hear a persistent reproval concerning its inefficiency. Furthermore, the market would not exist without government regulation. It is important to clear this question - otherwise the question about enforcement of innovative products is obsolete.

First of all, the problem of inefficiency is a problem of almost every new market. For most of the emerging markets, the costs of doing business are higher than in established markets. Proven structures in established markets promote speedy information flows between different groups of actors. Also the question of market transparency is an important issue. The solar PV market cannot clearly be declared as a transparent market. Different activities in the solar PV commercial process (e.g. searching for the products, services and companies) are very costly in terms of time. Consumers spend a lot of time in searching for reliable business partners able to install solar modules and to guarantee functionality of the products. The lack of precise information may lead to major contortions in the demand for solar modules. Producers spend a lot of time with finding trusted distributers in different countries. Innovative companies introducing a new product or product differentiation also require much time for presenting customers and business partners new technological alternatives.

The rapid growth of the global solar industry leads to incurrence of high market dynamics and at the same time tends to preserve non-transparency of markets. The reason for this is due to the fact that high dynamics implicate a constantly increased demand for solar products. Many installers have orders but cannot execute these orders as a consequence of shortage of solar products. Also silicon shortage limits the possibilities of solar manufactures to expand the supply of solar modules significantly. This causal chain increases the uncertainties in the market and does not result in more market transparency. For non-transparent market, it is difficult to react on market signals adequately.

A further aspect in photovoltaic market is government regulation which can also cause the emergence of certain market distortions. This a particularly critical point in the development of the PV market in general. In each country[6] which has a top position in the global solar market, there is a whole series of support programs which stimulate the market development. In Germany, the first national support program ("the German 1,000 roofs program") took place between 1991 and 1995. In 1999, another program ("the 100,000 roofs program") has been started offering interest-reduced loans for PV systems. In April 2000, the Renewable Energy Sources Act (EEG) became effective, which considerably increased the feed-in tariff. It is only a short list of government efforts which has made photovoltaic to a well-established electricity source in Germany. A similar situation can also be found in Japan and USA. At the first glance, the accusation that subsidies can create market distortions by artificially manipulating price signals does not seem to be arbitrary.

A direct comparison of government financial support and installed capacity is not possible because these indicators are measured in different physical units. Another possibility is to compare growth rates of these variables. Figure 4-1 illustrates the evolution of subsidies and financial supports compared to PV installation in Germany in the period of time between 1975 and 2005. Two noteworthy observations emerge. First, in the period of time 1975-1993 growth rates in public financial support and installed capacity run in parallel, although the rates of expansion for installed capacity seem to be higher. Second, the time after 1993 is a period of stagnation for subsidies. The PV sector seems to slowly overcome its

6 Japan, Germany and USA.

dependency from public support and exhibits positive tendency in growth rate (red line). Presumable, this period of time corresponds with the implementation of feed-in tariffs and other demand-oriented efforts of the German government in order to speed up the diffusion of PV technologies. However, the considerations above justify the analysis of market efficiency in a more detailed way.

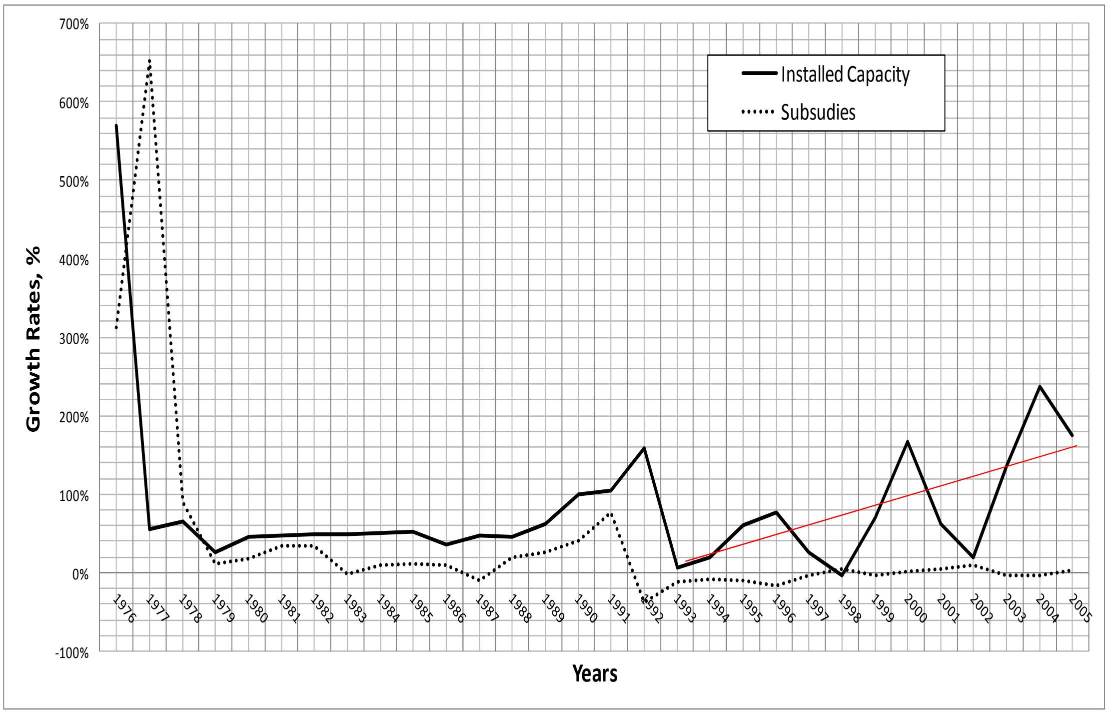


Figure 4-1: Growth of subsidies vs. growth of the PV installations in Germany[7].

At the University of Muenster (Germany) the coordination failure diagnostics (CFD) concept was established by Grossekettler in the 1980s. This approach can be used to test the presence of inefficiency in the German PV market. The CFD concept identifies five basic market processes which have to be analyzed in order to investigate the appearance of possible market distortions (see also Grosskettler 1999, Blanckenburg 2007).

The first is **the market clearing process** which analyses mid-term compensation between demand and supply, aiming at least detect a compensation tendency. This can be measured by existence of irremovable commodities (over-supply) or delivery delays (over-demand).

[7] Source: own computation; Data: DGS, BMBF, Förderkatalog.

The second is **the rate of return normalization process** which tries to ensure that firms which are active in the market survive. This survival ought to be ensured – yielding by high rate of returns- enterprises invest more and grow. The variation in capacities leads to achievement of a rate of return equalization and an efficient distribution of income.

The third is **the erosion of market power process** which impedes the emergence of long-term dominant positions in the market. This means that any market participant has the power to control or manipulate market processes according to own personal aims and against societal benefit.

The product innovation process should increase competitiveness in products and act with this in the interest of consumers. The measurement of this progress can be carried out using the metric re-scaling approach introduced in section 4.1.

Analogously to the product innovation process, **the technology innovation process** should select up-to-date technology and eliminate technology which is not state-of-the-art. Here those technologies can be analyzed that prevail in the market as a result of technological screening. In the next step, all these research questions have to be checked using the example of the solar PV market in Germany.

Analysis of the market cleaning process: The silicon shortage problem is well-known and has bothered PV industry experts for many years. As a consequence of this, demand for solar module exceeds supply. "Silicon feedstock supply constraints growth in the solar photovoltaic industry," notes Frost & Sullivan Industry Analyst Pramodh Panchanadam[8]. "Only by increasing production capacity and use of new technologies in both manufacturing and production processes will the effects of this constraint be mitigated" (ibid.).

However, regarding absolute numbers, the amount of silicon increases year after year. Additionally, the producers have learned to use available silicon more effectively. This has resulted in an increase of the output of solar cell manufactures. Between 2004 and 2006, solar cell manufacturing output doubled. But there are also numerous R&D efforts to solve the problem of silicon shortage using for example thin films technologies[9] which reduce silicon consumption in solar cell production.

---

8 http://www.energy.frost.com. Accessed on 18.01.2009.

9 Cp. also the analysis of the technology innovation process.

According to estimates from EuPD Research[10] global thin film production will grow faster than the global PV markets: thin film capacity will experience a 83% compound annual growth rate (CAGR) until 2010, compared to the CAGR crystalline capacity of 32%, and the CAGR global PV demand of 41%. Currently, 13% of the solar cells manufactured by Germany are thin-film solar cells[11]. Especially Eastern Germany has the potential to develop further as one of the most important production sites as well as technology clusters in thin film industry. Taken together, there are difficulties in compensation between demand and supply in PV solar cell production, but there are also positive tendencies despite efforts to close these differences.

Analysis of rate of return normalization process: It is difficult to argue the fact that photovoltaic systems are too expensive. For reasons, costs continue to be a decisive factor for the survival of solar companies. However, when the rising cost of energy and the cost of environmental damage would be taken into account, the costs for photovoltaic systems become more reasonable. Comparative evaluation for several technologies[12] which has been made within European Commission's 5th Framework Program ExternE project showed that Photovoltaic is the cleanest technology regarding the use. It has, however, considerable life cycle impacts (see Nisan and Benzarti, 2008).

Despite the PV industry is perceived as too expensive, it enjoys strong growth over the past 30 years. It reasonable to assume, that such development would not be possible for such long period of time by existence of substantial market failures. According to the study conducted by Photon Consulting, the profit margins in the solar industry are higher the closer the manufacturer is located to the beginning of the value-adding supply chain, i.e. silicium manufacturing. In the charge to reduce costs, the many firms try to control a larger share of the value added chain. For this reason, the PV market concentration of the PV industry is very high. This will be analyzed in the next market process.

Additionally, the dynamics of competition on the German PV market can be characterized by the number of market entrances. In the 1970s there were only three

10 http://www.eupd-research.com. Accessed on 18.01.2009.

11 http://www.renewableenergyworld.com. Accessed on 18.01.2009

12 Fossil fuels: coal and oil technologies, nuclear technology, renewable: solar photovoltaic, on-shore and off-shore wind, a wide range of biomass fuels (e.g. waste wood, crops).

developers of solar cells in Germany: Telefunken, Siemens and Nukem[13]. According to the PV report 2007[14], there are nine solar cell manufactures in Germany: Deutsche Cell in Freiberg, ErSol Solar Energy in Erfurt, EverQ and Q-Cells in Thalheim, Scheuten Solar in Gelsenkirchen, Schott Solar in Alzenau, Solarwatt Cells in Heilbronn, Solland Solar Cells in Aachen/Heerlen (NL) and Sunways in Konstanz and Arnstadt. Similar increase can be considered by number of silicon and wafer producers. Some large PV companies such as Solar World and RWE Schott Solar follow the strategy of integrated production from feedstock materials to module fabrication. This strategy allows firms to be independent from fluctuations of the PV market (see Marinova and Balaguer 2009). Additionally, the involvement in the upstream market gives the firms more scope of development in pricing for its products.

The downstream PV market (panel manufactures, thin film manufactures, distributors, and installers) has also grown significant. New application fields such as building integrated (façade) and roof solar cell systems arise. The successful development of new market segments requires further education of electricians, builders, façade makers and provides additional jobs.

Analysis of the erosion of market power process: The most frequently applied measures of market concentrations are k-concentration ratio (CR_k) and Herfindahl Hirschman Index (HHI). The CR_k takes the market shares of the k largest firms in the market and ignores the remaining actors. The main idea of this indicator is to test whether the market is dominated by a small group of actors. Table 4-1 shows the top two solar cell producer’s market shares in USA, Japan, and Europe. In general, the concentration ratio shows relatively high values. Especially the market of solar cell producers in the USA seems to be concentrated.

Year	1987	1988	1989	1990	1991	1992	1993	1994	1995
USA	82	77	82	84	85	81	85	80	76
Japan	49	55	51	56	59	62	66	65	68
Europe	51	31	28	28	30	37	47	44	49

13 Nukem as supplier of nuclear elements developed cadmium sulphide cells using its knowledge of heavy elements history. But due to environmental problems with the use of cadmium, this firm later concentrated on crystalline silicon cells (see Jacobsson et al. 2004).

14 IEA PVPS (2008)

Year	1996	1997	1998	1999	2000	2001	2002	2003	2004
USA	72	72	67	66	65	65	64	67	63
Japan	67	74	78	75	72	75	73	74	71
Europe	59	56	65	49	32	39	38	38	34

Table 4-1: Market share for the top two solar cell manufactures, %[15].

One can identify a declining trend from 1993 to 2004. But also at the end of the investigated period, the market in USA is dominated by only two solar cell producers. Until 2000 there were two top players: Siemens Solar Industries and Solarex. After 2000, Solarex was takenover by Astropower. Again this, for Japan the ratio concentration of the top two producers show continuous increasing trend beginning nearly by 50% in 1987 and achieve slightly over 70% in 2004. The values of the concentration ratio for Europe do not have any clear trend and fluctuate between 30% and 50%. Similar findings can be also derived using the HHI (see Figure 4-2).

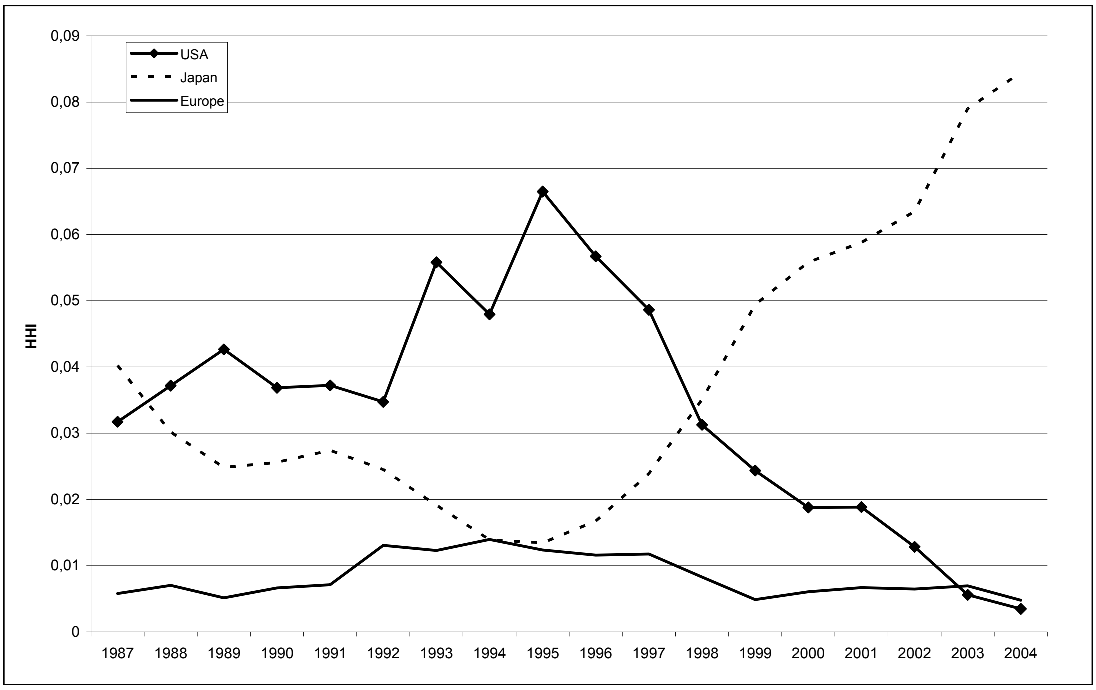


Figure 4-2: The HHI for solar cell producer market in USA, Japan, and Europe.

In summary, the European market of solar cell producers seems to be less concentrated in comparison with Japan, USA. Until 2004 Germany was the second world's largest PV market after Japan. After 2004 Germany took over Japan and is

15 Source: PV News, different issues.

now actually the world's largest PV market. For this reason, one can say that the German PV market can be also regarded as less concentrated one.

Analysis of the product innovation process is carried out using the metric re-scaling approach in 4.3. Considering the technology innovation process, PV materials can be categorized either as thick crystalline (e.g. monocristalline silicon (m-Si), polycristalline silicon (p-Si), and Ribbon Si) or thin films (e.g. amorphous silicon (a-Si), Cadmium Telluride (cdTe), and Copper Indium Diselenide (CIS)). Figure 4-3 shows PV technological development in the period of time between 1988 and 2006.

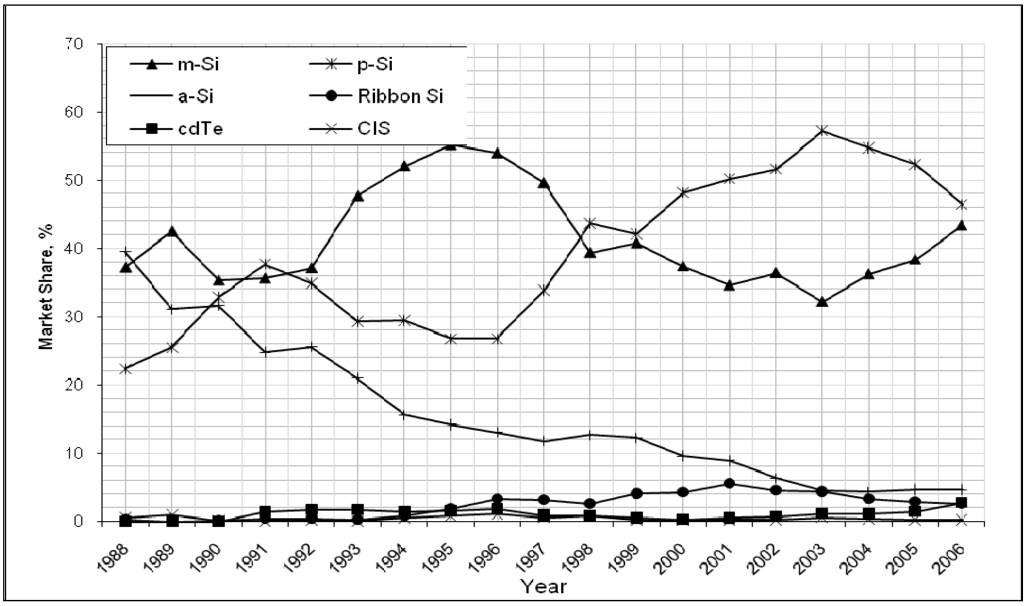


Figure 4-3: Market shares of different PV technologies[16][17].

At present, there are two main problems of modern PV solar cells: low efficiency[18] and high production costs. Different PV technologies try to solve these problems. Figure 4-3 shows a dominant position of thick crystalline materials. Amorphous silicon significantly decreased in production from ca. 40% market share in 1988 to 4.7% in 2006. Polycristalline silicon devices are generally less efficient than those of single-crystal silicon, but they can be less expensive to produce. This race between lower production costs and higher efficiency is reflected in time development of market shares of monocristalline and polycristalline silicon technologies.

---

16 Source: PV News, Photon different issues.

17 m-Si: Monocristalline Silicon, p-Si: Polycristalline Silicon, a-Si : Amorphous Silicon, cdTe : Cadmium telluride, CIS: copper indium selenide.

18 In power conversion, efficiency means the amount of electrical power generated by the solar radiation striking the surface of a photovoltaic cell in a given period of time (Lorenz et. al. 2008).

The other important PV approach is thin film technology which has existed for a long time, but has only recently proved that it can reach sufficiently high efficiency level at commercial production volumes. Thin film trades off lower efficiencies against a significantly lower use of materials - about 1% to 5% of the amount needed for silicon-wafer-based photovoltaics (m-Si and p-Si). The result is a cost structure roughly half that of wafer-based silicon. For this reason, this technology has significant headroom to extend the cost gap in the long term (see Lorenz et al. 2008).

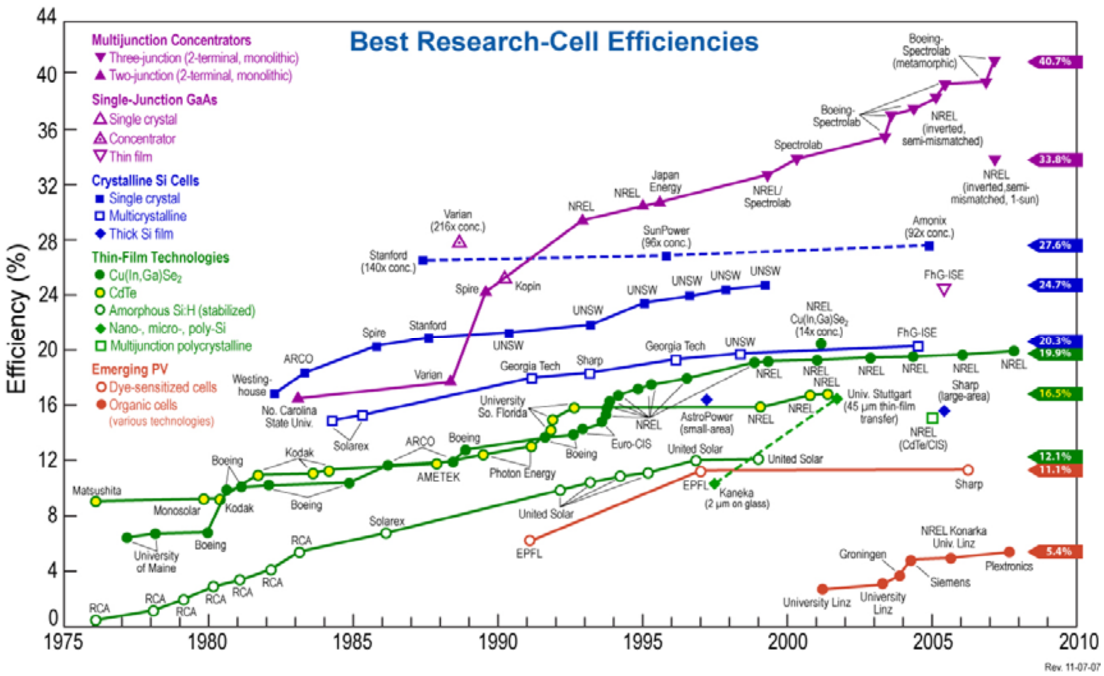


Figure 4-4: Efficiency development of solar cells from different producers[19].

In general, innovations are rapidly increasing the efficiencies of emerging cells (Figure 4-4). However, much development work and many field studies are needed, especially under difficult conditions, to fully exploit the promise of these cells in real world PV system.

In summary, the PV industry needs technological solutions for many issues. There is a large number of companies that have been established in recent times. It is a new fast-growing market, and government policies continue to influence the industry development heavily. Without such promotion, the high cost of generating solar power would prevent it from competing with electricity from other renewable energy sources. However, the industry is changing: over the last two decades, the cost of manufacturing and installing a PV solar-power system has decreased by about 20

[19] Source: Khan et al. 2008.

percent with every doubling of installed capacity. Grid parity (competitiveness with retail electricity prices) will be reached progressively from 2010 onwards in several European markets. Countries with the highest solar irradiation and higher electricity prices, such as Italy and Spain have the potential to reach grid parity starting in 2010 and 2012, respectively. Grid parity will be reached in Germany in 2015 and cover progressively most other EU countries up until 2020[20].

[20] European Photovoltaic Industry Association

4.3 Dynamics of Technical Performance of Solar Modules

In section 4.2 the assumption of open competition on the market of solar cell producers is confirmed using the coordination failure diagnostics (CFD) concept. This allows us to neglect government subsidies (of unknown quantity per firm and per country on an international level) in the analysis to follow. Although competition may work to drive out inefficient firms, it has still to be clarified whether the remaining manufacturers also provide superior products over a period of time. In order to quantify the technological competence of solar cell producers, the quality of its products has to be screened.

In this section, only selected solar modules offered by solar cell producers are considered. The selection decision is due to two reasons: first, the quality of solar modules is basically determined by the solar cells used. The interconnection of solar cells into cell strings and the embedding of the solar matrix in foil have a modest influence on the technical specifications of finished solar modules. Second, the solar cell producers give reliable data on their market shares in solar cell production for all their products together, not just for single variants. This information is required for the later investigation. Thus, it is very difficult to find technical information about intra-firm differentiation of the produced solar cells, since solar cells are converted into solar modules in-house or are delivered as customer-specific products (Original Equipment Manufacturing, OEM) for other solar module producers. The solar cell market differentiation can only be observed in-between firms. For this reason, the technical data of selected solar modules for the re-scaling approach are used.

In order to test the technological performance dynamics of solar modules, typical standard products of the following nine manufactures are selected: RWE Schott[21], BP Solar, Kyocera, Helios, Photowatt, Siemens Solar[22], Solarex[23], Solec/Sanyo[24], and Sharp[25].

The selection rule is to take one product per firm for household appliances (roof mounting) with the best available performance characteristics of that firm (if more

21 Former AEG and ASE.
22 Siemens Solar was bought out by Shell Solar in 2002.
23 When BP and Amoco merged in 1998, half of Solarex came with the deal. The companies merged completely in 1999, when BP Amoco took over the other half of Solarex from Enron.
24 Solec was sold to Sanyo in 2000.
25 It was not possible to find data of Sharp's solar modules before 1996.

than one is offered) in the range of 40 to 130 Wp (in 1987) to over 300 Wp (in 2006). Data sets covering the years 1987 to 2005 are included in the analysis. Four product characteristics (nominal power, efficiency[26], weight pro watt and warranty[27]) are selected. The restriction to only four characteristics is, for sure, very critical. However, the time frame of analysis from 1987 to 2005 makes it very difficult to collect a more extensive data set, which covers more technical characteristics.

There are different possibilities for the data acquisition of solar modules: Flyers, exhibition materials, technical literature, and information provided on the Internet. These sources mainly contain information about the actual deliverable solar modules. For the technical data of older solar modules other publications are needed. For this analysis, additionally the issues of the journal “Photon international” are used for the years after 1996. For the time before 1996, the publications of the Institute for Applied Ecology, Freiburg/Germany, are used. In the years 1987, 1991 and 1994/1995 this institute published a market review on the subject of photovoltaics (Meereis 1987, Leuchtner and Boekstiegel 1991, Leuchtner and Preiser 1995).

Using the re-scaling approach the following preference profile is applied:

$$\forall i, j : F(i, j) = 1/4. \tag{4-4}$$

i. e. all of the four considered technical characteristics of solar modules have the same importance (which, however, is a problematic assumption, see Grupp and Schubert 2008).

Figure 4-5 shows the values of the re-scaled quality indices for solar modules of nine solar cell producers, which are calculated using the approach describing in Section 4.1. By comparing the evolution of quality indices in the period of time between 1987 and 2005, continuous technical development in solar cell technology is clearly recognizable. With the exception of Photowatt and BP Solar, all producers have improved the quality of their products continuously until 2003.

An interesting technological development can be observed by Photowatt’s solar modules. Until 1996, Photowatt was a small company with approximately three MW

---

[26] A solar module's energy conversion efficiency renders the ratio between emitted power and irradiated power of a solar module based on the module surface.

[27] Warranty against power degradation: PV module show little degradation over many years of operation. Within period of warranty manufacturers provide that the module is still producing 80% or more of its name plate rating.

production capacity and 80 employees[28]. Correspondingly, the quality level of Photowatt remained almost unchanged between 1987 and 1998. In 1997, Photowatt was acquired by ATS[29]. As a result of this acquisition, Photowatt has rapidly expanded capacity. Also the quality of solar modules has been improved significantly achieving the leading levels in 2003.

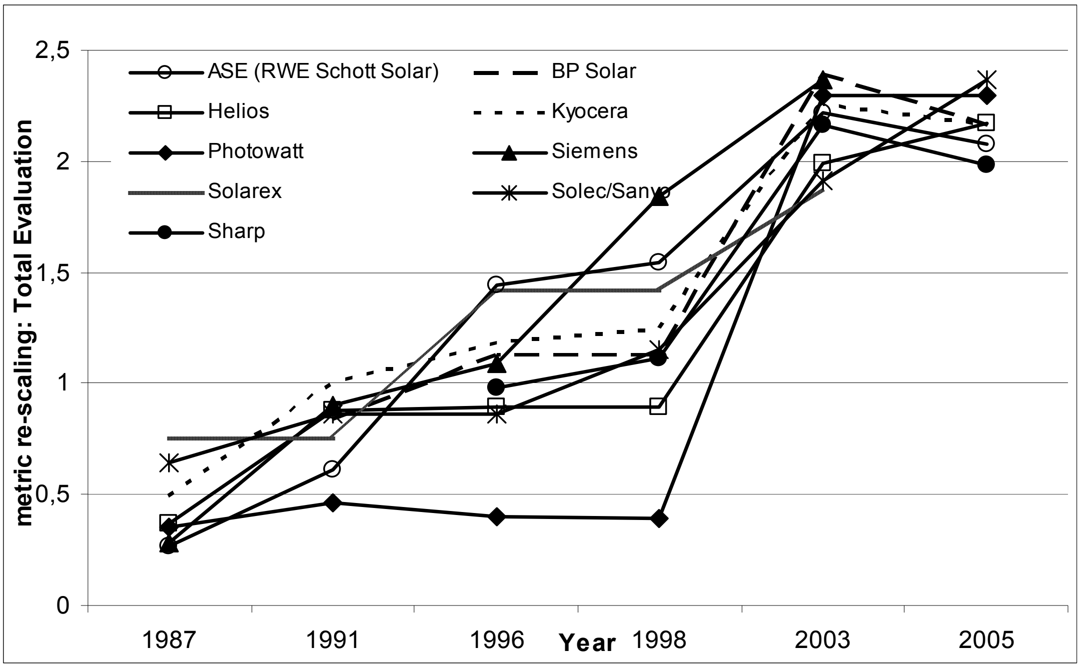


Figure 4-5[30]: Product quality according to the metric re-scaling approach[31].

In 1987, inter-firm product differentiation in the quality of solar modules was not at a high level. This is a typical property for an initial R&D-intensive market. At the beginning of market formation, there are only a few pioneer suppliers dominating the market. The high production costs caused a restriction in the variety of goods. Nevertheless, from eight solar modules which were offered on the market in 1987, ASE and Siemens had comparatively low performance values in quality. Solarex and Solec had significantly better values ($p=0.05$). Eleven years later, in 1998, some products improved tremendously. The product variety achieved its maximum. BP Solar was able to achieve the highest quality index of 2.4 in the sample due to high efficiency and low weight per watt proportion. The aggregated quality index is significantly better than the average at a significance level of 0.05. Later, quality and

28 See www.photowatt.com. Accessed on 18.11.2008.
29 Automation Tooling Systems Inc.
30 The data for solar modules produced by BP Solar are available from 1991 and by Sharp from 1996.
31 Note that years are not equidistant.

standardization became stricter for a wide variety of products, and the product selection mechanism reduced the variety of solar modules once more.

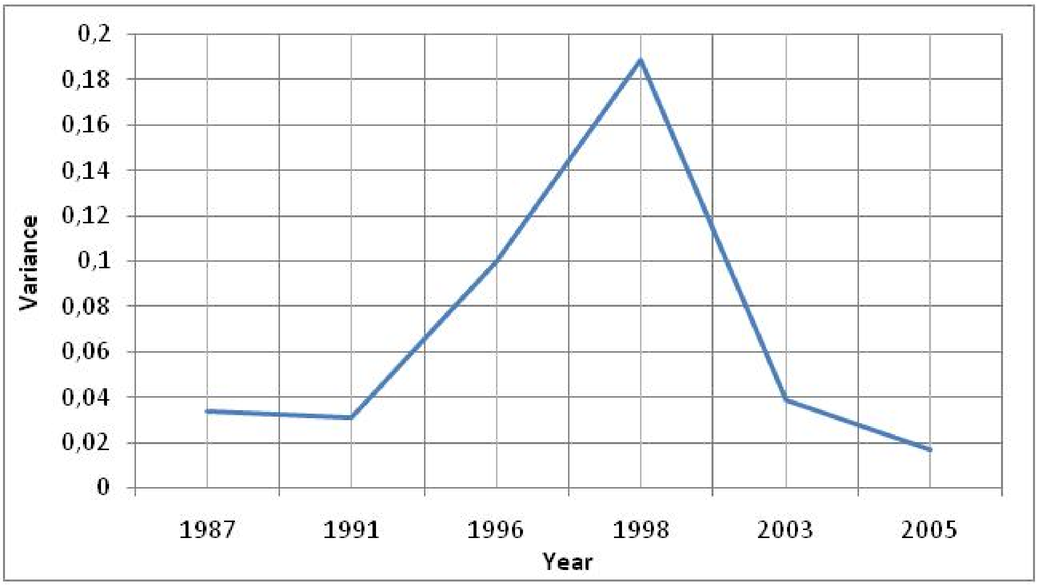


Figure 4-6: Variance of the quality index values[32] for solar modules[33].

Figure 4-6 shows the development of the inter-firm variance of the re-scaled quality index for the period of time from 1987 to 2005 and the nine producers, referring to the products displayed in Figure 4-5. Because the technical best PV module in the corresponding category has been selected for every producer, one can assume that the inter-firm variance in Figure 4-6 conforms to the German PV modules market variance in the respective year. Product differentiation in terms of product characteristics on the PV market went through different phases of the product life cycle. Since the PV market is a market with different customer preferences, customers were insecure in their decision about product choice. Consequently, the product characteristics were not homogeneous (see Earl and Wakeley, 2007). In this special case, Figure 4-6 shows that greater product differentiation occured in maturing phases of market development and then disappeared as a consequence of the standardization processes. The expansion and reduction phases of product variety seem to be related to dominant design formation and the double-hype effect in the PV technology cycle (see Haller and Grupp 2007).The total investigated period is

32 Calculated by metric re-scaling approach

33 Note that years are not equidistant.

divided in two phases: 1974 to 1990 and 1991 to 2005 (see chapter 3). This splitting does not conform with Figure 4-6. However, it is not unexpected that development in promotion activities and actual market happennings are separated in time.

This is in contrast to the common assumption that new product markets start with considerable differentiation which then matures one-way. This does not mean, however, that the observed pattern is universal. Further product differentiation may occur again, because producers "try to avoid the full brunt of price competition" (Vernon 1966) and identify more niche markets[34]. Moreover, manufacturers can move from standard products to product variety as a result of specialization (Saviotti 1996). Products are adapted to individual requirements of customers and become more specialized. Specific characteristics of niche products can bring additional value for some customers, so that they accept advanced prices in comparison with incumbent products. Sanyo[35] has reported, for example, that it will add a new type HIT[36] solar cell that uses multi-crystal silicon-based wafers in 2007 to meet a wide range of customer needs. In this way, Sanyo will not only increase production capacity but also offer solar cells with efficiencies higher than 22 %. Such solar modules can be products of choice in a residential application with limited roof area (Lunch 2007). At the same time, Sanyo will make thinner wafers in order to deal with the problem of the silicon shortage.

In the next step, the dynamics of the quality index over the time period between 1987 and 2005 are examined. There are two ways to do this: in the first option, the pure modifications in the quality indices can be calculated for each manufacturer. This method, however, does not reveal how far a technically underdeveloped producer was able to make up ground to the most innovative manufacture. In the second alternative, one can calculate relative changes in the quality indices in relation to the absolute technical progress in the industry. This approach measures the technical progress which is achieved by each producer in a reference period in relation to the absolute progress in the industry. The last possibility has obvious advantages and is chosen in the present study. The following equation is applied:

$$K^0 = \ln \frac{K_{t+1}^* / K_t^*}{K_{t+1}^{\max} / K_t^{\max}} \tag{4-5}$$

34 In this analysis, solar cells from developing countries are neglected due to lack of data.

35 http://www.sanyo.co.jp. Accessed on 25.09.2007.

36 HIT (Heterojunction with Intrinsic Thin-layer) hybrid solar cells are composed of a thin single-crystal silicon wafer surrounded by ultra thin amorphous silicon layers.

where K^0: change in quality index in time span *[t; t+1]*

K_t^* : value of quality index in year *t*

K_t^{max} : maximal value of quality index in year *t*.

The positive (negative) values for K^0 mean above-average (below-average) development of product quality of the respective producer in relation to the absolute technological progress in the sample. Indeed, it holds that $K_t^* \leq K_t^{max}$ for all *t*, but this relation does not stay in effect conclusively for the differences. Table 4-2 shows the results.

	1987-1991	1991-1996	1996-1998	1998-2003	2003-2006	1987-2006
RWE SCHOTT[21]	0.53	0.49	-0.18	0.10	-0.05	2.44
BP SOLAR		-0.07	0.12	0.12	-0.09	1.08[37]
HELIOS	0.57	-0.36	-0.25	0.55	0.10	1.84
KYOCERA	0.44	-0.20	-0.19	0.33	-0.03	1.41
PHOTOWATT	-0.02	-0.50	-0.28	1.52	0.01	2.06
SIEMENS	0.87	-0.18	0.28	-0.25		2.07[38]
SOLAREX	-0.28	0.27	-0.25	0.03		0.80[38]
SOLEC/SANYO[39]	0.01	-0.37	0.05	0.45	0.01	1.18
SHARP			-0.14	0.40	-0.07	1.21

Table 4-2: Quality changes of solar modules, 1987-2006

The solar modules of RWE SCHOTT SOLAR boast the greatest progress from a technical point of view. Through formation of the joint venture RWE SCHOTT SOLAR, it was possible to collect knowledge and experience in solar electricity over 50 years. Already in the late 1950s AEG[40] developed photovoltaic solar cells for space travel. The solar modules from Siemens have also undergone considerable progress from a technical point of view. Nevertheless, Siemens Solar was bought out by Shell Solar in 2002. The solar modules of Solarex show the lowest technical

[37] Changes are calculated for the period of time between 1991 and 2006.

[38] Changes are calculated for the period of time between 1987 and 2003.

[39] In 1994, Solec combined forces with Sanyo Electric, Japan. Both companies share a commitment to excellence and making strides toward improvements in the solar energy industry. Before 1994 only the solar modules from Solec were considered.

[40] AEG was later integrated in NUKEM and then in 2002 RWE SCHOTT Solar. More details can be found here http://www.schottsolar.com/. Accessed on 18.11.2008.

progress. Only these solar modules show a relatively lower technological improvement in comparison with the absolute technical progress.

This type of analysis has some caveats. It is an open question whether technical progress does not only change the characteristic parameters, but the set of characteristics itself. From reading of technical literature and market studies, there is no indication that this is so in the relatively short period of time between 1987 and 2005. It could well be the case that in the long run, i.e. in a period of time of more than 18 years covered, the issue of variance changes will look different. Further, despite of the overall demand for more efficient solar cells, limited product availability may be a problem for aggregate demand, if customer wants are not satisfied with certain preferred features (Tenn and Yun 2008).

In order to make a statement about innovation as a success factor, the results of Table 4-2 may be compared to the development of producer's market shares. This issue is the subject of the next section.

4.4 Development of Market Shares of Solar Cell Producers

By using the metric re-scaling approach and comparison of the results with changes in market shares of solar cell's producers, it can be investigated whether innovative solar module producers can disproportionately benefit from the growth rates in the PV market. On the global level, PV production is a rapidly growing market. The world production of solar cells soared to 2,536.6 MW, which equals a growth rate of 40 percent in 2006. Though it is less than 45 percent of the preceding year, 50 percent are listed for the year 2007. The question is how can the innovative solar cell's producers participate in this development?

In order to answer this question the same set of solar modules is selected and the change rate of the metric indicators between 1988 and 2006 is compared to the manufacturers' specific change rate in market share. In this way it is possible to test the hypothesis whether innovation is a factor of success in the PV industry. Thus in the current section again nine manufactures offering solar modules are considered. These producers account for over 50 percent of the world market in each year.

Figure 4-7 shows the development of the market shares. Until 2004, the rapid increase of Sharp's production capacity is remarkable. In 2000, the Japanese company eclipsed Kyocera (also a Japanese producer) and Siemens Solar took the top position among the global manufacturers. Since then, Sharp holds the position of the world's leading provider of solar cells. Still, in the years 2005 and 2006 Sharp was not able to increase its market share and lost 10 percent, by moving from 27 percent in 2004 to 17 percent in 2006.

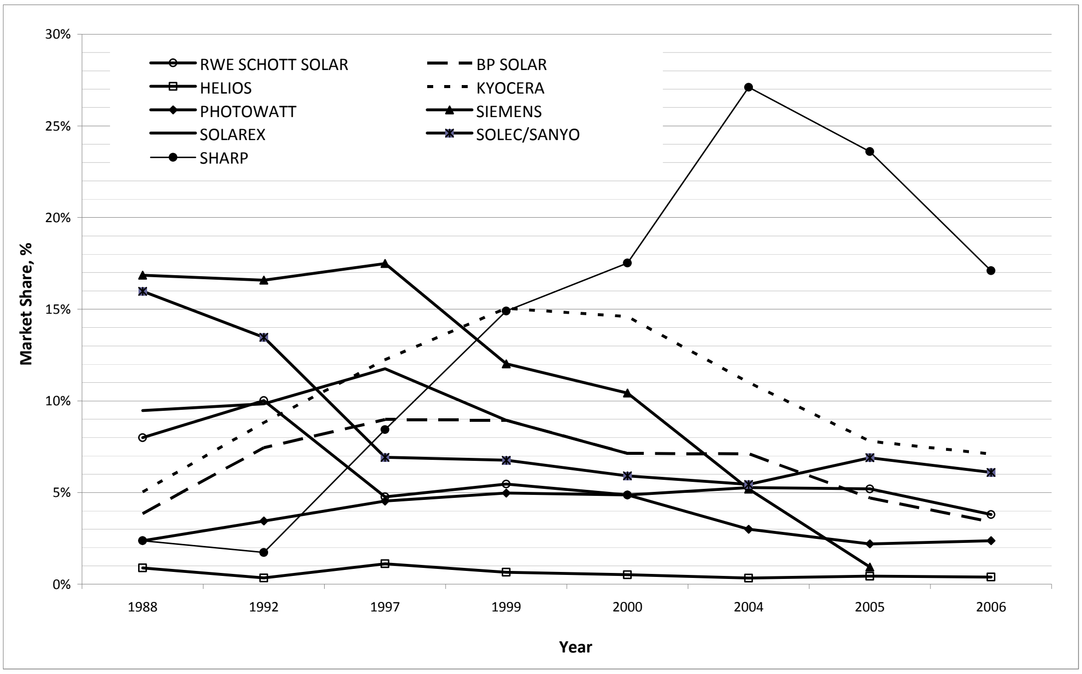


Figure 4-7: Share of PV Cell Production by Company, %[41].

Compared to the strong position of Japanese companies, the development of the production capacity in Europe and USA in the sample is rather moderate[42]. Over the complete time frame, neither European nor American companies had gained additional market shares. With the exception of Photowatt, all producers have lost some market shares. This finding, however, is valid only for companies in this sample. In general there are numerous new companies on the market which capture the market in relatively short time. Table 4-3 lists the top ten producing companies in 2006 and the first half of 2007. The actual top five producing countries are Japan, Germany[43], China, Taiwan, and the United States (see Figure 4-8). It is not possible, however, to include the new companies in a longitudinal study because the products of these firms have not been so long in the market.

41 Source: PV News, Photon International.

42 An exception constitutes, for example, the German solar cell producer Q-Cells. Founded in the end of 1999, the company is now the largest solar cell manufacturer in the world. But this firm is not included in the longitudinal analysis.

43 Figure 4-8 shows the share of Europe in worldwide photovoltaic product. However, the dominant position of Germany in Europe and in the World is well known. In 2006, 507.6 MW were produced by German PV solar producers. Therefore, the European lead position is the result of German production volume.

However, the number of new entrants to the market and the fact that such companies are able to increase their market share and compete with incumbent producers can be taken as an indicator for indication of the existence of competition in the market (see also section 4.2).

Company	Country	Production, %	
		2005	2006
Sharp	Japan	23.60%	17.10%
Q-Cells	Germany	9.10%	10.00%
Kyocera	Japan	7.80%	7.10%
Suntech	China	4.50%	6.30%
Sanyo	Japan	6.90%	6.10%
Mitsubishi	Japan	5.50%	4.40%
Motech	Taiwan	3.30%	4.00%
Schott Solar	Germany	5.20%	3.80%
Solar World	Germany	2.10%	3.50%
BP Solar	USA	4.70%	3.40%

Table 4-3: Photovoltaic production of the top ten producing companies[44].

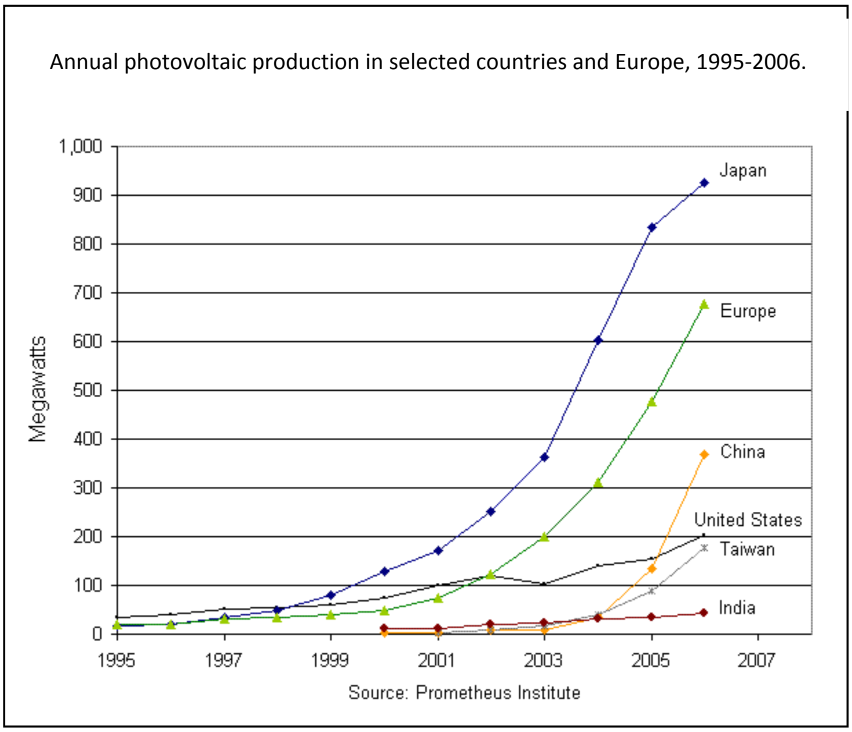


Figure 4-8: Annual photovoltaic production by country, 1995-2006[45].

[44] Source: Photon April 2007.

In the next step, changes in the market shares are calculated using data in Figure 4-7. The results of these calculations are listed in Table 4-4.

	1988-1992	1992-1997	1997-1999	1999-2004	2004-2006	1988-2006
RWE SCHOTT SOLAR	25.4	-52.4	14.6	-3.5	-27.9	-52.4
BP SOLAR	93.7	20.6	-0.5	-20.5	-52.2	-11.6
HELIOS	-60.7	217	-70.8	-49.2	18.2	-56.2
KYOCERA	75.1	39.0	23.0	-26.9	-35.5	41.2
PHOTOWATT	45.9	31.2	9.6	-39.6	-21.0	0.1
SIEMENS SOLAR	-1.7	5.5	-31.3	-56.8		
SOLAREX	3.9	19.5	-24			
SOLEC/SANYO	-15.7	-48.7	-2.3	-19.5	12.1	-61.8
SHARP	-27.0	387.9	76.9	81.9	-36.9	622.5

Table 4-4: Changes of the market shares for solar cell producers in %.

Table 4-4 shows a mixed picture. Siemens Solar was one of the big players in the solar industry for a long time. Unfortunately it was not possible for this company to keep this position because of high manufacturing costs which were a result of costly halfway manual solar cell production. As a consequence, Siemens was displaced from the market. The same applies to Solarex. Solec also lost market share. This can be explained by the fact that all Solec assets were purchased by Sanyo in 1994. This takeover led to a decline in production. Helios, RWE Schott Solar[21] and BP Solar also decreased their market shares. Sharp, Kyocera, and Photowatt gained additional market shares at the same time. These results are confronted with the results of Table 4-2 in five growth periods. Since it usually takes some time until advanced products find awareness by buyers, a lag of one year is introduced in comparing technical progress in solar modules and the development of market shares for the relevant producers.

[45] Source: Earth Policy Institute.

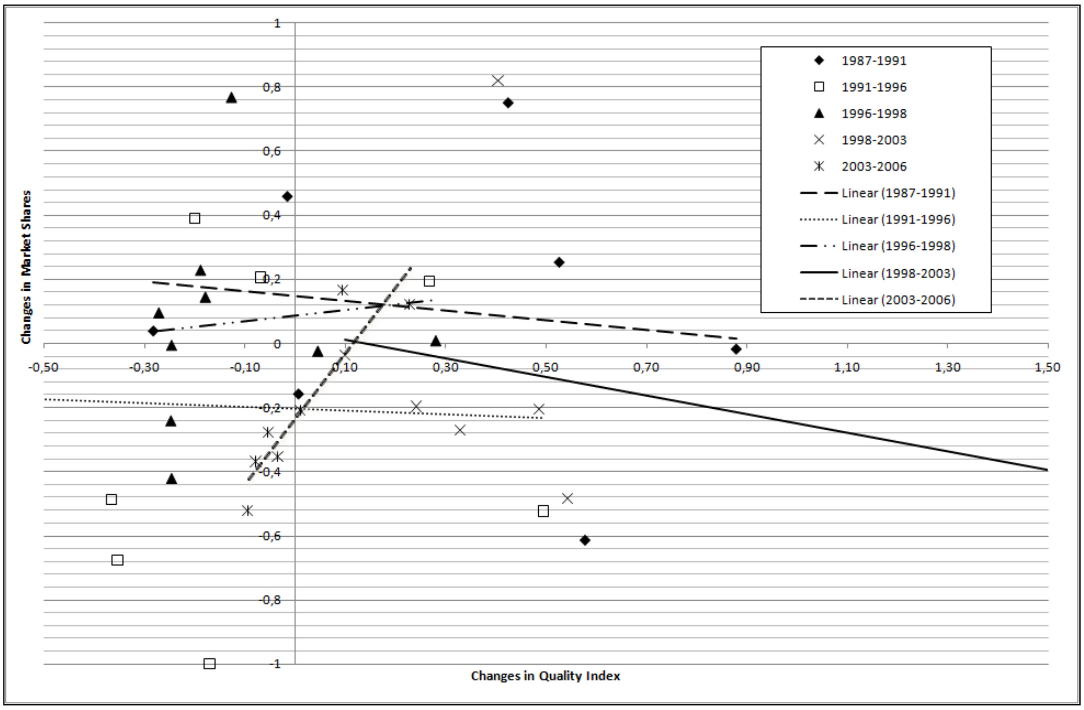


Figure 4-9: Comparison of quality and market position of solar cell producers.

Figure 4-9 shows a differentiated picture. To make some results clearly visible, linear regression lines are added. For the first two growth periods (1987-1991 and 1991-1996), there is a negative correlation between technical progress in solar modules and the development of market shares. This means that the producers with their early innovative products do not have competitive advantages on the beginning market contrary to Schumpeter's (1911) notion of quasi-monopolies of the pioneers. In the case of photovoltaics, firms may have erroneously oriented themselves towards professional customers (electricity utilities), but not to households (Grupp 1998). However, this has changed in the course of time. For the time periods 1996-1998 and 2003-2006, there was a positive correlation between technical progress and the gains on the market. It looks as though the continuous improvement of performance over a long period of time has been paying off. An exception from this pattern is the period 1998-2003. In this period, a negative correlation between market shares and improvement in quality indexes can be identified. This time span includes the stock market crashes in 2002, which had a negative influence on the semiconductor industry. As most of the silicon used for PV cells derives from low quality materials left over from the semiconductor industry, in this way the photovoltaic market was also affected.

One should never ignore the heterogeneity of enterprises with respect to their capabilities (Penrose 1959) which, however, is not possible to be covered in this investigation due to lack of intra-firm data. When interpreting the results one should be aware that a composite quality index was taken by equally weighting the different characteristics. Alternatives to this approach would be the calculation of hedonic prices by correlating each single feature to market shares. A customer-specific weighting could be done empirically, for instance by applying conjoint analysis.

It is desirable to use some sort of market- and firm-specific control variables. This could not be achieved for the scope of the analysis. First, it was observed an international sample of large firms being subject to various national R&D subventions and electricity price regulation as well as crude oil prices (the lead currency in energy markets subject to the Dollar exchange rate to national currencies). Second, firm strategies are difficult to discern if solar cell business production is only a small intra-firm business area alongside other energy business fields. These types of refinement remain on the agenda for future research.

4.5 Conclusions

In order to remain competitive, firms constantly strike for new possibilities of creating higher profits and larger market shares. On the one hand, production growth and growing market shares are traditionally an indicator of progress. On the other hand, the future success of industries is based on the ability of companies to innovate. Nevertheless, innovation and economic growth do not always go hand in hand. There are some examples in the empirical literature which reveal positive and negative relationships between innovation and growth. The problem is that it takes time until innovation can affect economic growth significantly. The major explanation for this paradox is that the usage of technology, and not the generation of technology is important. Only through diffusion of innovative products across markets the biggest impact on growth can be achieved.

For decades, the economic literature has been looking for a linkage between innovation and growth. The aim of this study is to check from product quality whether producers of innovative products have competitive advantages in order to meet market demand and to gain additional market shares. The market of solar modules in Germany was selected as an empirical example.

The investigation was carried out in three stages: first, the efficiency of the market was examined using the coordination failure diagnostics (CFD) concept established by Grossekettler. The existence of competition on the PV market is proven by analyzing five processes defined in this concept. Although competition may work to drive out inefficient products, it is still necessary to clarify whether the remaining leading manufacturers also provide superior products. In order to quantify the technical competence of solar cell producers the quality of its products has to be screened.

The technological improvement of solar modules from different producers was analyzed in stage two. The metric re-scaling approach was used in this part of the research. In general, further consistent technical development in solar cell technology is clearly recognizable for all the producers in the sample over the time period between 1987 and 2006. Different phases in market development can be identified. In the infant market, there is almost no product differentiation to be recognized in the quality of solar modules. All suppliers offer products of similar quality. The differentiation of quality occurred in maturing phases of market development and then disappearred as a consequence of standardization processes. In future research, these phases should be compared to the time of formation of dominant designs.

In order to make a statement about innovation as a success factor to meet demand, the product development phases had to be compared with the evolution of producer's market shares. The picture which appears here is mixed. Roughly speaking, in three of five stages of market formation any positive correlation between changes in market shares and the improvement in solar module's quality can be identified. It is the phase between 1987 and 1996 and the period of time between 1998 and 2003. The first phase can be explained by misunderstandings or non-interaction between supply and demand. The second was possibly a result of stock market crashes in 2002 which had a negative influence on the semiconductor industry.

In contrast, innovative producers seem to benefit more from the quality of their products in the second phase, from 1996 to 1998 and from 2003 to 2006. Market share growth was mostly, but not exclusively a consequence of technological advantages in that period. So, it can be shown that innovation in science-based markets is often a recipe for serving long-term growth in demand, but not necessarily.

The case study has shown that there are good prospects for renewed research on product characteristics thanks to better data bases and access: For selected markets and years, data is increasingly available through the Internet. The floor to enter a new kind of empirical demand research topic is open. The kind of new evolutionary and also general economic questions and puzzles that may be further elaborated on seems to be quite broad.

5 Time Series Analysis of Innovation Dynamics

The previous section deals with technological progress in the solar PV market in Germany. The driving force of market development here is the competition between solar cells' manufactures. As a result of combining this competition with scientific efforts, the producers achieve certain improvements in technological characteristics of solar modules. From this point of view, a competitive environment provides stimulus for manufactures to pursuit for innovation. This consideration is in line with Hayek's concept of competition. Following Hayek, the competition is:

> "[…] a procedure for discovering facts which, if the procedure did not exist, would remain unknown or at least would not be used." (Hayek 2002[1], p.9)

In theory this means that as a consequence of competition, producers that are more efficient have higher profits then inefficient ones. In reality, the marketing strategies and individual characteristics vary from manufacturer to manufacturer. Furthermore, the willingness of customers to pay more for progressive products does not remain stationary over time. As a natural consequence of this, some firms can capture additional market shares and others lose their market position. Moreover, it is not always a question of technological quality.

Dynamic market conditions present entrepreneurs with new challenges. Consequently, firms have to deal with new risks and adapt their marketing strategies permanently. Enterprises that have difficulties confronting the changed conditions can be driven out of the market. Survival of the fittest activates the natural selection in the market. In this way, it leads to the evolution of the market. Section 5 shifts the focus of analysis to the interaction of variables describing the market development among each other and to the reaction of the market to the exogenous factors.

Chapter 2 gives an overview of the empirical studies that deal with development of science-based technologies. Despite different scientific backgrounds and market potential the development of some science-based technologies can be generalized in a functional reference model. However, this model has a purely descriptive nature. The main goal of this section is to investigate the interaction between different indicators introduced in the stylized model. Additionally, it is interesting to see how

[1] The original paper dates back to 1968.

the market reacts to the exogenous factors affecting for the market development in general. These influences include development of oil price, political decisions about development programs, and subsidies. In this context, three research questions are relevant:

– Is there “growth equilibrium” between any of the indicators that measured different activities of market development in the stylized model?
– Is it possible to measure the extent of this “growth equilibrium” and the interdependencies between indicators describing the market development in general?
– To what extent can the equilibrium between different indicators be affected by exogenous factors?

In order to answer these questions an econometric approach is chosen. The representation of the PV market processes in this section is given by selected variables that are presented as time series in equal intervals and in a common measurement system of market formation.

The chosen econometric approach has some disadvantages. Time series data only give us a very crude numerical picture of the complex econometric decision making on various levels. This data describes the behavior of market actors (firms and consumers) on an aggregate level. The economic agents try to reach their aims with given factor endowment. The required decisions are based on individual or collective needs and run through complex evolutionary selection mechanism, which is not observable as a whole. Also, there may be problems with data due to measurement and compilation errors.

Additionally, the utilization of econometric methods gives adulterated results if, for instance, a relationship between variables is measured by wrong functionality. One supposes for example linearity, but the variables depend on each other differently. Another problem that occurs relatively often when using economic methods often leads to misrepresentation of data. While analyzing the relationship between two variables, it is often very difficult to explain which origin this relationship has. For this reason, the interpretation of results can vary significantly and has to be regarded very carefully.

Nevertheless, econometric methods have a very essential advantage. On the one hand, they give economic background to empirical data, and on the other hand, they give empirical content to economic theory. Madala (1992) gives the following definition of econometrics:

> "Econometrics is the application of statistical and mathematical methods to the analysis of economic data, with a purpose of giving empirical context to economic theories and verifying them or refuting them." (Madala, 1992 p.1)

Furthermore, there are methods in econometric approach that make it possible to estimate causal relationships in economic variables. The performance of econometric models can be checked using statistical tests.

Figure 5-1, with reference to Spanos (1995), introduces the components of the modeling process in a schematic way. The process of modeling is not a sequential process, as model construction requires many feedback mechanisms. The intention of using Figure 5-1 is rather to visualize the relations between the components.

The modeling process starts with describing of underlying economic phenomenon. In this case, it is the formation process of scientific-based markets. This theoretical concept has to be transformed into an economic model. The economic model[2] simplifies the complex reality and documents stylized facts of market formation. In this case, it is the stylized model described in Figure 2-8. This standardized reference scheme provides basis for selection of relevant indicators describing scientific and technological activities and market situation. These indicators are observed time series reflecting the development of real-world components of an economic model. After a careful analysis of the validity and reliability of the time series data from the case study of the solar PV market, a statistical model can be established. This model is the result of pure empirical-statistical investigation and is therefore used for an economic interpretation only in strict limits. In the next step, the statistical model is transformed into a general econometric model under the consideration of various restrictions given by a theory. In this transformation process, the stylized model of market formation, as outlined above, can repeatedly be used.

2 This model is described in detail in chapter 2.

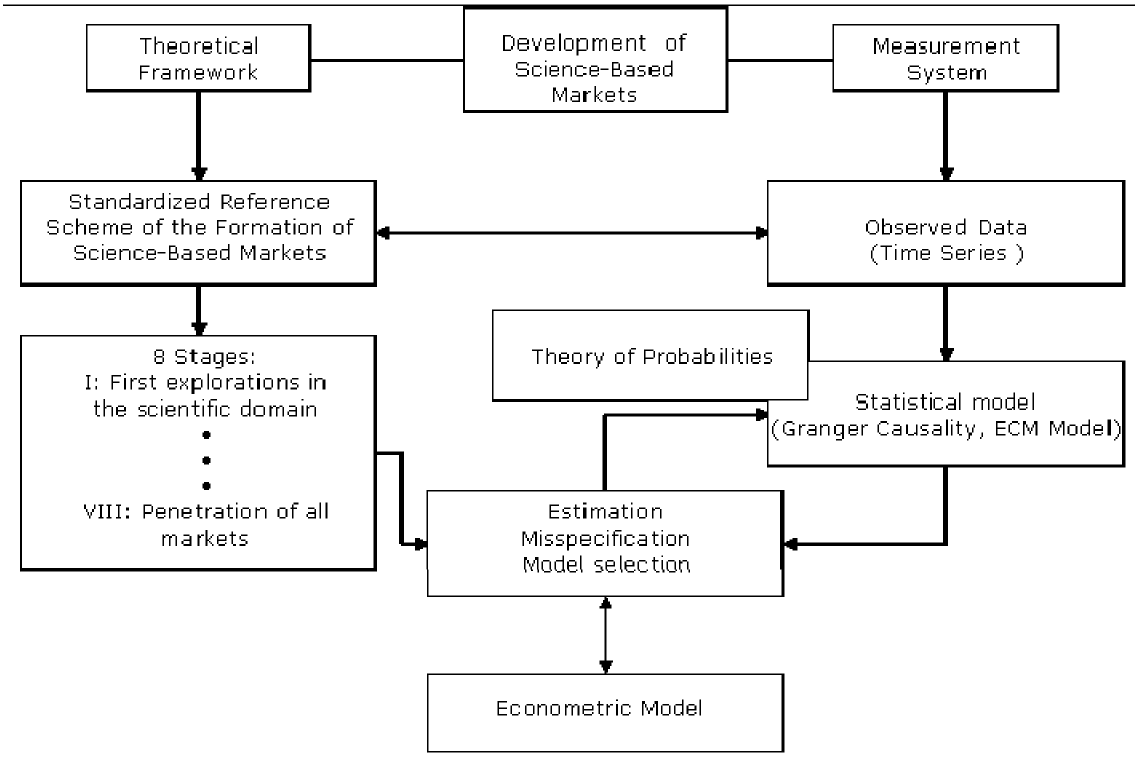


Figure 5-1: Components of the econometric modeling process[3].

The remainder of this section is organized as follows: In section 5.1 the main theoretical approaches applied in statistical modeling are outlined. There are unit root tests, granger test for causality, and co-integration concept. Section 5.2 gives a description of data collection. Section 5.3 deals with construction of statistical modeling and testing of test hypothesis. Finally, section 5.4 summarizes the results of this section.

3 Following Spanos (1995).

5.1 Statistical Methodology

5.1.1 Unit Root Tests

The time series in this section describe the evolution of the solar PV market over the course of 30 years. The impact of exogenous factors, for example fuel oil price, political decisions, technical progress in development of more innovative products, and individual decisions of economic actors (firms and consumers), creates new information and new conditions for the market development. This information may produce quick and irreversible changes. For this reason, it can be assumed that most of the time series describing long-term development are non-stationary. In contrast to stationary series, non-stationary series have typical characteristics: they do not return to a constant value or a given trend. Whenever there is non-stationarity in data, the using of classic regression techniques is not valid, because these methods are based on the assumption that the means and variances of variables are constants and do not change over time. Otherwise, the problems with spurious regression arise.

Rao (1994) gives the following explanation of the spurious regression problem. The means and variances of non-stationary times do not stay constant. Consequently, all computed summary statistics, in which these terms are used, are time dependent and do not converge to their true values as the sample size increases. Therefore, pre-testing the variables for presence of non-stationarity has become important in the applied econometric work.

There are different types of non-stationary time series. One can consider a deterministic linear trend process:

$$y_t = \gamma_0 + \gamma_1 t + u_t \qquad (5\text{-}1)$$

Or a stochastic trend:

$$y_t = \gamma_1 + y_{t-1} + u_t \qquad (5\text{-}2)$$

Whereby t is linear time trend and u_t a white noise[4] error term.[5] The process defined in equation (5-1) can be transformed into a stationary process by subtracting the trend $\gamma_0+\gamma_1 t$. For this reason, such processes are denoted as trend stationary

[4] White noise: a time series in which each element is independent draw from a distribution with mean zero and constant variance. (see also Kennedy 2008, p. 510)

[5] See Davidson and MacKinnon (1993): p. 700; Jungmittag (1995), p. 244.

processes. For the second processes that defined in equation (5-2) de-trending cannot achieve a stationary process. In this case a stationary process may be generated by differentiating the data:

$$\Delta y_t = y_t - y_{t-1} = \gamma_1 + u_t \qquad (5\text{-}3)$$

The process in equation (5-2) is called "differences-stationary" if the transformation in equation (5-3) generates a stationary process. In this context, a definition of integrated series can be given:

Series y_t is integrated of order one (denoted by $y_t \sim I(1)$) or contains a unit root if y_t is non-stationary, but Δy_t is stationary. Or, respectively:

Series y_t is integrated of order d (denoted by $y_t \sim I(d)$) or contains d unit root if y_t is non-stationary, but $\Delta^d y_t$ is stationary[6].

In order to answer the question: Is the trend stochastic or deterministic? It has to be clarified whether the underlying data-generating process (DGP) of a time series corresponds to equation (5-1) or equation (5-2). This classification can be carried out by consideration of general model. This model includes both cases (5-1) or (5-2). (cp. Jungmittag 1995, p. 246):

$$\begin{aligned} y_t &= \gamma_0 + \gamma_1 t + \upsilon_t \quad with \quad \upsilon_t = \alpha\upsilon_{t-1} + u_t \\ &= \gamma_0 + \gamma_1 t + \alpha(y_{t-1} - \gamma_0 - \gamma_1(t-1)) + u_t \end{aligned} \qquad (5\text{-}4)$$

Whereby u_t is a stationary process. The model in equation (5-4) can be transformed as follows:

$$\begin{aligned} y_t &= \gamma_0 + \gamma_1 t + \alpha y_{t-1} - \alpha\gamma_0 - \alpha\gamma_1(t-1) + u_t \\ &= \gamma_0 - \alpha\gamma_0 + \alpha\gamma_1 + \gamma_1 t - \alpha\gamma_1 t + \alpha y_{t-1} + u_t \\ &= \underbrace{\gamma_0(1-\alpha) + \alpha\gamma_1}_{\beta_0} + \underbrace{\gamma_1(1-\alpha)}_{\beta_1} t + \alpha y_{t-1} + u_t \end{aligned}$$

The result is:

$$y_t = \beta_0 + \beta_1 t + \alpha y_{t-1} + u_t \qquad (5\text{-}5)$$

Another notation from equation (5-5) is:

$$\Delta y_t = \beta_0 + \beta_1 t + (\alpha - 1) y_{t-1} + u_t \qquad (5\text{-}6)$$

6 where Δ is the difference operator.

If $\alpha < 1$ then equation (5-6) corresponds with equation (5-1); if $\alpha = 1$ then equation (5-6) conforms to (5-2). As a consequence, the null hypothesis $\alpha = 1$ can be defined. The alternative hypothesis is $\alpha < 1$. The corresponding test is known as unit root test. Dickey and Fuller (1979) and Fuller (1976) developed the basic test for unit roots and order of integration. This approach also includes the following equations:

$$\Delta y_t = (\alpha - 1) y_{t-1} + u_t \tag{5-7}$$

And

$$\Delta y_t = \beta_0 + (\alpha - 1) y_{t-1} + u_t \tag{5-8}$$

Equation (5-7) is very restrictive and is not common in economic theory. (5-8) can be applied only to time series without a trend.

Correspondingly, Dickey and Fuller proposed three models which are also known as τ, τ_μ, and τ_τ. The null (H_0) and alternative (H_1) models in each case are:

(i) H_0: $y_t = y_{t-1} + e_t$
H_1: $y_t = \varphi y_{t-1} + e_t$, $\varphi < 1$
This is a test for a random walk against a stationary autoregressive process of order one (AR(1)).

(ii) H_0: $y_t = y_{t-1} + e_t$
H_1: $y_t = \mu + \varphi y_{t-1} + e_t$, $\varphi < 1$
This is a test for a random walk against a stationary AR(1) with a constant term.

(iii) H_0: $y_t = y_{t-1} + e_t$
H_1: $y_t = \mu + \lambda t + \phi y_{t-1} + e_t$, $\phi < 1$
This is a test for a random walk against a stationary AR(1) with drift and a time trend.

The tests are based on the t-ratio of the y_{t-1} term in the estimated regression of Δy_t on y_{t-1} in equation (5-6) (with $\beta_0 = \beta_1 = 0$ in case (i), and $\beta_1 = 0$ in case (ii)). The test statistic is defined as:

$$\text{Test statistic} = \frac{\hat{\gamma}}{SE(\hat{\gamma})} \qquad \text{with } \gamma = \alpha - 1 \text{ in (5-6).}$$

The test statistic does not follow the usual t-distribution. Critical values are derived from Monte Carlo experiments in, for example, Fuller (1976, pp. 371, 373). However, the unit root test described above is valid if the time series is well characterized by an AR(1) with white noise errors. Many time series have a more complicated structure with the underlying AR-Process of a higher order. In this case, Dickey and Fuller[7] (1979) propose an alternative model to (5-5) that can be written as:

$$y_t = \beta_0 + \beta_1 t + \sum_{i=1}^{p} \alpha_i y_{t-i} + u_t \tag{5-9}$$

A problem now arises in determining the optimal number of lags p of the dependent variable. The order of p can be chosen by minimizing information criteria such as Akaike or Schwarz. Another possibility is a sequential search procedure to determine the best lag p. In practice, p has to be large enough to eliminate possible serial correlation in error term u_t.

5.1.2 The Granger Test for Causality

In section 5.1.1 univariate time series models were discussed. In the next step bivariate relationships between variables have to be investigated. The focus of this subsection is interdependencies between indicators describing the state of the market development. It is a question concerning a causal relationship between indicators. The problem is that causality concept is philosophically ambiguous and controversial (see Stier 2001, p. 83). There is an enormous amount of literature on causality, but it is popularly accepted that correlation between two variables is not equal with causality. Causality in common sense represents the principle of cause and effect including direction of this relation. The problem of this concept occurs by its application in economic context. The interdependencies between economic quantities arise in an uncontrolled experiment. It is difficult to imagine that economic theories can be tested under laboratory conditions[8].

Granger (1969) introduced a statistical concept of causality. This concept is focused on predictability of a time series x_t. According to Granger Causality:

[7] This test is also known as augmented Dickey-Fuller (ADF) test.

[8] There is a field of research called “experimental economics” which uses experimental methods to evaluate predictions of economic behavior. But this research also has limits because it is not possible to map all conditions of natural environments where normal economic decisions take place.

"If some other series y_t contains information in past terms that helps in the prediction of x_t and if the information is contained in no other series used in the predictor, then y_t is said to cause x_t" (Granger 1969, p.378).

The basic idea of this definition is quite simple and it is assumed that the cause cannot occur before the effect. But at the same time, it is the critical point of this approach. It should be noted that Granger causality is not equal to the causality in common sense of the term. In economy, it happens very often that the variables in the model react to some unmodeled factor (for example war or election result, etc.) and if the response of x_t and y_t is staggered in time, Granger causality can be observed though the real causality is different. Unfortunately, it is not possible to solve this problem. Granger causality measures whether one event happens before another event and helps predict it. However, one assumes that it partly conforms with real causality in the process (see Sørensen 2005).

Granger (2003, p. 365) gives in his Nobel Lecture the following statement about two components of causality:

1. The cause occurs before the effect; and

2. The cause contains information about the effect that that is unique, and is not present in any other variable.

Granger causality[9] can be described and tested in the context of a linear regression model. For illustration, consider the following autoregressive time series x_t with the prediction based on m past measurements:

$$x_t = \sum_{i=1}^{m} a_i x_{t-i} + \varepsilon_{x,t} \qquad (5\text{-}10)$$

Here $\varepsilon_{x,t}$ is the prediction error whose magnitude can be evaluated by its variance $var(\varepsilon_{x,t})$. Let y_t be another time series. In the next step, consider the following prediction of the current values of x_t based both on its own past values and on the past values from y_t:

$$x_t = \sum_{i=1}^{m} a_i x_{t-i} + \sum_{i=1}^{m} b_i y_{t-i} + \varepsilon_{x|y,t} \qquad (5\text{-}11)$$

9 In literature, one can find numerous references to Wiener (1956) definition of causality. Granger (2003, p. 365) says that this definition was taken as a basis for his concept.

If the prediction improves by incorporating the past values of y_t, this means that $var(\varepsilon_{x|y,t}) < var(\varepsilon_{x,t})$) in some suitable sense, then it can be assumed that y_t has a causal influence on x_t. Similarly, this concept can be used for y_t:

$$y_t = \sum_{i=1}^{m} b_i y_{t-i} + \varepsilon_{y,t} \tag{5-12}$$

$$y_t = \sum_{i=1}^{m} c_i x_{t-i} + \sum_{i=1}^{m} b_i y_{t-i} + \varepsilon_{y|x,t} \tag{5-13}$$

And again, one can say that x_t has a causal influence on y_t if $var(\varepsilon_{y|x,t}) < var(\varepsilon_{yt})$. Equations (5-11) and (5-13) can be noted together as a vector autoregressive model (VAR). One can also say that y_t (or x_t) granger causes x_t (or y_t) if coefficients b_i (or c_i) are significantly different from zero. This can be tested with the F-Test with the null hypothesis:

H$_0$: $b_1=b_2=\ldots=b_m=0$

and the alternative hypothesis

H$_1$: there is a coefficient b_i so that $b_i \neq 0$.

Let RSS_1 and RSS_0 be the respective sum of squared residuals defined as:

$$RSS_1 = \sum_{t=1}^{T} (\hat{\varepsilon}_{x|y,t})^2 \qquad RSS_0 = \sum_{t=1}^{T} (\hat{\varepsilon}_{x,t})^2$$

Then the test statistic is:

$$S_1 = \frac{(RSS_1 - RSS_0)/m}{RSS_1/(T-2m-1)} \sim F(m, T-2m-1) \tag{5-14}$$

If the test statistic is greater than critical values, reject the null hypothesis that y_t does not granger cause x_t. Model selection criteria (for example The Akaike Information Criterion (AIC)[10] or the Schwarz Criterion (SIC)[11]) can be used in order to determine the optimal order m.

[10] See Akaike (1974)

[11] See Schwarz (1978)

5.1.3 Co-integration Analysis

After considering bivariate relationships between variables, this subsection deals with the problem of non-stationarity in time series. In section 5.1.1 the diagnostic methods of non-stationarity are discussed. However, it is more important to find appropriate statistical technique which can handle this problem.

Generally, there are two possibilities for solving the problem of non-stationarity in time series:
– Calculation of differences
– Co-integration concept

The first possibility proposes generation of differences in data and is commonly used in some applications. In this approach, one regards the relationship between differences, i.e. rates of increase. From a statistical point of view, it is a valid approach if the rates of increase are indeed stationary. But from an economical point of view, this method loses a lot of information about investigated quantities. Considering the rates of increase, a statement about short-run dynamics in a process can be made. It has less to say about long-run covariation of the variables. This is unsatisfactory because the information about the development of variables in terms of levels is lost.

Therefore, it was a challenge to find other methods which deal with this problem and to find a method that could analyze both perspectives at the same time: short- and long-term. Granger, during the 1980s, suggests such methodology for statistical analysis. The key idea in this method is that a specific combination of two (or more) non-stationary series may be stationary. This finding corresponds with the general assumption in economic theory about the existence of an equilibrium relationship between two economic variables. It can happen that these variables deviate from the equilibrium in the short term. Nevertheless, in the long term they will adjust towards the equilibrium. There are some examples of variables that have long term relationships to each other: income and consumption, prices for substitute goods[12], price for one good in different markets. However, the great advantage of the co-integration concept is that this concept also offers powerful statistical methods for estimation and testing of a hypothesis.

[12] Classic example of substitute goods is petroleum and natural gas. Both goods are used for heating or electricity.

Engle and Granger (1987) outline the co-integration concept, estimation procedure and tests. According to Engle and Granger (1987) the co-integration for a set of *N* variables is:

> "Definition: The components of the vector x_t[13] are said to be co-integrated of order *d*, *b*, denoted $x_t \sim CI(d, b)$, if (i) all components of x_t are $I(d)$; (ii) there exists a vector a ($\neq 0$) so that $z_t=a'x_t \sim I(d-b),\ b>0$. The vector a is called the co-integrating vector." (Engle and Granger 1987, p. 253).

Engle and Granger (1987, p. 254) emphasize that a reduction of the integration order is a special characteristic of a co-integrating relationship. If the economic variables in x_t are stationary, then it is trivial that every linear combination of vectors generates a stationary process. However, this has nothing to do with the co-integration concept. In this case one can use the classic regression methods. A necessary condition of existence of a co-integration relationship is that the single economic variables in a vector have the same order of integration. However, this condition is not sufficient. This finding is also confirmed by Stier (2001, p. 315).

If $a'x_t =0$, then economic variables are in equilibrium. Nonetheless, in most of the time periods, x_t is not in equilibrium, so that the co-integration vector $z_t=a'x_t$ represents the short-run deviation from the long-term equilibrium. Engle and Granger (1987, p. 252) call it "the equilibrium error." This error should be small and the term z_t has to be stationary, if a co-integration relationship between economic variables exists.

In general, *N* variables can have maximum *N-1* co-integration relationships. For two variables x_t and y_t with integration order I(1), the co-integration relationship can be written as $y_t=\beta_0+\beta_1 x_t$ and the co-integration vector can be $(1,-\beta_1)$. The first coordinate of co-integration vector is usually normalized to one. In this way, the uniqueness in presentation can be achieved, because if z_t is stationary, then every product of z_t with a scalar $c \neq 0$ is also stationary.

However, the co-integration concept describes only long-term relationships between economical variables. Additionally, it can be very interesting to formulate a model that combines both long-term and short-term behavior of the variables. This characteristic has the error correlation model (ECM). In this model the

[13] x_t is a vector of economic variables.

disequilibrium from one period is partly corrected in the next period (cp. Engle and Granger (1987, p. 254). In this thesis, a bivariate ECM is used. According to the Granger representation theorem (Engle and Granger (1987, p. 255), two or more integrated time series that are co-integrated have an error correction representation. A bivariate ECM can be written in the following form:

$$\Delta y_t = \alpha_0 + \alpha_1 \underbrace{(y_{t-1} - \beta_1 x_{t-1})}_{=z_{t-1}} + \alpha_3 \Delta x_t + \alpha_4 s_t + \varepsilon_t \tag{5-15}$$

In this model, the current changes in y_t are the function of the disequilibrium in the previous period, changes in x_t, and an impact of an exogenous factor s_t. Specifically, α_3 captures any immediate effect that x_t has on y_t. Therefore, this term of the equation describes the short dynamics of the relationship between x_t and y_t. The coefficient β_1 reflects the equilibrium effect of x_t on y_t and is estimated in the co-integration vector. The absolute value of the coefficient α_1 can be interpreted as the speed of adjustment parameter. The coefficient α_4 displays the impact of exogenous variables.

In the next step, the testing procedure of co-integration analysis is discussed. Firstly, the necessary condition of co-integration existence has to be proven, e. g. one has to prove whether the variables have the same order (see unit roots tests in 5.1.1). Assume that all variables have an integration order one. It is the case most frequently. In order to examine whether the variables are co-integrated with each other, the test procedure developed by Johansen (1995) can be used. The basic model in this procedure is the Vector Autogressive Model (VAR) of order p:

$$y_t = A_1 \cdot y_{t-1} + \ldots + A_p y_{t-p} + B \cdot x_t + \varepsilon_t, \tag{5-16}$$

Whereby y_t is a k-vector of non-stationary I(1) times series. x_t is a d-vector of exogenous variables that can include a constant term, linear trend, and other crisis variables. The x_t vector takes into account short-run shocks to the system. The error terms ε_t are independent, identically distributed with a mean of zero. The VAR-model (5-16) can be rewritten as:

$$\Delta y_t = \Pi \cdot y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + B \cdot x_t + \varepsilon_t, \tag{5-17}$$

where

$$\Pi = \sum_{i=1}^{p} A_i - I \qquad\qquad \Gamma_i = -\sum_{j=i+1}^{p} A_j$$

According to the Granger representation theorem, there are co-integration relationships if the coefficient matrix *Π* has a reduced rank[14] r<k[15]. Consequently, the Johansen-trace-test for r co-integrating vectors proves the rank of matrix *Π*. If *Π= α' β*, where *α, β* two (*kxr*)- matrixes and *r<k*, and hence the matrix *Π* has a reduced rank. The parameters of matrix *α* are often referred to as "speed-of-adjustment" parameters. Each column of the matrix *β* contains the parameters of one of the co-integration relationships.

This test is based on the log-likelihood ratio $ln[L_{max}{}^{16}(r)/L_{max}(k)]$, and is conducted sequentially for $r = k\text{-}1,...,1,0$. The null hypothesis is that co-integration rank is equal to r against the alternative hypothesis that the co-integration rank is *k*. The alternative hypothesis means that the variables in y_t are stationary[15].

The rank of the matrix *Π* is equal to the number of its non-zero characteristic roots (eigen values). For this reason, a test for the number of co-integration relationships can be carried out by testing for the number of characteristic roots of the matrix *Π* (Johansen maximum eigen value test). This test is based on the log-likelihood ratio $ln[L_{max}(r)/L_{max}(r+1)]$, and is conducted sequentially for $r = 0,1,..,k\text{-}1$. This test has the same null hypothesis as the trace test, but the alternative hypothesis is that the co-integration rank is equal to $r+1$.

Given the rank of Π then (5-17) can be written as:

$$\Delta y_t = \alpha \cdot \beta \cdot y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + B \cdot x_t + \varepsilon_t, \qquad (5\text{-}18)$$

This equation can be taken as a basis for further discussions.

---

14 The number of linear independent rows and columns.

15 In case if r=k, e.g. the matrix Π has the full rank, then every linear combination variables in y_t generates a stationary process. This would imply that all times series are stationary. A VAR in levels is appropriate. If the rank is zero, no kombination of the variables is stationary, and so a VAR in differences is appropriate (cp. Kennedy 2008)).

16 The maximum likelihood $L_{max}(r)$ is a function of the co-integration rank r.

5.2 Empirical Data

This section gives a short description of the variables used in this model. The number of scientific publications is commonly used as an indicator to quantify the relevant scientific activities. In section 6, publication data is also used in order to analyze co-author's networks. In this section this data is used on a more aggregate level as absolute numbers.

In order to collect data for the publication statistics for the PV market, the online version of the Science Citation Index (SCI, host STN) was chosen. Thereby, the following keyword search strategy[17] in titles: solar cell or solar cells or photovoltaic#[18] was used. Between 1974 and 2005 23,390 scientific articles in the technology field 'solar cells' were identified. The curve of these scientific activities has a double peak structure. The first peak was year 1984, after the year 1992 the number of scientific articles increased rapidly, the second extreme occurred in 2005. As seen, the scientific activities in the PV field continued to rise.

Patents are frequently used as innovation indicators, because patent records are publicly available and easily accessible. Moreover, patent data is classified by technical fields, and patent time series allow for the convenient study of historical trends. There are a lot of free and commercially available patent databases which are potentially helpful in research. The data of this analysis is based on the EPO[19] Worldwide Statistical Patent Database version October 2007 (PATSTAT October 2007 database). The decision to work with the PATSTAT was guided by two factors:

Firstly, The PATSTAT provides a range of details for patent records including names, addresses, citations[20], and claims. The PATSTAT also allows search of the relevant patent records by using key-words. Secondly, it is a new worldwide database which has been developed by EPO in cooperation with OECD and Eurostat and it is now accessible for statistical producers and researchers. The database consists of raw data on some 60 million patents.

---

17 The same search strategy was used in section 6 with Web of Science.

18 # serves as a placeholder.

19 European Patent Organisation

20 This data is used for identifying of science dependency of solar PV technology in section 2.

Due to the founding of the EPO in the year 1978, there are two ways in which an applicant can file patent applications in Europe. The one possibility is to register an invention directly at the national office, such as the DPMA. As an alternative, the applicant may file an international application at the EPO, in which he can designate different European countries in which patent protection is desired. Each option has its advantages. The best solution depends on the invention and the market where the company operates in.

In order to create long time series for patent activities in photovoltaic patent applications at the DPMA and the EPO have to be considered. The patent sample includes documents which were researched with the following retrieval strategy:

1. Patent records with the IPC = H01L 031/04 or IPC= H01L 31/06 in the main group or subclasses.
2. Patent records with the IPC = H01L 31 and the keyword "solar#[21]" in the title.
3. Union of the sets 1, 2.

There are 4,667 retrieved documents between 1968 and 2005. The graph of patent activities referring to the technology field 'solar cells' clearly qualifies as a 'double-boom' cycle. The first maximum (127 documents) was reached in 1986, then a decrease of patent activities is apparent. The second peak (313 documents) occurred 17 years later, in 2000. Since 2001, patent applications in technological field Photovoltaic have been showing a stagnating tendency. This can traced back to the new market crisis in 2001. This financial shock[22] had an impact on the global semiconductor industry worldwide.

Subsidies given by the government are very important to the PV market. A time series of public subsidies is the third relevant variable in the econometric model. According to the Public Promotion Catalogue of the Federal Ministry of Education and Research nearly 850 million EUR were spent by the government between 1975 and 2006 on photovoltaic promotion. About 60% of all public expenditures on photovoltaic energy were spent in the years between 1987 and 1997. Since 1993 subsidies are stagnant at the constant value.

[21] # serves as a placeholder.
[22] Also known as the "dot-com bubble".

There are different programs that support the PV industry in Germany. For example there was the 100.000 Roofs Program, which aimed to install 300 MW of solar cells by the end of 2003. A total of 350 MWp PV capacities were installed on more than 60,000 roofs under the program. The empirical data of this program is going to be included in the model as dummy variable. IWR Solarenergie[23] offers more statistics on this program.

Replacing the Electricity Feed Act, the Renewable Energy Sources Act (Erneuerbare-Energien-Gesetz; EEG) regulates the prioritization of grid-supplied electricity from renewable sources. These two of Germany's principal renewable energy support instruments are also used in the model.

The price of crude oil has a wide influence on the development of the PV market. Especially the first oil crisis in 1973 and the second oil crisis in 1979 revealed the fragility of energy supply systems of industrialized countries. The oil price rose dramatically in 1973-1974. Consequently, this led to major government initiatives in renewable energy in general with a view to mitigating the risk associated with heavy use of oil. As a result of this, firm behavior regarding consumption of energy from renewable energy sources began to change as well. In the 21st century this tendency has also been maintained. For example, one of Germany's major conventional power producers and one of the biggest polluters in Europe, Vattenfall states an ambition to play a leading role in renewable electricity and heat production. In 2007, the total share of renewable energy in Vattenfall's electricity generation was around 22% and in heat generation it was 12%[24]. The company makes efforts to increase these shares by investing in the research and development of renewable energy sources.

Installed capacities of photovoltaic systems are represented by the last indicator in the model[25]. According to AGEE-Stat[26] (statistics organization of the Ministry of the Environment) the German Photovoltaic Market reached 635 MW (906 MW) of installed solar power in 2004 (in 2005), bringing the cumulated total of installed German capacity to 1,074 MW (1,980 MW). Based on data collection of Photon International magazine, in 2004 (2005) there was 618 (909) MW of new solar

23 http://www.iwr.de/solar/markt/pvprog.html (accessed on 27.01.2010)

24 See Vattenfall Corporate Social Responsibility Report (CSR) 2007

25 There were a lot of contradictory statements in empirical data for this indicator. But according the last comparison of data the deviations between the installed PV capacity according AGEE-Stat and PHOTON are minor.

26 http://www.erneuerbare-energien.de/inhalt/2720/ (accessed on 27.01.2010)

capacity installed which corresponds with a total cumulated solar power of 1,052.67 MW in 2004 and 1,961.197 MW in 2005.

Photon's statistics are grounded on information from grid operators and energy supply companies which are committed to purchase the electricity generated from renewable energies. The problem is that this data was collected only for the short period from 2000 to 2005. For this reason statistics from AGEE-Stat are used. The AGEE-Stat's data also conforms to statistics from IEA-PVPS[27], EPIA[28]. Figure 5-2 gives an overview on the variables included in the model.

In 2006, Germany represents 84.88% of the total capacity installed in the European Union (EurObserv'ER 2006, p.13). The actual data for installed capacity is also available for the year 2007, but the data for other time series are collected only until 2005. For this reason, the installed capacity is also regarded until 2005.

For further modelling these time series are tranformed in logarithmic scales. Table 5-1 summarizes the data description, notation, and their sources used in this section. It has to be emphasized that the variables in Figure 5-2 have different aggregation levels:

- The variable "installed capacity": The installed capacity refers to annually new installed power plants in Germany.

- The variable "scientific publications": The scientific publications include worldwide publications placed in reputable international journals in the photovoltaic research field. It would make no sense to restrict the variable "scientific publications" only to publications written by German scientists. This restriction would, on the one hand, unnecessary complicate the counting of international publications. On the other hand, it would imply that scientific foreign publications without German participation have no impact on the development of the German PV industry. This constraint is also not meaningful because scientific knowledge can develop its impact across geographic boundaries.

27 International Energy Agency - Photovoltaic Power Systems Programme
28 European Photovoltaic Industry Association

- The variable “patent applications”: Only patent applications at the EPO and the DPMA are included in the model. After establishment of the EPO, many inventors prefer to apply their inventions at the EPO because in this way they can more easily protect their intellectual property in several European countries using only one application. Nonetheless, Germany stays the biggest market regarding new installed PV capacity in Europe and even worldwide.

- The variable “Oil price”: In this investigation, I assume that the oil market is more or less homogenous and includes development of crude oil prices in $ in the money of 2005 as the exogenous factor which influences evolution of the PV market in general and the German PV market in particular.

- The variable “subsidies”: The Public Promotion Catalogue of the Federal Ministry of Education and Research provides detailed data on subsidies for funding of PV technology in Germany.

- The variable “compensation”: By contrast with the variable “subsidies”, the compensations are related to all renewable energy sources promoted by the StrEG resp. the EEG. Unfortunately, there is no compensation data for only energy fed from PV plants.

These differences in data can eventually negatively affect the result of investigation.

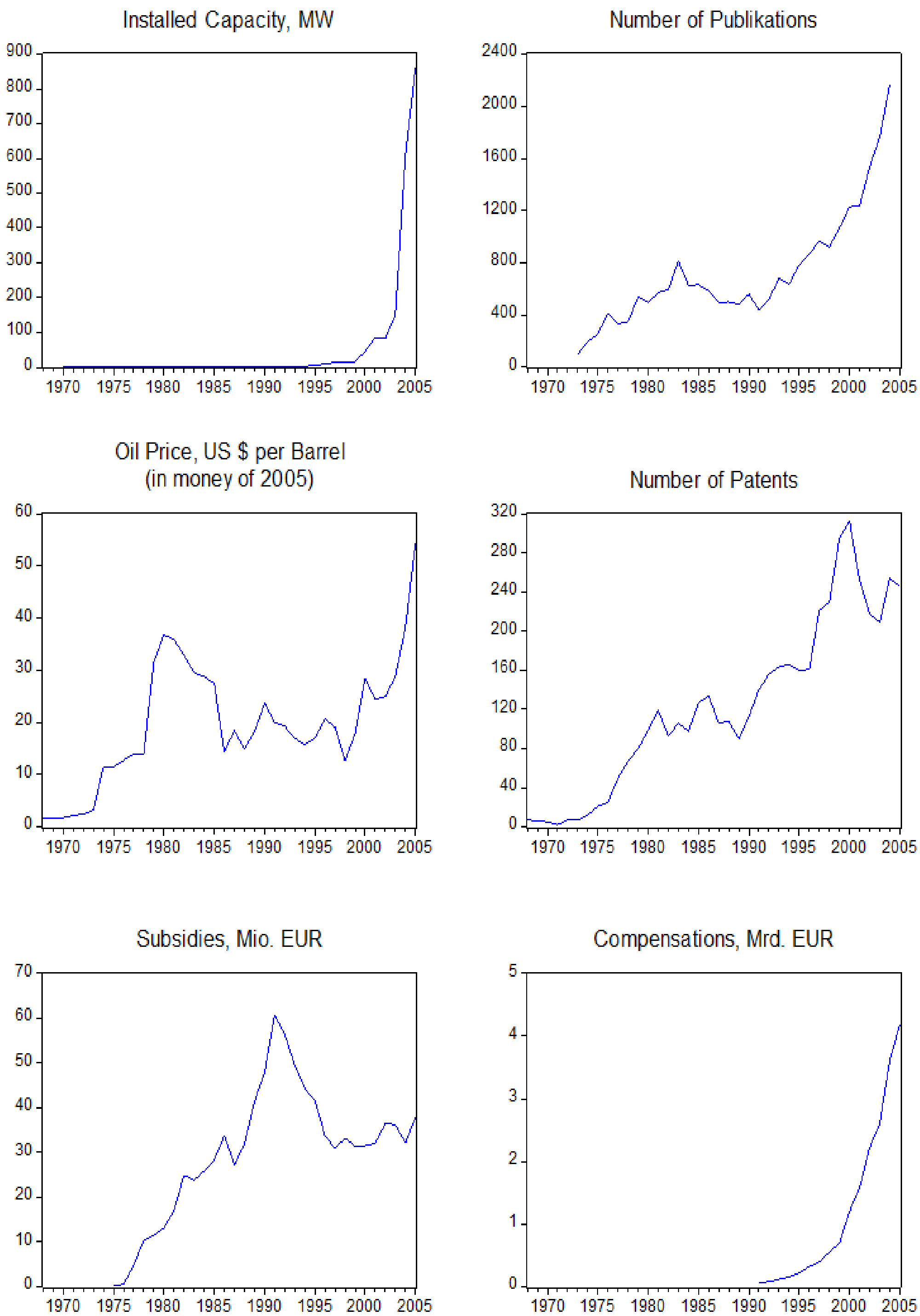


Figure 5-2: Summary of empirical data used for modeling.

Variables	Notation[29]	Period	Data Source
Number of Patent Applications	lpat	1968-2005	EPO Worldwide Statistical Patent Database version October 2007
Number of Scientific Publications	lpub	1974-2005	The Science Citation Index (SCI, host STN)
Compensation according to the StrEG resp. the EEG, Mrd. Euro	lcompens	1991-2007	The Federal Ministry for the Environment (BMU) Workgroup Renewable Energies-Statistic (AGEE-Stat)
Subsidies, Million Euro	lsubs	1975-2006	The Public Promotion Catalogue of the Federal Ministry of Education and Research
Installed PV Capacities (MWp)	lcap	1990-2007	The Federal Ministry for the Environment (BMU) Workgroup Renewable Energies-Statistic (AGEE-Stat)
Oil Price, in Prices 2005	loil	1968-2007	The BP Statistical Review of World Energy
the 100.000 Roofs Program	Dummy_100	1999-2003	IWR Solarenergy

Table 5-1: Summary description of model variables and their sources.

[29] For the further analysis all variables were logarithmed.

5.3 Results of Statistical Modeling

A short description of the main steps in statistic modeling is given here: firstly, the construction of the model starts with a univariate data analysis. The focus of this analysis lies in investigation of non-stationarity of time series using the ADF-test discussed in 5.1.1. In this investigation, the properties of single variables, like trend, and order of integration are checked.

Secondly, the Granger Causality Test is used in order to test the causal relationships between variables. The market formation process in a science-driven environment is characterized by many feedback loops, therefore a clear-cut demarcation of model variables into endogenous and exogenous units does not seem to be plausible. The mutual dependence of technology on science and at the same time of science on technology is continuously high. The profit-oriented enterprises look for opportunities to make the best use of scientific achievements. Interestingly, also the scientific research is very much dependent on the performance of modern laboratory technology. Therefore, in the next steps, the concept of endogenous and exogenous variables should be avoided. The results of Granger Causality Test help identify bidirectional causal relationships between the variables.

Thirdly, estimates of co-integrating relations are obtained using Johansen's multivariate procedure. A statistically significant co-integration vector is included in the estimation of the error correction model (ECM). The ECM is the result of this empirical study. An interpretation of achieved empirical results concludes the model construction.

5.3.1 Results of Unit Roots Tests

Table 5-2 and Table 5-3 present the ADF-test results for the levels and first differences of the variables. The results of the ADF test show that all time series with exception of subsidies are not stationary in levels. The absolute values of t-statistics are smaller than critical values on the 1% and 5% levels. After observing the first difference of the variables, the null hypothesis can be rejected with a significance level of 1%. This means that all variables are integrated of order one, I(1), in level forms. Since the variables are considered to be I(1), co-integration analysis using an error correction model (ECM) is appropriate to equilibrium model.

Variable	Specification[30]	ADF t-statistic	Critical Values		Results
			1%	5%	
lpat	Intercept, Trend	-1.36	-4.23	-3.54	non-stationary
lpub	Intercept, Trend	-2.07	-4.37	-3.6	non-stationary
lcompens	Intercept, Trend	-1.91	-4.67	-3.73	non-stationary
lsubs	Intercept, Trend	-6.69	-4.31	-3.57	stationary
lcap	Intercept, Trend	-1.1	-4.73	-3.76	non-stationary
loil	Intercept, Trend	-1.89	-4.23	-3.54	non-stationary

Table 5-2: Results of the ADF-tests for the level variables.

First Differences	Specification[31]	ADF t-statistic	Critical Value		Results
			1%	5%	
D(lpat)	Intercept	-6.64	-3.63	-2.95	Stationary
D(lpub)	Intercept	-6.77	-3.66	-2.96	Stationary
D(lcompens)	Intercept	-5.18	-3.96	-3.08	Stationary
D(lcap)	Intercept	-6.78	-3.96	-3.08	Stationary
D(loil)	Intercept	-5.49	-3.63	-2.95	Stationary

Table 5-3: Results of the ADF-tests for the first differences of variables.

[30] trend and constant: t, c; only constant: c
[31] trend and constant: t, c; only constant: c

5.3.2 Results of Granger-Causality Tests

The next step in the model construction is the identification of bidirectional causal relationships between variables using the Granger Test for Causality. In order to test whether the linkage between variables is stable or not, different lag lengths are tested. An investigation of bi-directional causality is used to prove the following hypotheses:

The first hypothesis proves there are, in fact, interdependencies between publication and patent statistics. Only a few decades ago it was a fact that patents were a matter of industrial firms and private inventors. Academic researchers preferred to publish their achievements in scientific papers. Today, there is not any traditional boundary between the industrial and academic research.

These close relationships between academic and industrial research have many positive aspects, but at the same time there are some doubts about consequences of these changes. The quality of fundamental research can suffer from this trend as research substance can become increasingly applied and field of research without marketing orientation can be disregarded (see Czarnitzki et al. 2007). It is difficult to do anything against this tendency. On the one hand, it becomes apparent that there is a clear trend toward commercialization of academic science, on the other hand, the industrial research is increasingly dependent on new acquisitions of the science. In either case, these considerations hold for the science-driven markets in general, and also for the PV market. Therefore, it can be accepted that there is a strong causal relationship between patent applications and scientific publications. This consideration can be taken as the basis to formulate the following hypothesis:

<u>*Hypothesis 1:*</u> The number of patent applications Granger-causes the number of scientific publications in the same technological field and the number of scientific publication Granger causes the number of patent applications.

The next hypotheses refer more to the special case of the German PV Market. R&D expenditures for PV may be used as a yardstick for willingness to establish the PV market. Under the terms of the priority that Germany gives to R&D related to PV, it is in third position after the USA and Japan. Funding of R&D projects has a positive effect on industry. Following the standard of the market failure argument (cp. Arrow 1962) firms do not invest enough in R&D because the benefits of innovative activities cannot be fully reaped due to incomplete appropriateness and knowledge spillovers between firms. The government spends public money in order to stimulate innovative activity. This can be successful because subsidies reduce marginal costs and increase the profitability of R&D projects. In this sense, the public subsidies can be seen as input indicators of innovation processes. Patent statistics are commonly used as a performance indicator. Consequently, it can be assumed that there is a link between subsidies and patent applications. Surely, there is a certain time lag between allocation of subsidies and patent applications. This time span has to be taken into consideration during analysis.

Similar reasons can be applied to the impact of subsidies on a number of scientific publications. However, research units in general can be at least partially considered as publishers of public knowledge. For this reason, the government grants subsidies to these organizations. These considerations can be summarized by the following hypothesis:

Hypothesis 2: The subsidies of the government Granger-cause the number of patent applications and scientific publications.

The next hypothesis deals with the Electricity Feed Act and the Renewable Energy Sources Act that were passed by government legislation in order to facilitate adoption of renewable energy and stimulate private R&D activity. These laws concern not only photovoltaic energies but also other renewable energy sources like hydrodynamic power, wind energy, geothermal energy, and biomass, and others.[32] The following causal relationship can be expected:

Hypothesis 3: Compensation from the Electricity Feed Act and the Renewable Energy Sources Act Granger-causes the number of patent applications.

Statuary feed-in tariffs made PV investments profitable for the first time. This facilitated the consumer decision for use of this technology. In this sense, one could explain the unidirectional causality compensations versus installed capacity. From a statistical point of view, the inverse direction is not interesting. According to the Renewable Energy Sources Act, the electricity that is generated by grid-connected solar power installations is bought by network operators at above market price. The

[32] More details can be found in Renewable Energy Sources Act published by the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety: http://www.bmu.de/files/pdfs/allgemein/application/pdf/res-act.pdf (accessed on 27.11.2008).

difference between the retail price of electricity and the price of electricity generated by renewable resources is spread over all customers of the utility company. Consequently, an increase of installed capacity leads to the increase of compensations. According to this consideration, the following hypothesis can be formulated as follows:

Hypothesis 4: Compensation from the Electricity Feed Act and the Renewable Energy Sources Act Granger-causes the installed PV capacities.

The last hypothesis also deals with the diffusion of photovoltaic modules. The rapid growth of the photovoltaic market measured by installed capacities and further growing demand lead to falling manufacturing costs by producers due to technological innovations (see chapter 4) and economies of scale. This makes solar energy competitive with conventional sources of electrical power. The firms can again expand production, improve technological quality of products, and invest more in new developments. Furthermore, sold and installed solar plants connote the refinancing of the investment costs for PV industry. This development is reflected in the increase of patents and publications. Just like in the third hypothesis, it can take some years until the increased installed capacities will have an impact on patent activities. The next hypothesis is:

Hypothesis 5: The installed PV capacities Granger-cause patent applications.

Pairs of Time Series	Lag's number			
	1	2	3	4
Patents vs. Publications	–	→**	→*	→*
Subsidies vs. Publications / Patents	–	–	–	–
Compensations vs. Patents	–	→***	→*	→**
Compensations vs. Installed Capacities	–	←*	–	–
Installed Capacity vs. Patents	–	–	–	–

Table 5-4: Results of the Granger Causality Test[33].

[33] significance levels *:p<0.1, **:p<0.05, ***:p<0.01.

Table 5-4 shows the results of Granger Causality test[34]. Not all hypotheses can be confirmed at a significance level of $p<0.05$. Regarding the first hypothesis, there is a unidirectional relationship between the number of patent applications and the number of scientific publications. The values of F statistic suggest that the number of patent applications Granger-causes the number of scientific publications, but number of publications does not cause the number of patent applications. Thus, it can be argued that past values of patent application numbers contribute to the prediction of the present value of publications even with past values of publication numbers. In general, one can assume that scientific publications as well as patent applications publish new technological achievements. Because novelty is one of the conditions for obtaining a patent, the novelty of scientific publication is more or less constituted by the fact that this paper was published in a journal included in SCI. In case of patent application, the inventor endeavors to stake out a claim for his invention at the patent office as soon as possible, assuming he chooses this kind protection of intellectual property. The complete examination of the patent application proceeds later. It looks quite different when considering a publication process. The complete peer review process, the manuscript revision process by the authors, and the technical editing is carried out before publication and causes a delay in publishing of papers. For the time lag of 2 years between priority date of patent applications and publication date of papers Table 5-4 shows the strongest causality with significance level of 5%. The causality becomes weaker for the lags of more than 2 years.

Regarding the second hypothesis one can see that there is no empirical confirmation for the causal relationship between subsidies and a number of patent applications respectively the number of scientific publications. This result can be explained as follows: firstly, there is no definite opinion about whether subsidies in general stimulate private R&D spending (see David et al, 2000). Secondly, the subsidies in this model include only expenditures of the German government. On the other side, the European patent applications and worldwide publications are regarded. This inconsequence of basis data can lead to such a result. Nonetheless, it is very difficult to collect complete data about total subsidies worldwide for solar Photovoltaic technology. It would be pointless to regard scientific activity only from German researchers, because published technological and scientific knowledge is not restricted by geographical borders.

[34] The detailed estimation output can be found in Appendix 9.4

Another result can be derived regarding causality between feed-in compensations and the number of patent applications. Because the feed-in compensation for electricity from renewable energy sources affect not only German producers, but also foreign companies that sell their products in Germany. The solar modules have to be installed in Germany, the origin of products is not important. Here, one can see that predicted values of compensations contribute to the prediction of the present value of European patent applications but with a time lag of 2-4 years. It is not totally unexpected, because Germany is the world's largest user of solar power ahead of Japan and USA. For this reason, it is understandable that the development of the German solar PV market has an impact on R&D activities at least on the European level.

An empirical investigation of fourth hypothesis confirms only a weak relationship between installed capacity and compensations according to the Electricity Feed Act and later the Renewable Energy Law (EEG). However, this relationship is not so interesting from a statistical point of view because it is based only on monetary flow.

Finally, there is no empirical confirmation for causality between installed solar PV capacities and the number of patent applications. The fifth hypothesis cannot be accepted.

5.3.3 Empirical Results of the Johansen Co-integration Analysis

In order to analyze long-term co-integration between model variables more deeply, co-integration analysis is applied. It is a third step of model estimation. This step is carried out for two variable pairs: the number of patent applications vs. compensations and the number of patents vs. the number of scientific publications. At first, the existence of a long-term relationship has to be tested. The results of the ADF-test show that all these time series have the same order of integration. It is the necessary, but not sufficient, condition for existence of co-integration relationship between two time series.

Using co-integration test, the existence of co-integration relationship can be proven. According to Table 5-5, there is a co-integration vector between the number of patents and the number of scientific publications. The trace and maximum Eigen value test indicate existence of one co-integration vector. Again, there is no evidence of existence of a co-integration vector between compensations and the number of patent applications. In this case, the null hypothesis of the absence of co-integration relations cannot be rejected at 5%.

Model	H_0	Trace Test		Max Eigen Value Test	
		Test Statistic	0.5 Critical Value	Test statistic	0.5 Critical Value
Patents–Publications	None At most 1	17.51 0.14	15.49 3.84	17.37 0.14	14.26 3.84
Compensation-Patents	None At most 1	15.13 3.22	15.49 3.84	11.90 3.22	14.26 3.84

Table 5-5: Results of the Johansen co-integration test.

Finally, the error correction model will be a test for variable pair patent applications and scientific publications.

According to the basic hypothesis (see introduction), there are two different phases in the development of a science-driven market: the “science-push” and “demand pull” phase. Regarding patent applications, the ANOVA F-statistic (78.8) suggests significant difference in means (p-value < 1%) for two periods of time: between 1968 and 1990 and between 1991 and 2005. A similar result can be achieved for scientific publications. Here, the ANOVA F-statistic (24.19) indicates significant differences in means of publication numbers between 1973 to 1990, and 1991 to 2005[35]. Correspondingly, the time between 1968 until 2005 is split up in two time periods: 1968 to 1990 and 1991 to 2005. The estimation of a error correction model is carried out for each time period separately. The fit of the model was improved in due consideration of exogenous variables. Different exogenous variables were included in the account. The best results were achieved, including the impact of the fuel oil prices for the first time period, and the impact of the StrEG and the EEG for the second time period. The models were compared using adjusted R^2 and the AIC and the SIC. Basically, almost all coefficients are significant. The comparison of the ECM's for two time periods follows in Table 5-6.

Error Correction:	$\Delta lpub^{1973-1990}$	$\Delta lpub^{1991-2005}$
Co-integration Vector (CointVec)		
lpat(-1)	1	1
lpub(-1)	1.25	3.63
	[3.27]	[14.41]
Intercept	-11.96	-30.09
$CointVec^1(-1)$[36]	-0.26	
	[-4.43]	
$CointVec^2(-1)$		-0.52
		[-7.72]
D(lpat(-1))	-0.16	0.11
	[-0.84]	[0.89]
D(lpub(-1))	-0.51	0.33
	[-2.76]	[2.18]
Intercept	-0.87	0.44
	[-2.31]	[11.06]
loil	0.33	

35 Same splitting is carried out by regarding co-authorship networks.

36 ECM(-1): the disequilibrium error in the previous period is equivalent to deviations from long-run equilibrium at t-1.

	[2.65]	
lcompens		0.69 [7.7]
R-squared	0.64	0.93
Adj. R-squared	0.53	0.89
Sum sq. resids	0.29	0.025
S.E. equation	0.16	0.05
F-statistic	5.73	29.11
Log likelihood	10.39	26.73

Table 5-6: Estimated error correction models. T-statistics in [].

According to the results in Table 5-6, the first log differences of publication numbers in the time span from 1973 to 1990 and from 1991 to 2005 can be written in the following way:

$$\Delta lpub_t = \alpha_0 + \alpha_1 \underbrace{(lpat_{t-1} - \beta_1 lpub_{t-1})}_{=\text{CointVec}} + \alpha_2 \Delta lpat_{t-1} + \alpha_3 \Delta lpub_{t-1} + \alpha_4 s_t + \varepsilon_{t,}$$

whereby s_t is an exogenous variable (*loil* for the first time span, and *lcompens* for the second time span). The estimated coefficients are α_1=-0.26 resp. -0.52, β_1=1.25 resp. 3.63, α_2=-0.16 resp. 0.11, α_3=-0.51 resp. 0.33, and α_4=0.33 resp. 0.69. The speed of adjustment is the time it takes to reach a new equilibrium after the initial shock is determined by the coefficient α_1. The data indicates magnitude variations across years, from a moderate speed of adjustment in 1973-1990 (-0.26), where 26% of the derivation from equilibrium is eliminated in the next time period to a relatively fast adjustment in 1991-2001 (–0.52) with 52% of deviation. The coefficient $\beta_{1,}$ interpreted as a long-term equilibrium coefficient, increases almost threefold. This means a 1% increase in the average number of patent applications leads to 1.25% resp. 3.63% increase in the average number of scientific publications. The short-term coefficient α_2 is insignificant for both time periods. The impact of the exogenous factors is positive and significant, e.g. increased oil price simulates more publications and increased compensations also have a positive impact on the number of scientific publications.

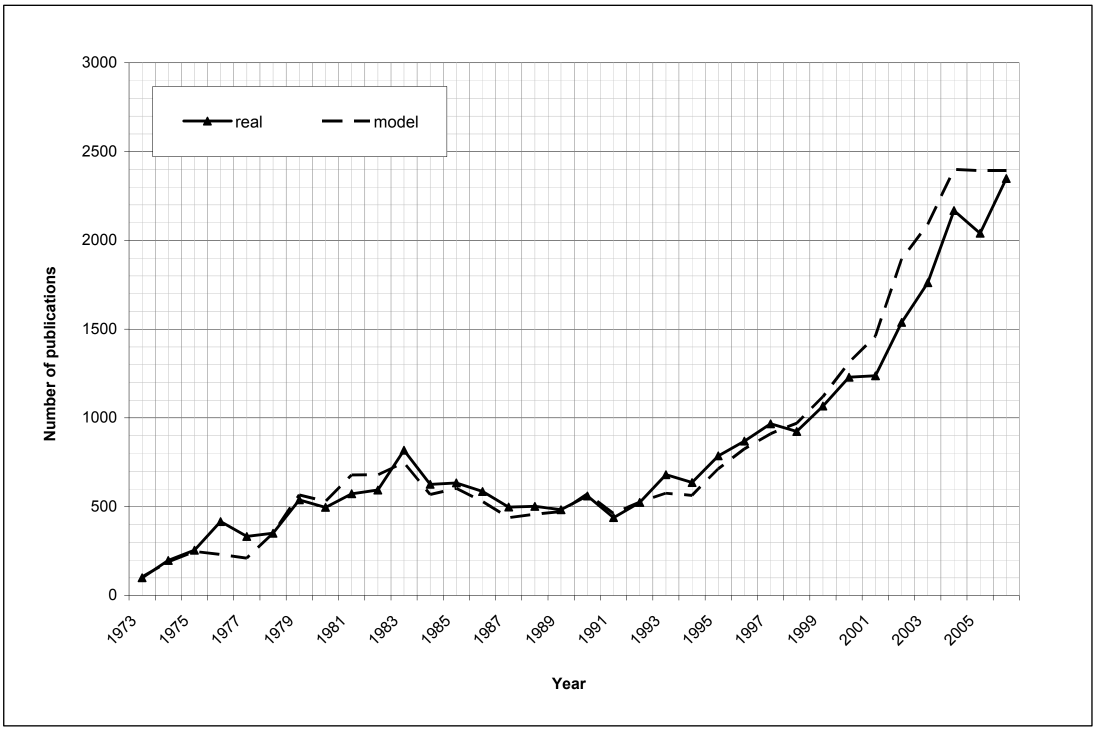


Figure 5-3: Model results and empirical data.

Using different goodness of fit tests, the validity of the models are checked. The adjusted R-squared statistic is 0.53 and 0.89 respectively, which is a relatively good explanation, especially for the second time period. The F-test for the significance of the goodness of fit is 5.73 and 29.11, respectively. The critical values are 5.67 and 15.98, respectively. Since the F-test statistics are greater, the goodness of fit of regression is significant for both time periods. In order to identify the presence of autocorrelation in the residuals, residual portmanteau tests for autocorrelations are carried out. The null hypothesis of no residual autocorrelations up to lag 7 cannot be rejected for both time periods. Summarizing, the diagnostic tests support the validity of the estimated model.

5.4 Conclusions

The subject of this paper is an empirical analysis and dynamic mathematical modeling of innovation processes in science-driven markets. Numerous empirical studies provide evidence that science-driven markets have different development patterns then consumer markets, in which the science base of the underlying technological development is rather unimportant. The purpose of this work is the construction of an econometric model which has the power to explain the dynamics of science-based innovation processes using a few relevant variables. The model is tested and validated with empirical data in terms of regression and time series analyses.

The stylized model of the formation of science-driven markets presented in the introduction is taken as theoretical background for development of the econometric model. In so doing, the main hypothesis is constructed, namely: in the development of science-based markets two quite different development phases can be observed due to basically different sets of determinants ("double-boom hypothesis"). This corresponds to the mathematical modeling of more than one steady state in the overall development of a new innovative market, rather than the usual diffusion modeling ("S-type curves").

The construction of the econometric model includes several steps: First, a univariate examination of statistical properties of the selected time series is done. The next step of model construction is the identification of bidirectional causality relationships between variables using the Granger Test for Causality. Based on these findings, the co-integration vectors are estimated. According to "double-boom hypothesis" the whole data set is split into two time periods: from 1973 to 1990 and from 1991 to 2005. For both time intervals the existence of a long-term equilibrium between publications and patents is verified. The speed of adjustment to the equilibrium for both time intervals varies strongly, from a relatively moderate speed of adjustment of -0.26 (-26%) in the first time period to very fast speed of adjustment of -0.52 (-52%) in the second time interval. Using an error correction model, the impact of different exogenous factors is tested. In the first phase of the PV market development, the fuel oil prices play an important part. The influence of the StrEG and the EEG is important in the second phase of market formation.

The current condition in the research of science-driven markets provides a lot of studies on either the science push or the demand pull side. Most of these papers are of qualitative nature. There are a few studies trying to reconcile both views into a formal mathematical model. This is the challenge of the proposed work: to come up with first solutions to this problem.

6 Investigation of Collaboration Relationships in PV Technology based on Social Network Analysis.

Scientific knowledge is an essential factor in the development of science-driven markets. Scientists who are active in corresponding fields can be seen as creators and carriers of technological knowledge. They publish their research results in scientific papers, some of which have more than one author. By using this bibliographic data, including co-authorship information, collaboration networks in the scientific community can be mapped. Analysis of these networks helps us to understand collaboration activity within an academic community.

In such networks, two agents (here authors or organizations) are connected by having written a joint publication. Correspondingly, there are two networks: co-author and inter-organizational. In both, the agents are the nodes and the joint publications are the links. The unit of analysis, be it author or organization, depends on the corresponding research question and will be discussed later. In order to understand the process of knowledge generation and diffusion, chapter 6 analyzes co-author and inter-organizational networks.

This chapter explains the occurrence of the two stages in technological development by investigating two research questions:

1. How does the network topology of the scientific collaboration community change in two periods of technological cycle: “science push” and “demand pull”?
2. What changes in the authors’ behavior can also explain the emergence of two different phases in the technology cycle?

At the beginning of the chapter, the importance of collaboration networks in the evolution of science-based technologies is discussed. Building networks in a scientific community is important for at least two reasons: generation and diffusion of knowledge.

Generation of knowledge:
In most cases, the generation of technological knowledge has to be regarded as a collective process. The interdisciplinary nature and increasing complexity of science-based technologies implies that, in these technological fields, knowledge generation can no longer be seen as a product of the individual. New knowledge is created as a

result of many social interactions rather than through the efforts of isolated players. Technological knowledge is often seen as

> "…a coherent stock of fragmental pieces of information, partially owned by a variety of economic agents" (Patrucco 2002, p. 405).

The complementary nature of knowledge is another reason networks in a scientific community are important. The complementary pieces of knowledge need to be accumulated and adopted by agents before being used in a specific context.

The capability to accumulate and recombine existing knowledge requires a certain level of competence. In this context, Cohen & Levinthal (1989) emphasize the dual role of R&D. R&D's primary goal is to develop new products and advance technological knowledge. The second goal is to, in general, increase learning and cooperative capabilities within a company. Furthermore, spreading costs and risks among market players increases the probability of innovation success (Sakakibara 2003). A firm can be regarded as a portfolio of core competencies (Prahalad and Hamel 1990). Inter-firm cooperation brings heterogeneous capabilities of partners together so that each partner can benefit from productive, non-duplicative collaboration (Hamel 1990).

The collective nature of knowledge generation is confirmed in scientific literature. Spence (1984) is one of the first authors to examine the advantage of cooperative R&D using a theoretical model. There are also several empirical studies that reveal the influence of social interactions on knowledge creation. Howells (1999) highlights the fact that even large multinational enterprises can no longer expect to be totally dependent on their own research and technical resources in order to retain their innovative performance. DeBresson et al. (1998) investigated the innovation activities of manufacturing industries in ten OECD countries and showed that more than 90% of the innovatively active European companies generally collaborated with at least one external partner. Drejer and J∅rgensen (2005) analyzed collaboration activities between private firms and public research institutions in Denmark. Their study also supports the finding that innovative activities are rarely carried out in isolation. However, they mention a low frequency of public-private research collaborations because of confrontations between the two different types of organizations. This confrontation needs to be overcome. In summary, collaboration activities can be regarded as essential for market players to keep pace with the development of new technologies. This statement can be tested with the following hypotheses:

Hypothesis 1: Collaboration networks derived from bibliographic data expand over time with the addition of new agents and links between new agents and agents already present in the network.

Scientific productivity can benefit from network growth and intensive cooperation. Thus, the second hypothesis is:

Hypothesis 2: As a result of network evolution, an intensification of cooperation and increased scientific productivity of authors can be observed.

Diffusion of knowledge:
Social interactions are the key factor facilitating the diffusion of technological knowledge. Above all, conferences and publications can be seen as potential ways for agents to share knowledge that they have generated. In this context, the structure of co-author networks plays a central role and has an effect on the speed and extent of knowledge diffusion. Details about who is connected to whom affect what type of technological knowledge is passed on, how much, and how efficiently. All these factors influence general development of investigated technology (see also Cowan and Jonard 2004). Network structure is also important because some not codified knowledge can be exchanged only through face-to-face interactions. For this reason, many empirical studies show that the rate of knowledge diffusion is highest in small world networks (e.g. Bala and Goyal 1998, Morone and Taylor 2004). A small world network is a network in which most nodes are not connected with each other, but the connected nodes can be reached from every other node by a small number of steps. According to the small world property, co-author networks will tend to form more highly connected clusters in order to guarantee faster and wider knowledge diffusion. This assumption leads to the formulation of the third hypothesis:

Hypothesis 3: The small world property of the co-author network in the demand-pull phase becomes more intense in comparison with the co-author network in the science-push phase regarding the same technological field.

Furthermore, an important aspect of network evolution is the dynamic development of cooperation intensity. The cooperation intensity changes as new technologies evolve. The first reason for this is the change in cooperation motives. Hagedoorn (1993) provides a very good overview of cooperation motives and proposes a simplified linear interpretation of the innovation process beginning with scientific discovery, through to market entry, and ending with the joint introduction of new products. Using the MERIT[1]-Cooperative Agreements and Technology Indicator's database, which contains detailed information on nearly 10,000 technology cooperation agreements involving some 3,500 different parent firms, Hagedoorn (1993) identifies the major motives behind firm's engagement in cooperative behavior. Hagedoorn (1993) extensively investigated 4,192 strategic technology alliances made in different sectors and fields of technology[2]. Three motives were most frequently mentioned: complementary technology (31%), reduction of the innovation time-span (28%), as well as market access and influencing the market structure (32%). All these categories become more important when technology achieves a certain level of complexity in it's evolutionary progress. As a consequence of this, cooperation intensity grows with the advancement of technological development.

Another observation regarding changes in cooperation intensity over time can be found in Dosi (1982) and is strongly associated with the term "technological trajectory". At the beginning of any innovation phase, there is a large set of possible solutions to a technological problem. A selection of more promising approaches is the result of general questions: "Is any practical application conceivable?", "Is there some possibility of the hypothesized application being marketable?", and so on. A tendency to try to solve the actual problem in isolation is relatively high because there are not enough collaboration partners, and the necessity for cooperation is relatively low. In the course of time, "the determinateness of the selection"[3] increases and "natural trajectories of technical progress"[4] establish. There is more standardization in solving technological problems, and there are subsets of actors moving together in one technological direction. In parallel with this progress, the cooperation possibility expands. Finally, the increased complexity and the inter-sectoral nature of new technologies lead to an augmented need for cooperation. In this case, it is the growing convergence between, for instance, subfields of chemistry, physics, and electronics; computer science and process technologies; or materials

1 Maastricht Economic Research Institute on Innovation and Technology

2 Biotechnology, New Materials Technology, Computers, Industrial Automation, Microelectronics, Aviation/Defense, Heavy Electric/Power, Instruments/Medical Technology, Software, Telecommunications, Other IT, Automotive, Chemicals, Consumer Electronics, Food and Beverages, and Others. The first 10 categories belong to high-tech industries according to R&D intensity based on OECD Observer, Juni-July 1990.

3 See Dosi (1982) p. 153.

4 See Dosi (1982) p. 154.

science, electronics, and chemistry - that forces agents to cooperate with each other (Hagedoorn 1993).

This dynamic network perspective implies possible changes in the entire network structure and changes in individual agent characteristics. These properties need to be proven in this chapter as well.

The remainder of chapter 6 is structured as follows: section 6.1 presents a short introduction into social network analysis with some measurement concepts. Section 6.2 presents data and some descriptive statistics. Section 6.3 analyzes the topological changes of collaboration networks. Section 6.4 provides an investigation of changes in collaboration behavior. Finally, section 6.5 concludes the chapter by summarizing the results.

6.1 Basic Terminology and Measurement Concepts of SNA

Social network analysis (SNA) is a branch of sociology which focuses on the mapping and measuring of relationships and information or resource flows among interacting units like individuals, organizations, groups, states, etc. (Berkowitz 1982; Wellman 1988; Wasserman & Faust 1994). Tindall and Wellman (2001) define the SNA in the following way:

> "Social network analysis is the study of social structure and its effects. It conceives social structure as a social network, that is, a set of actors (nodes) and a set of relationships connecting pairs of these actors" (p. 1-2).

According to the definition by Wasserman and Faust (1994, p. 89), social network data can be regarded as a social relational system characterized by a set of actors and their social ties. Thereby, there are two key components in the social network approach: actors and relationship (ties) between these actors. The main goal of the social network approach is "to understand how social structures facilitate and constrain opportunities, behaviors, and cognitions" (Tindall and Wellman 2001, p. 266). In this context, the overall behavior is more than the sum of individual behaviors; consequently, the network analysis approach assumes that some social phenomena cannot be explained only by individual characteristics, but also by the social structure influencing the actions of individuals. According to this, there are two strands in the SNA literature. Some analysts (cp. Wellman 1988) focus only on the social structure in which the actors are embedded; the others (cp. Doreian 2001) consider both aspects, structure and individual characteristics of actors, in order to explain their behavior. In the following investigation, structure and individual characteristics will be analyzed.

SNA has rapidly developed in the past 20 years. The growing interest in this method can be explained by two conventions[5] that the network approach employs:

1. It is guided by formal theory organized in mathematical terms
2. It is grounded in the systematic analysis of empirical data

This fruitful combination of theoretical concepts with the numerous application possibilities has attracted a lot of research described, for example, by Wasserman and Faust (1994, chap. 1)

---

5 http://www.insna.org/sna/what.html (accessed on 2.10.2008)

Advances in computer technology, availability of computer databases that cover published scientific papers, and the emergence of a range of software packages like UCINET[6], Pajek, NetDraw[7], has enabled the construction and analysis of scientific collaboration networks. The publication records include detailed information about authors and their affiliation. Newman (2001b) assumes that these networks are for the most part true acquaintance networks, since it is likely that a pair of scientists who have co-authored a paper know each other. For this reason, these networks provide a promising source of real-world data to fuel the current surge of research interest in social network structure.

SNA provides a number of methods for revealing patterns of interpersonal relationships and has frequently been applied to study collaboration through co-authorship networks (Newman 2001a, 2001b, 2001c; Otte and Rousseau 2002; Kretschmer and Aguillo 2004).

It is most meaningful to start the introduction of SNA-methods by establishing some formal and theoretical concepts (see also Roediger-Schluga and Barber 2006). The basic mathematical structure for visualizing networks is a graph. A graph consists of a set of nodes[8] and a set of links[9] that connect pairs of nodes.
The total number of links connected to a node n_i is the degree denoted by $d(n_i)$. Nodes with degree one or higher are connected and have at least one neighbor, while nodes with degree null are isolated.

The graph density is a simple measure of the network structure, and is calculated by dividing the number of linkages present, L, by the theoretical maximum in the indirect network[10], $n(n-1)/2$, where n is the number of nodes in the network (Wasserman & Faust, 1994, p. 102). Density scores range from 0 (empty network) to 1 (connected network[11]).

If a graph is connected, it is possible to establish a path from any node to any other node. A path is the sequence of nodes and links that connect two non-adjacent nodes, without repeating any of them. Paths are useful to measure distance, i. e. how far apart nodes are in a graph. The geodesic distance is the shortest path between two

6 http://www.analytictech.com/ucinet.htm, (accessed on 11.03.2009)
7 http://www.analytictech.com/Netdraw/netdraw.htm (accessed on 11.03.2009)
8 Nodes here are authors or organizations.
9 Links here are joint publications.
10 A network with undirected links
11 A network in which every node is connected to every other node.

nodes of the network. Average path length is defined as the average distance between all reachable pairs in the network.

A disconnected graph can contain one or more isolated (unreachable) nodes and can be partitioned into subgraphs (subsets of nodes and links). Empty sub-graphs including three nodes are called triads. If one node in a triad is connected to two other, this sub-graph is defined as connected triple. If three nodes are completely connected, they constitute a triangle.

Groups of nodes connected together are defined as components. A giant component (GC) is the largest component in the graph. Figure 6-1 shows an example of disconnected graph with one isolated node and one component. The giant component includes four of five nodes.

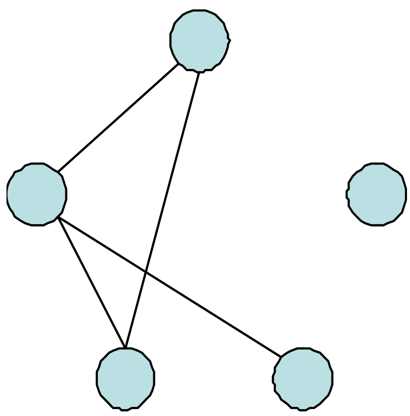


Figure 6-1: A simple disconnected indirect graph.

SNA provides several types of measures for assessment of properties for a particular node or the whole network. These measurement concepts are discussed in the next section.

6.1.1 Actor Level Measures

At the individual level, network measures describe neighborhoods and the position of an agent within a network. One example of an individual measure is the node degree discussed in 6.1. Freeman (1979) suggests the centrality concept in order to determine the most important agents in a group based on their network position. As it is a specific measurement of individual nodes, the centrality indices make it possible to check whether there are any differences between the various nodes in relation to their network position. The value of these indices varies from a minimum of 0 (unimportant nodes) to a maximum of 1 (extremely central nodes).

There are three commonly applied types of centrality: degree centrality, betweenness centrality, and closeness centrality.

The first measure is degree centrality, which simply counts the number of direct links for each agent. Since this measure depends on the size of the network, it is meaningful to consider this indicator in relation to the maximum degree that can be achieved in a network of the same size. In a network with g nodes the maximum degree is equal to $g-1$. The normalized degree centrality of node n_i can be expressed with the following equation:

$$C_D(n_i) = \frac{d(n_i)}{g-1} \tag{6-1}$$

Those agents with the most links have the highest centrality scores. Such players can use their position to influence others agents or to get more information. This indicator can be regarded as a measurement of communication activity of an agent.

The second measure is betweenness centrality, which counts the frequencies that one node is on the shortest path between two others. Similar to the first indicator, it can be normalized by dividing betweenness centrality by the theoretical maximum: the number of all shortest paths (geodesics) between all node pairs in a network of the same size. This indicator is regarded as a measurement of the potential for control of communication between certain players in the network (Freeman 1979) and is defined as:

$$C_B(n_i) = \sum_i \frac{g_{jk}(n_i)}{g_{jk}} \tag{6-2}$$

where g_{jk} represents the number of geodesics between nodes j and k; and $g_{jk}(n_i)$ represents the number of geodesics that involve the node n_i.

The third indicator is closeness centrality which focuses on the distances between an agent and all other nodes in the network. The concept of closeness centrality is available only for strongly connected graphs[12]. Actor closeness centrality is the

[12] Graphs without isolated nodes

inverse of the sum of geodesic distances from actor *i* to the *g-1* other actors, and can be normalized by *g-1*:

$$C_C(n_i) = \frac{g-1}{\sum_{j=1}^{g} d(n_i, n_j)} \qquad (6\text{-}3)$$

Since the amount of isolated nodes in this example is relatively high, this approach is not applied here.

6.1.2 Network Level Measures

There are some measuring concepts discussed above that can be calculated for the whole network: density, number of isolated nodes (authors and organizations), average path length, and the size of the giant component.

Two further indicators can be used for measuring global network properties on the network level. They are the Gini coefficient and the clustering coefficient. The Gini coefficient is the standard inequality measure, and it can also be used for testing inequality in the degree distribution[13] among the nodes. The Gini coefficient represents the deviation from a completely equal distribution of links (scientific papers) between nodes (authors or organizations). Gini values close to 0 indicate that every scientist has a similar number of co-authors; Gini coefficients closer to 1 imply greater inequality in scientific collaboration activity between authors.

The clustering coefficient measures the tendency of the network to be highly interconnected. The clustering coefficient[14] was introduced by Watts and Strogatz (1998) and is determined by the following equation:

$$Clustering\ Coefficient\ (C) = \frac{3 \times Number\ of\ Triangles\ on\ the\ Graph}{Number\ of\ connected\ Triples\ of\ Vertices} \qquad (6\text{-}4)$$

The factor of 3 in the numerator compensates for the fact that each complete triangle of three nodes contributes three connected triples, one centered on each of the three nodes, and ensures that C=1 for a completely connected graph (see Newman (2001a)).

13 The Gini coefficient can be also applied for the other two centrality concepts.

14 It is important to note that clustering coefficient only considers adjacent nodes.

The three centrality indicators discussed in 6.1.1 can be aggregated to the network level and are defined as network centralization scores. These indicators focus on the variability of individual actors centrality scores within the network. In general, network centralization can be considered as a rough measure of inequality between network agents. Network centralization calculates the extent to which a network is centralized or dominated by a few agents (see Freeman 1979). Alternatively, decentralized networks distribute control to many agents. Commonly, the centralization score is normalized by the maximum variation which is possible for the network of the same size:

$$C_N = \frac{\sum_{i=1}^{g}(c_{\max} - c_i)}{\max\left\{\sum_{i=1}^{g}(c_{\max} - c_i\right\}} \qquad (6\text{-}5)$$

Where c_i is the centrality of an agent i, and c_{max} is the largest observed normalized centrality in the network N.

Like centrality indicators centralization scores vary from a minimum of 0 to a maximum of 1. An increase in centralization score corresponds to an increase in differences between the centrality indices of the individual nodes. A minimum centralization score is achieved if all nodes in the network are equal with respect to their centrality. A maximum value is achieved if only one node has a very central position and all the other nodes have the same low centrality. Figure 6-2 shows an example of a cycle network with a minimal centralization score of 0 and a star network with a maximal centralization score of 1.

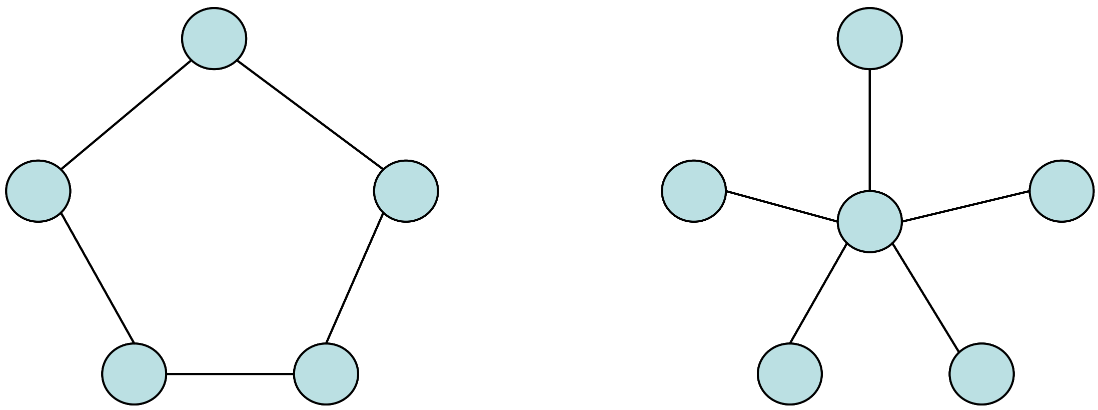


Figure 6-2: Cycle and star network.

Similar to the centrality concepts three centralization concepts exist: network degree centralization which is based on degree centrality, network betweenness centralization corresponding to betweenness centrality, and network closeness centralization corresponding to closeness centrality.

Degree centralization is defined as:

$$C_D^N = \frac{\sum_{i=1}^{g}(c_D^{\max} - c_D(n_i)}{(g-1)(g-2)} \tag{6-6}$$

Betweenness centralization can be expressed as:

$$C_B^N = \frac{2\sum_{i=1}^{g}(c_B^{\max} - c_B(n_i)}{(g-1)^2(g-2)} \tag{6-7}$$

The concept of closeness centralization is not used in this thesis.

6.2 Data and Analysis of Descriptive Statistics

The investigated collaboration network was derived from publication data, which was formed from keyword searches on the ISI Web of Science. For this kind of work, the following keyword search strategy in titles[15] was used: (solar cell or solar cells or photovoltaic#). The Web of Science limits the number of records, max 500, that can be downloaded at a time. For this reason, the records for this thesis were downloaded in batches i.e. 1-500, 501-1000, etc. Fortunately, there is no limit to the total number of records which can be downloaded. In the next step, a Java program (Appendix 9.3) was used to format the data for Pajek[16].

The period from 1974 to 2005 was investigated. The total number of authors was 16,960 and they appeared 50,967 times altogether in 12,570 scientific papers. The trends in the number of publications of scientific papers and in the number of authors for the two time periods: 1974-1990 and 1991-2005, are given in Figure 6-3 and Figure 6-4. The decision to divide into two time periods was based on analysis from chapters 2 and 5 and corresponds to the science-push and the demand-pull phases in the development of solar PV technology.

A substantial problem arose due to the different sizes of the collaboration networks in the two time periods. In order to have a more precise picture of the changes in the topological properties, it was useful to divide the investigation period of 32 years into six main periods: 1974-1980[17], 1981-1985, 1986-1990, 1991-1995, 1996-2000, and 2001-2005. The first three time windows belong to the first period of growth, the last three to the second. With the exception of the first time span, each time period comprises 5 years. This division solved the "size problem" and ensured that the collaboration networks were roughly stable during one time period.

---

[15] Since 1991, Web of Science provides the topic search along with the words/phrases in article titles, abstracts or keyword lists. For searches before 1991 only article titles are available. To ensure a stable basis of comparison which is required for a long-term study, data collection is confined to search in titles. Additionally, all the calculations from chapter 6 were carried out for the large networks derived from research strategy in titles, abstracts, and keywords lists. However, the tendency of network structure development and changes in behavior of actors stay the same as for small networks derived only from publication search in titles.

[16] Pajek (Slovenian: spider) is a popular open source Windows program for analysis and visualization of large networks. It can be downloaded from the site http://vlado.fmf.uni-lj.si/pub/networks/pajek/ (accecced on 23.12.2008).

[17] The fact that the first time span is longer than the others should not falsify the results significantly as the network in that time period was smaller in comparison with the networks describing collaboration activity of researchers in the subsequent time periods.

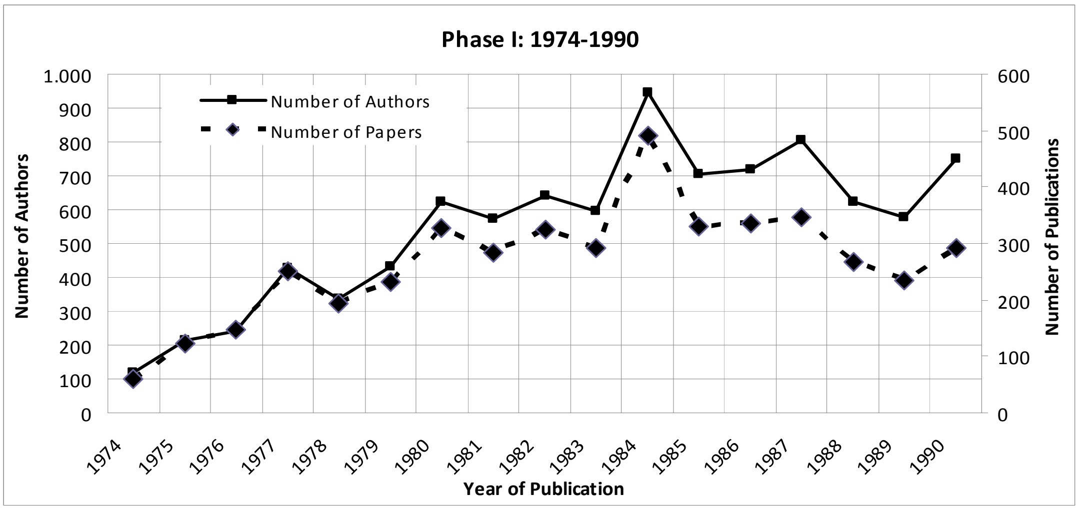


Figure 6-3: Number of publications and authors published in the PV field[18].

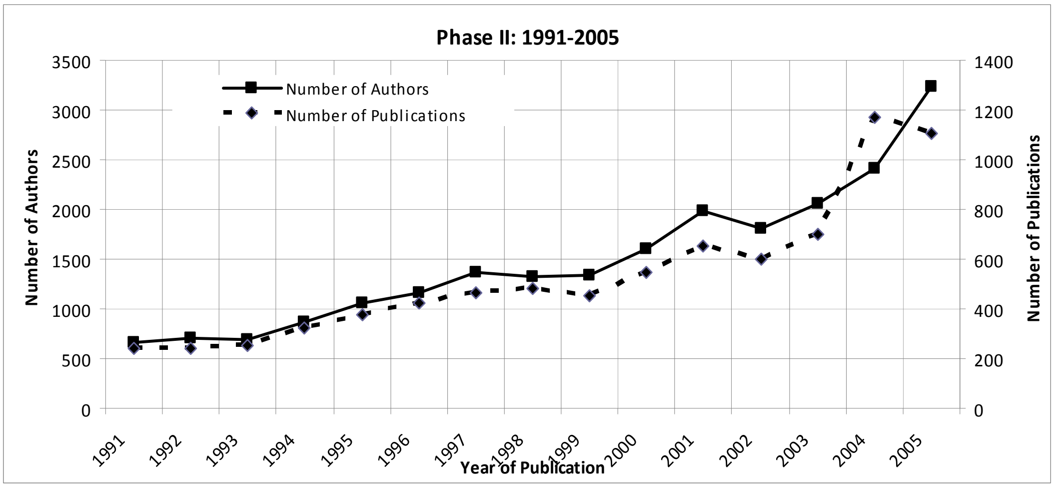


Figure 6-4: Number of publications and authors published in the PV field[19].

Table 6-1 gives some descriptive statistics for the scientific collaboration networks studied here. The first three indicators: the total number of publications, authors, and organizations were used for checking network expansion (see *Hypothesis 1*). The mean papers per author were used as an indicator for scientific productivity (see *Hypothesis 2*). The next six indicators reveal changes in cooperation intensity (see *Hypothesis 2*), and the last two indicators (clustering coefficients and average path length) help to quantify the small world property discussed in *Hypothesis 3*.

[18] Source: Web of Science, own computations.
[19] Source: Web of Science, own computations.

Indicators	1974-1980	1981-1985	1986-1990	1991-1995	1996-2000	2001-2005
Total No. of publications	1332	1724	1476	1437	2370	4232
Total No. of authors	1639	2458	2553	3006	4832	7607
Total No. of organizations	354	362	364	413	1067	1841
Mean papers per author	2.07	1.87	1.77	1.68	1.99	2.68
Mean authors per paper	2.55	2.66	3.06	3.51	4.05	4.59
No. of co-authors per author	3.33	3.79	4.13	3.36	3.97	5.36
No. of collaborators per organization	1.23	0.92	0.96	1.75	3.36	6.70
Share of isolated authors, %	7.69	8.1	5.8	4.19	2.5	1.37
Share of isolated Organizations, %	48.31	54.42	49.45	41.89	18.84	7.88
Size of the giant component[20], %	8.9	13.6	21.7	32	57.4	75.6
Average path length	4.05	6.35	8.98	8.02	6.19	5.43
Clustering Coefficients	0.63	0.66	0.71	0.77	0.79	0.81

Table 6-1: Descriptive statistics for the collaboration networks in the PV field.

In order to test the intensity of indicator changes in Table 6-1 linear regression was used. The results are presented in Table 6-2.

	No. of Papers	No. of Authors	No. of Organizations	Mean Papers per Authors	Mean Authors per Paper	Percentage of Isolated Authors
Slope in Phase I (17 years)	12.4*** (3.99)	36.3*** (27.3)	1.7 (5.7)	-0.01* (0.014)	0.05** (0.025)	-0.0028** (0.004)
Slope in Phase II (15 years)	57.3*** (29.4)	152.3*** (58.1)	51.3*** (13.00)	0.04*** (0.038)	0.11** (0.33)	-0.0035** (0.001)
	Av. Nr. of Co-Authors per Author	Av. Nr. of Collabor Ators per Organization	Percentage of Giant Comp. in the Authors NW	Percentage of Isolated Organiza-tions.	Cluster Co-efficient	Av. path length / Share of Giant Comp
Slope in Phase I (17 years)	0.09*** (0.05)	0.016** (0.027)	0.0012* (0.0024)	-0.008** (0.15)	0.011*** (0.007)	1.31* (2.39)
Slope in Phase II (15 years)	0.20*** (0.08)	0.19*** (0.07)	0.026*** (0.06)	-0.034*** (0.017)	0.006*** (0.0023)	-0.55 (1.86)

Table 6-2: Estimation of linear trends from 1974 to 1990 and from 1991 to 2005[21].

The calculation was carried out for data on a yearly basis for science-push and demand-pull phases. For each phase, the corresponding slopes are listed.

[20] In the co-author network.
[21] (Standard deviation in parentheses). Significance levels: ***: 0.01, **: 0.05.

Between 1974 and 1990, 4,531 papers were published from 5,438 authors who belonged to 834 organizations. Whereas, in the time span from 1991-2005, the number of papers, authors, and organizations increased rapidly: 8,039 publications published by 12,678 authors from 2,564 institutions. Significant growth of co-author network can be seen by comparing the slopes from the linear regression equations for the publication and author numbers for both phases. The number of active organizations in the first phase remained approximately constant; the slope obtained from linear regression was 1.7, which is positive, but not statistically significant. On the other hand, in the demand-pull phase from 1990 to 2005 the number of active organizations increased rapidly. This means that PV field seems to be more attractive for a broad spectrum of scientific market players in the second development stage. *Hypothesis 1* was confirmed with regard to the growth of publication and author numbers, but it could only be confirmed for the second phase regarding the number of organizations.

It was interesting to see that despite the rapid network evolution, the productivity of authors (i.e. the average number of papers per author) changed rather gradually over the investigated period. In the science-push phase, the productivity of researchers is considered to be stagnant if it is only at the low significance level of 0.1. However, in the demand-pull phase, a positive significant slope can be considered noteworthy although the value of the slope is relatively low. *Hypothesis 2,* with respect to author productivity, was not confirmed.

The average number of authors per paper, collaborators per author and collaborators per organization showed a continuous trend towards intensified collaboration. Additionally, the share of isolated scholars and organizations decreased. All corresponding slopes were positive and significant at least at the level of 0.05. Also, the next indicator points to merging processes within the network special to the second phase: the relative size of the giant component (GC) increased steadily. The largest sub-network in the period from 1974-1980 contained only 146 authors which constituted about 9% of all the scientists. In the time span from 2001-2005, the GC already contained 5,751 researchers which corresponds to almost 76% of the population. These results show clear confirmation of the second part of *Hypothesis 2* relating to intensified collaboration.

The examination of *Hypothesis 3* was carried out by considering the mean path length and clustering coefficients. The mean path length showed an interesting development. It increased in the first phase, but than declined in the second.

However, direct comparison of the network in the first phase with the network in the second phase is difficult because the mean path length is measured between reachable pairs. The network in the second phase has more reachable pairs. For this reason, it is more sensible to consider the average shortest distance in relation to, for example, the relative size of the giant component. In this way, the normalized mean path length can be calculated. Analyzing the results of Table 6-2, a slightly increased normalized mean path between reachable pairs was found for the time span 1974-1990. However, different results were found in the time from 1991-2005, as the negative slope was not significant. Nevertheless, although the size of the giant component increased from 32% (1991-1995) to 76% (2001-2005), the average shortest path drops from 8.02 hops to 5.43. The value of the clustering coefficient, which characterizes the local cohesiveness of a network, increased constantly over time reaching a very high value of 0.81 in the last time period, 2001-2005. This means that collaboration teams with three or more participants are common in the investigated network. This finding can be confirmed by inspecting the average number of authors per team with 4.59 authors per paper between 2001 and 2005.

The collaboration network demonstrates a tendency towards a shorter average distance between reachable pairs and higher clustering coefficient values. These results confirm *Hypothesis 3*. The small world property is characteristic for co-author networks and is very useful in networks for knowledge creation and knowledge diffusion. Nonaka (1994) emphasizes that the interaction between individuals is essential for the “knowledge conversion” from individual to inter-organizational knowledge via group and organizational levels. This transformation is very important and enables individual knowledge to be “amplified” and affects organizational knowledge “crystallization”.

In the next step, the significance of the differences between regression coefficients listed in Table 6-2 was checked using the t-test. To that end, the null hypothesis of no differences in slopes across phase I and phase II was assumed. Table 6-3 shows the results. Most regression coefficients revealed intensified tendency regarding network growth, cooperation intensity, and small world property in the demand-pull phase in comparison with the science-push phase. There are two exceptions. The percentage of isolated authors shows a slightly decreasing tendency. The significance of the difference in the number of organizations and in the ratio “mean path length to size of the giant component” can be assumed because in both cases one of the regression coefficients was not significant.

	Number of Papers	Number of Authors	Number of Organizations	av. Nr. of Papers per Authors	av. Nr. of Authors per Paper	Percentage of Isolated Authors
Difference between the slopes in Phase I and Phase II (t-testt)	5.24**	7.07**	-	4.80**	5.97**	0.66
	Av. Nr. of Partners per Author	Av. Nr. of Partners per Organization	LN (Share of SCC in the Authors NW)	Mean path length / Size of Giant Comp.	Cluster coefficient	Percentage of Isolated Organiza-tions
Difference between. the slopes in Phase I and Phase II (t-test)	4.81**	8.40***	5.21**	-	3.11**	4.24**

Table 6-3: Comparison of the regression coefficients.

To summarize, the PV collaboration network is growing. The growing number of investigated units, be it papers, authors, or organizations, supports this trend. The productivity of authors changes little, despite a trend toward increasing collaboration and the fact that small world development was observed. In the next section, different centrality measures for the co-author and inter-organizational networks will be discussed.

6.3 Analysis of Topological Characteristics

This section aims to analyze the structural properties of the collaboration network in the field of photovoltaic technology because network structure influences knowledge diffusion within a network (see also Newman 2001c). Two kinds of collaboration networks are studied here. The first possibility is to consider co-author networks. These networks allow individual relationships between scientists to be investigated. Another possibility is to investigate inter-organizational networks in which authors are pooled according to their affiliation. Analyzing these networks affords insights into cooperation between organizations.

In the previous section, general tendencies of network development were analyzed. The following investigation focuses on the analysis of the network centralization scores introduced in 6.1.2. Network degree and betweenness centralization scores will be calculated for both kinds of collaboration networks. Mote (2007) states that network centralization can be seen as a good indicator of the flow of knowledge and communication between individuals. Due to the existence of two phases in the PV technology cycle, one can assume that there are differences in the network structure of the two phases. These differences can be measured by centralization scores.

In general, it is difficult or impossible to say whether central or decentral structure is more beneficial. Each structure has advantages and disadvantages. High centralization improves efficiency by generating potential economies of scale. It also improves stability by minimizing opportunities for errors. Decentralized networks have potentially higher flexibility for reorganization. Consequently, it is interesting to see which structure is observed in science-push and demand-pull phases.

Table 6-4 shows two network centralization scores calculated for co-author and inter-organizational networks. Both indicators reveal similar development for both networks based on their centralization scores.

	1974-1980	1981-1985	1986-1990	1991-1995	1996-2000	2001-2005
Network degree centralization (co-author network)	0.013	0.012	0.011	0.011	0.012	0.022
Network betweenness centralization (co-author network)	0.004	0.009	0.023	0.049	0.045	0.054
Network degree centralization (inter-organizational networks)	0.15	0.034	0.042	0.043	0.069	0.24
Network betweenness centralization (inter-organizational networks)	0.085	0.007	0.006	0.006	0.109	0.28

Table 6-4: Network centralization indices in the field of PV.

Table 6-5 shows general centralization trends using linear regression on the basis of yearly centralization scores for the corresponding networks. Positive (negative) slope in regression is denoted as "⇑" ("⇓"resp.), i.e. the network becomes more central (decentral resp.). The significance of trends was derived from the significance of regression coefficients in the corresponding equations.

	1974-1990	1991-2005
Network degree centralization (co-author network)	⇓*	⇑
Network betweenness centralization (co-author network)	⇑*	⇑
Network degree centralization (inter-organizational networks)	⇓	⇑*
Network betweenness centralization (inter-organizational networks)	⇓	⇑**

Table 6-5: Trends in the development of centralization scores[22].

In the first phase (1974 to 1990), a slight overall decentralization trend for can be observed. At the organization level, there is no indication that a dominate core can control scientific cooperation activity within the network. At the author level, there is any clear tendency. The degree centralization score reveals decentralization, but the betweenness centralization indicates that the co-author network becomes more centrally structured.

The picture in the second phase (1991-2005) looks quite different. A consistent centralization tendency for both the networks becomes visible. In summary, the collaboration network at the author and organization level becomes more central in

[22] Note: *: at 0.05 significance level, ** at 0.01 significance level.

the sense of control over information flow (betweenness centralization) and in the sense of communication activity (degree centralization).

In the next step, network structure is analyzed using network density and inequality of degree distribution measured by the Gini coefficient[23]. Figure 6-5 shows roughly constant values of the Gini Coefficient in the time span from 1974 to 1990. There is no significant difference in the mean at the 0.05 level for this time span. After 1991 the Gini Coefficient increases significantly ($p < 0.05$).

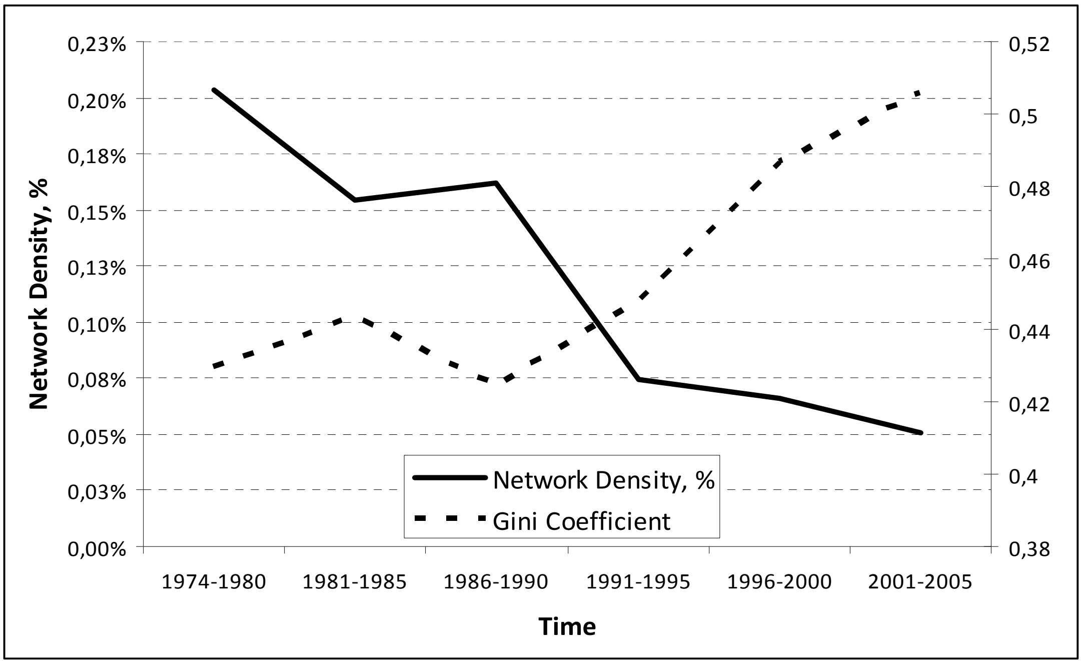


Figure 6-5: Gini coefficient and network density.

The Gini Coefficient values depict a tendency toward a more uneven distribution of linkages in the collaboration network in the second phase. At the same time the network density declines slightly from 1974 to 1990, and then it declines sharply again in the second phase. However, the density values are very low which indicate that the collaboration network in the field of photovoltaic is very loose.

Thus, two contrary trends relating to the level of interconnection (network density) and to the level of inequality in node degree (Gini Coefficient) can be observed. For the structure of the collaboration network, this implies that the network is less connected in the second phase of development from 1991 to 2005 and a number of additional links is divided under a small group of authors.

[23] The results of this calculation are given only for the authors' level, for the organizations' level similar statements can be given.

6.4 Analysis of Collaboration Behavior

So far, the network methods discussed considered the global topology of the network. All nodes were treated identically. This consideration is not realistic and is seldom the case for real-world networks. Nodes may differ in terms of categorical or numerical properties (see also Roediger-Schluga and Barber 2006). For example, the organization type to which an author belongs can influence cooperation strategy significantly. At the level of authors, the cooperation affinity is affected, for example, by the number of links that both agents have. The impact of organization type and author degree on the collaboration behavior will be analyzed in the next section.

6.4.1 Cooperation Behavior with Respect to Organization Type

The collaborative behavior is essentially impacted by the affiliations of scientists because the affiliation significantly influences the motives for cooperation. Hall, Link, and Scott (2000) investigated industry/university partnerships and identified differences in collaboration motives in universities and firms. The two core motives for industry partners were: firstly, access to complementary research activity and research results; secondly, access to key university personnel. The primary reason for universities to cooperate with industry is to find a partner which can finance the research.

In the context of this study, it seems appropriate to consider different organization types when forming affiliation groups. These are:

(1) Universities and higher educational facilities
(2) Research institutes
(3) Enterprises seeking profit
(4) Others.

First, the size of these four groups are considered. Figure 6-6 shows the group distribution by organization type and reveals that the relative size of the groups stays stable from 1974 to 2005. Roughly 50% of published organizations are universities, ca. 25% are research institutes, about 22% are firms, and only 3% are others. Universities clearly make up the largest part of organizations that published in the PV field, whereas research institutes and firms have a similar group size. What can one say about the position of these groups within the collaboration network? Which group is more important? Are there significant changes over time?

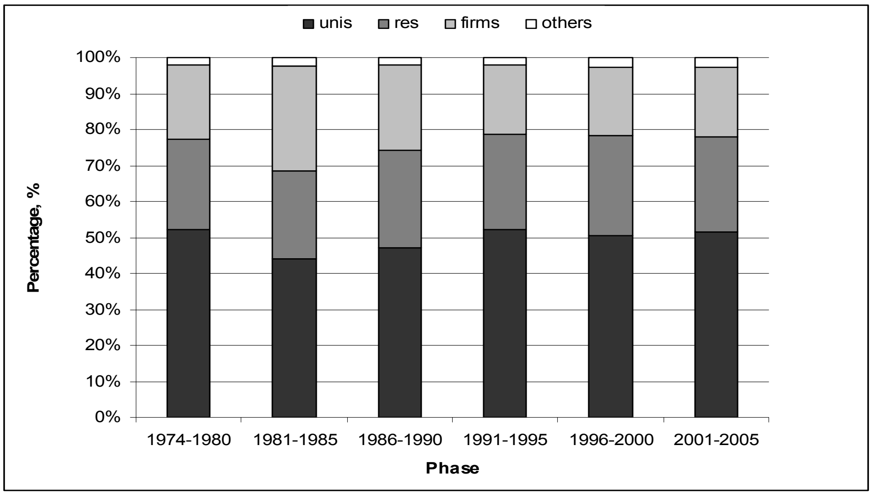


Figure 6-6: Distribution of organization types published in the PV field.

Dosi (1982, p. 147) documents the "degree of autonomy of the innovative activity" from market needs in science-push and demand-pull phases of market development. Consequently, it happens very seldom that the same actors determine the market situation over a long period of time. The role of universities and research institutes as collaboration partners is crucial, especially in the first phase.

Appendix 9.5 shows the top 20 organizations that published scientific papers in the field of photovoltaic technology from 1974 to 2005. Table 6-6 gives a summary of this top list. In the late 1970s and the early 1980s, 5 of the top 20 organizations were enterprises. After 1986, the percentage of firms fell continuously to 10 percent in 1991-1995, and finally in the last 10 years of the investigation period there were no private companies in the top 20 publishers. In contrast to this, the involvement of universities and research institutes together rose.

	Universities	Research Institutes	Firms
1974-1980	45%	30%	25%
1981-1985	35%	40%	25%
1986-1990	40%	40%	20%
1991-1995	45%	45%	10%
1996-2000	35%	65%	0%
2001-2005	65%	35%	0%

Table 6-6: Types of the top 20 organizations published in the PV field.

For example, Bell Laboratories[24] had a top position during the first phases 1974-1980 and 1981-1985. In 1954, the researchers of this company constructed the first

[24] Also known as Bell Labs or Bell Telephone Laboratories, part of the engineering department of the American Telephone & Telegraph company (AT&T).

silicon PV cell[25]. It was the first solar cell capable of converting enough of the sun's energy into power to run everyday electrical equipment. However, after 1986, Bell Laboratories has been active in the field of PV technology, but does not occupy one of the top positions regarding the number of scientific publications in the present sample.

Unfortunately, there was an announcement concerning the future of fundamental physics research by Bell Labs. In August 2008, Alcatel-Lucent, the parent company of Bell Labs, decided to pull out of basic research, material physics and semiconductor research and will instead be focusing on more immediately marketable areas such as networking, high-speed electronics, wireless, nanotechnology and software. The idea being to align the research work in the Lab closer to areas that the parent company is focusing on, according to Peter Benedict, spokesperson for Bell Labs and Alcatel-Lucent Ventures. It is only one example, but it is a well-known fact that private companies are more interested in fast product development and revenues than in enhancing general knowledge[26].

Hall et al. (2000) document that universities participate in research projects involving "new" science. The main objective of universities is research results. Firms can benefit from cooperation with universities by getting access to the latest technological achievements. Joint collaboration work between research units (universities and research institutes) and industry is required, so that new products can successfully capture markets. An innovation has a macro-economic impact only if it spreads quickly and widely. In the demand-pull phase, firms have to consider market needs more intensively. Academic institutions do not have the necessity to make research results ready for the market. In contrast, the main goal of companies is rapid commercialization of new technical knowledge in order to survive in competition.

Protection of intellectual property causes additional problems. Noll and Rogerson (1998) discuss the contracting problem which a government faces in supporting university research; however, similar considerations can be applied for private-public R&D cooperation efforts. Noll and Rogerson (1998, p. 108) consider a university as "a firm producing several outputs, including education and research". It is difficult

---

[25] One of the first applications for the Bell solar cell was to help run telephone lines in rural Georgia, where there were no other nearby power sources. Among the most famous inventions of Bell Labs are the transistor, laser and a series of contributions to computer and science technologies.

[26] http://blog.wired.com/gadgets/2008/08/bell-labs-kills.html (accessed on 28.12.2008)

enough to measure educational output, but measurement of research output is more complicated because "new ideas cannot be measured or weighted". As a consequence of these measurement problems, R&D agreements cannot be made based on measurable scientific output. Incomplete contracting creates an essential dilemma concerning intellectual property protection at a future date.

Strong disagreements over intellectual property, conflicting targets, differences in research cultures, and financial conflicts are only some of the possible reasons which make further collaborative efforts involving multiple partners difficult. According to these considerations two hypotheses can be derived:

Hypothesis 4: In the science-push phase of market development, structural positions of universities, research institutes, and firms within networks are balanced.
Hypothesis 5: In the demand-pull phase, the position of firms in scientific communities is more decentral in comparison with the position of universities and research institutes.

In order to test these hypotheses, the group centralities of each of the four groups has to be calculated. Unfortunately, the common centrality measures are defined for individual agents. Everett and Borgatti (1999) have introduced some possibilities for measurement of group centrality. According to the reduced model approach, all the members of an investigated group are replaced by one single "super" node. The "super" node is connected with another node if at least one member of the group was connected with it.

One disadvantage of the reduced approach is that the centrality measure is only one number. It is not possible to build a confidence interval for this indicator in order to make statements about statistical significance of values. In the context of this study, it is important to compare the centrality values between the four groups. For this reason, the average measurement of group degree centrality and group betweenness centrality were chosen. Using the t-test and ANOVA F-statistics the equality of means between group centralities was tested.

In order to consider differences between the science-push and demand-pull phases of market development, two sub-phases were selected, as an example: the period from 1974 to 1980 for science push, and the last part of the investigated period from 2001 to 2005 for demand-pull. There were no significant differences between mean group degree centrality and group betweenness centrality for universities, research institutes, and firms in the science-push phase.[27] The results given in Table 6-7 and Table 6-8 were not significant; the first number is the t-statistic, and the second is the ANOVA F-statistic. It was a relatively interesting finding; although the groups have different sizes, no one group seems to be dominant regarding scientific activity or control of information flow in the time period from 1974 to 1980.

1974-1980	Universities	Research Institutes	Firms
Universities		0.31 0.097	1.40 1.96
Research Institutes			0.93 0.87

Table 6-7: Comparison of the degree centrality, 1974-1980[28].

1974-1980	Universities	Research Institutes	Firms
Universities		1.08 1.17	1.18 1.40
Research Institutes			0.50 0.25

Table 6-8: Comparison of the betweenness centrality, 1974-1980[29].

A very different picture arose in the time span from 2001 to 2005. The results, which can be seen in Table 6-9 and Table 6-10, revealed strongly significant differences in means of group degree centrality ($p<0.01$) and in means of group betweenness centrality ($p<0.05$) between universities/firms and research institutes/firms.

2001-2005	Universities	Research Institutes	Firms
Universities		0.26 0.07	3.09*** 9.56***
Research Institutes			3.90*** 15.2***

Table 6-9: Comparison of the degree centrality, 2001-2005[30].

---

[27] Because of the small group size of “others”, this group was not included in calculations.

[28] Calculation: EVIEWS, the investigated period: 1974-1980.

[29] EVIEWS, the investigated period: 1974-1980.

[30] Calculation: EVIEWS, the investigated period: 2001-2005.

2001-2005	Universities	Research Institutes	Firms
Universities		1.41 1.98	2.43** 5.89**
Research Institutes			2.56** 6.57**

Table 6-10: Comparison of the betweenness centrality, 2001-2005[31].

Universities and research institutes have on average more collaborative partners and lie more often on the shortest paths (geodesics) between all reachable pairs of others nodes regarding their scientific publication activity. These results confirm *Hypothesis 4* and *5*.

6.4.2 Cooperation Behavior with Respect to Author Degree

Barabasi and Albert (1999) explain network evolution from a dynamic point of view by using two hypotheses: growth and preferential attachment. The first one argues that networks sequentially expand through the addition of new nodes and links between the new nodes and nodes already present in the network. This concept of collaboration networks can be found in 6.2.

The preferential attachment hypotheses suggests that nodes (here authors) enter sequentially and have certain preferences in choosing partners for attachment. Probability P that a new node will be connected to node i depends on connectivity k_i of that node. In mathematical terms, preferential attachment means that the probability that a node i with degree k_i acquires a link is:

$$P(k_i) = \frac{k_i}{\sum_i k_i} \qquad (6\text{-}1)$$

This attachment scheme is also known as the “rich get richer” effect. This means that new nodes tend to connect to nodes with a large degree. By comparing degree distributions of $P(k_i)$ over time, the phenomenon of preferential attachment can be investigated. Figure 6-7 illustrates discrete degree distributions for the six time periods in log scale for both axes.

[31] Calculation: EVIEWS, investigated period: 2001-2005.

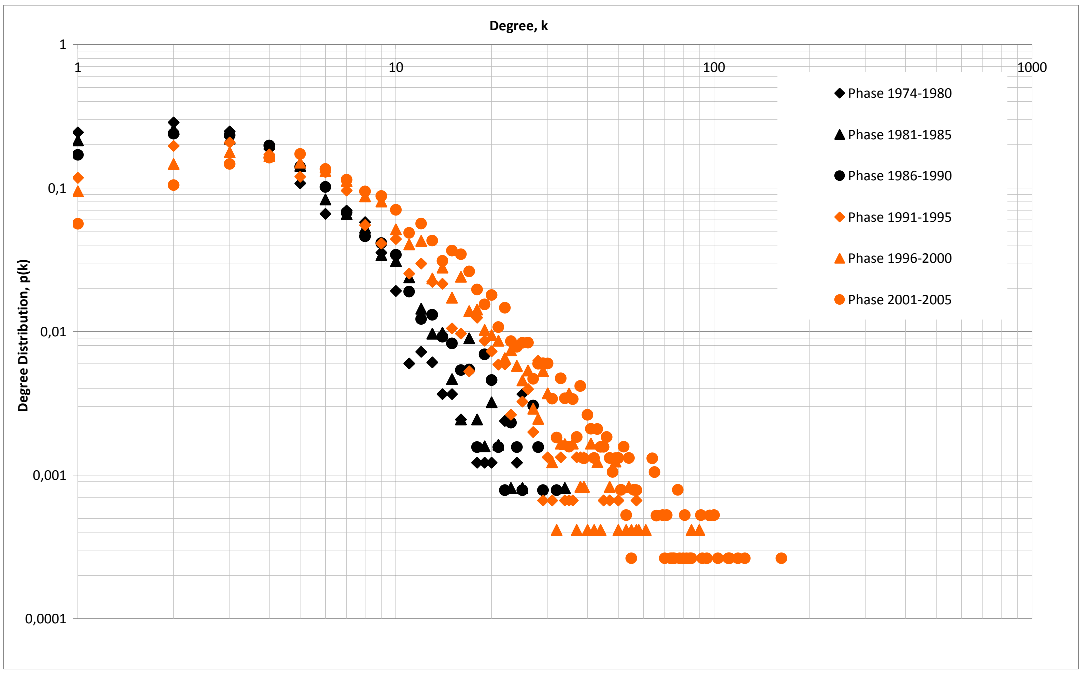


Figure 6-7: Co-authorship degree distributions in the PV field, 1974-2005.

The distributions remain relatively stable over time. However, one can see slight changes in the degree distribution for small *k*. The curves for the first three time periods (black curves) are flat for small k, i.e. the frequency of authors with 1 to 4 co-operation partners is relatively similar. The curves for the last three time periods (orange curves) have a positive gradient for small k. This means that the authors with only one link are rarer than the authors with two, three, or four links.

For $k>4$, the degree distributions decrease uniformly in k and follow a power-law tail:

$$P(k) \sim k^{-\gamma} \qquad (6\text{-}2)$$

where γ is a scale exponent. The population of authors with a lot of links grows according to the principle of "the rich get richer"; the "fat tail" of distribution shifts to the bottom on the right. There are two reasons that hinder preferential attachment (see also Amaral et al 2000):

The first reason is the aging of authors. In time, every author will stop publishing. For the network, this fact implies that even a very highly connected node will, eventually, stop receiving new links. The node is still part of the network and contributes to network statistics, but it no longer receives links. The aging of the nodes thus limits the preferential attachment.

The second reason is the costs of adding links to the nodes; the capacity of each author to create and maintain social ties is limited. One limitation is geographic distance. Certainly, internet use and computerization facilitate co-operation of researchers. New communication technologies relieve human-to-human communication considerably. Nevertheless, face-to-face contact is of great importance for people in scientific communities. Another limitation is time; each social contact requires time, and therefore, the remaining time for further research work and establishment of new contacts is reduced. Hence, physical costs of adding links and limited capacity of a node will limit the number of possible links attaching to a given node.

For occurrence of cooperation it seems to be important whether authors prefer to communicate with others of similar degree. As a consequence of these preferences, agents with similar social resources would come together. This property is also known as "assortative mixing". An analysis of degree correlation of connected nodes can detect occurrences of the "assortative mixing" phenomenon. This indicator can be derived from the joint probability *p(k, k')* that is defined as (see Ramasco et al. (2004)) the probability that a randomly selected pairs of nodes has degrees *k* and *k'*:

$$p(k,k') = \begin{cases} \dfrac{L(k,k')}{2L} & for \quad k = k' \\ \dfrac{L(k,k')}{L} & for \quad k \neq k' \end{cases} \qquad (6\text{-}3)$$

where *L(k, k')* is the number of links connecting nodes with degrees *k* and *k'*, and *L* is the total number of links in the network. Using this definition, one can calculate the average degree of the nearest neighbors of the nodes with degree *k*. Let $\overline{k_{nn}(k)}$ be the measure of the correlation between the degree of a node and that of its neighbors

$$\overline{k_{nn}(k)} = \frac{\bar{k}\sum_{k'} k'\, p(k,k')}{kp(k)} \qquad (6\text{-}4)$$

where $\bar{k}$ is the network's average degree. If large (small) degree nodes prefer to connect preferably to large (small) degree nodes, then the network represents assortative mixing. Figure 6-8 illustrates that the collaboration network in the field of PV technology is a typical assortative network where $\overline{k_{nn}}$ increases more or less monotonically with *k*. The tendency to cooperate with authors who have higher degrees in the demand-pull phase (after 1990) is stronger than in the science-push phase, since the correlation degree curves shift upwards. Figure 6-8 shows the existence of two scaling regimes. The divide lies roughly at 10 collaboration partners per author.

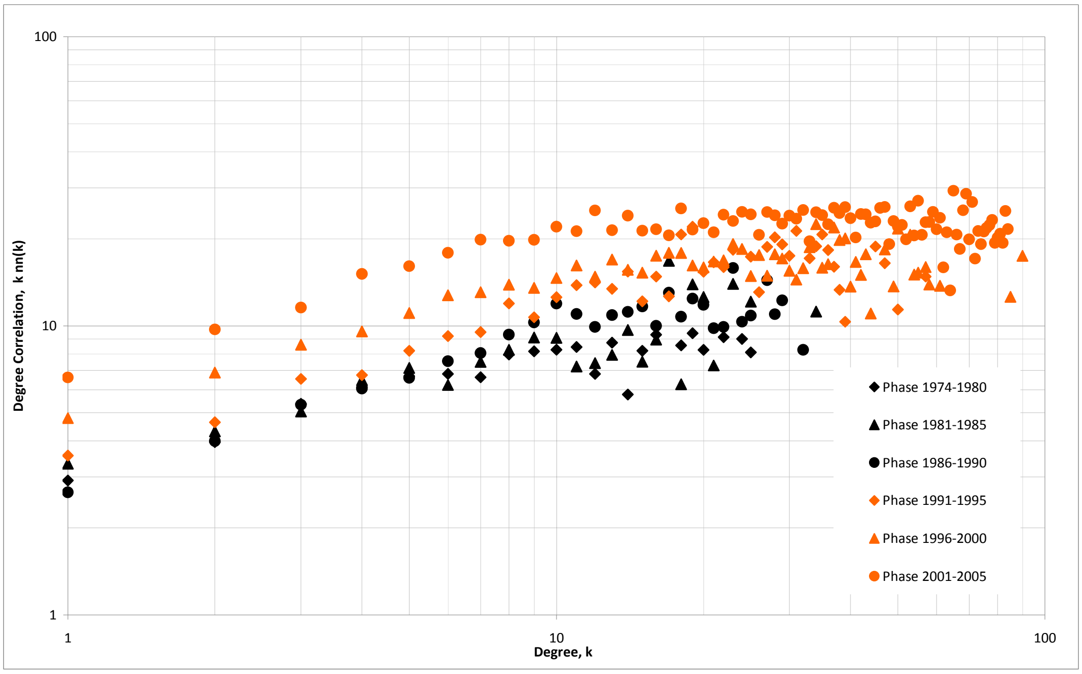


Figure 6-8: Degree correlation in the collaboration network for the PV field.

For *k≤10* a clear potential trend of the distributions can be identified (see Table 6-11). The gradient of this trend does not vary significantly over time; the intercept increases continuously with the exception of the time span 1986-1990.

	Slope	std. Dev.	intercept	Std. Dev.	r^2
1974-1980	0.45**	(0.09)	0.48**	(0.06)	0.97
1981-1985	0.44**	(0.10)	0.51**	(0.07)	0.96
1986-1990	0.62**	(0.08)	0.42**	(0.06)	0.99
1991-1995	0.55**	(0.10)	0.53**	(0.07)	0.98
1996-2000	0.49**	(0.06)	0.69**	(0.04)	0.99
2001-2005	0.52**	(0.07)	0.83**	(0.05)	0.99

Table 6-11: Results of the linear regression model for a degree ≤ 10.

The trend of the degree distribution for k>10 does not have this clear structure (see Table 6-12). One see instead a point cloud which moves to the right with time.

	Slope	std. Dev.	intercept	std. Dev.	r^2
1974-1980	0.21	(0.52)	0.66**	(0.18)	0.16
1981-1985	0.49	(0.95)	0.38	(0.32)	0.24
1986-1990	0.02	(0.50)	1.02**	(0.15)	0.01
1991-1995	0.05	(0.45)	1.13**	(0.12)	0.01
1996-2000	-0.03	(0.29)	1.26**	(0.07)	0.01
2001-2005	-0.04	(0.20)	1.42**	(0.04)	0.04

Table 6-12: Results of the linear regression model for a degree >10.

Wagner and Leydesdorff (2005) suppose that the fat tail distribution is formed by the existence of elite researchers whose are already very well connected within a network and compete to build networks of intellectual followers of the next generation. The slight negative, although non-significant, values of slope after 1996 show a disassortative cooperation tendency among prestigious scientists. The cooperation partners of prestigious researchers in the demand-pull phase are less connected than experienced researchers themselves.

6.5 Conclusion

The collaboration network of scientists and organizations in the technological field of solar photovoltaic power between the years 1974 and 2005 was considered in this chapter. Therein two scientists/organizations were considered to be connected with each other if they had published a common paper. The data for the analysis was collected from Web of Science using a keyword search strategy. Different network analytical techniques were used to examine changes in the structural properties of the network and modifications in cooperation behavior of agents over the whole investigated period. The results help to understand the pattern of connection and evolution of communication between authors. The main objectives were to analyze whether there are differences in network structure and changes in collaboration behavior for two phases of technological development: science-push (1974-1990) and demand-pull (1991-2005) phases. The following observations and conclusions can be drawn:

The collaboration network evolved rapidly, especially after 1991; the total number of papers, authors, and organizations increased significantly. This finding is in line with the fact that in the last decade of the 20th century there were numerous PV promotion programs which were introduced in various countries. For example, the German 1,000 Roofs program and the Electricity Feed-In Act, which ran in the early 1990s.

In order to get a more precise picture of network evolution in relation to changes in network topology over time; two level of analysis were considered: the level of individual researchers and the level of organizations that these researchers belonged to. In analyzing the descriptive statistics of the collaboration network, some general trends in network topology were identified.

First, the general trend towards a higher intensity of collaboration was proven by different indicators like the average number of collaborators per author or collaborators per organization, the relative share of isolated authors or organizations, and the average size of researcher groups that published a paper in the PV field. Thereby, in the second stage (demand-pull) a statisticaly significant increase in collaboration efforts at the author's and organization's level was revealed.

Second, contrary to the network expansion, only a moderate trend of growth in productivity of authors was observed. In the science-push phase, a slightly negative trend in author's productivity was discovered. A better development of scientific productivity was observed in the demand-pull phase. A statistically significant but modest positive slope was found.

Third, the investigated collaboration network demonstrated a tendency towards a broadening of its giant component, a shorter average node to node distance, and a slight increase in the clustering coefficient. It was a positive result because scientific knowledge can more easily diffuse between collaboration agents and can be more quickly adopted and converted into new products. On the other hand, learning processes between the agents would have been more effective if the agents were more heterogeneous. A highly connected network with short lines of communication loses diversity and therefore, the possibility to find new, radical solutions for technological problems. In contrast, a sparse network with longer average path length can potentially offer a better diversity of agents. For this reason, it can be quite useful for the agents to have different backgrounds, and still remain close enough to communicate with each other.

Two contrary trends relating to the level of interconnection (network density) and to the level of inequality in node degree (Gini coefficient) in the author network is observed. For the structure of the collaboration network, this implies that the network is less connected in the second phase of development from 1991 to 2005 and a number of additional links is divided among a small group of authors.

After analyzing general trends that reflect rather global tendencies in the network's development, the characteristic variability for individual actors relative to their position in the network is investigated. Since the intensity of collaboration activity reflects the objectives of the organizations that authors belong to, four organization types were considered: universities, research institutes, firms, and others. The analysis of centralization scores (degree and betweenness centralization on the author's and organization's level) revealed by and large a decentral tendency in network development in the science-push phase. However, the picture changed in the demand-pull phase. A continuous centralization tendency in the author and organization collaboration networks was identifiable.

The next research question related to changes in collaboration behavior of actors. Action of authors in searching for co-operation partners was decisively influenced by their affiliation. For this reason, in the first step, the inter-organizational network was considered.

In the science-push stage of market development, the innovative activity was more detached from market needs than in the demand-pull stage. Consequently, the role and position of research units and industrial actors did not remain constant over time. The question arises whether it is possible to measure this shifting using centrality measures on the group level of universities, research institutes, and firms. Here a positive answer was given.

Although the size of the groups stayed nearly constant over all investigated periods, a modification in position of each group was observed. In the science-push phase, there was a relatively balanced position between groups. According to size, universities were clearly the dominating group (~50% of all organizations). However, regarding group centrality score there was no dominating group. This finding holds for both approaches: degree centrality as a measurement of collaboration activity and betweenness centrality as an indicator for control over information flow.

For the demand-pull phase, the situation was different. The group betweenness and degree centrality scores of universities and research institutes were significantly higher than the corresponding group centrality values of enterprises. This finding supports the hypotheses that the position of firms concerning publication activity loses weight in the second phase. The publication of technological achievements was not the top priority for firms. Private companies were more interested in fast product development and revenues than in enhancing general knowledge.

In the second step, the evolution of the cooperative behavior at the author level was analyzed using the network degree distribution and the degree correlation, which were considered to be suitable approaches. Two additional network properties were identified, preferential attachment and assortative mixing.
The principle of “preferential attachment” i.e. the rich nodes get richer, showed that although the most authors in the network had a relatively small number of links, a few researchers had an enormous number of links. In general, the researchers with more links had more benefit and influence in the collaboration network.

The principle of ”assortative mixing” was investigated by analyzing degree correlation between connected authors. The results showed that authors preferred co-operation partners with just as many links as they had. Only the elite group of scientists, which were already very well connected within the network, did not care about competition for more co-authorship relations. The main goal of these researchers was rather to build networks of intellectual followers in the next generation.

In summary, the results of this chapter indicated that the topological characteristics of PV collaboration networks have changed to ensure quick information transfer. Furthermore, there was a shift in the role and importance of research and industrial actors over the investigated period. These findings support the hypotheses that some science-based technologies pass through different developmental stages: science-push and demand-pull. These phases are also caused by changes in the collaboration networks and the behavior of agents.

7 Final Conclusions

This dissertation aims to provide insights into the development pattern of science-based technologies. The transformation process of modern economies into science-based economies increases the importance of these technologies. However, the dynamics of the time scale of innovation processes in science-based markets has not been explored to any great extent. A number of investigations deal with the question of market formation in science-based sectors. Some of these are based on empirical data (e.g. Grupp and Schmoch, 1992, Schmoch, 2007); others use the rather qualitative experience of engineers and marketing experts (e. g. Gartner's Hype Cycle graph, Rickerby and Matthews, 1991). All these studies can be summarized as using a stylized model of the market formation in qualitative and "appreciative" terms (see Grupp, 1998). Three indicators are used in this model: (1) measurement of scientific activities by publication statistics, (2) measurement of technological development by patent applications or patent grants, respectively, and (3) measurement of installed or sold products (shipped respectively) in order to estimate diffusion. The basic hypothesis deriving from this model is that there are two quite different development phases in the technology cycle.

The first phase of development is denoted as the “science push” phase. The “voice of the market” is largely missing in this stage and the development goals are oriented towards internal success within scientific communities. During this period, consumer needs are more or less neglected. Discrepancies between demand and consumer preferences lead to a stagnation of the market. New solutions are thus required in order to make the innovative products technically and economically viable. Further improvements and investigations, but also exogenous factors such as government policy, facilitate the adjustment process between the scientific / technological potentials of new technologies and the demand side. The second phase of market development, known as the “demand pull” or “market push”, can then start.

However, up until the present time very little has been known about the real reasons for the occurrence of these phases. The stylized model is rather descriptive and does not provide any explanation for the observed trends within a technology cycle. The main goal of this thesis is thus to fill this gap and to identify determinants for this wave-like development.

In the first part of the thesis an operational definition of science-based technology is discussed. This definition is applied to 45 technological fields that are intuitively defined as science-based by Schmoch (2007). The calculation is carried out for patent applications with a priority date between 1989 and 2004 in relation to these 45 fields. The patent applications include at least one reference. Using Non-Patent Citation (NPC) Methodology and two indicators, the scientific dependence of these technologies is measured. According to the first indicator (the share of non-patent literature citations among all citations), 30 of the listed technologies have significantly higher values of science intensity in comparison with all patent applications at the EPO. According to the second indicator (the average number of non-patent literature references), 40 of the investigated technologies have above-average science intensity. However, both approaches yield strong correlated results regarding the ranking of technologies according to their science intensity.

Analyzing the development pattern of these technologies, the following result can be achieved: 23 of the investigated 45 fields graph a strong double-boom pattern, and for a further ten fields a weakly formed double boom can at least be considered. Two fields seem to be in the first period of development and only the remaining ten fields indicate no double-boom course, but continual increased scientific development. Still not all science-driven technologies are subject to the wave-development pattern, but quite a large share of these technologies do seem to follow this pattern.

In order to analyze the reasons for the occurrence of the wave pattern more deeply, the technological field of “solar photovoltaic (PV) cells” is analyzed explicitly. The investigation period encompasses the years from 1974 to 2005. According to the "double-boom hypothesis" the whole data set is split into two time periods: from 1974 to 1990 and from 1991 to 2005. The analysis of the PV field consists of three parts.

The first part deals with a longitudinal analysis of the solar module quality and the market-share development of the corresponding solar module producers. This part of the thesis is carried out in three steps: first, the existence of competition within the solar PV market is tested for using the coordination failure diagnostics (CFD) concept. This concept helps to detect market disequilibrium by analyzing five processes. The results point to the existence of competition in the PV market. Second, the technological progress of solar modules is investigated using the metric rescaling approach. Regarding the solar modules of nine producers (RWE Schott, BP Solar, Kyocera, Helios, Photowatt, Siemens Solar, Solarex, Solec/Sanyo, and Sharp), a continuous improvement in the technological performance of solar modules can be documented. The solar modules of RWE SCHOTT Solar perform very well and demonstrate the best progress from a technical point of view. The analysis shows that Solarex is the worst manufacturer in the group of the investigated companies in terms of its lower rate of technological improvement. Third, the results of the metric rescaling approach are compared with the changes in the market shares of solar cell producers. In this way, one can test whether the firms offering advanced solar modules have a better position in the market. The results provide a differentiated picture. In the earlier phases of market development, a lack of understanding of customer needs plus high production costs made it hard for the solar cell producers to meet market requirements. Additional complications caused by the stock market downturn of 2002 had negative effects on the semiconductor industry in particular, as a result of the absence of venture capital.

Summarizing the results of the first part, an analysis of the technological process of solar modules is possible only in the "demand pull" phase. In the case of photovoltaic field, the necessary data is available only after 1987. For this reason, only five years of the investigated period belong to the "science push" phase. During these five years, the presence of discrepancies between consumer needs and the existing characteristics of solar modules are visible. These differences remain, however in the later "demand pull" phase a positive impact of the technological improvement of the products on their market position is recognizable. The technological background of technologies is closely linked with its intrinsic characteristics. One crucial factor is the maturity of technological solutions in terms of market needs. If new technological solutions can be quickly adapted for new applications and the achievement of short-term market success is possible, then diffusion of technologies can proceed relatively quickly. Nevertheless, it should be kept in mind that the development of these fields has a long-term perspective.

The process of answering the following three research questions forms the focus of the analysis in the second part of this thesis:

– Is there "growth equilibrium" between any of the indicators that measured different activities of market development in the stylized model?

– Is it possible to measure the extent of this "growth equilibrium" and the interdependencies between indicators describing the market development in general?

– To what extent can the equilibrium between different indicators be affected by exogenous factors?

In order to answer these questions a time series analysis is used. The market development is measured using six time series: compensations, subsidies provided by the German government, patent and publication statistics, development of oil prices, and installed PV capacity data. Five of the six variables show significant non-stationary properties as a consequence of high dynamics in the PV market development. In this case, an application of standard regression techniques is non-valid due to the problems with spurious regression. Consequently, an application of the co-integration approach is more appropriate. This method deals with the non-stationary problem by searching for stationary relationships between the model variables.

Relatively persistent “growth equilibrium” is proven between logarithmic number of patent applications and number of scientific publications. In the “science push” phase of market development this relationship is affected by increased oil prices. The first and second oil price crises can be regarded as starting points that stimulated the search process for renewable energy sources in general. The costs of the fossil energy sources become less important with the further development of scientific and technological activities over the course of time. The “demand phase” is stimulated more by supportive government policy. With the exception of the variables pair patent applications and scientific publications, any co-integration relationship can be established. Using the Granger causality test, causal relationships for compensations vs. patents and compensations vs. installed capacities are shown. Both variable pairs reveal the driving force of the EEG on the development of the German PV market.

These relatively long cycles imply that exogenous factors such as the political, social, and economic environment can have different impacts, depending on the stage of development. It seems reasonable to suppose that certain factors, like oil or financial crises, can only be controlled to a limited degree. However, the choice of support measures has to be adapted to the development stage of the corresponding technological field.

The essential role of scientific knowledge in the development of the science-based technologies is widely accepted. The increased complexity of these technologies implicates the importance of collaboration networks in the scientific community. The evolution of these networks is the subject of the third empirical part of this thesis. Two research objects are analyzed using social network analysis: changes in the topological structure of the collaborator networks and changes in cooperative behavior of co-authors and those organizations to which the authors belong. The most important findings are as follows:

A rapid growth of collaboration networks especially in the “demand pull” phases of market development is obvious. The total number of papers, authors, and organizations increases considerably. Regarding cooperation intensity between authors and organizations, a statistically significant stepping-up of collaboration efforts on the author and the organization’s part is observable. Despite these trends there is no remarkable increase in the productivity of authors. However, the intensification of small-world properties in the “demand phase” facilitates a faster information transfer between researchers. Another important result is the significant shift in the network position of firms during the “demand pull” phase. In the “science push” phase industrial actors seek intensive cooperation with universities and research units in order to participate in the actual technological achievements. In the “demand pull” phase firms no longer take a central position in the collaborative scientific network. The competitive pressure for profit forces companies to pay more attention to consumer needs.

Finally, regarding degree distribution and degree correlation, two characteristic properties, “assortative mixing” and “preferential attachment” of the network, are analyzed. According to the mechanism of “preferential attachment” those authors who already have a lot of co-publications are favored in terms of being chosen for preparation of further publications. Such researchers derive more benefit from and have more control within the network. According to the principle of “assortative mixing” the authors in the network prefer to be connected with partners that have more links than they themselves have. However, authors with a lot of links tend to have other objectives. The main goal of the most “prominent” authors is to find intellectual followers. For this reason they are often less concerned about co-publications with other well-connected authors, preferring instead to work with young researchers. Both of these principles are more clearly observable in the “demand pull” phase of market development.

To sum up, determinants of the technology cycle in the case of science-based technologies differ in nature. The quality of data significantly affects the selection of suitable analyzing techniques. In this thesis, the application of different techniques of analysis proves to be very beneficial. As a consequence of the transformation process of modern economies towards science-based economies, the analysis of the long-term development of science-based technologies at the meso-level becomes increasingly important for innovation theory in general and will remain an important research field in the future.

8 References

Aigner D., C.A.K. Lovell, P. Schmidt (1977): Formulation and Estimation of Stochastic Frontier Production Function Models. Journal of Econometrics 6 (July): 21

Akaike H. (1974): A new look at the statistical model identification, in: IEEE Transactions on Automatic Control 19 (6): 716–723.

Amaral L.A.N., Scala A., Bartelemy M., Stanley H.E. (2000): Classes of small world networks, in: Proceedings of the National Academy of Science (PNAS) 97 (21), pp. 1181-11152.

Arrow K.J (1962): "Economic Welfare and the Allocation of Resources for Invention", in R. Nelson (ED), The Rate and Direction of Inventive Activity. Princeton University Press: Princeton, pp 609 – 625.

Bacchiocchio E., Montobbio F. (2004): EPO vs. USPTO citation lags, URL: ftp://ftp.unibocconi.it/pub/RePEc/cri/papers/WP161Montobbio.pdf

Bala V., Goyal S. (1998): Learning from neighbors, in: Review of Economic Studies, 65, pp. 595-621.

Barabási A.-L., Albert R. (1999): Emergence of scaling in random networks, in: Science 286, pp. 509-512.

Barabási A.-L., Jeong1 H., Néda Z., Ravasz E., Schubert A., Vicsek T. (2002): Evolution of the social network of scientific collaborations. Physica A 311, pp. 590-614.

Barro R.J. (1999): Notes on Growth Accounting, in: Journal of economic Growth Vol. 4(2), pp. 119-37.

Bass F. (1969): A new product growth model for consumer durables, in: Management Science 15 (5): pp. 215–227.

Bass F (1969): A New Product Growth Model for Consumer Durables; Management Science 15, pp. 116 – 131.

Baudisch A. (2005): Ko-Evolution von Konsumentenverhalten und industrieller Organisation, Beschreibung des Dissertationsvorhabens, Max-Planck-Institut für Ökonomik, Jena.

Berkowitz S.D. (1982): An introduction to structural analysis: The network approach to social research. Toronto: Butterworth.

Blanckenburg K. (2007): Methods of Testing the Workability of Markets, in: Prinz, A., Steenge, A.E. and Schmidt, J. (Eds.): The Rules of the Game: Institutions, Law and Economics, Wirtschaft: Forschung und Wissenschaft, Bd. 22, LIT: Muenster, 2007, pp.187-203.

Blind K. (2004): The Economics of Standards – Theory, Evidence, Policy, Cheltenham: Edward Elgar.

BMU (2007): Renewable Energy Sources Act. Progress Report 2007. URL: http://www.wind-energie.de/fileadmin/dokumente/Themen_A-Z/EEG/bmu_erfahrungsbericht_eeg_summary_en.pdf accessed on 21.12.2008

Bundesgesetzblatt (2000), Teil I, Nr. 13, issued Bonn on 31. März 2000 [BGBl. I, p. 305).

Bush V. (1945): Science, The Endless Frontier. Public Affairs Press, Washington, DC.

Cohen L., Levinthal D.A. (1989): Innovation and Learning: Two Faces of R&D. Economic Journal , 99 (397), pp. 569-596.

Cowan R., Jonard N. (2004): Network structure and the diffusion of knowledge, in: Journal of Economic Dynamics and Control, Vol. 28(8), pp. 1557-1575.

Coward H. R., Franklin J. J. (1989): Identifying the Science-Technology Interface: Matching Patent Data to a Bibliometric Model, in: Science, Technology & Human Values, Vol. 14 (1), pp. 50-77.

Czarnitzki D., Glänzel W., Hussinger K. (2007): Heterogenety of Patenting Activity and Its Implications for Scientific research. URL: ftp://ftp.zew.de/pub/zew-docs/dp/dp07028.pdf

David P. A., Bronwyn H. H., Andrew A. T. (2000): Is public R&D a complement or substitute for private R&D? A review of the econometric evidence, in: Research Policy. 29: 497–529.

Davidson R., MacKinnon J. C. (1993): Estimation and inference in econometrics, New York/ Oxford.

DeBresson C., Hu X., Drejer I., Lundvall B.-A. (1998): Innovative Activity in the Learning Economy – a comparison of systems in 10 OECD countries. OECD draft report.

Dickey D.A., Fuller, W.A. (1979): Distribution of the Estimators for Autoregressive Time Series with a Unit Root, in: Journal of the American Statistical Association, 74, pp. 427–431.

DOD (1969): Project Hindsight, Office of the Director Defense Research and Engineering. Washington, DC: DTIC No. AD495905, October 1969.

Doreian (2001): Causality in Social Network Analysis. Sociological Methods and Research 30(1), pp. 81-114.

Dosi G. (1982): Technological Paradigms and Technological Trajectories – A suggested interpretation of the determinants and directions of technical change, in: Research Policy Vol. 11, pp. 147-162.

Drejer I. and J∅rgensen B.H. (2005): The Dynamic creation of knowledge: Analysing public-private collaborations, in: Technovation 25, pp. 83-94.

Earl P.E., Wakeley T. (2007): Conjectures about future wants: some insights from evolutionary economics with reference to digital photography. Int. J. Technology Intelligence and Planning 3 (1): pp. 24-38.

Engle R.F., Granger C.W.J. (1987): "Co-integration and Error Correction: Representation, Estimation, and Testing," Econometrica, Econometric Society, vol. 55(2), pages 251-76, March.

EPIA (2008): Global Market Outlook for Photovoltaics until 2012: URL: http://www.epia.org/fileadmin/EPIA_docs/publications/epia/EPIA__MarketPublication_18feb.pdf (09.08.2008)

Everett M.G., Borgatti S. P. (1999): The centrality of groups and classes; in: Journal of Mathematical Sociology 3, pp. 181-201.

Fuller, W., (1976): Introduction to Statistical Time Series. Wiley, New York.

Franz P. (2008): Wie reagieren Hochschulen auf neue Technologien? Das Bespiel der Photovoltaik, in : Wirtschaft im Wandel 12(14), pp. 460-467.

Freeman L.C. (1979): Centrality in Social Networks Conceptual Clarification, in: Social Networks, 1 (1978/79), pp. 215-239.

Freeman C. (1982): The Economics of Industrial Innovation, London: Frances Printer.

Geels F.W. (2005): The Dynamics of Transitions in Socio-technical Systems: A Multi-level Analysis of the Transition Pathway from Horse-Drawn Carriages to Automobiles (1860-1930). Technology Analysis and Strategic Management, December 2005, v. 17, iss. 4, pp. 445-76.

Geels F.W. (2006): Non-linearity and Expectations in Niche-Development Trajectories: Ups and Downs in Dutch Biogas Development (1973-2003). Technology Analysis and Strategic Management, July-September 2006, v. 18, iss. 3-4, pp. 375-92.

Godin B. (2006): The Knowledge-Based Economy: Conceptual Framework or Buzzword? in: The Journal of Technology Transfer, Vol. 31, No. 1, pp. 17-30.

Granger C.W.J. (1969): Investigating Causal Relations by Econometric Models and Cross Spectral Methods, Econometrica, 37, pp. 424-428.

Granger C.W.J. (2003): Nobel Lecture, December 8, 2003. URL: http://nobelprize.org/nobel_prizes/economics/laureates/2003/granger-lecture.pdf

Grosskettler H. (1999): Das Koordinationsmängel-Diagnosekonzept als didaktisches Instrument, in: Lübke, E., Grossekettler, H. (ed.), Beiträge zur Wirtschafts- und Finanzpolitik- Festschrift für Ingolf Metze, Berlin Duncekr&Humblot.

Grupp H., Schmoch U. (1992): At Crossroad in Laser Medicine and Polyimide Chemistry. Patent Assessment of the Expansion of Knowledge. In: Grupp, H. (ed.): Dynamics of Science-based Innovation, Berlin/ Heidelberg/ New York: Springer-Verlag, p. 269-302.

Grupp H. (1994) The measurement of technical performance of innovation by technometrics and its impact on established technology indicators. Research Policy, 23:175-193

Grupp H. (1996): Spillover effects and the science base of innovations reconsidered: an empirical approach, in: Journal of evolutionary Economics, vol. 6, pp. 175-197.

Grupp H. (1998): Foundations of the Economics of Innovation. Theory, Measurement and Practice. Edward Elgar, Cheltenham, UK. Northampton, MA, USA.

Grupp H., Maital S. (2001): Managing New Product Development and Innovation. A Microeconomic Toolbox. Edward Elgar Publishing, Cheltenham.

Grupp H, Schubert T. (2008): Review and new evidence on composite innovation indicators for evaluating national performance. Forthcoming

Hagedoorn J. (1993): Understanding the Rationale of Strategic Technology Partnering: Interorganizational Modes of Cooperation and Sectoral Differences, in: Strategic Management Journal, Vol. 14, pp. 371-385.

Hall B.H., Link A.N., Scott J.T. (2000): Universities as Research Partners, NBER Working Paper No. 7643, April 2000.

Haller I., Grupp H. (2007): Mathematical Modelling of Innovative Dynamics: An Empirical Analysis of the Photovoltaic Market in Germany. Manuscript, IWW, Karlsruhe University, KIT, Karlsruhe, 2007.

Haller I., Grupp H. (2009): Demand by Product Characteristics: Measuring Solar Cell Quality over Time. Forthcoming.

Hamel G. (1991): Competition for Competence and Interpartner Learning within International Strategic Alliances, in Strategic Management Journal of International Economics Vol. 12, pp. 83-103.

Hatzichronoglou T. (1997): Revision of the High-Technology Sector and Product Classification. Paris: OECD; at http://www.olis.oecd.org/olis/1997doc.nsf/LinkTo/OCDE-GD(97)216.

Hayek F.A. (2002): Competition as a discovery procedure, in: the quarterly journal of austrian economics Vol. 5, No. 3 (Fall 2002): pp. 9–23.

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

Hirooka M. (2003): Nonlinear dynamism of innovation and business cycles, in: Journal of Evolutionary Economics 13: pp. 549-576.

Hoffmann V.U., Becker H., Decker B., Erge T., Grochowski J., Heilscher G., Riess H., Rindelhardt U. (1998): 1000-Dächer-Meß- und Auswerteprogramm. Wissenschaftlicher Endbericht. Fraunhofer Publica.

Höft U. (1991): Lebenszykluskonzepte. Grundlage für das strategische marketing- und Technologiemanagement, Berlin, Erich Schmidt Verlag.

Howells J.(1999): Research and technology outsourcing, in: Technology Analysis and Strategic Management 11 (1999) (1), pp. 17–29.

IEA PVPS (2008): Annual report 2007.Implementing Agreement on photovoltaic power systems. URL: http://www.iea-pvps.org/products/download/IEA%20PVPS%20Annual%20Report%202007-web.pdf (accessed 21.12.2008)

Illinois Institute of Technology Research Institute (1968): Technology in Retrospect and Critical Events in Science. Prepared for the National Foundation by the Illinois Institute of Technology Research, Contract NSF-C535. Report, December 1968. Technology in Retrospect and Critical Events in Science

Ironmonger D.S. (1972): New Commodities and consumer Behaviour. University Press, Cambridge

Jacobsson S., Sandén B.A., Bångens L. (2004): Transforming the Energy System – the Evolution of the German Technological System for Solar Cells, in: Technology Analysis & Strategic Management, Vol. 16, No. 1, pp. 3-30.

Jannsen S. (2005): Auf dem Weg in die solare Zukunft, Deutsche Gesellschaft für Sonnenenergie.

Jäger-Waldau A. (2008): PV Status Report 2008, URL: http://re.jrc.ec.europa.eu/refsys/pdf/PV%20Report%202008.pdf (accessed 17.12.2008)

Johansen, S. (1995): Likelihood based inference in cointegrated vector autoregressive models, Oxford University Press, Oxford.

Jungmittag A. (1995): Langfristige Zusammenhänge und kurzfristige Dynamiken zwischen Direkt-investitionen und Exporten – Eine mehrstufige Modellierung dynamischer simultaner Mehrgleichungsmodelle bei kointegrierten Zeitreihen, Dissertation, Berlin 1996: Volkswirtschaftliche Schriften, Duncker & Humblot.

Kemp R., Schot J., Hoogma R. (1998): Regime Shifts to Sustainability through Process of Niche Formation: The Approach of Strategic Niche Management, in: Technology Analysis & Strategic Management 10(2), p. 175 – 195.

Kennedy P. (2008): A Guide to econometrics. Sixth Edition, Blackwell Publishing.

Khan R., Merfeld D., Pearsall T. P., Geyer M., Dauskardt R.H. (2008): Innovations in solar power, in: Advanced Materials & Processes, Nov2008, Vol. 166 Issue 11, pp. 45-48.

Kondratieff N.D. (1926): Die langen Wellen der Konjunktur. Archiv für Sozialwissenschaft und Sozialpolitik 56: pp. 573-606.

Kostoff R. N. (1997): Symposium overview: Accelerating the conversion of science to technology, in: The Journal of Technology Transfer, Vol. 22 No.3, pages 3-8.

Kretschmer H., Aguillo I. (2004): Visibility of collaboration on the Web, in: Scientometrics, 61, pp. 405–426.

Kuhn T. (1970): The structure of scientific revolution(2nd ed..). Chicago: University of Chicago Press.

Kumbhakar S.C., C.A.K. Lovell (2000): Stochastic Frontier Analysis. Cambridge, UK: Cambridge University Press.

Lancaster K.J. (1966): A new approach to consumer theory, in: The Journal of Political Economy, Vol. 74, No. 2 (Apr., 1966), pp. 132-157.

Lancaster K.J. (1971): Consumer Demand, A New Approach, New York, London.

Lancaster KJ (1991): Modern Consumer Theory. Edward Elgar, Aldershot.

Laukamp H., Kiefer K., Kaiser R., Schmidt H. (2000): Reliability issues in PV systems – experience and improvements, 2nd World Solar Electric Buildings Conference.

Leuchtner J., Boekstiegel C. (1991): Photovoltaik – Marktübersicht: Informationen zur Stromerzeugung mit Solarzellen, 5th edition: Öko-Institut e.V., Institut für angewandte Ökologie, Freiburg.

Leuchtner J., Preiser K. (1995): Photovoltaik-Anlagen: Marktübersicht 1994/1995, 2nd edition: Öko-Institut e. V., Institut für angewandte Ökologie, Freiburg.

Lindqvist S. (1994): Changes in the Technological Landscape: The Temporal Dimension in the Growth and Decline of Large Technological System, in Granstrand, p. 271-288.

Lorenz P., Pinner D., Seitz T., Farrell D., Remes J.K. (2008): The economic of solar power, in: McKinsey Quartely, Issue 4, pp. 66-78.

Lunch P. J. (2007): Investing in PV Solar Energy: http://www.earthtoys.com/emagazine.php?issue_number=07.08.01&article=lynch

Lundvall B.A. (1992): National Systems of Innovation. Towards a Theory of Innovation and Interactive Learning. London: Pinter Publishers.

Lundvall B.A. (1996): The social Dimension of the Learning Economy, Aalborg, DRUID WP 96-01.

Machlup F. (1962): The Production and Distribution of Knowledge in the United States, Princeton University Press.

Madalla G.S. (1992): Introduction to Econometrics, second edition. New York: Macmillan Publishing Company.

Marinova D., Balaguer A. (2009): Transformation in the photovoltaics industry in Australia, Germany and Japan: Comparison of actors, knowledge, institutions and markets, in: Renewable Energy 34, pp. 461-464.

Marsili O. (2001): The Anatomy and Evolutions of Industries. Edward Elgar, Cheltenham and Nortamton.

Martin B.R. (1992): Research Foresight and the Exploitation of the Science Base, Report submitted to the Office of Science and Technology, Cabinet Office, SPRU, Sussex University.

Meereis J. (1987): Photovoltaik-Marktübesicht. Öko-Institut e.V., Institut für angewandte Ökologie, Freiburg.

Meyer M. (2000a): Does science push technology? Patent citing scientific literature, in: Research Policy 29, pp. 409-434.

Meyer M. (2000b): Patent citations in a novel field of technology. What can they tell about interactions between emerging communities of science and technology? Scientometrics Vol. 48 (2), pp. 151-178.

Mokyr J. (1990): The Lever of Riches, New York: Oxford University Press.

Morone P., Taylor R. (2004): Knowledge diffusion dynamics and network properties of face-to-face interactions, in: Journal of Evolutionary Economics, 14, 3, pp. 327-351.

Mote J. E., Jordan G., Hage J., Whitestone Y. (2007): New directions in the use of network analysis in research and product development evaluation, in: Research Evaluation, 16(3), September 2007, pp. 191–203.

Myrdal G. (1957): Economic theory and under-developed Regions. London, Ducksworth.

Narin F., Noma E. (1985): Is technology becoming science? In: Scientometrics, vol. 7, pp. 369-381.

Narin F., Hamilton K. S., Olivastro D. (1997): The increasing linkage between US technology and public science. in: Research Policy, Vol. 26, pp. 317-330.

Nelson R. R. (1959): The simple of economics of Basic Scientific Research, in: Journal of Political Economy, Vol. 67, pp. 297-306.

Newman M.E.J. (2001a): Scientific collaboration networks. I. Network construction and fundamental results, in: Physical Review E, Volume 64, 016131.

Newman M.E.J. (2001b): Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality, in: Physical Review E, Volume 64, 016132.

Newman M.E.J. (2001c): The structure of scientific collaboration networks. Proceedings of the National Academy of Sciences USA 98, pp- 404-409.

Newman M.E.J. (2002): Assortative mixing in networks, URL: http://arxiv.org/PS_cache/cond-mat/pdf/0205/0205405v1.pdf (26.10.2008)

Nisan S., Benzarti N. (2008): A comprehensive economic evaluation of integrated desalination systems using fossil fuelled and nuclear energies and including their environmental costs, in: Desalination 229 (2008) pp. 125–146.

Noll R. G., Rogerson W. (1998): The Economics of University Indirect Cost Reimbursement in Federal Research Grants, in: Challenges to Research Universities, Roger G. Noll, ed., Washington: Brookings Institution.

Nonaka I. (1994): A Dynamic Theorie of Organizational Knowledge Creation", in: Organization Science (5)1, pages 14-37.

de Nooy W., Mrvar A., Batagelj V. (2005): Exploratory Social Network Analysis with Pajek, Cambridge University Press.

Noyons, E.C.M., van Raan A.F.J., Grupp H., Schmoch U. (1994): Exploring the science and technology interface: inventor-author relations in laser medicines research, in: Research Policy, Vol. 23, pp. 443-457.

OECD (1986): OECD Science and Technology Indicators: R&D, Invention and Competitiveness.

OECD Observer (June-July 1990)

OECD (1994): The Measurement of Scientific and Technological Activities: Using Patent Data as Science and Technology Indicators Patent Manual, URL: http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/Annexes/pat_td_ntot_sm1_an2.pdf .

OECD (1996a): Technology, Productivity and Job Creation. Paris: OECD.

OECD (1996b): The Knowledge-based economy. Paris: OECD.

OECD (2002): Frascati Manual – Proposed Standard Practice for Surveys on Research and Experimantal Development, OECD Publishing 2002, available at www.oecd.org.

OECD (2007): OECD Science, Technology and Industry Scoreboard 2007.

OECD (2008): OECD Factbook 2008.

Otte E., Rousseau R. (2002): Social network analysis: a powerful strategy, also for the information sciences, in: Journal of Information Science, 28, pp. 441–453.

Partucco P.P. (2002): Review Article: Social and Contractual Interactions in the Production of Technological Knowledge, in: Information Economics and Policy 14, pp. 405-416.

Pavitt K. (1984): Sectoral patterns of technical change: Towards a taxonomy and a theory, Research Policy 13, p. 343-373.

Pavitt K. (1998): 'Do patents reflect the useful research output of universities?' Research Evaluation, 7(2) pp. 105–111.

Penrose ET (1959): The Theory of the Growth of the Firm. Wiley, New York

Perlin J. (2002): From space to earth. The story of solar electricity. Harvard University Press.

Pilbert C. (2006): Barriers to technology diffusion: the case of solar thermal technologies. International Energy Agency: http://www.oecd.org/dataoecd/46/14/37671704.pdf

Prahalad C.K., Hamel G. (1990): The Core competence and the corporation, in: Harvard Business Review 68: 79-91.

Van Raan, A.F.J. (1997): Scientometrics: State-of-the-Art, Scientometrics 38(1), p. 205-218.

Ramasco J. J., Dorogovtsev S. N., Pastor-Satorras R. (2004): Self-Organization of Collaboration networks, in: Physical Review E 70, 036106.

Rao B.B. (1994): Editor's introduction in: Rao, B.B. (Hrsg.): Cointegration for the applied economists, London/New York, pp. 1-8.

Reichmuth M., Hünnekes C. (2003): Monitoring results of the German 100 000 Roofs Programme. 3rd World Conference of Photovoltaic Energy Conversion, May 11-18, 2003 Osaka, Japan.

Rickerby D. S., A. Matthews (1991): Market Perspectives and Future Trends. In: Rickerby, D. S. and A. Matthews (eds.): Handbook of Surface Engineering. Glasgow / London: Blackie, p. 343-364.

Roediger-Schluga T., Barber M.J. (2006): The structure of R&D collaboration networks in the European Framework Programmes. UNU-MERIT Working Papers, 36.

Rothwell R., Zegveld W., (1985): Reindustrialization and Technology. Longman. Harlow.

Rothwell R. (1994): Towards the Fifth-generation Innovation Process, in: International Marketing Review, Vol. 11 No. 1, pp. 7-31.

Sahal D. (1985): Foundation of Technometrics. Technological Forecasting and Social Change, 27:1-37

Sakakibara M. (2003): Knowledge Sharing in Cooperative Research and Development, in: Managerial and Decision Economics Vol. 24, pp. 117-132.

Saviotti P.P. (1996): Technology Evolution, Variety and the Economy. Edward Elgar, Cheltenham and Brookfield

Saviotti P.P., Metcalfe J.S. (1984): A Theoretical Approach to the Construction of Technological Output Indicators. Research Policy, 13:141-151

Scherer F.M. (1982): Inter-Industry technology flows in the United States, in Research Policy, vol 11, pp. 227-245.

Scherer F.M. (1986): The World Productivity Growth Slump, in R. Wolff (ed.): Organizing Industrial Development, Berlin, New York: Walter de Gruyter, p. 15-27.

Scherer F.M. (1992): Schumpeter and Plausible Capitalism, in: Journal of Economic Literature, Vol. 30, pp. 1416-1433.

Schmoch U. (1993): Tracing the knowledge transfer from science to technology as reflected in patent indicators, in: Scientometrics Vol. 26 (1), pp. 193-211.

Schmoch U. (2007): Double-boom cycles and the comeback of science-push and market-pull, in: Research Policy 36, pages 1000-1015.

Schmookler (1966): Invention and Economic Growth, Cambridge, Mass.: Harvard University Press.

Schumpeter J. A. (1939): Business Cycles. McGraw-Hill, New York.

Schwarz G. (1978): Estimating the Dimension of a Model, in: The Annals of Statistics, Vol. 6, No. 2, (Mar., 1978), pp. 461-464.

Smith R. (1987): The roots of innovation, in: British Medical Journal, Vol. 295, pp. 1335-1337.

Solow R. (1957): Technical Change and the Aggregate Production Function, in: The Review of Economics and Statistics, Vol. 39, No. 3., pp. 312-320.

Sørensen B. E. (2005): Granger Causality. ECONOMICS 7395, Spring, 2005. URL: http://www.uh.edu/~bsorense/gra_caus.pdf.

Spanos A. (1995): Statistical foundation of econometric modelling. Cambridge University Press.

Spence M. (1984): Cost Reduction, Competition, and Industry Performance, in: Econometrica, Vol. 52, No. 1, pp. 101-121

Stier W. (2001): Methoden der Zeitreihenanalyse. Berlin: Springer.

Stokes D.E. (1997): Pasteur's quadrant: Basic science and technological innovation. Washington, DC: Brookings Institution.

Tenn S., Yun J.M. (2008): Biases in demand analysis due to variation in retail distribution. International Journal of Industrial Organization, 26:984-997

Tijssen R.J.W. (2002): Science dependence of technologies: evidence from inventions and their inventors, in: Research Policy 31, pp. 509-526.

Tindall D., Wellman B. (2001): "Canada as social structure: Social network analysis and Canadian Sociology," Canadian Journal of Sociology, 26 (2), 265-308.

Tornatzky L.G., Fleischer M. (1990): The Process of Technological Innovation, Lexington Books.

Van der Duin P., Ortt R., Kok M. (2007): The Cyclic Innovation Model: A New Challenge for a Regional Approach to Innovation Systems?, in: European Planning Studies Vol. 15, No. 2, February 2007.

Vernon R. (1966): International Investment and International Trade in the Product Cycle, Quarterly Journal of Economics 80:190-207

Wadman W.M. (2000): Variable Quality in Consumer Theory: Toward a Dynamic Microeconomic Theory of the Consumer, M.E Sharpe, Armonk, New York; London, England.

Wagner C.S., Leydesdorff L. (2005): Network structure, self-organization, and the growth of international collaboration in science, in: Research Policy 34, pp. 1608-1618.

Walsh J. (1973): Technological innovation:new study sponsored by NSF take socioeconomic, managerial factors into account, in: Science 180, pp. 846-847.

Wasserman S., Faust K. (1994): Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press.

Watts D.J., Strogatz S.H. (1998): Collective dynamics of "small-world" networks, in: Nature, 393: 440-442.

Wellman B. (1988): Structural analysis: From method and metaphor to theory and substance, in: B. Wellman & S. D. Berkowitz (Ed.), Social structures: A network approach (pp. 19-61). Cambridge: Cambridge University Press.

Wiener N. (1956): The theory of prediction. In Beckenbach, E.F. (Ed.), Modern mathematics for engineers, New York: McGraw-Hill, series 1, pp. 165–190.

WIPO (2008): World Patent Report – A statistical review, 2008. URL: http://www.wipo.int/export/sites/www/ipstats/en/statistics/patents/pdf/wipo_pub_931.pdf .

Wooding S. (2007): Setting the historical context for project retrosight, in: RAND Europe, WR-466-RS, URL: http://www.rand.org/pubs/working_papers/2007/RAND_WR466.pdf

Zweibel K., Hersch P. (1982): Basic Photovoltaic Principles and Methods, New York. Van Nosstrand Reinhold Company, Inc.

9 Annex

9.1 Java Code

```
/*The Program initializes the author, organizations, and countries
lists. Three tabels are generated:
ShortName x Orgas, ShortName x Author, ShortName x Country. The
total number of publications
can be determined by adding of columns.
The ShartName-Tables can be taken as a basis for the analysis of
the weighted networks.*/

import java.awt.*;
import java.awt.event.*;
import javax.swing.*;
import java.io.*;
import java.util.*;
import java.util.regex.*;
import java.nio.charset.*;


public class PubDataShortName_chooseYears extends JFrame
{
      private class MeinWindowListener extends WindowAdapter
      {
            public void windowClosing(WindowEvent arg0)
            {
            System.exit(0);
      }
      }
      private JComponent contentPane = (JComponent)
      getContentPane();
      private JButton button;
      private JLabel a=new JLabel("Jahre die zu berücksichtigen
      sind mit ; getrennt");
      private JTextArea e1=new JTextArea("");
      private JLabel b=new JLabel("AusgabePfad");
      private JTextArea e2=new JTextArea("");
      private JLabel c=new JLabel("EingangsPfad");
      private JTextArea e3=new JTextArea("");
      private JPanel panel=new JPanel();
      private JPanel buttons=new JPanel();


      private String eingang=new String();
      private String eingang2=new String();
      private String ausgang=new String();




      PubDataShortName_chooseYears() //Konstruktor - Fenster
      {
      super("Build Matrix");
      setSize(700, 300);
      initLayout();
      setLocationRelativeTo(null);
      addWindowListener(new MeinWindowListener());
            setVisible(true);
            setResizable(true);
      }
```

```
public static void main(String[] args)
{
      PubDataShortName_chooseYears frame = new
PubDataShortName_chooseYears();
frame.setLocation(150, 90);
frame.setResizable(false);
frame.setVisible(true);
}

void initLayout() //Layout
{
      contentPane.setLayout(new BorderLayout());

button=new JButton("START");
buttons.setLayout(new FlowLayout(FlowLayout.CENTER));
buttons.setVisible(true);
buttons.add(button);
contentPane.add("South", buttons);

panel.setLayout(new FlowLayout(FlowLayout.LEFT));
panel.setVisible(true);
panel.add(c);
e3.setPreferredSize(new Dimension(550, 20));
panel.add(e3);
panel.add(b);
e2.setPreferredSize(new Dimension(560, 20));
panel.add(e2);
panel.add(a);
e1.setPreferredSize(new Dimension(10370, 20));
panel.add(e1);


contentPane.add("Center", panel);

      button.addActionListener(new ActionListener()
  {
            public void actionPerformed(ActionEvent e)
//Knopffunktion
            {
                  eingang=e1.getText();
                  eingang2=e3.getText();
                  ausgang=e2.getText();
                  start();
            }
      });
}



public static ArrayList organisations; //Liste der
Organisationen
public static ArrayList fields; //Liste der Felder
public int [] jahre;
public BufferedWriter PajekOrgasBw;
public PrintWriter PajekOrgasPw;
public BufferedWriter PajekFieldsBw;
public PrintWriter PajekFieldsPw;


public void start ()
{
      try
      {
```

```
                PajekOrgasBw=new BufferedWriter(new
      FileWriter(new File(ausgang+"PajekOrgasFile.net")));
                PajekOrgasPw=new PrintWriter(PajekOrgasBw);
                      PajekFieldsBw=new BufferedWriter(new
      FileWriter(new File(ausgang+"PajekFieldsFile.net")));
                PajekFieldsPw=new PrintWriter(PajekFieldsBw);
                String[] dummy=eingang.split(";");
                jahre=new int[dummy.length];
                for (int i=0;i<jahre.length;i++)
                      jahre[i]=Integer.parseInt(dummy[i]);

                organisations= new ArrayList();
                fields= new ArrayList();
                initList();
                buildTableOne();
                buildTableTwo();

                System.out.println("Done");
                PajekFieldsBw.close();
                PajekFieldsPw.close();
                PajekOrgasBw.close();
                PajekOrgasPw.close();


          }
          catch(Exception e) {System.out.println(e.toString());}

      }




      public void initList()    //2 Listen erstellen
      {
          try
          {
                int counter=0;
                for (int j=0;j<jahre.length;j++)
                {
                      int jahr=jahre[j];
                      BufferedReader br1=new BufferedReader(new
FileReader(eingang2+jahr+".txt"));  //Inputstream für das jahr
erstellen
                      br1.readLine();  //Titelzeile lesen und
miussachten
                      while (br1.ready())
                      {
                            String[]
dummy=br1.readLine().trim().toUpperCase().split("\\t");  //Zeile an
Tabs trennen
                            counter++;
                            String[]
felder=dummy[46].toUpperCase().split(";"); //autorenteil bereinigen
                            for (int i=0;i<felder.length;i++)
//für alle in der Zeile vorkommenden Autoren
                            {
                                  if
(!fields.contains(felder[i].trim()))
                                        {

      fields.add(felder[i].trim());
```

```
                                          }          //Liste
ergänzen falls noch nicht vorhanden
                              }

                              if
(dummy[20].trim().toUpperCase().length()>5)              //Umgehen
von fehl/nicht vorhandenen Informationen
                              {

                                    String[]
orgaLand=dummy[20].toUpperCase().split(";");


                                    for (int i=0;
i<orgaLand.length; i++)
                                    {
                                          if
(orgaLand[i].indexOf("BREMERHAVEN")>0)
                                          {
                                          if
(orgaLand[i].lastIndexOf(",")!=-1)
                                          {

                                          String
orga=orgaLand[i].substring(0,orgaLand[i].indexOf(",")).trim().toUpp
erCase();
                                          //String
orga=orgaLand[i].substring(0,orgaLand[i].indexOf(",",(orgaLand[i].i
ndexOf(",")+1))).trim().toUpperCase();
                                                /*if
(orga.length()<4)
                                                {
                                                      String
orga1=orgaLand[i].substring(0,orgaLand[i].lastIndexOf(",")).trim();
                                                      if
(orga1.lastIndexOf(",")!=-1)

      orga=orga1.substring(0, orga1.lastIndexOf(","));
                                                      else
orga=orga1;
                                                }
                                          */
                                                if
(!organisations.contains(orga))

      organisations.add(orga);        //Wie bei Autoren



                                          }
                                          }
                                    }
                              }
                        }
                        br1.close();

                  }
                  PajekFieldsPw.println("*vertices
"+(counter+fields.size())+" "+counter);
                  PajekOrgasPw.println("*vertices
"+(counter+organisations.size())+" "+counter);

            }
            catch(Exception e) {System.out.println(e.toString());}
```

```
	}

	public void buildTableOne()   //Tabelle SortName/Organisation
erstellen
	{
		try
		{
			BufferedWriter bw=new BufferedWriter(new
FileWriter(new File(ausgang+"Orgas.txt")));
			PrintWriter pw=new PrintWriter(bw);
			BufferedWriter bw1=new BufferedWriter(new
FileWriter(new File(ausgang+"orgas.vec")));
			PrintWriter pw1=new PrintWriter(bw1);
			BufferedWriter bw2=new BufferedWriter(new
FileWriter(new File(ausgang+"orgas.csv")));
			PrintWriter pw2=new PrintWriter(bw2);
			BufferedWriter bw3=new BufferedWriter(new
FileWriter(new File(ausgang+"orgas_liste.txt")));
			PrintWriter pw3=new PrintWriter(bw3);


			for (int i=0;i<organisations.size();i++)

	pw.print(";"+organisations.get(i).toString());
			pw.println();
			int[] publnumber_orgas = new
int[organisations.size()];
			int counter=1;
			for (int j=0;j<jahre.length;j++)
			{
				int jahr=jahre[j];
				int counter2=1;
				BufferedReader br1=new BufferedReader(new
FileReader(eingang2+jahr+".txt"));
				br1.readLine();

				while (br1.ready())
				{

	System.out.println(jahr+"_"+(counter2++));
					String[]
dummy=br1.readLine().trim().toUpperCase().split("\\t");
					String
autor=dummy[1].toUpperCase().split(";")[0];
					String LastNm;
					String FirstNm;
					if (autor.split(", ").length>1)
					{
						LastNm=autor.split(",
")[0].replace(" ","");
						FirstNm=autor.split(",
")[1].substring(0,1);
					}
					else
					{
						LastNm="[ANON]";
						FirstNm="";
					}


					if (LastNm.length()>8)
						LastNm=LastNm.substring(0,8);
```

```
      pw.print(LastNm+"_"+FirstNm+"("+dummy[25].trim()+")"+dummy[26
].trim()+":"+dummy[31].trim());
                              PajekOrgasPw.println((counter++)+"
"+LastNm+"_"+FirstNm+"("+dummy[25].trim()+")"+dummy[26].trim()+":"+
dummy[31].trim());
                              for (int
i=0;i<organisations.size();i++)
                              {

      if(dummy[20].contains(organisations.get(i).toString().toUpper
Case()))
                                    {
                                          pw.print(";1");
                                          publnumber_orgas[i]+=1;
                                    }
                                    else
                                          pw.print(";0");
                              }
                              pw.println();
                        }
                        br1.close();
                  }
                  for (int i=0;i<organisations.size();i++)
                        PajekOrgasPw.println((counter++)+"
"+organisations.get(i).toString());

                  /*Ranglisten für Orgas ausgeben*/
                  pw1.println("*vertices "+organisations.size());

                  for (int i=0;i<organisations.size();i++)
                  {
                        pw1.println(publnumber_orgas[i]);

      pw2.println(organisations.get(i).toString().toUpperCase()+";"
+publnumber_orgas[i]);
                  }
                  for (int i=0;i<organisations.size();i++)
                  {
                        pw3.println((i+1)+"
"+"\""+organisations.get(i).toString().toUpperCase()+"\"");
                  }

                  pw.close();
                  bw.close();
                  pw1.close();
                  bw1.close();
                  pw2.close();
                  bw2.close();
                  pw3.close();
                  bw3.close();

                  BufferedReader br=new BufferedReader(new
FileReader(ausgang+"Orgas.txt"));
                  br.readLine();

                  PajekOrgasPw.println("*matrix");

                  while(br.ready())
                  {
                        String[] dummy=br.readLine().split(";");
                        for (int i=1;i<dummy.length;i++)
                              {
```

```
                                    PajekOrgasPw.print(dummy[i]+"
");
                              }

                        PajekOrgasPw.println();
                  }






            }
            catch (Exception e) {System.out.println(e.toString());}
            System.out.println("DONE1");
      }


      public void buildTableTwo()   //Tabelle ShortName/Felder
erstellen
      {
            try
            {

                  BufferedWriter bw=new BufferedWriter(new
FileWriter(new File(ausgang+"Felder.txt")));
                  PrintWriter pw=new PrintWriter(bw);
                  BufferedWriter bw1=new BufferedWriter(new
FileWriter(new File(ausgang+"feld.vec")));
                  PrintWriter pw1=new PrintWriter(bw1);
                  BufferedWriter bw2=new BufferedWriter(new
FileWriter(new File(ausgang+"feld.csv")));
                  PrintWriter pw2=new PrintWriter(bw2);

                  int[] publnumber_feld = new int[fields.size()];
                  for (int i=0;i<fields.size();i++)
                        pw.print(";"+fields.get(i).toString());
                  pw.println();
                  int counter=1;
                  for (int j=0;j<jahre.length;j++)
                  {
                        int jahr=jahre[j];
                        BufferedReader br1=new BufferedReader(new
FileReader(eingang2+jahr+".txt"));
                        br1.readLine();
                        while (br1.ready())
                        {
                              String[]
dummy=br1.readLine().trim().toUpperCase().split("\\t");;
                              String
autor=dummy[1].toUpperCase().split(";")[0];
                              String LastNm;
                              String FirstNm;
                              if (autor.split(", ").length>1)
                              {
                                    LastNm=autor.split(",
")[0].replace(" ","");
                                    FirstNm=autor.split(",
")[1].substring(0,1);
                              }
                              else
                              {
                                    LastNm="[ANON]";
```

```
                                FirstNm="";
                        }


                        if (LastNm.length()>8)
                                LastNm=LastNm.substring(0,8);

      pw.print(LastNm+"_"+FirstNm+"("+dummy[25].trim()+")"+dummy[26
].trim()+":"+dummy[31].trim());
                        PajekFieldsPw.println((counter++)+"
"+LastNm+"_"+FirstNm+"("+dummy[25].trim()+")"+dummy[26].trim()+":"+
dummy[31].trim());

                        for (int i=0;i<fields.size();i++)
                        {

      if(dummy[46].contains(fields.get(i).toString().toUpperCase())
)
                                        {
                                        pw.print(";1");
                                        publnumber_feld[i]+=1;

                                        }
                                else
                                        pw.print(";0");
                        }
                        pw.println();

                }
                br1.close();
            }


            /*Ranglisten für Autoren ausgeben*/
            pw1.println("*vertices "+fields.size());
            for (int i=0;i<fields.size();i++)
            {
                pw1.println(publnumber_feld[i]);

      pw2.println(fields.get(i).toString().toUpperCase()+";"+publnu
mber_feld[i]);
            }


            pw.close();
            bw.close();
            pw1.close();
            bw1.close();
            pw2.close();
            bw2.close();

            BufferedReader br=new BufferedReader(new
FileReader(ausgang+"Felder.txt"));
            br.readLine();

            PajekFieldsPw.println("*matrix");

            while(br.ready())
            {
                String[] dummy=br.readLine().split(";");
                for (int i=1;i<dummy.length;i++)
                        {
                                PajekFieldsPw.print(dummy[i]+"
");
```

```
                    }

                PajekFieldsPw.println();
            }

        }
        catch (Exception e) {System.out.println(e.toString());}
        System.out.println("DONE2");
    }

}
```

9.2 Results of Granger-Causality Test

Pairwise Granger Causality Tests
Date: 12/29/08 Time: 13:29
Sample: 1968 2008
Lags: 1

Null Hypothesis:	Obs	F-Statistic	Probability
PUBL does not Granger Cause SUBS	30	0.10063	0.75351
SUBS does not Granger Cause PUBL		0.15614	0.69584
PATENTS does not Granger Cause SUBS	30	0.05734	0.81256
SUBS does not Granger Cause PATENTS		0.33359	0.56834
COMPENS does not Granger Cause SUBS	14	0.10920	0.74726
SUBS does not Granger Cause COMPENS		0.40077	0.53963
INSTALL_CAP does not Granger Cause SU	15	0.03081	0.86359
SUBS does not Granger Cause INSTALL_CAP		0.05737	0.81475
OIL does not Granger Cause SUBS	30	0.43159	0.51677
SUBS does not Granger Cause OIL		0.35052	0.55875
PATENTS does not Granger Cause PUBL	32	0.80740	0.37629
PUBL does not Granger Cause PATENTS		0.33198	0.56894
COMPENS does not Granger Cause PUBL	14	0.21822	0.64952
PUBL does not Granger Cause COMPENS		0.70038	0.42047
INSTALL_CAP does not Granger Cause PU	15	13.1068	0.00344
PUBL does not Granger Cause INSTALL_CAP		5.23506	0.04108
OIL does not Granger Cause PUBL	32	1.41956	0.24314
PUBL does not Granger Cause OIL		6.86011	0.01387
COMPENS does not Granger Cause PATE	14	0.00240	0.96183
PATENTS does not Granger Cause COMPENS		1.77521	0.20968
INSTALL_CAP does not Granger Cause PA	15	0.00937	0.92450
PATENTS does not Granger Cause INSTALL_CAP		0.00139	0.97085
OIL does not Granger Cause PATENTS	37	0.04605	0.83136
PATENTS does not Granger Cause OIL		1.09787	0.30213
INSTALL_CAP does not Granger Cause CO	16	0.43871	0.51932
COMPENS does not Granger Cause INSTALL_CA		10.6396	0.00619
OIL does not Granger Cause COMPENS	14	0.06838	0.79853
COMPENS does not Granger Cause OIL		9.83008	0.00949
OIL does not Granger Cause INSTALL_CAP	15	1.49888	0.24434
INSTALL_CAP does not Granger Cause OIL		11.7552	0.00500

Pairwise Granger Causality Tests
Date: 12/29/08 Time: 13:30
Sample: 1968 2008
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Probability
PUBL does not Granger Cause SUBS	29	0.05810	0.94368
SUBS does not Granger Cause PUBL		0.22769	0.79807
PATENTS does not Granger Cause SUBS	29	0.03545	0.96522
SUBS does not Granger Cause PATENTS		0.14567	0.86521
COMPENS does not Granger Cause SUBS	13	1.96021	0.20286
SUBS does not Granger Cause COMPENS		1.22345	0.34388
INSTALL_CAP does not Granger Cause SU	14	1.31701	0.31500
SUBS does not Granger Cause INSTALL_CAP		0.05546	0.94637
OIL does not Granger Cause SUBS	29	0.39252	0.67961
SUBS does not Granger Cause OIL		0.26851	0.76679
PATENTS does not Granger Cause PUBL	31	4.04907	0.02945
PUBL does not Granger Cause PATENTS		1.30078	0.28946
COMPENS does not Granger Cause PUBL	13	2.69491	0.12743
PUBL does not Granger Cause COMPENS		0.94630	0.42768
INSTALL_CAP does not Granger Cause PU	14	10.9554	0.00388
PUBL does not Granger Cause INSTALL_CAP		0.38432	0.69157
OIL does not Granger Cause PUBL	31	0.63284	0.53907
PUBL does not Granger Cause OIL		3.57089	0.04264
COMPENS does not Granger Cause PATE	13	1.59201	0.26180
PATENTS does not Granger Cause COMPENS		8.68002	0.00990
INSTALL_CAP does not Granger Cause PA	14	0.58489	0.57702
PATENTS does not Granger Cause INSTALL_CAP		0.48342	0.63181
OIL does not Granger Cause PATENTS	36	0.07762	0.92549
PATENTS does not Granger Cause OIL		0.66779	0.52006
INSTALL_CAP does not Granger Cause CO	15	1.61964	0.24586
COMPENS does not Granger Cause INSTALL_CA		7.10763	0.01201
OIL does not Granger Cause COMPENS	13	0.97650	0.41739
COMPENS does not Granger Cause OIL		8.48382	0.01054
OIL does not Granger Cause INSTALL_CAP	14	0.38884	0.68870
INSTALL_CAP does not Granger Cause OIL		10.9588	0.00387

Pairwise Granger Causality Tests
Date: 12/29/08 Time: 13:31
Sample: 1968 2008
Lags: 3

Null Hypothesis:	Obs	F-Statistic	Probability
PUBL does not Granger Cause SUBS	28	0.19665	0.89749
SUBS does not Granger Cause PUBL		0.27873	0.84010
PATENTS does not Granger Cause SUBS	28	0.68910	0.56880
SUBS does not Granger Cause PATENTS		0.10566	0.95585
COMPENS does not Granger Cause SUBS	12	1.37238	0.35232
SUBS does not Granger Cause COMPENS		0.52058	0.68650
INSTALL_CAP does not Granger Cause SU	13	3.06444	0.11286
SUBS does not Granger Cause INSTALL_CAP		0.55960	0.66093
OIL does not Granger Cause SUBS	28	0.24073	0.86131
SUBS does not Granger Cause OIL		0.26888	0.84708
PATENTS does not Granger Cause PUBL	30	2.81283	0.06187
PUBL does not Granger Cause PATENTS		0.93435	0.44007
COMPENS does not Granger Cause PUBL	12	4.72710	0.06372
PUBL does not Granger Cause COMPENS		0.54458	0.67282
INSTALL_CAP does not Granger Cause PU	13	10.2589	0.00890
PUBL does not Granger Cause INSTALL_CAP		1.15350	0.40147
OIL does not Granger Cause PUBL	30	0.39163	0.76016
PUBL does not Granger Cause OIL		2.08092	0.13057
COMPENS does not Granger Cause PATE	12	1.47849	0.32700
PATENTS does not Granger Cause COMPENS		3.59502	0.10110
INSTALL_CAP does not Granger Cause PA	13	0.67412	0.59869
PATENTS does not Granger Cause INSTALL_CAP		2.42117	0.16426
OIL does not Granger Cause PATENTS	35	0.37654	0.77060
PATENTS does not Granger Cause OIL		0.45608	0.71511
INSTALL_CAP does not Granger Cause CO	14	4.00154	0.05958
COMPENS does not Granger Cause INSTALL_CA		15.2617	0.00187
OIL does not Granger Cause COMPENS	12	3.77862	0.09325
COMPENS does not Granger Cause OIL		2.62872	0.16217
OIL does not Granger Cause INSTALL_CAP	13	6.83949	0.02308
INSTALL_CAP does not Granger Cause OIL		5.46297	0.03761

Pairwise Granger Causality Tests
Date: 12/29/08 Time: 13:46
Sample: 1968 2008
Lags: 4

Null Hypothesis:	Obs	F-Statistic	Probability
PUBL does not Granger Cause SUBS	27	0.12089	0.97321
SUBS does not Granger Cause PUBL		0.32402	0.85817
PATENTS does not Granger Cause SUBS	27	0.48026	0.74989
SUBS does not Granger Cause PATENTS		0.09926	0.98132
COMPENS does not Granger Cause SUBS	11	2.78922	0.28092
SUBS does not Granger Cause COMPENS		0.63118	0.68866
INSTALL_CAP does not Granger Cause SU	12	9.19306	0.04945
SUBS does not Granger Cause INSTALL_CAP		0.49744	0.74504
OIL does not Granger Cause SUBS	27	0.26629	0.89577
SUBS does not Granger Cause OIL		0.34520	0.84387
PATENTS does not Granger Cause PUBL	29	2.39268	0.08486
PUBL does not Granger Cause PATENTS		0.33605	0.85041
COMPENS does not Granger Cause PUBL	11	480.612	0.00208
PUBL does not Granger Cause COMPENS		2.06068	0.35239
INSTALL_CAP does not Granger Cause PU	12	6.55092	0.07724
PUBL does not Granger Cause INSTALL_CAP		2.69086	0.22114
OIL does not Granger Cause PUBL	29	0.98994	0.43567
PUBL does not Granger Cause OIL		1.39720	0.27088
COMPENS does not Granger Cause PATE	11	0.49461	0.75270
PATENTS does not Granger Cause COMPENS		17.1141	0.05597
INSTALL_CAP does not Granger Cause PA	12	0.41586	0.79204
PATENTS does not Granger Cause INSTALL_CAP		1.09384	0.49027
OIL does not Granger Cause PATENTS	34	0.37131	0.82680
PATENTS does not Granger Cause OIL		0.98003	0.43622
INSTALL_CAP does not Granger Cause CO	13	6.26437	0.05163
COMPENS does not Granger Cause INSTALL_CA		2.75084	0.17534
OIL does not Granger Cause COMPENS	11	2.75582	0.28356
COMPENS does not Granger Cause OIL		1.86280	0.37845
OIL does not Granger Cause INSTALL_CAP	12	11.9744	0.03451
INSTALL_CAP does not Granger Cause OIL		4.27453	0.13126

9.3 Top 20 List of Organizations Published in the field of PV.

1974-1980	1981-1985	1986-1990
GE (3)	CALTECH (2)	OSAKA UNIV (1)
RCA LABS (3)	SOLAR ENERGY RES INST (2)	SOLAR ENERGY RES INST (2)
BELL TEL LABS INC (3)	JET PROP LAB (2)	UNIV DELAWARE (1)
CALTECH (2)	UNIV DELAWARE (1)	WEIZMANN INST SCI (2)
JET PROP LAB (2)	STANFORD UNIV (1)	CNRS (2)
IBM CORP (3)	EXXON RES & ENGN CO (3)	UNIV TEXAS (1)
SO METHODIST UNIV (1)	CEN (2)	SPIRE CORP (3)
UNIV DELAWARE (1)	CNRS (2)	INDIAN INST TECHNOL (2)
STANFORD UNIV (1)	MIT (1)	CALTECH (2)
RUTGERS STATE UNIV (1)	RUTGERS STATE UNIV (1)	FRAUNHOFER INST (2)
SANDIA LABS (2)	NASA (2)	HITACHI LTD (3)
MCMASTER UNIV (1)	APPL SOLAR ENERGY CORP (3)	ENTERPRISE RADIO COMPONENTS & SEMICOND (3)
UNIV CALIF (1)	INDIAN INST TECHNOL (2)	PURDUE UNIV (1)
NASA (2)	WEIZMANN INST SCI (2)	STANFORD UNIV (1)
ROCKWELL INT (2)	ARCO SOLAR (3)	UNIV ILLINOIS (1)
UNIV ILLINOIS (1)	ARIZONA STATE UNIV (1)	UNIV NEW S WALES (1)
BROWN UNIV (1)	BELL TEL LABS INC (3)	WESTINGHOUSE ELECT CORP (3)
CNRS (2)	CATHOLIC UNIV LEUVEN (1)	ELECTROTECH LAB (2)
COLORADO STATE UNIV (1)	MICROELECTR (3)	SOLID STATE PHYS LAB (2)
MOBIL TYCO SOLAR ENERGY CORP (3)	SUNY BUFFALO (1)	UNIV MONTPELLIER 2 (1)

1991-1995	1996-2000	2001-2005
OSAKA UNIV (1)	NATL RENEWABLE ENERGY LAB (2)	CHINESE ACAD SCI (2)
CNRS (2)	STATE UNIV (1)	NATL RENEWABLE ENERGY LAB (2)
NASA (2)	UNIV NEW S WALES (1)	OSAKA UNIV (1)
NATL RENEWABLE ENERGY LAB (2)	UNIV STUTTGART (1)	NATL INST ADV IND SCI & TECHNOL (2)
AIN SHAMS UNIV (1)	CNRS (2)	JOHANNES KEPLER UNIV (1)
GEORGIA INST TECHNOL (2)	INST PHYS ELEKT (2)	HAHN MEITNER INST BERLIN GMBH (2)
HITACHI LTD (3)	FRAUNHOFER INST SOLAR ENERGY SYST (2)	UNIV NEW S WALES (1)
TOKYO INST TECHNOL (2)	ECN (2)	SAKA UNIV (1)
ECOLE POLYTECH (1)	HAHN MEITNER INST (2)	UNIV GRONINGEN (1)
FRAUNHOFER INST SOLAR ENERGY SYST (2)	RUSSIAN ACAD SCI (1)	UNIV LONDON IMPERIAL COLL SCI TECHNOL & MED (1)
MICHIGAN STATE UNIV (1)	TOKYO INST TECHNOL (2)	EINDHOVEN UNIV TECHNOL (1)
PENN STATE UNIV (1)	IMEC (2)	SWISS FED INST TECHNOL (2)
UNIV NEW S WALES (1)	SANDIA NATL LABS (2)	FRAUNHOFER INST SOLAR ENERGY SYST (2)
BANARAS HINDU UNIV (1)	TOYOTA TECHNOL INST (2)	UNIV POLITECN MADRID (1)
INDIAN INST TECHNOL (2)	GEORGIA INST TECHNOL (2)	UNIV STUTTGART (1)
JADAVPUR UNIV (1)	INDIAN INST TECHNOL (2)	UNIV CALIF BERKELEY (1)
SANYO ELECT CO LTD (3)	NAGOYA INST TECHNOL (1)	TOHOKU UNIV (1)
SOLAR ENERGY RES INST (2)	UNIV KONSTANZ (1)	NAGOYA INST TECHNOL (1)
SVERDRUP TECHNOL (2)	ECOLE POLYTECH (1)	UNIV UPPSALA (1)
UNIV DELAWARE (1)	KERNFORSCH ZENTRUM JULICH (2)	TOKYO INST TECHNOL (2)

Comment: The numbers in parentheses refer to organization types: (1) Universities, (2) Research Institutes, and (3) Firms.