Network Dynamics and Field Evolution: The Growth of Interorganizational Collaboration in the Life Sciences[1]

Walter W. Powell
Stanford University and Santa Fe Institute

Douglas R. White
University of California, Irvine, and Santa Fe Institute

Kenneth W. Koput
University of Arizona

Jason Owen-Smith
University of Michigan

A recursive analysis of network and institutional evolution is offered to account for the decentralized structure of the commercial field of the life sciences. Four alternative logics of attachment—accumulative advantage, homophily, follow-the-trend, and multiconnectivity—are tested to explain the structure and dynamics of interorganizational collaboration in biotechnology. Using multiple novel methods, the authors demonstrate how different rules for affiliation shape network evolution. Commercialization strategies pursued by early corporate entrants are supplanted by universities, research institutes, venture capital, and small firms. As organizations increase their collaborative activities and diversify their ties to others, cohesive subnetworks form, characterized by multiple, independent pathways. These structural components, in turn, condition the choices and opportunities available to members of a field, thereby reinforcing an attachment logic based on differential connections to diverse partners.

INTRODUCTION

The images of *field* and *network* are common in both contemporary physical and social science. In the physical sciences, fields are organized by information in the form of geometric patterns. The study of the geometry

of fields has attracted considerable interest in the statistical mechanics of complex networks. Research by physicists interested in networks has ranged widely from the cellular level, a network of chemicals connected by pathways of chemical reactions, to scientific collaboration networks, linked by coauthorships and cocitations, to the World Wide Web, an immense virtual network of websites connected by hyperlinks (Albert, Jeong, and Barabási 1999; Jeong et al. 2000; Newman 2001; Watts and Strogatz 1998). Albert and Barabási (2002) and Newman (2003) provide excellent overviews of this burgeoning literature on the network topology of different fields, highlighting key organizing principles that guide interactions among the component parts.

In the social sciences, however, analyses of fields and networks have been oddly disconnected. We say oddly because the study of the macro-dynamics of networks should be central to the understanding of how fields evolve. This lack of connection is rooted in several features of contemporary research. An abundance of research in network analysis examines why ties form between two actors and the consequences of particular network positions. Salancik (1995) observed, however, that most network research has taken an individual-level perspective and missed out on the opportunity to illuminate the structure of collective action. McPherson et al. (2001) note that there are few studies that employ longitudinal data to analyze networks. Burt (2000) has voiced a similar concern that most studies of network structure are cross-sectional. In the most comprehensive text on network methods, there is only a paragraph on network dynamics in a section on future directions (Wasserman and Faust 1994). Thus while some progress has been made analyzing the dynamics of dyads (e.g., Lincoln et al. 1996; Gulati and Gargiulo 1998;

Stuart 1998), little attention has been given to the evolution of entire networks.

There are a number of excellent studies of the structuring of specific organizational fields (DiMaggio 1991; Thornton 1995; Dezalay and Garth 1996; Ferguson 1998; Scott et al. 2000; Hoffman 2001; Morrill and Owen-Smith 2002). An organizational field is a community of organizations that engage in common activities and are subject to similar reputational and regulatory pressures (DiMaggio and Powell 1983). Such fields have been defined as "a network, or a configuration, of relations between positions" (Bourdieu 1992) and as "centers of debates in which competing interests negotiate over the interpretation of key issues" (Hoffman 1999, p. 351). Fields emerge when social, technological, or economic changes exert pressure on existing relations and reconfigure models of action and social structures. But despite the relational focus on how different actors and organizations constitute a recognized arena of social and economic activity, studies of fields have not analyzed the interactions of multiple, overlapping networks or the regulated reproduction of network ties through time. This linkage between network dynamics and the evolving structure of fields needs to be made in order to make progress in explaining how the behavior of actors or organizations of one kind influence the actions of organizations of another kind.

The goal of this article is to account for the development and elaboration of the commercial field of biotechnology by showing how the formation, dissolution, and rewiring of network ties over a 12-year period, from 1988 to 1999, have shaped the opportunity structure of the field.[2] By mapping changing network configurations, we discern how logics of attachment shift over time, and we chart multiple influences on the varied participants in the field. Our effort is part of a more general move in the social sciences to analyze momentum, sequences, turning points, and path dependencies (see Abbott [2001] for an overview). By linking network topology and field dynamics, we consider social change not as an invariant process affecting all participants equally, but as reverberations felt in different ways depending on an organization's institutional status and location in the overall network as that structure evolves over time. Our aim is to illuminate how patterns of interaction emerge, take root, and transform with ramifications for all of the participants.

We develop arguments concerning how the topology of a network and

[2] We use the term *field* rather than *industry* or *population* intentionally. Biotechnology is not a separate industrial sector with well-defined boundaries. Universities, government labs, nonprofit hospitals, and research institutes are a critical part of the field; while on the commercial side, both established pharmaceutical firms and dedicated biotechnology companies are involved in bringing new medicines to market. Thus, *field* captures the diversity of organizations more aptly than any other term.

1134

the rules of attachment among its constituents guide the choice of partners and shape the trajectory of the field. As organizations enter a field and relationships deepen and expand, significant structural changes occur. To analyze and understand these emergent network structures, we use a triangulation of methods. We first analyze the expansion of the network to see if the process is random or uniform. Prior research suggests that as new organizations join the network, there will be an attachment bias that increases the probability of linking to an organization that already has ties (de Solla Price 1965, 1980; Barabási 2002a). We go further and assess whether other attachment processes are operative as well. We map the development of the field by drawing network configurations to create a framework with which to view network dynamics. We use the software package Pajek (de Nooy, Mrvar, and Batagelj 2005) for the representation of network dynamics. Pajek allows us to analyze the nearly 2,800 nodes in our sample and to identify cohesive subsets such as multiconnected components (White and Harary 2001, pp. 12–14). We present a small selection of these network visualizations to highlight both the evolving topology of the field and the processes by which new ties and organizations are added. (The full "movie," with year-to-year representations of the topology and new additions to the network, may be viewed in the online version of *AJS*). We then turn to a statistical examination of network formation and dissolution and assess the effects of alternative mechanisms of attachment. Using McFadden's (1973, 1981) discrete choice model, a variant of the conditional logit model, we test to see if the basis of attraction is accumulative advantage, similarity, follow-the-trend, or multiconnectivity.

The Topology of Large-Scale Networks

A variety of researchers in physics and sociology are studying the structure of large-scale networks with the intuition that complex adaptive systems evince organizing principles that are encoded in their topology. Large-scale networks typically have characteristic signatures of local structure, such as clustering, and a global structure, such as average distance between nodes. Local and global characteristics of networks help to define network topologies such as small worlds, which are large networks with both local clustering and relatively short global distance. Watts (1999) showed that adding only a handful of remote links to a large network where the level of local clustering is high (e.g., friends of friends are friends) is sufficient to create a small world network. Watts and Strogatz (1998) helped to revitalize the earlier line of research introduced by Milgram (1967) and developed by White (1970). The wide appeal of the small world idea had been portrayed in the arts, in John Guare's play *Six Degrees of*

Separation, and in the popular Kevin Bacon game, where virtually every Hollywood actor is linked through a few steps. Even a small proportion of randomly distributed ties can knit together diverse clusters of nodes to produce small world phenomena. Researchers have applied the small world concept to a wide range of activities, including scientific collaborations (Newman 2001), corporate board interlocks (Davis, Yoo, and Baker 2003), national ownership networks (Kogut and Walker 2001), invesment bank syndicates (Baum, Shipilov, and Rowley 2003), and Broadway plays (Uzzi and Spiro 2005).

Watts and Strogatz's (1998) formalization of the small world problem lacked any role for network hubs—nodes with an unusually large number of ties or edges. This limitation also held for early models of random networks, in which an equal probability of any given pair of nodes being connected generated only a mild tendency for nodes to differ in their number of edges.[3] But research on citation networks, however, has shown highly skewed distributions, with most articles having few citations, while a handful have an exceedingly large number. Lotka (1926) and de Solla Price (1965, 1980) showed that in degree distributions of citation networks, the proportion of nodes with degree k varies as a function of $1/k^{\alpha}$, that is, by the inverse power law $P(k) \sim 1/k^{\alpha}$, where alpha is the power coefficient. Barabási and Albert (1999, p. 510) and Barabási (2002*a*, p. 70) popularized the term "scale free" for such networks,[4] and they confirmed that network growth with preferential attachment according to degree

[3] Even when edges form with equal probability for all pairs of nodes in a network, there is considerable inequality in the distribution of the number of edges for different nodes. The tail of the degree distribution of a simple random process of tie-formation will be truncated, however, by exponential decay in the number of nodes at successively higher degree. The degree distribution of nodes in a highway network, e.g., tends to be exponential, more like a random network than the degree distribution of nodes in today's U.S. airlines network, with hubs that jump over many of the nodes connected by spokes. Note that these differences are, in part, a matter of intentional design and not solely a function of geography; there was a time when the airline network was more like a highway network.

[4] The meaning of *scale invariant* or *scale free* for a power law is that the coefficient does not vary from scale to scale as magnitude varies from 1 to 10 to 100, etc. Put differently, for a network with a power-law tail to the degree distribution, there is no characteristic number of edges per node, as in a bell-curve-shaped distribution or exponential decay. Barabási's (2002*b*) organizing motif is that networks with power-law degree distributions have a characteristic scale-free signature of self-organized systems. Analogy is made to the scaling of the frequency of earthquakes in relation to their energetic intensity, e.g., which tends to follow a power law. This would imply that there is no typical scale for earthquakes and suggests that the physical mechanism for large earthquakes is the same as that for the small ones. The theme of "the same mechanism" has not been established for social networks, however.

predicts a scale-free tail of the degree distribution.[5] The well-connected nodes that newcomers attach to become hubs that create short paths between many pairs of nodes in the network.

Preferential attachment to higher degree leads to a dynamic of rich-get-richer. Power-law tails of degree distributions are present in very diverse kinds of networks. In the movie actor network, for example, new actors tend to start their careers in supporting roles accompanying famous actors, and in science new publications cite well-known papers. Attachment bias in network formation bears strong similarity to the more general phenomenon of accumulative advantage (Merton 1973), in which those who experience early success capture the lion's share of subsequent rewards. We use early starter accumulative advantage and preferential attachment as baseline arguments, since they provide potential explanations for growing inequalities in the process of network expansion. But not all early entrants turn out to be winners, and some latecomers attain prominence. As the saying goes, the early bird may catch the worm, but it is the second mouse that eats the cheese. Similarly, Albert and Barabási's (2002) formalization of a "scale-free" class of networks, in which the probability that a new entrant will choose to link to an incumbent node is proportional to the number of links the incumbent already has, is elegantly simple but overgeneralized. Other attachment processes, or a combination of diverse mechanisms, can produce power law degree distributions. Our analyses test for multiple, probabilistic biases in the processes of network growth.

We enter the discussion of network dynamics with data from a field where social, political, economic, and scientific factors loom large in shaping patterns of attachment among the participants. In earlier work on the biotechnology industry, Powell, Koput, and Smith-Doerr (1996) found a liability of disconnectedness, in which older, less-linked organizations were the most likely to fail. Certainly, early entrants have more time than later arrivals to establish connections, but Powell et al. (1996) found that how connections were established and what activities were pursued were critical. Biotechnology firms had to both make news and be in on the news; that is, they needed to generate novel contributions to the evolving science while maintaining the capability to evaluate what other organizations were doing. The pathway to centrality in the industry network was through research and development collaborations. Other routes were ei-

---

[5] Here the actual attachment probability of new nodes with incumbents, $P(k)$, is proportional to k where (k) is the degree of the incumbent. The preferential attachment probability generates a degree distribution in which the frequency of nodes with a given degree d is a function $f(d)$ of $1/d^a$, where a is the power-law coefficient and can be calculated from the slope of the linear regression line on a log-log plot of d and $f(d)$.

ther ineffective or much slower in generating centrality. Moreover, in a highly competitive world in which it is not easy to rest on past accomplishments, firms that do not expand or renew their networks lose their central positions. At the same time, resource-rich participants are more capable of altering their positions by reconfiguring their networks.

Biotechnology is characterized by a high rate of formation and dissolution of linkages. Connections are often forged with a specific goal in mind, such as taking a company public or selling and distributing a new medicine. Once the task is completed, the relationship is ended and successful collaborators depart gracefully. There is a good deal of entry and exit into the field, with new entrants joining at particular times when financing is available and novel scientific opportunities can be pursued. The rate at which new nodes appear in the network is, in part, determined by the success that existing nodes have in making progress on a technological frontier. Moreover, many of the participants in the field are "multivocal," that is, they are capable of performing multiple activities with a variety of constituents (Burt 1992; Padgett and Ansell 1993; White 1985, 1992). But multivocality is not distributed evenly, those organizations that are more centrally located in the industry have access to more sophisticated and diverse collaborators and have developed richer protocols for collaboration (Powell et al. 1996).

To illustrate the questions we are pursuing, consider a contemporary dance club, where revelers compete to get inside; once inside, they may dance in groups or repeatedly with only one partner or serially with many partners during an evening. The mix of available partners changes as the evening goes on, and diverse styles of music are played in different rooms. While new partners may be chosen, the imprint of past choices often lingers. Some dancers may be highly sought after and some music may attract more dancers. Or, by way of contrast, consider a more formal setting, such as an early 20th century Swedish military ball, where the young officers would ride in a horse-drawn carriage around Stockholm in advance of the event with an official dance card and visit the homes of young women to obtain permission to dance with them from their parents. By the time of the dance, the aspiring young officers may have filled their dance cards and rehearsed their repertoire of conversation and dance.[6] In either complex setting, an analytically rich set of questions follows: Who dances which dance? With whom? When? To address these questions, one needs information about the cast of participants, the repertoire of activities performed, and the sequences linking partners and

---

[6] We thank Örjan Sölvell, Stockholm School of Economics, for the detailed description of the status dynamics of the social world of his grandfather, Otto Hammar, and for a photocopy of his grandfather's 1911 dance card.

activities. As the combinations of partners and dances unfold, collective dynamics emerge. Individual choices may cumulate into a cascade, resulting in everyone following similar scripts. Or trends may cluster and find coherence only in small, densely connected groups. Choices made early may strongly affect subsequent opportunities, but path dependence might be offset by a constant flow of new arrivals and departures. The challenge to understanding any such highly interwoven system is to relate the behavior and dynamics of the entire structure with the properties of its constituents and their interactions. The aim is to discern what types of actors and relationships are most critical in shaping the evolution of the field at particular points in time.

We assess different sources of attachment bias and test to see if these simple rules guide the process of partner selection, and if so, for which participants and at what points in time? We supplement the idea of accumulative advantage with alternative mechanisms that sociologists have repeatedly found to be important in the formation of social and economic ties and the evolution and replication of social structures. The first alternative to early advantage or rich-get-richer is *homophily* (McPherson and Smith-Lovin 1987), a process of social similarity captured best in the phrase, "birds of a feather flock together." A second alternative is based on following the trend, thus the participants observe others and attempt to match their actions to the dominant behavior of the overall population (White 1981; DiMaggio and Powell 1983). In this context, action is triggered by a sense of necessity, by a desire to keep pace with others by acting appropriately (March and Olsen 1989, pp. 160–62). This pattern may also arise from participants reacting in similar ways to common exogenous factors.

We contrast arguments based on rich-get-richer processes with mechanisms such as homophily and appropriateness, and propose a new model based on multiconnectivity (the multiple linking of partners both directly and through chains of intermediaries) and a preference for diversity. To pursue the dance imagery, homophily suggests that, when you select a new partner, he or she is someone with attributes similar to those of your previous partners. A rich-get-richer process involves competing for the most popular dancer. Following the trend entails choosing both a partner and a dance that are comparable to the choices of most other participants.

A preference for diversity, however, suggests a search for novelty and the inclination to move in different communities and interact with heterogeneous partners. Our ideas concerning multiconnectivity involve several intuitions. A cohesive network, with plural pathways, means participants are connected through different linkages. Thus many nodes must be removed to disconnect such a structure, meaning that such groupings are highly resilient. The larger the number of pathways for communication

and exchange, the more rapidly news percolates through the network. In turn, when more knowledge is exchanged, participants attend to their network partners more intensively (Powell 1990). The enhanced flow of ideas and skills then becomes an attraction, making the network more appealing to join. Rapid transmission and diverse participants enhance both the likelihood of recombination and the generation of novelty. In the language of organizational learning, diversity entails a preference for exploration over exploitation (March 1991). Of course, we do not necessarily expect that one mechanism dominates at all time periods and exerts equal gravitational pull on every participant. The very essence of dynamic systems is that they constantly change over time. The actors may well play by different rules at different points in time, depending on the experience of their partners and their position in the social structure. Moreover, alternative organizing principles may be dominant at different stages in the formation of the network. Framed more formally, the alternative mechanisms can be stated in a series of hypotheses:

HYPOTHESIS 1.—*Network expansion occurs through a process in which the most-connected nodes receive a disproportionate share of new ties* (accumulative advantage).

HYPOTHESIS 2.—*Network expansion follows a process in which new partners are chosen on the basis of their similarity to previous partners* (homophily).

HYPOTHESIS 3.—*Network expansion entails herdlike behavior, with participants matching their choices with the dominant choices of others, either in mutual response to common exogenous pressures or through imitative behavior* (follow-the-trend).

HYPOTHESIS 4.—*Network expansion reflects a choice of partners that connect to one another through multiple independent paths, which increases reachability and the diversity of actors that are reachable* (multiconnectivity).

Field Structuration: Science Meets Commerce

Our empirical focus is on the commercial field of biotechnology, which developed scientifically in university labs in the 1970s, saw the founding of hundreds of small science-based firms in the 1980s, and matured in the 1990s with the release of dozens of new medicines. This field is notable for its scientific and commercial advances and its diverse cast of organizations, including universities, public research organizations, venture capital firms, large multinational pharmaceutical corporations, and dedicated biotech firms (which we refer to as DBFs). Because the sources of scientific leadership are widely dispersed and rapidly developing and the relevant skills and resources needed to produce new medicines are broadly

distributed, the participants in the field have found collaboration essential. The evolving structure of these collaborative ties is the focus of our network study. Concomitant with changes in the network, an elaborate system of private governance has evolved to orchestrate these interorganizational relationships (Powell 1996), and the internal structure of organizations has changed accordingly, coevolving with the collaborative network.

In the early years of the industry, from 1975 to 1987, most DBFs were very small start-ups that relied deeply on external support. No DBF in this period had the necessary skills or resources to bring a new medicine to market; thus they became involved in an elaborate lattice-like structure of relationships with universities and large multinational firms (Powell and Brantley 1992). The large multinational firms, with well-established internal career ladders, lacked the cutting edge of university science. Lacking a knowledge base in the new field of molecular biology, the large companies were drawn to the start-ups, which had more capability at basic and translational science (Gambardella 1995; Galambos and Sturchio 1996). This asymmetric distribution of technological, organizational, and financial resources was a key factor in driving early collaborative arrangements in the industry (Orsenigo 1989; McKelvey 1996; Hagedoorn and Roijakkers 2002).

Many commentators at the time argued that these interdependent linkages were fragile and fraught with possibilities of "hold-up," in which one party could opportunistically hinder the other's prospects for success.[7] Some analysts argued that the field would undergo a shakeout. Thus large pharmaceutical companies would assume dominance, while founding rates for new firms would slow to a trickle (Sharpe 1991; Teece 1986). But as these observers and others came to recognize, a shakeout did not occur, nor did cherry-picking of the most promising firms by larger companies prove viable.[8] Instead, the ensuing period saw the give-and-take, the mutual forbearance of relational contracting (Macneil 1978), become institutionalized as common practice in this rapidly developing field. By the late 1980s, some of the DBFs had become rather large and formidable organizations in their own right, while many of the big pharmaceuticals created in-house molecular biology research programs (Henderson and Cockburn 1996; Zucker and Darby 1997). Even as mutual need declined

[7] See Holmström and Roberts (1998) for a useful review of the hold-up problem and a discussion of various alternatives to vertical integration as a solution.

[8] Hybritech was one of the better known DBFs of the mid-1980s. This San Diego–based firm was purchased in 1986 by the large pharmaceutical company Eli Lilly for $300 million. Within a year, no Hybritech employees remained with Lilly; meanwhile more than 40 firms have been founded by former Hybritech employees (Walcott 2002).

as a basis for interorganizational affiliations, the pattern of dense connectivity deepened, suggesting the original motivation of exchanging complementary resources had changed to a broader focus on utilizing innovation networks to explore new forms of R&D collaboration and product development (Powell et al. 1996). This is the period we focus on, with data from 1988 to 1999. Table 1, which lists the top-selling biotechnology drugs in 2001, illustrates the division of innovative labor that has typified the field. All 10 drugs were developed by biotech firms, but only five were marketed by biotech firms, and just four by their originator. In the other five cases, a large pharmaceutical company handled or guided the marketing in return for a hefty share of the earnings. Comparable data from the early 1990s indicate that marketing power and control of revenues was dominated overwhelmingly by the pharmaceutical giants (Powell and Brantley 1996). A notable feature of drug development is that there is no consumer loyalty to a company and limited brand loyalty as well. Combine these influences with a market structure that has many winner-take-most features, and the outcome is a volatile, fast-changing field.

A number of factors undergird the collaborative division of labor in the life sciences. No single organization has been able to internally master and control all the competencies required to develop a new medicine. The breakneck pace of technical advance has rendered it difficult for any organization to stay abreast on so many fronts; thus linkages to universities and research institutes at the forefront of basic science are necessary (Orsenigo, Pammolli, and Riccaboni 2001). The high rate of technical renewal is reflected in patent data. Figure 1 shows the brisk rise of life science patenting and highlights the similarities in the technological trajectories of universities and biotech firms.[9] Note the parallel climb of university and DBF patenting in the mid-1980s, a steady ascent for the next 10 years and a steep increase in the late 1990s. Universities start out ahead and then are passed by DBFs in 1997, but the more important point is the extent to which both become members of a common technological community (Owen-Smith and Powell 2001a and 2001b). This joint membership in a community greatly increases the frequency of interaction between universities and industry.

The availability of funding has also increased markedly as biomedicine has become a major force in modern society. The total budget of the U.S. National Institutes of Health (NIH), a key funder of basic research that allocates approximately 80% of its budget to external research grants to universities and firms, nearly doubled under the Clinton administration,

[9] The data are drawn from the National Bureau of Economic Research patent database, using our sample of biotech firms and Owen-Smith's (2003) sample of the most research-intensive U.S. universities.

TABLE 1
TOP TEN BIOTECHNOLOGY DRUGS, 2001

Product	Indicated Use	2001 Sales (in Millions)	Developer	Marketer
Procrit	Red blood cell enhancement	3,430	Amgen	Johnson & Johnson
Epogen	Red blood cell enhancement	2,109	Amgen	Amgen
Intron A/ Rebetron	Hepatitis C, certain forms of cancer	1,447	Biogen, ICN	Schering-Plough
Neupogen	Restoration of white blood cells	1,346	Amgen	Amgen
Humulin	Diabetes mellitus	1,061	Genentech	Lilly
Avonex	Relapsing multiple sclerosis	972	Biogen	Biogen
Rituxan	B-cell non-Hodgkin's lymphoma	819	IDEC	Genentech, IDEC
Enbrel	Rheumatoid arthritis	762	Immunex	Immunex, American Home Products
Remicade	Rheumatoid arthritis, Chron's disease	721	MedImmune	Johnson & Johnson
Cerezyme	Enzyme replacement therapy	570	Genzyme	Genzyme

NOTE.—Data are drawn from Standard & Poor's "Biotechnology," May 2002.
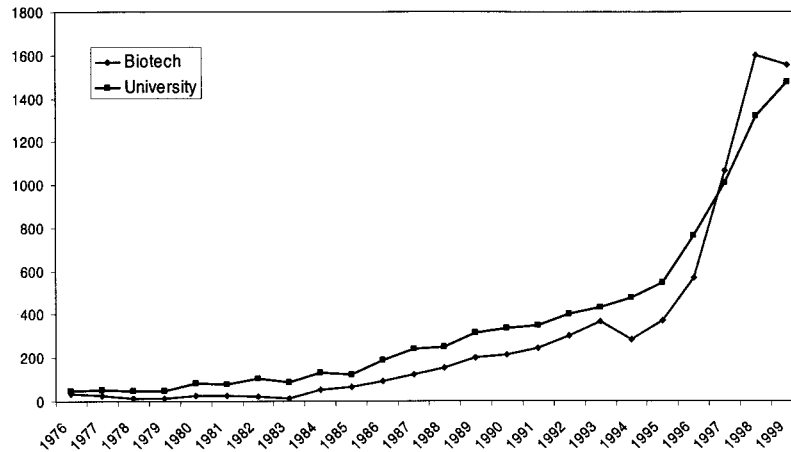
Fig. 1.—Life science patents assigned to biotechnology firms and research universities, 1976–99.

going from $8.9 billion in 1992 to $17.08 billion in 2000. The NIH plays a highly significant role in fostering exploration and variety on the research front. Internal R&D expenditures by biotech and pharmaceutical companies have also ramped up, from $6.54 billion in 1988 to $26.03 billion in 2000.[10] Venture capital disbursements, or seed money to biotech start-ups, have flowed into biotech, but more irregularly as the public equity markets have windows of opportunity when particular technologies are in vogue. As Lerner, Shane, and Tsui (2003) note, unexpected events affecting a single firm—notably the rejection or delay of a drug candidate by the U.S. Food and Drug Administration—can have pronounced effects on all firms' stock prices as well as their ability to raise capital. Consequently, venture funding of biotech is rather episodic, reaching $395.5 million in 1988, declining over the next three years, then jumping to $586.4 million in 1992, remaining around the half billion level for the next four years, then climbing to $1.1 billion in 1997, and staying above $1 billion in 1998 and 1999 (National Science Foundation 2002). Biotech financing by venture capital is also somewhat countercyclical; when there was great enthusiasm for Internet and telecom start-ups, interest in biotech waned. In recent years, with the burst of the Internet bubble and a precipitous decline in telecommunications, funding for biomedical ventures has been on the upswing. Biotech firms that are well positioned in the network

[10] Sources: NIH budget, http://www.nih.gov, 2001; Pharmaceutical R&D spending, Pharmaceutical Manufacturers of America annual surveys, http://www.pharma.org (downloaded November 2002).

with connections to basic research funding, industrial R&D support, and venture capital financing are not only able to obtain money from multiple sources, they also develop the capability to interact with varied participants. These experiences facilitate organizational learning and expand the scope and depth of an organization's knowledge base.

The different members of the field have varying catalytic abilities and competencies. Some of the participants are quite specialized, while others have a hand in multiple activities. Figures 2a and 2b provide simple count data to illustrate the relationship between functional activity and organizational form and suggests how the correspondence of activity and form has shifted over time. Figure 2a emphasizes an overall pattern of expansion in number of ties, and for three of the forms of organization, a branching out in terms of activity. Growth and diversity go hand-in-hand for DBFs, public research organizations (including public and private universities and nonprofit institutes and research hospitals; hereafter, PROs), and large pharmaceutical companies. Venture capital growth is notable as well, but as the percentages in figure 2b reveal, there is a strong co-occurrence of some forms and activities: government agencies specialize in R&D, venture capitalists (VCs) in finance. Some organizations, however, are able to shift their attention. While enlarging the number of ties, both dedicated biotech firms and public research organizations broadened their range of activities as well. Figure 2b, which reports the percentage of activity by organizational form, illustrates the pattern of specialization by government and VCs, and the diversification by biotech and pharmaceuticals, while PROs display a trend toward an R&D and licensing model. Venture capitalists come to dominate finance, and pharmaceuticals master commercial ties, while licensing and R&D are pursued by an array of participants.

Finally, as the biotech field gained coherence, and the pattern of reliance on collaboration proliferated, institutions emerged to both facilitate and monitor the process. Offices were established on university campuses to promote university technology transfer (Owen-Smith and Powell 2001*b*), law firms developed expertise in intellectual property issues in the life sciences, and venture capital firms provided financing along with management oversight and referrals to a host of related businesses. As these relations thickened and a relational contracting infrastructure grew (Powell 1996), the reputation of a participant came to loom larger in shaping others' perceptions. Robinson and Stuart (2002) argue persuasively that the network structure of the field becomes a "platform for the diffusion of information about the transactional integrity" of its participants. Centrality in the network increases the visibility of a participant's actions and reduces the need for more overt, contractual forms of control, such as an equity stake or dominance on the board of directors. We turn now

Fɪɢ. 2a.—Distribution of organizational forms and activities, 1988, 1993, and 1997 (COM = commercial, LIC = licensing, FIN = finance)

Percentage of Activity by Organzational Form 1988, 1993, 1997

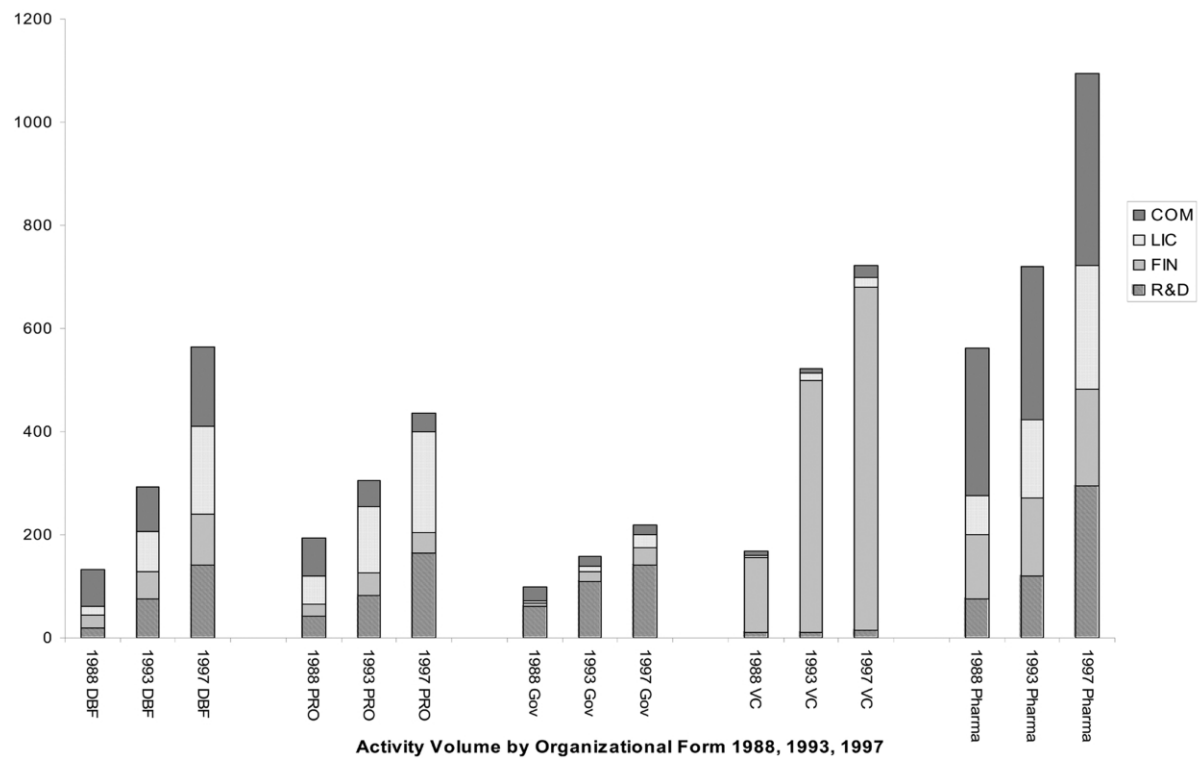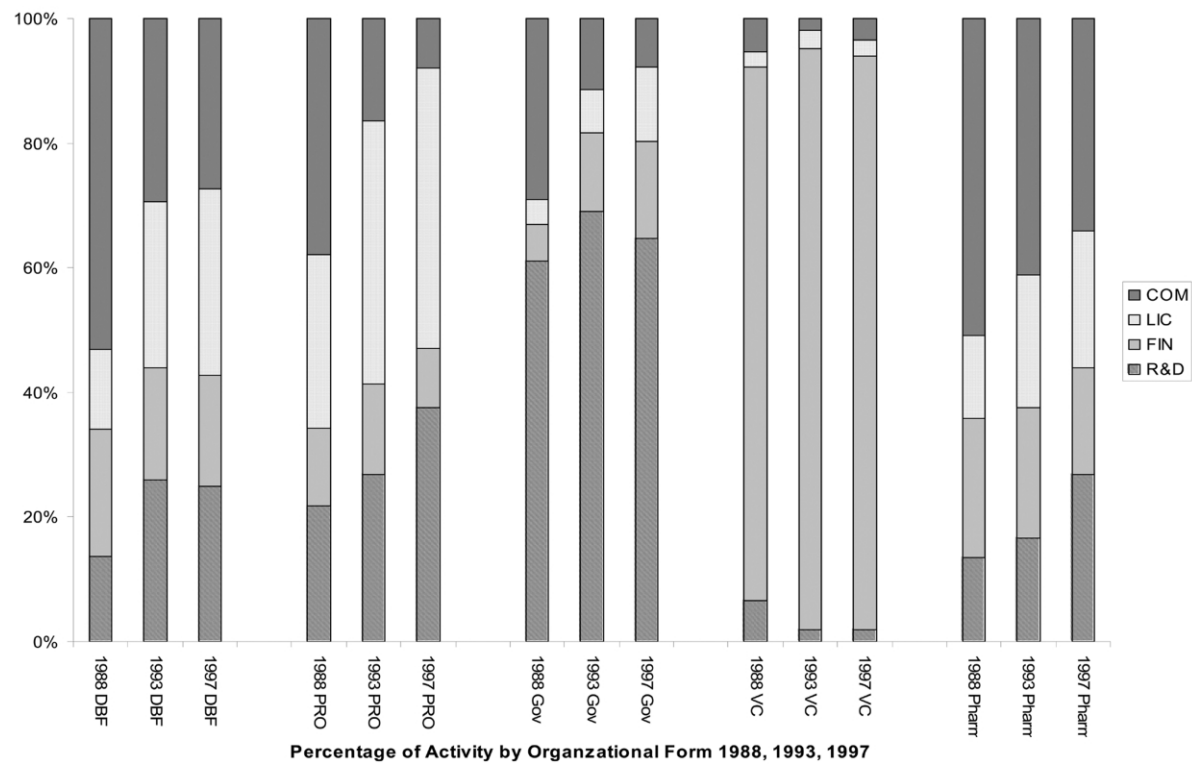Fɪɢ. 2b.—Distribution of organizational forms and activities, 1988, 1993, and 1997 (COM = commercial, LIC = licensing, FIN = finance)

to a discussion of the database we have developed to map the evolving structure of this field.

DATA AND METHODS

Our starting point in developing a sample is *BioScan*, an independent industry directory, founded in 1988 and published six times a year, which covers a wide range of organizations in the life sciences field.[11] Our focus is on DBFs. We include 482 companies that are independently operated, profit-seeking entities involved in human therapeutic and diagnostic applications of biotechnology. We omit companies involved in veterinary or agricultural biotech, which draw on different scientific capabilities and operate in quite different regulatory climates. The sample of DBFs covers both privately held and publicly traded firms. We include publicly held firms that have minority or majority investments in them by other firms as long as the company's stock continues to be independently traded. We exclude organizations that might otherwise qualify as DBFs, but are wholly owned subsidiaries of pharmaceutical or chemical corporations. Large pharmaceutical corporations, health care companies, hospitals, universities, or research institutes enter our database as partners that collaborate with DBFs. Our rationale for excluding both small subsidiaries and large, diversified chemical, medical, or pharmaceutical corporations in the DBF database is that the former do not make decisions autonomously, while biotechnology may represent only a minority of the activities of the latter. Their exclusion from the primary sample of DBFs eliminates serious data ambiguities.

The primary sample covers 482 DBFs over the 12-year period, 1988–99. In 1988, there were 253 firms meeting our sample criteria. During the next 12 years, 229 firms were founded and entered the database; 91 (of the 482) exited, due either to failure, departure from the industry, or merger. The database, like the industry, is heavily centered in the United States, although in recent years there has been considerable expansion in Europe. *BioScan* reports information on a firm's ownership, formal contractual linkages to collaborators, products, and current research. Firm characteristics reported in *BioScan* include founding data, employment

[11] The first volume of *BioScan* was released in 1987 by the biotech firm Cetus, but coverage was limited as many firms were reluctant to share data with a competitor. Oryx Press issued the first independent directory in 1988. To supplement *BioScan*, we consulted *Recombinant Capital* as well as including various editions of *Genetic Engineering and Biotechnology Related Firms Worldwide*, Dun and Bradstreet's *Who Owns Whom?* and *Standard &Poor's*. In addition, we utilized annual reports, Securities and Exchange Commission filings and, when necessary, made phone calls to companies.

levels, financial history, and, for firms that exit, whether they were acquired or failed. The data on interorganizational agreements cover the time frame and purpose of the relationship. Our database draws on *BioScan*'s April issue, in which new information is added for each calendar year. Hence the firm-level and network data are measured during the first months of each year. We define a collaborative tie, or alliance, as any contractual arrangement to exchange or pool resources between a DBF and one or more partner organizations. We treat each agreement as a tie and code each tie for its purpose (e.g., licensing, R&D, finance, commercialization) and duration. Some ties involve multiple stages of the production process. All such ties include commercialization activities, such as manufacturing or marketing, hence we code complex agreements as commercial ties. We say a connection, or link, exists whenever a DBF and partner have one or more ties between them.

We seek to explain the processes that attract two parties to one another, the evolving patterns of tie formation and dissolution, and the overall structure of the network. We code the dominant forms of partner organizations into six categories, representing those that populate the field: public research organizations, large multinational pharmaceutical corporations (as well as chemical and diversified health care corporations), government institutes (such as the National Cancer Institute or the Institut Pasteur), financial institutions (principally venture capital as well as banks and insurance companies); other biomedical companies (providers of research tools or laboratory equipment), and those DBFs that collaborate with other biotech companies. There are more than 2,300 non-DBFs in the partner database.

The four types of ties involve different activities, ranging from basic research to finance to licensing intellectual property to sales and marketing. Thus, the matrix of organizational forms and activities is 6 × 4, or 24 possible combinations of partner forms and functional activities of ties. Some of the cells are quite rare, but there are cases in every cell. Some of these activities involve the exchange or transfer of rights, while others require sustained joint activity. The latter obviously entails more integration of the two parties to a relationship. This difference is one reason we treat the four types of activities separately in most analyses.

Given the differentiation of organizational forms and types of ties, our approach has some limitations. In some, though not all, of our measures, we treat the type of tie or the form of the partner as equally important. Obviously, this is not altogether realistic; indeed, if the analyses were based on only a single year of data, this limitation would loom large. The benefit of this assumption is that it permits comparisons across time periods. There is also heterogeneity in the nature of participation of different organizational forms for a particular type of activity. At the extreme, an

R&D partnership between a global pharmaceutical company and a DBF may reach the scale of $1 billion, while a DBF's R&D alliance with a university laboratory may involve as little as $100,000 or $200,000. The salience of these limitations recedes as we add more years of observations to the data set. The advantages of 12 years of fine-grained data reside in capturing the length of relationships, the dissolution of ties to particular partners and the forging of ties to others, as well as the deepening of some ties. Issues of scale are assumed to be constant while we examine duration of ties and the extent to which the parties involved in a relationship share other partners in common at specific points in time. This approach allows us to speak to Salancik's (1995, p. 348) concern that network analysis should show how adding or subtracting a particular interaction in a network changes the coordination among the participants and either enables or discourages interactions between parties.

We do not collect data on the ties among the non-DBF partner organizations. In some cases, such ties would be very sparse or nonexistent (e.g., venture capital funding of universities or pharmaceutical companies); in other cases, they are more common (e.g., pharmaceutical support of clinical trials conducted at a university medical center).[12] The practical problem is that the data on a network of 2,310 × 2,310 disparate organizations would be very difficult to collect. Thus, we analyze the connections that DBFs have to partners, and the portfolio of DBFs with whom each partner is affiliated. To do this we use k-components to identify cohesive subsets of organizations.[13] This measure of multiconnectivity

[12] Our collaborators Fabio Pammolli and Massimo Riccaboni at the University of Florence have constructed a large database on R&D projects in the biomedical field that covers more than 10,000 external collaborations among participants in the life sciences throughout the decade of the 1990s. The most frequent type of partnership (approximately 45%) was between a biotech firm and a pharmaceutical company. The least common affiliation was between a pharmaceutical company and a public research organization (.05%).

[13] We are indebted to James Moody who applied to our network data his algorithm for k-components (Moody and White 2003) as our measure of cohesion. A k-component is a potent measure of cohesion because such a structure cannot be disconnected except by removal of k or more nodes. In contrast, a maximal subgraph (known as a k-core) in which all nodes have k degree or higher may be a disconnected graph (White and Harary 2001). The computation of k-components as units of cohesion showed a perfect (1.0) correlation for each time period between k-components and maximal k-core subgraphs. This result points to a highly cohesive network, with a well-connected core at the center. Save for a few small bicomponents on the periphery, there were singular rather than multiple overlapping k-components at each level (e.g., at each level of 8,7,6,5,4, and 3), *and* each unique k-component was embedded in a k-component at the next lower level of cohesion. Organizations in the eight-component for each time period, for example, were embedded in the seven-component, and so on down the stack. K-components are necessarily embedded, but they need not form a single hierarchy. Distinct hierarchies of k-components may form with as many as k-1 overlap-

does not require complete data on relations among nonbiotech partners, thus allowing us to analyze the total network of nearly 2,800 organizations. Our focus, then, is on cohesion, mediated exclusively by ties with DBFs.

Network measures such as the analysis of degree distributions, unlike k-components, require complete data. To this end, in the first and third of the analyses that follow, we separate our database into two parts: the complete one-mode network (482 × 482) of ties among DBFs, where we also have extensive data on the attributes of the 482 biotech firms, and the two-mode network that consists of complete data on the ties of DBFs to non-DBF organizations.

Over the past decade, we have also interviewed more than 200 scientists and managers in biotech and pharmaceutical firms, as well as university professors who are actively involved in commercializing basic research. Members of our research group have done participant observations in university technology licensing offices, biotech firms, and large pharmaceutical companies. Students working on the project have developed a large data set on the founders of biotech companies and analyzed the careers of scientists who joined biotech companies. In short, while the analyses presented here are based on data derived from industry sources, our intuition about the questions to address are grounded in primary data collection.

Analysis 1: Degree Distributions

Research on networks in the graph theoretic and statistical mechanics traditions often utilizes degree distributions as a diagnostic indicator of whether tie formation in a network (growth or replacement) is equiprobable (simple random) for all pairs of nodes or biased proportional to existing ties of potential partners. The degree of each node is measured as the number of other nodes directly connected to the focal node. Preferential attachment to already connected nodes is referred to as a popularity bias. Unlike the tail of a random bell curve whose distribution thins out exponentially as it decays, a distribution generated by a popularity bias has a "fat" tail for the relatively greater number of nodes that are highly connected. The fat tail contains the hubs of the network with unusually high connectivity.

Different types of degree distributions can be distinguished when plotted on a log-log scale, with log of degree on the x-axis and log of the number of nodes with this degree on the y-axis. The degree distribution for a network in which the formation of edges is governed by a popularity

ping nodes in common. Such was not the case for the cohesive biotech network (White et al. 2004).

bias, where nodes with more connections have a higher probability of receiving new attachments, would plot as a straight line on the log-log graph, indicative of a power law.[14] A power-law degree distribution is not sufficient by itself, however, to identify the actual mechanisms that facilitate tie formation. A power-law degree distribution can reflect not only preferential attachment by incumbency (degree of attracting node) but also preferences for attractiveness, legitimacy, diversity, or a concatenation of mechanisms. In our study, we test for the existence of a preferential attachment process for each year to the next rather than assume its existence. We use multivariate analysis to try to discern specific mechanisms that govern tie formation and the type of degree distributions that different substantive processes generate. Before we begin that analysis, however, it is useful to examine the degree distributions.

In figure 3, we plot the aggregate degree distributions of DBFs on log-log scales for the six types of partner organizations: other DBFs; pharmaceutical, chemical, and diversified health-care corporations; universities, nonprofit research institutes and hospitals; government agencies; venture capital firms and other financial entities such as banks; and biomedical companies that supply research tools and instruments. For each of the six plots, the x-axis reflects log-degree (aggregated over all time periods) and the y-axis the log of the number of partners of the plotted form having a given network degree (also aggregated over all time periods). A power-law distribution, as noted, would plot as a straight line. For a network of sufficient density, an exponential-decay degree distribution, mimicking the results of a simple random process of tie formation, would form a convex curve that bows to the upper right away from the origin in the log-log plot. An exponential distribution can be statistically rejected based on the data presented in figure 3, along with a simple-random attachment process that would generate degree distributions that decay exponentially.[15]

The least squares fitted linear slopes for the log-log plot of the degree distributions in figure 3 are in the expected range for power laws, between

[14] A power law is a mathematical expression for a distribution that is unlike a normal bell curve with a peak in the middle, where most nodes have a similar number of ties. A scatterplot or histogram plotting power-law decay is a continually decreasing curve where values on the Y axis approximate a linear function of those on the X axis raised to a fixed power α, hence $\log y = C + \alpha \log x$ where C is the intercept constant.

[15] The exponent in this function is not fixed, but varies with values on the X axis. Hence, the exponential $y = b + C^x$ is linear in $\log y$ and x, but is bowed outward in a log-log plot. Inequality resulting from a simple random process of tie formation is accidental, or, in the case of network growth, it may be due to early entry. Networks with exponential degree distributions do not have hubs that are extreme outliers in terms of very high degree.
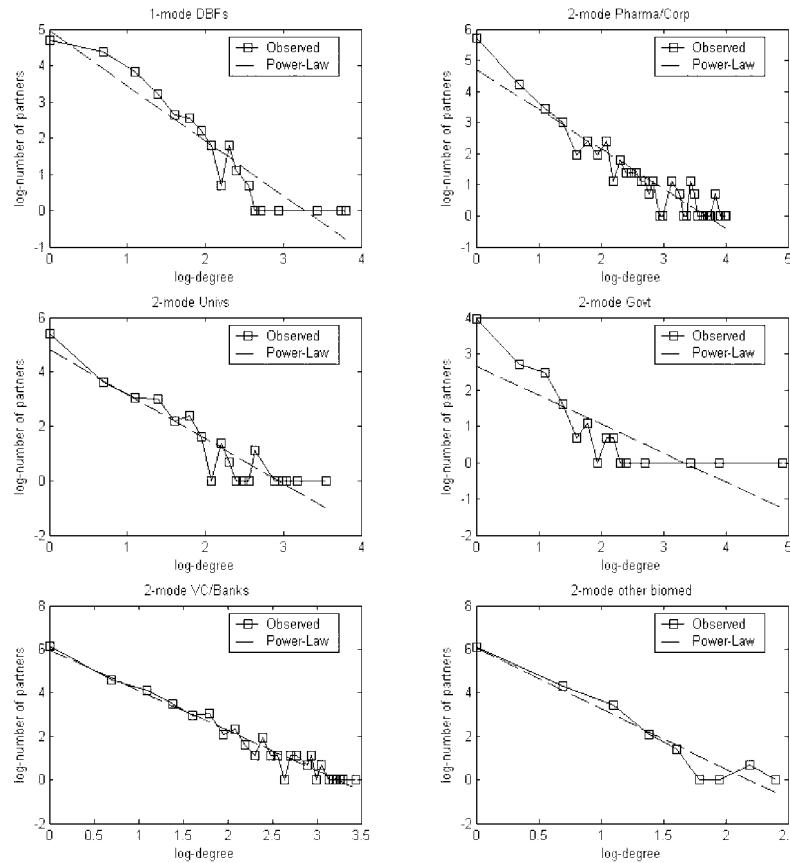
Fig. 3.—Degree distributions by type of partner, all years

1.1 and 2.7, over four orders of magnitude of variation in degree. Only the left half or low-degree half of one of the distributions, that for government agencies, has a slope anywhere near three, as expected from pure preferential attachment for large networks (Barabási and Albert 1999; Bollobás and Riordan 2002). This government degree distribution is interesting because of the highly pivotal central role of the NIH, which is a key funder of basic research. The NIH is the most active partner in the entire network. That the slopes of the five other distributions are considerably less than three may be an indication that processes other than pure preferential attachment are operative.[16]

[16] Whether there are attachment processes in the extremes of several of the graphs that go beyond power-law attachment biases cannot be determined due to the very small

Although these degree distributions are aggregate measures over all time periods, they give some hint of growth processes in attachment. The shape of the distributions mimics what would be expected for tie formation in the biotech network from a process of preferential attachment to degree, but power-law slopes that are closer to one than to three suggest that other substantive processes are operating. These static snapshots of the degree distributions, however, do indicate that the importance of different organizational forms varies with respect to patterns of affiliation. To explore the dynamics of the field and the changing impact of different organizations, we next present a series of network visualizations, followed by a statistical examination of the actual attachment processes.

Analysis 2: Discrete-Time Network Visualizations

We utilize Pajek, a freeware package for the analysis and visualization of networks, to present a series of discrete-time images of the evolution of the biotechnology network. Pajek employs two powerful minimum energy or "spring-embedded" network drawing algorithms to represent network data in two-dimensional Euclidian space. These algorithms simulate the network of collaborations as a system of interacting particles, in which organizational nodes repel one another unless network ties act as springs to draw connected nodes closer together. Spring-embedded algorithms iteratively locate a representation of the network that minimizes the overall energy of the system, by reducing the distance between connected nodes and maximizing the distance between unconnected nodes. (For more on Pajek visualizations, see app. A.)

We generate two sets of images for the time period covered by our database. To simplify the presentations, we include only those organizations in the main component each year, thereby removing the isolates from this large and expanding network.[17] The first set of images presents the full collaboration network, while the second set represents the new ties added each year. (Given space constraints here, we select a subset of

---

numbers of organizations with high degree. A degree distribution that plots as a concave curve, bent in toward the origin on a log-log scale, might indicate a super-power-law process in which more complex rules of attachment are operating and power-law inequality is accelerated. With one small exception, hardly systematic, there is no evidence in fig. 3 for a super-power-law process.

[17] The main component is the largest connected cluster in the network. It eliminates both isolates and small disconnected clusters. Most network measures are based on the main component, which is a connected graph for which measures can be generated. In substantive terms, the main component is the largest subset of organizations that can reach each other through direct and indirect paths. The percentage of organizations connected to the main component is 85.4% in 1988, dips slightly to 80.3% in 1992, and rises to 98.6% by 1999.

these pictures for inclusion in the article, but we urge readers to view the complete set in the online version of *AJS*.) The visualizations afford a multifaceted view of network evolution, including growth in the number of participants, changes in the purpose of ties, and the formation of new ties.

The Pajek figures are designed to reflect visually our hypothesized models of attachment. The graphics cannot adjudicate between the various models, which we do in statistical analyses below, but they do provide suggestive evidence, or existence proofs, if you will. Consider what the full network and new tie visualizations might look like if strong versions of our four mechanisms were operating. After a sketch of these stylized images, we then turn to the data and examine a series of empirical visualizations. If an accumulative advantage process drives attachment, the spring-embedded images would display a preponderance of star-shaped structures attached to large nodes. These stars would be positioned at the center of the image and continue to grow rapidly over time as newcomers (triangles) affiliate with the most-connected incumbent nodes. We scale the size of nodes to represent the number of ties an organization has. Under a rich-get-richer model, we would expect few small nodes to have numerous new partners, and many large nodes would be situated at the network's center. In contrast, if homophily strongly conditions tie formation, then we would anticipate images differentiated into coherent and loosely interconnected single-color clusters. These homophilous clusters should be fairly dense and dominated by characteristic organizational forms without necessitating a preponderance of stars tethered to large central nodes. New entrants (pictured as triangles) would move into the neighborhoods that most closely matched their type and profile.

Were a follow-the-trend logic dominant, new ties would be overwhelmingly uniform in color and the predominant color should reproduce the previous year's pattern of affiliation. If, for instance, R&D (red) ties dominated in a prior year, neophytes would generally enter the network through this established route and the images would be fields of red. A preference for diversity of partners and pathways implies less visual coherence than the other mechanisms. Multiconnectivity coupled with heterogeneous activities would show clusters in which all four colors would be evident in the springs. The nodes at the center of the network should be different colors, reflecting the array of organizational types. Some small nodes should be noticeable at the center of the field, while some large nodes should have few new partners. New entrants should be sprinkled throughout the network. Empirical reality is, of course, more messy than stylized models. We use these thought experiments to make the detailed visualizations of this large database more interpretable to readers not familiar with graphical representations of network dynamics.

Figure 4, showing the main component in 1988, serves as the starting point. The color of the nodes reflects their organizational form, with light blue a DBF, yellow a large pharmaceutical, chemical, or diversified medical corporation, and brown a government institute or agency. In subsequent years gray nodes become important, reflecting the growing imprint of venture capital. The springs are colored according to the functional activity reflected by a tie, with red an R&D partnership, magenta a licensing agreement, green a financial relationship, and dark blue indicating an alliance involving one or more stages in the commercialization process, ranging from clinical trials to manufacturing to sales and marketing. Node size is scaled to standardized network degree in the total network, reflecting variation in the extent of degree connectivity among the organizations. To return to our earlier metaphor of a dance, the representation captures dancers with different identities who may participate in different types of dances with one or many partners. The size of a node, given the power-law tendencies in the degree distributions, might reflect a rich-get-richer process. In the images that follow, particularly the visualizations of new ties, there are several large blue DBF nodes that are clearly the most attractive stars of the network. In later years, there are nodes that have multiple linkages for different activities, reflecting a preference for diversity. Looking at the overall population rather than specific nodes, we observe shifts in the dominant activities as well as changes in the composition of the nodes, which illustrate the overall trends in the field.

Several key features stand out for 1988 in figure 4. The predominant color is blue, and the most active participants are biotech firms, pharmaceutical companies, and several government agencies. The strong impact of commercialization ties is a clear indication of the dominant strategy of mutual need that characterized the industry's early years. Biotech firms lacked the capability to bring novel medicines to market, while large firms trailed behind in understanding new developments in molecular biology (Gambardella 1995; Powell and Brantley 1996; Henderson, Orsenigo, and Pisano 1999). Finance ties (green) are less prevalent and only a few venture capital firms (gray) are present, providing further evidence that most DBFs supported themselves by selling their lead product to large corporations, who subsequently marketed the medicine and pocketed the lion's share of the revenues. Clustered in the center are red (R&D) and magenta (licensing) ties, which show that DBFs with significant intellectual property and strong research capability are highly sought for collaboration. The large multiconnected nodes in the center of the representation are a small group of established, first-generation DBFs, major multinational firms, and government institutes (the NIH and the National Cancer Institute). In the detail shown in the lower right-hand corner of figure 4, we identify a handful of the largest nodes: NIH, which will serve as an

Key For All Figures
CYAN NODE = DBF
ORANGE = PRO
BROWN = Gov't
YELLOW = Pharma
GRAY = VC
WHITE = Other
Triangle = New Entrant
Circle = Incumbent
Node size = standardized network degree, constant within figures
RED TIE= R&D
GREEN = FINANCE
BLUE = COMMERCIAL
MAGENTA = LICENSING
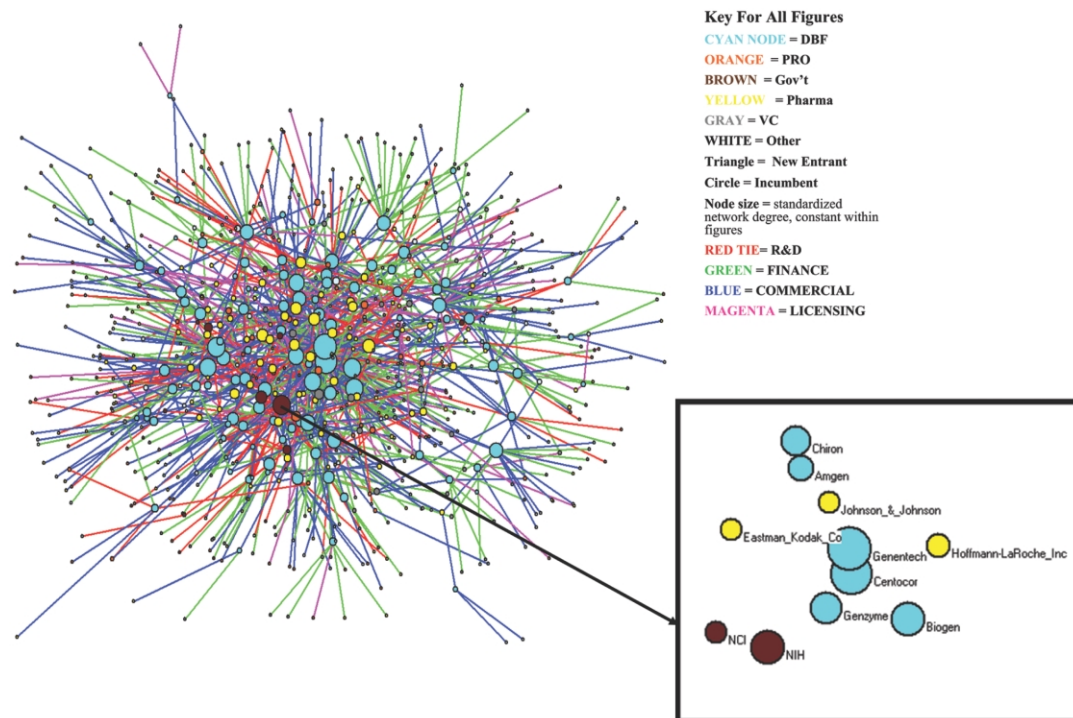
FIG. 4.—1988 main components, all ties

orienting node in the full network figures because of its centrality, and NCI; a group of first-wave biotechs founded in the 1970s and early 1980s (Genentech, Centocor, Amgen, Genzyme, Biogen, and Chiron, which became the largest and most visible firms by 1988) and three multinational firms (Eastman Kodak, Johnson & Johnson, and Hoffman La Roche). The close proximity of the Swiss firm Hoffman La Roche to Genentech is interesting because in 1990, the Swiss firm became the majority stockholder of Genentech. Kodak & Johnson and Johnson reflect the broad interest in biotech by a range of large firms in different technology sectors. Kodak soon drops out of the center of the field, as does Johnson & Johnson, but the latter returns in the late 1990s by acquiring two companies, first Centocor and then Alza, both established, well-connected DBFs. Kodak's subsequent loss of centrality and Johnson & Johnson's purchase of two incumbent DBFs illustrates that a first mover, rich-get-richer account does not always hold, even for some of the early resource-rich participants.

In figure 5, we present the new ties added in 1989. To return to the dance metaphor, the music has stopped temporarily while new partners are being chosen. We add shape to the presentation, with triangles representing new entrants to the network, while circles are the incumbents. Note the very active role of NIH (the largest brown node) in forging R&D ties with new entrants, and the appearance of many grey triangles, illustrating the growing involvement of venture capital in financing DBFs. Node size continues to be scaled for network degree in the prior year's full network, thus graphically representing how network position in one year may condition the addition of new ties in a subsequent year. The distribution of node sizes might indicate the importance of rich-get-richer models of attachment. In the initial years, we see visual affirmation of the positive effect of number of previous ties on new tie formation. Note that several large blue DBF nodes are at the center of the new tie network. The dominance of commercialization springs (blue) in 1988 recedes markedly in the 1989 new ties picture. R&D and finance are the main avenues generating growth in the network. Observe that large nodes with just a few ties have little diversity in their partners, while some of the smaller nodes adding many new ties have a wide variety of partners.

We fast forward to 1993, but encourage readers to follow the annual changes in the biotech field represented by the visualizations available on the AJS Web site.[18] Figure 6 portrays a large expansion in activity,

[18] We use annual changes as a matter of convention as publicly traded firms routinely provide accounts of their activities on a yearly basis and various data sources are organized in this manner. We have spent a good deal of energy studying the sequences of partners and activities, and have analyzed the time in monthly intervals to key events, e.g., first tie, first R&D collaboration, going public, etc. Because the data are more reliably reported annually, we use year-to-year changes in the visualizations.
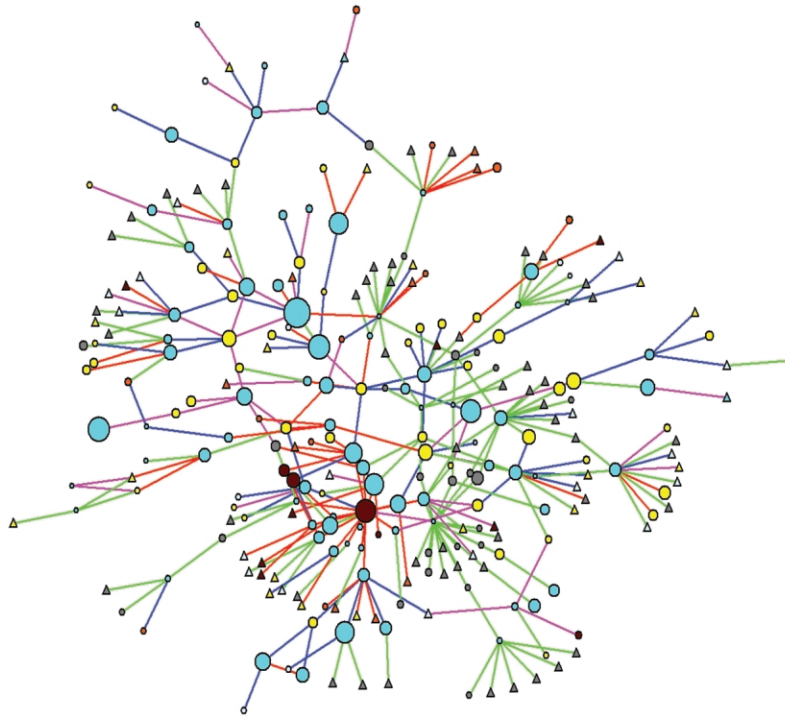
Fig. 5.—1989 main component, new ties

with green ties (finance) now much more prominent. We also observe a shift in the composition of the most connected participants. Put colloquially, the music has changed—from commercialization to finance, and accompanying the shift in the predominant collaborative activity is both an increase in the number of highly connected nodes (there are roughly three times as many larger nodes as in 1988) and the march-in of gray (venture capital) and orange (university) nodes. This shift in the primary locus of activity is important on a number of dimensions. Finance, as opposed to commercialization, has a powerful mobilizing effect, enrolling new types of actors (VC, and subsequently investment banks, pension funds, university endowments, etc.) in financing the expansion of the field. As we have shown elsewhere, the locus of venture capital-financed biotech start-ups was initially the Bay Area and Boston, but by the end of the 1990s had spread to a number of key regions in the United States and Europe (Powell, Koput, Bowie, and Smith-Doerr 2002; Owen-Smith et al. 2002; Sorenson and Stuart 2001). Thus growth in the number of new firms, new partners, and new ideas is enhanced by an increase in financial
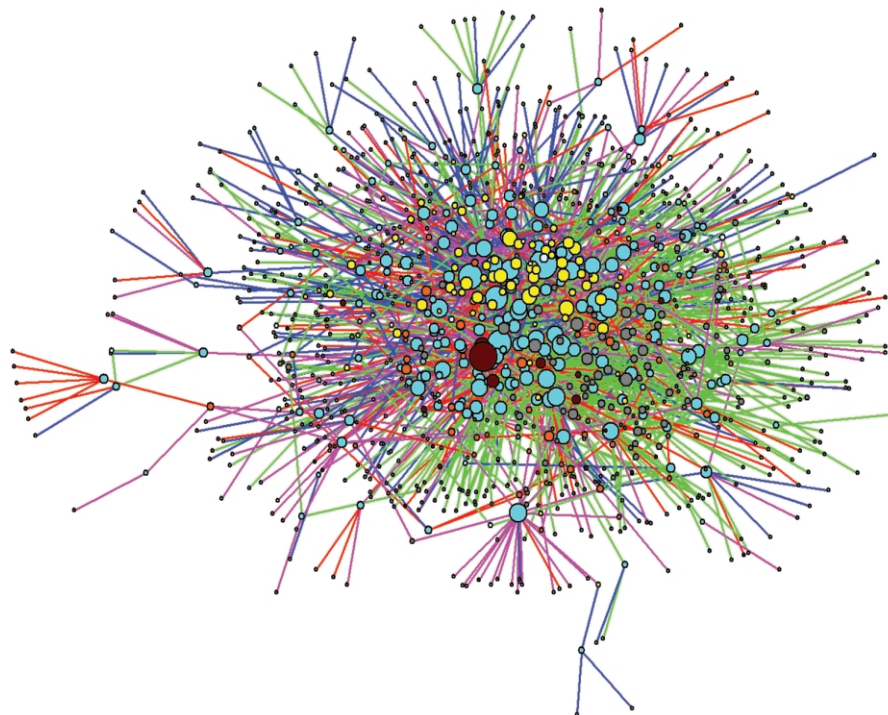
FIG. 6.—1993 main component, all ties

linkages, signaling the important role of the public equity markets in fostering growth. In contrast, commercialization is a more restrictive activity. The ability to manufacture new biomedical products was a relatively scarce skill, as was the ability to market and distribute a new medicine throughout the world. A relatively small number of large firms had these capabilities, and it took at least a decade before DBFs developed these skills. Consequently, during the first two decades of the field, commercialization ties flowed through a small set of dominant multinationals and a handful of established biotech firms. Thus, not only were the number of participants limited, commercialization is a "downstream" activity; indeed, when it involves the sale of a new medical product, it is the "last dance" in the product life cycle. Finance, in contrast, is an "upstream" activity, which, in turn, fuels R&D, licensing, and commercialization, and thus enrolls more participants in the industry network.

The organizational composition of the center of the field has shifted as well. Two research universities, MIT and Harvard, along with a handful of leading VC firms, are now at the center. The composition of multinational firms shifts from diversified chemical and medical companies to some of the giants of the pharmaceutical sector, for example, Schering-Plough and Merck. The diversity of organizational forms at the center of the network is notable in that these varied organizations—DBFs, pharmaceuticals, VCs, research universities, and government institutes—operate in distinct selection environments, subject to very different pressures and opportunities. Networks anchored by diverse organizational forms are more robust to both failure and attack than homophilous structures (Albert, Jeong, and Barabási 2000; Owen-Smith and Powell 2004). Such diversely anchored, multiconnected networks are less likely to unravel than networks reliant on a single form of organization for their cohesiveness.

The picture of new ties in 1994 (fig. 7) reflects growing complexity in the activity sets of the participants. On the right side of the network, finance is very pronounced, and there are many more gray nodes, which are growing in size. But note there are now yellow nodes (which are not triangles, so these are not new entrant pharmaceutical companies) involved in financing smaller biotech firms. On the left side of the figure, we see blue and green ties linked to well-connected DBFs. At the center of the network is the NIH, the anchor of R&D activity and the largest node, linked to both small as well as large DBF nodes. The picture of new ties in 1994 illustrates the growing multivocality of the industry, with both well-connected DBFs and pharmaceuticals developing the capability to finance younger firms, contribute to basic and clinical science, and commercialize new medicines. The overall picture has shifted from one in which commercialization and rich-get-richer were the dominant scripts
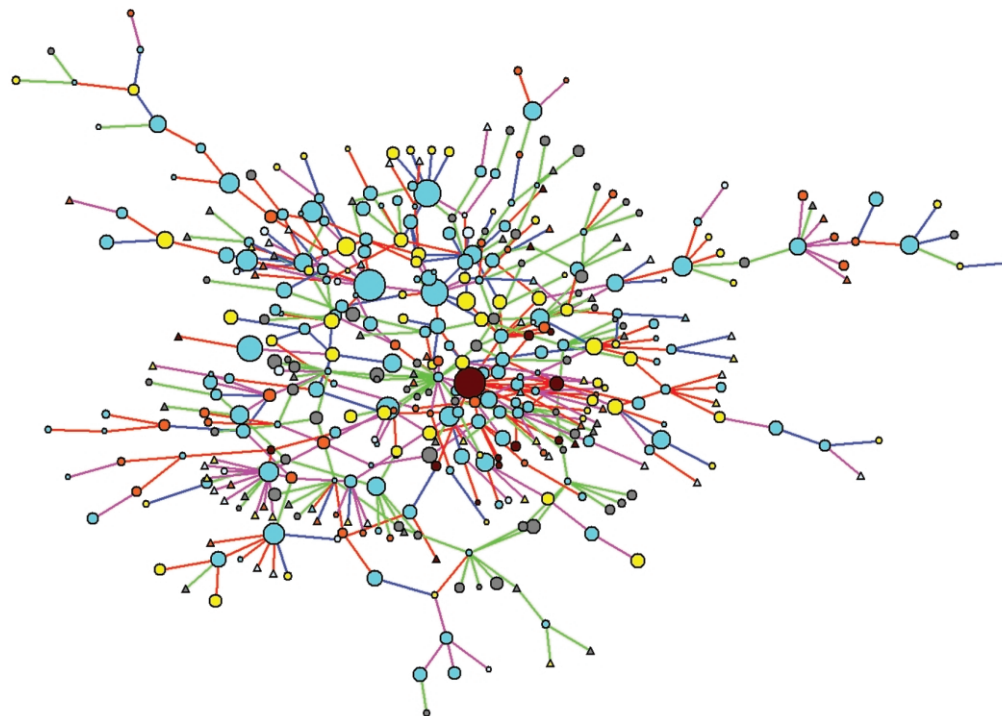
Fig. 7.—1994 main component, new ties

to one in which finance is generating much more diversity in activity and there is more heterogeneity in the makeup of the key participants. With greater numbers of new ties added in 1994, the more treelike structure of springs in 1989 has now changed to reflect multiconnectivity, that is, a more cohesive structure even among the new partnerships.

The density of the field and the number of participants continue to grow throughout the 1990s. The picture of all ties in the main component in 1997 (fig. 8) illustrates a growing number of large nodes, strong expansion of collaborative activity and a reaching out to new entrants, and more varied types of organizations at the center of the figures. In 1997 (fig. 8), we see a very cohesive hub of DBFs, pharmaceuticals, venture capital firms, universities, and the NIH complex. The colors of the ties are less discernable, reflecting the fact that the most active members of the network are now either engaged in multiple activities or connected to DBFs who are. The pharmaceutical sector underwent a period of consolidation in the mid to late 1990s, as mergers and acquisitions became commonplace. Novartis, formed out of the merger of the two large Swiss firms Sandoz and Ciba Geigy, and Glaxo Wellcome, the product of the acquisition of Burroughs Wellcome by fellow British giant Glaxo, are in the center of the network. Outside the center there are several specialist DBFs, one with a very active commercialization portfolio and another with a licensing cluster. In the new ties image for 1998 (fig. 9), finance continues to be generative in pulling in new entrants. For the first time, a handful of pharmaceutical giants, bolstered by a round of mergers and acquisitions, are central in the new tie network. While the NIH remains in the core, it is no longer clearly the largest node. Many more DBFs are found in the center and on the edges of the new tie network, reflecting the growing presence of second-generation firms who are active in the field. The consolidation in the pharmaceutical sector is producing a rich-get-richer effect among the largest multinationals, but these survivor corporations have learned to do more than just commercialize the lead compounds of the smaller DBFs. The big multinationals have become multivocal. Meanwhile, a combination of DBFs, universities, and government institutes are active in pulling in new participants.

The full network image for 1997 (fig. 8) shows a notable clustering of financial ties on the right side, with connections to smaller-size DBF nodes. This shift underscores the cyclical nature of venture capital, which involves taking a firm public, ending that relationship and moving on to finance new firms. The number of yellow nodes at the center has decreased, as consolidation shrinks the number of big pharmaceutical firms. In turn, the size of each yellow node increases, as their portfolio of alliances grows and diversifies. Recall that the NIH's budget for R&D grew markedly throughout the 1990s, and it remains hugely important as a funder
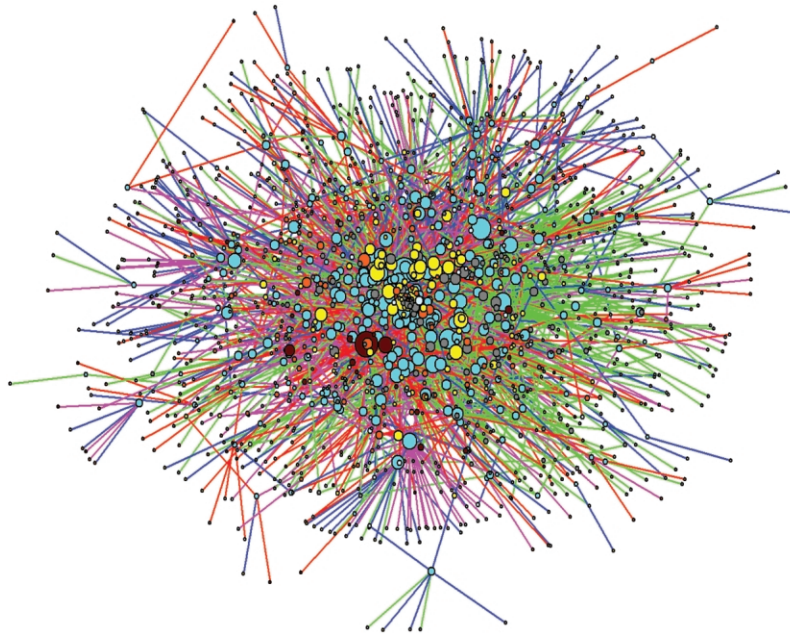
Fɪɢ. 8.—1997 main component, all ties

of basic research and as a participant in licensing the results of intramural NIH research done. The new tie network in 1998 (fig. 9) is the most expansive yet, with more than 1,100 new ties added. Multivocality is the dominant pattern here, as nearly every large node is engaged in multiple types of collaborations.

Pajek visualizations, which we use to show the extent of clustering as well as type of organization, type of tie, and network degree, can be considered to provide a visual goodness-of-fit test for the weighting of our hypotheses. As a supplement to the graphics, we provide count data on patterns of entry and exit into the network. We see in table 2 that the number of participants, both DBFs and partners, grows every year. But the rate of expansion for total number of ties and new ties outpaces the entry of organizations, suggesting a more connected field. The visualizations afford the opportunity to see the diverse types of organizations that are driving this connectivity.[19] The rate of tie dissolution grows, then wanes, then heads up, so there is considerable turnover in interorgani-

---

[19] There is a falloff in new ties and partner entry in 1999, which reflects incomplete reporting in 1999 annual reports. We have since added four more years of data and observe a downturn in 1999–2001 and a notable uptick in 2002 and 2003.
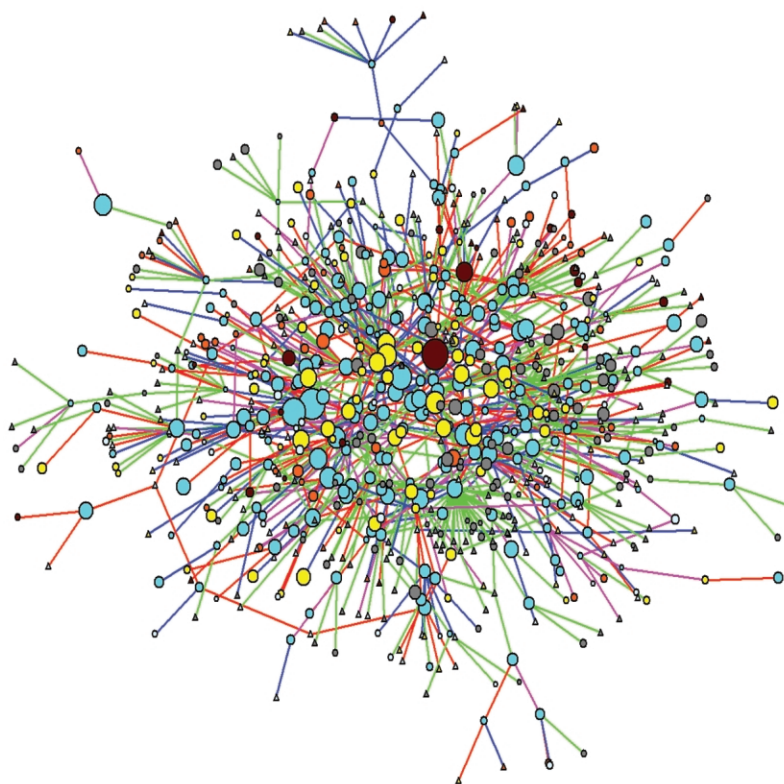
Fig. 9.—1998 main component, new ties

zational relations, reflecting both successful completion of some projects and dissolution due to lack of progress on others. The general picture is one of a continuing flow of new entrants into the field, alongside the forging of new collaborations, making for an increasingly dense network. To test our hypotheses, as well as the insights derived from the visualizations, we turn to a more fully specified analysis in which we take organizational variables, network portfolios, and time periods into account.

TABLE 2
PATTERNS OF ENTRY AND EXIT INTO THE NETWORK

YEAR	DBFs		Non-DBF PARTNERS			TIES			
	In Network	New Entrants	In Network	New Entrants	TOTAL	Initiated	New	Repeat	Discontinued
1988	155	. . .	579	. . .	1,565	. . .	. . .	. . .	. . .
1989	181	29	672	156	1,780	459	362	97	244
1990	199	28	747	146	1,954	472	379	93	298
1991	210	18	792	119	2,056	473	379	94	371
1992	233	31	800	123	2,162	544	429	115	438
1993	265	35	873	149	2,474	643	520	123	321
1994	297	35	938	119	2,783	634	508	126	325
1995	316	24	985	141	3,057	700	543	157	426
1996	340	34	1,058	165	3,373	912	737	175	596
1997	351	14	1,172	201	3,737	877	696	181	513
1998	360	13	1,313	251	4,295	1,121	957	164	563
1999	363	12	1,332	122	4,176	479	422	57	598
All years ...	450	273	2,265	1,692	8,818	7,314	5,932	1,382	4,703

NOTE.—Thirty two biotech firms never have any ties. We include these firms in subsequent analyses as part of the pool of potential partners. In addition, we exclude 45 organizations from the non-DBF partner list because we cannot determine their precise year of entry.

Analysis 3: Attachment Bias

We now turn to a statistical analysis of attachments between DBFs and various partner organizations. For analytic purposes, we distinguish between four categories of relationships:

1. New one-mode attachments, in which a DBF contracts with another DBF as a partner and is the first tie between this dyad.
2. Repeat one-mode attachments, in which a DBF contracts with a DBF partner and is not the first tie between this dyad.
3. New two-mode attachments, in which a DBF contracts with a non-DBF partner and is the first tie between the dyad.
4. Repeat two-mode attachments, in which a DBF contracts with non-DBF partner and is not the first tie between the dyad.

New attachments expand the structure of the network, whereas repeat ties thicken relations between existing dyads. One-mode attachments create a cooperative structure among competing biotech firms, while two-mode attachments engage different organizational forms to access resources and skills. Each relationship involves a focal DBF, a partner to which the attachment occurs, and a set of alternative partners with whom the DBF might have collaborated but did not. We refer to the set of partners to which the focal DBF might link for a particular observed attachment, including the partner to which the connection occurred, as the *risk set* for that attachment. For new one-mode attachments, the risk set is other DBFs, excluding those to which the focal DBF is currently or has previously attached. For repeat one-mode attachments, the risk set is, conversely, those current or prior partners of the attaching DBF. For new two-mode attachments, the risk set is partners other than DBFs, excluding those to which the attaching DBF is currently or has previously been linked. For repeat two-mode attachments, the risk set is, again conversely, all current or prior partners (other than DBFs) of the focal DBF.

MEASURES

We draw upon multiple measures to test each hypothesized mechanism across four classes of ties. We present and define the variables used in our statistical analyses, along with appropriate controls, in table 3. Our dependent variables are binary indicators of tghwhether an alliance of each of the four types occurs between a DBF and a partner, given that the DBF and partner are "at risk" of attaching. A DBF may forge connections to more than one partner in any given window of time, as long as alliances can be ordered such that a partner for which a linkage occurs at a specific moment is removed from (added to) the risk set for subsequent

TABLE 3
VARIABLES IN STATISTICAL TABLES

Label	Unit of Observation	Description
Dependent variables:		
New attachment	Tie	Computed as a binary indicator of whether an attachment occurs between a DBF and a partner for the first time
Repeat attachment	Tie	Computed as a binary indicator of whether an attachment occurs between a DBF and a partner, other than for the first time
Independent variables:		
Accumulative advantage:		
Firm degree	DBF	No. of ties of the DBF just prior to the attachment
Partner degree	Partner	No. of ties of the partner just prior to the attachment
Firm experience	DBF	No. of years since inception of DBF's first tie
New partner	Partner	Indicates whether this is the partner's first year in the network
Partner experience	Partner	No. of years since inception of partner's first tie
Prior ties	Dyad	No. of prior ties connecting the DBF-partner dyad
Prior experience	Dyad	Duration since first tie connecting the DBF-partner dyad
Homophily:		
Collaborative distance	Dyad	One-mode: Euclidean distance between the activity-type-by-partner-form profiles of the attaching DBF and partner, just prior to the attachment
Collaborative distance	Dyad	Two-mode: Euclidean distance between the activity-type profiles of the attaching DBF and partner, just prior to the attachment
Age difference	Dyad	One-mode: absolute difference in age between the attaching DBF and the partner, at time of attachment
Size difference	Dyad	One-mode: absolute difference in number of employees between attaching DBF and the partner, at time of attachment

Governance Similarity	Dyad	One-mode: dummy variable capturing whether both firms are publicly traded or privately held, at time of attachment
Co-location	Dyad	One-mode: indicator of whether the attaching DBF and partner are in the same region, at time of attachment.
Partner's partner collaborative distance	Partner's neighborhood	Average Euclidean distance between activity-type-by-partner-form profiles of attaching DBF and other DBFs attached to the partner
Partner's partner age difference	Partner's neighborhood	Average absolute difference in age between the attaching DBF and other DBFs attached to the partner
Partner's partner size difference	Partner's neighborhood	Average absolute difference in number of employees between the attaching DBF and other DBFs attached to the partner
Partner's partner governance similarity	Partner's neighborhood	Average absolute difference in whether publicly held between the attaching DBF and other DBFs attached to the partner
Partner's partner co-location	Partner's neighborhood	Average of indicator of whether the attaching DBF and other DBFs attached to the partner are in the same three-digit zip code region

Follow-the-trend:

Dominant trend	Field	% of other attachments up to the time of the attachment that are in same activity-type-by-partner-form category
Dominant type	Partner	% of a partner's ties that fall into the same activity-type category as the activity type of the attachment

Multiconnectivity:

Firm cohesion	DBF	DBF's maximum value of k for which the DBF is in a k-component, just prior to the attachment
Partner cohesion	Partner	Partner's maximum value of k for which the partner is in a k-component, just prior to the attachment
Shared cohesion	Dyad	Maximum value k of k-components occupied by both partner and DBF, just prior to attachment

TABLE 3 (*Continued*)

Label	Unit of Observation	Description
Firm tie diversity	DBF	Blau heterogeneity index over DBF's activity-type-by-partner-form portfolio, just prior to the attachment
Partner tie diversity	Partner	One-mode: Blau heterogeneity index over partner's activity-type-by-partner-form portfolio, just prior to the attachment
Partner tie diversity	Partner	Two-mode: Blau heterogeneity index over the partner's activity-type portfolio, just prior to the attachment
Prospective diversity	Dyad	Change in DBF tie diversity resulting from attachment
Partner's partner collaborative diversity	Partner's neighbor-hood	Average tie diversity of other DBFs attached to the partner in terms of activity-type-by-partner-form categories, at time of attachment
Controls:		
Period	Field	Period = 1: 1989–93, Period = 2: 1994–96; Period = 3: 1997–99
Timeline	Field	Linear time trend, computed as year of observation - 1987
Age	DBF	Duration in years since DBF's founding or first entry into biotech
Size	DBF	No. of employees of DBF
Governance	DBF	Indicates whether DBF is publicly or privately held
Location	DBF	Three-digit zip code
Form	Partner	Two-mode: indicates form of partner organization, e.g., biomedical corp, university, nonprofit, government, pharmaceutical or other for-profit
Type	Dyad	Indicates type of activity or exchange involved in collaboration, e.g. research, financing, licensing, commercialization.

new (repeat) attachments. The predictor variables operationalize the attachment mechanisms for each of the four hypotheses: accumulative advantage, homophily, follow-the-trend, and multiconnectivity. For each observed attachment, partner and dyad measures are computed for all partners in the risk set.

Accumulative advantage is reflected in the network *degree* and *experience* of both the attaching DBF and partner, which captures both the number of ties and years of experience with collaborations. For repeat ties, we also include *prior ties* and *prior experience* of the DBF-partner dyad. We also use *new partner*, an indicator of whether a partner has been in the network for less than a year, to capture the effect of being "a new kid on the block."

Homophily is assessed in several ways. We measure the *collaborative distance* between the alliance profiles of the DBF and the partner in the attaching dyad. For one-mode attachments, we measure the *age difference*, *size difference*, and *governance similarity* between the two DBFs forming the dyad, as well as measuring *co-location*. For two-mode attachments, we lack data to measure partner's *age* or *size* difference, and *governance similarity* does not apply to noncommercial organizations. We did measure homophily, however, between the attaching DBF and other DBFs in the partner's neighborhood, defined as the set of DBFs to which that partner has direct ties, with the variables *collaborative distance*, *age difference*, *size difference*, *governance similarity*, and *co-location*, all considered with respect to a partner's other affiliations. These "second order" measures account for the possibility that connections to partner organizations are conditioned by the prior experience of those partners with other DBFs.

Follow-the-trend is captured by the field-level variable *dominant trend* and the partner variable *dominant type*. Both measures reflect whether firms are engaging in activities that are comparable to those of others in the field. *Multiconnectivity* has two facets: cohesion and diversity. The former captures the extent to which firms are connected by multiple independent pathways, while the latter reflects whether firms are engaged in multiple types of activities. *Cohesion* is calculated using the maximum level of k-component for each firm and partner, measured separately as *firm cohesion* and *partner cohesion*, and jointly as *shared cohesion*, which is the maximum level for which both parties share a common k-component, if any.[20] For diversity, we use Blau's (1977) heterogeneity measure

---

[20] More precisely (White and Harary 2001, pp. 12–14), "The (*node-*) *connectivity* $\kappa(G)$ is defined as the smallest number of nodes that when removed from a graph G leave a disconnected subgraph or a single node. . . . A maximal connected subgraph of G with connectivity $k > 0$ is called a *k-component* of *G,* with synonyms *component* for 1-component, *bicomponent* for 2-component, *tricomponent* for 3-component, etc. A *cohesive block* of a graph G is a k-component of G where the associated value of

as an index of both the range of activities and multiple types of partners. We measure diversity in four ways: for a DBF (*firm tie diversity*), prospectively for a DBF (*prospective tie diversity*, i.e., the diversity of a firm's collaborative profile if the attachment were to occur), for a partner (*partner tie diversity*) and for the average of a partner's set of partners (*partner's partner tie diversity*). One might think of the last measure as an assessment of the heterogeneity of a DBF's partner's neighborhood.

The *control variables* include firm demographics measured at the time of attachment, including *age* (in years), *size* (number of employees), *governance* (whether publicly held), and *location*, measured by three-digit zip code region, or in the case of companies outside the United States, by the telephone prefix for nation and city. The key partner control is the variable *form*, which reflects the form of organization, and the key dyad control is *type*, which reflects the type of activity that is the focus of the collaboration. We also include three *time period* variables, which emerged from the discrete-time images as key inflection points in the pattern of affiliation, along with a linear time trend variable, *time line*. The descriptive statistics for all the variables are presented in appendix B, table B1.

STATISTICAL METHOD

Our challenge is to model a set of binary indicator variables, each of which can occur multiple times within a firm, and for each occurence we must update the risk set of alternative partners and the network measures. Consequently, our unit of analysis must be the attachment rather than either the individual firm or the dyad. Our choice of a statistical model for analyzing attachment bias is based on our unit of analysis, as well as empirical and theoretical considerations. Empirically, the design of our sample defined the population of DBFs and then identified all the partners engaged in alliances with them. Theoretically, our research question asks what mechanisms account for differential (as opposed to simple random) patterns of attachment. For these reasons, we use McFadden's estimator for multiprobability assessments, which is a variant of a conditional logit model that takes each event as a unit of analysis, which in our case are attachments, and distinguishes between a focal population and a set of alternatives (McFadden 1973, 1981; also see Maddala 1986; Ben-Akiva

connectivity defines the cohesion of the block." A graph G is *k-connected* if $\kappa(G) > k$, hence we use the term *multiconnected*. A fundamental theorem of graphs is that a multiconnected k-component is also equivalent to a maximal graph with k or more node-independent paths between every pair of its nodes, which adds significantly to the power of the concept of multiconnectivity as a measure of cohesion.

and Lerman 1989).[21] We set up the data so that the focal DBFs are the population and the partners in the risk set for an attachment are the alternatives, thus reflecting the one- and two-mode networks in our sample.

We conducted the analyses in three stages. First, we perform overall tests of the four hypotheses, on each type of attachment, by applying McFadden's estimator as follows. For each attachment, the probability of DBF i attaching to partner j, given that DBF i attaches at time t to some partner in the set $J_{i,t}$, is specified as a function of partner (X) and dyadic (W) variables:

$$\text{prob}\left(y_{ij,t} = 1 \Big| \sum_{J_{i,t}} y_{ij,t} = 1, \ X_{j,t}, \ W_{ij,t}\right) = \frac{\exp\left(\beta X_{j,t} + \gamma W_{ij,t}\right)}{\sum_{J_{i,t}} \exp\left(\beta X_{j,t} + \gamma W_{ij,t}\right)}.$$

Some of the partner characteristics are not defined for partners without ties prior to the attachment. We include these characteristics by interacting them with an indicator variable for whether the partner has any prior ties, and include the main effect of this indicator variable in the models. The main effect is always negative, showing that partners with no ties in the prior year are more likely to receive attachments.

Second, we explore the extent to which the mechanisms of differential attachment are contingent on various combinations of (1) characteristics of the focal, attaching DBF, (2) the form of a partner, (3) the type of activity involved in each case, and (4) combinations involving a partner's partners. We do so by interacting variables representing each of these categories with measures of the attachment bias mechanisms. For instance, to explore how the focal DBF's attributes (Z) may alter the "rules"

[21] Although McFadden's conditional logit is commonly known as the Discrete Choice model, this label describes a theoretical orientation and not an analytical approach. The model of multiprobability assessment is agnostic as to whether the process of attachment is calculative, a form of following the herd, conditioned by social structure, or random. The model is equally applicable to circumstances where DBFs are choosing partners, the partners select DBFs, the social structure of affiliation matches DBFs and partners, or any combination thereof. McFadden (2001) initially dubbed his estimator a conditional logit, but has since indicated he prefers the term multinomial logit. His model is more flexible than other conditional "fixed effects" estimators, and more general than standard multinomial estimators, converging to each with correct specification of dependent and independent variables. Hence, we refer to it more generally as a multiprobability model.

of attachment, where those rules are specified in terms of partner and dyadic variables (X,W), we estimate the following specification:

$$\text{prob}\left(y_{ij,t} = 1 \mid \sum_{J_{i,t}} y_{ij,t} = 1,\ X_{j,t},\ Z_{i,t},\ W_{ij,t}\right) =$$

$$\frac{\exp\left(\beta X_{j,t} + \gamma W_{ij,t} + \lambda Z_{i,t} X_{j,t} + \delta Z_{i,t} W_{ij,t}\right)}{\sum_{J_{i,t}} \exp\left(\beta X_{j,t} + \gamma W_{ij,t} + \lambda Z_{i,t} X_{j,t} + \delta Z_{i,t} W_{ij,t}\right)}.$$

Third, we examine how the sources of attachment bias shift over time by interacting our measures of the attachment mechanisms with (1) period effects, as indicated in table 3 and (2) a linear time trend. These measures give us purchase on how the logics of attachment may change as the overall structure of the network evolves.

We estimate conditional models of attachment, continually updating both the risk set of alternative partners and the network measures to lessen problems of observation dependence common in dyad-level analyses. Although we define DBFs as our focal population and partners as alternatives, both DBFs and partners are vying to connect with one another, and at any point in time a DBF or partner may have a finite capacity for connections.[22]

We estimate all models using Stata 8/SE. For the two-mode analyses, two features are notable. First, for two-mode new attachments, there are 5,087 events over our time period, with up to 1,600 alternative partners in any given year. Given the number of variables we tested, the size of the data set needed to analyze the full population of new two-mode attachments became cumbersome. Thus, we obtained a random sample of 1,500 two-mode new attachments to form the data on which estimates were obtained, and for each attachment we took a 25% random sample of alternative partners. We then repeated the random sampling process and reestimated the two-mode new attachment models 10 times as a test

---

[22] This limitation introduces a form of competitive interdependence in our observations. That is, assuming that organizations have some finite need or capacity for partners, not all firms may be able to attach to all partners they might otherwise have chosen. This interdependence is substantive, rather than statistical, in that it could, in theory, be specified and modeled. Such specification would involve many of the variables we have measured, such as prior degree and experience, but would also require some variables such as financial resources or managerial skill that are not easily available. This form of interdependence is, to varying degrees, present in many applications of McFadden's estimator. For instance, in the transportation studies (McFadden, 1984), each mode (car, bus, train) was vying for commuters, and not all individuals could choose the same mode, as doing so would crowd buses or jam highways. As a result, we do not believe that our results are compromised by this limitation. Further research into the statistical properties of McFadden's estimator in the face of such interdependence does, nevertheless, seem warranted.

of robustness. The results were nearly identical for every sample, and with no sample did the pattern of significance change; thus we present the initial results, noting that their significance levels would be much higher for the full two-mode sample.

In the two-mode analyses the comparison of organizations from different sectors (e.g., for-profit corporations, nonprofits, universities, government agencies, and venture capitalists) requires some comment. A nonprofit hospital and a venture capital firm are not substitutes for one another because they provide different resources. Yet, we have included them as alternatives in the risk set for the same attachments. From a practical standpoint, in this example the inclusion assumes that a focal DBF may have discretion as to whether its next alliance will be to obtain further funding or to seek clinical evidence of efficacy. We tested the robustness of this assumption on our findings in two ways. First, we stratified the two-mode risk sets by form, so that only organizations of the same form as the partner to which the attachment occurred were included in the risk set. Second, we included interactions of our other variables with a set of indicator variables for organizational form. Interestingly, while there was some variation in magnitude of coefficients, the pattern of direction and significance of predictor variables was constant across partner forms save for two exceptions: finance ties to venture capitalists and ties of any type to biomedical supply firms. These exceptions are not surprising, as venture capitalists are the most specialized and least multivocal partners (recall fig. 2a–2b, which illustrates that more than 95% of VC collaborations are financial.) Other biomedical firms provide equipment and research tools, thus they represent bilateral exchange relations, and do not participate in the development of new medical products in the same, mutual manner as do other participants. Based on these preliminary analyses, we collapsed the two-mode data across all organizational forms and included interactions for finance ties to venture capitalists and ties of any type to biomedical supply firms when exploring contingencies in the mechanisms for attachment bias.

RESULTS

We organize the presentation of results around our four hypothesized rules of attachment: accumulative advantage, homophily, follow-the-trend, and multiconnectivity. We utilize a number of variables to measure each mechanism. Thus, assessing the relative strength of a logic of attachment requires examination of an aggregate pattern of support or falsification. For our purposes, social mechanisms are more strongly supported when the bulk of the effects are in the hypothesized directions and those effects are

consistent across all four classes of ties. The full table of results is both elaborate and complicated; it is presented in appendix B. For the sake of exposition, we have excerpted the results that are the primary tests of each hypothesis into a series of four more manageable tables (see tables 4–7 below). To facilitate interpretation, we present the odds ratios, rather than the coefficients. The odds ratios are obtained from the coefficients by exponentiation. That is,

$$\text{odds ratio for } X_i = \exp(\beta_i).$$

The odds ratio gives the amount by which the odds of an attachment being biased toward a partner are multiplied for each unit increase in the level of the independent variable, X, for that partner (odds ratios less than one reflect negative effects).

Accumulative advantage.—The odds ratios for our tests of accumulative advantage, hypothesis 1, are presented in table 4. Note that we employ a narrow reading of accumulative advantage, focusing only on the structural feature of popularity bias. Odds ratios greater than one for the variables partner degree, partner experience, prior ties and prior experience would offer support for the accumulative advantage mechanism. For the variable new partner, hypothesis 1 would predict an odds ratio lower than 1. Overall, the evidence does not favor this hypothesis. Only the negative impact of partner experience is consistent across all four categories of ties. Whether the relevant experience resides with the partner or dyad, these findings indicate a preference for novelty. These results run sharply counter to an accumulative advantage argument. The pattern of results for the other measures of accumulative advantage varies by type of attachment.

For partner degree, we find support for accumulative advantage only for one-mode new ties. These linkages have a distinctive pattern, as they are biased to partners with more ties and less time in the network. This result implies that DBFs seek to balance novelty and visibility when they form an alliance with another DBF. For each additional tie, the probability that a new attachment will involve a partner with higher degree increases, on average, by 4%. The contingencies reported in appendix table B2 indicate that the increase is higher if the attaching DBF is younger and declines as the focal DBF matures. For each year of a partner's network experience, the differential likelihood of attachment decreases by 5%. These findings suggest that as DBFs age, the salience of partner degree decreases, and the importance of partner experience recedes as partners spend more time in the field. Both results emphasize that prospecting for new partners is common, particularly among older DBFs who opt for new partners at the expense of more visible ones. In sum, preferential attachment by degree operates only for new one-mode ties among DBFs.

TABLE 4
TEST OF ACCUMULATIVE ADVANTAGE: ODDS RATIOS FROM MCFADDEN'S MODEL

	HYPOTHESIS	ONE-MODE		TWO-MODE	
		New	Repeat	New	Repeat
Partner degree	>1	1.042**	.996	.917**	1.007
Partner experience ...	>1	.947**	.867**	.807**	.865**
Prior ties	>1	NA	1.389**	NA	1.262**
Prior experience	>1	NA	.742**	NA	.862**
New partner	<1	1.169	10.953**	1.662**	.788

NOTE.—Controls for homophily, follow-the-trend, and multi-connectivity were also included in these models; odds ratios for these mechanisms are presented in subsequent tables. All models include fixed effects for firm, year, and type, as well as main effects for form. The complete models are presented in app. B. NA indicates not applicable for new ties.

* $P < .05$, two-tailed significance.

** $P < .01$.

For all four classes of ties, partners that are more recent entrants are preferred.

Prior ties and prior experience offer dyad-level operationalizations of accumulative advantage that are relevant for repeat ties only (table 4). As with partner experience, the effects for prior experience refute the accumulative advantage hypothesis. A one-year increase in shared partnership experience lowers the probability of a repeat tie by 26% and 14% for one- and two-mode attachments, respectively. Prior ties at the dyad level do fit the hypothesis, however. An increase of one common tie ups the likelihood of a repeat connection by 39% and 26% for one- and two-mode attachments.

Finally, consider the new partner variable. Here an accumulative hypothesis predicts a negative effect. The results for repeat one-mode ties and new two-mode ties weigh against the hypothesis. A partner's arrival in the network for the first year produces a strong bias for both subsequent repeat one-mode and new two-mode attachments, increasing the differential attachment probability elevenfold and by 66%, respectively. Repeat one-mode affiliations reflect sponsored mobility, where a new DBF is assisted by another firm. New two-mode alliances are less likely to be pursued with an already popular partner than with a new arrival, suggesting these collaborations follow an exploratory logic. Table B2 shows a contingency for this preference for new two-mode partners to become particularly strong in the mid- to late 1990s. The fact that this trend becomes more pronounced over time suggests that rather than being a consequence of a young field, the importance of novelty in partner selection is an enduring feature of an industry characterized by rapid technical advance.

In our dance hall language, when a DBF finds a young (i.e., often newly

entered) DBF dance partner that has created a stir on the dance floor, the two partners sustain the relationship. New two-mode partners are also preferred if they are new to the dance floor, but the alliance is less likely to be renewed. These results suggest that well-positioned, veteran DBFs spot up-and-coming newcomers and escort them into the network. In some cases, these new firms are spinoffs from more established DBFs, which may account for the very strong support for repeat ties with new-comers. In sum, we find scant support for the hypothesis that repeat attachments are influenced by the accumulative advantage of the partner.

Homophily. Table 5 presents the odds ratios for our tests of homophily (hypothesis 2). We find evidence for homophily if the odds ratios exceed 1 for co-location and governance similarity and if the odds ratios are lower than 1 for the measures of distance and difference. The results for homophily vary both within and across classes of attachment. For the similarity measures (co-location and public/private governance), four of the odds ratios are significantly supportive, while four run contrary to the hypothesis. For the distance and difference measures, however, only three of the odds ratios offer strong support for the hypothesis, with 17 contrary. Overall, homophily holds more strongly for new attachments. Homophily is less prevalent for one-mode repeat ties, and absent in repeat two-mode attachments.

In the one-mode case, similarity looms large with regard to geographic co-location, but not for organizations of similar governance, age and size. Firms located close to one another are twice as likely to collaborate, but similarity in organizational age and size decrease the odds of collaboration. Thus, DBFs are more likely to attach to nearby DBFs that differ in age and size. This finding echoes the previous accumulative advantage result, with larger, veteran DBFs collaborating with smaller newcomers. The contingencies in table B2 show that the preference for age and size diversity moderates somewhat in later years. There is weak evidence for an attachment preference based on similar collaborative profiles and geographic propinquity in the partner's neighborhood for new one-mode ties. For repeat one-mode attachments, diversity rather than similarity is the rule with respect to collaborative distance.

Homophily receives some support in new two-mode attachments, where there is a tendency to involve partners whose neighborhood includes other DBFs with collaborative profiles similar to the focal DBF. A unit decrease in the average Euclidean distance between the collaborative profiles of the attaching DBF and other DBFs in the partner's neighborhood increases the likelihood of attachment by 8%. But the pull of diversity is also present, particularly with respect to a partner's own collaborative profile. Thus, the logic of homophily is attenuated by a preference for diversity in new two-mode partners.

1178

TABLE 5
TEST OF HOMOPHILY: ODDS RATIOS FROM MCFADDEN'S MODEL

		ONE-MODE		TWO-MODE	
	HYPOTHESIS	New	Repeat	New	Repeat
Co-location	>1	1.95**	2.091*	NA	NA
Partner's partner co-location	>1	1.294$^+$	.460	1.098	.129**
Governance similarity	>1	.864	2.037$^+$	NA	NA
Collaborative distance	<1	.973$^+$	1.053$^+$	1.196**	1.017$^+$
Partner's partner collaborative distance	<1	.997	1.689**	.920**	1.319**
Size difference	<1	1.018**	1.001	NA	NA
Partner's partner size difference	<1	.991	.995**	1.003*	.991
Age difference	<1	1.036**	1.137**	NA	NA
Partner's partner age difference	<1	.980	1.063	1.005	1.161**

NOTE.—Controls for accumulative advantage, follow-the-trend, and multi-connectivity were also included in these models; odds ratios for these mechanisms are presented in prior and subsequent tables. All models include fixed effects for firm, year, and type, as well as main effects for form. The complete models are presented in app. B. Size is divided by 100. NA indicates that the two-model data were not available for computation.

$^+$ $P < .10$, two-tailed significance.
* $P < .05$.
** $P < .01$.

In the tests of homophily as an attachment rule for repeat two-mode attachments, there is a distinct pattern of dissimilarity in terms of demography, location, and collaborative profile between the attaching DBF and the other DBFs attached to a prospective partner. For example, an average difference in age of one year increases the probability of collaboration by 16%. Repeat two-mode attachments are 87% less likely to occur when the focal DBF is co-located with a partner's other DBF allies. As was the case with accumulative advantage, two-mode tests of homophily also provide strong support for an attachment process driven by novelty and diversity.

Follow-the-trend.—The odds ratios presented in table 6 test for an attachment bias based on a logic of appropriateness. Hypothesis 3 is supported if the odds ratios exceed 1 for both the dominant trend of the field and the dominant type of the partner. Both variables are measured in percentage point units. To facilitate interpretation, we use a change of 10% to assess the effects of dominant type and trend For dominant type, the odds ratios paint a consistent picture supporting hypothesis 3 for all classes of alliances. For dominant trend of the field, new and repeat attachments offer divergent results, with repeat ties following a logic counter to the overall trend.

New one-mode ties manifest a strong tendency for DBFs to follow both

TABLE 6
TEST OF FOLLOW-THE-TREND: ODDS RATIOS FROM MCFADDEN'S MODEL

	HYPOTHESIS	ONE-MODE		TWO-MODE	
		New	Repeat	New	Repeat
Dominant type	>1	1.027**	1.039**	1.025**	1.029**
Dominant trend ...	>1	1.251**	.805*	1.064**	.884**

NOTE.—Controls for accumulative advantage, homophily, and multi-connectivity were also included in these models; odds ratios for these mechanisms are presented in prior and subsequent tables. All models include fixed effects for firm, year, and type, as well as main effects for form. The complete models are presented in app. B.

* $P < .05$, two-tailed significance.
** $P < .01$.

the dominant trend and dominant type of activity. A 10 percentage-point rise in the dominant type measure increases the differential attachment probability by 30%. The same rise in the dominant trend measure increases the differential attachment probability ninefold. The choice to follow the trend is also apparent in new two-mode ties, both in terms of the dominant trend in the field and the partner's dominant type. The conditional attachment probability is boosted by 30% and 85% for each 10% increase in the type and trend measures. In both classes of repeat ties, following the trend is apparent for the dominant type of partner, but there is now a strong preference to buck the field's overall trend. With each 10% increase in the trend measure, repeat one-mode attachments are nearly 90% less likely to occur in the same category, and repeat two-mode attachments are 70% less likely.

The follow-the-trend results suggest that "first dances" are conservative, matching the pattern on the floor, which signals that norms of propriety condition the logic of attachment. The reverse holds, however, for repeat ties where deeper linkages between already connected dyads offer the occasion to flout the dominant logic. This reinforces findings from tests of homophily where diversity drove repeat ties. The follow-the-trend results are the first to show a uniform pattern (all significant at $P < .05$; consistent with hypothesis 3) across one- and two-mode ties.

Multiconnectivity.—Table 7 reports tests for the multiconnectivity hypothesis. Hypothesis 4 predicts positive (> 1) effects for all variables. There is consistent support for the cohesion and partner tie diversity measures across all four classes of ties. Six out of the eight possible effects of partner and shared cohesion are significant ($P < .01$) in the predicted direction. Cohesion can be regarded as a strategic window where partners evaluate a radar screen of possible future partners, as opposed to sampling from the entire disparate field. The strength of shared cohesion in shaping two-mode attachments illustrates the value of linkages to multiple affiliates

TABLE 7
TEST OF MULTICONNECTIVITY: ODDS RATIOS FROM MCFADDEN'S MODEL

	HYPOTHESIS	ONE-MODE		TWO-MODE	
		New	Repeat	New	Repeat
Partner cohesion	>1	1.432**	2.589**	1.667**	1.080
Shared cohesion	>1	1.058*	1.103	5.272**	1.905**
Partner tie diversity	>1	1.025**	1.016	1.037**	1.020**
Prospective tie diversity ...	>1	.972	.738**	.883**	.952**
Partner's partner tie diversity	>1	1.042	.944**	.986**	.959**

NOTE.—Controls for accumulative advantage, homophily, and follow-the-trend were also included in these models; odds ratios for these mechanisms are presented in prior tables. All models include fixed effects for firm, year, and type, as well as main effects for form. The complete models are presented in app. B.
* $P < .05$, two-tailed significance.
** $P < .01$.

as a means to validate information. Cohesion entails more than access to information, however. Resources, skills, access to personnel, and a host of related benefits are also obtained through partners. The results indicate that as DBFs scan for potential collaborators, a partner's diversity of ties is a valuable marker of resources and information.

Both partner and shared cohesion influence new one-mode attachments. For each additional level of partner connectivity, the differential probability of a new link jumps by 43%. Similarly, each new independent pathway connecting a dyad (the increment in shared cohesion) increases the differential attachment odds by 6%. In addition, new one-mode affiliations evince a preference for partners that have more varied ties. A 10% increase in partner's tie diversity raises the differential probability of attachment by 30%.

For repeat one-mode attachments, partner cohesion continues to provide a positive influence, increasing the conditional attachment odds 2.6 times for each additional level of connectivity. Shared cohesion and partner tie diversity, from which the DBF already benefits, lose some of their appeal in repeat tie formation. Further, the coefficients on prospective tie diversity and partner's partner tie diversity turn negative. Relatively typical increases of 1% and 10%, respectively, in these tie diversity measures lower the differential probability of affiliation by 25% and 45% (see app. B, table B1, for means and SDs for these measures). Thus, repeat one-mode attachments serve to reinforce rather than expand prior collaborative activities. These results suggest a different evaluative metric for repeat ties, in which the parties forego increased diversity in favor of more cohesive (and possibly long-term, and deeper) relations. We lack detailed data on the content of specific relations, but one might consider that these

participants have come to trust one another more and now are bound together in a common fate with respect to a new medical product, or that the information being exchanged is tacit or sticky, and thus more easily shared in a strong-tie relationship.

Cohesion and partner's tie diversity positively influence attachment for new two-mode linkages. The effects of cohesion are quite strong: each increase in the level of cohesion of the partner bumps up the probability that a collaboration will form with this partner by 66%, while each additional path of connectivity for the dyad increases the differential probability of attachment more than fivefold. There is a bias, however, against partners whose allies have more diverse profiles of collaboration. This pattern seems to indicate that DBFs favor attachments to partners that are better positioned with more diverse collaborations, while, ignoring contingencies, partners prefer attaching to DBFs with more specialized collaborative profiles. Hence, comparative judgments about partner diversity reverberate through the network. The negative coefficient of prospective tie diversity might signal a preference for reinforcing the DBF's prior collaborative profile, rather than filling out the DBF's "dance card" with different forms of partners for diverse types of activities. A percentage point increase in the index of prospective diversity decreases the odds of attachment by 12%.

The results for multiconnectivity in repeat two-mode ties demonstrate that shared cohesion of the dyad and the partner's tie diversity are both sources of positive attachment bias. The impact of an increase of one pathway of connectivity between a DBF and a partner enhances the differential probability of attachment by 90%. A 10% increase in a partner's tie diversity boosts the probability of a repeat two-mode tie to the partner by 20%. The prospective tie diversity of a potential alliance and the partner tie diversity are both sources of negative bias, by 4%–5% for each percentage point increase in the respective index.

The odds ratios for prospective diversity run counter to hypothesis 4 for three of the four classes of attachments (all but new one-mode ties). These results may indicate that diversity does not represent a goal in and of itself. When interaction effects are added (see app. B, tables B2 and B3, cols. 2 and 3, for one- and two- mode results, respectively), the main effects of prospective tie diversity flip the sign for three of the four classes of attachments (all but repeat one-mode ties).[23] Hence, the search for

[23] For this pattern to obtain, it is sufficient to add interaction terms for cohesion (shared cohesion × firm age, firm cohesion × partner cohesion, firm cohesion × prospective diversity). The change in signs is stable with period and timeline effects added to the conditional logit model. Change in the signs of the coefficients do not occur in appendix tables B2 and B3 for any of the other variables listed in table 7, with one minor exception. The contingency of combined firm cohesion × prospective diversity, for

diversity is most operative among firms at low levels of structural co-hesion. At higher levels of cohesion, the bias toward diversity recedes in favor of a preference for partnerships with new entrants. Thus, including interactions between diversity and cohesion measures indicates a modi-fication of hypothesis 4. Those DBFs with higher cohesion, having ob-tained diversity through indirect k-connectivity, find a diminished need to increase diversity in new or repeat partners. Cohesion and prospective tie diversity represent countervailing forces at higher levels. If cohesion was the most powerful force, the radar screen of possible partners would become restricted and then ossify. The exploration of diversity precludes such lock in.

The process of searching for partners is both dynamic and recursive. Preferential attachment operates through the attractiveness of shared and partner cohesion, with firms apparently moving up a ladder in terms of the cohesiveness of their networks. At the lower rungs of the cohesion ladder, there is a preference for expanding diversity by linking to well-connected partners. At the higher rungs of the cohesion ladder, firms may forgo cohesion, opting to ally with recent entrants to the field. This re-lationship is suggestive of a systemic pumping action, with the most con-nected members pushing out in a diastolic search to pull in newcomers, and those less connected being pulled inward, in a systolic action, to attach to those with more cohesive linkages. The pumping process operates up-wardly from the bottom, level by level.

We illustrate this dynamic process with two figures. We return to Pajek visualizations, but now we use them to illustrate the econometric results with respect to the effects of cohesion and diversity. Figure 10 contains all members of the main component in 1997, with node size scaled to network degree and color reflecting levels of connectivity. The red nodes are members of the five component, the cluster with at least five node-independent paths between all pairs of members, which is a remarkably cohesive community. The red nodes, not surprisingly, tend to be bunched at the center of the figure. Green nodes are members of the three and four components, with blue nodes in the one and two components. Note the tiny blue nodes at the center of the figure. These are new entrants to the network, with few connections (reflected in their small size which represents low degree), but these nodes have "high quality" links to other well-connected organizations. This pulling in of newcomers reflects the process of sponsored mobility we referred to above. Figure 11 portrays new ties in 1998 among members of the main component. There were 1,121 new ties in 1998, of which 1,074 were forged by members of the

---

example, predicts lower attachment, but shifts the prediction for prospective diversity to higher attachment.

Node Key

Size scaled to degree

Red = five component

Green = 3 or 4 component

Blue = 1 or 2 component
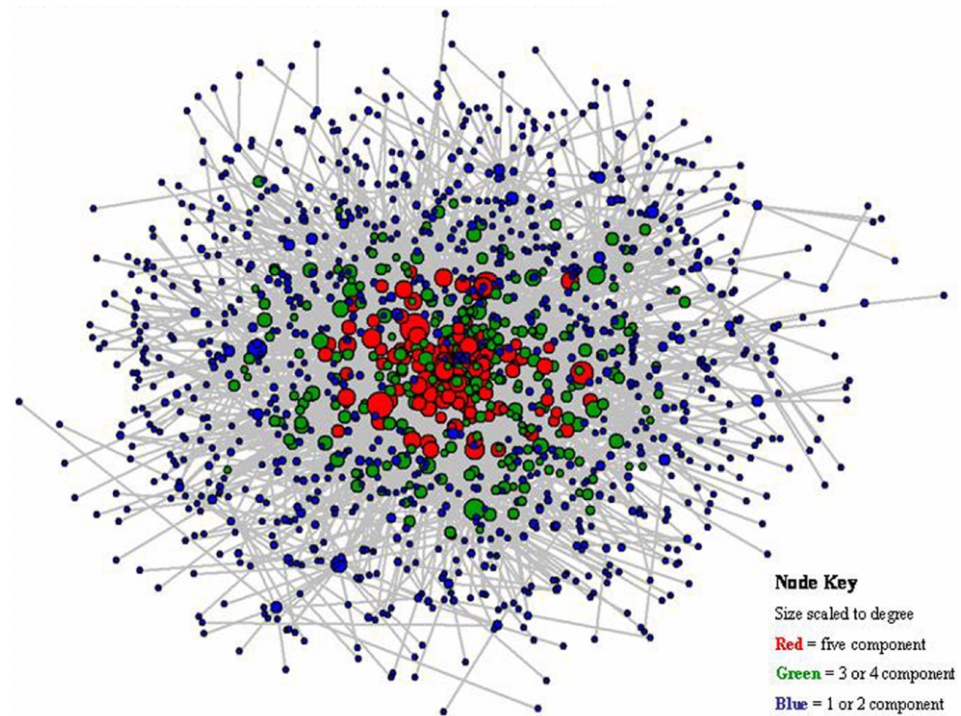
Fig. 10.—1997 all ties, main component, nodes colored by cohesion

Key

	# Nodes	# Ties
Red	112	634
Green	160	304
Blue	477	136
N	749	1074

Size scaled to 1997 degree

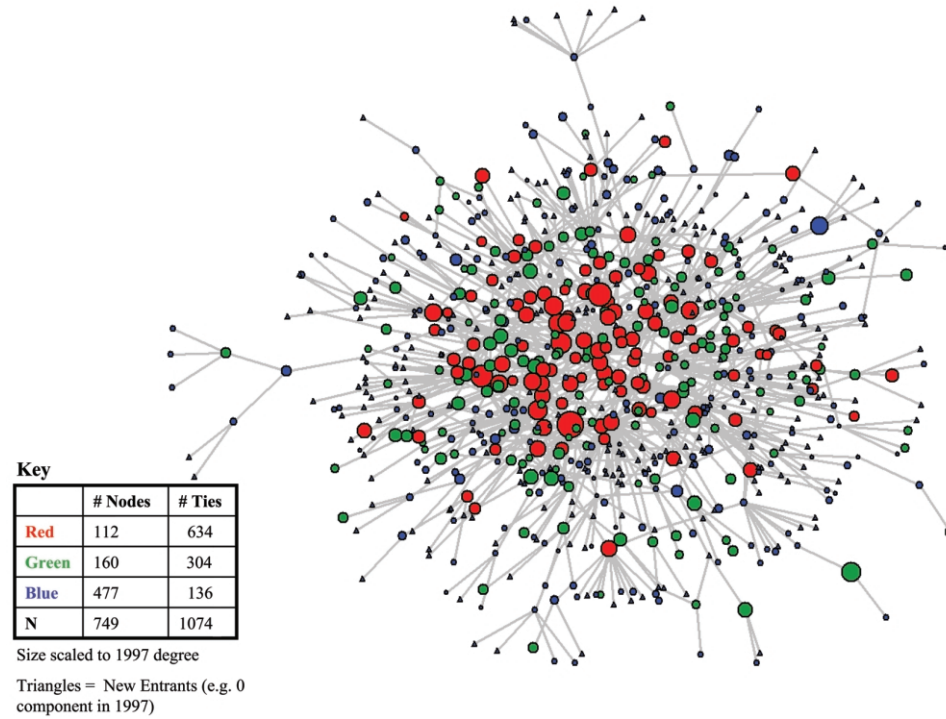Triangles = New Entrants (e.g. 0 component in 1997)

Fig. 11.—1998 new ties, main component, colored by cohesion

main component. Again, node size is scaled to degree in 1997 and color reflects levels of connectivity. There are but 112 organizations in the most cohesive component (red), but these organizations are responsible for 59% of all ties initiated in 1998. Note also that the red nodes are no longer exclusively clustered in the center but are dispersed. Both the physical placement of the red and green nodes and their links to a large number of small blue triangles (representing new entrants to the network in 1998) reflect the generative role of the most-connected organizations. The exploratory role of the "elite" red nodes may well account for the strong confirmation of the follow-the-trend hypothesis. Taken together, the two figures portray the processes of pulling less connected organizations into the network and the very active search undertaken by members of the most cohesive component. (The network images in the online version of this article include the full sequence of figures for this search process.)

The odds ratios for partner's partner tie diversity also appear to run counter to hypothesis 4 for three of the four types of attachments. A plausible account of these results is that if a potential partner organization has collaborations with organizations that have even more diversity than the ostensible target partner, that relationship may be foregone in favor of a direct tie with the more "attractive," distant organization (i.e., a partner's partner, or a friend of a friend). Such an interpretation is contrary to hypothesis 4, but does suggest a persistence of the preference for diversity. Note too that such calculations about diversity depend on cohesion to facilitate recognition of the types of affiliations that are held by various participants. Overall, the results of table 7 suggest that structural cohesion matters for the formation of multiple classes of ties. This relationship has two components. First, DBFs attach to organizations that are also members of their own component, thereby deepening multiconnectivity in those clusters. Second, biotechnology firms at the top of the cohesion hierarchy also reach out to partners located in more distant components.

There are a number of contingencies that add to the overall pattern of multiconnectivity. First, the older the focal DBF, the less the one-mode preference for shared cohesion surrounding the dyad. Second, among DBFs in low-cohesion networks, attachment is tipped toward partners that increase prospective diversity. The higher the firm cohesion, the greater the attachment bias from the partner's cohesion, but the less the bias from prospective diversity. This latter finding suggests that the pattern of new attachments among those DBFs in high cohesion networks is one of seeking out partners with a diverse portfolio of linkages rather than allying with other biotechs sequentially to complete the stages of the drug development process. Put differently, high cohesion firms are not trying to balance their dance card of affiliations by having a set of partners for each functional activity. In contrast, low cohesion DBFs are trying to

fill out their dance cards and this tendency looms largest in later years. Third, the effect of partner cohesion wanes over time, while that of partner diversity waxes, indicating that the influence of multiconnectivity is shifting from cohesion to diversity. This shift suggests a change in the topology of the network, from a power-law distribution toward a more exponential distribution, with the latter suggesting more distant and random search for new partners.

The results for the multiconnectivity hypotheses are generally positive, with nine out of twelve relationships supported significantly for partner cohesion, shared cohesion, and partner tie diversity. The results turn negative for prospective tie diversity and partner's partner tie diversity. Upon examining the contingencies, however, we were led to consider that these results indicate a more dynamic, reciprocal view of cohesion and diversity, in which low levels of cohesion trigger a preference for more connected partners and high levels of cohesion permit exploration in the form of search for new alliance prospects.

CONCLUSIONS

The tripartite set of analyses we have presented highlight a network structure in which multiconnectivity expands as the cast of participants increases, and, in turn, diversity becomes more important with time. At this point in the evolution of the field, a combinatorial or multivocal logic has taken root. Neither money nor market power—not even the sheer force of novel ideas—dominates the field. Rather, those organizations with diverse portfolios of well-connected collaborators are found in the most cohesive, central positions and have the largest hand in shaping the evolution of the field. This is a field in which the shadow of the future is long, as much remains to be learned about the functional aspects of molecular biology and genomics. The density of the network and the open scientific trajectory combine to enhance the importance of participants' reputation. The pattern of cross-cutting collaborations often results in a partner on one project being a rival on another. The frequent rewiring of attachments means that participants have to learn how to exit from relationships gracefully so as not to damage future prospects for affiliation. In a system where external sources of knowledge and resources are widely differentiated, a preference for diversity and affiliation with multiconnected partners has mobilizing consequences.

The three sets of analyses complement one another and provide insight into the dynamics of affiliation in this field. In the simplest terms, over the period 1988–99, the primary activities for which organizations engaged in collaboration shifted from commercialization to finance and R&D.

These two activities reinforce one another, as financial support fuels expensive research efforts and research progress attracts venture capital. In turn, as a smaller set of more diverse participants became skilled at pursuing multiple tasks, a more complex strategy emerged in which some of the most connected participants engaged in all four types of activity with a diverse set of partners. The spring-embedded graph drawings show a change in both the composition of participants and their activities over time. To return to the analogy of the dance hall, both the music and the dancers shift over time. The early dances are dominated by large multinationals and first-generation biotech firms, collaborating to the tune of commercialization of the lead products of the younger firms, with the bigger pharmaceutical firms garnering the lion's share of the revenues. Research progress, strongly supported by the stable presence of the NIH, attracts new participants to the dance and also enables incumbent biotech firms to deepen their product development pipelines and become less tethered to the giant pharmaceutical companies. Research progress attracts venture capital funding, which is excited by new possibilities. Thus, the music changes from commercialization to R&D partnerships and venture financing. Many of the larger multinationals are pushed to the periphery and some drop out of the network altogether.

Recall that this dynamic field is growing swiftly over this period, adding new entrants as progress is made along a broad scientific frontier in which no single organization can develop a full range of scientific, managerial, and organizational skills. A diverse set of organizations become cohesive, central players at the dance—a handful of research universities, key government agencies, a few elite research hospitals, a larger number of biotech firms, and a core of multinational giants. The multinationals go through an era of consolidation, with mergers and acquisitions commonplace, and their numbers are reduced considerably. Those that survive this period emerge as multivocal agents, no longer dancing only to the commercialization beat, but capable of executing R&D partnerships and startup financing. Venture capital remains at the dance, a critical specialist, but a somewhat fickle one that is easily attracted to other scenes, such as information and communications technologies. A small number of first-generation biotech firms grow into reasonably large organizations in their own right, and they do so by having learned how to engage in all forms of collaboration with a heterogeneous set of partners. Here we see strong affirmation of the finding from previous research that the organizations did not become "players" by virtue of being larger, but grew larger precisely because they became central players in the field (Powell et al. 1996).

When we consider the rules that guide attachment in this field, we observe that no single rule dominates over all time periods. Aside from the venture capital firms and biomedical supply companies, all of the

participants are engaging in multiple activities and these different pursuits are shaped by divergent rules of attachment. The multiprobability models tease out the mechanisms that undergird preferential attachment. These results lend the weakest support to the accumulative advantage or rich-get-richer hypothesis. The visualizations also highlight the importance of new connections and illustrate that growth in the network is often spurred either by new entrants or incumbents embarking on new activities. But the Pajek-generated figures suggest something the logit models cannot reveal—a very small core of perhaps one or two dozen organizations are routinely placed in the center, and their node size grows somewhat over the period. These organizations are the ones with the most diverse array of collaborations, suggesting that for this handful of central players, accumulative advantage is fueled by multiconnectivity. Thus, we see a combination of rich-get-richer and multiconnectivity at work for these core participants, and they may well set the pace for the dominant trend in the field. The new tie visualizations do not reveal large nodes at the center, growing even larger as the field matures. This suggests an open elite, accessible to novelty as the field expands.

The visualizations are somewhat limited in their ability to capture the mechanisms of homophily, and here the multiprobability models are much more effective at showing a preference for collaborating with organizations that are geographically proximate. If the unit of analysis is the dyad alone, there is little homophily operating. But when we consider second-order networks and ask about similarity of a class of possible partners compared to the chooser's existing allies, there are tendencies toward homophily.

Both the visualizations and the multiprobability models illustrate that, as a strategy for attachment spreads, it is quickly adopted. Whether this is through learning or imitation, or an indissoluble mixture of the two, we cannot say. But it is neither costless nor easy to follow the trend. Lots of ties are broken, and there is considerable exit from the field as well as the main component. So even if imitation is common, replication is not easily accomplished.

The most fundamental attachment biases are to multiconnectivity and diversity—that is, to aligning with varied partners who are more broadly linked or to new entrants who are sponsored by nodes that are well placed. This rule is robust in that it created a field with multiple, node-independent pathways that pulled in promising newcomers, while pushing out incumbents that failed to keep pace. This logic of attachment dominated in a period of overall expansion, however. One cannot assume these processes hold in all environments.

Speculation about whether this network will consolidate, and perhaps become an obstacle to rather than a catalyst of innovation, generates much

discussion among industry analysts and members of the scientific community. Clearly as long as the technological trajectory continues to generate new discoveries and opportunities, expansion is possible. Unlike some technologically advanced fields such as computing or telecommunications, there are few advantages for users in terms of common standards or brand recognition. As yet, no general purpose technology has emerged that would speed consolidation. A prolonged inability to raise money, due to declining research budgets or inhospitable public equity markets, would threaten undercapitalized firms and might spur consolidation. The diversity of institutional forms—public, private, and nonprofit—that are active in the field are located in different selection environments. This diversity offers some protection against unfavorable economic conditions. Preference for cohesion and diversity may well be effective search mechanisms for multiple, alternative solutions to problems in a field where the progress of human know how has been highly uneven. Network cohesion and diversity obviously have their limits, though these are not topics that have been well researched. We have shown in related work that there are declining returns to network experience (Powell et al. 1999). Moreover, repeated shocks to any system—whether in the form of a prolonged downturn in the business cycle, rapid inflation in costs such as in health care, or highly publicized product failures or corporate scandals—can destabilize it and result in a tip in the rules of affiliation and the resulting combinatorial possibilities. We are adding data now on the period 2000–2003, so we will soon see how stable both the rules of attachment are and the structure of the network is in the face of a period of economic uncertainty and decline.

Clearly, some aspects of the life sciences are rather idiosyncratic. There are a wide range of diverse forms of organizations that exert influence on the development of the field. In many other technologically advanced industries, universities were critical in early stage discovery efforts, but as the technology matured, the importance of basic science receded. In biotech, universities continue to be consequential, and career mobility back *and* forth between university and industry is now commonplace (Owen-Smith and Powell 2001*a*, 2004). While the institutional demography of the life sciences may be unusual, the rapid pace of development and the wide dispersion of centers of knowledge are more typical of other high-tech fields. The key story, in our view, is less the issue of the nature and distribution of resources and more how these institutional features promoted dense webs of connection that, once in place, influenced both subsequent decisions and the trajectory of the field. With detailed longitudinal data, we show how the topology of a network emerged, and generated novelty in an institutional system that has many conservative elements. The logics of attachment strongly shaped how both new infor-

mation and entrants were integrated into the field. The coevolution of science and commerce was marked by potent micro-macro linkages that altered the global properties of the field. We have demonstrated how a preferential bias for collaborating with actors that are either more multiply or differently linked reshaped the landscape of biotechnology.

APPENDIX A

Network Visualization in Pajek

Pajek, a freeware program for the analysis of large social networks, offers new opportunities for generating meaningful and replicable visualizations of complex network data.[24] This appendix offers an overview of the benefits and limitations of the software, while providing a more detailed discussion of the steps used to develop the images we presented.

Pajek's strengths and limitations both arise from its emphasis on visualizing and manipulating very large networks.[25] The most obvious benefit is the significant simplification of the analysis of large-scale network structures. As the algorithms are optimized for speed, computationally simple but powerful structural and ego network measures are also implemented.[26] More important for our purposes than such descriptive measures are Pajek's capacities to generate replicable images of complex networks. When turned to the analysis of discrete time "pictures" of the evolution of a network, Pajek offers the best approximation of dynamic visualization currently available.

Visualization.—Pajek includes a set of network drawing algorithms based on both graph theoretic conceptions of distance in a network and the physical theory of random fields (Guyon 1994). These minimum energy or "spring-embedded" network-drawing algorithms permit a robust representation of social network data in two-dimensional Euclidean space.[27]

---

[24] Pajek was developed by Vladimir Batagelj and Andrej Mrvar and is available online at http://vlado.fmf.uni-lj.si/pub/networks/pajek/. Pajek has been used in a number of disciplines to represent complex network data (Albert et al. 1999; Batagelj and Mrvar 2000; Moody 2001, 2004; Owen-Smith and Powell 2004; Owen-Smith et al. 2002; White and Harary 2001). We thank Andrej Mrvar for comments on this appendix.

[25] The program will handle networks of up to 1 million nodes. Due to limitations imposed by the computational capacities of our machines, we have never analyzed a network larger than 250,000 nodes.

[26] Consider simple centrality measures. Pajek calculates degree, closeness and betweenness measures for large nets, but does not implement more complex measures such as information or power centrality.

[27] One algorithm also offers the option to visualize a network as a three dimensional sphere. While such representations are valuable for some analyses, they raise difficulties for two dimensional presentation. Thus, we emphasize two-dimensional optimizations here.

We draw first on the Fruchterman-Reingold algorithm (FR; Fruchterman and Reingold 1991), which optimizes network images without reference to the graph theoretic distance among nodes, to develop initial positions for all organizations (connected and unconnected). We then turn to a second algorithm, the Kamada-Kawai (KK; see Kamada and Kawai 1989) to reposition the connected nodes in the network. Where the FR algorithm positions all nodes by analogy to a force-directed physical system, the KK algorithm locates connected nodes adjacent to one another and makes Euclidean distances among nodes proportional to graph theoretic distances. Put differently, the KK algorithm visually represents a system where the distance between nodes is a function of the shortest network path between them. Taken together, these two algorithms generate substantively interpretable visual representations of networks, which place isolated organizations on the periphery of the image, while capturing the pattern and density of collaborative activity and reflecting the extent to which these linkages generate meaningful clusters of organizations.[28]

We find that the two optimization procedures are most effective when used sequentially. Such a strategy takes advantage of the best features of both algorithms. Fruchterman-Reingold's capacity to reposition isolated nodes and to separate densely connected clusters from one another provides a useful first step for visualization.[29] As it is based on a mathematical analogy to a physical system, the FR algorithm positions nodes without reference to the graph theoretic properties of networks that underlie most common network methods. One consequence of this style of optimization is a tendency to visually overlay unconnected nodes that share ties to a common partner. The FR algorithm, for instance, does not effectively represent the "star" configurations that are commonplace in growing network structures. Because the KK algorithm takes graph theoretic distances into account, visualizations generated using it are potentially of greater substantive interest to social network theorists. We generate the images presented in this paper using three one-minute-long optimizations. An FR optimization from random starts is used first and taken as a starting point for a pair of KK optimizations.

A lexicon for Pajek visualizations.—Pajek's visual flexibility offers several options to convey complex information through images. Consider our figures, which draw upon a simple "lexicon" of four parameters: position,

[28] While the repelling forces of nodes are determined by a constant (though manipulable) factor, the attractive strength of ties can vary with the observed value of ties. In this paper, our images are created with constant tie strengths.

[29] As implemented in Pajek, the Kamada-Kawai algorithm (KK) only alters the position of connected nodes as network distances among unconnected portions of a graph are undefined. By the same token, KK tends to overlap unconnected components, making the visual analysis of network evolution difficult.

color, size, and shape. We use position (the outcome of FR and KK optimizations) to discuss the clustering of organizations in the macronetwork. We use color to distinguish among forms of organizations and the types of activities they are involved in. We turn to node size, scaled to reflect standardized network degree, to communicate variations in connectedness across organizational forms and to suggest the effect of prior network degree on the propensity to form new ties. In principle, node size can be scaled to reflect variation on any real number variable and need not be limited to information generated by Pajek. For coordinating statistical analyses with the visualization, it is useful to scale node size to reflect differential levels of network measures of node attributes that provide explanatory purchase.

In figures 5, 7, and 9, shape expresses the distinction between incumbent (circle) and entrant (triangle) organizations, while maintaining the same color variations across organizational form. The shape distinction, then, along with tie and node color and node position, visually suggests differences in the attachment profiles of entrants and incumbents across organizational forms. Taken together, these four parameters offer a wide range of possibilities for exploring data and conveying complex results in network visualizations.

Limitations.—Despite the many novel and useful features of Pajek, there are also limitations. Pajek cannot perform most statistical tests of significance, and cannot be used to calculate some commonly used network measures. More importantly, features of the KK and FR optimization procedures bear upon the interpretation of node position, the comparability, and replicability of images. The key difficulty lies in the algorithm's probabilistic optimization procedures. The danger of such procedures is that they may, in repeated trials, converge on different local minima rather than finding a global minimum energy by which to position nodes. As a result the position of extreme outlying nodes can shift from optimization to optimization. This variability occurs most commonly in the location of isolates and nodes with single connections.[30] The relative position of nodes toward the center of a given network are vastly more stable and reliable. Likewise, the relative position of different components as positioned in the FR algorithm may vary.

This limitation has several substantive implications. First, standards for the robustness of optimizations cannot be couched in terms of the exact Euclidean position of nodes. We rely instead on the aggregate features of the network and the position of core organizations (note, e.g., our

[30] To test this variation we optimized the same network multiple (30) times from random starting points and took the mean and variance of node coordinates, finding the least variation in the position of well-connected (high degree) nodes.

use of the NIH, a very well connected node, as a touchstone in the pull out presented in fig. 4). More important, the position of nodes cannot be understood in absolute terms. Instead node position must be interpreted relative to the position of other nodes in the network. The emphasis on relative position does make it meaningful to consider the coincidence of nodes in close proximity across images.[31]

---

[31] It is possible to generate more strictly comparable images by fixing the position of nodes across visualizations. Nevertheless, choosing an analytically appropriate constant position raises a new set of issues and such a strategy limits the package's visual flexibility by fixing one of the four parameters that can convey information.

APPENDIX B

<div align="center">

TABLE B1

DESCRIPTIVE STATISTICS

</div>

| | ONE-MODE ATTACHMENTS | | | | TWO-MODE ATTACHMENTS | | | |
| | New | | Repeat | | New | | Repeat | |
VARIABLE	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Accum. advantage:								
Firm degree	18.36	14.66	41.10	18.69	14.46	10.80	23.23	15.79
Partner degree	7.01	7.94	17.05	13.93	.99	2.49	7.41	11.20
Firm experience	7.80	4.38	12.24	3.98	6.22	3.76	9.36	3.96
New partner	.57	.50	.10	.30	.37	.48	.06	.24
Partner experience	5.69	4.49	8.57	4.05	3.36	3.51	5.62	3.26
Prior ties			6.10	4.18			5.81	4.33
Prior experience			4.84	2.79			4.95	3.56
Homophily:								
Collab. distance	8.74	5.87	15.72	7.05	8.83	7.47	14.80	11.07
Age difference	6.32	5.23	6.99	5.21				
Size difference	581	1,215	2,437	2,100				
Gov. similarity	.51	.50	.77	.42				
Co-location	.11	.32	.27	.44				
Partner's partner:								
Collab. distance	4.38	6.58	9.27	6.95	4.47	6.45	6.92	6.14
Age difference	2.49	4.39	3.76	4.15	2.46	4.02	2.59	3.50
Size difference	407	1,131	1,425	1,794	152	631	399	1,175
Gov. similarity	.86	.31	.85	.23	.82	.37	.83	.24
Co-location	.21	.38	.56	.38	.24	.41	.59	.38
Follow-the-trend:								
Dominant trend	.04	.02	.04	.02	.07	.09	.10	.11
Dominant type	.20	.27	.27	.22	.13	.32	.17	.55
Multiconnectivity:								
Firm cohesion	4.24	1.11	4.85	.74	3.88	1.22	4.49	1.00
Partner cohesion	2.42	1.86	4.11	1.20	.74	1.11	2.78	1.84
Shared cohesion	1.24	1.67	2.86	2.06	.69	1.01	2.67	1.76
Tie diversity:								
Firm	.73	.18	.85	.07	.66	.21	.77	.12
Partner	.47	.34	.75	.14	.05	.16	.27	.30
Prospective	.77	.12	.85	.06	.71	.15	.77	.12
Partner's partner	.32	.41	.79	.26	.34	.39	.73	.28
Controls:								
Type research	.31	.46	.21	.41	.22	.42	.33	.47
Finance	.16	.37	.28	.45	.47	.50	.27	.44
Comercial	.27	.44	.25	.43	.18	.38	.24	.43
Licensing	.26	.44	.25	.44	.13	.34	.16	.37
Form biomedical					.25	.43	.10	.29
Univ./nonprofit					.15	.36	.15	.35
Government					.04	.20	.10	.30
Venture capital					.33	.47	.25	.44
Pharmaceutical					.23	.42	.40	.49
Age	9.71	5.61	14.93	5.38	7.67	4.57	11.10	5.14
Size	569	1,200	2,588	2,165	192	584	681	1,449
Governance	.77	.42	.95	.21	.68	.47	.88	.33

TABLE B2

One-Mode Attachments: McFadden Multiprobability Model Estimates

	New Attachments			Repeat Attachments	
	(1)	(2)	(3)	(4)	(5)
Independent variables:					
Accumulative advantage:					
Partner degree	.0412**	.0618**	.0603**	−.0037	.0014
	(.0058)	(.0085)	(.0086)	(.0107)	(.0151)
New Partner	.1558	.1463	.1550	2.3936**	2.4770**
	(.1177)	(.1172)	(.1172)	(.5917)	(.6080)
				−.1429**	−.1529**
Partner experience	−.0537**	−.0474**	−.0477**		
	(.0128)	(.0128)	(.0130)	(.0420)	(.0447)
Prior ties				.3286**	.3799**
				(.0556)	(.0602)
Prior experience				−.2984**	−.3632**
				(.0947)	(.1027)
Homophily:					
Collaborative distance	−.0270+	−.0297*	−.0262+	.0513+	.0684*
	(.0142)	(.0149)	(.0153)	(.0293)	(.0311)
Age difference	.0353**	.0336**	.0492**	.1288**	.1117**
	(.0088)	(.0097)	(.0115)	(.0337)	(.0344)
Size difference	.0182**	.0163**	.0303**	.0011	.0012
	(.0045)	(.0046)	(.0062)	(.0011)	(.0012)
Governance similarity	−.1458	−.1238	−.1323	.7119+	.4917
	(.0892)	(.0901)	(.0906)	(.3909)	(.4140)
Co-location	.6681**	.6576**	.6566**	.7379*	.8580**
	(.0947)	(.0952)	(.0954)	(.2891)	(.2960)
Partner's partner:					
Collaborative distance	−.0027	−.0023	−.0021	.5245**	.5079**
	(.0144)	(.0145)	(.0146)	(.0584)	(.0601)
Age difference	−.0192	−.0116	.0830*	.0612	.0893*
	(.0118)	(.0119)	(.0331)	(.0409)	(.0437)
Size difference	−.0089	−.0083	−.0504*	−.0052**	−.0425*
	(.0056)	(.0056)	(.0247)	(.0018)	(.0188)
Governance similarity	−.0237	−.0143	−.0160	.6094	1.0179+
	(.1429)	(.1443)	(.1450)	(.5605)	(.5770)
Co-location	.2579+	−.2374+	.2210+	−.7763	−.9146+
	(.1334)	(.1340)	(.1343)	(.4950)	(.5146)
Follow-the-trend:					
Dominant trend	22.4007**	16.4288**	15.3775**	−21.5871*	−25.0997**
	(2.4351)	(3.3135)	(3.3216)	(10.0363)	(10.3612)
Dominant type	2.7150**	2.1221**	2.1213**	3.8823**	3.9045**
	(.1515)	(.1617)	(.1632)	(.6569)	(.6931)
Multiconnectivity:					
Partner cohesion	.3594**	.0419	.0543	.9512**	1.0824**
	(.0457)	(.1481)	(.1466)	(.1919)	(.2093)
Shared cohesion	.0566*	.1518**	.1458**	.0978	.0783
	(.0271)	(.0535)	(.0535)	(.0799)	(.0824)
Partner tie diversity	2.0492**	2.8440**	2.1283*	1.5821	1.9676
	(.2785)	(.9773)	(1.0408)	(1.4550)	(1.4345)
Prospective tie diversity	−2.7991	20.2502**	82.7609**	−30.3778**	−29.5611**
	(1.5170)	(5.6211)	(21.4090)	(10.6104)	(10.5284)
Partner's partner tie					
diversity	.4118	1.0324+	.8897	−5.8047**	−6.8265**
	(.4286)	(.6031)	(.6018)	(1.4601)	(1.6196)

1196

TABLE B2 (*Continued*)

	NEW ATTACHMENTS			REPEAT ATTACHMENTS	
	(1)	(2)	(3)	(4)	(5)
Contingencies:[a]					
Firm age × partner degree		−.0021**	−.0021**		
		(.0005)	(.0006)		
Firm age × shared cohesion		−.0107*	−.0103*		
		(.0048)	(.0048)		
Firm cohesion × partner cohe-					
sion		.0975**	.1275**		
		(.0346)	(.0357)		
Firm cohesion × prospective di-					
versity		−8.9460**	−27.6005**		
		(2.0969)	(6.4403)		
Period and timeline effects:					
Period 3 × age					
difference			−.0423*		
			(.0186)		
Period 3 × size					
difference			−.0184**		
			(.0068)		
Period 3 × partner					
cohesion			−.2068*		
			(.0835)		
Period 3 × partner					
diversity			4.0306*		
			(1.9870)		
Period 3 × prospective					
diversity			−69.9582**		
			(22.3349)		
Period 3 × firm cohesion ×					
prospective diversity			21.4625**		
			(6.9130)		
Timeline × partner's partner					
age difference			−.0106**		
			(.0037)		
Timeline × partner's partner					
size difference			.0044[+]		
			(.0024)		
LRχ²(*df*)	1,498	1,692	1,767	498	522
	(21)	(27)	(35)	(23)	(25)
Pseudo-*R*²	.17	.19	.21	.60	.64
N	966			373	

NOTE.—All models include fixed effects for firm, year, and type. Measure of firm size is divided by 100.

[a] Contingencies for activity type were included in models 2 and 3. They are available from Kenneth Koput.

[+] *P* < .10.

* *P* < .05.

** *P* < .01.

TABLE B3
TWO-MODE ATTACHMENTS: MCFADDEN MULTIPROBABILITY MODEL ESTIMATES

	NEW ATTACHMENTS			REPEAT ATTACHMENTS		
	(1)	(2)	(3)	(4)	(5)	(6)
Independent variables:						
Accumulative advantage:						
Partner degree	−.0859**	.0128	.0142	.0072	.0122	.0062
	(.0100)	(.0100)	(.0106)	(.0079)	(.0081)	(.0082)
New partner	.5081**	.7974**	.6150**	−.2385	−.2313	−.2323
	(.1167)	(.1350)	(.1669)	(.5891)	(.5772)	(.5799)
Partner experience	−.2140**	−.1535**	−.1381**	−.1453**	−.15276**	−.1819**
	(.0150)	(.0173)	(.0179)	(.0366)	(.0369)	(.0379)
Prior ties				.2327**	.2448**	.2470**
				(.0162)	(.0168)	(.0168)
Prior Experience				−.1480**	.0921	.1108
				(.0265)	(.0717)	(.0726)
Homophily:						
Collaborative distance	.1794**	.0386**	.0449**	.0168+	.0120	.0188+
	(.0124)	(.0126)	(.0127)	(.0100)	(.0102)	(.0104)
Partner's partner:						
Tie distance	−.0836**	−.0089	−.0083	.2766**	.2806**	.3077**
	(.0124)	(.0124)	(.0124)	(.0221)	(.0234)	(.0248)
Age difference	.0048	.0435**	.0430**	.1497**	.1525**	.1498**
	(.0133)	(.0163)	(.0162)	(.0229)	(.0235)	(.0237)
Size difference	.0033*	.0441**	.0415**	−.0088	−.0048	−.0139
	(.0013)	(.0122)	(.0123)	(.0157)	(.0157)	(.0162)
Governance similarity	.2690*	.2316	.2685	−.6362**	−.6019*	−.6873**
	(.1232)	(.1898)	(.1903)	(.2451)	(.2536)	(.2550)
Co-location	.0938	−.0934	−.1475	−2.0462**	−2.1407**	−2.1807**
	(.1327)	(.1763)	(.1765)	(.2758)	(.2857)	(.2870)
Follow-the-trend:						
Dominant trend	6.2637**	3.8608**	−3.7647**	−12.2855**	−12.1924**	−12.3164**
	(.4811)	(1.0541)	(1.0656)	(1.1006)	(1.3070)	(1.3330)
Dominant type	2.5418**	1.5736**	1.5962**	2.9147**	3.0110**	3.0718**

	(.1288)	(.1533)	(.1541)	(.2070)	(.2136)	(.2169)
Multiconnectivity:						
Partner cohesion	.5114**	2.1824**	3.29161**	.0771	.3877*	1.5000
	(.0622)	(.1907)	(.5708)	(.1064)	(.1871)	(.7250)
Shared cohesion	1.6624**	4.3605**	4.423**	.6444**	1.0000**	2.1678**
	(.0701)	(.2038)	(.2183)	(.1239)	(.1889)	(.5974)
Partner tie diversity	3.6399**	1.0723	2.9030**	1.9538**	1.8883**	−1.303$^+$
	(.2189)	(.7584)	(.2002)	(.3255)	(.3400)	(.7559)
Prospective tie diversity	−12.4229**	4.8200**	1.1967	−4.8813**	7.8716$^+$	9.9173*
	(.6607)	(1.6773)	(1.8764)	(1.6806)	(4.2148)	(4.6065)
Partner's partner tie diversity	−1.3459**	−8.0500**	−8.3912**	−4.1571**	−.9626	−5.2848$^+$
	(.3385)	(.7161)	(.7372)	(.8235)	(2.3109)	(2.7551)
Contingencies:[a]						
Firm cohesion × prior experience					−.0607**	−.0641**
					(.0162)	(.0163)
Firm cohesion × partner cohesion		−.6429**	−1.1419**		−.1186*	−.4833*
		(.0654)	(.2234)		(.0572)	(.2400)
Firm cohesion × shared cohesion		−.1650**	−.1655*			
		(.0624)	(.0648)			
Firm cohesion × partner tie diversity		−.1523	−.9175**			
		(.1834)	(.2418)			
Firm cohesion × prospective diversity		−5.1134**	−3.3183**			
		(.8105)	(.9022)			
Firm cohesion × partner's partner tie diversity		.9610**	1.0233**			
		(.1631)	(.1673)			
Firm tie diversity × partner tie diversity		−2.4258**	−1.5786$^+$			
		(.9173)	(.9318)			
Firm tie diversity × prospective diversity		−13.3864**	−17.1266**		−22.3118**	−26.5231**
		(2.6567)	(2.9355)		(7.7478)	(8.3263)
Firm tie diversity × partner's partner tie diversity		−1.9205**	−1.7864*		−7.0456*	−7.5804*
		(.7013)	(.7050)		(3.1050)	(3.2204)
Period and timeline effects:						
Period 3 × new partner			.5656*			
			(.2230)			
Period 3 × prospective diversity			11.7024**			
			(4.0933)			
Period 3 × firm cohesion × partner diversity			1.2202**			
			(.3217)			
Period 3 × firm cohesion × prospective diversity			−6.0453**			

			(1.8644)		
Period 3 × firm diversity × prospective diversity			14.2706*		
			(6.9727)		
Timeline × partner cohesion			−.1365*		−.1247[+]
			(.0610)		(.0748)
Timeline × firm cohesion × partner cohesion			.0573*		.0398[+]
			(.0231)		(.0241)
Timeline × shared cohesion					−.1336*
					(.0666)
Timeline × partner tie diversity					.4078**
					(.0895)
Timeline × partner's partner tie diversity					.6417**
					(.2020)
LR χ^2 (*df*)	5,970	8,893	9,059	2,231	2,289	2,347
	(17)	(40)	(48)	(18)	(24)	(29)
Pseudo-R^2	.32	.49	.51	.47	.49	.51
N	1,500			1,068		

NOTE.—All models include fixed effects for firm, year, and type, as well as main effects for form. Measure of firm size is divided by 100.

[a] Form and type contingencies were also included in the model. They are available from Kenneth Koput.

[+] $P < .10$.

* $P < .05$.

** $P < .01$.

REFERENCES

Abbott, Andrew. 2001. *Time Matters*. Chicago: University of Chicago Press.

Albert, Réka, and A. L. Barabási. 2002. "Statistical Mechanics of Complex Networks." *Reviews of Modern Physics* 74 (1): 47–97.

Albert, Réka, H. Jeong, and A. L. Barabási. 1999. "Diameter of the World Wide Web." *Nature* 401:130–31.

———. 2000. "Error and Attack Tolerance in Complex Networks." *Nature* 406:378–82.

Barabási, Albert-László. 2002*a*. "Emergence of Scaling in Complex Networks." Pp. 69–84 in *Handbook of Graphs and Networks: From the Genome to the Internet*, edited by S. Bornholdt and H. G. Schuster. Berlin: Wiley-VCH.

———. 2002*b*. *Linked: The New Science of Networks*. Cambridge, Mass.: Perseus.

Barabási, Albert-Lázló, and Réka Albert. 1999. "Emergence of Scaling in Random Networks." *Science* 286:509–12.

Batagelj V., A. Mrvar. 2000. "Drawing Genealogies." *Connections* 21:47–57.

Baum, Joel A. C., Andrew Shipilov, and Tim Rowley. 2003. "Where Do Small Worlds Come From?" *Industrial and Corporate Change* 12 (4): 697–725.

Ben-Akiva, M., and S. R. Lerman. 1989. *Discrete Choice Analysis*. Cambridge, Mass.: MIT Press.

Blau, Peter. 1977. *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. New York: Free Press.

Bollobás, Bela, and Oliver Riordan. 2002. "Mathematical Results on Scale-free Random Graphs." Pp. 1–34 in *Handbook of Graphs and Networks: From the Genome to the Internet*, edited by S. Bornholdt and H. G. Schuster. Berlin: Wiley-VCH.

Bourdieu, Pierre. 1992. "The Logic of Fields." Pp. 94–114 in *An Invitation to Reflexive Sociology*, by P. Bourdieu and L. Wacquant. Chicago: University of Chicago Press.

Burt, Ronald S. 1992. *Structural Holes*. Cambridge, Mass.: Harvard University Press.

———. 2000. "Decay Functions." *Social Networks* 22:1–28.

Davis, Gerald F., Mina Yoo, and Wayne Baker. 2003. "The Small World of the Corporate Elite, 1982–2001." *Strategic Organization* 1:301–36.

de Nooy, Wouter, Andrej Mrvar, and Vladimir Batagelj. 2005. *Exploratory Social Network Analysis with Pajek*. New York: Cambridge University Press.

de Solla Price, Derek J. 1965. "Networks of Scientific Papers." *Science* 149:510–15.

———. 1980. "A General Theory of Bibliometrics and Other Cumulative Advantage Processes." *Journal of the American Informatics Society* 27:292–306.

Dezalay, Yves, and B.G. Garth. 1996. *Dealing in Virtue*. Chicago: University of Chicago Press.

DiMaggio, Paul J. 1991. "Constructing an Organizational Field as a Professional Project: U.S. Art Museums, 1920–1940." Pp. 267–92 in *The New Institutionalism in Organizational Analysis*, edited by W. Powell and P. J. DiMaggio. Chicago: University of Chicago Press.

DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collectivity Rationality in Organizational Fields." *American Sociological Review* 48:147–60.

Ferguson, Priscilla Parkhurst. 1998. "A Cultural Field in the Making: Gastronomy in Nineteenth-Century France." *American Journal of Sociology* 104:597–641.

Fruchterman, T., and E. Reingold. 1991. "Graph Drawing by Force-Directed Replacement." *Software—Practice and Experience* 21:1129–64.

Galambos, Louis, J. Sturchio. 1996. "The Pharmaceutical Industry in the Twentieth Century." *History and Technology* 13 (2): 83–100.

Gambardella, Alfonso. 1995. *Science and Innovation: The U.S. Pharmaceutical Industry during the 1980s*. Cambridge: Cambridge University Press.

Gulati, Ranjay, and Martin Gargiulo. 1998. "Where Do Interorganizational Networks Come From?" *American Journal of Sociology* 104 (5): 1439–93.

Guyon, X. 1994. *Random Fields on a Network*. Berlin: Springer.

Hagedoorn, John, and Nadine Roijakkers. 2002. "Small Entrepreneurial Firms and Large Companies in Inter-Firm R&D Networks—the International Biotechnology Industry." Pp. 223–52 in *Strategic Entrepreneurship*, edited by M. A. Hitt et al. Cambridge, Mass.: Blackwell Publishing.

Henderson, Rebecca, and Iain Cockburn. 1996. "Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery." *Rand Journal of Economics* 27 (1): 32–59.

Henderson Rebecca, Luigi Orsenigo, and Gary Pisano. 1999. "The Pharmaceutical Industry and the Revolution in Molecular Biology: Interactions among Scientific, Institutional, and Organizational Change." Pp. 267–311 in *Sources of Industrial Leadership*, edited by D. C. Mowery and R.R. Nelson. New York: Cambridge University Press.

Hoffman, Andrew J. 1999. "Institutional Evolution and Change: Environmentalism and the U.S. Chemical Industry." *Academy of Management Journal* 42 (4): 351–71.

———. 2001. *From Heresy to Dogma: An Institutional History of Corporate Environmentalism*. Stanford, Calif.: Stanford University Press.

Holmström, Bengt, and John Roberts. 1998. "The Boundaries of the Firm Revisited." *Journal of Economic Perspectives* 12:73–94.

Jeong, Hawoong, B. Tombor, R. Albert, Z. Oltvai, and A. L. Barabási. 2000. "The Large-Scale Organization of Metabolic Networks." *Proceedings of the National Academy of Sciences* 98:404–9.

Kamada, T., and S. Kawai. 1989. "An Algorithm for Drawing General Undirected Graphs." *Information Processing Letters* 31:7–15.

Kogut, Bruce, and Gordon Walker, 2001. "The Small World of Germany and the Durability of National Networks." *American Sociological Review* 66 (3): 317–35.

Lerner, Josh, H. Shane, and A. Tsui. 2003. "Do Equity Financing Cycles Matter? Evidence from Biotechnology Alliances." *Journal of Financial Economics* 67 (3): 411–46.

Lincoln, James, Michael Gerlach, and Christina Ahmadjian. 1996. "Keiretsu Networks and Corporate Performance in Japan." *American Sociological Review* 61:67–88.

Lotka, A. J. 1926. "The Frequency Distribution of Scientific Productivity." *Journal of the Washington Academy of Science* 16:317–323.

Macneil, Ian R. 1978. "Contracts: Adjustment of Long-term Economic Relations under Classical, Neoclassical, and Relational Contract Law." *Northwestern University Law Review* 72:854–905.

Maddala, G. S. 1986. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

March, James G. 1991. "Exploration and Exploitation in Organizational Learning." *Organization Science* 2:71–87.

March, James G., and Johan P. Olsen. 1989. *Rediscovering Institutions*. New York: Free Press.

McFadden, Daniel. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior." Pp. 105–42 in *Frontiers in Econometrics*, edited by P. Zarembka. New York: Academic Press.

———. 1981. "Econometric Models of Probabilistic Choice." Pp. 198–272 in *Structural Analysis of Discrete Data: With Econometric Applications*, edited by C. Manski and D. McFadden. Cambridge, Mass.: MIT Press.

———. 1984. "The Measurement of Urban Travel Demand." *Journal of Public Economics* 3:303–28.

———. 2001. "Economic Choices." *American Economic Review* 91:351–78.

McKelvey, Maureen. 1996. *Evolutionary Innovations*. Oxford: Oxford University Press.

McPherson, Miller, and L. Smith-Lovin. 1987. "Homophily in Voluntary Organizations." *American Sociological Review* 52:370–79.

McPherson, Miller, L. Smith-Lovin, and J. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415–44.

Merton, Robert K. 1973. "The Normative Structure of Science." Pp. 267–78 in *The Sociology of Science*, by Robert K. Merton. Chicago: University of Chicago Press.

Milgram, Stanley. 1967. "The Small World Problem." *Psychology Today* 2:60–67.

Moody, James. 2001. "Race, School Integration, and Friendship Segregation in America." *American Journal of Sociology* 107 (3): 679–716.

———. 2004. "The Structure of a Social Science Collaboration Network: Disciplinary Cohesion from 1963 to 1999." *American Sociological Review* 69:213–38.

Moody, James, and Douglas R. White. 2003. "Social Cohesion and Embeddedness: A Hierarchical Conception of Social Groups." *American Sociological Review* 68 (1): 103–28.

Morrill, Calvin, and Jason Owen-Smith. 2002. "The Emergence of Environmental Conflict Resolution." Pp. 90–118 in *Organizations, Policy, and the Natural Environment*, edited by A. J. Hoffman and M. J. Ventresca. Stanford, Calif.: Stanford University Press.

National Science Foundation. 2002. *Science and Engineering Indicators, 2002*. Appendix table 6-19, pp. A6–64. National Science Board.

Newman, Mark. 2001. "The Structure of Scientific Collaboration Networks." *Proceedings of the National Academy of Sciences* 98 (January): 404–9.

———. 2003. "The Structure and Function of Complex Networks." *SIAM Review* 45: 167–256.

Orsenigo L. 1989. *The Emergence of Biotechnology*. New York: St. Martin's Press.

Orsenigo L., F. Pammolli, M. Riccaboni. 2001. "Technological Change and Network Dynamics: Lessons from the Pharmaceutical Industry." *Research Policy* 30:485–508.

Owen-Smith, Jason. 2003. "From Separate Systems to a Hybrid Order: Accumulative Advantage across Public and Private Science at Research One Universities." *Research Policy* 32 (6): 1081–1104.

Owen-Smith, Jason, and Walter W. Powell. 2001*a*. "Careers and Contradictions: Faculty Responses to the Transformation of Knowledge and Its Uses in the Life Sciences." *Research in the Sociology of Work.* 10:109–40.

———. 2001*b*. "To Patent or Not: Faculty Decisions and Institutional Success in Academic Patenting." *Journal of Technology Transfer* 26 (1): 99–114.

———. 2004. "Knowledge Networks in the Boston Biotechnology Community." *Organization Science* 15 (1):5–21.

Owen-Smith, Jason, M. Riccaboni, F. Pammolli, and W. W. Powell. 2002. "A Comparison of U.S. and European University-Industry Relations in the Life Sciences." *Management Science* 48 (1): 24–43.

Padgett, John F., and Chris Ansell. 1993. "Robust Action and the Rise of the Medici, 1400–1434." *American Journal of Sociology* 98:1259–1319.

Powell, Walter W. 1990. "Neither Market Nor Hierarchy: Network Forms of Organization." *Research in Organizational Behavior* 12:295–336.

———. 1996. "Inter-Organizational Collaboration in the Biotechnology Industry." *Journal of Institutional and Theoretical Economics* 120 (1): 197–215.

Powell, Walter W., and Peter Brantley. 1992. "Competitive Cooperation in Biotechnology: Learning through Networks?" Pp. 366–94 in *Networks and Organizations*, edited by R. Eccles and N. Nohria. Boston: Harvard Business School Press.

———. 1996. "Magic Bullets and Patent Wars: New Product Development in the Biotechnology Industry." Pp. 233–60 in *Managing Product Development*, edited by T. Nishiguchi. New York: Oxford University Press.

Powell, Walter W., K. W. Koput, J. I. Bowie, and L. Smith-Doerr. 2002. "The Spatial

Clustering of Science and Capital: Accounting for Biotech Firm-Venture Capital Relationships." *Regional Studies* 36 (3): 291–306.

Powell, Walter W., Kenneth Koput, and Laurel Smith-Doerr. 1996. "Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology." *Administrative Science Quarterly* 41 (1): 116–45.

Powell, Walter W., Kenneth Koput, Laurel Smith-Doerr, and Jason Owen-Smith. 1999. "Network Position and Firm Performance: Organizational Returns to Collaboration." *Research in the Sociology of Organizations* 16:129–59.

Robinson, David, and Toby E. Stuart. 2002. "Just How Incomplete Are Incomplete Contracts?" Working paper. Columbia University, Graduate School of Business.

Salancik, Gerald. 1995. "WANTED: A Good Network Theory of Organization." *Administrative Science Quarterly* 45 (1): 1–24.

Scott, W. Richard, Martin Reuf, Peter J. Mendel, and Carol Caronna. 2000. *Institutional Change and Health Care Organizations*. Chicago: University of Chicago Press.

Sharpe, Margaret. 1991. "Pharmaceuticals and Biotechnology: Perspectives for the European Industry." In *Technology and the Future of Europe*, edited by C. Freeman et al. London: Pinter.

Sorenson, Olav, and Toby Stuart. 2001. "Syndication Networks and the Spatial Distribution of Venture Capital Investment." *American Journal of Sociology* 106: 1546–88.

Stuart, Toby E. 1998. "Network Positions and Propensities to Collaborate." *Administrative Science Quarterly* 43:668–98.

Teece, David. 1986. "Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy." *Research Policy* 15 (6): 185–219.

Thornton, Patricia. 1995. "Accounting for Acquisition Waves: Evidence from the U.S. College Publishing Industry." Pp. 199–225 in *The Institutional Construction of Organizations*, edited by W. R. Scott and S. Christensen. Thousand Oaks, Calif.: Sage.

Uzzi, Brian, and Jarrett Spiro. 2005. "Collaboration and Creativity: The Small World Problem." *American Journal of Sociology*, vol. 111, no. 2 (September). In press.

Walcott, Susan M. 2002. "Analyzing an Innovative Environment: San Diego as a Bioscience Beachhead." *Economic Development Quarterly* 16 (2): 99–114.

Wasserman, Stanley, and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.

Watts, Duncan. 1999. *Small Worlds*. Princeton, N.J.: Princeton University Press.

Watts, Duncan, and Stephen Strogatz. 1998. "Collective Dynamics of 'Small-World' Networks." *Nature* 393:440–42.

White, Douglas R., and Frank Harary. 2001. "The Cohesiveness of Blocks in Social Networks: Node Connectivity and Conditional Density." *Sociological Methodology* 31 (1): 305–59.

White, Douglas R., Jason Owen-Smith, James Moody, and Walter W. Powell. 2004. "Networks, Fields, and Organizations: Scale, Topology, and Cohesive Embeddings." *Computational and Mathematical Organization Theory* 10 (1): 95–117.

White, Harrison. 1970. "Search Parameters for the Small World Problem." *Social Forces* 49:259–64.

———. 1981. "Where Do Markets Come From?" *American Journal of Sociology* 81: 730–79.

———. 1985. "Agency as Control." Pp. 187–214 in *Principals and Agents*, edited by J. Pratt and R. Zeckhauser. Boston: Harvard Business School Press.

———. 1992. *Identity and Control*. Princeton, N.J.: Princeton University Press.

Zucker, Lynne, and Michael Darby. 1997. "Present at the Revolution: Transformation of Technical Identity for a Large Pharmaceutical Firm after the Biotechnological Breakthrough." *Research Policy* 26 (4): 429–47