The creation of knowledge: motivations, mechanisms, and determinants for cooperation

A literature review with empirical tests on spatial determinants of collaboration in knowledge creation in the Dutch water industry.

Pieter W. Heringa
Supervisor: Prof. dr. W.J.M. Heijman
ECH-80724
Wageningen University, June 2010

The image at the front page is created using Wordle (http://www.wordle.net).

Preface

It is generally accepted nowadays that knowledge is a quintessential condition for economic growth. However, our insights in the mechanisms behind that are far from complete, and so is the knowledge on the configuration of knowledge production and its determinants. This thesis points to some of the most important findings in the literature so far, and meanwhile it shows some blind spots that seem to deserve more research in the future.

I am very grateful those who assisted me in writing this thesis. First of all, my sincere thanks to all who were involved in my appointment as a junior researcher at the Rathenau Institute, and in particular to Peter van den Besselaar. I highly appreciate your confidence, appointing me before I finished my Masters, ánd providing me with the conditions to write this piece during working time.

I greatly benefitted from the daily supervision of Edwin Horlings and Mariëlle van der Zouwen. Thanks for the inspiration and motivation, for all the interesting talks we had and of course for reading and commenting my preliminary versions of this text. Although at a much lower frequency, the meetings with Wim van Vierssen were very important in the development of my ideas. I enjoyed our conversations in the car a lot, hopefully we can continue them in the future.

I cordially thank Wim Heijman for his help. I am most happy that you allowed me to write this thesis under your supervision, for arranging a place to work in Wageningen, for your extreme flexibility in supervising me (while I was not around that often) and in meeting the formal requirements. I also benefitted a lot from your idea to employ a gravity model to test whether we could predict the intensity of collaboration with distance. I am very glad that we can continue our cooperation by writing some articles together. Martijn van der Heide I would like to thank for acting as a second examiner.

The advantage of working at three different organisations is that you get to meet a lot of different people. I would like to thank my colleagues at KWR, Wageningen University and the Rathenau Institute for socializing. A special word of thanks to my roommates Tjerk and Chris. Hopefully we will remain sharing rooms for four more years.

This thesis has come to an end, but my PhD project has just started. I am looking forward to investigating the topics discussed here in more detail in the coming years.

Summary

Before asking why knowledge is generated, it is worthwhile to first ask oneself what knowledge actually is. Many authors distinguish between knowledge and information. Information only turns into knowledge when users are capable of making sense of the information. Knowledge that is no information (i.e. does not consist of well codified messages) is sometimes referred to as know-how. Another commonly used difference is between articulated and tacit knowledge. Tacit knowledge is knowledge that is difficult to codify, and hence only acquired through practice. There are however processes to turn tacit knowledge into explicit knowledge and vice versa. Depending on the context, many other categorisations of knowledge can be made.
One of the first models analysing the relation between knowledge and the production of innovations was the 'science-push model'. However, this model is a oversimplification of the process in reality. The causality is more complicated than the (linear and one-way) model suggests. Because of the large uncertainties in research, it is often difficult to attribute inputs to a specific output. The scope in time and space of innovations also make it difficult to judge the impact of research for innovations.

At different scale levels there are different reasons for the production of knowledge. At a macro (often national) level, it is rather recently shown that economic growth rates are dependent on the accumulation of knowledge. The production of knowledge generates positive externalities, and because of that it is beneficial at macrolevel to promote the creation of knowledge. This can take place via different mechanisms, such as increases in the stock of common (public) knowledge, the training of skilled labour, or the creation of new instrumentation or of firms. Knowledge may also be very relevant for growth at international level, for example if it is embodied in traded goods. Although it is difficult to prove the relation between the level of knowledge and economic growth empirically, it can be shown that an (inter)national spillover effect of knowledge is very likely to exist.
Not only at the macro level, but also at the micro level there are obvious benefits from investments in research and knowledge creation, although the private returns to R&D are strongly dependent on the institutional environment (e.g. protection of intellectual property). Yet, markets fail to coordinate the production of knowledge effectively, because the perfect protection of information is impossible, and part of the benefits hence accrues to agents not involved in the knowledge creation process. There are more reasons why private firms are not likely to invest in R&D to optimal level from a social point of view, for example the fact that the failure rate of R&D projects is very high and that many small firm lack capacity to develop technologies by themselves. However, there is no agreement so far on how the government could stimulate the development of knowledge. This is mainly due to a lack of a good knowledge-based theory of the firm. Also non-commercial (e.g. academic) organisations do have benefits from carrying out research. One of the incentives the government has to bring the level of knowledge to its social optimum, is by producing the knowledge more or less 'in-house'. A variant is stimulating academia to produce knowledge that is relevant for the public.
Individuals in organisations also have incentives to create knowledge. Apart from an intrinsic motivation, they are stimulated by the incentive system of the organisation. In science, this is usually built on reputation and prestige, in private organisations it revolves around monetary rewards.
At industry level, very little is known about incentive systems. It might be worthwhile to make this operational as a market, and investigate the underlying mechanisms of the 'transactions' at this market.

Apart from the determinants of producing knowledge as such, one may also look at the factors that determine the content of this knowledge: who or what sets the research agenda?
In publicly funded research, several approaches and models are employed to study the priority setting process. Economic models essentially are either a cost-benefit approach (formulating a preference function for e.g. social, economic, environmental goals, and determining what research contributes to that goal), an interest group approach (a political market, where collective goods are demanded in return for support) , or a supply and demand approach (where expected pay-offs and latent demand are brought together by socio-economic and politico-bureaucratic structures). The political science approach also views the phenomenon as a matter of supply and demand, but concentrates on asymmetry of information between those who carry out the research and those who govern and fund it. In the philosophy of science approach, the concept of enlightened democracy plays a central role. In this view, decisions on funding should be made by a group of people, tutored by scientific experts. The approaches in systems theory (userbased model, institutional model and political model) are basically variants of the models described above, put in a framework of systems theory.
In industrial research, there are different mechanisms for agenda setting. Again there are quite some approaches to study them. Benefit measurement methods are similar to cost-benefit approaches. Mathematical programming methods optimize an objective function under certain constraints. Simulation models do more or less the same, but in a more sophisticated manner, using probability distributions for all uncertain elements of a research programme. Decision and game theory focus on uncertain future events and reaction from the environment of the firm. Heuristic modelling concentrates on acceptable rather than optimal solutions. In cognitive emulation approaches, it is tried to model the actual decision-making process in an organisation, building on past experiences.

Individual organisations are rarely able to innovate (or conduct research) independently. There is hence a lot of attention for collaboration in research and knowledge production. Collaboration can take many forms. Both public and private partners can be involved. The cooperation can be formal or informal. It is dependent on the content of the knowledge whether or not (and what) partners are involved in knowledge creation. Firms have two reasons to enter into a research partnership: overcoming internal resource constraints, and reducing the risks associated with conducting research. The benefits of research are assessed in various studies. Results indicate higher productivity and broader scope of the in-house research, higher market share, and a reduction in costs. It should be noted however that cooperation is not always beneficial from a social point of view. Widespread cooperation might lead to a decrease in competition. This holds especially if the knowledge is (nearly) perfectly appropriable.
Several characteristics explain the propensity to collaborate. These include: size of the firm, size of the in-house R&D, educational level of the employees. There is also a relation with business success, but the direction of causality is difficult to determine in that case.
A large branch of literature investigates the spatial determinants of research cooperation. Here as well, several methods and concepts are developed. The idea of externalities suggests that some part of the knowledge generated will spill over to others, and more easily so if they are in the geographical proximity. This can be shown empirically. In the network approach, the interdependent relations between actors (resulting from information linkages between them) are central. The concept of innovative milieux focuses on the cultural characteristics of regional development. The idea of regional innovation revolves around two ideas: an institutional infrastructure and a production structure. The most important conditions for spatial dispersion are divisibility of the production process and transportability of the end product. Actual patterns can be described with a knowledge production function.

A possible model to test the spatial patterns of collaboration empirically, is the gravity model. It relates the size of two institutes and the geographical distance between them with the intensity of cooperation. The gravity model is applied in many different fields, but it can also be used to analyse patterns of co-authorship, an indicator of research cooperation. This is done using a dataset containing publications on the water cycle, published by organisations within the Netherlands, from 2006 to 2008. The results show that both the mass of the institutes (as proxied by the total number of publications) and the distance between them are highly significant explanatory variables. Expanding the model with variables to indicate whether both co-authoring institutes are from the same category (e.g. both universities) hardly adds to the explained variance of the model.

Table of Contents

1. What is knowledge and how does it lead to innovation?

1.1 What is knowledge?

The general stock of knowledge has increased considerably over the years, as a result of the efforts and investments of many various actors. This raises the question why knowledge is created. However, before that, one may first ask oneself what knowledge actually is. This issue received attention from many renowned thinkers. Clear consensus is lacking so far. Yet, making some distinctions is helpful in investigating the reasons for knowledge development. Many authors distinguish knowledge from information[1]. Individuals and organisations need skills and resources to gather and understand information. Information only turns into knowledge when the users are capable of making sense of the information. Information consists of knowledge, reduced and converted into (codified) messages that can easily be shared. If one knows the relevant syntactical rules, information can be shared without any loss (see e.g. Salter and Martin, 2001; Dasgupta and David, 1994; Kogut and Zander, 1992). The part of knowledge that is no information according to this definition is sometimes referred to as know-how. Know-how may be defined as the accumulated skills or expertise that allow to carry out a specific task efficiently. As the word 'accumulated' implies, know-how is learned and acquired over time (Kogut and Zander, 1992).

Another commonly used difference is between explicit (or articulated) and tacit knowledge. This idea stems from Polanyi (1966), who stated that people apparently know more than they can explain. Although the concept of tacit knowledge is widely used, it is rarely strictly defined. It refers to the fact that (parts of) knowledge sometimes cannot be codified and explicated in blueprints or instructions, and therefore can only be acquired trough practice (see e.g. Stephan, 1996; Kogut and Zander, 1992). Dasgupta and David (1994) use the term somewhat broader as knowledge that is for whatever reason not codified. According to them, this is not only a matter of epistemology, it might be determined by economic considerations, namely the costs and benefits of tacitness (and hence secrecy) versus codification. As Nonaka (1994) points out, knowledge can be converted from tacit to explicit and vice versa, and that both types can be shared with other individuals. These processes are called socialization (building and sharing tacit knowledge while it remains tacit), combination (building and sharing explicit knowledge while it remains explicit), internalization (converting explicit knowledge into tacit knowledge), and externalization (converting tacit knowledge into explicit knowledge).
Much more categorizations are possible however. Faulkner (1994) provides a synthesis of the contributions of various others, reproduced in table 1.1.

[1] However, Salter and Martin (2001) state that many innovation surveys use the concepts of information and knowledge interchangeably, and to a lot of firms the distinction between the two is rather academic.

Table 1.1. Various categorisations of knowledge in the literature. Source: Faulkner, 1994.

Authors	Fleck&Tierney (1991)	Winter (1987); Dosi (1988)	Vincenti (1991)	Gibbons& Johnston (1974)	Faulkner et al. (1994)
Perspective	Social shaping of technology	Evolutionary Economics	History of technology	Innovation studies	Innovation studies
Aim	Conceptualise knowledge in terms of sociocognitive structures that relate the content of knowledge to how it is distributed among individuals and organisations	Distinguish features of technological knowledge that impinge on the ease of technology transfer between firms	Develop an epistemology of engineering, in particular to relate categories of knowledge to knowledge-generating activities	Establish the extent and character of knowledge flows from public sector research into industrial innovation by investigating the full range of knowledge inputs to innovation	
Categories	Metaknowledge Milieu Contingent knowledge Tacit knowledge Informal knowledge Formal knowledge Instrumentalities	Tacit-articulated Teachable-nonteachable Articulate-Nonarticulate Observable-Nonobservable Complex-simple Elements of a system-Independent Specific-general	Knowledge categories: Fundamental design concepts Criteria and specifications Theoretical tools Quantitative data Practical considerations Design instrumentalities Knowledge-generating activities: Transfer from science Invention Theoretical engineering research Experimental engineering research Design practice Production Direct Trials	*Content of information:* Theories, laws, general principles Properties, composition, characteristics of materials and components. Operating principles or rules. Required specifications, technical limitations. Design-based information. Test procedures and techniques. Existence of equipment materials with particular properties. Existence of specialist facilities or services. Location of information.	*Broad knowledge types:* Knowledge of particular fields Technical information Skills Knowledge related to artifacts *Impact on company activities:* New product ideas Articulation of user needs Feedback on existing products Scouting for new applications Scanning the research frontier. Underpinning knowledge. Routine problem solving. New research equipment. New R&D procedures. Skills in experimentation and testing. New process technology. New production methods. Technical backup.

The studies mentioned in the table differ in framework and approach. The study of Fleck and Tierney (1991) is developed from the perspective of a social shaping of technology. It links the social context in which expertise is used and produced with its cognitive character. The sociocognitive structures they recognize are analyzed along two dimensions: the components of knowledge (as described in the table) and the distribution of knowledge among its carriers. Winter (1987) and Dosi (1988), both active in the field of evolutionary economics, separately came up with a categorization of technological knowledge. The two appear to be both overlapping and complementary. They are therefore presented together in the table. The next perspectives draw more on an empirical base. Vincenti (1991) based his work on (historical) case studies in aeronautics prior to 1950. The study by Gibbons and Johnston is based successful innovations spread over several sectors in the early 1970s. The work by Faulkner et al. (1994) rather builds on promising technologies from the 1980s and 1990s that had not yet brought significant innovations. The latter two studies are both based on interviews.

From the synthesis of the works above, Faulkner (1994) identifies five main types of knowledge, namely knowledge related to:

- *the natural world*, e.g. theories and knowledge of the properties of materials.
- *design practice*, consisting of design criteria, specifications (on the basis of technological considerations), concepts (operating principles and configurations) and the design itself.
- *experimental R&D*, e.g. experimental and test procedures, specific research competences, test data.
- *the final product*, operating performance of the product, performance of components and materials, production competence.
- *knowledge*, consisting of things unknown the people developing a product, and yet are necessary for the development.

Grant (1996) points to the importance of common knowledge. Knowledge cannot be integrated if there is no shared common knowledge. Whereas ideas form in the minds of individuals, interaction often plays a crucial role in further developing those ideas (see also Nonaka, 1994). Once individuals share a certain amount of common knowledge, this enables the sharing of knowledge that is not common to all of them. One can think of several types of common knowledge, e.g. language, other symbolic communication (numeracy, basic software skills), within an organisation widely shared specific knowledge, shared meanings (common frameworks), etc. This is closely related to the idea of absorptive capacity, coined by Cohen and Levinthal (1990). It is quintessential for a firm to be able to recognize the value of new, external information, adopt it to its own situation, and apply it for its own (commercial) ends. This ability is known as absorptive capacity. In the next section, we will examine this process of applying knowledge for the generation of innovative outputs.

1.2 From knowledge to innovations

One of reasons for the creation of knowledge, is that it brings about innovations[2]. It is however not directly obvious how the one results in the other. The easiest model is a simple linear model, also known as the science-push model, shown in figure 1.1.

[2] For more on the motivations for generating knowledge, see chapter 2

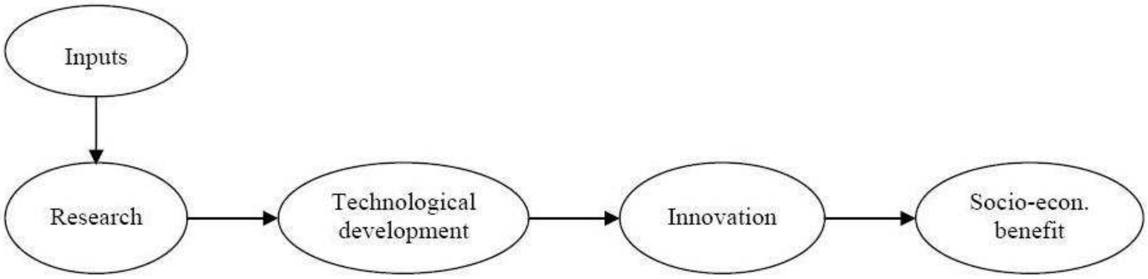

Figure 1.1. Simple linear model (science-push model) of relation between research and innovation (Martin and Tang, 2007).

However, reality is much more complicated than this model suggests. Firstly, the causation is more complicated than the arrows indicate, as there are important feedback loops present (e.g. the demand pull, which is at least as important as the science push). The introduction of the steam engine is a famous example where technology led science, rather than the reverse (see also Faulkner, 1994). Secondly, all elements mentioned have much more inputs than the ones mentioned here. For example, innovation is not only dependent on technological development, but also on non-technical inputs like markets, customer perception, organisational matters, etc. Thirdly, because of the large uncertainties in research and innovation processes, it is very difficult to attribute specific inputs to an output. Fourthly, the inputs often differ in geographical scope. Inputs to research are often at a national level, while effects of innovation are more often at an international level (country B may well benefit from the research that is carried out in country A). Comparable arguments apply for the time scale. As it may take decades before the potential benefits of new knowledge have materialized, it is extremely difficult to assess the impact of research. Bibliometric studies, measuring the citation of scientific publications in patent applications, found that the differences between fields are large, but that in some cases the time lag may be huge. Fifthly, in the simple linear model it is suggested that the output of research is new, codified, scientific knowledge. However, as is extensively shown in the literature, the newly produced knowledge may well be tacit in nature, embodied in people (Martin and Tang, 2007; Faulkner, 1994). The types of knowledge that are used in innovation are examined more closely in chapter four. We will first turn to the reasons that economic agents have to develop new knowledge.

2. Why do organisations develop knowledge?

The general stock of knowledge has increased considerably over the years, as a result of the efforts and investments of many different actors. Producing and disseminating knowledge is very costly, still even competitive firms are often engaged in these activities. This raises the question why knowledge is created. The answer to this question depends on the unit of analysis chosen. In the first section of this chapter we will look at the macro (or national) level. In section 2, we will pay attention to the micro (or firm) perspective. In section 3, the reasons for individual researchers are analysed. Section 4 pays attention to reasons at industry level to generate knowledge. Section 5 concludes with some empirical findings.

2.1 Macro level; relevance of knowledge for economic growth

Welfare and living standards vary strongly among (and within) countries. Economists often use income per capita (or a derived measure) to indicate how well off people are.[3] Income per capita is also distributed very unequally. Moreover, so is the rate of its growth. Traditionally, the accumulation of both physical and human capital are seen as main contributors to this growth. These are important production factors, and the concept of *total factor productivity* (TFP) is used to indicate the aggregate effectiveness of all production factors. Of course, differences in TFP also explain differences in growth rate. In other words, the amount and the effectiveness of the inputs of production together correlate with variation in the size of the output[4]. It is hence crucial to understand the determinants of TFP in order to explain the economic growth of nations. Technological change is seen as one of these major determinants.

Seminal works in this field include the studies by Solow in 1956 and 1957. They form the foundation of what is nowadays known as the neoclassical growth model. This model is based on several strong assumptions, the most important here being a constant rate of improvement in labour productivity. It leads to a striking conclusion: the growth rate of income per capita is negatively related to the capital-labour ratio. This implies that the more capital intensive a country becomes over time, the lower its growth rate. In the long run, capital intensity would stabilize, and output per worker grow at a constant rate, equal to the constant rate of technological progress. Most studies in the stream of research that emerged after the studies by Solow assumed technological change to be exogenously determined (Helpman, 2004).

Much later, Romer (1986) showed in a classical article that growth rates do not decrease over time. Using data from eleven countries for the period 1700-1978 (averaged per decade), he showed that all these countries were much more likely to experience increasing growth rates than declining ones. Romer concluded that models which assume a constant, exogenously determined rate of growth (like the one of Solow) do not suffice to explain economic trends in the long run. As an alternative, he came up with a model that assumes externalities in the accumulation of knowledge. The output of a firm is not only dependent on conventional inputs like labour and physical capital, but also on knowledge. In this model, knowledge is even seen as the basic form of capital. Knowledge is assumed to have three important characteristics. Firstly, it is produced by research technologies, with diminishing returns (i.e.

[3] Although well-being and welfare of people is constituted by much more factors than income alone, the latter is generally accepted as a rough measure of living standards (see e.g. Helpman, 2004).
[4] In fact, the causality of this relation is somewhat more complicated, as high growth rates may also induce investments in capital accumulation (see e.g. Helpman, 2004).

doubling the investment in research does not double the production rate of new knowledge). Secondly, the production of knowledge generates positive externalities. This idea rests on the much older work of Marshall (1920), who introduced the notion of increasing returns that are external to a firm, but internal to an industry. Other firms in an industry will profit if a firm makes investments in knowledge production: the new knowledge cannot be protected or kept secret completely. Thirdly, and technically speaking most important, the input of knowledge in the production of consumption goods is expected to have increasing marginal productivity. In other words, it is never optimal to stop generating knowledge, even if all other inputs are held constant. (Nevertheless, this model has an equilibrium, because of the first two properties mentioned earlier). This relates to the findings of Arrow (1962a), who assumed that the productivity of a firm is an increasing function, as new knowledge is generated in the process of production in a firm ("learning by doing").

Lucas (1988) uses similar notions. However, in his model, the externalities are in human capital rather than in knowledge. The empirical evidence is complicated, as externalities in general, and those in human capital in particular, are difficult to measure. For an operationalization of human capital as years of schooling, there is evidence that it plays a role in economic growth. The evidence on the existence of externalities in human capital is mixed (see Helpman, 2004 for an overview).

The simpler, earlier models predicted that countries with more labour available (or, the larger countries) would grow faster than others with less labour. Empirical evidence did not support these models (Helpman, 2004). Later models accounted better for the effects of scale. Some models introduced crowding effects, to dampen scale effects. Others, such as Young (1998), develop a model where productivity growth is caused by improvements in product quality. Yet, in larger economies more varieties of products emerge, so the larger amount of available resources has to be spread over a larger amount of varieties.

Although it is clear that conducting research contributes to economic growth, the exact mechanisms that bring about this growth are much less obvious. Salter and Martin (2001) distinguish six main types of benefits that stem from publicly funded research:

- Increase in the stock of common knowledge
- Training of skilled graduates
- Creation of new scientific instrumentation
- Formation of networks and stimulation of social interaction
- Increase in capacity of (scientific and technological) problem-solving
- Creation of new firms

Although the use of commonly available knowledge is not always straightforward[5], firms rely to some extent on past (publicly funded) research a source for new ideas or technological knowledge. In that way, past research contributes to the performance of firms and hence to technological growth. New graduates who enter the firm are also identified as a benefit to the firm. Although the entrance of students into industry is rarely a smooth process, and firms have to invest in training them, they bring very valuable skills and insights. Publicly funded research is also shown to have contributed to the development of new instruments and methodologies. It is however hard to evaluate this with e.g. innovation surveys, as industrial managers may have difficulty in pointing out what the contribution of publicly funded

[5] The use of information may be costly if effort is required to exploit it. This may for instance be the case for tacit (non-codified) knowledge. For more on this, see chapter 1 of this thesis.

research to instruments was. In addition, government-funded research forms an entry point to communities of expertise. It provides means to enter networks of research and development. Funding also stimulates the establishment of new networks. Nevertheless, the precise economic benefits are hard to measure. The increase in capacity of technological problem-solving that Salter and Martin (2001) distinguish, is largely overlapping with the increase in the stock of knowledge and the entrance of new graduates. Government-funded research provides a large pool of knowledge in many forms, which can assist in solving the complex technical problems that many firms face. Last but not least, the creation of new firms is sometimes mentioned as a benefit of publicly funded research. The empirical evidence on this point turns out to be mixed. While there is some evidence of a correlation between university research and firm generation for some sectors, for other sectors no significant relation could be found (Salter and Martin, 2001).

As the discussion above makes clear, scale issues are of paramount importance in developing an appropriate model of how knowledge generation affects growth. This is closely related to the effects of (international) trade. Research and development link the productivity levels of countries through various channels. First of all, trade increases the size effect. Access to larger markets raises the profitability of innovations. Second, there is a competition effect. This can have negative consequences, if competition causes a decrease in profits on innovation, and hence makes firms less likely to conduct research. However, it can also induce more R&D, if it enlarges the gap between technical leaders and followers (the leaders will put in more effort to stay ahead). The third effect can also go either way. Trade may change the factor prices. This may either make R&D relatively cheaper, or more expensive. Protection measures (trade barriers) may reinforce the effect. Fourth, trade prevents unnecessary duplications of R&D. If trade is completely open, firms will try to differentiate their products from all other products in the world, instead of from all other products in their home country. This results in a higher growth rate of the knowledge stock, and hence in stronger productivity growth. Fifth, trade opens up the access to the (intermediate) products of foreign suppliers. This increases the supply and variety of inputs for firms, which will raise factor productivity. Sixth, the stock of knowledge itself may be shared with other countries (international knowledge spillover). This is dependent on the content of the knowledge, as some knowledge is country specific (Helpman, 2004).

Whether this sort of spillovers indeed exist is a question of crucial importance to countries that lag behind in technological level. Since by far the largest part of the R&D in the world is carried out in a few industrialized countries, R&D may well be a strong factor in the divergence in development of richer and poorer countries, if knowledge does not (or only very slowly) flow from the countries that conduct the research to those that do not. However, even if the technical knowledge itself does not spill over to lagging countries, that does not imply that these countries do not benefit from the technical advances in more developed countries. Via trade it can benefit from new products invented and made elsewhere (Helpman, 2004). The international spillover effects of R&D are hence examined by various empirical studies. Coe and Helpman (1995) for example show a robust relation between the TFP of a country and the R&D capital stock of its trading partners for data on 21 OECD countries and Israel in the time period of 1971-1990, while controlling for the effect of domestic R&D. These findings are confirmed in a more recent study by Coe et al. (2009), using an updated (to 2004), expanded (to 24 countries) and revised version of the dataset from Coe and Helpman (1995). This study also included a few institutional factors as explanatory variables: the ease of doing business, the quality of tertiary education, the strength of intellectual property rights, and whether the origins of the national legal system are in French, German, English or

Scandinavian law. For this analysis, the countries were ranked and divided into three groups per variable (high, middle and low level in the variable under consideration). It turns out that countries where it is relatively easy to do business, profit more from the R&D carried out in the own country, from international R&D spillovers and from own investment in human capital formation. The same holds for the quality of tertiary education. The variance in protection of intellectual property rights turned out to be small between countries, and it increases over time in all countries that were compared. Finally, it appears that countries with a legal system that originates in French or (to a lesser degree) Scandinavian law, profit less from own efforts and spillovers in R&D than other countries (Helpman, 2004).

As has become clear from the arguments above, there are not only social returns to investments in R&D and knowledge, but also private ones. These (at least partially) accrue to the investing firm. Therefore, firms have their own incentives to invest in accumulating knowledge. Those incentives are at the core of the next section.

2.2 Micro level: relevance of knowledge for firms and other organisations

Essentially, the benefits of investment in R&D for firms are more or less as those for countries, they just apply at a different level. Next to his work about the national level, Romer (1990) also wrote a classical article on the perspective of the firm. In his model, firms invest in R&D to develop new products. They protect the production technology (blueprints to produce the new product) with patents. This yields them monopoly powers, and hence additional profits. These profits then form the incentive to invest in R&D. The private returns to R&D thus depend to a large extent on institutional features, such as the coverage of the patent protection, the efficacy of legal protection, etc (Helpman, 2004). Yet, no system can provide complete protection. During the process of invention, some useful information will be generated that does not end up in the blueprints, and some of it will be accessible for other organisations as well. To give an example, the description of a patent used for the registration gives important information on the technology used. Jaffe and Trajtenberg (2002) have shown that this is an important way for the diffusion of knowledge. It follows logically that the stock of knowledge that is commonly available to potential innovators is a function of R&D activities in the past. The more R&D was carried out, the larger the common pool of knowledge, and the easier (and hence cheaper) it is to conduct new R&D. On the other hand, the cheaper it is to carry out R&D, the more new products will be provided, and the smaller the profits will be. One of the main assumptions of the model of Romer (1990) is that technological changes emerge for a large part because of intentional reactions to market incentives. Based on this premise, technological change is endogenized in the economy. Romer shows that under the conditions of his model, economies with higher saving rates will grow faster, as they will spend more money on innovative activities.

It is thus clear that firms have obvious advantages from the production of knowledge. Some authors even state that knowledge is the critical production factor of firms. All human productivity is dependent on knowledge, and machines (physical) are basically embodiments of knowledge (see e.g. Grant, 1996). The production of goods and services usually requires different types of knowledge from various specialists. Markets are not able to coordinate this knowledge, because of a fundamental paradox in information exchange pointed out by Arrow (1962b): the value of information cannot be determined by a buyer until it is known what the information entails. Once he knows what the information contains, it is already disclosed to him and there is no reason to pay for it anymore. Firms however can integrate the knowledge of their specialists. They avoid the involved transaction costs of knowledge exchange. The stream of literature that adopts this view on the way firms work is therefore somehow similar to the much better known transactions cost theory of the firm (as described by e.g. Williamson) and is also known as the knowledge-based theory of the firm (see e.g. Grant, 1996).

There are more reasons why markets fail to coordinate the production of knowledge. Invention and knowledge production are risky activities. It can never be perfectly predicted from the inputs what the expected outputs are, or whether the outcome will provide a solution to the perceived need. And even if the outcomes are technologically successful, they may fail to become a commercial success (Audretsch et al., 2002). Markets would therefore only be able to allocate resources to knowledge production in an efficient way if there would be some sort of insurance against this sort of risks. However, insuring against failures in new product development would strongly reduce the incentives to proceed in this process. An (imperfect) approach to circumvent this is to have large firms conduct many research projects, all relatively small compared to the size of the firm. The risk is spread and the firm acts more or less as its own insurer (Arrow, 1962b). More importantly, information has properties of a public good[6]. To use economists' terms, it is non-rival and non-excludable. Non-rivalry refers to the fact that information is not depleted if it is shared with others. Non-excludability means that once information is generally available, it is hardly possible to exclude someone from use; it is difficult to appropriate the benefits of the results (see e.g. Stephan, 1996; Dasgupta and David, 1994). In short, there are three reasons why private firms are not likely to invest as much in R&D as would be optimal from a social point of view:

- The chance of failure of R&D projects is high
- Especially small firms may not have the capacity to develop a technology; cooperation with others can be costly
- The appropriability of the outcomes is small, i.e. the profits of the findings accrue to the firm only to a small extent.

Governments have several mechanisms at their disposal to overcome these problems, and prevent underinvestment in R&D. Basically, the remedies are:

- Providing incentives to stimulate the private production of knowledge.
- "In-house" production of knowledge by the government

The former can be done in various ways. Common policies include: patent laws, tax incentives, subsidies and the stimulation of research collaborations. Patents provide protection for the inventor of a new technology. In fact, a patent grants a partial monopoly. This increases the private returns to investments in R&D. A so-called tax credit decreases the private costs of a firm, by giving tax reductions if the firm is involved in knowledge production. Subsidies work in a similar way. Stimulating research partnerships may help overcome aspects of market failure. It reduces the risks of the participants by enlarging the common knowledge stock, and possibly also by the emergence of technology standards (increasing the chance of commercial success). Moreover, if it prevents redundant research, collaboration may yield cost savings and a decrease in the time to market (Audretsch et al., 2002).

[6] Knowledge is only a public good if it is well codified. Some authors therefore draw a distinction between knowledge, defined as the output of research, and information, the codified part of knowledge (see e.g. Dasgupta and David, 1994). For more on this, the reader is referred to chapter 1 of this thesis.

Theoretically, a government could also steer on the production of knowledge by stimulating specific forms of governance. However, the views on the effects of a specific governance form on the production of knowledge are contradictory. Some state that a hierarchy (a governance form based on the direction of authority) essentially avoids knowledge transfer. Others say that the very same hierarchy facilitates knowledge transfer. What is obviously lacking is a sound theory of alignment: a theory that indicates under what circumstances hierarchies prevail over markets or the other way around, or where a form in between these two extremes is preferred. In short, a good knowledge-based theory is not yet developed, let alone that there is consensus on the underlying logic (Nickerson and Zenger, 2004).

So far, the arguments presented here focus entirely on the perspective of the firm. However, there are much more organisations (such as universities) that produce knowledge and are much less prone to competitive incentives than firms are (see e.g. Romer, 1990). And although it is very difficult to infer a reliable, empirical relationship between technological change and economic growth (let alone an analysis of the process linking them), several studies found a large and positive contribution of universities or publicly funded research to economic growth (see Salter and Martin, 2001).

Roughly speaking, one could term the knowledge creation and application by firms 'technology' and the knowledge production by non-commercial institutes (such as academic organisations) 'science'. In a sense, many non-commercial institutes are a kind of in-house knowledge creation by the government. Although a lot of the institutes are formally speaking no governmental organisation, they are publicly funded. Depending on the science policy and the institutional infrastructure in a country, the government may steer in the programming of the institutes. According to Dasgupta and David (1994) there are three important differences between these realms of science and technology, namely the nature of goals accepted as legitimate, the norms of behaviour (especially where sharing and diffusion of knowledge is concerned) and the features of its rewards systems. Faulkner (1994) adds that despite the fact that science and technology seem to interact and overlap increasingly, the institutional separation has become sharper and more distinguished[7]. The quintessence of the differences seems to be in the incentives the systems provide to disclose knowledge. Researchers in science have to cope with priority rights (i.e. they need to be the first to publish a certain finding to get rewarded and will hence reveal the content of their findings as quickly as possible), while researchers in technology deal with proprietary rights (such as patents) and although this sometimes involves disclosing of the findings, there is no incentive to this quickly (see e.g. Stephan, 1996).

However, also on this point there is no consensus so far in the literature. There are two important strands. The first is the one just described (joined by e.g. Dasgupta, David, Stephan). They build on information theory and see the informational properties of science as a tool to analyse the pay-offs of public research. The second strand focuses much more on the importance of tacit knowledge[8], also in science. Tacit knowledge is difficult to transfer to others (which sheds different light on the idea that knowledge is a non-rival and non-exclusive good). This also implies that, in order to assess the economic benefits of public research, one should focus on the learning capabilities generated by investments in such research (see e.g. Rosenberg, 1990; see Salter and Martin, 2001 for further elaboration of the two strands of research).

[7] There is an extensive discussion in the literature on the differences and similarities of commercial and academic systems of research. This goes beyond the scope of this thesis. One could read e.g. Owen-Smith (2003).

[8] For more on tacit knowledge and other characteristics of knowledge, see chapter 1.

2.3 Incentives for the individual researcher

In the previous section attention was paid to the incentives firms and organisations have to invest in the production of knowledge. Yet, knowledge creation is an individual activity[9] (Grant, 1996). As there is quite some difference between the incentives for individual researchers in academia ('science') and for those in firms ('technology'), they will be discussed separately.

2.3.1 Incentives for individuals in science

The incentive regime for individual scholars in academia (loosely speaking, one may consider academia to be the place where science is carried out) consists of a competitive system with both monetary and non-monetary awards. The incentive system for research workers is usually a relatively subtle one. The outcomes of research are uncertain, cumulative, and often the result of team work (Cockburn et al., 1999). The system is fundamentally built on reputation building. The reputation of an individual for his contributions acknowledged by reference groups (others in the academic community) is the 'currency' of this reward structure. An important award is the recognition of being the first to communicate new findings. This point was first brought up by Merton in the late 1950's, who argued that the goal of scientists is to establish the priority of discovery (e.g. Merton, 1957). There is no award for being the second or third (note that this is in line with the desirable outcome for society, as there is no social value added for the second or third time the same invention is made). The recognition for being first can come in several forms (e.g. eponymy or giving the name of the researcher to a discovery, prizes, publications[10]). This system of performance-based reward sets up a race for scientific discovery. It ensures that a lot of effort is exerted to achieve the desired outcomes (new findings), while the effort itself need not be monitored. However, there is one obvious problem to this system of priority: if the unsuccessful researcher (in terms of first discoveries) would not receive anything at all, all risks would be put at the individual. As this individual will presumably be risk-averse, this system cannot be efficient (as people would be strongly hesitant to enter such a system at all). A researcher must therefore be paid a sort of flat rate, regardless of achievements. This is done in the monetary reward, that usually at least consists of a certain fixed amount, and sometimes bonuses coupled with achievements in the priority system (Dasgupta and David, 1994; Stephan, 1996). Apart from a personal monetary reward, there can also be rewards in the research budget, e.g. by leaving the decision power on part of the budget to high-performing research groups (e.g. Cockburn et al., 1999).
The rule of priority, apart from constituting an efficient incentive system, serves three social goals: first, as it presses researchers to disclose and share their findings quickly, it increases the social value of knowledge, by reducing the risk that the knowledge will reside in people who lack the willingness or abilities to exploit it. Second, the disclosing enables peers from the academic community to test and evaluate the findings. This increases the reliability and hence the social value of the new knowledge. The third advantage stems from the same point. As the findings are quickly disclosed to peers, a cumulative system of information is built,

[9] In the so-called approach of 'organizational knowledge' it is a common assumption that knowledge is accumulated by organizations, learning from its members. Yet, taking the organization as unit of analysis does not reveal the mechanisms through which (organizational) knowledge is created by the interaction of individuals (Grant, 1996).
[10] For a mathematical elaboration on the specific pay-offs of incentives like patents and prizes, see Wright (1983). For an empirical analysis of patenting in academic institutes, one is referred to Carayol (2004).

where people use the findings of colleagues to produce more new knowledge, and claim priority findings of their own (Dasgupta and David, 1994).

2.3.2 Incentives for individuals in technology

An important tradeoff between academia and private firms is the one of creative control versus focus. Whereas individual scientists in academia have rights to decide on what projects to select and what methodologies to use, in firms these decisions stay to a large extent at the level of the owner or manager of the firm (Aghion et al., 2005). Private firms tend to reward employees for subordinating private goals to the greater mission of the organization (Howitt, 2003). Both systems have advantages and disadvantages. Scientists presumably value their creative control, and will demand a monetary premium to give up on this freedom. Stern (2004) provides empirical support for this. He used multiple job offers to recent PhDs in biology, accounting for differences in the abilities of the scientists. This research indeed shows that wages are much lower in jobs where the candidates have some control on their individual research agenda or are stimulated to publish their findings. The advantage for academia is hence that they can hire their employees relatively cheaply; the disadvantage (from a social point of view) is that scientists will follow their own interests, which may not be the projects with the highest economic values. Firms on the other hand may have to pay their researchers more, but they can direct them to the projects with the most promising returns for the firm (Aghion et al., 2005).

This is not to say, however, that wages are the only incentives available to stimulate researcher in forms to exert full effort. Given that the outcomes of research are not only uncertain and cumulative, but also the result of group work, it is difficult from the perspective of the firm management to provide individual incentives. This can be overcome for example by giving rewards by changes in the research budget, adding a ‘bonus’ to the group budget and leave it to the group how this will be spent. Firms which also conduct basic research have yet another incentive at hand. They can encourage their employees to publish in scientific journals, attend conferences, or more in general to be an active member of the ‘public’ scientific community. This can be done for instance by basing personal promotion on ranking or standing in the hierarchy of science. Encouraging the involvement in public science may be costly to the firm (it inhibits intellectual property protection and may encourage employees to leave more applied activities), on the other hand ensures the connection of the firm to leading research developments in science. This may prove a great advantage, especially in sectors where technical advances are important for the competitive position (Cockburn et al., 1999). Research productivity in drug discovery for instance is argued to be positively related to the practice of promoting employees in a firm on the basis of their scientific standing (Henderson and Cockburn, 1994).

2.4 The meso level: incentives in industries

Most studies on the production of knowledge focus either on the national level or on the firm level, and those concerning innovation theories usually take the firm level. In that case, the firm is taken as exogenous, and its performance in generating technological changes is seen as endogenous. This is very obvious in the literature on so-called knowledge production functions, which describe a relation between the amount of innovative output on the one side, and inputs consisting of firm R&D and public research on the other side.[11] In the most basic form, such a knowledge production function presupposes the exogenous existence of the firm and its engagement in generating innovative activities. However, it turns out that empirical assessments of a knowledge production function give the strongest results at higher levels of

[11] For further elaborations on knowledge production functions, see chapter 4.4.

aggregation. It is therefore remarkable that, except for some case studies, hardly any research is conducted at a meso level. An exception is the work by Audretsch (1995). He finds that industries where R&D is an important input (such as the computer industry, pharmaceutics and instruments) are high in innovative output. On the other hand, industries with relatively very little R&D (wood products, paper) hardly have any innovative output. We know hence that the idea of a knowledge production function empirically makes sense at the industry level. However, what the underlying mechanisms and incentives of this production function are remains unclear. The exact functional links between knowledge sources and the resulting innovative output are unknown (Audretsch and Stephan, 1999). In the business literature, some ideas are developed on how firms in an industry can cope with competitive forces. A famous example is the article by Porter (2008; building on earlier work). However, the focus there remains on the perspective of the firm. In the few words that Porter mentions about the relevance of innovation, he just states that advanced technology is not enough to make an industry structurally attractive, as there are low-tech, bulk product industries with a far better *profitability for the firm* than in some very advanced high-tech industries.

Although some insights are gathered on collaboration patterns and determinants of knowledge production in specific industries[12], the broader picture of why knowledge is gathered in an industry and how this is done remains ambiguous at best. To improve this, it could be helpful to define industries as a market where transactions take place (information is exchanged). One could investigate the determinants of the transaction patterns (who exchanges information with whom, and why). It is likely that at least social (e.g. sharing attitude or expertise, speaking the same language, literally or e.g. technical 'language') and spatial determinants (being in the same area) play a role. An obvious problem in this way of making knowledge production operational at industry level however is that externalities (knowledge spillovers) play an important role in the diffusion of knowledge at industry level (see e.g. Audretsch and Stefan, 1999). Externalities are hard to define as a transaction, since an externality, by its very definition, is not transferred on purpose to the economic agent(s) that it affects. On the other hand, the metaphor of the transaction might shed more light on the awkward phenomenon that a lot of knowledge spills over to parties that did not invest in the production of this knowledge.

2.5 Empirical findings

Many scholars have tried to investigate the benefits from research (as sketched out above) empirically. Martin and Tang (2007) provide a very convenient overview of the findings of many studies on the benefits of (mainly public) research. There is an overwhelming amount of research on this topic, only a small selection of outcomes is given below. Griliches (1995) combines the results of earlier studies for a period of thirty years, and finds rates of return to public R&D between 20 and 50%. Mansfield (1991,1998) asked R&D managers in the United States about the relevance of academic research for the (product and process) innovations in their firms. In 1991, Mansfield found that about 10% of the innovations could not have taken place (or only with significant delay) without academic research. In his later study he found that both the importance of academic research was growing, and the time lag between basic research and its application in industry was decreasing (Mansfield, 1998).

Verspagen (1992) for example shows that technology is of great importance for the growth at a macro-level, although he also points out that the rate of technological progress is stronger related to the initial level of labour productivity than to factors like efforts in R&D and

[12] For more on this, see chapter 5.

investments. This is explained by the effect of knowledge spillovers and the capacity to acquire those.

The six channels mentioned in section 2.1 through which knowledge is exploited to the benefit of society, are all tested empirically. According to several studies, the transfer of skilled researchers and graduates is the most important channel for firms to acquire benefits from public research (see e.g. Roessner et al., 1998; Zellner, 2003). Zellner (2003) also specifically shows the benefits of the (relatively small) flow of scientists to the private sector. Related to this point is the increase (and enhancement) of problem-solving capacity. Publicly funded research yields large pools of resources for solving complex technical problems. E.g. Zellner (2003) showed that scientists who migrate from science, rather than bringing the latest theoretical insights, they bring in knowledge elements required for complex problem solving to the commercial sector. The third type of exploitation, an increase in the stock of knowledge of firms, is shown to exist indeed. Firms draw to some extent on new scientific ideas, and public research stimulates the private R&D of firms (Narin et al., 1997; Nelson and Rosenberg, 1994). It should be stressed however, that firms need to invest considerable amounts to be able to absorb the knowledge that is generated by the public sector (the so-called absorptive capacity of firms, see Cohen and Levinthal, 1989). There may also be significant time lags (up to decades) before scientific insights lead to commercial innovations (Martin and Tang, 2007). Large European firms are shown to rely on scientific publication as the primary source of public research (Arundel et al., 1995). A fourth channel of exploitation consists of new scientific instruments and methodologies. The effect of this is rarely empirically examined (Martin and Tang, 2007). However, e.g. Arundel et al. (1995) shows that large firms in Europe mention new instrumentation as the second important source of benefits from publicly funded research. The development of networks in which both scientists and researchers from industry participate are a fifth channel. The precise economic benefits of such networks are difficult to indicate. Some authors (e.g. Cooke and Morgan, 1993) take the intensity of the networks itself as indicator of the functioning of a regional or national innovation system. Others, such as Arundel et al. (1995) show that large firms themselves indicate informal networks as an effective means of learning about new research developments. A last benefit from public research is that it contributes to the creation of new firms. The empirical evidence on this point is mixed. For some sectors (e.g. biotechnology, nanotechnology) it holds that the presence of a university can be advantageous for the emergence of new firms in a region. However, for other sectors this correlation is not statistically significant (Zucker and Derby 2002; Fontes, 2005, Bania et al., 1993).

3. What factors determine the research agenda?

Before studying the patterns of actual collaboration in knowledge creation, we may first want to look at the determinants of the research agenda. It has become clear in the earlier chapters that knowledge, in the many forms that it may take, is a *sine qua non* for development, both from the perspective of nations and of organisations. However, given the wide range of possible research topics, and the limited resources that can be devoted to it, one may wonder how selections are made and priorities are set for the topics of research. It is this question that is at the core of this chapter. Even given a particular social or scientific goal (or set of goals), how can one know if a specific research portfolio is more effective than any other? As the methods of priority setting differ quite strongly between public sector research and private research, the first section of this chapter discusses several models to assess the phenomenon in the public sector. The second section pays attention to models and methods in the private sector.

3.1. Priority setting in publicly funded research

A specific characteristic of publicly funded research is that those who provide the funding are a different party than those who carry out the research. This can be examined in terms of supply and demand – although not in their meaning commonly used by economists, where markets can be cleared with prices. Demand may be seen as the (research) interests of the users of the research outcomes, whereas these research outcomes themselves constitute the supply. A complicating matter is that this supply essentially consists of public goods. Decisions regarding the supply and demand of public goods can only be solved by political institutions, not by markets. Decisions need to be made concerning the "weight" of different public research interests (Dalrymple, 2006; Sarewitz and Pielke, 2007). Various frameworks are used to study the process of priority setting in publicly funded research. Four of them are introduced below: a view from economics, one from political science, one from the philosophy of science, and one from systems theory[13].

3.1.1. Economic approach

There are several models in economics for the provision of public goods in general. The 'efficiency model', the 'interest group model' and a 'supply-demand model' are discussed here.

The efficiency model states that the quantity and distribution of public goods is determined by cost-benefit considerations, and is therefore also known as benefit-cost model. This idea is among others based on the work on agriculture of Hayami and Ruttan (1971). Government agencies act as to maximize economic efficiency, and adjust to prices of products and inputs. However, assessing the costs and benefits of projects ex ante is a difficult process in general, and even more so in the case of research. Setting the priorities for publicly funded research at a national level complicates the tasks even further.

[13] The selection of models used and their sources is strongly inspired by the work of Dalrymple (2006).

As Stewart (1995) points out, applying appropriate cost-benefit models for publicly funded research at a national level would require:

- Formulating a preference function for trade-offs between (amongst others) social, economic and environmental objectives for research (and this function should hold for the whole community)
- It should be known what type of research contributes most to the selected set of objectives.

As an alternative for the cost-benefit approach, Stewart comes up with a systems theory approach. This is discussed in section 3.1.4.

The interest group model stems from the literature on economic regulation. It postulates that government agencies distribute public goods with the goal of maximizing political support for the government. A political market is assumed, where collective goods are demanded in return for political support. The two models have some overlap. Variables that explain an individual's demand for public goods belongs to both. Moreover, the efficiency model is essentially a special case of the interest group model, namely one where consumers are the most important interest group (Guttman, 1980).

De Janvry (1978) developed an economic framework for the supply and demand of innovations specifically. This was originally meant for application in an agricultural context, but can well be generalized to the public sector (see Dalrymple, 2006). The model is represented in figure 3.1. Actions and perceptions of individuals are represented by circles, those of institutions by boxes. The expected pay-offs (in the upper right corner) represent the net economic gains and losses that interest groups expect from the implementation of a specific set of latent innovations. Although the figure only indicates one-directional flows, it may well be that there are several feedback loops in reality (see also Dalrymple, 2006).

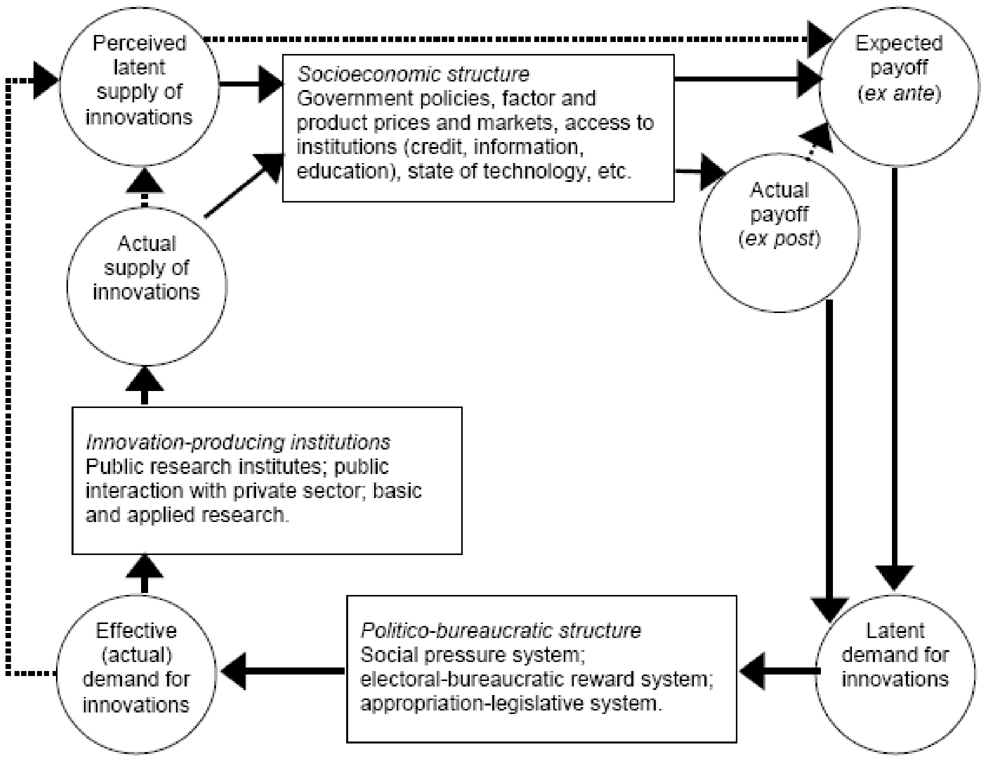

Figure 3.1. Generalized model of supply and demand for technological innovation in the public sector. Source: Dalrymple, 2006 (adapted from De Janvry, 1978).

3.1.2. Political science approach

The political science approach also departs from the point that the supply of research output is with a different party than the demand. However, here the focus is more on the asymmetry of information between those who carry out research (also called the performers) and those who govern and fund it (the patrons). Guston (2000) stipulates that this is the central problem of science policy. The patrons, often unaware of the actual (content of) the research, have to ensure that the research is worth their investment. The performers have to show that their performance is sufficient. Guston phrases this problem in the terminology of principal-agent theory. This is visualized in figure 3.2. This scheme shows the interactions between the funder, the researcher, and the end user.

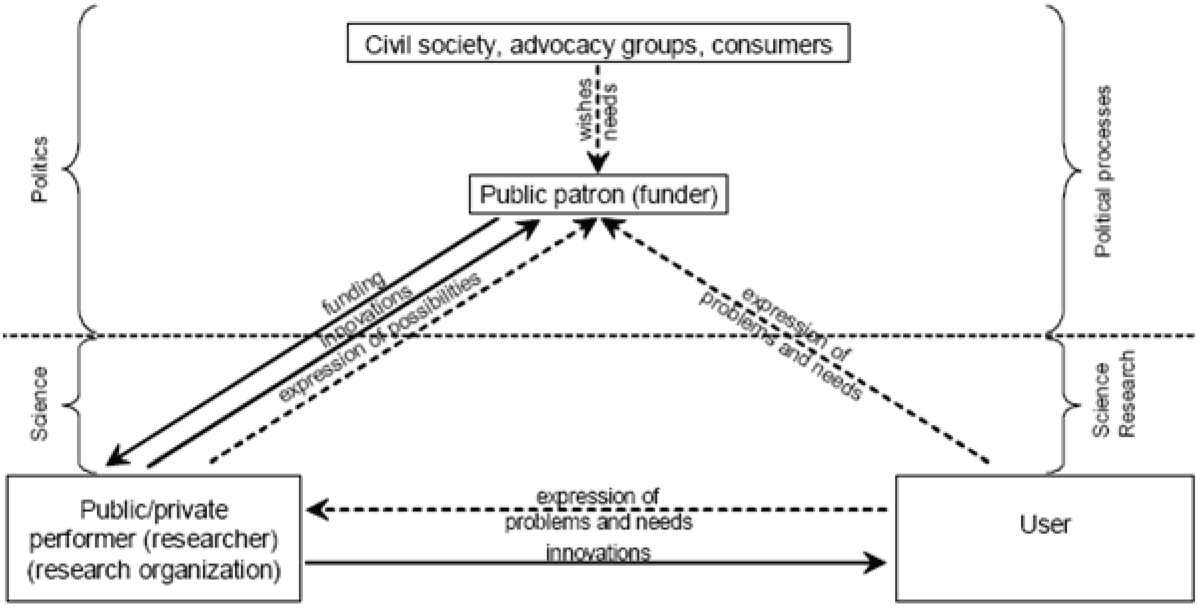

Figure 3.2. Principal-agent theory to publicly funded research. Source: Dalrymple, 2006 (based on Guston, 2000).

3.1.3. Philosophy of science approach

The problem of determining research problems of social relevance is also discussed in literature in the field of philosophy of science. Kitcher (2001) identifies three (somewhat differing) problems:

- The preferences of a large part of the public are ignored
- The untutored preferences of outsiders lead to an ignorance of problems with high significance
- If there were a consistent systematization of preferences widely shared by the public, strongly different priorities would be recommended.

Kitcher sketches a danger that may emerge from these problems: science areas that strongly depend on public funding may be shaped by governmental decisions. This may lead to a situation where projects with a more obvious social value may be preferred over those with a theoretical (but possibly strong epistemical) significance. Kitcher comes up with the concept of ‘enlightened democracy’ as a possible solution to this issue. Decisions on funding would be made by a group of people, tutored by scientific experts, and using inputs from all perspectives with a relatively broad base in society (Kitcher, 2001).

3.1.4. Systems theory approach

The focus of the systems theory approach is not so much on the actual choices for priority areas, but rather on the way these choices are made (Stewart, 1995). Three models can be derived from this perspective: a user-based model, an institutional model and a political model. Each model can be characterised by three questions: who makes the choices, with what incentives are they confronted, and at what level are the choices made? An overview of the characteristics is given in table 3.1.

Table 3.1. Typology of models in systems theory. Source: Stewart, 1995.

	User-based	Institutional	Political
Who chooses	Users (firms, citizens)	Scientists	Organised interests
Level of decision-making	Decentralised	Decentralised	Centralised or decentralised (pluralist)
Incentives to choosers	User needs	Rewards for research	Group benefits and costs.

The user-based model is closely related to the model of supply and demand described above. Its main focus is on the demand side of that model. Sometimes users can exert influence on the priorities of public research by (co-)financing the research (e.g. charity funds that finance health research). Yet many users lack adequate funding for this. It is possible to make the public research sector more aware of perceived user needs (e.g. by including citizens/ end users in advisory boards), however users usually are not in the position and capacity to judge the long term impact and importance of much of the research carried out with public funds (Stewart, 1995).

The institutional model focuses to a large extent on the internal way of working of public sector organisations. Researchers are influenced by their own interests and motivations, and research organisations differ in the ways to steer them. The incentives provided may well influence the topics of research. Priorities are set as much by this sort of institutional relationships as by formal selection processes. Explicitly analysing these institutional relations may reveal latent priority-setting mechanisms that have no clear rationale. As an example, many university lecturers spend a considerable part of their time on research. As changes in student numbers are usually reflected in number of staff employed, there is an apparent institutional link between student enrolment and research that cannot be explained from any formal selection mechanism. An institutional perspective also brings the influence of decision-makers within organisations to the fore. Reward structures for instance may very well (co-)determine the priority ranking of research topics (Stewart, 1995).

The political model in systems theory very closely resembles the model from political science discussed earlier. Scientists are reluctant to involve outsiders (like policy makers and lobby groups) to closely in their work. Yet on the other hand, these outsider groups may have considerable influence on feasibility of their work (through division of research funds, regulations, etc). However, the priorities of politics change all very frequently. It is also dependent on the personal preferences of the people with authority. This observation could be used in employing consultation processes in priority setting. By making sure that the input of consultations is as wide as possible (entails as much different ideas and opinions as possible), the more likely it is that the true societal preferences are reflected (Stewart, 1995).

3.2 Priority setting in industrial research

The problem of priority setting is somewhat different in industrial research. In this case, often within the firm itself decisions need to be made on how to allocate resources to research themes. This selection of projects needs consideration of several criteria, and many of them are subjective and/or uncertain, and subject to measurement problems. Often mentioned criteria include: probability of technical success, future profitability and potential market (Liberatore, 1988).

For quite some time already (since about the 1960s) models are developed which aim to assist in R&D project selection. These models need hence to be able to cope with subjective and uncertain information, and address a large variety of project types and measurement criteria, adopted to the specific context of the organization under consideration (Liberatore, 1988). Heidenberger and Stummer (1999) provide an extensive overview of the different categories of models available. They distinguish the following approaches:

- Benefit measurement methods
- Mathematical programming approaches
- Decision and game theory approaches
- Simulation models
- Heuristic models
- Cognitive emulation models

The benefit measurement methods are similar to the efficiency models in publicly funded research. Projects with the highest expected benefits are selected within overall budget constraints. Benefit measurement models can be divided in several sub-categories. *Comparative models* compare one proposal to (a subset of) other proposals. Depending on the number of comparsisons that need to be made, this can be time-consuming. There are two common ways to do so: the Q-sort approach is a psychometric method that orders a set of items in accordance with the individual sets of options of members of a decision group. The Analytical Hierarchy Process (AHP) allows to structure the comparisons, by employing a hierarchy of objectives at different levels (e.g. firm mission, firm objectives, technical strategies). By using pairwise comparisons a matrix is developed, which can be used to give the different objectives different relative weights (see also Liberatore, 1988). *Scoring approaches* score each project proposal on a limited number of criteria. The combined score per project is used as overall benefit measure. Multi-attribute utility analysis (where a utility function is maximized using several assessing criteria for each project) is included in the scoring approaches. *Traditional economic models* are meant to calculate the present value of expected costs and benefits of a project, or its option value and/or financial risk. This includes discounted cash-flow methods and options approaches. *Group decision techniques* combine the judgements of several experts from their respective fields. It is sometimes used as a tool to obtain data for more complicated models (Heidenberger and Stummer, 1999).

Mathematical modelling approaches essentially optimize an objective function under certain constraints. They frequently appear in the R&D project selection literature, but are rarely found in real applications. This might be due to a preference for simpler methods, and/or because of a lack of willingness to provide the detailed quantitative data required for mathematical models (Heidenberger and Stummer, 1999). Sub-methods include linear programming, non-linear programming, integer programming, goal programming, dynamic programming, stochastic programming and so-called fuzzy mathematical programming (see also Heidenberger and Stummer, 1999).

Both decision theory and game theory approaches explicitly take into account uncertain future events and reactions from the environment of a firm. The difference is that decision theory assumes changes in the environment to be exogenous to the actions of the firm, whereas game theory assumes rationally acting competitors (Heidenberger and Stummer, 1999).

Simulation models are more sophisticated than optimization techniques and can resemble the real world closer, yet they only provide the outcomes for one specific set of parameter values per simulation run. Popular are e.g. Monte Carlo simulations. They use the probability distributions of all stochastic elements of an R&D programme, to provide the overall probability distribution of the objectives and means of the programme (Heidenberger and Stummer, 1999).

Heuristic modelling is not focused on optimal, but rather on acceptable solutions. They are increasingly found in the literature on project selection. It is a possible solution if the problem under consideration contains many interactions between different elements of a model. This implies that it is extremely difficult to analyse all possible solutions. Heuristic modelling may save time in such cases (Heidenberger and Stummer, 1999).

In cognitive emulation approaches, it is tried to model the actual decision-making process in an organization, building on comparable circumstances in earlier processes. This can be done by statistical approaches (e.g. regressions or discriminant analysis). A second category are expert systems. An expert system is a computer program that is meant to replace the human inference process. It analyses all potential projects on certain criteria. A third variant is decision process analysis. This model represents a hierarchical organisation with several groups involved in project selection. It does not provide answers to the selection problem per se, but is meant to give insight in the managerial processes. These models reflect the coordination and information mechanisms between the different groups involved (Heidenberger and Stummer, 1999).

However, Heidenberger and Stummer (1999) also state that the discussed methods and models are mainly appropriate for later phases of the R&D process. According to them, for basic research the so-called ‘genius award’ (provide grants to researchers with proven capabilities) may yield results as good as using complex selection methods.

It should be stressed that neither in science, nor in technology any organisation sets its research agenda completely by itself. There are continuous feedback loops between different actors (from both science and technology, but also e.g. policy and consumers). Successful innovation is characterised by ongoing communication between producers and users of the knowledge required (see e.g. Sarewitz and Pielke, 2007). The next chapter will pay more attention to the patterns and determinants of such cooperation.

4. With whom is knowledge produced?

A dominant idea of the network theory of innovation holds that individual organisations are rarely able to innovate independently, and hence never innovate ‘in a vacuum’ (Freel, 2003). Network theorists stress that an increasing division of labour among organisations induced collaboration and interaction between them. Many firms concentrate on less core activities and competencies, using trade or cooperation to acquire other skills (Sternberg, 2000; Archibugi et al., 1999). The assumed benefits from collaboration mentioned in the literature revolve around two main themes: solve for an internal lack of specific skills or competences, or overcome internal resource constraints (see e.g. Freel, 2003; Tether 2002) and risk reduction (Tether 2002). Yet, the empirical evidence is mixed, both on the actual patterns of collaboration and on the perceived benefits. In section 1, we will look at actual collaboration patterns. In section 2, several frameworks are introduced that attempt to explain why organisations would collaborate. In section 3, the determinants and effects of the configuration of collaboration networks will be elaborated upon. In section 4, specific attention is paid to the role of spatial factors in the configuration of networks of cooperation.

4.1 What are the actual patterns of research collaboration?

Cooperation between organisations in the field of research is described from a range of different perspectives in the literature. There are hence quite some terminologies in use, often similar to each other or even overlapping, sometimes with differences in definitions. It seems therefore worthwhile to first develop a simple characterisation of research collaboration[14].

Broadly speaking, the partners involved in a collaboration can be public or private. Sometimes universities are considered a third category, since they can be owned either privately or publicly. Yet, as most universities receive a considerable part of their budget from public funding, it might be convenient to classify them as public (see e.g. Hagedoorn et al., 2000). Consequently, partnerships can either be public, private, or mixed.

The organisational structure of research collaboration can range from informal to formal. According to Link and Bauer (1989), 90% of the partnerships from a sample of cooperative-research active firms in the US manufacturing sector was informal. Firms do not only link up with other firms in informal agreements, but also with universities, where the universities often serve as a short-term research subcontractor, for specific research interests. More in general, very little is known about informal research partnerships. By their very nature, they are difficult to trace, let alone study in a systematic, quantitative way (Hagedoorn et al., 2000).

Of course, research collaboration can also take place in a formalized agreement. This can basically take either of two forms: an equity joint venture that focuses solely on research (and development), also known as research corporations, or a research joint venture (RJV), which is more of a contractual agreement between organisations.
A research corporation essentially is a separate firm, initiated by two or more firms that combine skills and resources in R&D. The newly established firm carries out R&D activities that are usually in line with the general research agenda of the parent organisations. Although this form of collaboration is rather popular, the stability is uncertain (Hagedoorn et al., 2000).

[14] I use the terms research collaboration and research cooperation interchangeably.

Somewhat older studies (Berg et al., 1982; Kogut, 1988) show that around half of all R&D related equity joint ventures do not meet expectations or are dismantled altogether.[15] In the other main category of formal research collaboration, RJVs, the risks are smaller, since no new organisational entity is created. It is hence much easier and less costly to end this form of partnership. The success of these contractual arrangements is strongly dependent on the actual commitment of the participating organisations.

The attention to patterns of collaboration in research might be high, firms indicate in surveys that a large part of their innovations is a result from their own (in-house) R&D efforts. Rothwell (1977) provides an overview of results of some major innovation studies in the 1960s and 1970s. He finds that the differences in the outcomes of these studies are rather small. About two third of the knowledge used by firms for an innovation process is said to stem from own R&D work. The second supplier consists of other firms, mainly users and suppliers, to a lesser extent also competitors. Between 5 and 20 percent is attributed to research from universities and government laboratories. Later research by e.g. Faulkner et al. (1994) confirms the findings that most of the knowledge inputs to innovative activities are considered to come from an internal source. More detailed outcomes of this study are provided in table 4.1.

Table 4.1. Impact of knowledge inputs on innovative activities, by source (in percentages of responses per question; actual number in parentheses). Source: Faulkner et al., 1994.

Activity	Source			
	Internal	Other companies	Public Sector	Non-response
Future innovations	57% (52.0)	33% (30.0)	10% (9.0)	(5)
Search activity				
Scouting for new applications	45% (13.5)	27% (8.0)	28% (8.5)	(2)
Scanning research frontier	30% (8.5)	18% (5.0)	52% (14.5)	(4)
Subtotal	37% (22.0)	22% (13.0)	40% (23.0)	(6)
Ongoing R&D				
Underpinning knowledge	40% (12.5)	2% (0.5)	58% (18.0)	(1)
Routine problem solving	88% (28.0)	9% (3.0)	3% (1.0)	(0)
Subtotal	65% (40.5)	6% (3.5)	30% (19.0)	(1)
Instrumentalities				
Research equipment	18% (5.5)	52% (15.5)	30% (9.0)	(2)
R&D procedures	47% (14.5)	24% (7.5)	29% (9.0)	(1)
Skills in experimentation and testing	55% (16.5)	17% (5.0)	28% (8.5)	(2)
Subtotal	40% (36.5)	31% (28.0)	29% (26.5)	(5)
Production	51% (23.5)	46% (21.0)	3% (1.5)	(18)
Technical backup	73% (14.5)	10% (2.0)	18% (3.5)	(12)
Overall total	51% (189.0)	27% (97.5)	22% (82.5)	(47)

Much earlier, Gibbons and Johnston (1974) already showed that it is also dependent on the content of knowledge to what extent the knowledge is acquired from external sources. They distinguish six types of content and show the respective importance of internal knowledge production, other companies and public sector research. The findings are given in table 4.2. Faulkner (1994) later comments that many researchers she interviewed indicated that tacit knowledge (skills) is to a large extent acquired in-house, and is of greater importance to

[15] For more on the advantages and drawbacks of research partnerships, the reader is referred to section 4.2.

innovation than more formalized forms of knowledge. She adds that further questioning however revealed that other companies and public sector research did partially contribute to the stock of tacit knowledge.

Table 4.2. Sources of knowledge for different types of content (percentages of information units; actual numbers in parentheses). Source: Gibbons and Johnston, 1974.

Content of Knowledge		Source		
	Info Units	Internal	Other companies	Public Sector Research
Theories, laws, general principles	8% (69)	52% (36)	16% (11)	32% (22)
Properties of materials and components	32% (270)	74% (200)	16% (42)	10% (28)
Design-based information, operating principles	24% (205)	81% (165)	15% (30)	5% (10)
Test procedures and techniques	10% (78)	80% (62)	12% (9)	9% (7)
Knowledge of knowledge	26% (217)	57% (124)	30% (66)	12% (27)
Total	100% (839)	70% (587)	19% (158)	12% (94)

Tether (2002) empirically examined the cooperation arrangements for innovation between firms and various types of external partners in the UK (although this is not strictly the same as cooperation in research, it is very closely related). Only projects with active participation are included in this research (i.e. pure contracting out without active participation is not regarded cooperation). Other firms from the same supply chain of a product (i.e. users and suppliers) are most often mentioned as partners in cooperation. More detailed findings on the type of partners firms collaborate with are summarized in table 4.3.

Table 4.3. Innovating firms with collaboration on innovation, by partner type. Source: Tether, 2002.

Type of partner	N	Amongst the innovating firms (%)
Any external partner	541	42
Supplier	283	22
Customers or clients	298	23
Competitors	188	15
Universities	207	16
Consultants (N=195) and private research institutes (N=51)	206	16
Others (government institutes and research and technology organisations)	206	16

Fritsch and Lukas (2001) investigate the propensity of German manufacturing enterprises to maintain R&D cooperation with customers, suppliers, other firms and public sector research institutes. They manage to identify a couple of variables that predict whether or not a firm will be involved in R&D cooperation. Generally speaking, firms engaged in R&D collaboration tend to be relatively large (in number of employees), have a relatively high share of R&D employees, employ a socalled ‘gatekeeper’(someone who monitors external developments that are relevant to the firm), and are relatively ambitious in product innovation (i.e. aimed at new product development rather than improvement of existing products). Moreover, they find that for suppliers, cooperation tends to be considered as a substitute for own research activities. As both Fritsch and Lukas (2001) and Tether (2002) point out, the motivations of firms to enter a research partnership are dependent on characteristics of the firm. It is to these motivations that we now turn.

4.2. Motivations, benefits and drawbacks of research collaboration

In innovation studies, it is recognised for a long time already that innovation to some extent is an interactive and hence distributed process. However, the specific forms of distribution, such as collaboration agreements and alliances, are looked into only rather recently. Relative little is known on the determinants of cooperation forms. Moreover, the character of the relationships between firms have become more complicated, as the boundaries of firms are increasingly fuzzy (Tether, 2002).

As stated before, there are broadly speaking two reasons for firms to enter into a research partnership: overcoming internal resource constraints (including knowledge constraints), or reducing the risks associated with innovation. These two may be interrelated (e.g. Tether, 2002). These main reasons can be disentangled into more specific possible (strategic) motives. This can include: gaining the technical abilities for horizontal diversification (new product lines), for vertical integration (i.e. integrate more steps of a product chain into the firm), or for leap-frog competition in the core business (i.e. leave competitors behind in technical position)[16]. Link and Bauer (1989) and Link (1990) empirically tested to what extent these motivations are indeed influencing a firm in its decision whether or not to enter a research collaboration agreement. It turns out that firms that face strong international competition consider the opportunities for horizontal diversification as a motivation for partnerships, whereas firms where the threat of international competition is much weaker see research collaboration as a possibility for vertical integration. They did not find empirical evidence for the factor of leap-frog competition, although others did (see Hagedoorn et al., 2000).

Research partnerships with universities are usually considered as a special case in the literature. Although the history of (informal) relationships between science and industry is relatively well documented, it is only in more recent literature that the motivations for these partnerships are investigated (Hagedoorn et al., 2000). Several researchers inquired what type of research partnerships would ask a public sector institute (usually a university or a government laboratory) to join (see e.g. Baldwin and Link, 1998; Leyden and Link, 1999). They find that only partnerships with a relatively large number of participants tend to invite a public institute. The underlying argument is that the appropriability of research is small in these large partnerships already, so a joining public institute does not induce new problems in that respect.[17] Motivations to ask a public institute to engage in collaboration are twofold: The first reason is the chance to access research and research results that are complementary to the research of the firm (or the existing partnership) itself. The second reason is the access to central members of the academic world. From the perspective of the university, it is suggested that the growing financial pressure forms an incentive to link up with partners in applied commercial research Hagedoorn et al., 2000).

[16] The factor of diversification and integration both also show the interrelatedness of dealing with resource constraints and risk reduction. To broaden the scope of a firm requires more internal resources, whereas it also implies a spreading of risks.
[17] For more on the appropriability of research, see section 2.2.

It is attempted in various studies to empirically show the benefits of research cooperation. Link and Bauer (1989) identified a positive correlation between the amount of cooperative R&D conducted by a firm, the market share of this firm, and the productivity of its own in-house R&D (i.e. the more cooperative R&D is carried out by a firm, the larger its market share and the more productive its R&D projects that are not shared with others). It is suggested that the latter effect derives from an enhanced absorptive capacity of the firm, resulting from its cooperative R&D projects.[18]

Moreover, several case studies suggest that research partnerships do contribute to a broadening of the scope of the own R&D of a firm (e.g. Scott 1996, see also Hagedoorn et al., 2000 for an overview). There also is some incidental evidence of increases in R&D efficiency, mainly due to a reduction in the duplications of research costs. However, it should be stressed that the empirical evidence of the last two points draws to a large extent on case studies. There may well be selection bias in these studies, as usually successful collaboration projects are selected as cases (Hagedoorn et al., 2000).

However, firms need a large in-house capacity to be able to reap rewards from collaboration. In order to recognize promising potential partners, evaluate their qualities, adapt the outcomes of joined projects, etc, one needs substantial 'own' knowledge and skills before one can enter a partnership. In other words, although collaboration brings more absorptive capacity to a firm, one needs a certain amount of it before starting collaboration (Freel, 2003; Dosi 1988).

One should be aware that cooperation in research is not always beneficial from a social point of view. Widespread cooperation might lead to a decrease in competition, or collusion. This was assessed theoretically rather early (in the mid 1980s). Studies stated that especially in concentrated product markets (i.e. markets with relatively few suppliers), collaboration improves the incentives for firms to conduct relatively inappropriable R&D. Strong competition on the product market on the other hand weakens the incentives to cooperate, as the benefits in such a market tend to shift the (expected) benefits of research from the producers to the consumers (d'Aspremont and Jacquemin, 1988; Katz, 1986). However, as Vonortas (1994) points out, these studies treated all R&D activities as one single category, or at best distinguish between (for the innovating firm) relatively appriopriable research (generic research) and relatively inappropriable research (the development part). According to Vonortas, this approach limits the applicability of the outcomes of these studies for three reasons. First, the output of R&D cannot be treated as a single product (knowledge). It may consist of a mix of uncodified, incomplete knowhow, and highly codified information[19]. Second, theoretical models generally assumed that when there are strong spillovers, firms conduct generic research only, whereas if there are hardly spillovers, they would only conduct very applied (development) research. In practice, most firms will do both (it is especially unlikely that a firm would not conduct any development research). Third, the older models would either predict that firms cooperate, or that they compete. In practice, many will compete with their products and yet cooperate for research purposes (probably in separate joint ventures).

Vonortas therefore proposes several extensions. In the model, generic research and development research are distinguished, where development outcomes are firm-specific and costly to transform for application elsewhere. All research can be conducted either in-house or together with others. Cooperation can take two forms: so-called secretariat RJVs, where the level of expenditures is determined together (ex ante), but the research is conducted independently; there are hence no changes in knowledge externalities (spillovers) compared with the situation where generic research is conducted completely independently. In the

[18] The reader is referred to section 1.1. for more on absorptive capacity.
[19] For more on this, see section 1.1.

second case, called operating RJVs, generic research is carried out together, and results are shared immediately. Framing the model as a game-theoretic problem where firms maximize their pay-offs (joined pay-offs in case of cooperation), he finds that both forms of cooperation are more efficient than non-cooperation *if* the generic knowledge created is a perfect public good. However, if the produced generic knowledge is perfectly appropriable (is a perfect private good), than both forms are less efficient than non-cooperation. Yet, the expected performance in a non-cooperative situation are also strongly dependent on the operating environment of the industry. This should be good for technological innovations. Not only should the spillover of knowledge be low, the technological opportunities need to be good for development research, and there should be good opportunities to increase the market share by price competition. In addition, Vonortas shows that operating RJVs are more efficient than non-cooperation in much more settings than secretariat RJVs.

It is now obvious that, despite some drawbacks and conditions, many firms have an obvious advantage in cooperating with other firms. One may wonder what factors influence the configuration of such networks (who collaborates with whom?). This question will be elaborated upon in the next section.

4.3 Determinants and configuration of cooperation

In the current era of globalization and modern information technologies, one might be tempted to think that organisations can initiate partnerships with any interested party in the world. This may be true to some extent (as there are often no regulations that would prohibit to do so). However, the very existence of Silicon Valley as a dense cluster of activities in modern information technology shows very clearly that a factor like physical distance is still of great importance in determining potential (research) partners. In this section it will be investigated what factors play a role in the configuration of networks among firms and other organisations.

First of all, it turns out that several characteristics of the firm itself can explain part of the propensity to collaborate. Factors that are often mentioned in the literature (and empirically tested), include: size of the firm (in employees or turnover), size of the in-house R&D (both in absolute and in relative numbers, the latter often being mentioned as “R&D intensity”), existence of a separate R&D department, and the educational level of the employees (see e.g. Arndt and Sternberg, 2000). To give an idea of the relevance of such aspects, table 4.4 reproduces some findings of the ERIS (European Regional Innovation Survey), a large survey conducted between 1995 and 1999 in eleven European regions. In this research, cooperation was defined as working together with other innovative actors (like suppliers, customers, service providers, competitors or research institutes), “*above and beyond normal business relations*” (Arndt and Stenberg, 2000, p. 470). However, these results are not undisputed. The relevance and impact of internal characteristics of the firm is assessed very differently in other studies. E.g. Oerlemans et al. (1998) suggest that sectoral differences may cause differences in factors like size and educational level. Together with other studies (e.g. Freel, 2003; Fritsch and Lukas 2001) this yields a somewhat diffuse picture.

Table 4.4. Characteristics of firms, disaggregated for cooperation and tested for significant differences. Source: Arndt and Sternberg, 2000.

Characteristics	Cooperating businesses (N=3166)	Non-cooperating businesses (N=448)	t-Value
Age (years)	34.8	30.4	1.89
Employees	146	54	7.48**
Turnover (Mio Euro)	24.0	7.5	5.99**
R&D-expenditures (% of turnover)	6.8	2.8	9.73**
R&D employees	8.8	1.9	4.84**
R&D employees (% of total employment)	7.8	3.9	7.86**
Employees with university degree (% of total employment)	12.5	7.4	7.41**
** Significant at p< 0.001			

It is also tried to infer a relation between cooperation and business success. Although a correlation between these two variables can be shown to exist, it is difficult to determine the direction of the causality (Arndt and Sternberg, 2000). For an example of the correlations that can be found, see table 4.5.

Table 4.5. Cooperation and business success. Source: Arndt and Sternberg, 2000.

Characteristics	Cooperating businesses	Non-cooperating businesses	t-Value
Share of turnover with new products (%)	27.2	17.1	4.60**
Annual employment change (%)	7.5	4.6	1.96*
Annual growth of turnover (%)	15.4	12.4	1.82
Ration of exports of turnover (%)	26.3	14.3	6.77**
* Significant at p=0.05; **significant at p<0.001.			

Research and development is increasingly an international phenomenon (see also section 4.4). Surveys revealed three main determinants of the internationalisation of R&D (see Meyer-Krahmer and Reger, 1999 for more elaboration):

- The early linkage of the R&D activities to a 'lead market' (i.e. conducting the R&D in continuous interaction with leading (potential) clients for the innovations).
- Coordination with scientific research. Scientific findings in that case drive the internationalisation of commercial research programmes.
- Strong linkages between production and R&D.

It turns out that not all factors are equally important in all industries. Meyer-Krahmer and Reger (1999) report a ranking of importance for three branches: pharmaceuticals, semiconductor technology and telecommunications. The results are reproduced in table 4.6.

Table 4.6. Relative importance of internationalisation factors in different industries.

Importance of R&D link to	Pharmaceuticals		Semiconductor technology		Telecommunications technology	
	Preclinical	Clinical research	Process technology	Product development	Hardware	Software
Lead market	Low	Very high	Low	Very high	Low	Very high
Science/research system	Very high	High	High	Low	High	Low
Production	Low	Low	High	Low	High	Low

However, the spatial dispersion of research is the result of a much more complicated set of factors. In the next section, these underlying factors of spatial patterns will be considered in more detail.

4.4 Spatial patterns

4.4.1 Theoretical approaches

There is a large branch of literature that documents the relevance of spatial dimensions for patterns of collaboration. In general, the reasons why distance matters to business configurations have changed over time. Earlier, the costs of capital (investment) determined competition patterns, and locational advantages could consist of for example local supply of cheap labour or high-quality infrastructure. Now, competitiveness is much more based on the costs and use of inputs. This requires continuous innovation (Porter, 1998). More specifically for R&D, the geographical dispersion has also changed over time. In the 1970s there seemed to be one world centre for each important field of technology (Western Europe for some specific fields such as chemistry, the United States for most others), and most large (technological) companies conducted their the largest part of their R&D in their country of origin. In the 1980s, an era began of increased internationalisation and decentralisation, and many new alliances emerged, also in technological fields. The trend of internationalisation continued in the 1990s, however, the trend of 'dislocation' was replaced for increasing concentration (also in space), focus and strategic emphasis. In particular for large multinationals that does not imply that the R&D is now again conducted in the country of origin. Rather, one or a few locations in the world are deliberately selected, because of their excellent conditions. The processes of production and R&D are decoupled (see Meyer-Krahmer and Reger, 1999).

Location choice received a lot of attention in recent years, with a lot of attention for the emergence of clusters. As a result, a full range of approaches and concepts revolving around this theme has developed, for instance:

- The concept of externalities
- The network approach
- The concept of innovative milieux
- The cluster approach
- The concept of regional innovation systems

These approaches have two ideas in common: first, all share the view that cooperation and interaction will affect innovation activities in a positive way. Second, there is agreement that geographical space plays a role in the processes of research and innovation. The firm is part of a structure (or network) with others (suppliers, customers, competitors, etc). These networks have two functions: they give pressure to the firm to remain innovative. Meanwhile, they also provide resources, necessary to innovate. Spatial proximity (i.e. being located at small physical distance of others in the network) can strengthen the ties of the network, as it facilitates and promotes face-to-face contacts (Arndt and Sternberg, 2000; Oerlemans et al., 2001).

The idea behind the concept of externalities dates back to the industrial district approach by Marshall. The use and interpretation of the concept has changed over time, nowadays regional scholars often refer to Marshall-Arrow-Römer (or MAR) externalities. Externalities are unintended side-effects from production (or occasionally consumption) processes; they are often external to the producing agent himself (i.e. they affect some other actor). An important externality in the production of knowledge are the so-called spillovers: some of the knowledge generated becomes available to other organisations too, even if it is not intendedly shared[20]. There is quite some empirical evidence that suggest that knowledge spillovers indeed exist. A striking result is for example that small firms appear to be the engine of innovative outputs, whereas they conduct relatively very little R&D themselves. Apparently they gather the knowledge somewhere else, and it is most likely that it spills over from other organisations (see Audretsch, 2003). MAR externalities are external to the producer, however they always remain inside the industry within certain geographic bounds (see e.g. Deidda et al., 2002). A competing theory, the Jacobs externalities, assumes that the most important spillovers rather occur across different industries, and hence advocates that organisations from different domains cluster together, rather than those with similar expertise. There have been many attempts to assess empirically whether MAR or Jacobs externalities are closer to reality. Two large empirical studies (Glaeser et al., 1992 and Feldman and Audretsch, 1999) tend more towards Jacobs externalities, but the results cannot reject MAR externalities and remain inconclusive (see Deidda et al., 2002 or Audretsch, 2003 for overviews).

In the network approach, the interdependent relationships between actors form the nucleus of the theory. They result from information linkages between participants. They develop long-term relations based on trust. A network reduces the costs and risks inherent to research and innovation activities. It is often assumed that SMEs (small and medium-size enterprises) benefit most from such networks, as it overcomes their resource constraints. Specific spatial issues can be discussed within this approach (Arndt and Sternberg, 2000).

The concept of innovative milieux focuses on cultural characteristics of regional development; therefore often attention is paid to rules, conventions and social relations. Although a milieu need not be bounded regionally per se, spatial proximity contributes a lot, as many of the elements studied have strong spatial (regional) boundaries Arndt and Sternberg, 2000).

The concept of regional innovation systems is derived from the ideas of a national and a sectoral innovation system. In essence, it consists of two elements: an institutional infrastructure and a production structure. The institutional infrastructure entails all organisations in a region involved in the activity under consideration (firms, research institutes, government facilities, associations, educational systems, etc). The production structure is formed by the development of innovations and new technologies. Spatial concentrations can reinforce the mechanisms of this model (Arndt and Sternberg, 2000).

[20] For more on this, see chapter 2.

A large part of the literature on regional innovation and clustering highlights the potential benefits of spatial proximity. However, there are some scholars with other views. According to them, regional structures can promote or reinforce fixed, rigid structures; they can enable top-down approaches (e.g. from regional government layers) that form barriers to development, and they can result in over-proportionally high costs (Fürst and Schubert, 1998). The transfer of knowledge in a cluster may induce imitation, cutting the profitability of innovations. Moreover, the probability that all complementary resources can be acquired from within the cluster are low (Oerlemans et al., 2001).

4.4.2 Mechanisms and conditions

So far, the main focus is on patterns of spatial dispersion and analytical tools to analyse them. To push the analysis a bit further, one may want to look at the underlying mechanisms: what are the reasons and determinants for spatial distributions? Steinle and Schiele (2002) observed that not all industries have the same propensity to cluster. They identified six conditions that indicate what kind of industries are likely to cluster. Two conditions are necessary:

- The production process should be divisible (i.e. along the value chain of a product, it is possible to specialise into different activities). If specialisation is not possible (for technical or organisational reasons), the organisational form is more or less naturally predetermined, and unlikely to change.
- The final product must be transportable, for obvious reasons: if the end product cannot be transported, the location of the customers determines the location of the production site.

The other four conditions are 'sufficient'. These are characteristics that foster the emergence of clusters. The first two of these are based on the idea of the complementarity and similarity of activities (where complementary activities often succeed each other in a value chain). Complementary but dissimilar activities often require a lot of coordination efforts. This is where clusters have a potential advantage, as they can reduce the costs of coordination. The two conditions are hence:

- A long, fragmented value chain. Such a chain depends on many complementary actors, and therefore on many interactions and interfaces.
- Dissimilarity of the activities in the value chain. If the activities are strongly different, it is more difficult to carry all of them out in-house. More actors will be involved, and the required coordination effort increases.

The two other sufficient conditions mentioned by Steinle and Schiele (2002) are:

- The importance of innovation. Although the relation between innovation and collaboration is a complicated and disputed one, there seems a common expectation that especially in industries with a high proportion of implicit knowledge, and where innovations emerge in a network of different actors, there are clear benefits from clustering.
- Volatile markets. In markets where the timing of reactions is crucial, organisations need to adopt very quickly. This implies additional interactions between actors, thus fostering spatial clustering.

With these views on the mechanisms and conditions in mind, the next question is: is there any empirical evidence for spatial patterns of collaboration in research? This is dealt with in the next section.

4.4.3 Actual spatial patterns

The relation between R&D, spatial clustering and knowledge spillovers is extensively tested in the literature. Feldman (1999) identifies four topics in the spillover research: the estimation of knowledge production functions, the linkages between patent citations, the mobility of skilled labour, and the trade of knowledge embodied in goods. An overview of the main results on these issues is given below.

In its general form, a knowledge production function with a spatial dimension could be written as:

$$I_{si} = IRD^{\beta_1} * UR_{si}^{\beta_2} * \eta_{si}$$

Where I is the amount of innovative output, IRD is the amount of corporate R&D, and UR is the amount of research by universities and public research agencies (both expressed in e.g. expenditure levels). The s indicates the geographical level (e.g. the state[21]), the i the specific industry. The η is an error term[22]. Jaffe (1989) was the first to carry out an empirical analysis of this sort. He found out that patents tend to be claimed in the states with high levels of private and public research. This result was later confirmed by others, also at a lower geographical level than states. Moreover, it was shown that the degree of geographical concentration is stronger in knowledge intensive industries, or domains where commercial R&D, public research and skilled labour are important inputs (see Feldman, 1999 for an overview of empirical studies). This does give some insight in the geographical patterns of knowledge production, but it does not reveal the underlying mechanisms. Even more, the direction of the causality cannot be discerned. In other words, these analyses do not show whether organisations tend to locate at a favourable location because it used to be an advantageous location already, or the other way around: firms locate together in the same area and hence create a favourable place (see Feldman, 1999).

The idea behind patent research as a means to find spatial patterns in research is that patents leave a ‘paper trail’. Patent applications include the location of the applicant(s) (the inventor or her/his affiliation). Major patents are often cited in other patents. By localizing the stream of citations of a patent, it can be shown whether the knowledge described in the patent spreads spatially over time. Jaffe et al. (1993) found indeed a localisation of citations: patents cite patents applied for in the same city more frequently than other patents. This effect seems to fade away over time; especially the first year after application the patent is cited in the same city more frequently, the knowledge later diffuses.

Later studies try to uncover the mechanisms that cause knowledge spillovers. One idea is that knowledge is embodied in persons, and hence can only spill over to others if people share it. There are some studies (especially on the field of biotechnology) that assess the role of so-called star scientists (highly productive scientists who discovered a major breakthrough in their field). There are also a large group of other scientists who co-authored articles with these star scientists. It is shown indeed that these scientists are concentrated in relatively few universities, research institutes and firms, in a small number of locations. Moreover, it can be shown that new start-ups in biotech tend to reside in those locations. The ties between firms and the scientists was also assessed. Firms that have access to leading scientists (determined by co-authored publications) appear to have a higher firm productivity than those that lack

[21] Note that there is no general consensus on the most appropriate geographical level for this type of analysis. In research on America, states are often used because of data availability and reliability (see Feldman, 1999).
[22] The function can be specified in more sophisticated ways as well. Acs and Varga (2002) for example add a term to measure the geographic coincidence of public and private research.

such access (see Zucker and Darby, 1996 for an overview of studies). Almeida and Kogut (1997) extend this approach. They investigate the interfirm mobility of star patent holders. They find that interfirm mobility cause the spread of ideas (as indicated by subsequently assigned patents), and this form of knowledge spillovers is spatially confined.
It is suggested however that one should not (only) look at leading scientists, but rather to average levels of human capital. Bartel and Lichtenberg (1987) showed with national level data that higher average skill levels in the labour levels are correlated with higher levels of innovation. Glaeser et al. (1992) found a relation between human capital level and economic growth rates of cities.

There are also studies that revolve around the idea that some knowledge spillovers are embodied in traded goods (where the innovator is not able to appropriate the surpluses). Coe and Helpman (1995) show that there are strong and significant knowledge spillovers mediated by trade in goods. This would imply that knowledge spillovers via trade are not (strongly) bounded by geographical distance. However, Branstetter (1996) uses firm level data to assess both national and international spillovers. He finds that knowledge spillovers are mainly national in scope, suggesting that distance does play a role in spread of spillovers. The results in this strand of literature so far are hence inconclusive and there remains a lot to be understood and assessed (see e.g. Feldman, 1999).

5. Clustering in the Dutch water sector: is geographical proximity related to collaboration between organisations?

In the first part of this thesis, several aspects of the dynamics of knowledge production and diffusion are discussed. The focus is on social and spatial determinants of collaboration in knowledge creation. One of the main findings from the literature research is that various dimensions of distance play a role in the configuration of networks involved in research and knowledge production. This does not only concern physical distance, but also but also e.g. social and technological distance (see e.g. Peri, 2005). In this part of the thesis, some of the ideas from the literature will be tested empirically for the Dutch water domain.[23] This selection was made because the knowledge production in the water domain within the Netherlands entails a wide variety of different actors, ranging from drinking water companies to large industrial users, to medical institutes, to academic faculties. Moreover, expertise from the field was available. This is important, for example to be able to identify the boundaries of the domain. Last but not least a large part of the required data preparation was already carried out for this field.

5.1 Gravity model

It is clear from the literature that a variety of conceptual frameworks and methods can be employed to test hypotheses on the configuration of research networks. Such research networks are an essential constituent of the knowledge production in the water industry. This research hence sheds light on the question who collaborates with whom. For this purpose a bibliometric dataset is used for statistical analysis. The dataset is obtained via ISI Web of Knowledge. It contains variables on articles, published in a large set of scientific journals. The set is filtered, to ensure that only papers from (co-)authors in the water sector in the Netherlands appear. The information on each paper includes details on the institutes the authors are affiliated with, among which the location of the institute and the type of organisation (university, firm, etc). It is hence possible to create a matrix containing the distances between all organisations involved. Moreover, a matrix can be created indicating the strength of the relations between the organisations involved, using the number of co-authored articles as a proxy (i.e. the indicator for the strength of the tie between institute A and B consists of a count of the number of papers written by (at least) both an co-author from institute A, and one from institute B).

This model closely resembles a so-called gravity model in economics. Several gravity models exist in the social sciences. The general idea is usually to relate aspects of mass or size and distance to some form of interaction (which is also why they take their name from the notion of gravity by Newton). The best known model is on international trade. Seminal work by Jan Tinbergen in the early 1960s induced a large amount of both theoretical and empirical work (e.g. Santos Silva and Tenreyro, 2006). It seems worthwhile to have a look at this literature here, as there is an obvious analogy between the gravity models for trade and the model that I propose here. Both use the size of actors (in the one case countries, in the other organisations) and the geographical distance between them to explain the strength of the interactions (in trade flows and co-authored publications respectively).

[23] To be more specific, the domain that concerns the drinking water cycle. In the Netherlands, there also is a vast amount of knowledge produced on water management (coastal management, protection against floods, etc.), but that is not considered here.

The most general form of the gravity model of trade can hence be described as follows:
$T_{ij} = \alpha_0 Y_i^{\alpha 1} Y_j^{\alpha 2} D_{ij}^{\alpha 3} \eta_{ij}$

Where T_{ij} is the trade flow between country i and j, Y_i and Y_j denote the size of the respective countries (usually measured in GDP), and D_{ij} is the distance between them. The symbols α_0, α_1, α_2 and α_3 denote unknown parameters (that can be estimated with empirical data). The factor η_{ij} is an error term, with an expected value $E(\eta_{ij}| Y_i, Y_j, D_{ij}) = 1$, and statistically independent of the regressors. For analytical convenience, it is custom to log-linearize the equation (i.e., at both sides of the equation the logarithm of all terms is taken; this makes the task of estimating the unknown parameters a lot easier, as all the exponents disappear because of standard mathematical rules). Doing so yields:
$\ln T_{ij} = \ln \alpha_0 + \alpha_1 \ln Y_i + \alpha_2 \ln Y_j + \alpha_3 \ln D_{ij} + \ln \eta_{ij}$

It should be noted that the assumptions on the error term are still required to hold after this modification. Moreover, it is important to realize that the logarithm of a value ‘zero’ is not defined, hence if there is no trade flow at all between two countries, this equation cannot be used (see also Santos Silva and Tenreyro, 2006). However, this can be solved by adding an (arbitrary) value to all observations.

Others have also attempted to adapt the gravity model for use in the analysis of co-autorship patterns. Sutter and Kocher (2004) seem to be the first. They try to find geographical patterns in the distribution of co-authorships in the economic departments of universities in the United States. The underlying rationale of using a gravity model to assess this phenomenon is that it seems intuitively straight-forward that geographical distance plays a role in the interactions between co-authors (of different institutions), and hence in the frequency of such papers. Moreover, the gravity model is analytically convenient. Its intuitive basic specification can easily be enriched by adding other relevant variables, such as size and quality indicators of the organisations, cultural variables (e.g. main language), etc (see also Sutter and Kocher, 2004).

The most basic specification of the gravity model for co-authorship patterns is as follows:

$$co_{ij} = \beta_0 + \beta_1 D_{ij} + \beta_2 m_i + \beta_3 m_j + \mu_{ij}$$

Where co_{ij} is the number of co-authorships between institute i and j, D_{ij} is the geographical distance between these institutes, and μ_{ij} is an independent and identically distributed (i.i.d.) error term. The unknown parameters to be estimated are β_0 and β_1, where β_0 is a constant.

5.2 Datasets

5.2.1 Dataset publications

The dataset was obtained with the Web of Science (WoS) of the Institute for Scientific Information (ISI) [24]. ISI maintains and provides bibliographic database services, including citation indices with detailed information per article (e.g. on authors, affilitions, keywords, etc). As a first step, a topic search was carried out for the search terms “drinking water”,

[24] Edwin Horlings (Rathenau Institute) constructed this dataset. I gratefully acknowledge his willingness to share data and provide extensive instructions and help. Any remaining errors are the responsibility of the author.

“water treat” and “desalinat*” (where the asterisk is a boolean operator for unknown characters). This is confirmed by experts as a set of key words that cover the research on the water cycle very generally, but also precisely. Moreover, the timeframe of the articles was delimited from 1969 to 2008. To refine the topic search, it was determined what journals published the largest part of these publications. These were: Desalination, Water Research, Environmental Science & Technology, Water Science and Technology, and the Journal of the American Water Works Association. All articles published in these journals were downloaded. The keywords mentioned in the articles were used to develop a more refined set of keywords for topic search. Based on this topic search the final set of publications was generated. To keep the amount of data within workable limits, only the publications from 2006 to 2008 were used. Moreover, only publications where all authors had an affiliation in the Netherlands were used filtered out (this can be done easily with a system of country codes). The result is a set of 2280 from 307 organisations. Out of those, 1254 are co-authored publications. Although all 2280 publications were used to determine the size (mass) of the institutes, only the 1254 co-authored publications were used in determining the relevance of distance. This is done to avoid noise of publications written by several people from the same institute.
All publications in the dataset described above contain details on the affiliations of the authors, including the address of the organisation. Often, there are minor differences in the way names and addresses are written. Therefore, all publications were organized per institute, and all institutes got one unique name and reference code. This is a difficult process. Because of the large number of publications, one has base the names on the information as provided by ISI WoS. Given that some of the organisations have a rather complicated structure (e.g. universities with own faculties, research labs with own names but within the same organisation, mixed use of English and Dutch name of organisation and/or department, etc.) it can be hard to determine whether a certain publication belongs to an institute, or to a separate entity (e.g. a spin-off without formal relations with its “source”). One of the advantages of giving all institutes a unique reference code is that the total number of publications per institute is known. This is used as a proxy variable for the (capacity) size of the institute as far as water-related knowledge production is concerned.

5.2.2 Dataset distances

After generating the dataset with the organisations, numbers of publications, and number of co-authored articles between any possible combination of two organisations, the addresses were used to determine the town where the institute is located. Sometimes one organisation appeared to have locations in several places. In such cases, the town that was most frequently mentioned on the articles was selected as location for publications from that organisation. This was done since the analysis is carried out at institute level, hence it would yield biases if some organisations were split up. It turned out that the 307 institutes had their locations in 97 towns throughout the Netherlands. These locations were georeferenced (i.e. longitude and latitude were collected). The result can be projected on a map, see figure 5.1. The quickest route between any combination of the locations was determined, and collected in a distance matrix. This distance matrix is linked to a matrix containing all possible ($307^2/2$) combinations of co-authorships between organisations (so including the ones that have zero co-publications). Moreover all institutes were categorised to their type of institute (university/academic research, applied research, medical research, consultancy, production industry, government or other). In as far as the name did not reveal a category, the website of the organisation was accessed to get information for categorisation.

Figure 5.1. The 97 towns where the institutes of the co-authors are located. Source: own data. Visualisation with gpsvisualizer.com

5.3 Results

As a first, illustrative result, the data from the dataset on institutes and co-authorship can be visualized, see figure 5.2. The size of the lines between organisations indicates the number of co-authorships they have together. Indicating all 307 organisations in one diagram gives a very messy view. That is why only relationships consisting of 6 or more co-authorships are projected in the figure. The algorithm employed is a so-called ‘directed-force’ algorithm. This algorithm ensures that the lines between institutes are all of more or less equal length, and that there are as little crossing lines as possible. The colours indicate the type of institute: red are academic research institutes and universities, green are medical institutes, blue are applied research institutes, and black are governmental institutes.

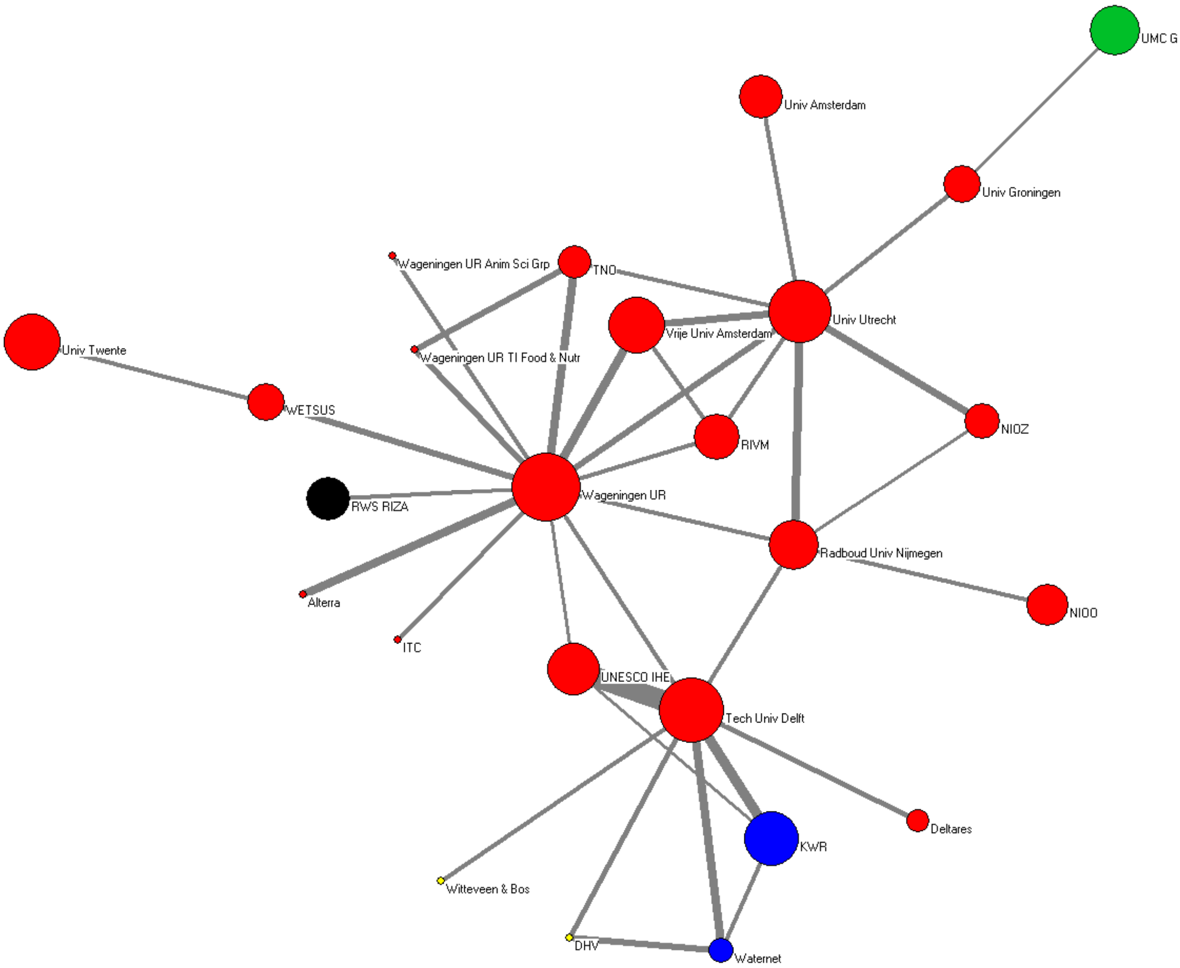

Figure 5.2. Co-authorshiprelations between institutes, only relationships consisting of 6 or more articles. Source: own analysis. Visualization with Pajek software.

More importantly, it is possible to test the validity of a gravity model for collaboration between institutes empirically. First, the most basic, linear regression was estimated:

$$co_{ij} = \beta_0 + \beta_1 D_{ij} + \beta_2 m_i + \beta_3 m_j + \mu_{ij}$$

The results of this model are reproduced in table 5.1.

Table 5.1. Linear gravity model for co-authored publications on water between Dutch institutes.

Source	SS	df	MS
Model	645.88483	3	215.294943
Residual	6775.58973	46677	.145159066
Total	7421.47456	46680	.158986173

Number of obs	46681
F(3, 46677)	1483.17
Prob > F	0.0000
R-squared	0.0870
Adj R-squared	0.0870
Root MSE	.381

pubs_coaut~d	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
distance_ab	-.0001649	.0000267	-6.19	0.000	-.0002172	-.0001127
mass_a	.0060072	.0001188	50.56	0.000	.0057744	.0062401
mass_b	.0020661	.0000488	42.34	0.000	.0019705	.0021618
_cons	-.0030679	.0035767	-0.86	0.391	-.0100783	.0039426

As one can see, all variables of this model (except for the constant) are statistically significant, even at the $\alpha = 0.01$ level. In other words, there is a chance of less than one percent that the relation found between the explanatory variables used and the dependent variable is the result of coincidence in the data sample rather than a real relationship. The explained part of the variance (R^2) is rather low (0.087). This is not very surprising as the model only accounts for two variables (distance and size of institutes), whereas the actual pattern of interactions is the result of a very complicated interplay of much more factors.

As shown before, the original gravity model is not a linear model. We test whether the original specification yields (even) better results than the linear one, by using a loglinearized version:

$$\ln(co_{ij}) = \ln(\beta_0) + \beta_1 \ln(D_{ij} + \beta_2 \ln(m_i) + \beta_3 \ln(m_j) + \ln(\mu_{ij})$$

Where ln stands for the natural logarithm.[25] The results of this specification are shown in table 5.2.

Table 5.2. Log-linear gravity model for co-authored publications on water between Dutch institutes.

Source	SS	df	MS
Model	56.1585717	3	18.7195239
Residual	615.671975	46677	.01319005
Total	671.830547	46680	.014392257

Number of obs	= 46681
F(3, 46677)	= 1419.22
Prob > F	= 0.0000
R-squared	= 0.0836
Adj R-squared	= 0.0835
Root MSE	= .11485

ln_pubs_pl~1	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
ln_distanc~b	-.0078308	.0006384	-12.27	0.000	-.009082	-.0065796
ln_mass_a	.0228293	.0005316	42.95	0.000	.0217874	.0238712
ln_mass_b	.0188231	.0004019	46.84	0.000	.0180354	.0196108
_cons	.018783	.0029531	6.36	0.000	.0129949	.024571

Note that again all explanatory variables are significant at the 0.01 level, and especially the t-value for the distance is much higher than in the linear specification. The R^2-values are also a little bit better. This specification apparently fits better to the actual pattern of collaboration.

5.3.1 A more sophisticated model

One of the appealing characteristics of the gravity model is that it allows for the incorporation of other relevant variables very easily. Before testing, we hypothesized that organisations would tend to have co-authored publications together with organisations from the same ‘type’, for example universities with universities, consultancy bureaus with consultancy bureaus, etc. However, as shown in table 5.3, empirical tests reveal that, although most categories are significant (only the group government is clearly insignificant), together they hardly explain an additional part of the total variance (R^2-values of 0.09 instead of 0.08). Changing the specification (linear or loglinear), or selecting one specific category instead of the whole set does not change these findings.

[25] The reader may have noted that the value of the dependent variable is exactly zero for the majority of the observations (namely for any combination of two institutes in the sample that did not have a co-authored publication in the ISI-dataset). As the log of 0 is not defined, all these observations would be dropped. To prevent this, the value of the dependent variable was increased by one for every observation before taking the log.

Table 5.3. Log-linear gravity model for co-authored publications on water between Dutch institutes with dummies for different categories of organisations.

```
Source |       SS       df       MS              Number of obs =   46681
-------------+------------------------------           F(  9, 46671) =  528.42
       Model |  62.1286663     9  6.90318515           Prob > F      =  0.0000
    Residual |  609.701881 46671  .013063827           R-squared     =  0.0925
-------------+------------------------------           Adj R-squared =  0.0923
       Total |  671.830547 46680  .014392257           Root MSE      =   .1143

------------------------------------------------------------------------------
ln_pubsplus1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ln_distanc~b |   -.007899    .000636   -12.42   0.000    -.0091456   -.0066525
   ln_mass_a |   .0221089   .0005324    41.53   0.000     .0210654    .0231523
   ln_mass_b |   .0179621    .000405    44.35   0.000     .0171683    .0187558
        univ |   .1423026   .0069726    20.41   0.000     .1286362    .1559691
     applied |   .0110676   .0034581     3.20   0.001     .0042897    .0178454
     medical |   .0152734   .0035997     4.24   0.000     .0082178    .0223289
 consultancy |   .0095163    .002786     3.42   0.001     .0040556     .014977
    industry |   .0081689   .0026177     3.12   0.002     .0030382    .0132997
         gov |   .0083869   .0065809     1.27   0.203    -.0045117    .0212855
       _cons |   .0180117   .0029507     6.10   0.000     .0122284     .023795
```

One might argue that all the observations with a value 0 for the number of co-authored publications are in a sense artificial observations, as they simply result from generating a set of all possible combinations of institutes in the dataset.[26] To make sure that the results are not dependent on this methodological issue, all tests were repeated for a dataset that only included ‘actual cooperation’: all observations with at least on co-publication. Interestingly, this results in a model with lower t-values (although all variables are still significant), but much higher R^2-values (this paradoxical result may well have to do with the fact that the number of observations is much smaller in this case). The results for the loglinearized model are in table 5.4.

Table 5.4. Log-linear gravity model for co-authored publications on water between Dutch institutes, at least one co-publication per institute.

```
Source |       SS       df       MS              Number of obs =     646
-------------+------------------------------           F(  3,   642) =   81.91
       Model |  74.0088095     3  24.6696032           Prob > F      =  0.0000
    Residual |  193.362863   642  .301188261           R-squared     =  0.2768
-------------+------------------------------           Adj R-squared =  0.2734
       Total |  267.371673   645  .414529725           Root MSE      =  .54881

------------------------------------------------------------------------------
     ln_pubs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ln_afstand~B |   -.072664   .0191384    -3.80   0.000    -.1102453   -.0350826
  ln_massa_a |   .1765108   .0130269    13.55   0.000     .1509303    .2020914
  ln_massa_b |   .1065617   .0117967     9.03   0.000      .083397    .1297264
       _cons |  -.0794641   .0877567    -0.91   0.366    -.2517889    .0928607
```

5.4 Discussion

The data are checked visually for ‘suspicious’ observations by using plots (e.g. QQplots, plotting residuals against fitted values, etc). A few outliers were found. Most of them were the result of a typing error in the distance matrix and could easily be corrected. The last one is no error, but appeared an influential observation (no less than 40 co-authored publications

[26] On the other hand, it can well be defended to have the ‘zero-observations’ included in the set: distance might be a factor in the outcome that two institutes have no joined publications.

between two institutes). However, omitting this observation (in whatever specification of the model) hardly changes the results and certainly does not change the interpretation and conclusions of the models.

Another potential source of noise is the distance between two institutes in the same town. To limit the size of the distance matrix, only “intertown” distances were estimated with geo-referencing. The “intratown” distances were set at one value. However, by running the models for several different values of this “intratown” distance, it is made sure that the value chosen do not alter the interpretation of the models.

As explained above, all organisations were categorised in one of six categories. The categories were a bit crude, and not perfectly mutually exclusive. To give an example, academic medical centres could be put under either universities or under medical institutes (they are in the latter). Moreover, it should be stressed that many of the organisations in the dataset are at the boundary of several categories. Most of the firms from the ‘production industry’ for instance produce very knowledge intensive products (e.g. highly specific systems for data collection and processing in various fields). The distinctions between the categories are hence quite fuzzy. Despite all that, we would have expected that it would have explained a larger part of the variance.

It is rather difficult to validate the results, as there is no large body of literature with comparable research. In fact, we have only come across one article where a similar model is employed. Sutter and Kocher (2004) use a dataset with publications from fifteen journals in economics with a high impact factor, and select those with at least one author affiliation in the United States. In addition to information on the geographical location of the institutes, they also have data on a quality ranking. They did not employ any proxy for the size of the investigated institutes. They use an linear model to test a gravity model. Surprisingly, and contrary to their own expectations and hypotheses, they find that all geographical variables employed (distance, being in the same or an adjacent state in the US) appear insignificant. The only significant variables appear to be the ones derived from the quality ranking (and those do not yield very high R^2-values). Intuitively, one can come up with a few possible reasons for the strong contrast between their findings and the results reported above. First of all, Sutter and Kocher only have top-ranked journals in their dataset. It may well be that especially for such more prestigious publications, co-authoring with someone from a renowned institute is more important than the distance to that institute. In our dataset are publications from a much broader set of journals and organisations, including many organisations for which scientific research is not the core business. Probably for them it is more important to have a partner at a convenient location than to have one from high (academic) standing. Moreover, both the domain (water versus economics) and the geographical scale (the Netherlands versus all of the US) are completely different. This may well result in different findings. It would require much more research to find out whether our results have any validity beyond the investigated domain and geographical area.

It should be stressed that the method used has strong limitations if one is to deduce patterns of collaboration. The most important one does not stem from the concepts used in the model as such, but rather from limitations in data available that could not be overcome within the scope and timeframe of this research. As explained in the first part of the thesis[27], knowledge can exist in many forms, and research collaboration can hence have many different forms of

[27] See also chapter 1.

output. A publication in scientific journals is just one of them, and probably not even the most important one[28]. This does not only imply that the method used strongly underestimates the actual intensity of interaction and collaboration in the water domain; it may well give a bias in the collaboration patterns found (if, for example, the propensity to transform knowledge into scientific papers is higher among universities than among commercial firms). Despite all this, it is worthwhile to carry out the analysis. Journal articles form a very reliable paper trail, and the data are relatively easy to gather. This enables to build a solid model and test it. The model allows for other indicators of collaboration as well, so once a rigorous model is developed, its validity can be ameliorated by using outputs other than journal articles.
An advantage of using authorship of articles as a variable is that it may pick up many informal relationships between organisations that are nowhere registered, and may hence be missed in many other approaches.

Other restricting limitations are in the selection of the dataset. Only organisations within the Netherlands are part of our sample[29]. Especially in the academic world, collaboration is increasingly a phenomenon that goes beyond state boundaries. It is very likely that authors with an affiliation to the organisations in our sample also have a lot of co-authored publications with partners outside the Netherlands. It would be worthwhile to investigate if geographical distances have the same influence at several scale levels (e.g. sub-country level, country level, continent level, world level). This is however beyond the scope of this thesis.
A related issue is that many researchers have a personal (often international) network of peer scientists with whom they interact. It is likely that the emergence of such networks is (partially) determined by personal characteristics of the researchers involved (e.g. drive to collaborate, reputation, etc). Using data that are disaggregated to the personal instead of the institutional level would enable the investigation of the relevance of such factors and the role of geographical distances therein.

One should realise that geographical distance is not the only relevant form of distance in this context. One might also think of e.g. cognitive or technological distance. The model would benefit considerably if it would be able to incorporate such forms of distance. However, most of them are hard to make operational, let alone to measure. This is beyond the scope of this thesis, but it may well prove worthy as a path for future research.

It is important to point out that the direction of causality in our model is more ambiguous than it may seem at first sight. Intuitively, one may be tempted to consider the physical location of organisations as exogenous (or given). However, in our dataset are at least a few organisations (often relatively young and small ones) that seem a spin-off of a larger organisation. In that case, it may well be that its location is deliberately chosen near the larger institute, for the very reason of easier collaboration. Testing whether this is indeed a significant effect would require the use of appropriate instrumental variables.

[28] For details on the relevance of different forms of knowledge in research collaboration, see section 4.1.
[29] More strictly speaking, the affiliation address stated on the publication should be within the Netherlands.

6. Conclusions and recommendations

6.1 Conclusions

There is general consensus that many different forms of knowledge exist. The form may have implications for the transferability of the knowledge. That in turn has consequences for a large variety of things, from the motivation to produce knowledge to the propensity to collaborate in knowledge production.
There is also consensus that the science-push model (explaining the relation between research, innovations and socio-economic benefits) is an oversimplification of reality as it lacks many feedback loops. However, there is no general agreement on a more realistic model.

At different scale levels there are different reasons for the production of knowledge. The production of knowledge generates positive externalities, and because of that it is beneficial at macrolevel to promote the creation of knowledge. Not only at the macro level, but also at the micro level there are obvious benefits from investments in research and knowledge creation, although the private returns to R&D are strongly dependent on the institutional environment (e.g. protection of intellectual property). Yet, private firms are not likely to spend as much money on research as would be optimal from a social point of view. There is no agreement so far on how the government could stimulate the development of knowledge. Individuals in organisations also have incentives to create knowledge. Apart from an intrinsic motivation, they are stimulated by the incentive system of the organisation.

Both in commercial and in public sector research, several approaches and models are employed to study the research agenda setting process. However, these models seem work mainly as a guide in the process, not as an instrument to study the process of agenda setting.

Individual organisations are rarely able to innovate (or conduct research) independently. There is a lot of attention for collaboration in research and knowledge production. Firms have two reasons to enter into a research partnership: overcoming internal resource constraints, and reducing the risks associated with conducting research. It should be noted however that cooperation is not always beneficial from a social point of view. Several characteristics explain the propensity to collaborate. These include: size of the firm, size of the in-house R&D, educational level of the employees. A large branch of literature investigates the spatial determinants of research cooperation. Again, there is hardly any general consensus on the underlying mechanisms, or on an appropriate model.

A possible model to test the spatial patterns of collaboration empirically, is the gravity model. It relates the size of two institutes and the geographical distance between them with the intensity of cooperation. It was employed in this study to examine patterns of co-authorship in the water sector in the Netherlands. The dataset used contains publications on the water cycle, published by organisations within the Netherlands, from 2006 to 2008. The geographical location of the organisations is referenced, and a distance table created. The results show that both the mass of the institutes (as proxied by the total number of publications) and the distance between them are highly significant explanatory variables. This result is robust to changes in specification of the model. However, the proportion of explained variance is rather small. This suggest that the pattern of collaboration is the result of a much more complicated interplay of different factors. Yet, expanding the model with variables to indicate whether both co-authoring institutes are from the same category (e.g. both universities) hardly adds to the explained variance of the model.

6.2 Recommendations

Despite the enormous amounts of research on knowledge production, there appears to be very little convincing empirical evidence on major hypotheses and ideas in this field. Of course, there is a lot of empirical research available, but the results seem fragmented and often inconclusive or contradictory to findings of others. As this thesis has covered a very broad part of the literature, the recommendations only point to a few striking issues in the status quo of the research in this area.

Although there is quite some insight in the incentives of organisations to engage in knowledge production, remarkably little is known on such incentives at industry level. The concept of externalities is illustrative in this respect. It is shown that externalities play an important role in the benefits of knowledge production. However, the old dispute whether these externalities are more important within or among industries, is not resolved so far (see e.g. Deidda et al., 2002 for an overview). It may prove worthwhile to develop a new conceptual approach to examine the incentives for organisations to share knowledge (or conduct research together). Explicit attention should be paid to the way organisations select their research partners and whether or not type of industry plays a role in that respect.

The selection criteria for research partners are interesting for more reasons. The various results suggest that at least social, institutional and spatial issues can explain part of the observed collaboration patterns. It may be interesting to assess whether or not the perceptions of firms on the respective importance of these determinants is in line with empirical observations. Developing a sound foundation for the concepts of social, institutional and spatial proximity may prove a sine qua non in this line of research.

Most of the literature so far seems to focus on knowledge production and research as processes. The content of knowledge is much less considered. There are quite some models available for research agenda setting (what topics should be researched), but these are often meant as instruments to guide the process of research agenda setting in the organisation. Models to analyse the actual process of agenda setting as it takes place in institutes are scarce.

Perhaps most important of all, the literature so far is very scattered. If large steps forward are to be made, a more common approach is probably required. This might be reached by developing a general framework, consisting of a few building blocks to bring several branches of literature together, and confront researchers with more insights and results of others. This should include insights on the motivations for economic agents to create knowledge, how the benefits of conducting research accrue to the performers and to others, and the motivations and mechanisms for collaboration. Such a synthesis could constitute the starting point for a well-developed knowledge-based theory of economic growth at various levels.

List of References

Acs, Z. J. and A. Varga, 2002, Geography, endogenous growth, and innovation. *International Regional Science Review* 25, 132.

Aghion, P., M. Dewatripont and J. C. Stein, 2005, Academic Freedom, Private-Sector Focus, and the Process of Innovation, American Law & Economics Assiocation Annual Meeting (Berkeley Electronic Press), 39.

Almeida, P. and B. Kogut, 1997, The exploration of technological diversity and geographic localization in innovation: Start-up firms in the semiconductor industry. *Small Business Economics* 9, 21-31.

Archibugi, D., J. Howells and J. Michie, 1999, Innovation Systems and Policy in a Global Economy, in: D. Archibugi, J. Howells and J. Michie, eds., Innovation Policy in a Global Economy (Cambridge University Press, Cambridge) 1-18.

Arndt, O. and R. Sternberg, 2000, Do manufacturing firms profit from intraregional innovation linkages? An empirical based answer. *European Planning Studies* 8, 465-485.

Arrow, K. J., 1962a, The Economic Implications of Learning by Doing. *The Review of Economic Studies* 29, 155-173.

Arrow, K. J., 1962b, Economic Welfare and the Allocation of Resources for Invention, in: NBER, ed., The Rate and Direction of Inventive Activity: Economic and Social Factors (UMI) 609-626.

Arundel, A., G. v. d. Paal and L. Soete, 1995, Innovation Strategies of Europe's largest Industrial Firms; Results of the PACE Survey for Information Sources, Public Research, Protection of Innovations and Government Programmes, (MERIT, Maastricht) 185.

Audretsch, B., 1998, Agglomeration and the location of innovative activity. *Oxf Rev Econ Policy* 14, 18-29.

Audretsch, D. B., 1995, *Innovation and Industry Evolution* (MIT Press, Cambridge, MA), 205.

Audretsch, D. B., 2003, Innovation and spatial externalities. *International Regional Science Review* 26, 167.

Audretsch, D. B., B. Bozeman, K. L. Combs, M. Feldman, A. N. Link, D. S. Siegel, P. Stephan, G. Tassey and C. Wessner, 2002, The Economics of Science and Technology. *The Journal of Technology Transfer* 27, 155-203.

Audretsch, D. B. and P. E. Stephan, 1999, Knowledge spillovers in biotechnology: sources and incentives. *Journal of Evolutionary Economics* 9, 97-107.

Baldwin, W. L. and A. N. Link, 1998, Universities as research joint venture partners: does size of the venture matter? *International Journal of Technology Management* 15, 895-913.

Bania, N., R. W. Eberts and M. S. Fogarty, 1993, Universities and the Startup of New Companies: Can We Generalize from Route 128 and Silicon Valley? *The Review of Economics and Statistics* 75, 761-766.

Bartel, A. P. and F. R. Lichtenberg, 1987, The comparative advantage of educated workers in implementing new technology. *The Review of Economics and Statistics* 69, 1-11.

Berg, S. V., J. Duncan and P. Friedman, 1982, *Joint Venture Strategies and Corporate Innovation* (Oelgeschlager, Gunn and Hain, Cambridge).

Branstetter, L., 1996, Are Knowledge Spillovers International or Intranational in Scope? Microeconometric Evidence from the Japan and the United States. *NBER Working Papers*.

Carayol, N., 2007, Academic Incentives, Research Organizatin, and Patenting at a Large French University. *Economics of Innovation and New Technology* 16, 119-138.

Cockburn, I., R. Henderson and S. Stern, 1999, *Balancing Incentives: the Tension between Basic and Applied Research* (NBER, Cambridge, MA) 47.

Coe, D. T. and E. Helpman, 1995, International R&D spillovers. *European Economic Review* 39, 859-887.

Coe, D. T., E. Helpman and A. W. Hoffmaister, 2009, International R&D spillovers and institutions. *European Economic Review* 53, 723-741.

Cohen, W. M. and D. A. Levinthal, 1990, Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly* 35, 128-152.

Cooke, P. and K. Morgan, 1993, The network paradigm: new departures in corporate and regional development. *Environment and Planning D: Society and Space* 11, 543-564.

d'Aspremont, C. and A. Jacquemin, 1988, Cooperative and noncooperative R & D in duopoly with spillovers. *The American Economic Review* 78, 1133-1137.

Dalrymple, D. G., 2006, Setting the agenda for science and technology in the public sector: the case of international agricultural research. *Science and Public Policy* 33, 277-290.

Dasgupta, P. and P. A. David, 1994, Toward a new economics of science. *Research Policy* 23, 487-521.

Deidda, S., R. Paci and S. Usai, 2002, Spatial externalities and local economic growth. *Contributi di Ricerca CRENoS* 2.

Dosi, G., 1988, The nature of the innovative process, in: G. Dosi, C. Freeman, R. Nelson, L. Soete and G. Silverberg, eds., Technical change and economic theory (Pinter, London).
Faulkner, W., 1994, Conceptualizing Knowledge Used in Innovation: A Second Look at the Science-Technology Distinction and Industrial Innovation. *Science Technology Human Values* 19, 425-458.

Faulkner, W., J. Senker and L. Velho, 1994, *Knowledge frontiers: Industrial innovation and public sector research in biotechnology, engineering ceramics and parallel computing.* (Clarendon Press, Oxford).

Feldman, M. P., 1999, The New Economics Of Innovation, Spillovers And Agglomeration: A Review Of Empirical Studies. *Economics of Innovation and New Technology* 8, 5-25.

Feldman, M. P. and D. B. Audretsch, 1999, Innovation in cities: Science-based diversity, specialization and localized competition. *European Economic Review* 43, 409-429.

Fleck, J. and M. Tierney, 1991, The management of expertise: Knowledge, power and the economics of expert labour. *Edinburgh PICT Working Paper*, 29.

Fontes, M., 2005, The process of transformation of scientific and technological knowledge into economic value conducted by biotechnology spin-offs. *Technovation* 25, 339-347.

Freel, M. S., 2003, Sectoral patterns of small firm innovation, networking and proximity. *Research Policy* 32, 751-770.

Fritsch, M. and R. Lukas, 2001, Who cooperates on R&D? *Research Policy* 30, 297-312.

Fürst, D. and H. Schubert, 1998, Regionale Akteursnetzwerke. *Raumforschung und Raumordnung* 56, 352-361.

Gibbons, M. and R. Johnston, 1974, The roles of science in technological innovation. *Research Policy* 3, 220-242.

Glaeser, E. L., H. D. Kallal, J. A. Scheinkman and A. Shleifer, 1992, Growth in cities. *Journal of Political Economy* 100.

Grant, R. M., 1996, Toward a Knowledge-Based Theory of the Firm. *Strategic Management Journal* 17, 109-122.

Griliches, Z., 1995, R&D and productivity: econometric results and measurement issues, in: P. Stoneman, ed., Handbook of the economics of innovation and technological change (Blackwell, Oxford) 52-89.

Guston, D. H., 2000, *Between Politics and Science: Assuring the Integrity and Productivity of Research.* (Cambridge University Press, New York), 213.

Guttman, J. M., 1980, Villages as Interest Groups: The Demand for Agricultural Extension Services in India. *Kyklos* 33, 122-141.

Hagedoorn, J., A. N. Link and N. S. Vonortas, 2000, Research partnerships. *Research Policy* 29, 567-586.

Hayami, Y. and V. W. Ruttan, 1971, *Agricultural Development: An International Perspective* (John Hopkins University Press, Baltimore) 367.

Helpman, E., 2004, *The Mystery of Economic Growth* (Harvard University Press, Cambridge, MA) 223.

Henderson, R. and I. Cockburn, 1994, Measuring Competence? Exploring Firm Effects in Pharmaceutical Research. *Strategic Management Journal* 15, 63-84.

Howitt, P., 2003, The Economics of Science and the Future of Universities, Timlin Lecture (University of Saskatchewan, Saskatoon).

Jaffe, A. B., 1989, Real effects of academic research. *The American Economic Review* 79, 957-970.

Jaffe, A. B. and M. Trajtenberg, 2002, *Patents, Citations & Innovations; a Window on the Knowledge Economy* (MIT Press, Cambridge, MA), 478.

Jaffe, A. B., M. Trajtenberg and R. Henderson, 1993, Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics* 108, 577-598.

Janvry, A. d., 1978, Social Structure and Biased Technical Change in Argentine Agriculture, in: H. Binswanger and V. Ruttan, eds., Induced Innovation: Technology, Institutions and Development (John Hopkins University Press, Baltimore).

Katz, M. L., 1986, An analysis of cooperative research and development. *The Rand Journal of Economics*, 527-543.

Kitcher, P., 2001, *Science, Truth and Democracy* (Oxford University Press, New York), 240.

Kogut, B., 1988, Joint ventures: Theoretical and empirical perspectives. *Strategic Management Journal* 9, 319-332.

Kogut, B. and U. Zander, 1992, Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization Science* 3, 383-397.

Leyden, D. P. and A. N. Link, 1999, Federal laboratories as research partners. *International Journal of Industrial Organization* 17, 575-592.

Liberatore, M. J., 1988, An expert support system for R&D project selection. *Mathematical and Computer Modelling* 11, 260-265.

Link, A. N., 1990, Strategies for Cooperation in R&D, in: T. Khalil and B. Bayraktar, eds., Management of Technology (Institute of Industrial Engineering, Miami).

Link, A. N. and L. L. Bauer, 1989, *Cooperative Research in US Manufacturing* (Lexington Books, Lexington).

Lucas, R. E., 1988, On the mechanics of economic development. *Journal of Monetary Economics* 22, 3-42.

Mansfield, E., 1991, Academic research and industrial innovation. *Research Policy* 20, 1-12.

Mansfield, E., 1998, Academic research and industrial innovation: An update of empirical findings. *Research Policy* 26, 773-776.

Marshall, A., 1920, *Principles of Economics: an introductory volume* (McMillan, London), 872.

Martin, B. R. and P. Tang, 2007, The benefits from publicly funded research, SPRU Electronic Working Paper Series, Paper no. 161 (SPRU, University of Sussex, Brighton, UK).

Merton, R. K., 1957, Priorities in Scientific Discovery: A Chapter in the Sociology of Science. *American Sociological Review* 22, 635-659.

Meyer-Krahmer, F. and G. Reger, 1999, New perspectives on the innovation strategies of multinational enterprises: lessons for technology policy in Europe. *Research Policy* 28, 751-776.

Narin, F., K. S. Hamilton and D. Olivastro, 1997, The increasing linkage between U.S. technology and public science. *Research Policy* 26, 317-330.

Nickerson, J. A. and T. R. Zenger, 2004, A Knowledge-Based Theory of the Firm: The Problem-Solving Perspective. *Organization Science* 15, 617-632.

Nonaka, I., 1994, A Dynamic Theory of Organizational Knowledge Creation. *Organization Science* 5, 14-37.

Oerlemans, L. A. G., M. T. H. Meeus and F. W. M. Boekema, 1998, Do Networks Matter for Innovation? The usefulness of the economic network approach in analysing innovation. *Tijdschrift voor economische en sociale geografie (Journal of Economic and Social Geography)* 89, 298-309.

Oerlemans, L. A. G., M. T. H. Meeus and F. W. M. Boekema, 2001, Firm clustering and innovation: Determinants and effects. *Papers in Regional Science* 80, 337-356.

Owen-Smith, J., 2003, From separate systems to a hybrid order: accumulative advantage across public and private science at Research One universities. *Research Policy* 32, 1081-1104.

Peri, G., 2005, Determinants of knowledge flows and their effect on innovation. *Review of Economics and Statistics* 87, 308-322.

Porter, M. E., 1998, Clusters and the new economics of competition. *Harvard business review* 76, 77-90.

Porter, M. E., 2000, Location, competition, and economic development: Local clusters in a global economy. *Economic Development Quarterly* 14, 15-34.

Roessner, D., C. P. Ailes, I. Feller and L. Parker, 1998, How Industry benefits from NSF's Engineering Research Centers. *Research Technology Management* 41, 5.

Romer, P. M., 1986, Increasing Returns and Long-Run Growth. *Journal of Political Economy* 94, 1002.

Romer, P. M., 1990, Endogenous Technological Change. *Journal of Political Economy* 98, S71.

Rosenberg, N., 1990, Why do firms do basic research (with their own money)? *Research Policy* 19, 165-174.

Rosenberg, N. and R. R. Nelson, 1994, American universities and technical advance in industry. *Research Policy* 23, 323-348.

Rothwell, R., 1977, The characteristics of successful innovators and technically progressive firms (with some comments on innovation research). *R&D Management* 7, 191-206.

Salter, A. J. and B. R. Martin, 2001, The economic benefits of publicly funded basic research: a critical review. *Research Policy* 30, 509-532.

Sarewitz, D. and J. R. A. Pielke, 2007, The neglected heart of science policy: reconciling supply of and demand for science. *Environmental Science & Policy* 10, 5-16.

Scott, J. T., 1996, Environmental research joint ventures among manufactures. *Review of Industrial Organization* 11, 655-679.

Silva, S. and S. Tenreyro, 2006, The log of gravity. *Review of Economics and Statistics* 88, 641-658.

Spender, J. C. and R. M. Grant, 1996, Knowledge and the Firm: Overview. *Strategic Management Journal* 17, 5-9.

Steinle, C. and H. Schiele, 2002, When do industries cluster?: A proposal on how to assess an industry's propensity to concentrate at a single region or nation. *Research Policy* 31, 849-858.

Stephan, P. E., 1996, The Economics of Science. *Journal of Economic Literature* 34, 1199-1235.

Stern, S., 2004, Do Scientists Pay to Be Scientists? *Management Science* 50, 835-853.

Sternberg, R., 2000, Innovation networks and regional development—evidence from the European Regional Innovation Survey (ERIS): theoretical concepts, methodological approach, empirical basis and introduction to the theme issue. *European Planning Studies* 8, 389-407.

Stewart, J., 1995, Models of priority-setting for public sector research. *Research Policy* 24, 115-126.

Sutter, M. and M. Kocher, 2004, Patterns of co-authorship among economics departments in the USA. *Applied Economics* 36, 327-333.

Tether, B. S., 2002, Who co-operates for innovation, and why:: An empirical analysis. *Research Policy* 31, 947-967.

Verspagen, H. H. G., 1992, Uneven growth between interdependent economies: an evolutionary view on technology gaps, trade and growth, Faculty of Economics and Business Administration (University of Limburg, Maastricht) 248.

Vincenti, W., 1991, *What engineers know and how they know it: Analytical studies from aeronautical history.* (John Hopkins University Press, Baltimore).

Vonortas, N. S., 1994, Inter-firm cooperation with imperfectly appropriable research. *International Journal of Industrial Organization* 12, 413-435.

Winter, S. G., 1987, Knowledge and competence as strategic assets, in: D. Teece, ed., The competitive challenge: Strategies for industrial innovation and renewal (Ballinger, Cambridge, MA).

Wright, B. D., 1983, The Economics of Invention Incentives: Patents, Prizes, and Research Contracts. *The American Economic Review* 73, 691-707.

Young, A., 1998, Growth without Scale Effects. *Journal of Political Economy* 106, 41-63.

Zellner, C., 2003, The economic effects of basic research: evidence for embodied knowledge transfer via scientists' migration. *Research Policy* 32, 1881-1895.

Zucker, L. G. and M. R. Darby, 1996, Star scientists and institutional transformation: Patterns of invention and innovation in the formation of the biotechnology industry. *Proceedings of the National Academy of Sciences of the United States of America* 93, 12709.

Zucker, L. G., M. R. Darby and J. S. Armstrong, 2002, Commercializing Knowledge: University Science, Knowledge Capture, and Firm Performance in Biotechnology. *Management Science* 48, 138-153.