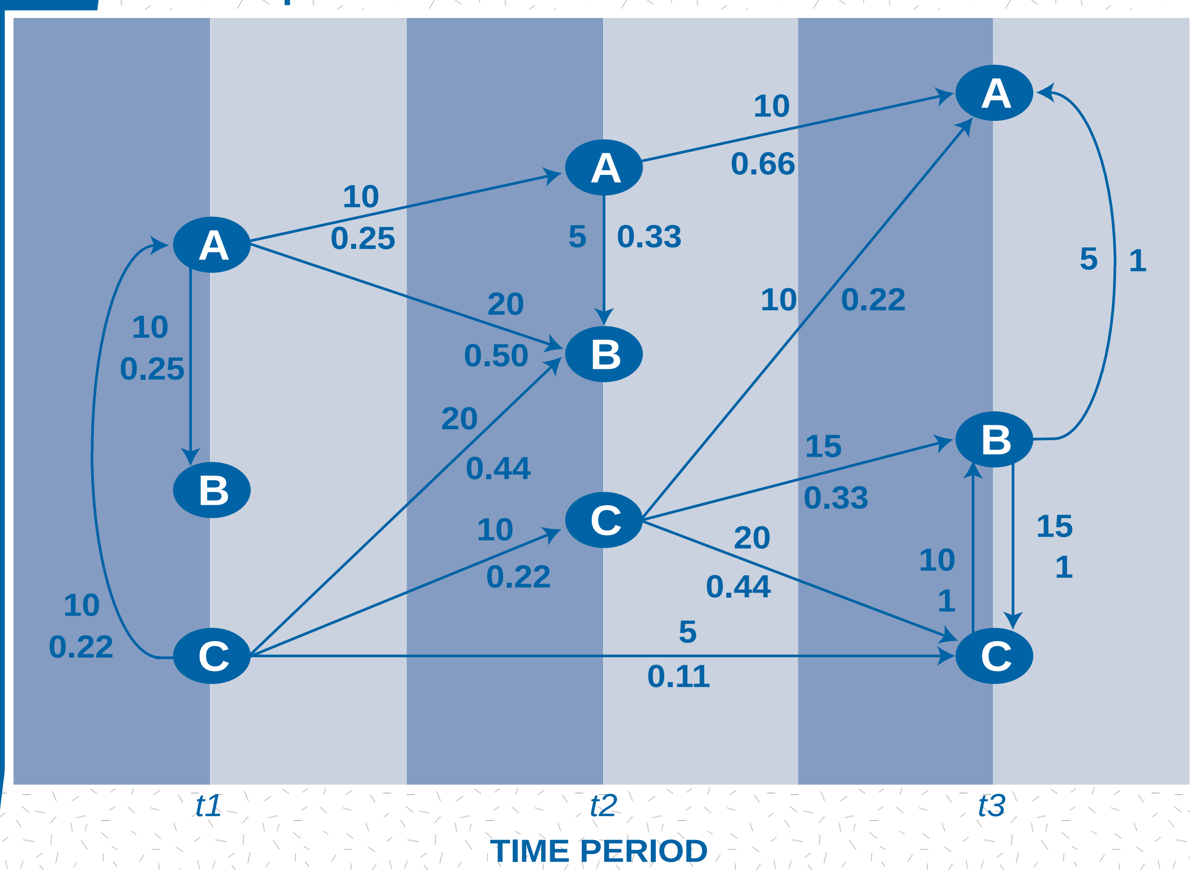
NIST GCR 06-895

ATP and the U.S. Innovation System: A Methodology for Identifying Enabling R&D Spillover Networks

Michael S. Fogarty
Amit K. Sinha
Adam B. Jaffe

R&D Spillover Network: *Cited Behavior Fractions*



October 2006

National Institute of Standards and Technology • Technology Administration • U.S. Department of Commerce

About ATP's Economic Assessment Office

The Advanced Technology Program (ATP) is a partnership between government and private industry to conduct high-risk research to develop enabling technologies that promise significant commercial payoffs and widespread benefits for the economy.

Since the inception of ATP in 1990, ATP's Economic Assessment Office (EAO) has performed rigorous and multifaceted evaluations to assess the impact of the program and estimate the returns to the taxpayer. To evaluate whether the program is meeting its stated objectives, EAO employs statistical analyses and other methodological approaches to measure program effectiveness in terms of:

- Inputs (program funding and staffing necessary to carry out the ATP mission)
- Outputs (research outputs from ATP supported projects)
- Outcomes (innovation in products, processes, and services from ATP supported projects)
- Impacts (long term impacts on U.S. industry, society, and economy)

Key features of ATP's evaluation program include:

- Business Reporting System, a unique online survey of ATP project participants, that gathers regular data on indicators of business progress and future economic impact of ATP projects.
- Special Surveys, including the Survey of Applicants and the Survey of Joint Ventures.
- Status Reports, mini case studies that assess ATP projects on several years after project completion, and rate projects on a scale of zero to four stars to represent a range of project outcomes.
- Benefit-cost analysis studies, which identify and quantify the private, public, and social returns and benefits from ATP projects.
- Economic and policy studies that assess the role and impact of the program in the U.S. innovation system.
- Data Enclave to allow for analysis of innovation and entrepreneurship (Spring 2007).

EAO measures against ATP's mission. The findings from ATP surveys and reports demonstrate that ATP is meeting its mission:

- Nine out of 10 organizations indicate that ATP funding accelerated their R&D cycle.
- An ATP award establishes or enhances the expected value in the eyes of potential investors, which is called a "Halo Effect."
- ATP stresses the importance of partnerships and collaborations in its projects. About 85 percent of project participants had collaborated with others in research on their ATP projects.

Contact ATP's Economic Assessment Office for more information:

- On the Internet: www.atp.nist.gov/eao/eao_main.htm
- By e-mail: atp-eao@nist.gov
- By phone: 301-975-8978, Stephanie Shipp, Director, Economic Assessment Office, Advanced Technology Program
- By writing: Economic Assessment Office, Advanced Technology Program, National Institute of Standards and Technology, 100 Bureau Drive, Stop 4710, Gaithersburg, MD 20899-4710

NIST GCR 06-895

ATP and the U.S. Innovation System:
A Methodology for Identifying Enabling R&D Spillover Networks

Prepared for
Economic Assessment Office
Advanced Technology Program
National Institute of Standards and Technology
Gaithersburg, MD 20899-4710

By

Michael S. Fogarty
College of Urban and Public Affairs
Portland State University
fogartym@pdx.edu

Amit K. Sinha
Institute of Genomics and Integrative Biology
Mall Road, Delhi University Campus
Delhi-110007 INDIA
amit.sinha.sm@gmail.com

Adam B. Jaffe
Office of the Dean
College of Arts and Sciences
Brandeis University
ajaffe@brandeis.edu

Contract 50SBNB6C9080

October 2006

U.S. DEPARTMENT OF COMMERCE
Carlos M. Gutierrez, Secretary

TECHNOLOGY ADMINISTRATION
Robert Cresanti, Under Secretary for Technology

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
William Jeffrey, Director

ADVANCED TECHNOLOGY PROGRAM
Marc G. Stanley, Director

Abstract

Spillovers serve a central role in justifying public support for R&D, but are difficult to identify and to measure. Improving methods of identifying and quantifying R&D spillovers is an important goal for public R&D programs like ATP. Research networks or systems—the patterns of interaction and communication among firms, universities, and other laboratories—reveal the generation and exchange of scientific and technological knowledge. An implicit hypothesis links the concepts of spillovers and research networks: the closer and denser the system of linkages is among various organizations, the greater the likelihood of knowledge spillovers. This new method builds on past research of identifying and quantifying knowledge spillovers using patent citations.

Where traditional approaches treat the spillover process as linear and additive, this emerging methodology uses systems analysis and fuzzy logic to analyze R&D knowledge spillovers within networks of R&D organizations, locating researchers within a mutually interconnected network. Spillovers that occur from one researcher in an organization to another researcher in a second organization depend on both the communications between those two researchers and their organizations, as well as communications that each engages in with other researchers at other organizations. This report's approach identifies spillover patterns across organizations, technological areas, geographic regions, and industries. It is illustrated by the mapping of two research networks, one underlying micro-electromechanical systems (MEMS), and the second underlying short wavelength sources for optical recording. For example, the methodology was able to identify the top five MEMS technologies, the most influential segment of the MEMS network, the key universities, and the key regions.

The most straightforward application of this new methodology to measure knowledge spillovers is ex post evaluation. Other potential uses include analysis of the evolution of networks surrounding particular industry-based technologies, or to evaluate the evolving status of different network members.

Acknowledgments

Several ATP staff members contributed to the publication of this paper. Stephanie Shipp, Director of the Economic Assessment Office, oversaw the editing and preparation of the paper. Connie Chang and Prasad Gupte provided input into the editing of the paper. Lorel Wisniewski, ATP Deputy Director, and Brian Belanger, former ATP Deputy Director, reviewed the final version of the paper. Kathleen McTigue edited the final versions of the paper, addressed questions raised by the reviewers, and updated portions of the material. She and Janet Brumby oversaw the desktop publishing and final publishing of the report. Jo Ann Brooks designed the cover.

Executive Summary

Scholars who study the process of innovation and technological change generally agree that the public or social benefits of commercial research exceed the private benefits. Private benefits are benefits that can be captured by the firms that undertake the research. This excess of the public benefits of research over the private benefits is commonly referred to as research "spillovers, " and their existence constitutes an important part of the public policy motivation for ATP.

An important category of research spillovers is *knowledge spillovers*. Projects generating a large volume of knowledge spillovers are especially attractive candidates for ATP funding because many of those projects by their very nature may not produce sufficient profit to provide enough incentive for the private sector to undertake the work. Plus, companies that will benefit from the knowledge spillovers produced by the project may be competitors of the originating firm, and thus may erode the competitive edge and reduce the profits of the originating firm even further, making these projects less likely to attract private investment. Thus, if ATP can identify projects that are likely to produce above-average knowledge spillovers, then ATP's ability to create high social returns from its technology investments would be further enhanced. Moreover, if ex post evaluation shows that ATP-funded projects result in actual substantial knowledge spillovers, then the public benefit of the program would be clearly demonstrated.

This study explores the potential for patent citation data to be used as a proxy for research spillovers, on both a prospective and retrospective basis. The growing empirical evidence showing that patent citations can be used as a proxy for flows of scientific and technical knowledge is summarized. By using citations data, ATP would obtain a quantitative indicator of the spillover potential of proposed projects, as well as assess the magnitude of spillovers generated by funded projects after the fact. This methodology can be used for technologies that are protected by patents. It may not be as useful for technologies that are protected by trade secrets and industries that do not patent their intellectual property.

Spillovers from a specific research project do not occur in a vacuum. Rather, research occurs within an existing research and innovation system, in which established patterns of communication and influence change slowly. Organizations or technological areas or regions that are well connected within this system have generated large spillovers in the past, and are likely to continue to do so in the future. Conversely, researchers who are not well connected within the innovation system are less likely to generate large research spillovers. The spillover potential can be assessed by using systems analysis methods. For this reason, the study goes beyond previous work by using patent citations as a proxy for knowledge spillovers to explore the usefulness of systems analysis methods for mapping patterns or networks of communication and influence within the innovation system. The systems approach maximizes the value of patent citations for understanding research spillover patterns.

The systems methodology that is developed identifies research entities organized as a network existing across organizations, technologies, and geographic regions. A node in this network is a particular organization (e.g., IBM or MIT) engaged in research in a particular technology (defined by U.S. Patent and Trademark Office technology classes; for example, Dynamic Information Storage and Retrieval) in a particular geographic region (e.g., a metropolitan area, a state, or a country, depending on the focus of interest). The extent of communication between nodes is initially measured by the extent to which patents assigned to each node cite patents assigned to other nodes. The structure of the resulting communications network results from assigning the nodes to clusters in organizational, technological, and geographical space by using *fuzzy* clustering methods. That is, the membership of a node in each cluster is a matter of degree rather than an absolute assignment.

Once nodes have been organized into networks, the *system influence* of each node is measured. Each node's system influence is based on the strength of its communication with other nodes, weighted by the strength of communication of the interacting nodes with the rest of the system. The use of this system influence measure as an indicator of the spillovers likely to be generated by research performed at a given node is then explored. This measure of influence can be aggregated in any of the dimensions of interest, which allows for the ranking of the spillover potential within a particular R&D network by the overall organization, a given technology class, or a given region.

The study then uses the methodology to construct mappings of the research networks underlying two emerging technologies: micro-electromechanical systems (MEMS) and optical recording. In each case, the structure of the research network is examined in technological, geographic, and organizational space. The United States and Japan, for example, are shown to be both major sources and recipients of MEMS spillover flows. The San Francisco, Los Angeles, and Boston metropolitan regions are the primary U.S. spillover generators and recipients for MEMS research. The important role played by particular organizations in the MEMS network, such as MIT, is also illustrated.

The structure of the underlying R&D network for optical recording and its relationship to a significant joint venture funded by ATP in this area is then examined. The optical recording network is found to be dominated by U.S. firms, two of which were members of the ATP-funded joint venture. Moreover, the ATP-funded joint venture was itself managed by the National Storage Industry Consortium (NSIC), among whose members are additional influential organizations within the optical recording network. The network analysis also reveals that university and government laboratories are influential network members. Given the public research entities' influence within the network, the likelihood of spillovers due to ATP's funding of optical recording is greater if they are involved in the project. Involvement of a number of the key public research institutions in NSIC may have provided an important mechanism for the transmission of spillovers resulting from ATP cost-shared research.

The use of the methodology for looking systematically at possible patterns of spillovers across broadly defined industrial sectors is also illustrated. The MEMS research is shown, for example, to likely create significant spillovers to the auto and aerospace sectors and, to a lesser extent, to the biomedical devices industry. Optical recording research appears to produce spillovers for the aerospace and information technology sectors.

Finally, the study concludes with some initial suggestions about how the methodology might be used by ATP. First, it could be used on an ex-post basis to quantify the spillover impacts of ATP-funded projects, individually or collectively. Such an impact analysis could, in principle, measure the direct spillovers from awardees, which would be reflected in an increased rate of system citation to the awardees' patents. It could also examine the broader influence of ATP funding on the relevant R&D network by looking at overall changes in the flows of knowledge through the network. For example, ATP funding of joint ventures, in addition to generating spillovers related to the joint ventures' research output, may enhance the overall spillover capability of the network. In principle, such an increase in the overall quality of communication links could be measurable in terms of increased overall network knowledge flows.

Prospectively, quantitative measures of system influence (in the appropriate R&D networks) of the organizations, regions, and technologies represented by project proposals could provide part of the basis for assessing the likely spillover potential of proposals. This kind of analysis could be used to identify ways in which the proposed joint venture might be made more effective by expanding the membership or roster of subcontractors to ensure connection to important nodes in the relevant R&D network.

At the end of the day, the concept of knowledge spillovers remains elusive and difficult to measure. It is not possible to do better than to provide numerous proxies that have been shown to be connected to the underlying knowledge flows. There is now significant evidence that patent citations constitute one such proxy. Use of such data, particularly in conjunction with a methodology that places spillover analysis in the context of an overall R&D system, offers significant potential for providing more quantitative, systematic information about spillovers.

Note: The research and data used in this report were collected in the late 1990s.

Contents

Figures

Tables

1. Introduction

One economic rationale for the Advanced Technology Program (ATP) rests on the assumption that the public or social benefits of the research projects ATP funds far outweigh the private benefits that can be captured by the firms that undertake the research. This excess of the public benefits of research beyond the private benefits is commonly referred to as research *spillovers* (Jaffe, 1996). An important aspect of investment in R&D is that the knowledge benefits tend to spill over to others not directly involved in the original R&D work. When one company conducts research, other companies also receive benefits because the results of R&D often become more generally known through patents, publications, and other means of industry dissemination of knowledge spillovers. Because of such spillovers, when one company conducts R&D, the overall benefit is greater than what this one company receives. Since other companies also benefit, there is a strong policy rationale for encouraging R&D investment through a public program such as ATP.

Spillovers have a central role in justifying public support of R&D, but are difficult to identify and measure. Improving methods of identifying and quantifying R&D spillovers is an important goal of a public R&D program. If it is assumed that commercial research typically generates some level of spillovers, this implies that any government program seeking to foster commercial R&D will generate some spillovers. Following that logic, if ATP can identify commercial R&D projects that would not otherwise be undertaken, or would be pursued less vigorously, in the absence of ATP support, ATP could reasonably assume that funding such projects would generate R&D spillovers. On the other hand, if all commercial R&D generates some level of spillovers, it seems more straightforward and less likely to cause economic distortions in the economy to subsidize commercial R&D using tax breaks or other tax incentives.

As described in an earlier ATP study (Jaffe, 1996), there are several distinct mechanisms by which commercial R&D generates spillovers. One of these, *knowledge spillovers*, occurs because the knowledge created by research is used by other researchers to facilitate their own research projects. Unfortunately, knowledge spillovers are extremely difficult to quantify because knowledge itself is essentially unobservable, the mechanisms or pathways by which knowledge moves are not well understood, and the impact that a particular bit of knowledge has on a subsequent research project is difficult to separate out from other influences.

Perhaps for this reason, ATP analyses of spillovers have focused mainly on *market spillovers*. Market spillovers result when a new product or process generated by research creates benefits for consumers or downstream industries. For example, new materials, new tools or methods, or better parts reduce costs and improve the performance of the industries that use them. To justify projects supported by ATP, market spillover analysis has been the common method used to present the broad-based economic benefits that these projects have delivered. Economic benefits of this kind have been identified and studied in the after-the-fact evaluation of ATP projects (Long, 1999).[1]

The emphasis on market spillovers, however, raises the question of the need for public suport of technologies that create large market benefits. Why wouldn't they be supported by the private sector? That is, while it is easy to show theoretically that the firm developing a new product or process will only capture a fraction of the benefits created for consumers, if the overall benefit is very large, then that fraction may generate a large profit. Projects generating large knowledge spillovers may or may not generate profits, and thereby be attractive candidates for private financing. To the extent that the spillover recipients are competitors of the spillover-generating firm, projects that generate large knowledge spillovers are not likely to be funded privately.

An emerging method using systems analysis and fuzzy logic to analyze R&D spillovers within networks of R&D organizations is introduced in this study. This novel method holds promise for retrospective evaluation as well as prospective selection of projects with above-average spillover potential. Furthermore, the method identifies spillover patterns across organizations, technological areas, geographic regions, and industries, and permits separation of knowledge spillovers into those realized by the United States and those realized by other countries. By funding projects involving particular organizations and technologies, ATP would be able to implicitly pick networks with implications for expected social benefits. An important theoretical aspect of this methodology is that it highlights the fact that the firm's value as a source of knowledge spillovers depends on its ability to learn from its external environment.

The potential for patent citation data to be used as a proxy for research spillovers, on both a prospective and retrospective basis, is explored in some detail. As will be discussed more fully, there exists growing empirical evidence that patent citations can be used as a proxy for flows of scientific and technical knowledge (Jaffe, Trajtenberg, and Fogarty, 2000a, 2000b).

---

1. The ATP uses a prototype evaluation tool, the Composite Performance Rating System (CPRS), to measure ATP projects' success through indicator metrics. A mini-case study of every ATP project is written about five years after project completion. One of the CPRS metrics for project achievement includes knowledge creation and dissemination—that is, evidence that the project has created a significant body of new knowledge, and that it is being disseminated. Knowledge creation and dissemination is a proxy for potential knowledge spillovers. Other metrics include measures of commercialization. See ATP (2001, 2005); Long (1999); Ruegg (2006).

Although there exists a fair amount of noise in the relationship between citations and knowledge flows, the overall magnitude of knowledge flows between organizations or regions appears to be correlated in a statistical sense with the underlying knowledge flows.

Thus, analysis of patent citations can provide a valuable window on the otherwise unobservable process of research spillovers between and among research organizations. By using citations data, ATP may be able to develop a quantitative indicator of the spillover potential of firms proposing projects, as well as assess the magnitude of R&D spillovers generated by funded proposals after the fact.

Spillovers from a specific research project do not occur in a vacuum. Rather, research occurs within an existing research and innovation system, where established patterns of communication and influence change slowly. Organizations or regions that are well connected within this system have generated large spillovers in the past, and are likely to continue to do so in the future. Conversely, researchers who are not well connected within the innovation system are less likely to generate large research spillovers.

System analysis methods can be used to assess spillover potential. For this reason, this study goes beyond previous work by using patent citations as a proxy for knowledge spillovers to explore the usefulness of systems analysis methods for mapping patterns or networks of communication and influence within the innovation system. This systems approach attempts to maximize the value of patent citations for understanding patterns of research spillovers.

Note: The research and data used in this report were collected in the late 1990s.

2. Evidence on Patent Citations and Research Spillovers

PATENT AND PATENT CITATION DATA

Micro-level data on patents that include a detailed technology field, for instance, or citations to other patents, number of claims, and inventor's geographical location are becoming increasingly available. These data have enormous potential for economists who study technical change and innovation, and are the only measure of technological and innovative activity available at this level of detail. The data have been used to explore questions involving spatial spillovers (e.g., Jaffe et al, 1993), knowledge flows among firms in a research consortium (e.g., Ham, Linden, and Appleyard, 1998), spillovers from public research (Jaffe and Trajtenberg 1997), and differences between corporate and university patenting (Trajtenberg, Henderson, and Jaffe, 1997).

Some of this literature has been self-referential. While the results from using measures based on patents have been internally consistent, measures like patent counts, citation counts, "basicness" (identifying basic inventions), and so forth have not generally been validated by external economic criteria, such as revenues or profits. Some exceptions exist, however. For example, Pakes and Schankerman (1984), Schankerman and Pakes (1986), Pakes (1986), and Schankerman (1998) use patent renewal data and fees to uncover the distribution of patent values; Trajtenberg (1990b) correlates consumer welfare measures for CT scanners with citation counts; Shane (1993) conducts an analysis of the market value of a small number of large semiconductor firms and their intangible assets; Austin (1993) employs an event study of biotechnology patents and their citations; and Harhoff et al. (1999) analyze individual German inventions that are patented in the United States and Germany. All of these authors have found that patenting is correlated with economic value, albeit with substantial error, and some have found that citation counts are even more highly correlated.

The use of patent data in the economic analysis of technological change has a fairly long, if somewhat unsatisfactory, history, stretching back to the path-breaking analyses of

Schmookler (1966).[1] A patent, as a matter of definition, is a temporary legal monopoly granted to inventors for the commercial use of an invention. In principle, in order for the inventor to receive this right, the invention must be nontrivial in the sense that it must not be obvious to a skilled practitioner of the relevant technology, and it must be useful, meaning that it has potential commercial value. If a patent is granted, then an extensive public document is created that contains detailed information about the invention, the inventor, the organization to which the inventor assigns the patent property right, and the technological antecedents of the invention.

This last category of information is the focus of our study. The citations that appear in a patent document serve to identify earlier inventions whose claims are sufficiently close to the claims of the citing patent that the inventor or the patent examiner deems it necessary to identify them. By identifying the prior art that is not covered by the property right granted in the citing patent, the citations that appear in a patent serve the important legal function of delimiting the property right granted by the patent. Thus, citations contained within a patent convey information about the technological antecedents of the new technological innovation embodied within the patent document.

Viewed optimistically, patent citations can be seen as providing direct observations of knowledge spillovers in that one technological innovation explicitly identifies others as constituting the technological state-of-the-art on which it builds. Unfortunately, this optimistic view is somewhat clouded by the reality that there is substantial noise in the patent citations data (Jaffe, Fogarty, and Banks, 1998; Jaffe et al., 2000a, 2000b). In addition to cites that are included in the patent document because the inventor actually learned something important from the cited invention or its inventor, there are also citations that are placed in the patent document for strictly legal reasons. These typically arise when the patent attorney wants to avoid potential lawsuits for patent infringement by citing a patent even when the inventor does not consider it part of the prior art. There are also after-the-fact citations, which are citations to relevant patents added to the patent document after the actual invention, and teaching cites that everyone considers basic even if they are old. Finally, the patent examiner may also require the inventor to add relevant citations to further bound the scope of intellectual property rights conferred by the patent, even though the inventor may not have been aware of the patent that was added to the citations.

EMPIRICAL EVIDENCE ON CITATIONS AND SPILLOVERS

These problems notwithstanding, economic research indicates that patent citations, particularly at levels of aggregation higher than the individual patent, yield valid, if sometimes noisy, information on real knowledge spillovers. Previous research has utilized citations data in a number of ways, but this prior research can perhaps best be summarized

1. Another pioneer who deserves mention is F. M. Scherer, who also brought the study of patents to the attention of the American economics research community with his 1965 article in the *American Economic Review*. For a comprehensive review of the promise and the problems of patents as economic indicators, see the survey by Griliches (1990).

under two themes: the use of ex-post citations to infer the quality or importance of the cited inventions and the use of citation patterns to make inferences about the nature and direction of knowledge spillovers.

Trajtenberg (1990b) demonstrated that within a particular class of medical instruments (computerized tomography or CT scanners), there was a strong correlation between the ex-post citations received by certain patents and the estimated social surplus attributed to the inventions based on those patents. This result validated the link between citations as a measure of invention quality and the social welfare generated by those inventions. Trajtenberg's method for calculating social surplus is quite demanding in terms of data requirements. It is, however, a powerful vindication of the potential usefulness of citation-weighted patents as a measure of innovative output. If we assume that such a link holds in general, then we can get a potentially much less noisy measure of research output at the organization, industry, or regional level at relatively low cost simply by weighting patents by the ex-post citations received. A simple weighting by ex-post citations is not the only measure of quality that could be developed using citations. For example, in a slightly different manner, Trajtenberg et al. (1997) investigated the basicness of inventions by looking at the breadth of ex-post citations across patent classes.[2]

Interesting work by Lanjouw and Schankerman (1998) also uses citations as a proxy for patent quality. They explore the relationship between (1) citations, ownership, the number of claims, and various technology class measures and (2) the probability that a patent will be litigated. They find that the ratio of citations to claims for a patent is positively correlated with the probability that that patent will be litigated and interpret this finding to imply that more valuable patents (as measured by citations per claim) are more likely to find themselves in court.

Another line of work examines the relationship between estimates of profits generated or licensing revenues received by the owner of the patent and the number of citations subsequently received by the same patent. Several researchers have recently begun or completed analyses of this kind: Harhoff et al. (1999) surveyed German patent holders of 962 U.S. patents that were also filed in Germany. The patent holders were asked to estimate at what price they would have been willing to sell the patent right in 1980, about three years after the date at which the German patent was filed. Harhoff et al. find, first, that more valuable patents are more likely to be renewed to full term and, second, that the estimated value is correlated with subsequent citations to that patent. The most highly cited patents are very valuable, "with a single U.S. citation implying on average more than $1 million of economic value" (Harhoff et al., 1999).

---

2. The term actually used by Trajtenberg et al. is "generality." This is calculated by examining the distribution of citations received by the patent across 3-digit patent classes. To be precise, generality was defined as unity minus the Herfindahl index of concentration of citations across classes. This generality index is 0 for patents whose entire body of subsequent citations is located in the same patent class and approaches 1 for patents whose citations are extremely dispersed.

Another line of citations-related research has looked directly at citations as an indicator of knowledge spillovers. Jaffe et al. (1993) investigated the extent to which knowledge spillovers are geographically localized by examining the location of inventors of cited and citing patents. They found that the tendency of subsequent inventors to cite previous local inventions was much higher than could be explained simply by the distribution of research activity across U.S. states. Jaffe and Trajtenberg (1999) have extended this work to analyze knowledge flows between countries, adding to the evidence of geographic and firm localization of citations. In an ambitious attempt to test theories of endogenous growth based on knowledge spillovers, Caballero and Jaffe (1993) estimated a complex structural model in which knowledge spillovers increase the innovative output of current R&D efforts. Finally, Jaffe et al. (1998) have attempted to measure the private sector benefits from research undertaken by the National Aeronautics and Space Administration (NASA). Their study examines citations to a set of NASA-generated inventions, finding again that knowledge spillovers tend to occur more frequently when firms are geographically proximate to NASA labs. In all these cases, knowledge spillovers were proxied by citation patterns in the U.S. patent data.

SURVEY EVIDENCE ON CITATIONS AND SPILLOVERS

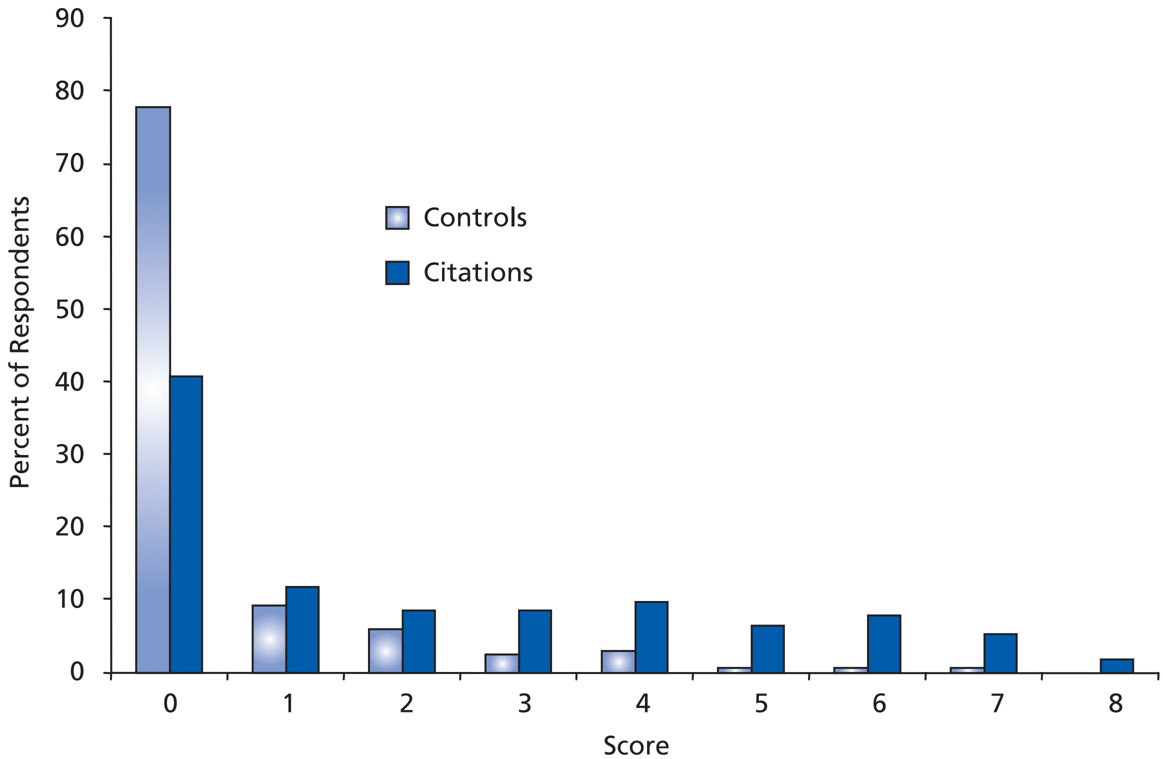
This statistical evidence has recently been supplemented by survey results exploring the extent to which inventors' perceptions of research spillovers are correlated with patent citations (Jaffe et al., 2000a, 2000b). A national survey of recent patentees was conducted to elicit their perceptions regarding the importance of their inventions, the extent of their communication with other inventors, and the relationship of both importance and communication to observed patent citations. A cohort of 1993 patentees were asked specifically about two patents they had cited, and a third "placebo" patent that was similar but which they did not cite. One of the two cited inventors was also surveyed.

The surveyed inventors reported significant communication, at least some of which suggests spillovers from the cited inventor to the citing inventor. In particular, the citing inventors reported that they were more likely to have heard of the cited patents than the placebos; they learned about them sooner; they learned of them by more information-rich modes of communication; and they were more likely to have benefited from them. The contrasts revealed by the results are illustrated in Figure 1, which is taken from Jaffe et al. (2000a). In more than half of the cases, inventors reported that they had benefited in some way from their knowledge of the cited inventions. In contrast, they reported deriving such benefits from the "placebo" patents only about 20 percent of the time. This difference and others found in the survey were statistically significant.

Figure 1 also shows that in 40 percent of the cases, the citing inventor stated that nothing useful was learned from the cited patent. One might expect that the results are affected by a human tendency to minimize the dependence of personal creations on those of others; in fact, survey responses from the *cited* inventors do provide some support for the proposition that the citing inventors' statements minimize their dependence on the cited inventions. Nonetheless, the results clearly show that a significant fraction of citations do not correspond

FIGURE 1

Distribution of Composite Spillover Score



to research spillovers. Indeed, it is clear that in a significant fraction of cases, the citing inventors did not even know about the cited inventions, or learned about them only while completing the patent application process.

Overall, then, it is clear that for any given patent citation there is a nontrivial chance that no spillover occurred. It is still true, however, that the probability of a spillover, conditional on a citation being observed, is significantly greater than the unconditional probability. This is quite consistent with there being a correlation, but not a perfect correspondence, at the level of organizations or regions, between the rate of citation and the rate of research spillover.

CITATIONS FROM A SYSTEMS PERSPECTIVE

All of the above analysis considers spillovers and citations as a “pair-wise” phenomenon between researchers, research organizations, or regions. But innovation is an inherently cumulative phenomenon that occurs within the context of an overall innovation system. The spillovers that one organization gets from another depend not only on the communication between the two organizations but also on the communication that each engages in with other organizations.

One way to incorporate this idea is to look at the cascading sequences of citations. That is, to determine an impact we would look not only at the inventors that cite the work but also at the inventors that cite the work of those who cite the original work, and so on. This idea was originally proposed by Trajtenberg et al. (1997). The importance of an invention was taken to be the number of citations it received plus l times the number of citations received by the patents that cited it. The parameter l is a weighting factor between 0 and 1 that determines the relative importance of subsequent generations of citations. This approach is inherently somewhat ad hoc since there is no particular basis for choosing the value of l.

A more satisfactory approach would be to begin by recognizing that researchers working on a technology or set of related technologies constitute a network. If so, then established methods of network analysis can be used to measure the communication flows within that network and, as a consequence, to measure the influence of the different nodes in the network.

3. A Systems Methodology for Analyzing R&D Spillover Networks

OVERVIEW AND MOTIVATION

There exists a growing literature developing the use of patent citations as an indicator of technological impact or knowledge spillovers. Here we will describe the methodology that we have developed to push this idea further. Just as the information contained in citations can be used to analyze differences in the importance or impact of patents, our systems approach allows us to show that *citations themselves likely differ in their significance*. For example, we would probably expect that a patent receiving five cites from Hewlett-Packard, Raytheon, MIT, Hughes Electronics, and Toshiba—known leaders in a field—might be more important than an equally cited patent with cites from less well known companies like Pegasus Technologies, Crystalloid, Cintas, Infocision, and Ball State University.

Although it is easy to speculate that citations from so-called influential organizations are more indicative of spillovers than citations from so-called less important organizations, we must discuss what distinguishes influential organizations, and how such distinctions can be quantified. Fundamentally, the systems methodology is based on the idea that the influence of organizations can be understood—and measured—by examining the place of each organization within an R&D network. From this perspective, influential organizations are those that are connected to and communicating with a lot of other organizations, and particularly a lot of other influential organizations. They represent important *nodes* in the communication system, meaning that a large fraction of the overall information flow in the network passes through them.

We believe that this systems perspective can broaden and deepen our understanding of spillover phenomena. For example, in their paper on localized spillovers, Podolny and Shepard (1997) present an interesting puzzle. They find that patent citations are more likely to be external (i.e., come from organizations outside the geographic region) when there exists a high level of *local* inventive activity. To the extent that geographic proximity facilitates spillovers, this result is surprising. We would expect that a high level of local activity would make dependence on more distant knowledge sources less necessary. From a systems perspective, however, R&D labs are connected, and a change in one lab will induce further change among all of the labs. We can view regions with a high level of local inventive activity as influential nodes in a worldwide R&D spillover network. Such places are highly connected to the world's regions, all poised to tap into influential parts of the R&D network, whether local or external.

The existence of spillover networks may also explain an interesting finding on CT scanner technology. Trajtenberg et al.'s (1997) study provides quantitative evidence showing the inequality of patent cites. He finds that the social value of CT patents increases nonlinearly with the number of patents (i.e., the information content increases disproportionately with the number of citations). The study findings hint at a much broader issue, that a larger number of citations may reflect diffusion through increasingly influential nodes in the R&D network. More generally, viewing R&D organizations as nodes in a network suggests the importance of feedback phenomena that can be a source of increasing returns. Because influential organizations benefit more from spillovers, their own R&D is more productive, making them more successful, which reinforces their influence. Such *virtuous cycles* can operate at the level of firms, technologies, or regions. Conversely, firms, technologies, and regions can be caught in a vicious cycle, because nodes that are not well connected find R&D less productive, and hence become less influential over time.

SYSTEMS METHODOLOGY IN RELATION TO THE EXISTING LITERATURE

The systems methodology builds upon the existing understanding and analysis of R&D and the role of spillovers in the research process. It develops, however, a somewhat different conception of how that process works. It also postulates a different set of relationships between the underlying R&D spillover concepts and the data that are observed. Before developing the new methodology's elements in detail it is helpful to describe the relationship between our new conception of the process and earlier work.

At the heart of both approaches is the notion that knowledge spillovers among organizations foster the research process, and that there are cases where the path and magnitude of these spillovers can be somewhat proxied using patent citations data. But the two approaches describe different concepts of the spillover process and relationship to patent citations. The traditional concept characterizes each spillover as a distinct bit of knowledge that flows from researcher A to researcher B, benefiting researcher B. Sometimes the flow of this bit of knowledge is subsequently recognized by a patent citation and sometimes not. Also, sometimes patent citations occur when no such bit of knowledge has actually flowed. In a sense these bits are additive or cumulative (i.e., researchers who receive a lot of them get a lot of spillover benefits; researchers who generate a lot of them created a lot of spillover externalities for others). By using this approach we can count the total citations received by the patents of researcher A as a (noisy) indicator of the quantity of spillovers researcher A is producing. And we can count the total number of citations made by researcher B's patents as a noisy indicator of the spillover externalities to B from other researchers.

In contrast, the systems approach does not conceive of spillovers as taking the form of *individual* knowledge bits that flow from or accumulate at particular locations. Rather, we think of spillovers as reflecting a communications *process*. In other words, research is an inherently collaborative and communicative process. Consequently, it is not so much that researcher B gets a discrete benefit from a knowledge bit garnered from researcher A. Instead, it is that researcher B will generally be more productive because that researcher communicates easily and frequently with other researchers within his sphere. (Researchers are less likely to communicate directly where proprietary knowledge is involved.) By using the systems approach, patent citations are seen as a proxy for this communication process, indicating the degree of openness and regularity of communication generally. Because the citations process provides only a noisy indicator, we think of the communications proxied by our systems analysis of patent citations as an indicator of the overall health of communication rather than as counting the "amount" of knowledge that has flowed.

The systems approach conceives of the relationship between pairs of inventors as embedded in a larger notion of a research communications network. On the one hand, the traditional approach treats the spillover process as a linear and additive. If researcher B gets 10 spillover units from researcher A and 5 spillover units from researcher C, then researcher B has received 15 spillover units. In contrast, using the systems approach we think of researcher B as residing within a mutually interconnected network. Researcher B's productivity will be affected by the overall vitality of knowledge flow in that network and the strength of the connections to the network generally. A researcher is well connected to the network by being well connected to other researchers, who are themselves well connected to the network.

This network conceptualization of the research and spillover process leads to specific methods for measuring which researchers are the most crucial. We utilize results from network theory and fuzzy set theory[1] to develop these methods. Fuzzy theory yields a particularly valuable methodology because it provides a rigorous way to model the reality that any given researcher belongs, in varying degrees, to multiple amorphous groups of researchers with whom that researcher communicates. Fuzzy theory gives us a way to make operational the notion of a researcher being well connected to the research network.

We have undertaken several case studies to understand how the systems approach leads to results different from traditional methods. We do not think the differences in results are due to random variations. Rather, they derive from the underlying differences in the concept of the R&D spillover process. In particular, because the systems approach derives from communication, what we call "influence"—the overall impact that a particular lab has on the network—depends on two-way communication. This means that a lab whose patents receive a lot of citations will not necessarily be influential if its researchers do not in turn

---

1. Fuzzy logic is a formal system of logic in which numbers on a scale from 0 to 1 are used instead of the values "true" and "false" as absolutes, to accurately represent the fact that some questions do not have a straightforward yes or no answer. With fuzzy logic, propositions can be represented by degrees of truthfulness and falsehood. For example, the statement, today is sunny, may be 100 percent true if it is a cloudless day, 80 percent true if there are some clouds, 50 percent true if it is hazy, and 0 percent true if it rains the entire day. Fuzzy logic was developed by Lofti Zadeh of the University of California, Berkley. See <http://www.cs.berkeley.edu/~zadeh/pripub.html>.

communicate with and draw on the work of others. In effect, a lab cannot be influential simply by sending out a lot of spillovers while receiving very few. If it is not engaged in two-way communication, then it is not an active and thereby influential member of the network.

Another important consequence of the systems approach is that it provides a rigorous framework for aggregation. Under the traditional approach, if there are 10 IBM labs around the country that received 100 citations in the aggregate, then our measure of the importance of IBM would simply be the overall citation count of 100. By using the systems approach, if we want to know the influence of IBM as an organization, then we must look at the overall interaction of the 10 labs within the network. The whole will be *greater than or less than* the sum of the parts, depending on the nature of the interactions among the individual parts, and between each part and the network.

Conversely, if we know a particular lab has received 10 citations, but belongs to a company that has received 100, we might suspect that its association with this larger entity has important consequences for its research productivity, and its role in generating spillovers. The traditional approach does not provide a clear way to measure effects of this kind. The systems approach explicitly incorporates a lab's links with other labs *within* the same organization. In principle, these aggregation effects can be analyzed across organizations, across technologies, and across geographic regions.

Later we will introduce the specific concepts from the network literature used in our analyses. Although these concepts are likely to be unfamiliar to many readers, by drawing from this explanation of the differences between the two spillover approaches, these concepts can be "mapped" onto more traditional concepts.

At the heart of the systems method is the concept of a *node*: a specific research lab, characterized by the organization that owns it, the technology it focuses on, and its geographic location. Nodes are the generators and recipients of spillovers. As was discussed, we focus on their overall connection to the network rather than the individual bits of knowledge they use or generate.

In principle every node in a network communicates with every other node. Following the fuzzy set terminology, we characterize the strength of these pair-wise communication interactions between nodes with the concept of a *truth value*. Finally, the overall importance or impact of a node is its *system influence*. System influence is analogous to the overall extent of spillovers generated by a lab under the traditional conception. We can calculate system influence at any level of aggregation (for an individual lab, a company, a technological area, or a geographic region).

THE R&D LAB

The basic unit in our model is the R&D lab located in a specific region, working on a specific technology, in a particular time period. The R&D networks are constructed from interactions among R&D labs. We analyze these interactions by using patent citations, which are interpreted as a form of communication. Although communication is an unfamiliar term in economics, laden with ambiguity, it is nonetheless at the heart of R&D spillovers. It can take many forms, including reading of papers, attendance at conferences, hiring of leading research consultants as windows on technology developments, indirect word-of-mouth learning, analysis of a patent database, hiring of key researchers from competitor labs, placing students in top graduate programs, and industrial espionage. In fact, sometimes it even involves personal conversation among inventors and R&D lab directors.[2] We conceive of patent citations as being a proxy or indicator variable and do not believe that one inventor's reading of another inventor's patent document is necessarily a particularly important mode of communication. Rather, consistent with the survey results discussed above, we expect that the likelihood of substantive knowledge flow from inventor A to inventor B is higher if inventor B's patent cites inventor A's patent than if no such citation occurred.

This communication interpretation incorporates learning, which strengthens the interpretation of patent citations as communication. For example, an inventor or R&D organization may not be aware of an important technology source until the patent examiner adds the citation to the application. If the cited organization is truly an important source, however, then the inventor is alerted and will monitor the newly found source for continuing developments. This is one reason why information content (communication) increases for strongly interacting R&D organizations.

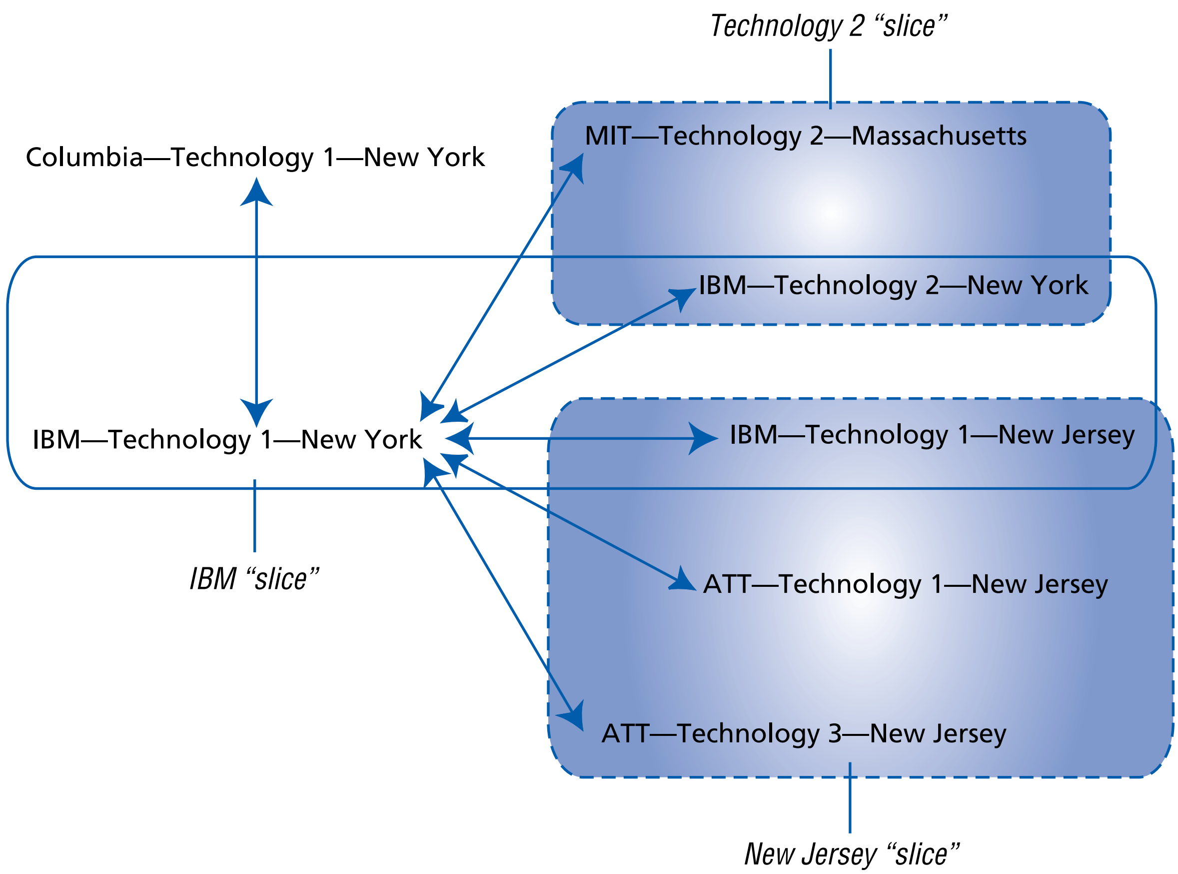
We distinguish each lab by the organization that owns it, its technological focus, and its geographical location. This means that the R&D network exists in a multidimensional space of organization-technology-region. As will be discussed, once the network has been conceived it is often useful to focus on two of these three dimensions and look at a "slice" of the network along one dimension. That is, we can choose to ignore the technological and geographic complexity and look only at the relationships among organizations, or we can ignore the organizations and geography and look only at relationships among technologies, or again we can focus on the relationships among regions, ignoring organizations and technologies.

Figure 2 illustrates the idea that each node has a location in organizational, technological, and geographic space. Thus, as shown in the figure, the activities of the organization IBM in technology area 1 in New York constitute a node distinct from IBM's activities in other technologies at the same location, and in technology 1 at other locations. With respect to the patent data, the "organization" is associated with the patent assignee, the "technology" is associated with the patent class, and "geography" is based on the address of the primary inventor. Conceptually, then, the node "IBM—Technology 1—New York" corresponds to that portion of the IBM organization that lives in New York and works in technology area 1.

2. Based on evidence from the semiconductor industry, geographic localization of citations across firms partially results from the tendency of inventors who change firms to take their research tracts with them, combined with a greater likelihood for an inventor to move to another firm in the same geographic area. See Almeida and Kogut (1999).

FIGURE 2

Construction of Influence Measure from Hypothetical Pair-wise Citation Pattern



Operationally, this node is represented by the body of patents assigned to IBM, in the patent class corresponding to technology 1, for whom the primary inventor resides in New York.

For the sake of illustration, suppose that for a particular network the nodes shown in Figure 2 comprise the entire network. We calculate a truth value, which is a fuzzy logic measure of the magnitude of interaction, for every pair of nodes. This interaction measure is based on the patent citations between the two nodes, *in both directions*. These pair-wise truth values are then used in the fuzzy clustering algorithm to map the full R&D network. If the node IBM—Technology 1—New York has strong interactions with other nodes, and/or the nodes that it interacts with themselves have strong interactions with the rest of the network, then IBM—Technology 1—New York will be well connected to the overall network. (Ultimately, the fuzzy logic measure of interaction between any two nodes incorporates the system position of each node; that is, the measures of interaction reflect direct and indirect effects.) Our hypothesis is that new knowledge generated at well-connected nodes will create large knowledge spillovers because such nodes are well connected to the rest of the relevant innovation system.

Nodes that are highly connected to each other are presumed to communicate with each other more intensely. The implication is that spillovers generated at highly connected nodes will flow more completely to other nodes in the system. Conversely, highly connected nodes will *receive* greater benefit from spillovers created within the system.

Once we have mapped the strength of interaction among the nodes in the network, it is then possible to accumulate the *system influence* of each node. The system influence of each node results from the strength of its interaction with other nodes, compounded by the strength of interaction of *those* nodes with other nodes. Crudely, an influential organization is one that is well connected to other influential organizations. The apparent circularity of this definition is resolved by the fuzzy-logic algorithm, which builds the measures of system influence with an iterative procedure.

Although the organization-technology-geography node is the fundamental building block of the network, for many purposes we are interested in the nature of the network interactions at a higher level of aggregation. To do this we can construct slices of the multidimensional system along any dimension of interest. We can, for example, estimate the overall system influence (within this particular technology network) of IBM as an organization. For this purpose we would sum the measures of interaction (truth values) for the three IBM nodes shown in the figure. In a completely analogous way we can create slices of the network along the geography or technology dimensions. This allows us to examine the influence of particular technologies or particular regions within the technology network we have singled out for examination.

Note that in this formulation, the treatment of the three dimensions—organization, technology, and geography—is completely symmetric. It is important to be clear about the implications of this symmetry. We are *not* assuming that the magnitude of spillovers within organizations is somehow the same as the magnitude of spillovers within regions. The citations data themselves tell us the extent to which nodes that share a region or nodes that share a technology are more likely or not to cite each other. In the hypothetical example illustrated in Figure 2, for example, we allow the citations data to tell us how the interaction between IBM—Technology 1—New York and IBM—Technology 1—New Jersey compares in magnitude with the interaction between IBM—Technology 1—New York and Columbia—Technology 1—New York or IBM—Technology 2—New York. We *are* assuming, however, that the relationship between observed citations and actual communication or spillovers is similar across these different dimensions. That is, we are relying on the maintained hypothesis (supported by the empirical regularities discussed above) that a greater volume of patent citations between node pairs is associated, on average, with greater communication between them. Moreover, we are assuming that the tightness of the link between citations and communications does not differ systematically across the different dimensions of organization, technology, and geography.

A NUMERICAL FUZZY SYSTEMS EXAMPLE

We now formally develop an abbreviated spillover systems model utilizing hypothetical patent data and patent citations. Appendix A develops the R&D spillover networks algorithm. By a system we mean that there exists a hierarchy of technologies, R&D organizations, and regions connected by a communication network. System effects exist when a change in any component diffuses throughout the network.[3]

The system can be described by four key components:

- An invention (embodied in the patent).
- Units of the system—a unit is a source of innovation as well as a destination for the knowledge flows communicated in patent citations. In our model a *unit* is a particular R&D lab, in a particular technology, region, and year.
- Channel of communication—citations are taken as a proxy.
- Flow of knowledge—depends on the position of the two units within the system and strength of interaction.

Source: Constructing the Set around Cited Behavior

Figure 3, using patent citations, illustrates a one-way knowledge flow from organizations in one time period to organizations in the following three time periods. We denote A-B to refer to the interaction between two labs, A and B. Our example includes three organizations, A, B, and C, and three time periods (*T* = t1, t2, and t3). Patents have been assigned to A, B, and C. (In practice, the basic unit of analysis is an R&D lab, by technology, location, and time.) The term "center" refers to the organization from whose perspective the members of the set are developed. A-B_t1 denotes A being cited by B in t1. The whole number in the chart represents the number of times the source node is cited by the destination (use) node. (The source node corresponds to the cited organization in a particular time period, while the destination or "use" node is the citing organization in a particular time period.)

Fractions

The corresponding fraction in Figure 3 is simply the number of times the cited organization in a particular time period is cited by the citing organization, divided by the total number of times the organization is cited in that time period. For example, based on time period 1's patents, A is cited a total of 40 times (10 by B in t1, 20 by B in t2, and 10 by A itself in t2).

---

3. Chatterji (1996) points out the increasing interest in external sources of technology. In fact, the Industrial Research Institute has formed a working group to identify and document "best practices" being used by companies to acquire external knowledge. The follow quote suggests the importance of R&D networks as a vehicle for firms' external learning: "… firms do not search in isolation; rather, they search as members of a population of simultaneously searching organizations." See Podolny and Stuart (1995). One measure of the significance given to the technology search process is the existence of a national organization called the Society of Competitive Intelligence Professionals.

Thus A_t1 is cited 10 times by A_t2, 10 times by B_t1, and 20 times by B_t2. The corresponding fractions are 0.25, 0.25, and 0.50, respectively.

Truth of Interaction

The fractions indicate the extent to which a cited organization contributes knowledge to citing organizations. If we were to associate with each fraction a measure of the percentage of such fractions that are below a particular value, then we would have a ranking of *flow out* of the cited organization to the citing organization in a particular time period *relative to* all such flows. In the fuzzy logic literature this term would be referred to as a *truth value*. Table 1 constructs truth values representing the cited numbers contained in Figure 3. Here we associate with each fraction the corresponding truth value.

FIGURE 3

R&D Spillover Network: Cited Behavior Fractions

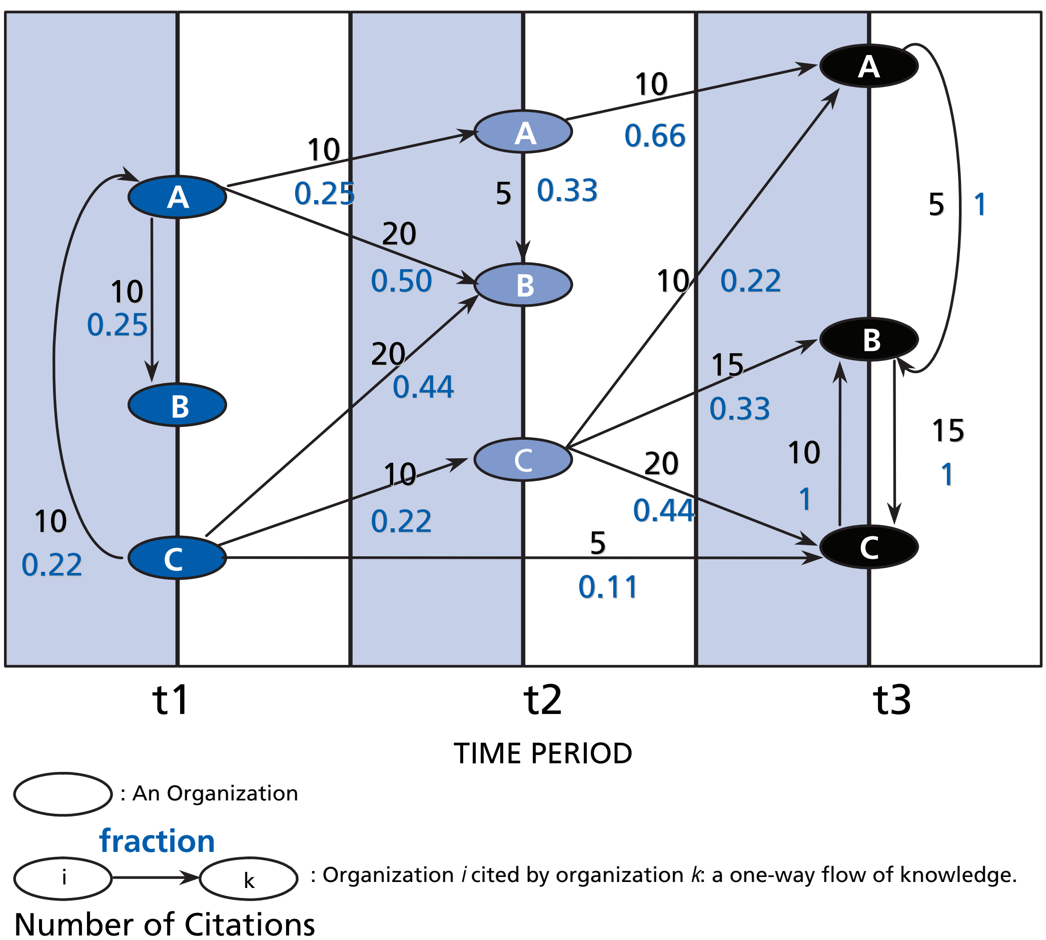


TABLE 1

Truth of Interaction (on Cited Behavior)

Interval	0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–0.5	0.5–0.6	0.6–0.7	0.7–0.8	0.8–0.9	0.9–0.10
(1) Number of Fractions	0	1	7	2	1	1	0	0	0	3
(2) Area = (1/total)	0	0.0667	0.46666	0.13333	0.0667	0.0667	0	0	0	0.2
Cumulative Area	0	0.0667	0.53336	0.6666	0.7333	0.8000	0.800	0.800	0.800	1
Area Below = Truth	0	0	0.667	0.53336	0.6666	0.7333	0.8	0.8	0.8	0.8

First, our illustration sorts fractions into 10 intervals. Second, the area represented by each fraction is simply the percentage of all fractions that lie within a particular interval (e.g., one cited fraction lies in the interval 0–0.1, so (2) is simply 1/15 = 0.0667, where 15 is the total number of fractions). Third, we next calculate the cumulative area, which represents the percentage of all such fractions up to a specific interval (e.g., the interval up to and including 0.3–0.4 includes 10 of the 15, or 0.6666 cited fractions). Finally, the truth value is defined as the area below a particular interval. For example, the truth value for 0.3–0.4 is 0.53336 (the area below 0.3–0.4). The resulting truth values are shown in Figure 4 along with fractions.

Constructing the Fuzzy Set

These data can now be used to identify a cluster or network of citing organizations by time period, around each cited organization in a particular time period. The (one-way) interactions with citing organizations belong to the set with varying truth values. In our example, the set around A_t1 on its being cited behavior is shown in Figure 5. The three members of the set are the interactions with B_t2, B_t1, and A_t2. The membership value of the interaction with B_t2 with respect to this set is 0.7333.

How true is the statement that the interaction with B_t2 belongs to the set of "being cited" behavior constructed around A_t1? The interaction with B_t2 belongs to the set with a membership value (truth) of 0.7333. Therefore, the truth values are a measure of the confidence we have in making the statement that A_t1-B_t2 interaction belongs to the set being cited around A_t1.

Systems Value

It is important to remember how the truth measure and, therefore, the set were constructed. At the first step we place the flow from a cited unit A_t1 to citing unit B_t2 relative to all units citing A_t1 (which is the fraction of A_t1 being cited by B_t2 = 0.50). In the second step, we place the relative measure in relation to all such relative measures and found that .7333 percent of all such fractions across all one-way flows are below the fraction of A_t1 being cited by B_t2. The *sum* of such a measure for cited organizations by time period would yield a ranking of the organizations on the knowledge being contributed by the organization. The sum would

FIGURE 4

Truth of One-way Knowledge Flows On Cited Behavior

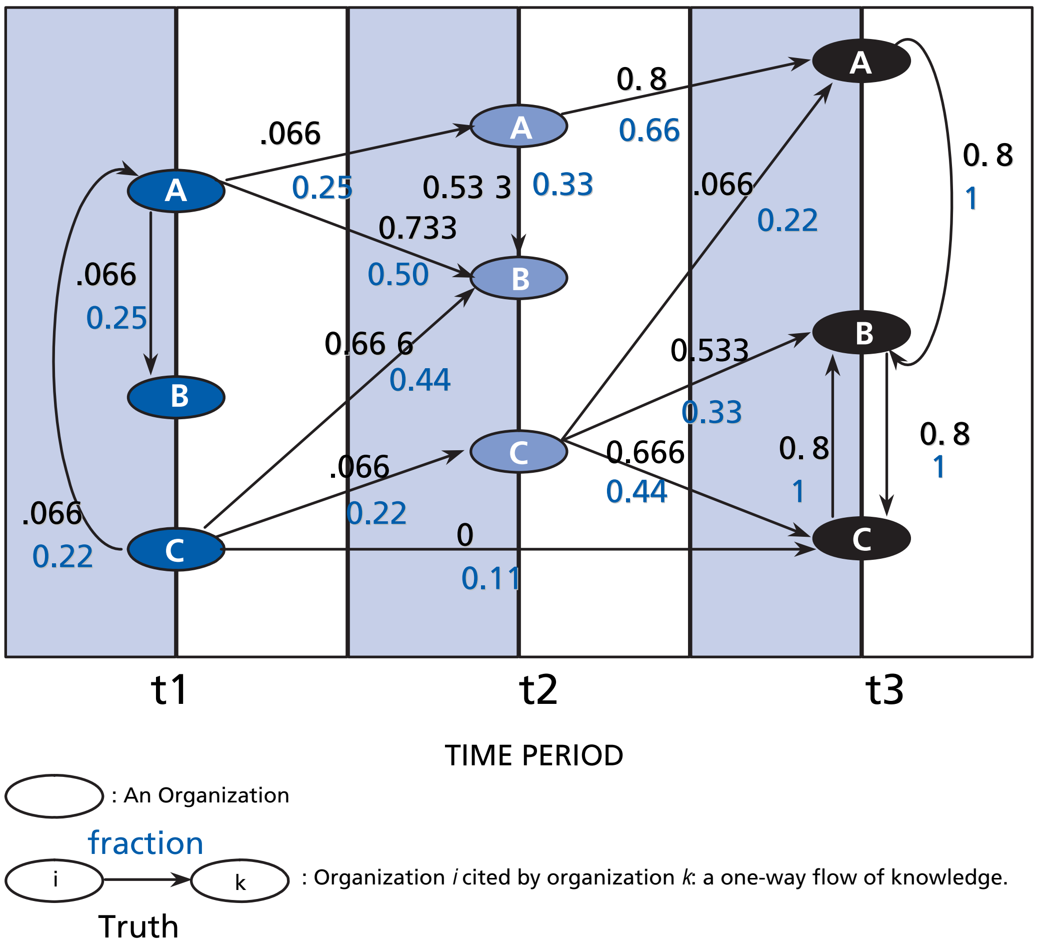
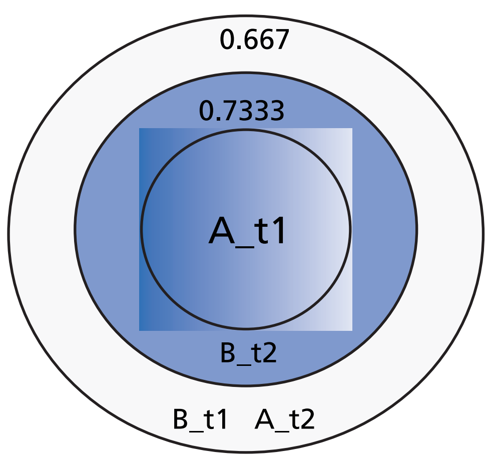


FIGURE 5

The Cluster around A_t1 on Being Cited Behavior



be high if the cited organization contributes knowledge to a very large number of similarly influential organizations. The high sum would reflect a rich knowledge source.

We also rank organizations on the variable *value*, which is defined as the relative ability of the organization to disburse knowledge or its influence. For an organization, the variable value is the sum of its truth values of one-way flows with other organizations divided by the total of the truth value of all one-way flows for *all* organizations. (See Table 2.) *(As discussed previously, this is a way to measure the system influence of each node, which is based on the strength of its communication with other nodes, weighted by the strength of communication of the interacting nodes with the rest of the system. This measure of influence can be aggregated in different dimensions of interest, and allows for the rankings of the spillover potential within a particular R&D network by the overall organization.)*

Learning: Constructing the Set around B on Citing Behavior

We follow a parallel process for constructing the set around B as citing (learning) behavior.[4] (See Figures 6, 7, and 8, which illustrate the methodology when applied to citing patents in a parallel fashion. Table 3 is based on the calculations included in Figure 6). To summarize, for each citing organization and cited organization, we first compute the relative fraction of times that the citing organization cites the cited organization. Next, we compute for each such fraction the percentage of citing fractions that are below the particular fraction. This yields the truth of citing behavior. We then define the citing organization's "learning ability" as the sum of its citing truth values divided by the total across all citing organizations (see Table 3). A ranking of organizations by this variable provides a hierarchical list of organizations as learners. If the citing (learning) organization is doing R&D that closely parallels the source organization's R&D, then we expect a greater flow of knowledge to occur. Our measure value captures the importance of a cited organization's R&D.

TABLE 2

Organizations Ranked on Influence

(1) Organization_Time Period	(2) Sum of Truth	(3) Value = (2)/total
A_t2	1.333	0.22988
C_t2	1.2666	0.21838
C_t1	0.8	0.137933
A_t3, B_t3, C_t3	0.8	0.137933
Total	5.799	

4. The previous section uses Figures 3, 4, and 5 based on Table 1 to explain the methodology when applied to cited patents behavior. This section shows the parallel side of the description for citing behavior in an abbreviated fashion.

FIGURE 6

R&D Spillover Network: One-way Knowledge Flows as Evidenced by Citing Pattern

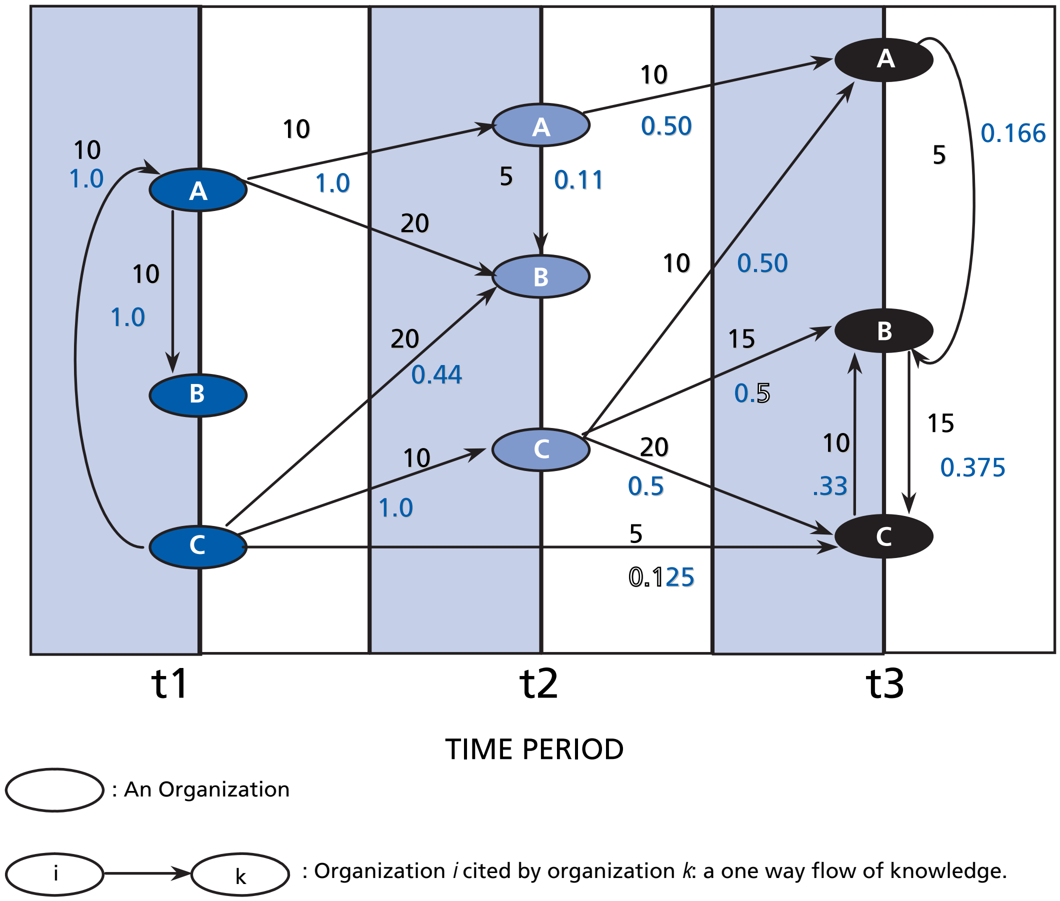


TABLE 3

Truth of Interaction (on Cited Behavior)

Interval	0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–0.5	0.5–0.6	0.6–0.7	0.7–0.8	0.8–0.9	0.9–0.10
(1) Number of Fractions	0	3	0	0	2	4	0	0	0	6
(2) Area = (1/total)	0	0.2	0	0	0.13333	0.2667	0	0	0	0.5
Cumulative Area	0	0.2	0.2	0.2	0.333	0.600	0.600	0.600	0.600	1
Area Below = Truth	0	0	0.2	0.2	0.2	0.333	0.600	0.600	0.600	0.600

FIGURE 7

Truth of One-way Knowledge Flows as Evidenced by Citing Patterns

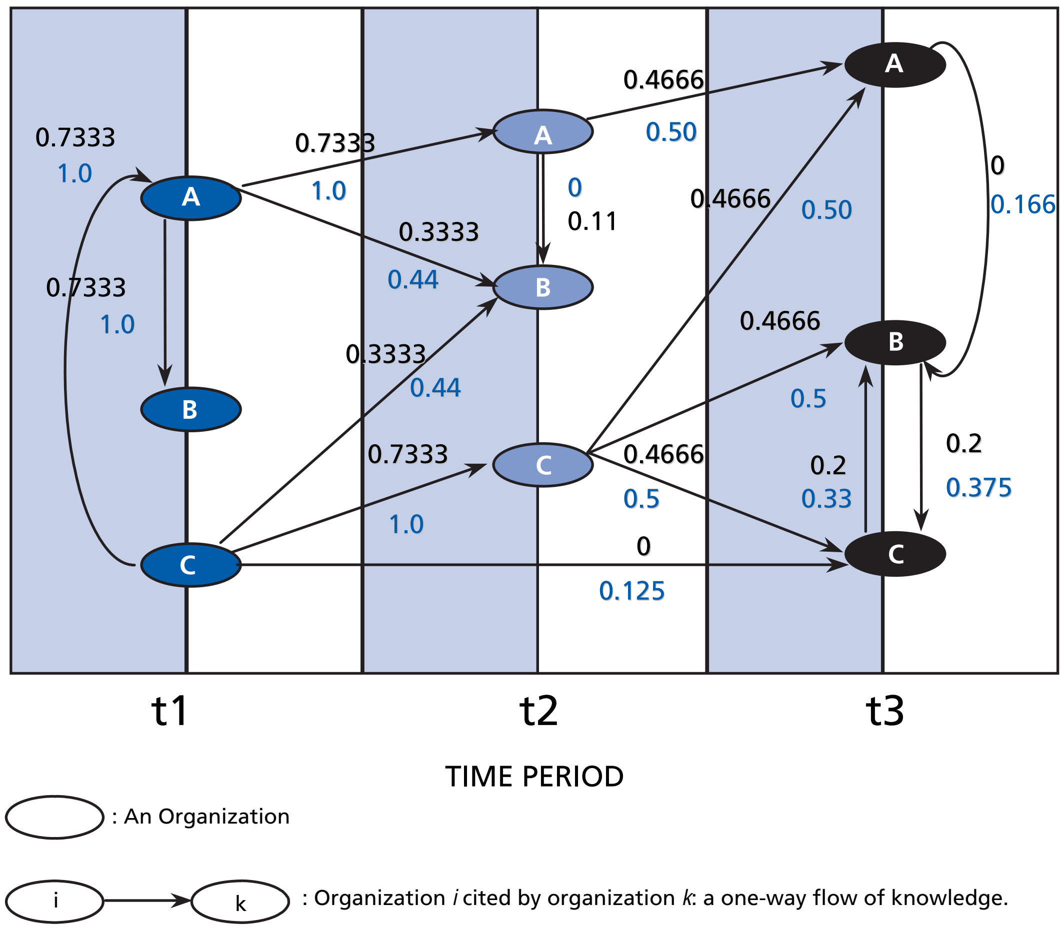
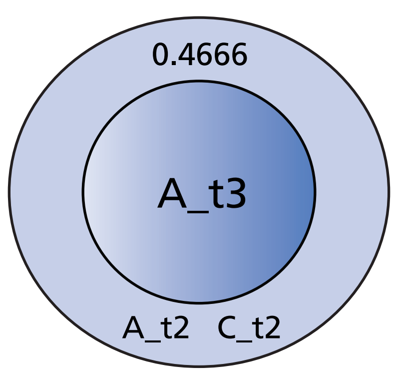


FIGURE 8

The Cluster around A_t3 on Citing Behavior



The Truth of Interaction

To summarize, the truth of A's cited behavior with respect to citing organization B is the membership value that A-B interaction has in the set of *cited* behavior around A. Similarly, the truth of B's citing behavior with respect to cited organization A is the membership value that A-B interaction has in the set of *citing* behavior around B. *It follows from fuzzy set theory* that the membership value of the A-B interaction to both the cited set around A and the citing set around B is the *minimum* of the truth values.

In our example, the (one-way) interaction A_t1-B_t2 belongs to the cited set around A_t1 with a truth of 0.7333 and the A-B interaction belongs to the citing set around B_t2 with a truth of 0.44 (see Table 4). It follows that the interaction belongs to both sets (the intersection of the two fuzzy sets) with a truth of 0.44. In other words, on the scale of 0–1 we are 0.44 sure that the interaction belongs to both sets of cited around A_t1 and citing around B_t2. Keep in mind that the truth value is developed with reference to the whole space of one-way flow interactions.

Why the Minimum?

Our analyses take the truth of the interaction between *i* and *j* to be the minimum of the truth based on citing and cited behavior from *i*'s perspective. Of course we could have employed any number of possible operators (e.g., the maximum, the average of the citing truth and cited truth, or even one or the other of citing/cited truths). Why the minimum?

We began by looking for an operator that distinguishes strong two-way communication. We chose to employ the minimum assumption for three reasons: first, the minimum assumption

TABLE 4

Organizations Ranked on Ability to Learn

(1) Organization_Time Period	(2) Sum of Truth	(3) Value = (2)/total
A_t3	0.9332	0.159
A_t1, A_t2, B_t1, C_t3	0.7333	0.125
B_t2, B_t3, C_t3	0.6666	0.1135
C_T1	0.0	0
Total	5.8662	

is consistent with fuzzy set theory (i.e., the interaction between two organizations A-B can be measured by the intersection of two fuzzy sets, which is the minimum of the truth of the cited set around A and the truth of the citing set around B; for a fuller treatment, see Sinha (2000)); second, the choice of the minimum provides a strict indicator of a two-way relationship, which we assume to reflect more influential forms of communication (the choice of the maximum could produce one-way flows); and, third, in experimentation with three operators (the maximum, the average, and the minimum) the minimum yielded the best results. We used knowledge of Cleveland's R&D base and its technologies to gauge the reasonableness of the results derived from each assumption. In essence, we used the algorithm utilizing three alternative assumptions to generate separate results specific to Cleveland (at three levels: technology, organization, and region). In all cases, the minimum assumption yielded results that were consistently the most reasonable.

Other possible constructions of the truth value are possible. For example, we could have used the maximum truth value, the average of the two values, or simply the truth on citing or cited behavior. Future work should analyze the many possible operators for fuzzy intersections and experiment with each of them (see Sinha, 2000).

We interpret the quantity (the value of A) multiplied by the second term (the membership value of A-B interaction with respect to the cited set around A and the citing set around B) as an indicative measure of the flow from A to B. The first term captures the importance or influence of the cited (source) organization, whereas the second term reflects the confluence of A and B's R&D. The product of the two terms suggests that a larger flow results when a larger knowledge pool is mediated by the truth of interaction. We follow this approach to develop an indicative measure of the flow of knowledge from cited to citing organizations across time periods.

Diffusion to Second and Third Level

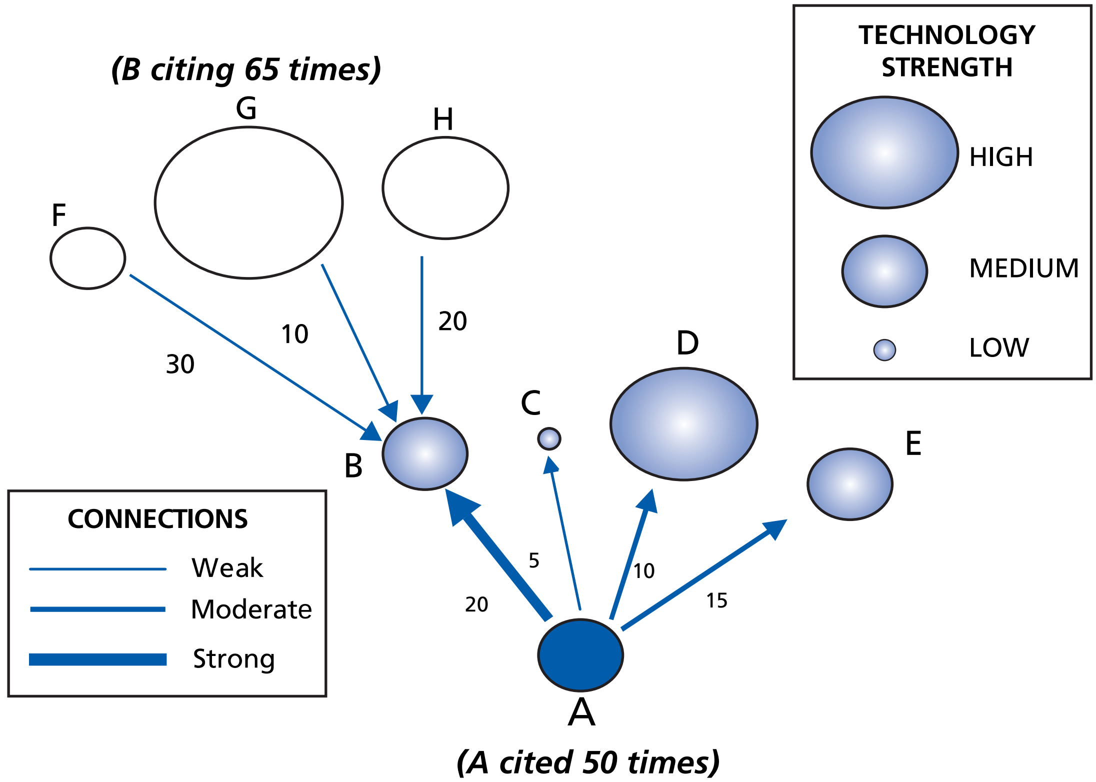
Our model also assumes that when an organization A interacts with an organization B, B benefits not only from A's knowledge but also from the organizations that A interacts with (see Figure 9). Thus, the knowledge flow from A to B is determined by the truth of the interaction between the center and the member, as well as the *group value* associated with the center. And, the group value multiplied by the interaction between the center and a member measures the flow from the group around the center to the member. This group value captures second-level diffusion effects. The algorithm used in this paper allows for third-level diffusion. In other words, the sum of group-level interaction truth values gives us a hierarchical communication structure on second-, then third-level, diffusion effects.

Summarizing Steps

Here we illustrate the use of patent citations to construct the system that represents the R&D network. A formal mathematical description of the procedure is given in Appendix A. The fuzzy methodology allows us to develop indicative membership measures between 0 and 1 representing the strength of interaction between any pair-wise combination of R&D labs,

FIGURE 9

All Citations Are Not Equal: A One-way Flow Illustration (B Citing 65 times)



specific to organization, technology, and region. Our systems model then builds the system iteratively incorporating the first, second, and then the third level of diffusion of spillovers. The result is a hierarchical R&D network system.

At the first step we construct sets on cited and citing behavior around each R&D lab, working in specific technologies at a specific geographic location, by year. The R&D lab is called the center of the set. At this stage the set members are other R&D labs working in specific technologies by year that cite or are being cited by the center. In classical set theory, a member either belongs to a set or does not belong to the set. In the first case, the member has a membership value of "1" and in the second a membership value of "0." Fuzzy set theory develops indicative membership measures of how true the statement is that a member belongs to a particular set. These measures are called truth values. The truth measure lies between 0 and 1. Different methods for constructing truth measures have been developed in the fuzzy set literature.

To model spillovers, in the following we develop a method to construct truth values, estimate the strength of R&D lab interactions, and construct a measure of system influence. The first four steps describe a hierarchical system defined on first-level diffusion. In developing the fuzzy measures, using data from Figure 9 we take the following steps:

Step 1. Compute cited fractions:

fA→= 5/50 = 0.1 (the fraction of citations received by A coming from B)

fB←A = 5/65 = 0.08 (the fraction of citations made by B going to A)

Step 2. Normalize fractions relative to all such flows: Is the flow a lot or a little relative to all flows in the dataset? Evaluate the cumulative distribution of all such fA→B. Then calculate the area under the cumulative distribution up to fA→B and compute this area as a percent of the total area. At this step, we have constructed the set around A on the flow out from A based on citations to A. The computed percentage is an indicative measure of how true the statement is that B belongs to the set around the center A on flow out. Computed for each citing member, the truth value represents the relative importance of B, E, C, and D as *learners from A from A's perspective*. We denote truth values on the flow-out property as α. So B belongs to the set around A on the flow-out pattern with a truth of $\alpha_{AB.}$ Next evaluate the cumulative distribution of all such fB←A. Calculate the area under the cumulative distribution up to fB←A and compute this area as a percent of the total area. At this step, we have constructed the set around B on the flow into B based on citations made by B. This is termed the truth of fB←A interaction. Computed for each cited organization, the truth value represents the relative importance of A, F, G, and H *as sources to B from B's perspective*. We denote the truth based on the flow-in property as β. So A belongs to the set around B on the flow into B with a truth of $\beta_{BA.}$

Step 3. For a specific pair A and B, we take the *minimum* of the two truths α_{AB} and β_{BA} as a measure of the strength of interactions between any two R&D organizations by technology and year.

In set theory terms, by taking the minimum we have constructed the intersection of the sets on flow out of A and flow into B. *Interaction corresponds to the intersection of sets. We denote the minimum as* δ. The strength of interaction captures the distinction between communication links of differing intensity. We choose the minimum to reflect the view that interaction between any two parties is likely to be constrained by the lesser of the two communications (the source and use). Using a classroom analogy, imagine four teaching-learning situations: "A" teacher, "A" student; "A" teacher, "C" student; "C" teacher, "A" student; and "C" teacher, "C" student. Where the two are different, our assumption takes the minimum of A-C and C-A.

Step 4. The truth between a center and any single member measures the system influence of the link (the interaction) with that member. *By summing the interaction truths of a center across all its members we get the center's total system influence*. For each center we have two system values, one based on the flow out of the center (the technology source) and a second based on flow into the center (the technology user). The value of the center, which is what the center has to offer, is estimated as the minimum of the term (the sum of truth of the center as a source and the truth of the center as a learner). The center also incorporates a time dimension; therefore, the analyses reveal evolution of spillover networks.

Calculating System Value

We can then determine the influence of particular R&D labs by technology and region. The sum of the interaction truths (values between 0 and 1) across an R&D network is the lab's total system value (SV) for the technology. The SV is a function of the number of links in a network and the strength of each pair-wise interaction. Therefore, an individual R&D lab's SV is high when there are a large number of intensive links with other influential labs.

For each pair of R&D labs in a network we develop three measures: the relative importance of the unit (system value); the strength of interaction (the truth value); and the relative systems influence of the link between the two units (the system value of the source multiplied by the truth of interaction between the source and learning R&D labs). The results can be aggregated across the three dimensions *by year:* technology, organization, and region.

Illustrating the Results of the Systems Analysis

Here we will present full-blown analyses of the R&D networks underlying two technologies that have been of interest to the ATP. Before turning to the two cases, it is useful to get a sense for the nature of the systems "maps" by examining small segments of larger systems. We present two such illustrations here. The first focuses on the structure of a network in organizational space, and a second focuses on the structure of a network in technology space.

The Micro-electromechanical Systems (MEMS) Network Surrounding MIT

Figure 10 illustrates one segment of a specific R&D organizational network for micro-electromechanical systems (MEMS) centered on the Massachusetts Institute of Technology (MIT) over the period 1985 to 1995. The full network is much larger and is discussed later. (See Figure 16 for the top 25 MEMS organizations.) This segment is centered on MIT, which is ranked twentieth in the full network. It illustrates that R&D networks operate as both sources and learners. (In Figure 10, an arrow pointing away from a lab indicates the lab as a *source* of spillovers; an arrow pointing to a lab indicates the lab as a *learner* from spillovers.) One implication is that a firm's value as a source depends on its ability to learn from its external environment. The nodes in the R&D network correspond to the R&D labs and the arrows indicate spillover flows between labs. The color bar indicates importance of the R&D lab and the systems influence of interactions. The most influential node is brown; the least (not shown) is dark blue. Arrows follow the same scheme.

The analysis, which is based on real data, identifies AT&T, Honeywell, and Xerox as the most important MEMS organizations in this part of the network. The full network contains more nodes, incorporating interactions through third-order diffusion of spillovers. Next are GTE, MIT, and the U.S. Army. The figure incorporates the most important interactions (i.e., those closest members of MIT's R&D MEMS neighborhood). The interactions include first- and second-order diffusion spillovers. For example, the flow from MIT to Xerox is about 0.5. This means that about 50 percent of all R&D pair-wise interactions in the full MEMS network are less intense than that between MIT and Xerox. Also shown, GTE has a second-order influence on MIT. This occurs through its direct influence on AT&T.

To illustrate, MIT's system value for MEMS technologies is the sum of estimated interaction truth values across all links in the network. If Figure 10 represents the full MEMS network, then MIT's system value for MEMS would simply be the sum of the value of each link in Figure 2 (values between 0 and 1). Of course, the full network is much larger.

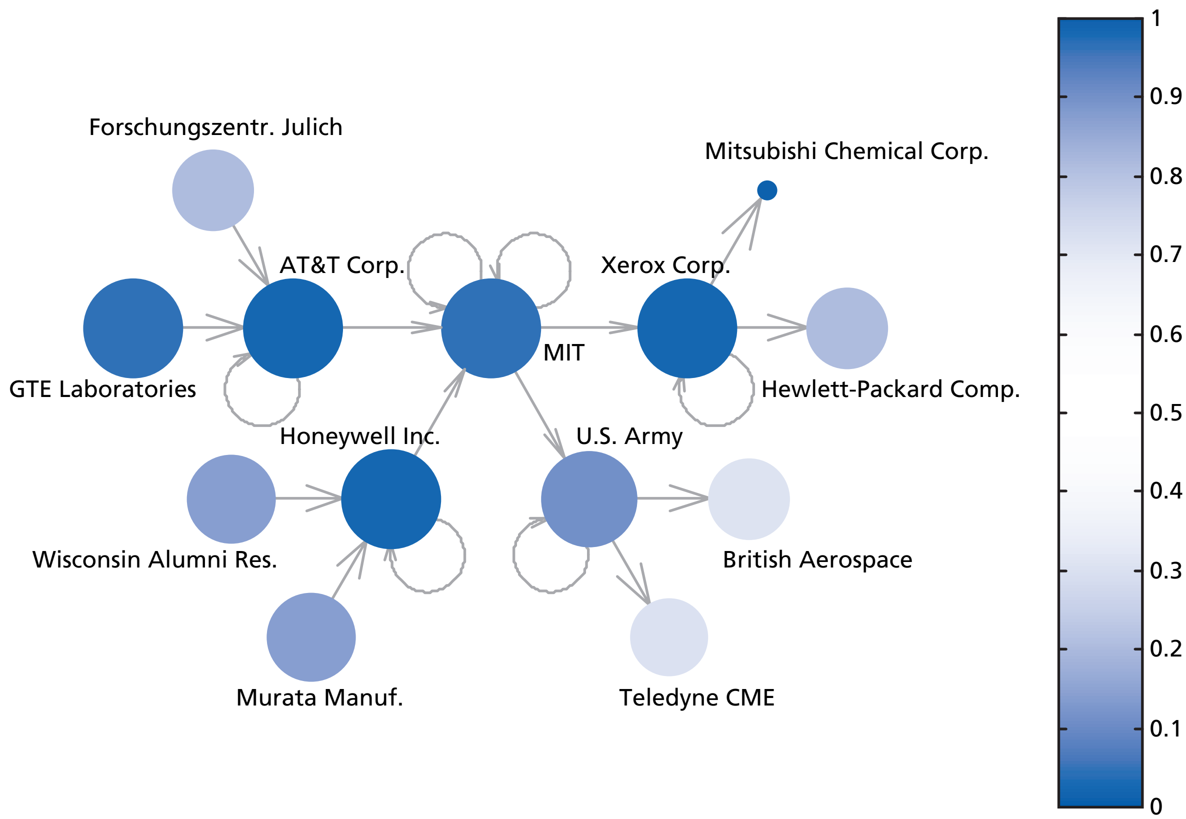
The methodology tells us that funding projects involving particular organizations and technologies inadvertently leads to specific networks. The expected social benefits can vary by industry (depending on the underlying network interactions), and regional beneficiaries can easily vary greatly as well.

The Polymer Network

Our methodology also permits us to analyze networked clusters of technologies. For instance, the MEMS organizational network surrounding MIT simultaneously identifies networks by organization, technology, region, and year. The technology we refer to as MEMS consists of a hierarchical set of specific patent classes connected by the network we model with patent

FIGURE 10

MEMS R&D Network around the Massachusetts Institute of Technology

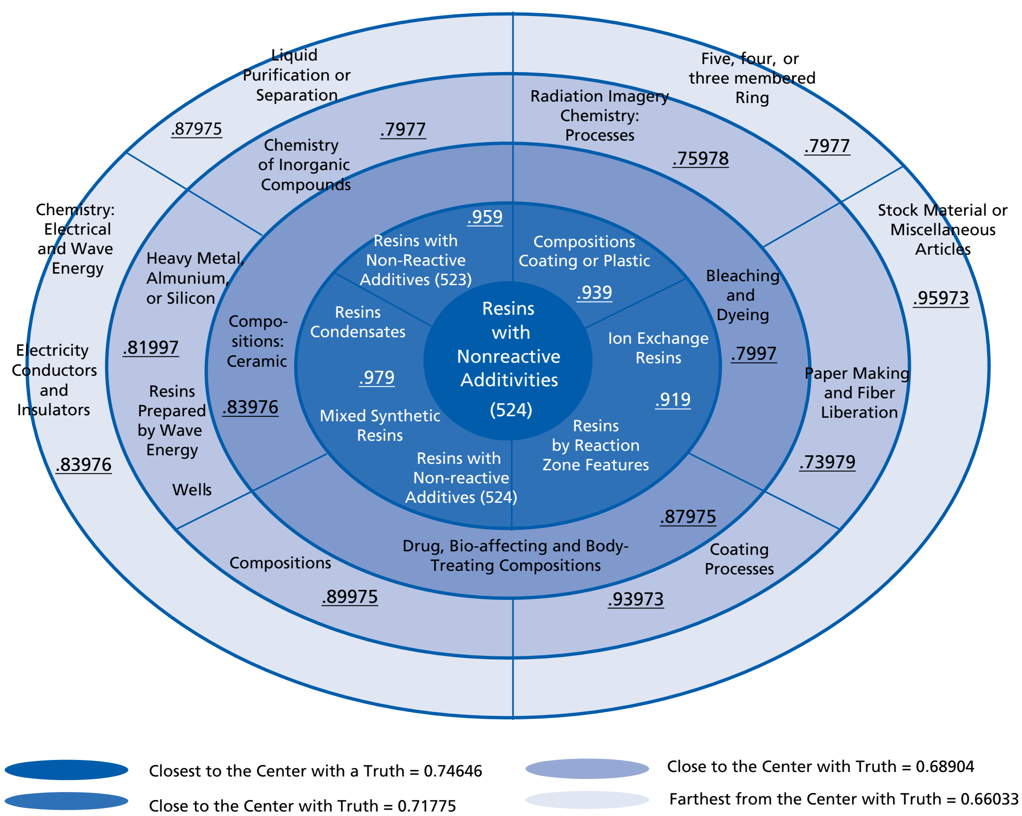


citations and the fuzzy systems methodology. Within the full MEMS network, some technologies contribute more to the full network MEMS system value and, therefore, are ranked higher within the broader technology called MEMS.

By using patent data covering the period 1985 to 1995, we can illustrate this point with an application of the methodology to polymers. Figure 11 shows the cluster components comprising the network around one important polymer patent class, Resins with Nonreactive Additives (class 524). In effect, the chart depicts the technology space for polymers centered on Resins and Nonreactive Additives, and R&D spillovers related to this specific patent class (524) involve the components shown in the chart. The chart shows a hierarchy from the perspective of the center based on interactions incorporating three levels of diffusion. One way to think about the chart is to imagine the full polymer network, which is not shown. Each member of the set around 524 is a center for its own network cluster, which in turn interacts with its members in the same way as depicted in this figure. In effect, there exists another network cluster which looks like Figure 11 for each of Figure 2's components, and another for each component at the third level of diffusion. Obviously, all of the components cannot be presented in one figure. The measures contained in Figure 11, however, include their effects.

FIGURE 11

Technology Network around the Polymer Patent Class 524



The hierarchy is represented by the four rings surrounding 524. Technologies in the first ring interact more intensely with the center and, therefore, might be considered as 524's core technologies. Note that the figure includes two values. The first value defines each ring and, therefore, the position of each component technology within the full polymer network. This value is based on a pair-wise interaction (i.e., the ring value represents the interactions of the *specific* technology with the center (524)). For example, the outer (fourth) ring includes technologies with an interaction truth value of 0.66033. Notice, however, that Stock Materials by itself interacts with 524 with a value of 0.95973. In other words, this individual interaction exceeds that of 96 percent of all such interactions within the full network, indicating that this class is very closely associated with 524. The second measure, which places Stock Materials in the outer ring, incorporates additional interactions that Stock Materials has with the set of technology members around it (i.e., it has its own network). As a consequence, when allowing for diffusion of R&D spillovers, Stock Materials is farther from the center than one might expect based on its individual interaction with 524. This can be contrasted with Ion Exchange Resins, which, despite its lower degree of interaction with 524, contributes more to 524. This happens because the network components surrounding Ion Exchange Resins are more important to 524 than the network components surrounding Stock Materials. One implication is that the indirect interactions picked up by this systems methodology, as R&D spillovers occur through three levels of diffusion, can significantly alter the measure of a technology's broader influence.

4. ATP and the Case of MEMS

In this section we focus on MEMS. This is an emerging technology that appears to be potentially enabling, at least in the sense of having apparently wide applicability to a number of different industries.[1]

MEMS combines computation, sensing, and actuation with miniaturization to make mechanical and electrical components.[2] The bulk of applications are pressure sensors, optical switching, inertial sensors, fluid regulation and control, and mass data storage. These applications cut across a number of manufacturing industries, including sensors, industrial and residential controls, electronic components, computer peripherals, automotive and aerospace electronics, analytical instruments, and office equipment. It is also likely that MEMS interacts extensively with other technologies. The industry list suggests a potential for generating a large, broad-based volume of R&D spillovers.

Our analysis suggests that R&D networks generating precompetitive, enabling technologies may have certain characteristics, such as that the universities and government labs play significant roles as sources, the network is sparse and evolving, the technology is new (the cited patents are relatively current), the total system spillovers increase significantly and technology gets diffused rapidly, influential companies perform significant basic research, and the technologies become geographically concentrated in important regions serving as incubators.[3]

---

1. When we began this research project, MEMS was under consideration for the formation of an ATP-focused program. Although this is possibly indicative of its enabling potential, none of our research or conclusions were tied to the focused-program analysis.

2. For a description of MEMS, see U.S. Department of Defense (1995). For an interesting description of MEMS, see *Discover* (March 1998).

3. Serious R&D organizations invest a portion of R&D to actively acquire external knowledge and aggressively search what Stuart and Podolny (1995) call the technological landscape for ideas, knowledge of competitors' technology, etc. See Cohen and Levinthal (1990). Clearly, a firm's learning process is not random. Evidence from our interviews suggests that firms are focusing more resources on acquiring technology from external sources. See Jaffe et al. (1998).

MEMS is an emerging technology worldwide. About 200 firms are actively engaged in MEMS R&D: roughly 80 are U.S. firms, and Japan is the second major player. According to the U.S. Department of Defense, the MEMS market was $1 billion in 1994. Projections for the year 2000 ranged from $8 to $14 billion.[4]

The U.S. industry investment in MEMS so far has been fairly modest (about $120 million in 1995). In contrast, in the same year federal R&D support of MEMS was a large component (about $35 million), $30 million of which came from the U.S. Department of Defense (mainly the Defense Advanced Research Projects Agency (DARPA)).[5] Suggestive of the technology's emerging character, about 30 universities and government labs are actively pursuing MEMS technologies. The National Science Foundation's (NSF) MEMS support was $3 million. National labs contributed about $2 million. Between 1989 and 2000, NSF sponsored 124 MEMS-related projects at 61 organizations (mainly universities), with funding of about $25 million. Approximately $1.4 million consists of Small Business Innovation Research (SBIR) grants.[6]

DATA AND MEMS PATENTS

The data are drawn from the universe of patents granted by the U.S. Patent Office from 1963 through 1995.[7] Information on patent citations begins in 1977. Electronic data on the assignee is available beginning in 1969. We geographically locate patents using the inventor's address, which means that the location in our analysis is the R&D lab's location and not the headquarters' (assignee's) location. In addition to country and state, inventors have been sorted first into counties and then metropolitan areas.

As a foundation for analyzing MEMS, starting with a short list of key inventors and federally funded MEMS projects, we developed a core database of about 1,200 MEMS patents. Citations to these initial patents were used to identify additional MEMS candidate patents. Each candidate patent's abstract and exemplary claims were read to ensure that the patent was a MEMS technology.[8]

---

4. The data for this report was compiled for the period 1985 to 1995.

5. Our research shows that DARPA has funded 62 projects at 48 organizations (17 universities, 5 government labs, 18 large companies, and 8 small firms). DARPA funded five SBIR projects at four companies; they previously funded an additional five SBIR projects. The U.S. Army has funded 17 MEMS-related projects at 14 firms through its SBIR program. The projects amount to nearly $2 million. NASA has sponsored 20 MEMS-related SBIR projects. (No dollar amount was available.) However, the MEMS working group at NASA-Lewis in Cleveland supported $2.5 million MEMS R&D by 17 scientists and engineers. Moreover, Ohio MEMS-Net has funded $2.4 million for capital investments in 1995 and 1996.

6. Ranked by total NSF support of MEMS projects, the top 10 institutions include: Stanford, UC Berkeley, University of Michigan, Cornell, University of Utah, University of Pennsylvania, University of Illinois Chicago, Case Western Reserve University, University of Minnesota, and University of Hawaii. Most of the MEMS university projects are associated with fairly extensive patenting. The 61 MEMS universities currently account for 312 MEMS patents.

7. The study uses the comprehensive patent database developed jointly by the Center for Regional Economic Issues at Case Western Reserve University and the National Bureau of Economic Research.

8. The analysis of MEMS patents was done by David Hochfelder, a research assistant on the ATP project.

There is considerable international competition involving MEMS technology. The technology is concentrated in a few countries. As shown in Figure 12, our systems analysis of MEMS technologies ranks the United States first, Japan second, followed by Germany, France, and Great Britain. Rank is based on each country's systems value as a MEMS source (i.e., our fuzzy estimate of each country's contribution to MEMS technologies built up from the nodes in the MEMS network). Spillovers occur across international boundaries. Figure 13 shows the balance of MEMS spillover flows for seven countries with the largest MEMS concentrations (flow in minus flow out). What the data show for MEMS is that higher-order countries are net exporters of the technology. The balance of MEMS knowledge flows between the United States and Japan favors the United States.

The results suggest the potential for ATP to use the methodology to become more knowledgeable in its project support by incorporating knowledge of spillover networks. For example, one important consideration for ATP could be to use understanding of R&D networks to measure any knowledge spillover trade flows from its projects. Although a technology's importance can be evaluated by using total R&D network spillovers, the methodology permits separating these spillovers into two components, those benefiting the United States and those benefiting other countries.

THE MEMS R&D NETWORK

We can describe networks in detail: about 400 technologies (patent classes), individual R&D labs by organization, and metropolitan region location. Each member's position is simultaneously located in the three spaces by year. The data cover the period 1985 to 1995. We can also analyze change (i.e., the evolution of R&D networks). For example, in the case of MEMS, given the importance of investing in enabling networks, an important issue is: Are

FIGURE 12

World's Top Five Sources of MEMS

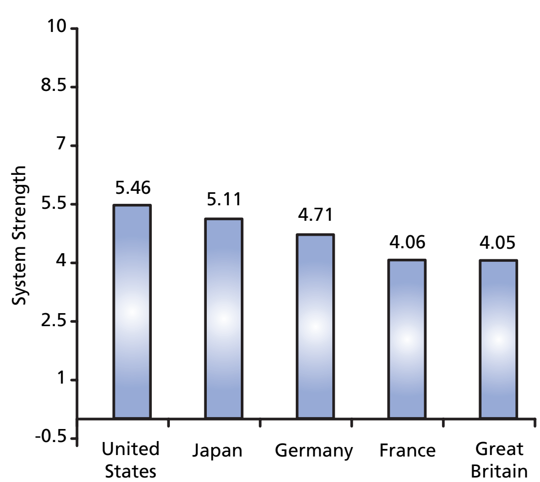


FIGURE 13

Balance of International Flows in MEMS

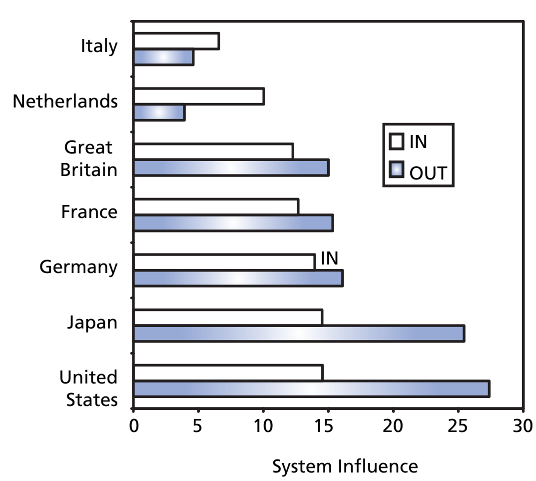
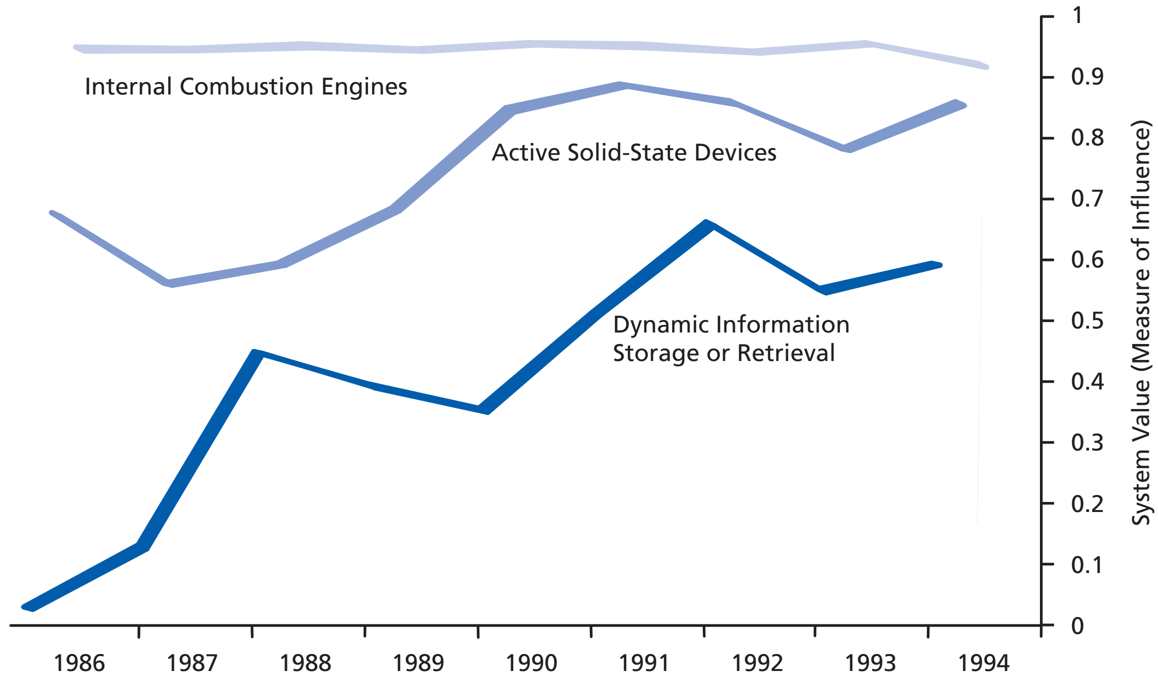


FIGURE 14

The Increasing Role of Computer-Related Technology in the Auto Network



university members becoming less important sources of the technology while companies become increasingly important network members?

Also, as mentioned earlier, it is possible to analyze evolution of networks surrounding particular industry-based technologies. (These are developed in greater detail later.) We construct industry R&D networks by first identifying R&D companies whose R&D is predominantly focused on a particular industry using either CorpTech or more specialized technology information, as in the case of MEMS.[9] The second step involves identifying both significant sources and users of the technology generated by the initial data set. This step is repeated up to the third level of diffusion. The system source and user interactions are then normalized with respect to all interactions contained in the specific industry R&D network. We can then aggregate these measures in various ways to analyze how regions, organizations, and technologies interact over time as sources and learners in developing new technology by industry.

Take one specific example, auto technologies. See Figure 14. What the figure shows is the growing importance of computer technology as an auto ingredient. This result reflects the

9. The CorpTech database is a directory of about 50,000 high-tech companies provided by Corporate Technology Services, Inc. The file used here is 1996. CorpTech estimates that the file contains 99 percent of companies over 1,000 employees, 75 percent of companies with 250–1,000 employees, and 65 percent of companies with fewer than 250 employees.

increasing significance of nontraditional sources of auto technology (sources other than the major assemblers and component suppliers, such as IBM). The system value is normalized with respect to all interactions in the auto R&D network, so that the different technologies can be compared in one figure. (In normalizing the value for each separate auto technology, the minimum equals "0" and the maximum equals "1". Automotive technologies with values closer to 1.0 are more influential within the set of all automotive technologies. The procedure parallels the calculation that was made for the individual patent citations, which also range from 0 to 1).

Each R&D network's spillovers can be analyzed across all three dimensions. This capability might provide ATP with the means to select projects associated with enabling MEMS R&D networks, or to measure knowledge spillovers associated with ongoing or completed projects. One hypothesis could be that ATP's funding would be more "enabling" (i.e., draws on more basic, early-stage research, and stimulates broader, more influential spillovers across a wider spectrum of important uses), if it leverages other important federal R&D funding.

Only the findings for each network's most influential members as influential *sources* are shown. These particular examples characterize the R&D network *sources* of MEMS technologies. A parallel network exists for spillover use networks.

Technologies

Here we illustrate the system importance of leading technologies in the MEMS network. See Figure 15. The top five MEMS technologies, which we can think of as the MEMS core, are Semiconductor Device Manufacturing Process, Metal Working, Electricity: Electrical Systems and Devices, Incremental Printing of Symbolic Information, and Optics: Systems (including Communication) and Elements. Our analysis shows that each of the top five MEMS technologies plays an influential role in both auto and aerospace industry technologies. ATP-funded projects might be more likely to create greater R&D spillovers if projects focused on core MEMS technologies.

Organizations

Figure 16 shows the top MEMS organizations ranked by system influence. An organization's position on the list is determined by the magnitude of spillovers generated by its associated R&D network. In other words, organizations that are members of the most influential segments of the MEMS network are higher ranked. The analysis shows that IBM ranks first as a source, followed by U.S. Phillips, NEC, General Electric (GE), Texas Instruments, and United Technologies. Some of the key universities are (in order of their influence) MIT, Stanford, and Berkeley. The full organization list (not shown) would reflect the prominence of federal labs and universities as prominent MEMS sources. The ATP's projects are more likely to create more R&D spillovers if they involve organizations or joint ventures whose R&D is core to MEMS (i.e., they operate within influential portions of the MEMS network; that is, interacting with influential neighbors).

FIGURE 15

Fuzzy MEMS Technology R&D Networks

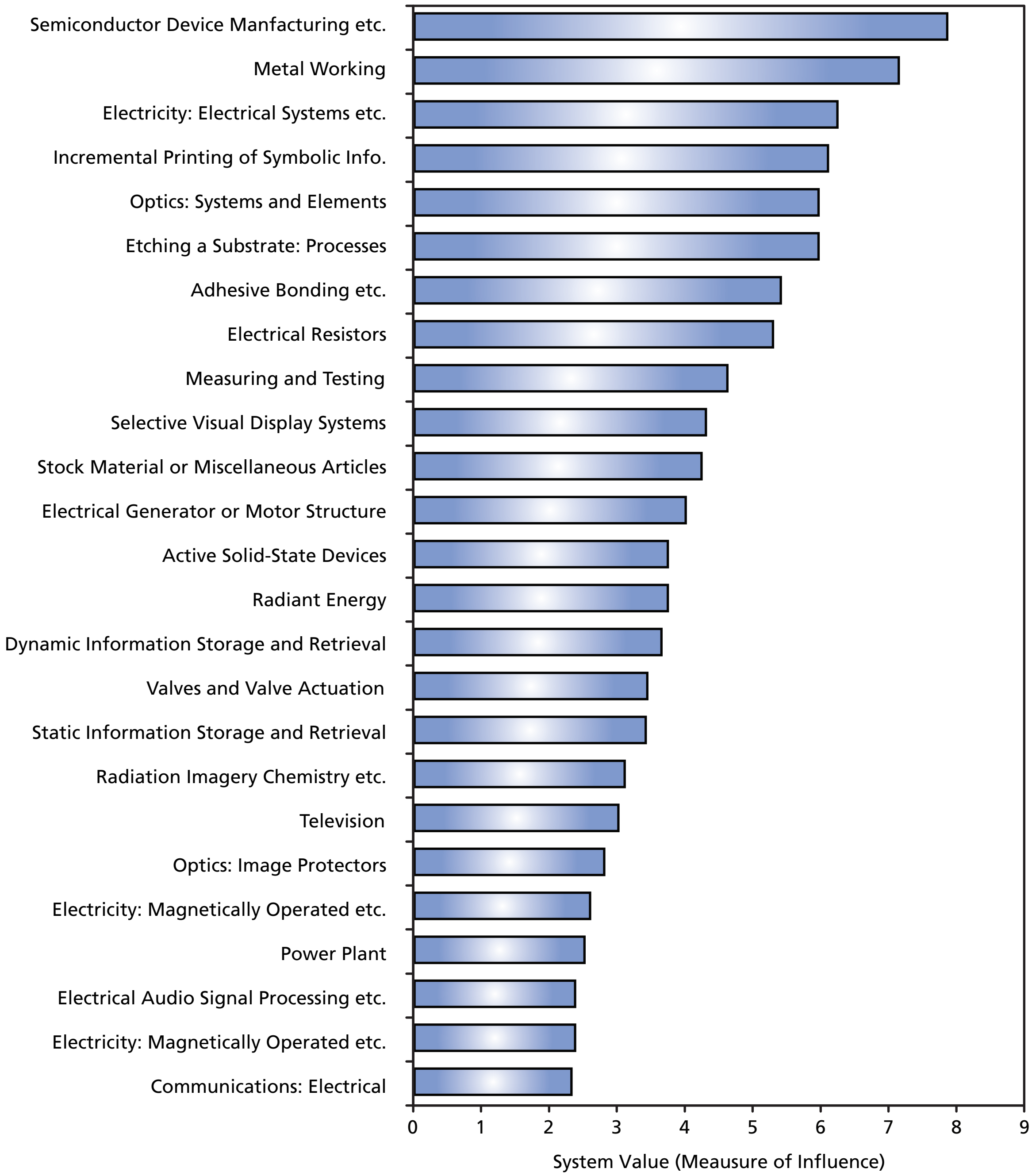
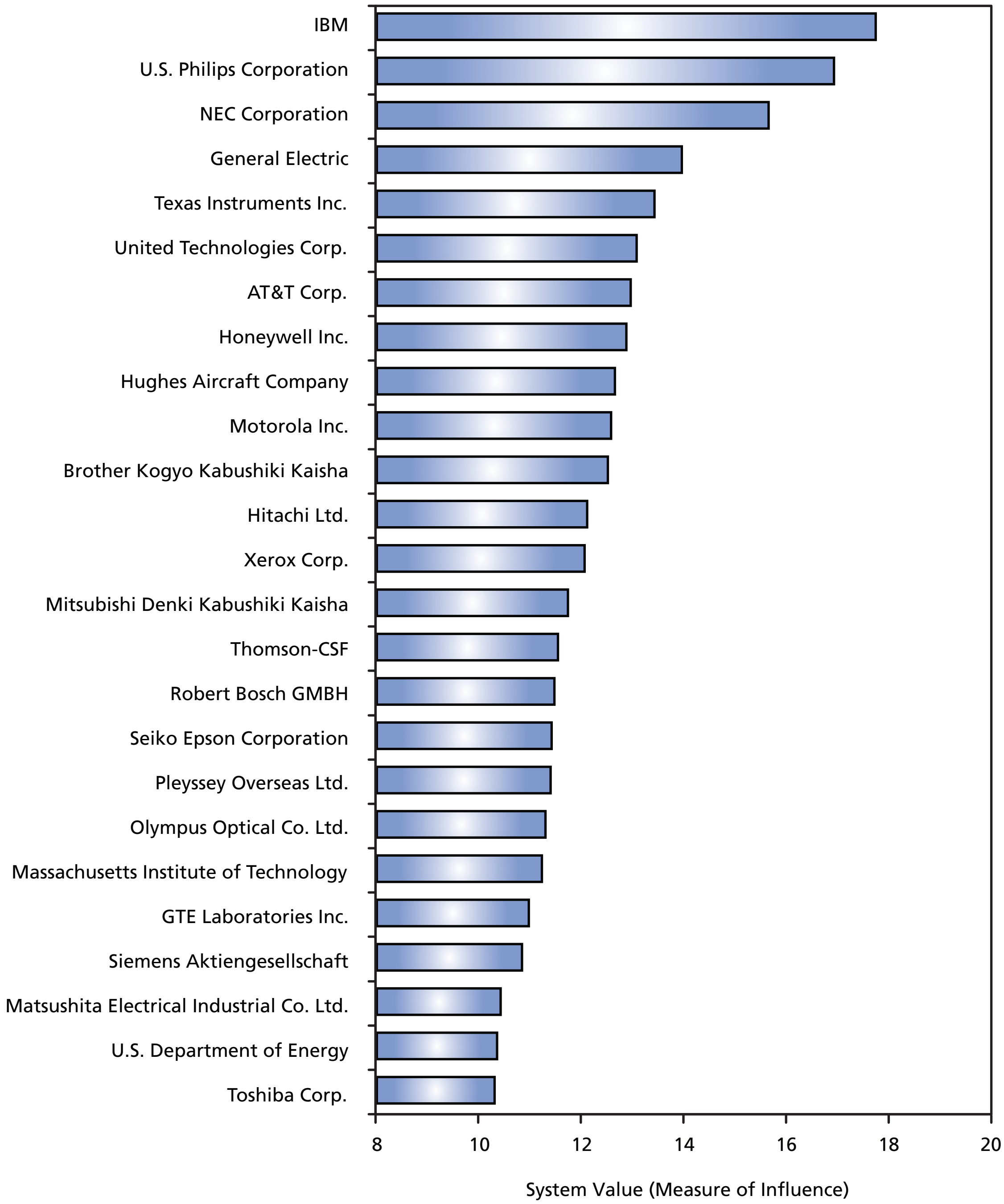


FIGURE 16

Fuzzy MEMS Organization R&D Networks

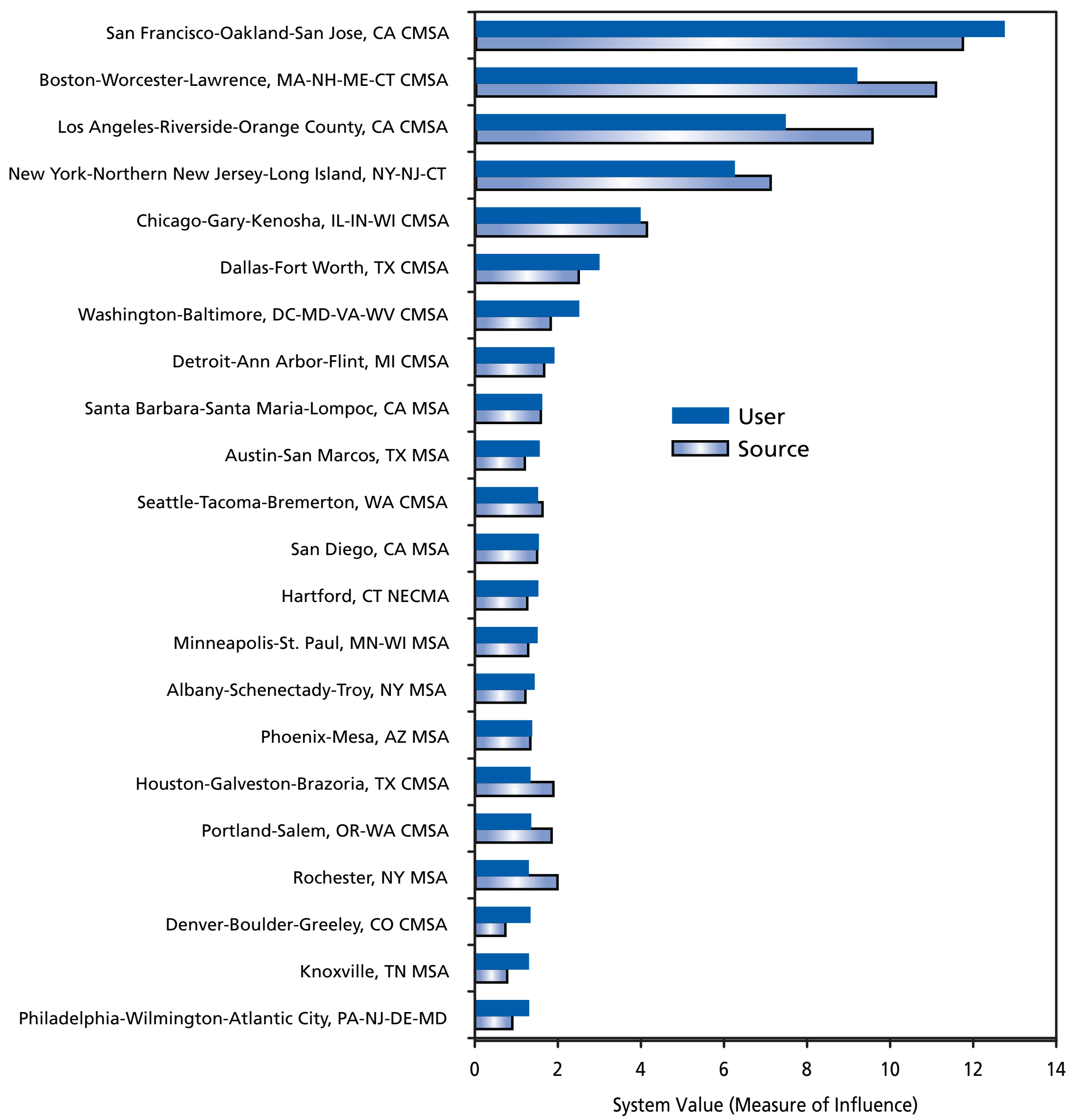


Regions

When ATP selects a project, it is also inadvertently selecting the geography of R&D spillovers and, consequently, the regional beneficiaries of ATP-funded R&D. Previously we saw that MEMS is highly concentrated in a small set of leading countries. As shown in Figure 17, an analysis of U.S. regions reveals that development of MEMS is also geographically concentrated in a few top regions: San Francisco, Boston, Los Angeles, New York, Chicago, and Dallas account for the largest share of the technology.

FIGURE 17

Regional Distribution of MEMS Technology



Our hypothesis is that an important characteristic of enabling networks is location in a successful regional agglomeration supportive of new technology development. Although influential R&D organizations search in a global R&D network, in a precompetitive, incubation phase, geography plays a critical function; that is, the accumulation of a critical mass of strong network connections that speed growth of the enabling technology. In general, we expect that as local activity increases, the volume of important spillovers grows but an increasingly large fraction becomes external as a region develops higher-order R&D organizations with worldwide connections. Because R&D labs have a specific location, an agglomeration of strong R&D networks serves a dual function; that is, good regional sources are also good learners. Local R&D networks are strong only if they are solidly linked to the global network.

Both San Francisco and Boston are important MEMS sources—both regions are influential external sources of MEMS. Only San Francisco, however, appears to be developing significant local spillovers. Based on this particular enabling characteristic of networks, it is reasonable to anticipate that ATP investments in projects with strong connections to San Francisco's MEMS network may produce more spillovers and faster development of the technology. Investing in ATP projects located in these large R&D agglomerations will likely produce a higher social rate of return because the investment builds from critical mass, creating increasing spillover returns to R&D. The ATP's investments in leading regions may also result in capture of a greater share of spillovers by the United States. The reason stems from the expected increasing returns to R&D coupled with a faster rate of technology diffusion and commercialization. Even if ATP's investments are concentrated in leading regions, spillover benefits would be shared by a much larger set of U.S. industries and regions that draw from incubator regions.

COMPARISON OF SYSTEMS RESULTS TO SIMPLE CITATION COUNTS

To illustrate the spillover process, we have utilized the systems approach to generate information along a variety of different dimensions. It is worth asking, however, whether the systems approach presents any differences in its basic ranking of organizations as potential sources of spillovers, relative to what we would conclude by simply counting the patent citations of different organizations. In the case of MEMS, the answer is that the two ranks are correlated but several interesting and important differences emerge.

Table 5 ranks leading MEMS organizations based on both measures. For example, United Technologies ranks considerably lower by the total citations (12) than by the systems source rank (6). The reverse is true for Xerox, which has a citations total rank of 2 and a systems methodology source rank of 13. If ATP were evaluating MEMS projects involving these two organizations solely based on "expected" R&D spillovers, the simple citation count would prefer Xerox (ranked 2 versus 12) while the systems methodology would prefer United Technologies (ranked 6 versus 13). In other words, our analysis of MEMS R&D networks tells us that United Technologies is more influential than Xerox as a source of MEMS technologies, although this difference is not visible in the simple counts.

TABLE 5

Comparison of Organizational Ranks by Existing (Citation) and System Methods

Organization	Citation Method Rank	System Source Rank	System Use Rank
IBM	3	1	1
U.S. Philips	5	2	3
NEC	7	3	9
General Electric	6	4	6
Texas Instruments	4	5	4
United Technologies	12	6	22
AT&T	14	7	17
Honeywell	15	8	30
Hughes Aircraft	21	9	10
Motorola	18	10	2
Brother	16	11	46
Hitachi	10	12	26
Xerox	2	13	13
Mitsubishi	20	14	37
Thomson-CSF	22	15	20
Robert Bosch	27	16	21
Seiko Epson	33	17	36
Plessey Overseas	49	18	79
Olympus Optical	9	19	11
MIT	8	20	15

Differences in the two ranks are due to differences in R&D networks. The systems analysis assigns higher value to the intensity and spread of spillovers through the network's three dimensions: technology, organization, and geography. If an organization's system rank is higher (lower) than its citation count rank, this means that the higher-order interactions (the influence of the nodes citing the patents of the organization of interest) are higher (lower) than average.

5. ATP and the Case of Short-wavelength Sources for Optical Recording

NSIC AND SWAT

We have chosen to illustrate the usefulness of the R&D spillover systems methodology with an important ATP-funded project: the Joint Venture on Short-wavelength Sources for Optical Recording (SWAT). This joint venture was one of ATP's initial (1991) awards. It received a grant of $5.4 million as part of a $14.6 million cost-shared five year project budget. SWAT's purpose was "to develop a short-wavelength integrated laser source."[1] Importantly, the joint venture was managed by the National Storage Industry Consortium (NSIC), which was formed as a response to increased Japanese competition in the recording industry.[2]

In 1994 the data storage device industry was a $50 billion per-year market. Roughly two thirds of the market belonged to U.S. companies and one third to Japanese firms. Although optical recording technology represented a relatively small share of the market, in 1991 it was viewed as the most promising new technology. The fact that Japanese companies controlled 80 percent of the optical recording market represented a threat to U.S. industry dominance in the recording industry. As a response, the SWAT joint venture proposed to develop an integrated short-wavelength laser source for optical recording.

NSIC, which managed the SWAT joint venture, had 39 companies and 35 universities and government laboratories as members in 1994. The ATP-funded joint venture involved a handful of NSIC members: Applied Magnetics, Bernoulli Optical Systems, Eastman Kodak,

---

1. Information concerning the SWAT Joint Venture comes from an ATP study by Link (1994). According to NIST, "Multiple lasers (for multichannel recording), solid-state components to increase frequency, and a nonmechanical scanning system for tracking the beams will be fabricated in a single device. Besides greatly advancing the art of diode laser sources and optical modeling, these new heads would revolutionize the industry—data storage four times as dense, data read and write speeds twice as fast or better, in significantly small, more rugged devices." See http://jazz.nist.gov/atpcf/prjbriefs/prbrief.cfm? ProjectNumber=90-01-0231.

2. NSIC's objective was "…to enhance the competitiveness of the U.S. recording industry through a strategic plan to form joint venture programs on precompetitive technologies and to coordinate technology developments among corporations, universities, and governmental organizations." See http://www.nsic.org/members.html.

IBM, Maxoptix, and the Optical Storage Center of the University of Arizona. Importantly, SWAT membership changed considerably over time, even though most of the current members were in the original group. The current participants consist of Carnegie-Mellon University; Eastman Kodak, Mass Memory Division, Research Labs (Rochester, New York); IBM (San Jose, California); Uniphase (San Jose, California); and University of Arizona. Only Uniphase was not among the original joint venture members.[3]

A SYSTEMS ANALYSIS OF OPTICAL WAVEGUIDE TECHNOLOGIES

The ATP case study presents an opportunity to illustrate the R&D spillover systems analysis as applied to an ATP-funded joint venture. Our purpose is to illustrate the potential of the methodology, and not to evaluate it . An evaluation would require considerably more information and more current data. Our data and systems analysis covers the period 1985 to 1995. (The joint venture grant was awarded in 1991.)

We began our analysis of Optical Waveguide (OWG)[4] technologies by first developing a core set of OWG patents. We started with patents held by the SWAT members as well as OWG patents assigned to the other NSIC members. We then expanded our core set of OWG patents by using the Science Citations Database to identify leading researchers in the field and their OWG patents. Citations to the patents were used to identify additional OWG candidate patents. Each candidate patent's abstract and exemplary claims were read to ensure that the patent was an OWG technology.[5] Using our systems analysis methodology, we built the full OWG network from the initial set of OWG patents. The full OWG network was analyzed in four spaces: technology, organization, region, and time. Each member's network position is simultaneously located in the four spaces. These particular examples characterize the R&D network as *sources* of OWG technologies.

Technologies

The resulting top 20 most influential OWG technologies are shown in Figure 18. The top five OWG technologies are Coherent Light Generators (372); Optics: Systems (Including Communication) and Elements (359); Optical Waveguides (385); Dynamic Information Storage or Retrieval (369); and Compositions (252). (The number in parenthesis is the U.S. Patent and Trademark Office's patent class number.) The technology slice of the network could be used to analyze the technology area where research activity is most intensely occurring. For instance, by analyzing the system value associated with each component technology, it may be possible to identify which technologies (patent classes) are growing or declining in importance within the OWG network. Also, with more current data, ATP might

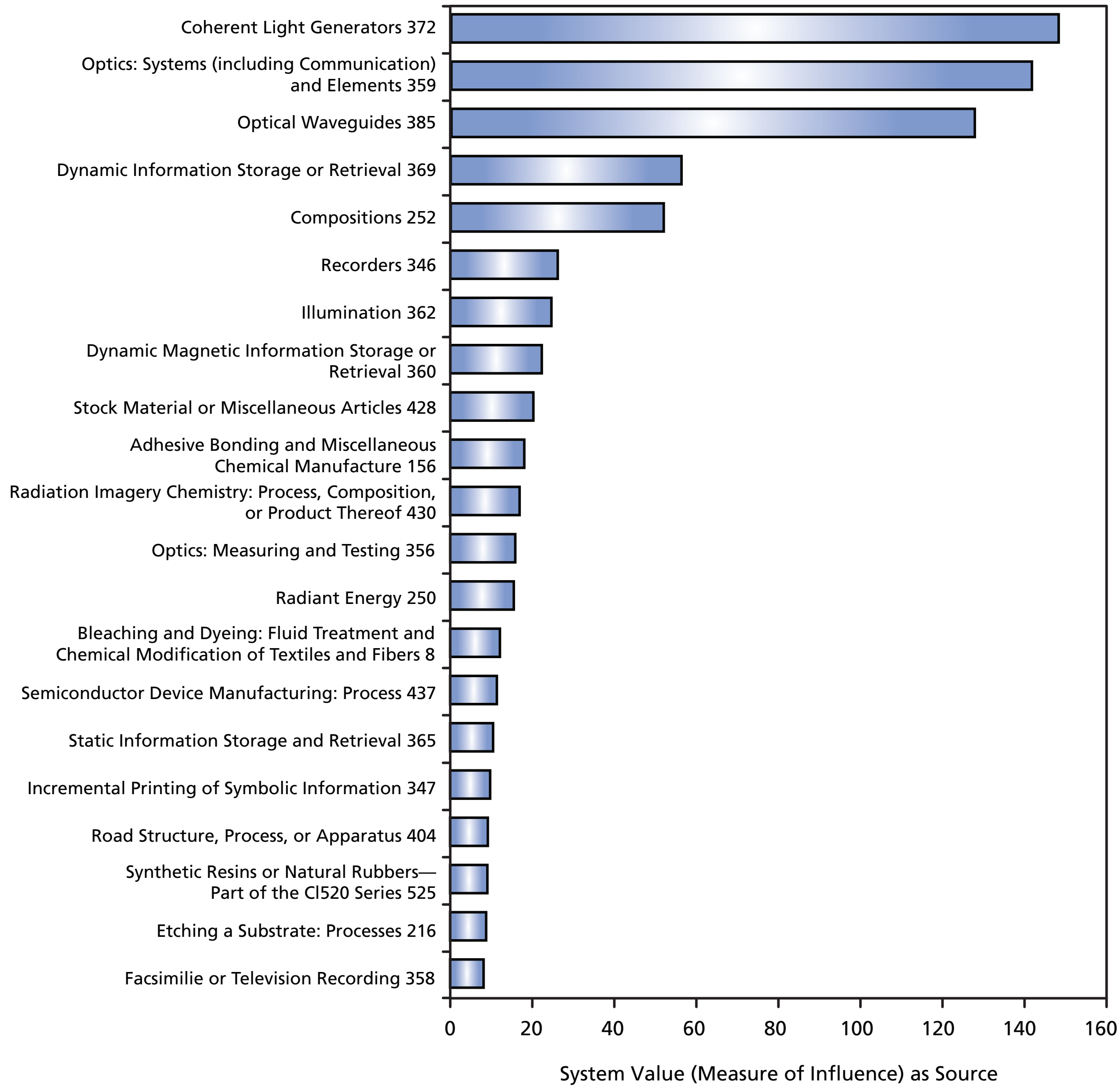
3. Other companies had shown strong interest: 3M, DEC, HP, KOMAG, Quantum, Storage Technology, Carlisle Memory Products, and Visqus. CMU's Data Storage System Center also expressed early interest, and eventually became a SWAT member.

4. Optical waveguide is any structure having the ability to guide optical energy. Optical waveguides may be (a) thin film deposits used in integrated circuits or (b) optical fibers. See ATIS website <http://www.atis.org>.

5. The analysis of OWG patents was done by David Hochfelder, a research assistant on the ATP project.

FIGURE 18

Top 20 Optical Waveguide Technologies



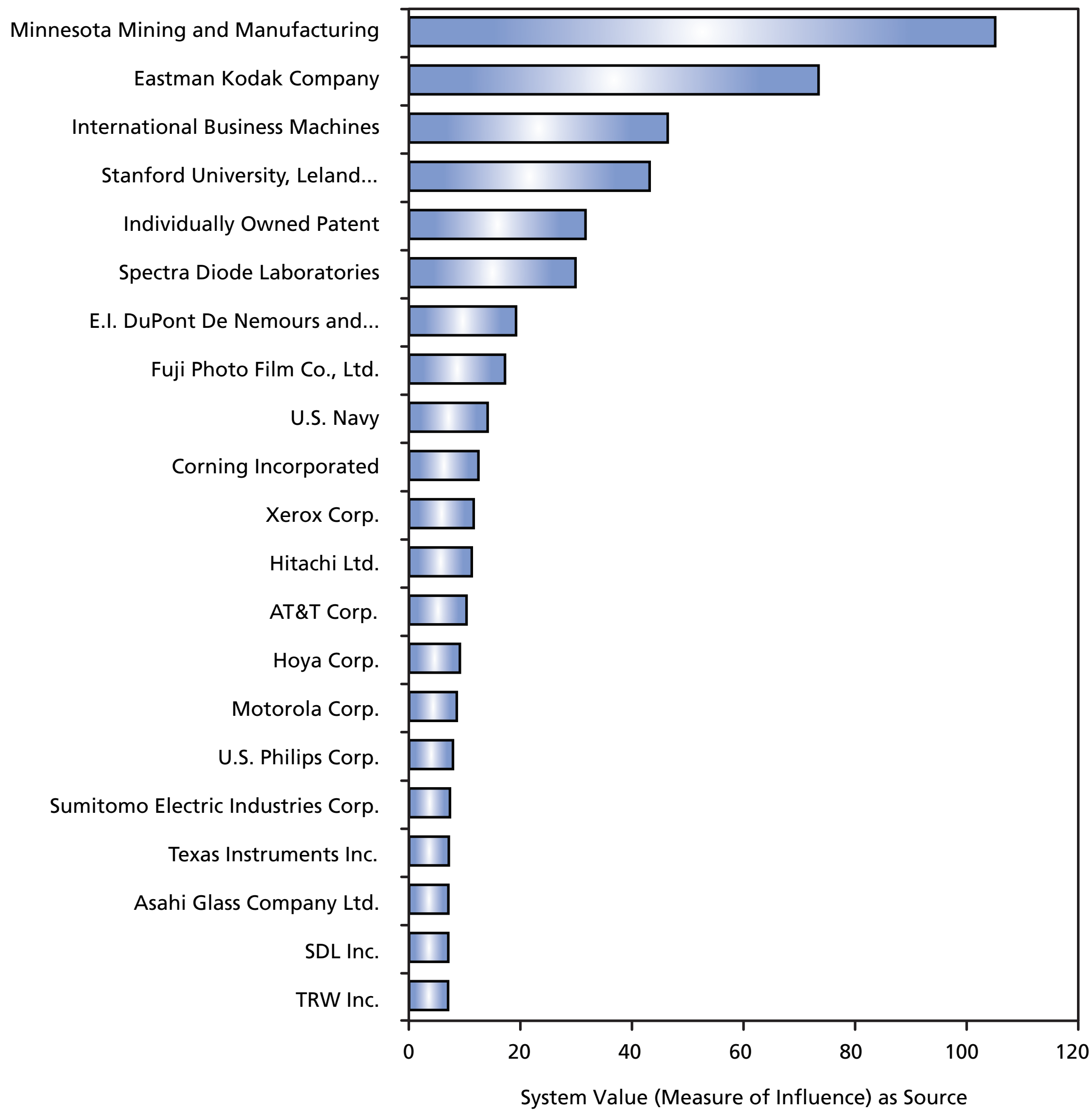
be able to determine what joint venture participants were generating the most activity relative to the development of the proposed technology's most important components.

Organizations

Figure 19 shows the top 20 OWG organizations ranked on system influence (i.e., each organization's total system influence within the OWG network). An organization's position on the list is determined by its R&D spillover contribution to the OWG network. For example, the top five ranked organizations are 3M, Eastman Kodak, IBM, Stanford, and

FIGURE 19

Top 20 Waveguide Organizations



Spectra Diode Laboratories. Suggestive of OWG's emerging character, several universities as well as government labs are ranked among the top 50 organizations: Stanford University and MIT, and three government labs (one each of the U.S. Navy, U.S. Air Force, and U.S. Army). As a gauge of the degree of competition within the recording industry, Japanese companies are among the influential members of the OWG network, including Fuji Photo, Hitachi, Hoya, Sumitomo Electric, and Asahi Glass.

Table 6 lists the original SWAT members, the current SWAT members, and other NSIC members determined to be ranked among the top 50 most influential organizations within

TABLE 6

OWG System Influence of Joint Venture and Other NSIC Members

Original Members	Other Top 50 Current Members	System-Ranked NSIC Members	Value Rank
Applied Magnetics			Not ranked
Bernoulli Optical Systems			Not ranked
Eastman Kodak	Eastman Kodak		2
IBM	IBM		5
Maxoptix			Not ranked
University of Arizona	University of Arizona		Not ranked
Carnegie-Mellon University	Carnegie-Mellon University		Not ranked
	Uniphase		Not ranked
		3M	1
		MIT	13
		Polaroid	17
		HP	27
		Energy Conversion Devices	28
		Texas Instruments	30

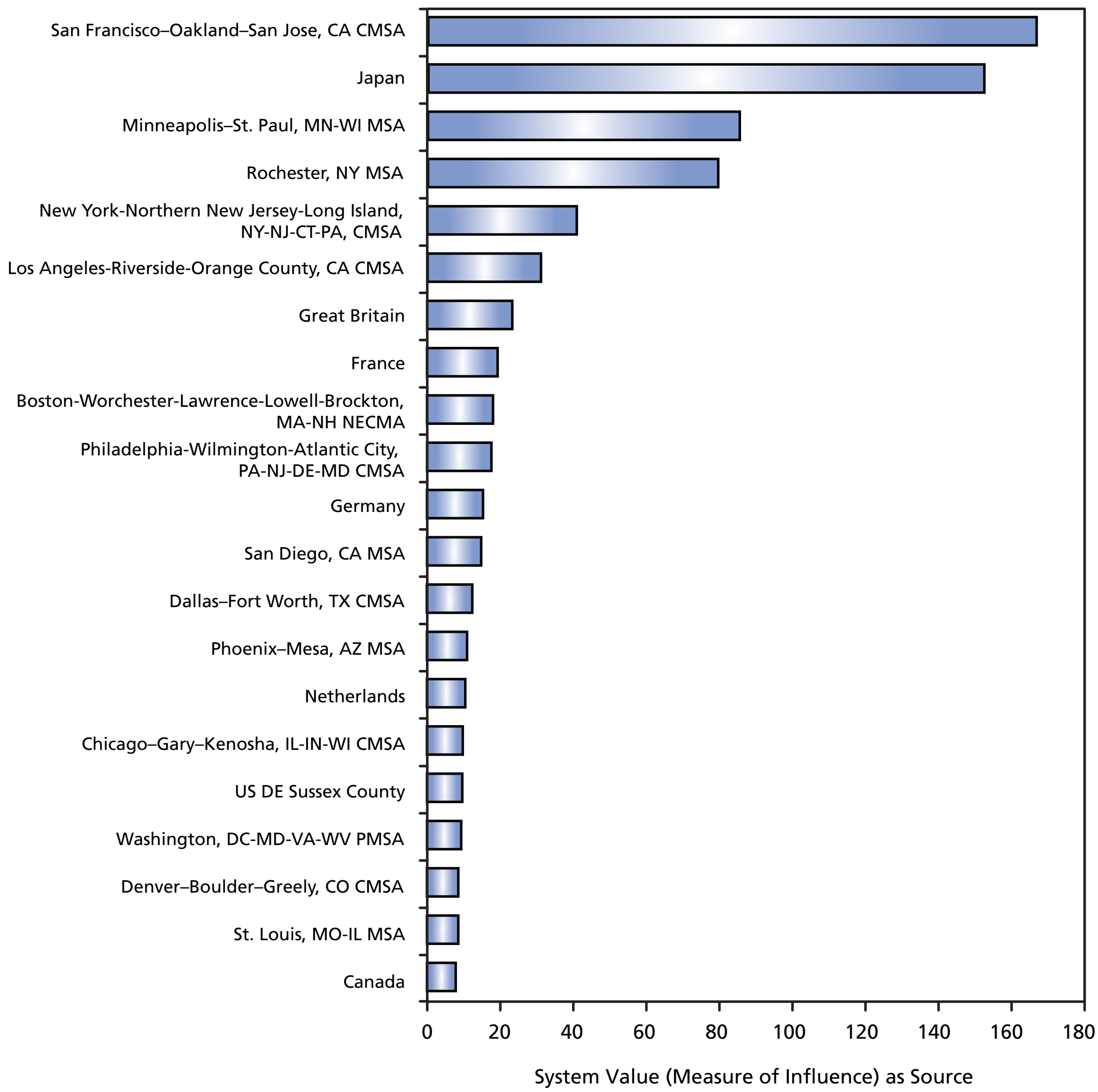
the OWG network but not members of the joint venture. The table also provides each organization's system value rank. Of note, the top four OWG sources are NSIC members; the fifth-ranked organization in Figure 19, Spectra Diode Laboratories, was a subcontractor to the project, but not an NSIC member. (There are also a number of highly important sources of OWG technology that are not members of the NSIC. Included within the top 20 listed in Figure 19 are Corning, Xerox, AT&T, Motorola, U.S. Phillips, and TRW.)

Regions

As shown in Figure 20, OWG technology in the United States is concentrated in a few regions: San Francisco is the first, followed by Minneapolis, Rochester, New York-New Jersey, and Los Angeles. Our hypothesis is that an important characteristic of enabling networks is location in a successful regional agglomeration supportive of new technology development. It is not surprising that the NSIC formed in San Francisco, which is the top OWG agglomeration. Two of SWAT's five active members are located in the San Francisco Bay Area: IBM's Almaden Research and Uniphase are both located in San Jose.

FIGURE 20

Top 20 Optical Waveguide Regions



LIKELY SPILLOVERS FROM THE ATP-FUNDED PROJECT

The results suggest that ATP could reasonably have expected its investments in the joint venture to generate a large volume of R&D spillovers for three reasons. First, the OWG network is dominated by U.S. organizations. The top six most influential organizations are U.S. owned: 3M, Eastman Kodak, Stanford University, IBM, Corning, and the U.S. Navy. Two of these are joint venture members. Second, several top-ranked companies within the OWG network are NSIC members and joint venture participants. The inclusion of at least one top-ranked member within the network might lead one to expect significant R&D

spillovers stemming from an ATP grant. By including at least one top-ranked member, the R&D spillover contribution of an unranked joint venture member would be enhanced by connecting it to more intense segments of the network. Third, because the joint venture was lodged within NSIC, ATP could also have reasonably expected its investments to produce even broader spillover benefits (i.e., effects beyond the immediate joint venture members).

Although only a few members participated in the joint venture, NSIC represents a potentially powerful mechanism for magnifying R&D spillovers from the joint venture. It is likely that other partnerships and linkages already existed or developed between joint venture members and nonparticipating NSIC members. For instance, ATP's requirement that ownership of the technology be assigned to a for-profit joint venture member might reduce a university's incentive to participate in the joint venture. Universities also might participate and contribute to the technology's development indirectly outside SWAT (e.g., through other forms of university-industry interaction, such as consulting, company-funded applied research at universities, and through their graduates taking jobs with NSIC companies). Stanford University, for instance, the first ranked university within the OWG network, probably has multiple research and consulting relationships with various NSIC members, whether SWAT participants or not. If the data were available, a reasonable hypothesis might be that Stanford's graduates represent another R&D spillover mechanism as they take jobs with NSIC members, forming even stronger OWG network connections, later traceable to patents. Such informal mechanisms associated with other ATP-funded projects could also be analyzed.

OTHER OBSERVATIONS

ATP's support of SWAT through the NSIC increases the likelihood that its investment in the OWG will generate a large volume of R&D spillovers. The ATP's support likely influences the R&D decisions of several highly connected OWG network members, both directly (joint venture members) and indirectly (through NSIC). The OWG technology also exhibits several enabling technology characteristics. In particular, universities and government labs play significant roles, and the technologies are highly concentrated geographically.[6] By investing in a technology in which universities and government labs play a prominent role, ATP's investment can leverage millions of dollars of federal government support of basic research funding and likely increase the odds of commercialization and result in capture of economic benefits in the United States.

What made the OWG a good joint venture candidate? First, it presents a large market, dominated by U.S. companies, in which the most promising, emerging technology is

6. We tentatively identify the characteristics of enabling technologies as the following: universities and government labs play significant roles, the network is sparse and evolving, the technology is new (cited patents are relatively current), total system spillovers increase significantly and get diffused rapidly, influential companies perform significant basic research, and technologies become geographically concentrated in important regions serving as incubators. (For startups, with an important technology but too small to be picked up by our methodology, we may look for venture capital funding.)

controlled by Japanese companies. Second, U.S. companies representing highly influential sources of technology in the recording industry formed NSIC as a collective response to increased competition (in other words, a partnership among industry participants was forming). Third, a number of universities and government labs with research related to optical recording became associate members of NSIC, indicating that the consortium and the joint venture would likely leverage the millions of dollars of federal R&D support to universities and government labs. Finally, several key companies from NSIC were among the joint venture participants. Although the joint venture consists of a small subset of NSIC organizations, facilitated by increased interaction among NSIC members, ATP's investment in the joint venture may be substantially magnified by increased R&D spillovers and a faster rate of technology development.

6. R&D Spillovers from MEMS and Optical Recording

One of the ATP's mission goals is to support projects that deliver broad-based economic benefits. To what extent are the R&D spillovers widely distributed across industry sectors? An industry-wide distribution of spillovers suggests that a firm or joint venture will be less able to capture all of the benefits. It also suggests that such projects will contribute to the development of the new technology in more geographic regions. This section illustrates the application of the fuzzy methodology to the analysis of spillovers that cut across five broad industries: automotive, aerospace, information technologies, advanced materials, and biomedical devices.

Imagine an ATP-funded MEMS R&D joint venture involving MIT, Honeywell, and Xerox. As shown earlier in Figure 10, Honeywell and Xerox are part of the tightly knit MEMS network centered on MIT, which was the top-ranked university in MEMS during the period 1985 to 1995. Honeywell is ranked 8 and Xerox is ranked 13 in this full MEMS network. Suppose that the ATP-funded collaboration supports MEMS research that advances technology in dynamic information storage and retrieval. This patent class is highly ranked in both MEMS and in the auto industry. (Recall that Figure 14 illustrates the growing relative importance of computer-related technology in the auto network.) It is reasonable to anticipate that, with diffusion, the auto industry would be a significant beneficiary of MEMS spillovers from the hypothetical project.[1] In fact, because organizations will vary in the extent to which their networks overlap the networks underlying different industries, ATP's project selection has the potential to significantly influence the distribution of spillovers across industries.

---

1. Even though this patent class grew significantly in importance within the auto industry, its average rank over the period 1985 to 1995 was not high. One implication is that it may be possible (and desirable) to analyze emerging technologies that trigger the evolution or change in the mix of influential technologies within particular industry sectors, such as autos.

METHODOLOGY

One way to examine this issue is to analyze the overlap between each of the two networks—MEMS and OWG—and the network of technologies associated with several broad industries. To accomplish this we separately analyzed the R&D networks related to specific technologies vital to the five major industries discussed.[2] This step involved identifying companies whose R&D was predominantly focused on a particular industry. For example, we developed an initial list for the auto industry using several industry directories to identify major auto assemblers and auto suppliers. We also drew extensively from the CorpTech database, a directory of about 50,000 high-tech companies, in developing our initial list. (CorpTech develops broad industry technology categories consistent with four of our five industries.) The company lists were then used to develop a company-level patent database for each broad industry. Finally, we constructed the R&D networks associated with each industry using our fuzzy systems methodology.

By using the same methodology applied to MEMS and OWG, we identify the most influential technologies underlying each industry as the aggregate of their system influence as sources by industry. For example, Table 7 lists the top 25 technologies (patent classes) based on the aggregation of their system influence for information technology as sources. As before, the system value is the sum of interaction values across all links representing the network for each broad industry. Although we ignore location and organization in our presentation, their effects are incorporated in the analysis and system measures.

To make the methodology more specific, assume that we have two R&D networks: MEMS and biomedical devices. Our objective is to determine the extent to which the two networks overlap. The hypothetical intersection is illustrated in Figure 21.[3]

In essence, our analysis involves determining the truth value of membership within both technologies. We evaluate the joint truth value as the product of the two truth values. We then sum the product of the two truth values across all technologies within each R&D network. To summarize, we evaluate the intersection of the two fuzzy networks in several steps:

1. Let: i = organization and j = technology
2. T_{ijm} = the truth of technology j specific to organization i in "MEMS"
3. T_{ijb} = the truth of technology j specific to organization i in "biomedical"
4. $T_{ijm} \times T_{ijb}$: the truth of technology j specific to organization i in both "MEMS" and "biomedical"
5. Sum over i ($T_{ijm} \times T_{ijb}$) for all i, j, k (all earlier truths by i, j, k).

Thus our analysis occurs at the i, j, k, t level (a specific organization's lab), in a particular region, working within specific patent class, by year.

2. See Fogarty et al. (1999).

3. The picture is somewhat misleading because the R&D networks are hierarchical, with the core (most influential) components located closest to the center of each network, and less influential component technologies located further from the center.

TABLE 7

The Top 25 Technologies: System Influence by Source in the Information R&D Network, 1985–1995

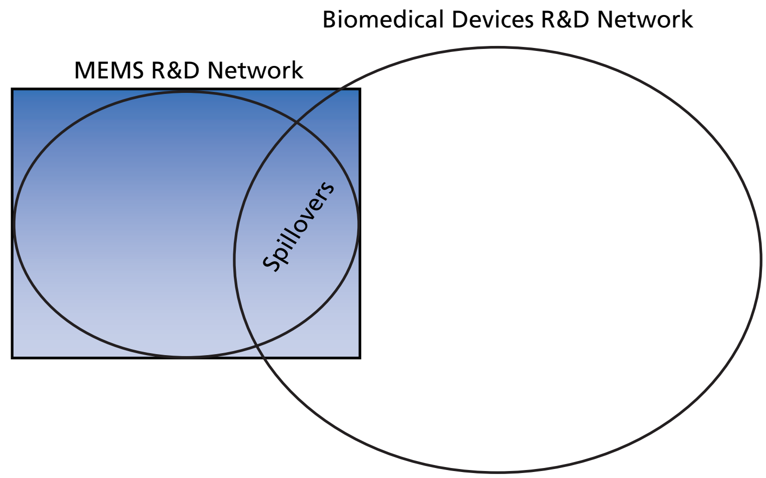
Class	Patent Class	Source
395	Information Processing System Organization	1,586.9
364	Electrical Computers and Data Processing Systems	1,351.3
361	Electricity: Electrical Systems and Devices	892.4
324	Electricity: Measuring and Testing	860.5
371	Error Detection/Correction and Fault Detection/Recovery	739.3
250	Radiant Energy	708.6
428	Stock Material or Miscellaneous Articles	674.9
370	Multiplex Communications	667.1
327	Miscellaneous Active Electrical Nonlinear Devices, Circuits, and Systems	653.9
359	Optics: Systems (including Communication) and Elements	642.7
375	Pulse or Digital Communications	641.2
360	Dynamic Magnetic Information Storage or Retrieval	620.7
326	Electronic Digital Logic Circuitry	588.8
430	Radiation Imagery Chemistry: Process, Composition, or Product Thereof	566.7
345	Selective Visual Display Systems	540.3
437	Semiconductor Device Manufacturing: Process	533.6
439	Electrical Connectors	533.2
257	Active Solid-State Devices (e.g., Transistors, Solid-State Diodes)	517.3
380	Cryptography	514.5
347	Incremental Printing of Symbolic Information	504.5
156	Adhesive Bonding and Miscellaneous Chemical Manufacture	475.3
340	Communications: Electrical	471.7
235	Registers	464.3
348	Television	457.6
400	Typewriting Machines	445.2
341	Coded Data Generation or Conversion	426.6

RESULTS

Tables 8 and 9 contain the spillover calculations for MEMS and OWG across the five industries. The first number represents estimated spillovers, and the second number in parenthesis provides the system value rank of the patent class *within* each industry's R&D network. For example, spillovers from MEMS' top patent class (semiconductor device manufacturing (437)) to the same class within information technology is 63. The (16)

FIGURE 21

Hypothetical MEMS Spillovers across Biomedical Devices



indicates that this class was ranked sixteenth in importance (influence) within the information technology network.

Not surprisingly, because of the defense and aerospace origins of the technology, MEMS has disproportionately created spillovers for the aerospace industry (roughly 60 percent). The second most important recipient of MEMS spillovers was biomedical devices (about one-fifth of the total). The remaining 20 percent of spillovers were somewhat evenly distributed among the other three industries. The implication is that MEMS spillovers have influenced many of the core technologies of four of the five industries (the exception was biomedical devices). This is shown by the relatively large number of top 25 technologies for each industry that were influenced by MEMS R&D spillovers.[4, 5]

One limitation of these calculations is that the quantities represent an average covering the period 1985–1995. Ideally, the analysis would examine trends. It is very likely that an analysis of the trend in MEMS system influence by industry would find a broader diffusion of the technology. For instance, MEMS spillovers contributed relatively little to the core technologies in biomedical devices (as Table 8 shows, no high-ranked technologies within biomedical devices were among the high-ranked MEMS technologies). Our result reflects that the MEMS R&D network, however, is based on all MEMS, even though, as pointed out earlier, MEMS technologies can be grouped into several categories, one of which is in the rapidly growing bio-MEMS field. Moreover, as the auto example points out, special attention in future research should be focused

---

4. The MEMS totals and percent by industry are auto: 1,700 (8.4 percent); aerospace (12,185 (60.0 percent); advanced materials: 1,422 (7.0 percent); information technology: 1,117 (5.5 percent); and biomedical devices: 1,117 (19.1 percent). Amounts for OWG are auto: 49 (11.5 percent); aerospace: 174 (40.9 percent); advanced materials: 53 (12.4 percent); information technology: 59 (13.8 percent); and biomedical devices: 91 (21.4 percent).

5. Appendix B provides 6 tables that show the Top 25 Technologies in six industries: (1) aerospace, (2) information technology, (3) automotive, (4) advanced materials, (5) bio-medical devices, and (6) Mems.

TABLE 8

R&D Spillovers from MEMS to Selected Broad Industries, 1985–1995

Class	Top 25 MEMS Technologies	Auto	Aerospace	Advanced Materials	Information Technology	Biomedical Devices
437	Semiconductor Device Manufacturing: etc.				63 (16)	
29	Metal Working	191 (7)	1321 (7)	179 (18)	65	
361	Electricity: Electrical Systems etc.	189 (6)	943 (6)		120 (3)	474
347	Incremental Printing of Symbolic Info.				76 (20)	
359	Optics: Systems and Elements	132 (12)	940 (10)		86 (10)	521
216	Etching a Substrate: Processes					
156	Adhesive Bonding etc.	118 (18)	995 (25)	138 (15)	57 (21)	565
338	Electrical Resistors					
73	Measuring and Testing	205 (5)	1664 (9)	251 (7)		832 (19)
345	Selective Visual Display Systems				66	
428	Stock Material or Miscellaneous Articles	259 (1)	1635 (4)	332 (1)	85 (7)	866 (17)
310	Electrical Generator or Motor Structure	166 (10)	678 (11)			
257	Active Solid-State Devices		(23)		67 (18)	
250	Radiant Energy	161 (13)	1352 (2)	243 (10)	102	620
369	Dynamic Information Storage or Retrieval					
251	Valves and Valve Actuation					
365	Static Information Storage and Retrieval					
430	Radiation Imagery Chemistry: etc.				77 (14)	
348	Television		543		57 (24)	
353	Optics: Image Projectors					
324	Electricity: Measuring and Testing	130 (16)	1157 (5)	166 (17)	133 (4)	
60	Power Plants	149 (8)	(14)	113		
381	Electrical Audio Signal Processing etc.					
335	Electricity: Magnetically Operated etc.					
340	Communications: Electrical		957 (17)		63 (22)	602

on the task of identifying emerging technologies within specific R&D networks, including those that cut across industries. This analysis might lead us to ask if ATP's funding of a project will speed the diffusion of the technology across industries.

As shown in Table 9, the OWG fostered fewer spillovers over the period 1985 to 1995 (the spillover calculation is much smaller); spillovers, however, are more evenly distributed across the five industries than was the case for MEMS. Our spillover estimates produce a smaller number largely because OWG consists of a much smaller slice of technology than MEMS. In this case, aerospace is also the primary beneficiary. This sector received approximately two-fifths of OWG spillovers. Biomedical devices, with one-fifth of total spillovers, was the next most important spillover recipient. The remaining two-fifths was evenly distributed among the other industries: auto, advanced materials, and information technology.

TABLE 9

R&D Spillovers from OWG to Selected Broad Industries, 1985–1995

Class	Top 25 OWG Technologies	Auto	Aerospace	Advanced Materials	Information Technology	Biomedical Devices
372	Coherent Light Generators	13.1	26.8	13.4	15.0	12.3 (18)
359	Optics: Systems (including Communication) and Elements	5.5 (12)	31.6 (10)	6.9 (19)	9.3 (10)	17.8
385	Optical Waveguides	8.3 (23)	30.2 (8)	15.7 (12)	10.1	22.1
369	Dynamic Information Storage or Retrieval	2.0	16.2	4.4	6.0	11.4
252	Compositions	5.0	14.0	2.4 (3)	2.5	7.2 (25)
346	Recorders	2.7	7.9	0.7	1.7	5.6
362	Illumination	1.7	9.1		1.3	0.1
360	Dynamic Magnetic Information Storage or Retrieval	1.3	2.3		1.8 (12)	1.4
428	Stock Material or Miscellaneous Articles	2.8 (1)	3.2 (4)	1.7 (1)	0.5 (7)	1.2 (17)
156	Adhesive Bonding and Miscellaneous Chemical Manufacture	0.2 (18)	4.3 (25)	(15)	0.1 (21)	2.4
430	Radiation Imagery Chemistry: Process, Composition, or Product Thereof	1.0	2.4	0.9	0.7 (14)	1.0
356	Optics: Measuring and Testing	0.4	3.1 (12)	2.9	0.1	1.1
250	Radiant Energy	0.0 (13)	3.9 (2)	0.7 (10)	3.1 (6)	2.2
8	Bleaching and Dyeing; Fluid Treatment and Chemical Modification of Textiles and Fibers		2.5		1.0	
437	Semiconductor Device Manufacturing: Process		2.6	0.9	0.9 (16)	0.4
365	Static Information Storage and Retrieval		1.7	0.1	0.2	0.5
347	Incremental Printing of Symbolic Information			0.8	1.1 (20)	0.9
404	Road Structure, Process, or Apparatus		1.8			1.8
525	Synthetic Resins or Natural Rubbers—Part of the Cl520 Series	1.0	1.3	0.2 (6)		0.6
216	Etching a Substrate: Processes	2.0	1.9	0.1		
358	Facsimile or Television Recording		2.2	0.5	0.4	
427	Coating Processes	0.9 (24)	1.3	0.1 (16)	0.3	
257	Active Solid-State Devices (e.g., Transistors, Solid-State Diodes)	0.1	(23)	0.2	1.0 (18)	
355	Photocopying	0.6	1.5		1.6	0.6
528	Synthetic Resins or Natural Rubbers—Part of the Cl520 Series		1.8	(13)		

7. Conclusion

We have explored the use of patent citation information, analyzed in the context of R&D networks, to measure potential R&D spillovers. A significant quantity of empirical and survey evidence exists suggesting that patent citations are an indicator of knowledge spillovers, albeit one with considerable noise. This study's key contribution is the identification of R&D networks as the main spillover mechanism. We have developed a systems methodology using fuzzy logic methods to construct mappings of R&D-based networks. The mappings are based on the assumption that patent citations between any two R&D organizations occurring over a period of time identify communication and interaction. We illustrated the methodology with two applications to two technologies that have been of interest to the ATP (micro-electromechanical systems and optical recording).

Our fuzzy system methodology offers a new and potentially useful way to analyze and interpret R&D spillovers. The spillover literature conceptualizes spillovers as pools, thereby abstracting from specific diffusion pathways and specific spillover mechanisms. We introduce the idea of R&D spillover networks. We argue that our methodology more closely approximates a system that exists in the real world where the quality of network connections matters a good deal. A comparison of spillover ranks based on the existing method and our systems method shows that, while the ranks are correlated, there are significant differences in ranks that can be interpreted as stemming from variation in the strength of the networks.

The new fuzzy logic methodology makes identification of the system possible. Previously, researchers studying spillovers were constrained by methodology and were unable to discover and analyze patterns in the complex patent data set. If the spillover pool metaphor is the right one, then the statistical methods and interpretations found in the literature make sense. If there is an underlying spillover network structure embedded in the patent data, however, then the existing methods fall short. The new methodology helps frame new research on spillovers, provides a framework for developing ATP strategies to maximize spillovers, and suggests one approach for evaluating ATP projects.

Even though it is possible to provide some interesting, new evidence on spillover patterns and R&D networks, currently the methodology does not permit us to draw statistical inferences and formally test hypotheses about the network structure. For example, we would like to know if differences in the measured system influence between two organizations are statistically significant. Statistics assume a certain structure satisfying the condition of measurability. Since fuzzy measures do not necessarily satisfy classical measure properties, fuzzy models cannot be readily used for standard statistical hypothesis testing. Without new research, it remains unclear whether it would be possible to develop a fuzzy measure theoretic basis for our methodology, which could then be used to develop the associated statistical tests.

USE OF PATENT CITATIONS AND SYSTEMS METHOD BY THE ATP

The ATP would like to maximize spillovers produced by the research it funds and measure the extent to which its awardees have generated spillovers. Both of these tasks are difficult because spillovers are hard to see and harder to measure quantitatively. We believe that the analysis here indicates that more systematic analysis of patent citations would potentially be useful to the ATP.

Experience with the methodology is still limited. Because we are on new ground, the findings are essentially exploratory. Therefore, the suggestions are very much in the form of conjectures or hypotheses rather than demonstrated results.

Ex-post evaluation

The most straightforward possible application would be in ex-post evaluation of the impact of ATP funding. Such evaluations can, with relative ease, measure direct outputs of the ATP-funded project, such as patents applied for by awardees, products introduced, sales of those products, and so forth. But such direct measures do not address measuring and generating spillovers and, more broadly, enabling the development of commercially valuable new technology.

The patent citations literature suggests that a logical next step (beyond counting patents, new products, and revenues) would be to examine the citations to ATP grantees' patents. For example, it would be useful to know whether the rate of citations made and citations received increase in association with the ATP grant. If so, the citations would indicate communication between the grantees and its surrounding R&D network.

The systems analysis methods described in this paper offer the potential for detecting indirect effects more rigorously than previous methods. Moreover, because the system influence measures and total knowledge flow are time-specific, it might be possible to undertake before-and-after analysis of the R&D network in the technology area where an ATP project lies. If ATP funding helped to generate spillovers, then not only should the system influence of the funded organization(s) rise (i.e., the funded organizations become more influential players within the network), but the overall intensity of knowledge flow through the system should also have risen. In principle, then, such an analysis could be used to look for the ultimate effects that ATP is seeking to stimulate; that is, enabling of the overall R&D network. Examination of the awardees' performance alone would probably not fully capture such effects. Examination of how the *system* or network has changed could, in principle, however, provide useful evidence of the *system* impacts associated with an ATP-funded project.[1]

In addition, the systems analysis approach could be applied to different settings in which ATP funding was provided (and perhaps also settings in which it was not provided). If so, then we could eventually undertake comparisons of systems that were substantially strengthened with systems that were not. These comparisons might include some analysis of where ATP grants strengthened network influence in some industries, as well as the relative impact of ATP grants of different sizes and forms on certain technology areas.

Project selection

The potential value of these methods for project selection is less clear, but perhaps still worth exploring. As things stand, ATP has no quantitative measures of the spillover potential of applicants.[2] *(The source selection boards are comprised of experts in their fields who readily recognize potential for spillover benefits in an applicant's proposal.)* Objective criteria and judgment regarding the industrial background, the nature of the applicants, and the nature of the technology is used to assess the relative likelihood of R&D spillovers. But it might be useful to supplement the use of such qualitative tools with quantitative indicators, even if those indicators are noisy and imperfect.

A starting point might be to analyze the citation intensity of the project applicant's patents in the relevant technological area. Spillovers are more likely to be generated by applicants who are connected to the relevant R&D network, and patent citations made and received are indicative of such connections. With greater investment of time and effort, software and a database might be developed that would provide a capability for routine use of the fuzzy-logic methodology to measure the system influence of applicants within the relative technological area of their proposal. Having a quantitative indicator of each applicant's position within the relevant R&D network would help evaluate the likelihood of spillovers resulting from a proposed project and be used as one possible measure of the applicant's strength in this area. The software and database might provide a new tool for learning more about the fundamental characteristics of enabling R&D networks.

---

1. Of course, the question of *causation*—whether the system changes that have occurred would or would not have occurred without the ATP funding—is not resolved by this analysis. That is a problem that is endemic to the evaluation setting.

2. The project selection process occurs in a compressed timeframe, thus limiting the ability and opportunity of staff to conduct such research in an effective and timely manner.

In addition, the methodology offers the possibility of a broader set of uses for joint venture proposals. By definition, multiple members of a proposed joint venture occupy a wider spectrum of positions within an R&D network. Perhaps the proposed joint venture members represent the most influential organizations, technologies, and geographic regions in the relevant R&D network represented. This methodology might expose any gaps within the joint venture team, and ATP could discuss any noticeable gaps with the applicants, which could lead the joint venture to seek additional members or subcontractors, or to a better explanation to the ATP why their proposed membership is already sufficiently broad. It could also potentially lead to concluding that the proposal is not likely to be enabling.

Another potentially valuable application to joint ventures would be the modeling of the joint venture's formation itself. One way to think of joint venture formation is that it increases communication among the joint venture members. It may be possible to simulate such effects by examining the impact of the joint venture on total system knowledge flows (the increased communication between the specific nodal pairs that constitute the joint venture). Again, it may be possible to identify weak links within the relevant R&D network, which, if strengthened, would generate particularly beneficial impacts on the network as a whole.

Broader analyses

In addition to evaluation of specific project proposals, it is possible that the methodology could be useful in providing broader background to the ATP's planning and policy process. For example, the methodology could be used to construct an overall mapping of broad technologies into industries. Such a mapping could be done periodically, providing a basis to gauge the broad-based economic impacts of projects in various technological categories. Such a structure would also provide an additional framework for thinking about cumulative or systemic impacts of ATP's funding of a number of related projects. In the 1990s, ATP sought to increase the impact of its funding by funding a number of technologically related projects in its focused programs. At present, there are not many methods for assessing synergistic or cumulative impacts of multiple projects in a systematic way.[3] If ATP were to evaluate a number of its projects within a given technological area, the overall broad-based economic impacts of technological advances in that area could be examined in the same way as the MEMS and optical recording were done in this study. Individual projects could then be evaluated within the context of the estimated overall economic impact of the technological area.

POSSIBILITIES FOR FURTHER RESEARCH

This report suggests several possible avenues for future research, each of which would involve interaction among the fuzzy network analysis, case studies (e.g., drawing from ATP's database and cases), and surveys focused on enabling networks.

3. The ATP's Economic Assessment Office has conducted benefit-cost studies for its projects in specific technology areas, and has been able to assess some of the related economic impacts in systematic ways.

- Develop the database and software to permit ATP to routinely evaluate the spillover potential of project participants (focused on several dimensions, such as technology, industry, and geography), retrospectively evaluate spillovers generated by funded projects, and explore the full patent database for particularly enabling networks. Development of this generic capability would permit ATP to monitor, on an ongoing basis, the interaction of awardees with the relevant R&D networks.

- Use the methodology, coupled with case study and survey methods, to model the formation of joint ventures. For example, how does the formation and operation of joint ventures affect communication (and, therefore, spillovers) among members and the full network? Can network analysis be used to design more effective joint ventures?

- Use the methodology to develop analogous/parallel networks based on citations to the scientific literature. Under what conditions do the science and technology networks intersect and interact? How does each influence the other? Do certain networks exhibit characteristics that are associated with the network drawing extensively from science? How can ATP use this knowledge to leverage early-stage federal R&D support of basic science? Such an analysis would require integrating our patent networks with networks derived from science citations data.

- Advance the fuzzy logic methodology for analysis of R&D networks. Advancing the methodology would require combining development of the underlying mathematics (e.g., what is the best way to model two-way flows; that is, learning from and using knowledge spillovers) with case studies and surveys to add institutional knowledge to the picture, especially for cases involving emerging, enabling technologies.

- Develop the measurement theory basis for our fuzzy logic measures. Given the nature of fuzzy measures, currently the methodology does not permit us to draw statistical inferences and formally test hypotheses about the network structure. New research is necessary to determine whether it is possible to develop a fuzzy measure theoretic basis for our methodology, which could then be used to develop the associated statistical tests.

Appendix A: A New Fuzzy and Possibilistic Partitioning Algorithm: An Application to Modeling R&D Networks*

1 A MODEL OF THE U.S. INNOVATION SYSTEM

In this section we summarize the steps taken to model the U.S. innovation system. This section develops the system requirements that satisfy the objectives of an economic analysis of the patent data set.

1.1 System Specifications of the Innovation System

The U.S. innovation system was described as a network of R&D labs where the strengths of interaction between labs are indicated by citations made by one lab to the other. These citations can be further specified by technology classes of cited and citing patents, and their respective grant years. The model specifications are as listed below.

<u>System Specifications</u>

S1. The system is defined on a network of nodes, where each node represents the inventions of an R&D lab, in a specific technology and year.

S2. There are knowledge flows from cited node to citing node, as evidenced by citations.

S3. Measures of connectivity between nodes reflect system-wide diffusion effects of the interaction, and system-wide diffusion effects of the nodes. A direct citation is interpreted as a one-step diffusion.

S4. Measures of connectivity of nodes satisfy the assumption that the influential (strongly connected) is what is strongly interconnected with the influential, which is what is strongly interconnected with the influential, and so forth.

S5. The system of networks can be partitioned into overlapping subnetworks of strongly interconnected nodes.

* Additional details are contained in Sinha (2001).

S1 specifies a node as representing the inventions of an R&D lab, in a specific technology and year because the R&D lab is where the invention occurs and because the issues of interest are the organizational, technological, and geographical spread of knowledge over time.

S2 states a commonly used but widely discussed assumption that citations are evidence of knowledge flows. This issue had been discussed in more detail in the text of the report.

The system significance of knowledge flow from one node to another can be properly gauged only if measured relative to knowledge flows between all pairs of nodes. Further, it is reasonable to assume that an R&D lab has greater potential to influence another lab if the R&D lab is innovative. An indicator of an R&D lab's innovativeness is the organizational, technological, and geographical spread of citations made and received. Another factor is the innovativeness of the citing and cited R&D labs. S3 and S4 model the hypotheses discussed.

Patents draw on existing knowledge classified by technology, organization, region, and time and embody usable knowledge distinguishable from prior art. This is seen from the distribution of citations on a single patent across the three dimensions. In other words, nodes influence the innovative activity of other nodes in varying degrees of significance.

So intuitively, and drawing on domain knowledge, a multivalent, logic-based approach is in this case more likely to provide *a more true* model of the reality than a bivalent approach. One objective of analysis is to identify broad categories of technology, R&D labs, and regions, and also to evaluate the significance of specific technologies, R&D labs, and regions within and across the categories. By requiring that the model enable a partition of the system into overlapping partitions, S5 ensures the use of a multivalent logic in structuring overlapping networks which can then be characterized.

2 A SYSTEMS APPROACH TO CLUSTERING

In this section, we develop a new approach to the problem of clustering. Grouping objects into self-similar groups involves evaluation of individual pairs of proximities in relation to all pairs of proximities in the data set. It follows that proximities amongst data points behave like a system, where change in proximities in a sector diffuses, changing orders in the neighborhoods. So, from the perspective of clustering, all data can be described as a system irrespective of the objective of the analysis. We propose a new definition of a partition matrix that attempts to mathematically approximate this thinking. The definition, although motivated by the specific problem of developing a model of the innovation system based on the U.S. patent data, is valid for all problems of clustering.

We also develop an algorithm for generating the partition matrix and a graph-theoretic procedure to construct the clusters. The definition of the partition matrix and the algorithm assume that the data has a network structure. This is a nonrestrictive assumption. Any given weight matrix, including a distance matrix on n data points, can be viewed as a network defined on n objects with weights between the data points as the flows on the arcs, suitably scaled depending on whether the problem is that of minimization or maximization. We establish results on validity and convergence of the approach. The proofs show that the measures generated by the proposed algorithm satisfy the system requirements of the U.S. innovation system laid out in the previous section. The generality of the new definition and the algorithm is evident and supported by our computational experience with several benchmark data sets.

At the outset, we rephrase the system requirements in more general terms.

S1´ The system is defined on a network of nodes, where each node represents the objects in the data.

S2´ There exists a relationship on (dis)similarities between nodes as evidenced by the weight on the arcs, which correspond to the given relational matrix.

S3´ Measures of connectivity between nodes are determined by the significance of the relation relative to the system, and the system-wide connectivity of the modes, with a direct relation being interpreted as a direct connection.

S4´ Measures of connectivity of nodes satisfy the assumption that the highly modal is what is strongly interconnected with the highly modal, which is what is strongly interconnected with the highly modal, ad infinitum.

S5´ The system of networks can be partitioned into overlapping subnetworks of strongly interconnected nodes.

2.1 Notation and Definitions

$G = (V = \{i\}, E = \{(ij)\}, i, j = 1, 2, \ldots, n$ a network defined on the nodes $\{i\}$ corresponding to the objects of the system. $(ij) \in E$ if $i \in V$ and $j \in V$ and there is a flow from node i to node j

$\alpha \in (0, 1)$: a predefined intersection parameter

$\beta \in (0, 1)$: a predefined algebraic sum parameter.

Traversal: an arc (lm) is said to be traversed from a node i in I steps if there exist a path $P_{i \to \ldots l \to m}$ of length $l \leq I$. A node l is said to be traversed from node i in I steps if the arcs incident to node l are traversed in I steps. The direction of traversal of the arcs depends on the networks considered as defined below.

$N^{I,1}(i)$: *I-step Outflow Network* of a node i is the set of all nodes and arcs traversed in I steps from i, the first step being restricted to be in the direction of the arcs.

$N^{I,2}(i)$: *I-step Inflow Network* of a node i is the set of all nodes and arcs traversed in I steps of diffusion from i, the first step being restricted to be the direction reverse that of the arcs.

$N^I(i) = N^{I,1}(i) \cup N^{I,2}(i)$: *I-step Neighborhood* of a node i is the union of the I-step outflow network and I-step inflow network of node i, where the union of networks is defined as the set union of the set of arcs in the networks, incident to set union of sets of nodes in the networks.

An I-step Neighborhood Parameter

Any parameter constructed on I-step neighborhoods is said to contain *full information on I-step neighborhoods* if and only if a simple ordering of the parameter values orders the I-step neighborhoods on the property being measured.

Any parameter constructed on I-step neighborhoods is defined to be a *I-step Neighborhood Parameter* if the parameter contains full information on the I-step neighborhoods and partial or no additional information on higher order, i.e.; on $I + k$, $k = 1, 2, \ldots, n$ step neighborhoods.

I-step Neighborhood Parameters

$$p_i^{I,1}(j):$$

measure of connectivity on the arc (ij) evaluated on I-step neighborhood of node j.

$$p_j^{I,2}(i):$$

measure of connectivity of the arc (ij) evaluated on I-step neighborhood of node i.

$$f_i^{I,1}(j):$$

fraction of $\sum_j p_i^{I,1}(j)$, the I-step out-connectivity of i to j .

$$f_j^{I,2}(i):$$

fraction of $\sum_i p_j^{I,2}(i)$, I-step in-connectivity of j from i.

$$t_i^{I,1}(j) = \frac{\sum_{(l,m)\in E} I_{f_i^{I,1}(j)}(f_l^{I,1}(m)) * f_i^{I,1}(j)}{\sum_{v=0}^{v=1} \sum_{(l,m)\in E} I_v(f_l^{I,1}(m)) * v}:$$

typicality of $f_i^{I,1}(j)$, the fraction of connectivity.

$$t_j^{I,2}(i) = \frac{\sum_{(l,m)\in E} I_{f_j^{I,2}(i)}(f_l^{I,2}(m)) * f_j^{I,2}(i)}{\sum_{v=0}^{v=1}\sum_{(l,m)\in E} I_v(f_l^{I,2}(m)) * v}:$$

typicality of $f_j^{I,2}(i)$, the fraction of connectivity.

$$O^I(i):$$

set centered on *i* on the property of I-step out-connectivity of *i*.

$$I(j):$$

set centered on *j* on the property of in-connectivity of *j*.

$$\mu_i^{I,1}(j) = \sum_{t_l^{I,1}(m)=0}^{t_i^{I,1}(j)} \sum_{(l,m)\in E} t_l^{I,1}(m):$$

the membership value of the arc (*ij*) in $O^I(i)$.

$$\mu_j^{I,2}(i) = \sum_{t_l^{I,2}(m)=0}^{t_j^{I,2}(i)} \sum_{(l,m)\in E} t_l^{I,2}(m):$$

the membership value of the arc (*ij*) in $I(j)$.

$$\mu_i^{I,3}(j) = \alpha * \min(\mu_i^{I,1}(j), \mu I,2_j(i)) + (1-\alpha) * avg(\mu_i^{I,1}(j), \mu_j^{I,2}(i):$$

the membership value of the arc (*ij*) in the intersection of $O^I(i)$ and $I^I(j)$.

$$\omega^{I,1}(i) = \sum_j \mu_i^{I,1}(j):$$

I-step outdegree of node *i*.

$$\omega^{I,2}(j) = \sum_i \mu_j^{I,2}(i):$$

I-step indegree of node *i*.

$$V_i^{I,1} = \sum_j \mu_i^{I,3}(j)\omega^{I,2}(j):$$

total connectivity of node *i* evaluated on I-step outflow networks. Also referred to as I-step outmodality.

$$V_i^{I,2} = \sum_i \mu_j^{I,3}(i)\omega^{I,1}(j):$$

total connectivity of node *i* evaluated on I-step inflow networks. Also referred to as I-step inmodality.

$$V_i^I = \beta * V_i^{I,1} + (1-\beta) * V_i^{I,2}:$$

total connectivity of node i evaluated on I-step neighborhoods. Also referred to as I-step modality.

2.2 Partition Matrix—A New Definition and An Associated Algorithm

At the outset we introduce some more definitions:

Defn. Weak Equivalence Neighborhoods $N(i)$ and $N(k)$ are said to be *weakly equivalent* if the sum of the modality parameters of the nodes in the two neighborhoods are *approximately equal*. $N(i)$ and $N(k)$ are said to be β-weakly equivalent if the modality sums on the two neighborhoods differ by no more than β.

Defn. Strong Equivalence Neighborhoods $N(i)$ and $N(k)$ are said to be *strongly equivalent* if they are weakly equivalent and if $\{v \in N(i)\} = \{v \in N(k)\}$. $N(i)$ and $N(k)$ are said to be *β*–strongly equivalent if they are β-weakly equivalent and if $\{v \in N(i)\} = \{v \in N(k)\}$.

Defn. A Fuzzy-Possibilistic Partition Matrix Given $X = \{x_1, x_2, \ldots, x_n\}$, a set of n objects, and an arbitrary $n \times n$, asymmetrical relational matrix $R = [r_{ij}]$, the set of $n \times n$ matrices,

$$\left\{\begin{matrix} U^1 = [\mu_i^1(j)] \\ U^2 = [\mu_i^2(j)] \\ U^3 = [\mu_i^3(j)] \end{matrix}\right\} i = 1,2,\ldots,n; j = 1,2,\ldots,n$$

describe the fuzzy-possibilistic network partition of X, where

1. $\mu_i^k(j) \in [0,1] \forall i, j, k = 1,2,3$

   $\mu_i^1(j)$: the membership value of the arc (ij) in $O(i)$

   $\mu_j^2(i)$: the membership value of the arc (ij) in the $I(j)$

   $\mu_i^3(j)$: the membership value of the arc (ij) in the intersection of $O(i)$ and $I(j)$

2. $\mu_i^3(j) > r \Leftrightarrow (ij)$ is r-significant to i and j in the system.

3. $$\left[\sum_j \mu_i^3(j) \sum_l \mu_j^2(l) \sum_j \mu_j^3(i) \sum_l \mu_j^1(l)\right] - \left[\sum_j \mu_k^3(j) \sum_l \mu_j^2(l) + \sum_j \mu_j^3(k) \sum_l \mu_j^1(l)\right] \leq \beta.$$

   $\Leftrightarrow N(i)$ and $N(k)$ are *b-weakly equivalent*.

As before S1 and S2 are satisfied by construction and assumption respectively. Property (3) stipulates that if the membership values of the arcs in the I-step neighborhoods of two nodes are the same then the nodes have weakly equivalent neighborhoods, implying that measures of interaction between nodes, and that of the nodes, reflect system-wide diffusion effects. As such, the partition matrix as defined satisfies S3 and S4.

Appendix B:
The Top 25 Technologies in Six Industries

TABLE B.1

System Influence by Source and Destination in the Aerospace R&D Network

Class Number	Patent Class Name	Source	Destination
364	Electrical Computers and Data Processing Systems	1,986.18	2,029.79
250	Radiant Energy	1,510.77	1,509.65
439	Electrical Connectors	1,413.83	1,329.40
428	Stock Material or Miscellaneous Articles	1,406.01	1,322.59
324	Electricity: Measuring and Testing	1,349.58	1,449.16
361	Electricity: Electrical Systems and Devices	1,321.60	1,246.72
29	Metal Working	1,311.32	1,342.37
385	Optical Waveguides	1,309.06	1,285.97
73	Measuring and Testing	1,267.97	1,304.24
359	Optics: Systems (including Communication) and Elements	1,119.24	1,097.68
310	Electrical Generator or Motor Structure	1,044.80	1,112.91
356	Optics: Measuring and Testing	938.10	828.17
327	Miscellaneous Active Electrical Nonlinear Devices, Circuits, and Systems	894.43	885.59
60	Power Plants	893.97	894.84
74	Machine Element or Mechanism	887.60	904.72
318	Electricity: Motive Power Systems	840.93	901.25
340	Communications: Electrical	828.52	725.94
395	Information Processing System Organization	798.87	946.76
342	Communications: Directive Radio Wave Systems and Devices (e.g., Radar, Radio Navigation)	792.38	825.96
375	Pulse or Digital Communications	779.70	803.05
363	Electric Power Conversion System	761.69	730.05
219	Electric Heating	737.44	729.00
257	Active Solid-State Devices (e.g., Transistors, Solid-State Diodes)	734.32	743.47
333	Wave Transmission Lines and Networks	710.74	631.87
156	Adhesive Bonding and Miscellaneous Chemical Manufacture	709.11	737.48

TABLE B.2

System Influence by Source and Destination in the Information Technology R&D Network

Class Number	Patent Class Name	Source	Destination
395	Information Processing System Organization	1,586.95	1,844.16
364	Electrical Computers and Data Processing Systems	1,351.31	1,376.68
361	Electricity: Electrical Systems and Devices	892.40	852.58
324	Electricity: Measuring and Testing	860.51	907.52
371	Error Detection/Correction and Fault Detection/Recovery	739.36	730.13
250	Radiant Energy	708.61	727.75
428	Stock Material or Miscellaneous Articles	674.95	554.22
370	Multiplex Communications	667.12	700.07
327	Miscellaneous Active Electrical Nonlinear Devices, Circuits, and Systems	653.92	650.87
359	Optics: Systems (including Communications) and Elements	642.79	581.73
375	Pulse or Digital Communications	641.22	592.28
360	Dynamic Magnetic Information Storage or Retrieval	620.77	515.57
326	Electronic Digital Logic Circuitry	588.88	547.80
430	Radiation Imagery Chemistry: Process, Composition, or Product Thereof	566.71	546.93
345	Selective Visual Display Systems	540.31	522.99
437	Semiconductor Device Manufacturing: Process	533.61	514.86
439	Electrical Connectors	533.23	501.35
257	Active Solid-State Devices (e.g., Transistors, Solid-State Diodes)	517.32	582.45
380	Cryptography	514.51	535.19
347	Incremental Printing of Symbolic Information	504.52	485.20
156	Adhesive Bonding and Miscellaneous Chemical Manufacture	475.39	435.55
340	Communications: Electrical	471.79	403.74
235	Registers	464.36	446.12
348	Television	457.61	415.89
400	Typewriting Machines	445.21	440.34
341	Coded Data Generation or Conversion	426.65	431.28

TABLE B.3

System Influence by Source and Destination in the Auto R&D Network

Class Number	Patent Class Name	Source	Destination
428	Stock Material or Miscellaneous Articles	1,924.79	1,745.27
364	Electrical Computers and Data Processing Systems	1,481.10	1,581.37
123	Internal Combustion Engines	1,475.58	1,490.37
74	Machine Element or Mechanism	1,475.04	1,519.00
73	Measuring and Testing	1,212.05	1,292.44
361	Electricity: Electrical Systems and Devices	1,188.98	1,062.54
29	Metal Working	1,120.25	1,189.90
60	Power Plants	1,073.36	1,057.41
439	Electrical Connectors	1,029.13	991.11
310	Electrical Generator or Motor Structure	1,017.47	1,066.06
192	Clutches and Power-Stop Control	996.58	970.70
359	Optics: Systems (including Communication) and Elements	958.74	947.27
250	Radiant Energy	958.04	1,012.16
318	Electricity: Motive Power Systems	935.89	995.40
340	Communications: Electrical	924.63	857.48
324	Electricity: Measuring and Testing	898.79	951.23
264	Plastic and Nonmetallic Article Shaping or Treating: Processes	895.94	918.31
156	Adhesive Bonding and Miscellaneous Chemical Manufacture	834.32	878.75
210	Liquid Purification or Separation	825.75	860.03
180	Motor Vehicles	794.88	699.18
137	Fluid Handling	785.30	729.30
280	Land Vehicles	768.29	759.44
385	Optical Waveguides	762.29	750.78
427	Coating Processes	751.91	814.03
188	Brakes	709.20	736.74
252	Compositions	700.93	734.35

TABLE B.4

System Influence by Source and Destination in the Advanced Materials R&D Network

Class Number	Patent Class Name	Source	Destination
428	Stock Materials or Miscellaneous Articles	2,728.03	2,595.63
210	Liquid Purification or Separation	1,730.55	1,770.28
252	Compositions	1,631.23	1,710.80
439	Electrical Connectors	1,574.50	1,506.97
364	Electrical Computers and Data Processing Systems	1,557.07	1,619.67
525	Synthetic Resins or Natural Rubbers—Part of the Class 520 Series	1,471.39	1,319.59
73	Measuring and Testing	1,415.93	1,422.15
524	Synthetic Resins or Natural Rubbers—Part of the Class 520 Series	1,415.71	1,387.30
264	Plastic and Nonmetallic Article Shaping or Treating: Processes	1,387.61	1,380.16
250	Radiant Energy	1,341.32	1,363.10
423	Chemistry of Inorganic Compounds	1,289.52	1,239.35
385	Optical Waveguides	1,286.07	1,261.73
528	Synthetic Resins or Natural Rubbers—Part of the Class 520 Series	1,284.77	1,172.02
502	Catalyst, Solid Sorbent, or Support Thereof: Product or Process of Making	1,116.04	990.85
156	Adhesive Bonding and Miscellaneous Chemical Manufacture	1,098.47	1,142.23
427	Coating Processes	1,027.39	1,080.86
324	Electricity: Measuring and Testing	998.70	1,049.80
29	Metal Working	973.10	967.49
359	Optics: Systems (including Communications) and Elements	948.26	888.89
361	Electricity: Electrical Systems and Devices	936.56	881.53
166	Wells	885.83	931.75
568	Organic Compounds—Part of the Class 532–570 Series	775.44	715.17
514	Drug, Bio-Affecting and Body Treating Compositions	754.61	707.84
521	Synthetic Resins or Natural Rubbers—Part of the Class 520 Series	747.43	659.59
219	Electric Heating	745.03	734.25

TABLE B.5

System Influence by Source and Destination in the Biomedical Devices R&D Network

Class Number	Patent Class Name	Source	Destination
606	Surgery	240.15	249.46
604	Surgery	239.84	253.75
623	Prosthesis (i.e., Artificial Body Members)	175.09	156.33
607	Surgery: Light, Thermal and Electrical Application	156.85	167.83
600	Surgery	83.11	84.80
280	Land Vehicles	74.67	74.09
424	Drug, Bio-Affecting and Body Treating Compositions	63.16	65.07
514	Drug, Bio-Affecting and Body Treating Compositions	58.67	58.73
422	Chemical Apparatus and Process Disinfecting	58.42	48.32
227	Elongated-Member-Driving Apparatus	56.39	48.20
222	Dispensing	52.02	50.21
435	Chemistry: Molecular Biology and Microbiology	50.74	45.46
429	Chemistry: Electrical Current Producing Apparatus	48.92	40.43
264	Plastic and Nonmetallic Article Shaping	43.67	42.21
206	Special Receptacle or Package	38.78	48.15
439	Electrical Connectors	37.80	27.75
428	Stock Materials or Miscellaneous Articles	36.92	26.31
372	Coherent Light Generators	33.70	36.46
73	Measuring and Testing	33.04	29.81
601	Surgery: Kinesitherapy	28.23	26.93
378	X-ray or Gamma Ray Systems or Devices	26.89	32.86
180	Motor Vehicles	23.86	21.73
29	Metal Working	23.29	22.35
72	Metal Deforming	22.89	24.52
252	Compositions	21.46	22.26
427	Coating Processes	20.89	26.10
310	Electrical Generator or Motor Structure	19.68	9.33
219	Electrical Heating	19.55	21.25
351	Optics: Eye Examining, Vision Testing and Correcting	18.98	17.50
549	Organic Compounds—Part of the Class 532–570 Series	18.95	14.08
525	Synthetic Resins or Natural Rubbers—Part of the Class 520	17.82	17.38

TABLE B.6

System Influence by Source and Destination in the MEMS R&D Network

Class Number	Patent Class Name	Source	Destination
359	Optics: Systems (including Communications) and Elements	459.68	176.34
385	Optical Waveguides	353.79	208.52
73	Measuring and Testing	330.39	287.77
310	Electrical Generator or Motor Structure	312.26	204.82
437	Semiconductor Device Manufacturing: Process	250.74	106.89
257	Active Solid-State Devices (e.g., Transistors, Solid-State Diodes)	188.62	138.65
347	Incremental Printing of Symbolic Information	124.30	49.21
216	Etching a Substrate: Processes	117.86	70.90
156	Adhesive Bonding and Miscellaneous Chemical Manufacture	112.25	52.82
369	Dynamic Information Storage or Retrieval	106.24	51.95
428	Stock Material or Miscellaneous Articles	97.40	34.38
348	Television	90.02	55.67
60	Power Plants	86.36	43.55
345	Selective Visual Display Systems	84.20	28.28
136	Batteries: Thermoelectric and Photoelectric	70.55	40.08
365	Static Information Storage and Retrieval	63.48	47.30
361	Electricity: Electrical Systems and Devices	56.62	54.27
367	Communications, Electrical: Acoustic Wave Systems and Devices	54.58	31.67
356	Optics: Measuring and Testing	47.67	56.02
353	Optics: Image Projectors	42.31	24.59
251	Valves and Valve Actuation	26.92	26.22
29	Metal Working	23.48	31.54
324	Electricity: Measuring and Testing	13.88	3.67
338	Electrical Resistors	11.48	7.85