CESIS

Electronic Working Paper Series

Paper No.108

HOW CAN RESEARCH NETWORKS IMPROVE THE INNOVATION PROCESS?

Mats Wilhelmsson

December 2007

HOW CAN RESEARCH NETWORKS IMPROVE THE INNOVATION PROCESS?

By

Mats Wilhelmsson*

December, 2007

* Royal Institute of Technology, 100 44 Stockholm, Sweden, matsw@infra.kth.se. Funding from Sweden-America Foundation and Wenner-Gren Foundation is highly appreciated. I also thank the Haas School of Business and Goldman School of Public Policy (UC Berkeley) for their hospitality. Valuable comments from discussions with Professor Roland Andersson and Professor John Quigley have been, as always, fruitful. I also thank the participants at Homer Hoyt Institutes May Weimer School 2007 for valuable comments, two anonymous referees, and the editor Lawrence Smith. Finally, I am grateful to Sandy Nixon for her editorial help with the manuscript.

How can Research Networks Improve the Innovation Process?

August, 2007

Abstract: Accumulation of human capital is essential for economic growth. An important question is how knowledge spillover into innovations and production. One way of knowledge diffusion is within innovation networks. We investigate innovative networks in patent data in Sweden from 1994-2001. We define research networks with the help of direct and indirect ties among inventors. The main result clearly indicates that those researchers that collaborating, in innovation networks, improves the efficiency of the innovation process by getting more patents applications approved. The odds getting a patent application approved are in the range 1.1 to 1.5 times better if an application is a result from research collaboration. Moreover, the result suggests that collaboration is more important in the IT sector than in the mechanical engineering sector. Finally, the empirical outcomes indicate that networking is more important in less dense areas compared to the denser labor markets. Thus, networks in such areas might be a substitute for agglomeration advantages.

Keywords: Innovation network analysis, patents, success and failure in innovation.

1. Introduction

Out of all filed patent application in Sweden in the Year 1994, less than 50 percent were awarded by the end of Year 2001. In some industries, the success rate is higher, and in other, it is lower. What determine success and failure in innovation? It is not unreasonable to expect that corporation is more likely to get a patent application approved compared to private persons, and it maybe more likely that large corporations are more successful. Hence, size perhaps matters, but it seems that the size of the research project is more important than the size of the company. For example, some very small companies in the biotech sector are just one research project, hence, the size of the company is relatively small, but the size of the research project is large. Moreover, what is less obvious is that the number of researchers in R&D projects has an impact on the success rate (see Freeman and Soete, 1997). That is to say, given R&D resources, the size of the project team appears to increase the probability that the innovation will succeed.

There has been a growing interest in research networks and its implications on the creation of new knowledge. For example, there seems to be a consensus that those "scientists who collaborate with each other are more productive, oftentimes producing 'better' science, than are individual investigators".[1] According to Acs (2000), networks are also associated with a greater degree of innovativeness. Hence, innovative networks generate more knowledge, better innovations, and, therefore, higher profits and more wealth.

What explains that the size of the research team and the network matter? Storper and Venables (2004) give one explanation. They analyzed the concept of face-to-face contacts as a way of transfer knowledge. Within networks, they argue, knowledge is more efficiently diffused among the inventors and, thereby, the innovation is more likely to succeed.

The aim of the present research is to contribute to a better understanding of how important innovative networks are for the efficiency of research. Our objective is to test the hypothesis that innovative networks perform research more efficiently in the sense that they are more likely to get their patent application approved. In particular, we want to analyze and test the hypothesis on existing innovative networks in Sweden over the period 1994-2001.

[1] Stephan (1996, pp 1221-1222). See also "The Wisdom of Crowds" by Surowiechi (2004).

Section 2 presents a brief literature review on economic growth, human capital, and innovation, as well as networks and social network analysis. In Section 3, we discuss the used methods in the paper including the basic concepts of the social network analysis and the used discrete-time hazard model, and in Section 4, we present the data together with some descriptive and social network statistics. In Section 5, we give the econometric analysis, and Section 6 ends this paper with a conclusion.

2. A Brief Literature Review

Externalities flowing from human capital in economic growth had a scientific revival with the endogenous growth models starting with Romer (1986). However, the precise linkage between academic research, knowledge, spillovers, and economic growth remains unclear. In the words of Jaffe et al (2002), the "transport mechanism" is not well understood.

Empirical studies have attempted to quantify these knowledge transfers from research to innovating firms, through various proxies, for example, by investigating the patenting of innovations. Griliches (1979), Jaffe (1986), and others have all modeled this knowledge transfer effect in a production function framework and found a significant and positive effect of university research on output.

In a different approach, Storper and Venables (2004) argue that face-to-face contact is the "missing aspect of mechanisms that are considered to generate agglomeration." Why do networks come up with 'better' projects? One reason that collaboration on average produce 'better' research with a higher aggregated return is that networks and face-to-face contacts is efficient when it comes to communication and that it can solve incentive potential problems, as well as it can make socialization and learning easy, and provides psychological motivation. An alternative reason why networking projects have a higher success rate is that the networks in it self is a screening process, and the networks only include the best researchers. Hence, the networks will include a higher than average shares of competent researchers, that works harder, and that have a higher probability of undertaking successful projects.

Owen-Smith and Powell (2004) make a distinction of innovation networks as channels and as conduits. The former is characterized as a personal tie among inventors making knowledge spillovers possible and the latter is more a legal arrangement between two companies. By using firm data within the biotechnology industry in the Boston region, they showed that a "membership in a geographically collocated network will positively effect innovation." Earlier results showing similar conclusions are, for example, Walker et al (1997) and Stuart (2000). For example, Stuart (2000) analyzes technological collaboration among firms and its outcome. His findings indicate that collaborating firms both performs better research and are more innovative than non-collaborating companies are.

What determine "success and failure in industrial innovation"? Freeman and Soete (1997) have an extensive review and discussion on this issue. For example, Rothwell et al (1974) analyzed 58 pairs of innovations (29 successes and 29 failures) within the chemical and instrument industry in the so-called SAPPHO-project. The instrument industry was mostly electronic instruments and the chemical industry was mainly related to petroleum products. Among other things, their result revealed that the size of the project team is important for success. The size is important both in an initial stage of the research project and at the peak of the project. Furthermore, the size of the project team appears to be more important within the chemical industry than in the instrument industry. Other characteristics of success are the size of the project, but not necessarily the size of the firm, and that the project is linked to the outside scientific community, that is, that there exist a university-industry knowledge transfer.

How can we measure and characterize innovative networks? Some recent empirical studies investigating innovative networks and discuss how networks could be measured are Balconi et al (2004) and Ejermo and Karlsson (2004), as well as Singh (2005) and Breschi and Lissoni (2006). They all use the social network analysis from sociology as a tool to elaborate the concept and deepen our understanding of networks.

Balconi et al investigate the role of academia in innovation networks. They do that by performing a social network analysis of Italian patent data. Their conclusion is that the formation of networks is very scattered and fragmented in Italy. The exception is within science-oriented technology fields, such as in the chemical industry. Their results also indicate that a substantial portion of the innovations within high-technology sectors is a result of university-industry collaboration.

In a recent paper by Ejermo and Karlsson (2006), the interregional structure of inventor networks in Sweden was investigated. They measure how close the relationship is between two regions. They found out that the relationship is highly affected by the distance between the two labor markets. Fritsch (2001) also emphasized in his study in Germany that spatial proximity is important for collaboration among firms.

Recently, Singh (2005) and Breschi and Lissoni (2006) investigated, with he help of social network analysis, the question about knowledge diffusion (spillovers). The investigations use patent data from the U.S. and Italy, respectively. As anticipated, Singh's findings indicate that the flow of knowledge is stronger within firms and within regions than flows across firm and regional boundaries. Breschi and Lissoni's result suggests that citation patterns are bounded in space if the inventors are relatively immobile and if the social network is not so spatially fragmented and dispersed.

3. Methodology and Variable Selection

Our proposition is that the likelihood to get a patent awarded is higher if the research forgoing the application has been a result of networking, everything else equal. The importance of networks may vary among industries and size of company.

We are using two very distinct methods in this paper to test this hypothesis. First, in an attempt to investigate the innovative network we are going to use the toolbox of social network analysis. Here, only a limited descriptive part of the toolbox is used. By utilizing social network analysis, our aim is to construct a measure characterizing the size of the innovative networks. The second method is that of econometric analysis or more precisely the estimation of a discrete-time hazard (duration) model in order to answer the research question whether networked innovations of better quality and more successful when it comes to approval rates.

3.1. Social network analysis

Innovation networks will be investigated by using some basic concepts from the social network analysis toolbox. Social network analysis is not very common in economic literature. However, the method has become more and more accepted and used. Examples using empirical social analyses within Economics are Owen-Smith et al (2002), Balconi et al (2004), Singh (2005), and Ejermo and Karlsson (2006), as well as Breschi and Lissoni (2006) and Cantner and Graf (2006).

The basic social network analysis examines the nodes and the ties, and the relationship between them. In the context of innovation networks, the nodes are the inventors and the ties are the relationship between the inventors. The ties (or edges, links) show the interconnectedness and the distance between the innovators. It is possible to use several different measures to characterize a network. Here we will especially use measures such as; Network Density, Geodesic Distance, Network Centrality, Isolates, Components and Size of the Largest Component. Details concerning the methodology can be found in Wassermann and Faust (1994) and Marsden (1990).[2]

For example, assume that there are seven inventors (see Figure 1). Three of them (inventor 1, 2, and 3) collaborate and have patent A. That is, there exist direct ties between them. One of the inventors (3) also has patent C by himself and patent D with another inventor (7). Hence, besides the direct ties between inventors 1, 2, and 3 there also exist an indirect tie between inventors 1 and 7, and 2 and 7 through inventor number 3. Two other inventors (4 and 5) have patent B. One inventor (6) does not collaborate with anybody and have one patent (E).

FIGURE 1 IN HERE

The network density is defined as the number of existing ties between nodes divided by the maximum number of ties (nodes*(nodes-1)). This means that there are eight ties between four nodes and two ties between two nodes. The total number of ties is ten; hence, the network density is equal to almost one quarter. The number of components is equal to three; hence, two disconnected networks exist plus one isolate. Note that two one-inventor patent exists but only one isolate. The size of the largest component is equal to the number of nodes in the largest network, in this case four. The geodesic distance between two nodes in a component is the minimum number of ties between them. Centrality of the network is a measurement of how much the network revolves around a node.

3.2. The measurement of networking

We are using three different measurements (proxies) to estimate the degree of networking (see the lower part of Figure 1). The first measurement (*NW1*) is a variable indicating if the number of inventors, with a direct tie to one another, is larger than 1. Hence, it is a binary variable that is equal to one if the patent is not a one-inventor patent, else zero.

The second measurement (*NW2*) is the exact number of inventors with a direct tie to one another. Of course, these measures for networking (direct ties between inventors based on applicant information on patents applications) could be weak proxies, as they do not say anything about the strength of collaboration, or the breadth of the network that was involved in achieving the innovation. On the other hand, Singh (2005) argues that inventors on the same patent application works together intensively over a long period and, therefore, the co-inventor information on the patent application capture the most important ties between all (formal or informal) inventors.

The third measurement (*NW3*) tries to overcome some of that problem in the two first proxies, by utilizing the concept of social network analysis. The measure is equal to the number of inventors with both a direct and an indirect tie to each other, that is, the size of the component. The *NW1* variable for patent A, in our example, is equal to 1 and the *NW2* variable is equal to 3. The *NW3* variable is on the other hand equal to 4 as also inventor 7 is included in the network, even if he is not involved directly in patent A.

3.2. Discrete-time hazard model

---

[2] Ucinet has been used as software for the social network analysis (see Borgotti et al, 2004).

In our data set, patent status (approved vs. not approved) is given annually. Accordingly, we specify a discrete-time duration model to represent the patent approval process. As the data are limited in time, we have a right-hand censoring problem. The duration model provides an efficient solution to this problem. The basic discrete-time duration model can be regarded as a sequence of binary choice problems and is relatively simple to estimate (Kennedy, 2003). We are using the procedure outlined in Jenkins (1995).[3] After the data set has been re-organized, a logit regression model can be used and usual interpretation is applied (Jenkins, 1995, and Shumway, 2001). The logit regression model is equivalent as estimating a Prentice-Gloeckler (1978) hazard model without taken care of potential unobserved heterogeneity.[4] More precisely, the discrete-time hazard rate *h(t)* can be generalized as (see Dor and Friedman, 1994, and Jenkins, 1995):

$$h_{i,t}^{*} = \left(\frac{h_{i,t}}{1 - h_{i,t}} \right) = \alpha_t + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_n X_{n.i} + \varepsilon \qquad (1)$$

where the dependent variable h is the probability of obtaining a patent in period t. The relative weight of the factors in predicting the outcome is given by the coefficients. The interpretation of the coefficient is as a multiplicative effect on the odds ratio. In the case of a dichotomous explanatory variable, for instance product type, e^{β} the estimate of the odds-ratio of having the outcome for, say, biochemistry is compared with the default sector. The baseline hazard can be represented as a continuous variable or as separate intercept, one for each period. The parameters α, β_1, ..., β_k, in the model, are estimated by maximum likelihood method.

3.3. The independent variables

The *X*'s are vectors of explanatory variables that in our case does not vary in time. As discussed, the first variables included in the matrix *X* are different proxies for networks. However, to be able to isolate the networking effect we need to control for other determinants.

[3] We are using the pgmhaz routine in STATA created by Stephen P. Jenkins (Jenkins, 1997).
[4] However, we have also used a Prentice-Gloeckler (1978) hazard model with unobserved heterogeneity, but there are no differences in parameter estimates.

As firms and industry differ in their patenting propensity (Griliches, 1990), but not necessarily in their approval rates, we are controlling for the product types. Furthermore, Fritsch and Lucas (2001) stress that industrial sector is an important determinant of networking, that is, some industries are more likely to cooperate than others are. Hence, if we are not controlling for the product types there is a potential risk we interpret the product type effect as a networking effect. The product type variable is represented as a binary variable.

Moreover, Freeman and Soete (1997), Adams et al (2001), Cassiman and Veugelers (2002), and Freel (2002) all highlight that firm size and in-house R&D are drivers of networking and determinants of success in innovation. We do not have critical information of firm level R&D expenditures. However, we do have the information if it is a private person or a corporation that have filed the application. The latter is a proxy for higher R&D expenditures. Moreover, we have the information of the approximate size of the corporation measured in total turnover. Our proposition (not testable) is that larger companies have larger R&D expenditures. Hence, we are going to use different proxies for networking and size of the company together with product types as independent explanatory variables in our discrete-time hazard model.

4. Data and Descriptive Analysis

4.1. The data

For the purpose of this study, we define innovation as commercial patents applications or awarded patents in Sweden. The data are based on applied or approved patents registered to the Swedish Patents and Registration Board (PRV) or the European Patent Office (EPO). Each patent, in our database, has the information on the application firm and their address and all the inventors with home address. As in

Balconi et al (2004), Ejermo and Karlsson (2006), and Cantner and Graf (2006), we implicitly assume that the inventors, on the patent application, know one another and share knowledge. Furthermore, we assume they share information with other researchers in the broader definition of network that we are using.

TABLE 1 IN HERE

4.2. The descriptive statistics

The data set is based on the patent and includes only applied patents in 1994. Some of the applied patents have been approved, and we have recorded all approvals over the years 1994 to 2001. This data set has been supplemented with data concerning the product type of the patent, number of inventors with direct and indirect ties, information whether or not it is corporation, and size of the firm (dummy for large capitalization firms and turnover).[5] The total number of application amounts to 3,815.

FIGURE 2 IN HERE

In total, 42 percent of the total number of applications in the year of 1994 was approved over the period 1994 to 2001. Most of them were awarded within 3-4 years from the application year (see Figure 2).

TABLE 2 IN HERE

The number of inventors per patent is around 1.5 in 1994. This number is lower compared to Ejermos and Karlssons (2006) and Balconis et al (2004) estimates concerning Sweden and Italy (around 1.75-1.90 inventors per patent). However, they all investigated a longer period and the variation over time can be substantial. There is a substantial variation in the number of inventors among industry sectors. For example, the average number of inventors per patent in the sector of *Organic chemistry* is 2.3 compared to only 1.2 in the sector of *Sports and amusements*.

[5] The definition of product type is original based on the International Patent Classification system created in 1997. However, we will use the same classification as in Andersson et al (2005), which is a classification that is closer related to economic activities. The classification we are using can be found in the Appendix and in Table 1.

The number of one-inventor application is around 70 percent. That means that around one third of the applications are a result of collaboration. Most of them are intra-firm collaborations. Corporations file more than half of the applications and around one third of the applications come from the large companies.

4.3. The social network analysis

The social network analysis presented in the present paper is limited. Our main objective is to estimate a measure of networks. In the table below, some network statistics are presented.

TABLE 3 IN HERE

The total number of inventors is not the unique number of inventors. The unique number of inventors has been identified by using the name of the inventors. If the inventor has a unique name, she/he is considered to be a unique inventor. If two inventors, with the same name, have a different job and home address, they are considered to be unique. If the have the same job address and/or home address, they are not considered to be distinct different. The total number of inventors is equal to 5,630 inventors. Almost 75 percent of them are unique.

The number of patents is equal to 3,815 and the total number of components (including the isolates) amounts to 2,674. Around 50 percent of the applied patents are isolates, that is, much fewer than the one-inventor patents. The percentage isolates are the same among the approved patents. The size of the components is similar regardless if the patent application is approved or not. However, this is not true if we look at different industries. For example, within the *IT* sector, almost 7 percent of the components include more than five inventors, but with the sector *Mechanical Engineering,* it is as low as 1 percent.

The network density is very low. However, the statistics reveals that the network density is higher among those applications that were approved than those that have not been approved. It is also interesting to note that the network density varies substantially among industries. The density is highest within the *IT* sector.

The average geodesic distance is almost the same regardless if the patent has been approved or not and regardless industry. However, the network centralization seems to vary. Not approved patents are more centralized than approved patents, that is, network centrality is inversely related to approval rates. Moreover, it seems that the centralization is larger within the *IT* sector.

4.4. Networks and approval rates

What about the relationship between network, product type and approval rates? In Table 4 are the three measurements of networking related to the product types.

TABLE 4 IN HERE

The relationship ship is clear. More than 36 percent of the approved patents are a result from networking compared to less than 25 percent among the not approved patents. The average number of inventors in the research team on the application is around 1.6 for the approved patents but only 1.4 for the not approved. Using the broader network definition, the numbers are 2.5 and 2.0, respectively.

The variation among industries is substantial. For example, within the *Metallurgy* sector, almost 57 percent of the approved patents were a result from collaboration with project teams with more than one inventor compared to only 24 percent of the not approved applications. Moreover, within the *IT* sector, the average component size is equal to 3.3 inventors per approved patent but only 1.9 inventors per not approved patent. The above giving us reason to believe that networking do have an effect on the probability to succeed. However, in an attempt to control for such thing as firm size and handling the censoring problem, a discrete-time duration (hazard) model will be estimated in the next section.

5. Econometrics Analysis

5.1. The basic model

In the discrete-time duration models, presented below, the dependent variables will indicate whether the patent application has been successful. As baseline hazard is both a continuous year variable and separate year intercept tested. We transformed, by natural logarithm, all continuous variables because of apparent outliers.

TABLE 5 IN HERE

One interesting result of the analysis is of course that networking has a positive affect on approval rates. The result indicates that patent application with more than one inventor have a higher probability to be approved. The parameter estimates concerning the dummy for networking (*NW1*) is highly significant. Furthermore, it is not only the existence of network that increases the likelihood to be approved. More inventors seem to increase the likelihood that the application will be approved. However, the estimated parameter concerning the *NW2* variable (number of inventors with a direct tie to each other) is not statistically significant. On the other hand, the coefficient concerning the *NW3* variable (number of inventors with direct and indirect ties to each other) is significant suggesting that the broader definition of networking is more important in explaining the success rates. Hereafter, we will use the *NW3* variable together with baseline hazard rates represented as separate year effect.

The estimates are robust in the sense that the magnitude and the significance are stable even if we control for the proxies concerning R&D expenditures (the size variables). With no exception, the likelihood to succeed is higher for corporations (or corporations are more likely to file patent application of better quality that is more likely to be approved). Even though private persons are less likely to collaborate with others, controlling for corporations and large company effects, networking do increase the likelihood to be approved.

The interpretation of the individual estimate (*NW1*) is that, everything else equal, that the odds getting the application approved are 1.21 (the odd ratios = exp(0.19)) times higher if the patent application is a result of research collaboration. An application from a corporation is more than twice as likely to be approved. The odds getting a corporation application approved are 2.41 times higher than a non-corporation application.

5.2. Variation among industries

The models, presented in Table 6, explore the possibility that product type have an influence on the likelihood to be approved (the first model) and that product type have an influence how networking is related to approval rates (model two to four).

TABLE 6 IN HERE

The first model includes all observation and a dummy variable for each product type. Overall, it does not appear to be any difference among product types when it comes to the likelihood to get an application approved. The only exception is the *Organic Chemistry* and *Biochemistry* sectors, where it is less likely to succeed compared to the other product types. It is interesting to observe that the parameter concerning the *NW3* variable is larger in magnitude and more significant if we are controlling for product type. The economic interpretation is that if the number of inventors increases by one percent, the hazard of getting the application approved will increase by 13 percent.[6]

The following three models explore the relations between success rates and networking within three specific fields – *Performing operations*, *Mechanical engineering*, and *IT*. Research collaboration seems to be most important in the sectors *Performing operations* and *IT* and of less importance in *Mechanical engineering*. The latter result is in line with, for example, Rothwell et al (1974).

5.3. Regional variation and variation by size of company

In the models presented in Table 7, we have divided the data by size of the labor market and the corporation. We are analyzing the relationship within the three largest metropolitan areas with the rest of Sweden. Moreover, we are exploring the relationship between networking and success rate among corporations and among large corporations.

TABLE 7 IN HERE

---

[6] As the independent variable *NW3* is in the form natural logarithm, the interpretation of the parameter is equal to the percentage change in the hazard, given a percent change in number of inventors in the research network.

The results suggest that innovative networking is not that important in the three large metropolitan areas. Research projects, with more than one inventor, do not increase the probability to get the application approved. However, the parameter is highly significant in the rest of the country.

Furthermore, the result also appears to indicate that large corporations have a larger effect on the probability in the denser labor markets. The results seem to support the hypothesis by Johansson and Quigley (2004) that networking in dispersed areas can act as substitutes to agglomeration economies in other more dense areas.

If we analyze only corporations or only large firms, the results are not altered. Networking seems to be important regardless of size of the firm.

6. Conclusion

Both earlier empirical findings and theoretical analyses indicate that the firm performance is better among companies that cooperate in research. Moreover, some results suggest that cooperative companies are more likely to be more innovative than non-cooperative companies are, and they participate in more projects. Finally, earlier result suggests that large project team is more likely to succeed, that is, intra-firm networks are important for the success of the innovation.

Our hypothesis is that intra-firm collaborative research produces research of higher quality. We investigated patent data and approval rates from 1994-2001 in Sweden. The main result clearly shows that researchers collaborating in innovation networks improve the efficiency of the innovation process by getting more patents applications approved even after controlling for product type and size of the company. The odds getting a patent application approved are in the range 1.1 to 1.5 times better if an application is a result from research collaboration. Furthermore, the results indicating that the size of the research team is more important in the *IT* sector and outside the three large metropolitan areas. For example, the hazard getting a patent application approved within the IT sector is 52 percent higher if the team size increase by 1 percent compared to only 13 percent within all sectors.

One implication from the result is that innovation hubs can play an important role and both increasing the quality of the research and increase the innovativeness, especially outside the largest metropolitan areas and in markets where the *IT* sector is strong. However, as Meagher and Rogers (2004) point out, it can be hard to duplicate the success of Silicon Valley.

Further research should focus on elaborating the concept of networking and not only using the information on the patent application. For example, further research should focus on collecting new information about the nature of the networks, both informal and formal ties, directly from the inventors by a questionnaire. Furthermore, the questionnaire could be used to collect information on individual projects R&D expenditures. In the present research, we only have the information about the size of the company. Although this is important in explaining the success rates, the size of the research project could be very important.

Appendix: Classification

Industry	New-code	IPC-code
Human necessities	C1	A-A61-A63+C05
Medical or veterinary science; hygiene	C2	A61
Sports, games; amusements	C3	A63
Performing operations	C4	B-B60-B61-B62-B63-B64-B82
Transporting	C5	B60 to B64
Nanotechnology	C6	B82
Chemistry; metallurgy	C7	C-C05-C07-(C12M to C12S)
Organic chemistry	C8	C07
Biochemistry	C9	C12M to C12S
Textiles; paper	C10	D
Fixed constructions	C11	E
Mechanical engineering; lighting, heating; weapons	C12	F
Physics	C13	G-G05-G02-G06-G09C-G11
Information technology	C14	G02+G06+G09C+G11+H04
Controlling; regulating	C15	G05
Electricity	C16	H-H04

References

Acs, Z., 2000. *Regional Innovation, Knowledge and Global Change*. Pinter, London.

Adam, J., Chiang, E. and Starky, K. (2001). Industry-University Cooperative Research Centers. *Journal of Technology Transfer*, 26, 73-86.

Andersson, R., Quigley, J. and Wilhelmsson, M. (2005). Agglomeration and the Spatial Distribution of Creativity. *Papers in Regional Science*. Vol. 84(3), 445-464.

Balconi, M., Breschi, S. and Lissoni, F. (2004). Networks of inventors and the role of academia: an exploration of Italian patent data. *Research Policy*, Vol 33, 127-145.

Borgotti, S.P., Everett, M.G., and Freeman, L.C. (2002). Ucinet for Windows: Software for Social Network Analysis. Analytic Technologies, Harvard.

Breschi, S. and Lissoni, F. (2006). "Cross-firm" Inventors and Social Networks: Localised Knowledge Spillovers Revisited. *Annals d'Economie et de Statistique*, 79-80.

Cantner, U. and Graf, H. (2006). The Network of innovators in Jena: An application of social network analysis. *Research Policy*, vol 35, 463-480.

Cassiman, B. and Veugelers, R. (2002). R&D Cooperation and Spillovers: Some Empirical Evidence from Belgium. *The American Economic Review*, 92(4), 1169-1184

Dor, A. and Friedman, B. (1994). Mergers of Not-for-profit Hospitals in the 1980s: Who Were the Most Likely Targets? *Review of Industrial Organization*, Vol.9, 393-407.

Ejermo, O. and Karlsson, C. (2006). Interregional Inventor Networks as Studied by Patent Coinventorships. *Research Policy*, vol 35, 412-430.

Freel, M.S. (2002). Sectoral patterns of small firm innovation, networking and proximity. *Research Policy*, vol 32, 751-770.

Freeman, C. and Soete, L. (1997). *The Economics of Industrial Innovation*. The MIT Press, Cambridge, Massachusetts.

Fritsch, M. (2001). Co-operation in Regional Innovation Systems. *Regional Studies*, 35(4), 297-307.

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

Griliches, Z. (1979). Issues in assessing the contribution of R&D to productivity growth. *Bell Journal of Economics*, Spring, Vol.10, 92-116.

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature*. Vol.27, 1661-1707.

Jaffe, A. (1986). Technological opportunity and spillovers of R&D: Evidence from firms patents, profits and market value. *American Economic Review*, Vol.6, 984-1001.

Jaffe, A. and Trajtenberg, M., ed. (2002), *Patents, Citations & Innovations. A Window on the Knowledge Economy,* The MIT Press, Cambridge, Massachusetts, London, England.

Jenkins, S.P. (1995). Easy Estimation Methods for Discrete-Time Duration Models. *Oxford Bulletin of Economics and statistics*, Vol.57, 129-138.

Jenkins, S.P. (1997). Discrete time proportional hazards regression. *STATA Technical Bulletin*. STB-39.

Johansson, B. and Quigley, J. (2004). Agglomeration and networks in spatial economies. *Papers in Regional Science*. Vol.83, 165-176.

Kennedy, P. (2003). *A Guide to Econometrics*. The MIT Press, Cambridge, Massachusetts.

Marsden, P.V. (1990). Network Data and Measurement. *Annual Review of Sociology*, Vol.16, 435-463.

Meagher, K. and Rogers, M. (2004). Network density and R&D spillovers. *Journal of Economic Behavior & Organization*. Vol.53, 237-260.

Meyer, B.D. (1990). Unemployment insurance and employment spells. *Econometrica*, Vol.58(4), 757-782.

Owen-Smith, J. and Powell, W.W. (2004). Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community. *Organization Science*. Vol.15:1, 5-21.

Prentice, R. and Gloeckler, L. (1978). Regression analysis of grouped survival data with application to breast cancer data. *Biometrics*, Vol.34, 57-67.

Romer, P. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, Vol.94(5), 1002-1037.

Rothwell, R., Freeman, C., Horsley, A., Jervis, V.T.P., Robertson, A.B. and Townsend, J. (1974). SAPPHO updated - project SAPPHO phase II. *Research Policy*, Vol.3, 258-291.

Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business*, Vol.74(1).

Singh, J. (2005). Collaborative Networks as Determinants of Knowledge Diffusion Patterns. *Management Science*, Vol.51(5), 756-770.

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

Storper, M. and Venables, A. (2004). Buzz: face-to-face contact and the urban economy. *Journal of Economic Geography*, Vol.4, 351-370.

Stuart, T.E. (2000). Interorganizational Alliances and the Performances of Firms: A Study of Growth and innovation Rates in a High-Technology Industry. *Strategic Management Journal*, Vol21, 791-811.

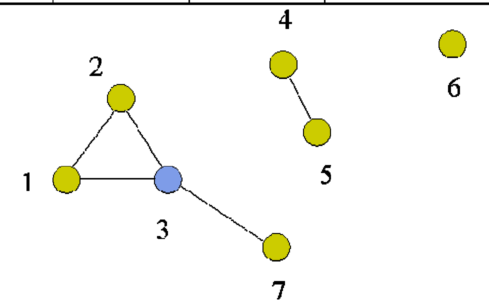
Surowiecki, J. (2004). *The Wisdom of Crowds*. New York: Random House.

Walker, G., Kogut, B. and Shan, W. (1997). Social Capital, Structural Holes and the Formation of an Industry Network. *Organization Science*, Vol. 8(2), 109-125.

Wassermann, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications.* Cambridge University Press, Cambridge.

Figure 1. Social Network Analysis.

Patent	A	B	C	D	E
Inventor	1,2,3	4,5	3	3,7	6



Patent	NW1	NW2	NW3
A	1	3	4
B	1	2	2
C	0	1	4
D	1	2	4
E	0	1	1

Table1. Variable definition.

Variable name	Description	Unit	Total	Percent
Patents				
Number of patents	Patent application in 1994	Number	3815	
Approval rates	Approved patents 1994-2001 out of application 1994	Binary	2533	66.4
Proxies for networking				
NW1	Dummy if more than 1 inventor	Binary	1113	29.2
NW2	Number of inventors per patent	Number	1.476	
NW3	Number of inventors per component	Number	2.212	
Size				
Corp	Corporation	Binary	2211	57.9
Turnover	Turnover (the 100 largest corporations)	SEK (million)	109,951	
Large	Large market capitalization firms	Binary	832	21.8
Product type				
C1	Human Necessities	Binary	398	10.4
C2	Medical science	Binary	409	10.7
C3	Sport; games	Binary	70	1.8
C4	Perf. Oper.	Binary	759	19.9
C5	Transporting	Binary	298	7.8
C6	Nanotechnology	Binary	0	0
C7	Metallurgy	Binary	133	3.5
C8	Organic chem.	Binary	38	1.0
C9	Biochemistry	Binary	30	0.8
C10	Textiles, Paper	Binary	99	2.6
C11	Constructions	Binary	346	9.1
C12	Mechanical Engineering	Binary	433	11.3
C13	Physics	Binary	321	8.4
C14	IT	Binary	256	6.7
C15	Controlling	Binary	17	0.4
C16	Electricity	Binary	205	0.5

Figure 2. Approvals over time.

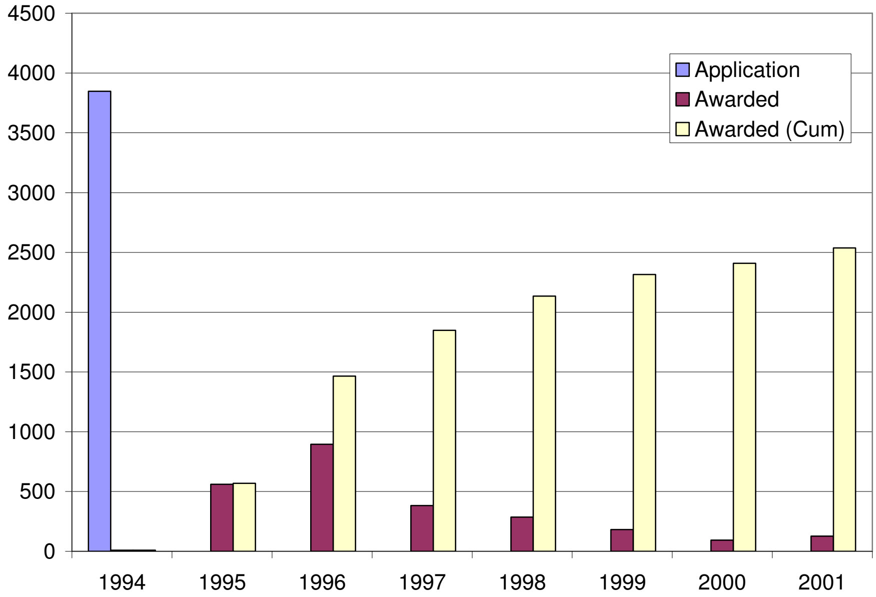


Table 2. Descriptive statistics.

Product type	Patents	Approved	Approval rates (percent)	Inventors	One-inventor	Corporation	Large Cap firms
Human Necess.	398	134	33.7	491	329	138	22
Medical science	409	164	40.1	713	227	287	148
Sport; games	70	18	25.7	84	57	16	1
Perf. Oper.	759	374	49.3	1065	544	473	141
Transporting	298	111	37.2	361	248	126	40
Metallurgy	133	66	49.6	214	84	88	30
Organic chemistry	38	5	13.2	92	14	27	19
Biochemistry	30	1	3.3	62	15	19	11
Textiles, Paper	99	57	57.6	205	49	78	42
Constructions	346	145	41.9	432	286	158	19
Mech. Engin.	433	197	45.5	589	326	267	87
Physics	321	133	41.4	514	211	184	54
IT	256	118	46.1	454	156	186	134
Controlling	17	9	52.9	25	12	13	6
Electricity	205	91	44.4	325	142	149	77
Total	3815	1625	42.6	5630	2702	2211	832

Table 3. Social network descriptive.

	All	Approved patents	Performing operations	IT	Mechanical
Patents	3,815	1,625	759	256	433
Inventors	5,630	2,533	1,065	454	589
Unique inventors	4,213	2,092	889	352	507
Ties	6,586	2,854	888	874	436
Network density (x100)	0.021	0.045	0.078	0.425	0.126
Components	2,674	1,274	618	190	364
Isolates	1,904	834	445	283	273
One-inventor patents	2,702	1,034	544	156	326
Largest size	28	13	13	10	8
Number of components with more than 5 inventors	84	46	10	13	4
Average geodesic distance	1.057	1.053	1.049	1.023	1.027
Network centralization	0.19%	0.09%	0.18%	1.35%	0.53%

Table 4. Descriptive statistics – Network size.

Product type	Approved			Not Approved		
	NW1 (percent)	NW2	NW3	NW1 (percent)	NW2	NW3
Human Necess.	26.9	1.38	1.74	12.5	1.16	1.34
Medical science	54.4	1.62	4.36	45.2	1.82	3.58
Sport; games	24.8	1.28	1.78	11.9	1.17	1.83
Perf. Oper.	31.0	1.51	2.28	19.5	1.30	1.63
Transporting	41.4	1.29	1.57	29.3	1.17	1.42
Metallurgy	56.8	1.74	2.71	23.9	1.48	2.79
Organic chemistry	22.2	2.80	3.40	37.5	2.36	5.24
Biochemistry	38.5	3.00	3.00	24.6	2.03	2.93
Textiles, Paper	43.3	2.02	3.21	45.3	2.14	2.71
Constructions	22.2	1.35	1.54	17.3	1.17	1.49
Mech. Engin.	34.8	1.46	2.22	22.1	1.28	1.69
Physics	25.2	1.57	2.28	11.8	1.62	2.57
IT	45.5	2.12	3.27	28.4	1.48	1.86
Controlling	60.0	1.33	1.33	63.6	1.62	1.75
Electricity	100.0	1.61	2.54	48.3	1.56	1.89
Total	36.4	1.56	2.46	23.8	1.41	2.03

Table 5. Basic discrete-time hazard model.

	Model A Coefficient	Model B Coefficient	Model C Coefficient	Model D Coefficient
Baseline hazard rates				
Year	-0.0479	-	-	-
	(-4.10)			
T94	-	Default	Default	Default
T95	-	4.2541	4.2534	4.2537
		(10.32)	(10.32)	(10.32)
T96	-	4.9708	4.8697	4.8704
		(11.84)	(11.84)	(11.84)
T97	-	4.1210	4.1186	4.1196
		(9.95)	(9.94)	(9.94)
T98	-	4.0153	4.0118	4.0130
		(9.66)	(9.65)	(9.66)
T99	-	3.7056	3.7020	3.7040
		(8.86)	(8.85)	(8.85)
T00	-	2.7710	2.7674	2.7696
		(6.41)	(6.40)	(6.40)
T01	-	3.1133	3.1095	3.1123
		(7.30)	(7.29)	(7.29)
Networking				
NW1	0.1912	0.1936	-	-
	(3.35)	(3.31)		
Ln(NW2)	-	-	0.0842	-
			(1.46)	
Ln(NW3)	-	-	-	0.0858
				(2.29)
Size				
Company	0.8591	0.8739	0.8880	0.8829
	(13.42)	(13.45)	(13.67)	(13.61)
Ln(Turnover)	0.0186	0.0185	0.0208	0.0190
	(3.38)	(3.26)	(3.66)	(3.30)
Constant	92.3521	-7.1810	-7.1572	-7.1672
	(3.96)	(-17.45)	(-17.40)	(-17.42)
R^2	0.0319	0.1198	0.1191	0.1193

Note: t-values within parentheses.

Table 6. Variation among industries.

	All		Performing Operation (C4)		IT (C14)		Mechanical (C12)	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff	t-value
Baseline Hazard Rate								
T94	-		-		-		-	
T95	4.2612	(10.33)	3.8118	(5.31)	2.6234	(4.22)	2.7748	(6.37)
T96	4.8890	(11.88)	4.4309	(6.19)	3.5742	(5.90)	3.1845	(7.35)
T97	4.1435	(10.00)	3.8691	(5.36)	2.5918	(4.04)	2.3853	(5.18)
T98	4.0466	(9.74)	3.8852	(5.37)	-		2.2944	(4.87)
T99	3.7434	(8.94)	3.9183	(5.40)	1.0409	(1.26)	2.1052	(4.32)
T00	2.8108	(6.50)	2.7578	(3.63)	1.7827	(2.49)	0.9879	(1.61)
T01	3.1562	(7.37)	1.9337	(2.36)	3.3883	(5.40)	-	
Networking								
Ln(NW3)	0.1353	(3.47)	0.2025	(2.27)	0.5217	(3.81)	0.1385	(1.16)
Size								
Corporation	0.8713	(13.20)	0.8158	(5.98)	1.0628	(2.85)	1.0634	(5.28)
Ln(turnover)	0.0277	(4.58)	0.0270	(2.07)	0.0189	(0.82)	0.0453	(2.69)
Product type								
C1	-0.6104	(-0.73)	-		-		-	
C2	-0.9237	(-1.11)	-		-		-	
C3	-0.9060	(-1.05)	-		-		-	
C4	-0.4188	(-0.51)	-		-		-	
C5	-0.5408	(-0.65)	-		-		-	
C6	-		-		-		-	
C7	-0.5706	(-0.68)	-		-		-	
C8	-2.3871	(-2.53)	-		-		-	
C9	-3.6906	(-2.84)	-		-		-	
C10	-0.3541	(-0.42)	-		-		-	
C11	-0.4152	(-0.50)	-		-		-	
C12	-0.4881	(-0.59)	-		-		-	
C13	-0.6197	(-0.75)	-		-		-	
C14	-0.7852	(-0.94)	-		-		-	
C15	-0.3848	(-0.43)	-		-		-	
C16	-0.7546	(-0.91)	-		-		-	
Constant	-6.6233	(-7.19)	-20.9448	(-32.15)	-6.3716	(-9.78)	-5.6355	(-12.82)
R^2	0.1277		0.1907		0.1835		0.1466	

Note: t-values within parentheses

Table 7. Regional variation and variation among different size of company.

	Three largest		Rest		Companies		Large	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff	t-value
Baseline Hazard Rate								
T94	-		-		-		-	
T95	5.1822	(5.16)	3.9037	(8.60)	3.9087	(9.44)	4.5820	(4.55)
T96	6.0116	(6.00)	4.3404	(9.58)	4.7092	(11.42)	5.7986	(5.78)
T97	5.0813	(5.05)	3.7713	(8.24)	3.8885	(9.33)	4.8588	(4.82)
T98	5.0054	(4.97)	3.6515	(7.94)	3.7591	(8.97)	4.8312	(4.78)
T99	4.7316	(4.69)	3.3210	(7.14)	3.4633	(8.19)	3.9091	(3.82)
T00	3.8522	(3.78)	2.3347	(4.75)	2.5558	(5.77)	3.9235	(3.83)
T01	4.4971	(4.45)	2.2375	(4.50)	3.1309	(7.28)	4.6094	(4.54)
Networking								
Ln(NW3)	0.0787	(1.54)	0.1461	(2.66)	0.1153	(2.84)	0.1127	(1.97)
Size								
Corporation	0.8598	(8.76)	0.8520	(9.80)	-		-	
Ln(turnover)	0.0342	(4.34)	0.0137	(1.56)	0.0216	(3.69)	0.0527	(2.02)
Product type								
C8	-1.2649	(-2.75)	-		-1.8560	(-3.64)	-3.9662	(-3.04)
C9	-2.8484	(-2.83)	-		-2.8995	(-2.88)	-2.5069	(-2.48)
Constant		(-8.33)	-6.5536	(-14.48)	-6.0601	(-14.77)	-7.3851	(-7.07)
R^2	0.1305		0.1907		0.1099		0.1349	

Note: t-values within parentheses