

NBER WORKING PAPER SERIES

**COLLABORATION STRUCTURE AND
INFORMATION DILEMMAS IN
BIOTECHNOLOGY: ORGANIZATIONAL
BOUNDARIES AS TRUST PRODUCTION**

**Lynne G. Zucker
Michael R. Darby
Marilynn B. Brewer
Yusheng Peng**

Working Paper No. 5199

**NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1995**

This research has been supported by grants from the Alfred P. Sloan Foundation through the NBER Research Program on Industrial Technology and Productivity, the National Science Foundation (SES 9012925), the University of California Systemwide Biotechnology Research and Education Program, the University of California Systemwide Pacific Rim Research Program, the UCLA Center for American Politics and Public Policy, and the UCLA Institute of Industrial Relations. We acknowledge very useful comments received from participants in the Stanford Conference on Trust in Organizations, May 1994, and encouraging comments on a preliminary version from participants in the West Coast Conference on Small Group Research, April 1993. The contributions of Maximo Torero to graphics development and SAS programming are especially appreciated; Kerry Knight also made substantive contributions to the analysis. We are indebted to a remarkably talented team of post-doctoral fellows Zhong Deng, Julia Liebeskind, and Hiromi Ono and research assistants Paul J. Alapat, Jeff Armstrong, Cherie Barba, Lynda J. Kim, Kerry Knight, Edmundo Murrugara, Amalya Oliver, Alan Paul, Erika Rick, Maximo Torero, Alan Wang, and Mavis Wu. This paper is part of NBER's research program in Productivity. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

© 1995 by Lynne G. Zucker, Michael R. Darby, Marilyn B. Brewer, and Yusheng Peng. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

COLLABORATION STRUCTURE AND
INFORMATION DILEMMAS IN
BIOTECHNOLOGY: ORGANIZATIONAL
BOUNDARIES AS TRUST PRODUCTION

ABSTRACT

Scientists who make breakthrough discoveries can receive above-normal returns to their **intellectual capital**, with returns depending on the degree of **natural excludability** - that is, whether necessary techniques can be learned through written reports or instead require hands-on experience with the discovering scientists or those trained by them in their laboratory. Privatizing discoveries, then, only requires selecting trusted others as collaborators, most often scientists working in the same organization. Within organizational boundaries, incentives become aligned based on repeat and future exchange, coupled with third-party monitoring and enforcement. We find that high value intellectual capital paradoxically predicts both a larger number of collaborators and more of that network contained within the same organization. Specifically, same-organization collaboration pairs are more likely when the value of the intellectual capital is high: both are highly productive "star" scientists, both are located in top quality bioscience university departments, or both are located in a firm (higher ability to capture returns). Collaboration across organization boundaries, in contrast, is negatively related to the value of intellectual capital and positively related to the number of times the star scientist has moved. Organizational boundaries act as **information envelopes**: The more valuable the information produced, the more its dissemination is limited. In geographic areas where a higher proportion of coauthor pairs come from the same organization, diffusion to new collaborators is retarded.

Lynne G. Zucker
Department of Sociology
University of California, Los Angeles
Los Angeles, CA 90095-1551
and NBER

Michael R. Darby
Anderson Graduate School of Management
University of California, Los Angeles
Los Angeles, CA 90095-1481
and NBER

Marilynn B. Brewer
Department of Psychology
Ohio State University
1883 Neil Avenue
Columbus, OH 43210-1222

Yusheng Peng
Department of Sociology
Chinese University of Hong Kong
Shatin, N.T.
HONG KONG

COLLABORATION STRUCTURE AND INFORMATION
DILEMMAS IN BIOTECHNOLOGY:
ORGANIZATIONAL BOUNDARIES AS TRUST PRODUCTION

Lynne G. Zucker, Michael R. Darby, Marilyn B. Brewer, Yusheng Peng

Science norms urge sharing of information, science rewards require publication in refereed journals, and science training generally includes student access to new information as active components of on-going research teams. In most sociology of science, it is assumed that the information being created has value, but is treated as a public good: scientists both contribute to and draw from a common resource pool that consists of discoveries and refinements of those discoveries. In most economic treatments, it is assumed that scientific discoveries have only fleeting value unless formal intellectual-property-rights mechanisms are used to prevent use of the information; i.e., absent patents, trade secrets, or actual secrecy, the value of a discovery erodes quickly as the information diffuses.

We have quite a different view. Scientific discoveries vary in the degree to which others can be excluded from making use of them. Inherent in the discovery itself is the degree of **natural excludability**: if the techniques for replication are not widely known prior to the discovery, then any scientist wishing to build on the new knowledge must first acquire hands-on experience.[2] If he or she cannot gain access to a research team or laboratory setting with that know-how, then working in that area may be difficult if not impossible. We argue that it is primarily in collaborations that the fine details and hands-on knowledge to conduct cutting-edge bioscience is transmitted; only in collaborations are all biologic materials freely shared, though requirements for making them more widely available once research is published are growing (Eisenberg 1987: 197-205, 229-231).

Trust is extraordinarily important in communicating discoveries in biotechnology because

of their high scientific and commercial value. The resulting intense competition produces an information dilemma, with contradictory incentives to communicate the new knowledge and to withhold it (Schneider and Brewer 1987; Schneider 1990).[3] In brief, if a scientist communicates usable information about a new discovery, the benefits associated with exclusive access to that information are compromised.[4] But withholding information about the new discovery may slow progress in the field as a whole.

Information dilemmas, usually couched as conflict between individual self-interest and group interest, can be resolved by relying on close-knit collaborations, sharply limiting with whom the new discovery is shared. While the information is not shared with the field as a whole, it is shared with a group of collaborators that tends to grow over time. The information boundaries that these collaboration structures define determine the extent of diffusion of the new discovery. Because organizations have both established internal exchange relations and enforcement mechanisms, we expect that trust among members of the same organization will be significantly higher than trust between members of different organizations, and thus that organizational boundaries are efficient **information envelopes**. In general, the higher the value of the intellectual capital, the more likely organizational boundaries are used to limit its diffusion.

Thus, we can extend our argument one step further, to the effects of organizational boundaries on diffusion of information. If trust is produced, and information flow is in fact restricted along organizational lines, then diffusion should slow differentially. Specifically, within a geographic area, the higher the proportion of same organization pairs of co-authors, the less information should diffuse within that area. Indeed, our final model explains nearly all of the variation in diffusion to new co-authors of scientific articles between geographic areas, with significant amounts explained by variables related to value of intellectual capital and the resultant

patterns of collaboration within or between organizations.

Trust Production in Information Dilemmas

Trust production can occur when an individual is **open to social influence** from another individual, or when a third party with whom both individuals are open to social influence intervenes to mediate (Zucker 1986). This reframes the trust problem substantially, focusing attention on the mechanisms by which an individual becomes more or less open to social influence from another individual (or the interaction becomes mediated). Predictability or certainty is **not** sufficient for trust production, contra the Williamsonian (1979) argument; for example, a narcissistic individual will behave predictably in his or her self-interest, but this consistent behavior will on average produce distrust, since much of the self-interested behavior is not open to influence by the other person in the exchange and ignores that other's interests.

Trust, by defining the group of others who are likely to be open to social influence reciprocally, determines where the information boundaries will be drawn. The higher the value of the information the more likely it is that trust production will be a central concern. Trust production is often based on institutional mechanisms, including in-group preference and formal rules and procedures defined by formal organization boundaries (Zucker 1983 and 1986; Brewer and Silver 1978).[5] We expect that organizational boundaries will heavily determine with whom to share recombinant DNA (rDNA) sequence discoveries.

In the economic literature, Darby and Karni (1993), Klein, Crawford, and Alchian (1978), and Darby and Lott (1989), have studied the use of third party experts as one means to create trust for parties to act more to maximize group values rather than simply to pursue myopic self interest. Trusted agents receive a higher income from their "brand name capital" which is reduced or lost if they do not behave consistent with that trust. Also, when a third party is not available to monitor the exchange, anthropologists (e.g. Geertz 1978) find that a

combination of repeat exchange and expected future exchange best produce trust (see also Kollock 1994).

Within organizations, of course, both repeated and future exchange are common, as is third party intervention, or at least the possibility of it. Senior managers, for example, are trusted experts with organizational and self interests aligned who potentially can apply substantial sanctions in the face of self-seeking behavior. To the extent that within-organization collaborations involve a third party, whether explicitly or only implicitly, organizational involvement will help to increase the self-enforcing range, and thus produce a higher rate of collaboration within organizational boundaries.[6] Though we do not measure these effects directly, our indirect measurement provides substantial support for this line of argument.

Production of trust, then, involves information boundaries that are at least partially constructed by collaborating within the same organization (university, research institute/hospital, or firm). We in fact demonstrate that there are specific characteristics of other scientists, generally those indicating the potential value of their discoveries, that lead them to be included more often in scientific collaborations, and to be more often in collaborations with scientists working in the same organization. Our finding is consistent with expectations derived from the transaction cost approach, where the boundaries of the firm are determined by the relative costs and benefits of using markets or the firm's own hierarchy to govern each exchange (Coase 1937 and 1988; Williamson 1979 and 1991). If there exist similar exchanges both within and between organizations, the higher costs involved in transacting across organizational boundaries should imply that these transactions will have higher value, or not occur.

Cost and Demand in Trust Production

Social agency is required to **produce** trust. Human action is required to form repeat collaborations or collaborations within an organization and to exclude alternative forms of action.

As Zucker has argued elsewhere (Zucker 1986; Zucker and Kreft 1994; Tolbert and Zucker 1995), this social agency involves activity that is costly, requiring human time, attention, and resources, and thus there must be some **demand** for trust before it will be produced. Yet social scientists often unrealistically treat social process and the resulting structure as if they were simply a byproduct of human activity, as ubiquitous as air, and therefore costless (Granovetter 1985). In contrast, we argue that deciding to produce trust incurs drawing the opportunity cost of resources away from other activities. It is unlikely that such resource flow will occur in the absence of identified demand.[7] We will first examine costs involved in production of trust in the biosciences and then effects of the value of the information on generating demand for trust.

Costs may not be simply the human energy and money expended, but also other opportunities lost. In the case of bioscience collaborations, for example, a colleague in the same institution may become a coauthor for the reasons outlined above, yet not be the ideal choice as a coauthor in terms of potential intellectual contribution to the project. Other bioscientists who might make a much more significant contribution to the project are not included because their trustworthiness is too costly to establish compared to the alternatives.

Costs also are incurred because defining who is trusted sufficiently to include in a collaboration simultaneously defines a much larger group of scientists who are excluded. While science norms call for **inclusion** of a large circle of trusted colleagues in an "invisible college" model of both high volume and high velocity exchange of scientific information/discovery (Merton 1938 and 1957; Crane 1969 and 1972; Gaston 1973), there is considerable evidence that whenever the discoveries have significant **value**, whether as pure science or as a commercial product, some scientists will exploit nonpublic knowledge for personal gains -- monetary or nonmonetary (Watson 1968; Taubes 1986). If the excluded scientists recognize the value, then they become angry and voice complaints; much time and energy is spent responding to these

complaints and either resolving specific difficulties or creating new structures designed to define the amount of exclusion permitted. In a thorough review of the early controversies in bioscience, Rebecca Eisenberg documents the kinds of disagreements that occurred over access to cell cultures that, if access is denied, make actual replication of the published research and later extension of it close to impossible (1987: pp. 197-205, 214-216, and 229-231, and examples throughout).

In the face of these costs of trust production, some positive demand must exist in order for trust to be produced. Our hypothesis is that the major demand for trust in scientific research is derived from the potential value of the specific discoveries.

Value of Intellectual Capital and Demand for Trust

Let us define "intellectual capital" as the value of nonpublic information possessed by an individual in excess of the costs of learning the information (see Zucker, Darby, and Brewer 1994). We conceive of translating nonmonetary returns to the information such as prestige and professional advancement to monetary equivalents and then taking the present value of the sum of the monetary value of all future monetary and nonmonetary returns and subtracting the cost of learning the information.[8] So defined, intellectual capital is the wealth value of the knowledge to an individual who makes a significant discovery or to whom the information discovered is transmitted before it has diffused sufficiently to earn only the normal returns to the cost of learning the information.

As information diffuses after a discovery, the associated intellectual capital of a person who embodies the information declines both because the supranormal returns decline as there are more scientists using the information competitively and because there is less time remaining until the information is part of routine science and thus no longer capable of earning supranormal returns. The value of intellectual capital created by a new discovery increases as the discovery

involves techniques that must be learned first-hand through collaboration or apprenticeship; that is, to the extent that they possess the "natural excludability" necessary for relatively long-lasting supranormal returns.[9]

Scientists incorporate the information -- transiently both nonpublic and yielding supranormal returns -- as part of their human capital. The returns to this information may come from other scientists, in the form of citations, promotions, job offers, and so on. These returns may also come from commercialization of the discovery, in the form of consulting income, patent royalties, ownership interest in the firm, and so on. In Zucker, Darby, and Brewer (1994), we demonstrate that intellectual capital in conjunction with active publishing is a strong predictor of founding new biotechnology enterprises. This reward structure encourages other scientists to invest in learning the know-how to obtain the supranormal returns, although as more come to know the information its capital value falls.

Students have especially strong incentives to work with scientists on the leading edge, from whom they can gain knowledge that is not available from other scientists. As Harriet Zuckerman discovered in her study of Nobel laureates (1967 and 1977), scientists working at the intellectual frontier can obtain the best students, another form of capital, by diffusing the knowledge to them. Zuckerman's path-breaking work establishes as a central research question the differences in scientific production between the first rank scientist in terms of quality and those lower down in the prestige hierarchy of science.

We argue that the higher the value of the intellectual capital, the greater the demand for production of trust and the use of information boundaries to create it. The value of intellectual capital varies along many dimensions, four of which we will examine empirically in this paper:

- o **Quality of the individual scientist:** We identify "star" scientists in terms of their productivity in biotechnology, and study these stars and their collaborators. The

quality of intellectual output generally will be higher if two star scientists collaborate than if one star scientist collaborates with graduate students, postdoctoral students, and/or other less distinguished scientists.

- o **Quality of the university:** We identify scientists in terms of their location at one of 18 top quality universities, defined by having exceptionally high ranking biochemistry, microbiology, and/or molecular biology programs (see below). Working in one of these top quality university departments indicates expected high quality intellectual output and greater use of information boundaries.
- o **Time:** As the discovery diffuses, the value of the intellectual capital declines. Over time, then, we expect that information boundaries will be relaxed.
- o **Appropriation regime:** Firms are generally able to appropriate the value of new discoveries more successfully than universities or research institutes, and thus we expect greater use of information boundaries in firms.

Scientific Discovery and Collaboration Structure in Biotechnology

Discovery: Gene splicing is an extremely significant discovery that set off a cascade of research in biotechnology (see Cohen, Chang, Boyer, and Helling 1973). Though there were other very significant discoveries made in biotechnology at about the same time, only genetic sequence discoveries were exhaustively cataloged in a data file, GenBank, created for bioscientists.[10] Gene splicing has the distinction of being a scientifically valuable innovation that has a virtually complete record of subsequent discoveries of genetic sequences, so that the timing, extent, and other characteristics of its diffusion can be tracked. We extensively reprogrammed the original GenBank files to identify the 327 "top producing" scientists: those scientists who discovered at least 41 genetic sequences from 1968 to 1990. Because some sequences are harder to unravel, we added to these "star" scientists another 22 who published

20 or more articles identifying gene sequences, which made them among the most productive scientists, but reported fewer genetic sequence discoveries per article. Thus, the **star** scientists we identify world-wide total 327, listed as authors on 4,061 distinct published articles in major journals.[11] The distribution of these publications over time can be seen in Figure X.1, along with the distribution for the 208 stars who ever work in the U.S.

At the beginning of the time period, in 1967, it was exceedingly difficult to make a genetic sequence discovery. Over our time period, it gradually became easier to sequence genes, until toward the end of our time period these discoveries become part of normal science. Between 1967 and 1990, there is both increasing diffusion of the ability to gene sequence and a decreasing value of the related techniques. Indeed, by 1987 or 1988 discovering a gene sequence could no longer earn a Ph.D. at any of the major U. S. research universities. At the same time, machine-based technology was becoming more reliable and relatively inexpensive.[12]

The reorganized GenBank data files contain a great deal of information on the characteristics of the discoveries, including the list of authors who collaborated in making the discoveries. These scientists are co-authors of our 327 star scientists, but do not themselves meet the "star" criteria. These **collaborators** number 7,825 with 4,012 of them ever working in the U.S. Because we are examining the effects of collaborating within the same organization on fine-grain diffusion patterns, we will be focusing only on scientific collaborations in the U.S. in our analyses in this paper.

Without adding to the original GenBank files we are able to study the general pattern of diffusion. But information on the organizational membership (same versus different; type of organization, including university, research institute or hospital, and firm) and on the geographic location of the scientists is not contained in GenBank, nor are they available except for the first

author in any on-line data base (see MedLine). If star scientists always appear first, the use of MedLine affiliation would at least permit the study of the top producers. But our informal information proved accurate: star scientists most often appear last, so that among articles with U.S. stars, 71.2 percent of last authors are stars, compared to 16.4 percent of first authors.

Organizational Location and Characteristics: Since our major hypotheses concerning the role of trust in collaborations rest on organizational location, we pulled each one of the 4,315 articles identified as containing "high quality" intellectual capital and coded institutional affiliation and location of each author, including any dual affiliations and changes in affiliation, from the headings of the articles. From these data, we find that intellectual capital is rather concentrated: in the U.S., only 264 distinct organizations were given at any time through 1990 as affiliations of any of the stars or collaborators in our articles data set. There are 149 universities, 71 research institutes and hospitals, and 44 firms.

Our university data consist of all U.S. institutions listed as granting the Ph.D. degree in any field in the Higher Education General Information Survey (HEGIS), Institutional Characteristics, 1983-84.[13] Each university is assigned an institutional ID number, a university flag, and located by zip code based on the HEGIS address file. Additional information was collected for those universities granting the Ph.D. degree in biochemistry, cellular/molecular biology, and/or microbiology which we define as "**biotech-relevant**" fields (see Jones, Lindzey, and Coggeshall 1982). In our analyses here, we rely only on the National Academy of Science measure of university department quality, a scholarly reputation rating based on responses from approximately fifteen percent of the faculty in the fields studied. Since we were interested in identifying the very best programs, we considered only the highest rated of the biochemistry, cellular/molecular biology, and/or microbiology programs offered by a particular university. Whether or not any specific star or collaborator in the U.S. is located in

one of the most highly rated programs (rated above 4) is our variable TQU, or "Top Quality University." [14]

Firms and research institutes (including hospitals) listed as affiliations in the article data set, received an institutional ID number and an institute/hospital flag, and obtained an address with zip code as required for geocoding. For a review of additional data and sources not used in the analyses presented in this paper, including non-U.S. information, see "Data Sources" at the end of the reference list for this paper and Appendix A in Zucker, Darby, and Brewer (1994).

Collaboration Structure: We now introduce a further refinement to our measurement of collaborations. Since our sample is selected based on star scientists, and sequence-reporting articles by collaborators appear in our sample only if one of our star scientists is an author, we define the collaboration as all possible pairs of coauthors that have at least one star in it. Our basic unit is thus the number of **coauthor pairs** in a collaboration that have at least one star. By examining coauthor pairs, we are able to model explicitly the selection criteria used in generating our sample: each pair must have at least one star. Most of these coauthor pairs consist of one star and one collaborator, for a worldwide total of 20,595, with 9,025 of those having both scientists located in the U.S. and 10,514 having at least one scientist in the U.S. Two-star pairs constitute 2,044 of the worldwide pairs and 904 pairs with at least one star in the U.S (820 both in U.S.).

The distribution of coauthor pairs and the increasing number over time may be attributed both to a dense interaction network of lab-based collaborations and to the diffusion of the new technology. [15] But this is where our story really begins.

Organizations as Information Envelopes: Value, Trust, and Diffusion

Changes in patterns and relationships among bioscientists occur over time, reflecting

changes in the value of the information and the consequent demand for trust production. Two aspects of the information flow identify gradual changes in use of information boundaries consistent with the initial high value of intellectual capital and subsequent decline in its value:

- o The number of coauthor pairs increases over time in part as a result of increasing average collaboration size, such that wider access is being given to the new scientific discoveries as the area develops. This suggests that the value of the intellectual capital produced is falling. Figure X.2 shows the increasing average number of authors per article worldwide, smoothed to reduce the effects of very small numbers of articles published prior to the mid-1970s.
- o More coauthor pairs are located in the same organization early in the process, suggesting that within-organization collaboration is being used to limit information flow more early in the diffusion process. The number of same-organization pairs decline steadily over time, again pointing to the declining value of intellectual capital. Figure X.3 shows the decrease in the percentage of same institution coauthors.

Specification of Variables and Models: While these over time changes are suggestive of the relationships we hypothesize, we turn now to multivariate analysis of the collaboration structure in the biosciences. Specifically, as the value of the genetic sequence discoveries being reported increases, we expect that there will be more frequent same organization collaboration, and less frequent collaboration with scientists from different organizations, and that same organization collaborations will slow diffusion relative to collaborations outside the same organization. Because of our focus here on the diffusion process, we will limit our analyses primarily to the U.S., studying relative diffusion within the 183 different BEA areas.

The variables designed to operationalize these concepts, including the value of the intellectual capital being produced, the collaboration structure that results, measures of time

trends, scientist mobility, and the "birth" of new collaborators are all described in Table 1, along with several control variables. All of the variables are based on characteristics of coauthor pairs, where at least one scientist in each pair must be a star scientist and both must be from the U.S. (except for one analysis, as indicated in the notes to the table), reflecting our sampling frame and U.S. focus for these analyses.

The first seven variables listed in Table X.1 are indicators of high value in terms of number of discoveries (star), quality of the university program (in biochemistry, molecular biology, and/or microbiology), and high ability to capture rents from discoveries in firms relative to universities and research institutes. Our measures of organizational and geographic location are defined next in the table. We expect more valuable intellectual capital to involve more collaboration within the same organization. The number of times a scientist moves between organizations is expected to widen collaborations to include a larger number of scientists. We define geographic location in terms of the 183 functional economic areas within the U.S. as defined by the U.S. Bureau of Economic Analysis (here called BEA areas). Our diffusion measures examine both the size of a star's collaboration network and the "birth" of new collaborators in the same BEA. Control variables include measures of the size of collaborations, years of entry into genetic sequence publishing, and time trends to capture the declining value of intellectual capital.

For the most part, we select familiar statistical models. But because we are attempting to explain "count" data, that is data that involves counting the number of new collaborators, first from the perspective of each star (CLBNET) and then by BEA (NCOLLAB), we rely on a slightly less familiar poisson regression technique, as suggested by Hausman, Hall, and Griliches (1984). The poisson process is consistent with count variables: non-negative integers, often with significant mass at zero. The poisson process assumes that births of new collaborators in a BEA

area with a probability λ_{it} per unit time. The logarithm of λ_{it} is a linear function of the explanatory variables included in the regression. We estimate these regressions using the LIMDEP package (Greene 1992: 539-49).

Collaboration and Trust Production: We begin by examining the predictors of the number of new coauthors in the stars' coauthor network (CLBNET) in Table X.2. As might be expected from the cosmopolitan/local models of scientific collaboration (Gouldner 1957-58), star scientists that possess higher value intellectual capital generally have larger networks of new coauthors, though the larger the number of collaborators who are also stars (BSTARCLB), the smaller the total number of collaborators (CLBNET). At first glance the strong positive effects of both TQCOLNET and FIRMLAST on the network size seem inconsistent with our argument that higher value intellectual capital is more protected; indeed, even scientists whose last affiliations are with firms (where capturing returns might mean limiting dissemination of the discoveries) have a significant positive effect on network size.

However, these results do not reflect the differential numbers of collaborators drawn from inside the same organization. The number of collaborators from the same or different organizations, SAMEOCLB and DIFFOCLB, both increase the size of collaboration networks significantly, but having more coauthors from different organizations is a much stronger predictor. Also, as expected, the number of times the star scientist moves from one job to the next increases the size of his/her network. All of these effects are measured controlling for the total number of publications by each U.S. star.

Table X.3 shows that over 75% of collaborations with the same type of organization occur within the organizational boundaries, increasing to nearly 95% for firms. Firm scientists very rarely coauthor with scientists at other firms. Most collaboration outside the boundaries of the same organization takes place between university scientists and other scientists located in

research institutes and firms. Since universities are the "source," that is the location of many of the initial discoveries and talent, it is not surprising that both firms and research institutes collaborate frequently with scientists in universities. If we take into account all of these external collaborations regardless of organization type, the balance shifts somewhat: 43 to 44% of all collaborations involving research institutes and firms take place across organizational boundaries, while 64% of collaborations involving university scientists remain inside the organizational boundaries. Overall, our findings counter assumptions about the open structure of scientific discovery across different universities.

What predicts whether any given coauthor pair is found within the organization's boundary or spans it? In Table X.4, we find a strong positive effect of higher value intellectual capital, controlling for collaboration size and time, on the probability that the collaboration pair come from the same organization. Again, we are examining only pairs in which at least one author is a star and at least one author is from the U.S. We find strong support for our hypothesis that intellectual capital of high value--where both coauthors are stars, from top quality universities, or from firms that can capture value better--is more likely to be within the **information envelope** created by working in the same organization. As expected, the larger the collaboration size the less likely that any particular pair of coauthors is from the same organization. Time does not have a significant effect, net of the other variables in the equation.

A minor difficulty in interpreting Table X.4 arises because the three variables BOTHFIRM, BOTHSTARS, and BOTHTQU all inevitably increase the probability that both coauthors in the pair are from the same institution since all potential authors at a different type of institution are by definition excluded.[16] This effect, however is relatively small, and much smaller than required to explain the large estimated coefficients for these variables in Table X.4. To take the most extreme case, there are 18 top quality universities, so if pairs are randomly

assigned there is a probability of $1/18 = .0556$ that they would be from the same university. Compared to the average probability that the authors are from the same organization (0.639), adding the full .0556 (an over-estimate of the increased probability due to this factor) would increase the logit dependent variable from 0.571 to only 0.822. This increase of 0.251 is only a tenth of the size of the coefficient of BOTHTQU. The corresponding estimates of maximum upward bias for the coefficients of BOTHFIRM and BOTHSTARS are 0.0916 and 0.0207, both of which are trivial relative to the size of the estimated coefficients.

Transaction costs provides an alternative -- possibly complementary -- explanation to protection of valuable intellectual capital for the high frequency of collaborations inside organizational boundaries. Generally, both protection is lower and *other* transaction costs are higher for collaborations across organizational boundaries, as compared to within one's own organization. Therefore, all else equal, transactions across organizational boundaries should occur only when there are sufficient benefits to offset the additional costs involved. Looking back at Table X.3, it appears that the reward/cost ratio of transacting with organizations of the same type seldom make it worthwhile. This is especially so for firms, with very rare collaboration with other firms, perhaps in part because of problems concerning property rights that don't emerge when firms collaborate with scientists at universities or research institutes.

But if we examine the average benefits of collaborations in terms of citations that the research receives, occurring within or across organizational boundaries as shown in Table X.5, the data appear to support the hypothesis that transaction costs (including distrust) increase in inter-organizational collaborations: For authors from the same type of organizations, collaborations across organizational boundaries are more highly cited than for those within the same organization, significantly so universities and research institutes. Interestingly, there is a much greater apparent citation payoff for scientists from universities and research institutes who

collaborate with those from firms, consistent with the hypothesis that the difference in cultures further reduces trust, but also possibly due to selection of the best academic scientists for collaborations by the generally more highly cited firm scientists (some empirical support for selection as a significant factor is provided in Zucker, Darby, and Armstrong 1994). Our innovation above is that we can use indicators of the value of the information produced to identify collaborations for which reduced trust has a greater impact on the total transaction costs.

Organizational Boundaries and Diffusion: Our final analysis of the paper is in many respects the most important. The two maps of the geographic location of stars and collaborators in 1980 and 1990 in the U.S., displayed as Figure X.4, provides a clear picture of the rapid diffusion of rDNA techniques. These maps also depict both the dispersion and the local agglomeration of intellectual capital in the biosciences. In Table X.6 the effects of collaboration with trusted others within the same organization on diffusion of research reporting genetic sequence discoveries, measured as the number of new collaborators who coauthor a publication in that year, are estimated within U.S. BEAs. As predicted, coauthoring within the same organizational boundaries (SAMEORG) has a **negative** effect, reducing the number of new coauthors "born" in that BEA, while coauthoring across BEA boundaries generally increases the number of new coauthors "born" in that BEA. These results generally hold whether or not the lagged number of collaborators born in that BEA are entered into the equation, although the positive effect of coauthorship across BEAs (DIFFBEA) is only significant when the lagged dependent variable is included in the equation. The control variable, number of stars in the BEA, has a positive effect on the "birth" of new collaborators. Effective information boundaries are in fact constructed: Collaboration across geographic boundaries, thus between different organizations, increases diffusion to new collaborators, while in geographic areas where a higher proportion of co-author pairs come from the same organization, diffusion to new collaborators

is retarded.

Conclusion

Summary: Scientists who make breakthrough discoveries can receive above-normal returns to their **intellectual capital**, with returns depending primarily on the degree of "natural excludability," so that at the extreme privatizing discoveries requires only selecting trusted others as collaborators. Most often these trusted collaborators are scientists working in the same organization where incentives become aligned based on repeat and future exchange that is monitored and enforced by managers serving as third-parties to the exchange. We find that high value intellectual capital paradoxically predicts both a generally larger number of collaborators and more of that network contained within the same organization. Specifically, same-organization collaboration pairs are more likely when both are highly productive "star" scientists, both are located in top quality bioscience university departments, or both are located in a firm (higher ability to capture returns). Collaboration across organization boundaries, in contrast, is negatively related to the value of intellectual capital, predicted by the number of times the star scientist has moved. As expected, organization boundaries act as **information envelopes**, such that the more valuable the information produced, the more its dissemination is limited.

At the same time, when collaborations do occur across organizational boundaries, they tend to involve higher valued output, as measured by citations. Also, as we predicted, diffusion to new collaborators is significantly decreased when collaborations occur within organizational boundaries, effectively slowing the spread of new techniques and related discoveries.

Implications: There are a number of implications of our argument and results, some of which have already been introduced above. First, belonging to the same organization as a potential coauthor appears to be a powerful and effective means of generating trust:

collaboration within organizations is significantly more likely, compared to collaborating across organizational boundaries, especially when the information produced from the collaboration is potentially of high value (both star, both from top quality university, both in firm). Second, our results suggest that distrust is one of the major costs involved in transacting across organizational boundaries.

Last, and most important, we have evidence that organizational boundaries operate effectively to restrict diffusion. Collaborations across organizational boundaries increase the "birth" of new collaborators in the local geographic area, while collaborations within the same organization significantly retard it. The organizational literature marginalizes organizational boundaries relative to the technical core and the surrounding organizational environment; without making invidious citations, a survey of seven major textbooks on organizations revealed no more than a page or two on organizational boundaries but a chapter or more on the organizational environment. In sharp contrast, we argue that is the organizational boundary, and associated repeat/future interactions and monitoring by managers, that identifies where trust is being actively produced.

REFERENCES

- Bilofsky, H.S., and C. Burks (1988). The GenBank (R) genetic sequence data bank. *Nucleic Acids Research*, 16, 1861-1864.
- Brewer, M. B. and M. Silver (1978). Ingroup bias as a function of task characteristics. *European Journal of Social Psychology*, 8, 393-400.
- Coase, R. (1937). The nature of the firm. *Economica*, 4, 386-405.
- Coase, R. (1988). The nature of the firm: origin, meaning, influence. *Journal of Law, Economics and Organization*, 4, 3-47.
- Cohen, S., A. Chang, H. Boyer, and R. Helling (1973). Construction of biologically functional bacterial plasmids *in vitro*. *Proceedings of the National Academy of Sciences*, 70, 3240-3244.
- Coleman, J. S. (1986). Social theory, social research, and a theory of action. *American Journal of Sociology*, 91, 1309-35.
- Crane, D. (1969). Social structure in a group of scientists: A test of the invisible college hypothesis. *American Sociological Review*, 34, 335-52.
- Crane, D. (1972). *Invisible colleges: diffusion of knowledge in scientific communities*. Chicago, IL: University of Chicago Press.
- Darby, M. R., and E. Karni (1973). Free competition and the optimal amount of fraud, *Journal of Law and Economics*, 16, 67-88.
- Darby, M. R., and J. R. Lott, Jr. (1989). Qualitative information, reputation, and monopolistic competition. *International Review of Law and Economics*, 9, 87-103.
- Dawes, R.M., A. van de Kragt, and R. Orbell (1987). Not me or thee but we: the importance of group identity in eliciting cooperation in dilemma situations. Paper presented at the Public Choice Meeting, Tucson, AZ.

- Eisenberg, R. S. (1987). Proprietary rights and the norms of science in biotechnology research. *The Yale Law Journal*, 97, 177-231.
- Gaston, J. (1973). *Originality and competition in science: a study of the British high energy physics community*. Chicago, IL: University of Chicago Press.
- Gebhart, F. (1983). Quality and consistency improved, the automated gene synthesizer market is on the rebound. *GEN*, 3:1, 10.
- Geertz, C. (1978). The bazaar economy: information and search in peasant marketing. *American Economic Review*, 68, 28-32.
- Gouldner, A. W. (1957-58). Cosmopolitans and locals: toward an analysis of latent social roles. *Administrative Science Quarterly*, 2, 281-306 and 444-480.
- Granovetter, M. (1985). Economic action and social structure: A theory of embeddedness. *American Journal of Sociology*, 91, 481-510.
- Greene, W. H. (1992). *LIMDEP: user's manual and reference guide, version 6.0*, Bellport, NY: Econometric Software, Inc.
- Hausman, J., B. H. Hall, and Z. Griliches (1984). Econometric models for count data with an application to the patents-R&D relationship. *Econometrica*, 52, 909-938.
- Klein, B., R. Crawford, and A. Alchian (1978). Vertical integration, appropriable rents, and the competitive contracting process. *Journal of Law and Economics*, 21, 297-326.
- Kollock, P. (1994). The emergence of exchange structures: An experimental study of uncertainty, commitment, and trust. *American Journal of Sociology*, 100, 313-345.
- Kramer, R.M. and M. B. Brewer (1984). Effects of group identity on resource use in a simulated commons dilemma. *Journal of Personality and Social Psychology*, 46, 1044-1056.

- McCarthy, J. D., and M. N. Zald (1973). *The trends of social movements in America: professionalization and resource mobilization*. Morristown, N.J.: General Learning Press.
- McCarthy, J. D., and M. N. Zald (1977). Resource mobilization and social movements: A partial theory. *American Journal of Sociology*, 82, 1212-39.
- Merton, R. M. (1957). Priorities in scientific discovery. *American Sociological Review*, 22, 635-59.
- Merton, R. M. (1938). Motive forces of the new science. In R. M. Merton, *Science, Technology and Society in Seventeenth-Century England* (pp. 80-111). Bruges, Belgium: Saint Catherine Press.
- Porter, L.W. and K.H. Roberts (1976). Communication in organizations. In M.D. Dunnett (Ed.), *Handbook of Industrial and Organizational Psychology*. Chicago, Ill.: Rand McNally.
- Sato, K. (1988). Trust and group size in a social dilemma. *Japanese Psychological Research*, 30, 88-93.
- Schneider, S. K. (1990). Communications dilemmas: the effects of social categorization on communication networks. Unpublished doctoral dissertation, UCLA. [*Dissertation Abstracts International*, 1990 Feb, v50 (n8-B):3763.]
- Schneider, S. K. and M. B. Brewer (1987). Effects of group composition on contributions to a public good. Unpublished manuscript, UCLA.
- Shapiro, Susan P. (1987). The social control of impersonal trust. *American Journal of Sociology*, 93, 623-658.
- Taubes, G. (1986). *Nobel dreams. power, deceit and the ultimate experiment*. New York: Random House.

Tolbert, P. and L.G. Zucker (1995) Institutional analyses of organizations: Legitimate but not institutionalized In S. Clegg, W. Nord, and C. Harley (eds.), *Handbook of Organizational Theory*. London, England: Blackwell.

Watson, James D. (1968). *The double helix: a personal account of the discovery of the structure of DNA*. New York: Norton.

Williamson, O. (1979). Transaction cost economics: the governance of contractual relations. *Journal of Law and Economics*, 22, 3-61.

Williamson, O. (1991). Comparative economic organization: the analysis of discrete structural alternatives. *Administrative Science Quarterly*, 36.

Zucker, L.G. (1983). Organizations as institutions. *Research in the Sociology of Organizations*, 2: 1-47.

Zucker, L. G. (1986). Production of trust: institutional sources of economic structure 1840 to 1920. *Research in Organizational Behavior*, 8, 53-111.

Zucker, L. G., M. R. Darby, and J. Armstrong (1994). Intellectual capital and the firm: The technology of geographically localized knowledge spillovers. Working Paper No. 4946. Cambridge, MA: National Bureau of Economic Research.

Zucker, L. G., M. R. Darby, and M.B. Brewer (1994). Intellectual capital and the birth of U.S. biotechnology enterprises. Working Paper No. 4653. Cambridge, MA: National Bureau of Economic Research.

Zucker, L. G., and I. G. G. Kreft (1994). The evolution of socially contingent rational action: effects of labor strikes on change in union founding in the 1880s. In J. A. C. Baum and J. V. Singh (Eds.), *Evolutionary Dynamics of Organizations* (pp. 294-313). Oxford, UK: Oxford University Press.

Zuckerman, H. (1967). Nobel laureates in science: patterns of productivity, collaboration, and authorship. *American Sociological Review*, 32, 391-403.

Zuckerman, H. (1977). *Scientific elite: Nobel laureates in the United States*. New York, NY: Free Press.

DATA SOURCES

Bioscan, volumes 3-7, 1989-1993.

CITIBASE: Citibank Economic Database, machine readable data base, 1946-June 1993, New York: Citibank, N.A., 1993.

GenBank, Release 65.0, machine readable data base, Palo Alto, CA: IntelliGenetics, Inc., September 1990.

Institute for Scientific Information, *Science Citation Index*, ISI Compact Disc Editions, machine readable data base, Philadelphia: Institute for Scientific Information, various years through 1993.

Jones, L. V., G. Lindzey, and P. E. Coggeshall (Eds.) (1982). *An Assessment of Research-Doctorate Programs in the United States: Biological Sciences*, Washington, DC: National Academy Press.

North Carolina Biotechnology Center, *North Carolina Biotechnology Center U.S. Companies Database*, machine readable data base, Research Triangle Park, NC: North Carolina Biotechnology Center, April 16, 1992.

[Pre-Bioscan], Biotechnology company data base, predecessor source for Bioscan, provided in printout form by a major biotechnology company, 1988.

Pratt, S. E. (1982). *Guide to venture capital sources*, 6th ed., Englewood Cliffs, NJ: Prentice-Hall, Inc.

U.S. Department of Commerce, Bureau of the Census (1982). *County Business Patterns, 1980: U.S. Summary, State, and County Data*, machine readable data base, Washington: Bureau of the Census. [Ann Arbor, MI: Inter-University Consortium for Political and Social Research, 1986 (distributor)]

U.S. Department of Commerce, Economics and Statistics Administration, Bureau of Economic Analysis (1992a). *National Income and Product Accounts of the United States, Volume 2, 1959-88*, Washington: U.S. Government Printing Office.

U.S. Department of Commerce, Economics and Statistics Administration, Bureau of Economic Analysis (1992b). *Regional Economic Information System, Version 1.3*, CD-ROM, machine readable data base, Washington: Bureau of Economic Analysis, May 5, 1992.

U.S. Department of Education, National Center for Education Statistics, *Higher Education General Information Survey (HEGIS), Institutional Characteristics, 1983-84*, machine readable data base, ICPSR 8291, Ann Arbor, MI: Inter-University Consortium for Political and Social Research, circa 1985.

ENDNOTES

1. This research has been supported by grants from the Alfred P. Sloan Foundation through the NBER Research Program on Industrial Technology and Productivity, the National Science Foundation (SES 9012925), the University of California Systemwide Biotechnology Research and Education Program, the University of California Systemwide Pacific Rim Research Program, the UCLA Center for American Politics and Public Policy, and the UCLA Institute of Industrial Relations.
2. Indeed natural excludability has created some problems for making the "enabling disclosure" that is required for a valid U.S. patent application. In order to obtain the seventeen-year monopoly granted by a patent the applicant must make a disclosure that will enable the public to practice the innovation once the patent expires. After some litigation and legislation, patents are now obtainable by biotech inventors who disclose their invention by placing a culture in a recognized public depository. (See Eisenberg 1987 for a discussion of this history.) Disclosure by deposit eliminates the inherent difficulty in disclosing the art used to obtain the invention so that it can be readily replicated.
3. Very little research has examined trust production in organizational settings; most has focused on the effects of trust once it exists (for an excellent review see Porter and Roberts 1976). Brewer and Silver (1978) found that ingroup members were rated as more trustworthy, honest, and cooperative than their outgroup counterparts. Some laboratory research on small groups, simulating organizations, has found that members of the same organization are more likely to communicate freely with each other, assuming that reciprocity of communication is more likely (Schneider 1990; see also Sato 1988, Kramer and Brewer 1984). But this increased communication occurs only when individuals believe that the benefits are going exclusively to members of their own group (Dawes, van de Kragt, and Orbell 1987).

4. In biotechnology, withholding of information has been acknowledged as a common strategy used by both academic and commercial scientists to "retain the exclusive benefits of a discovery for themselves" (Eisenberg 1987:204). As in many other fields of science, recognition and other financial and social awards accrue to the scientist who solves the problem first. As a result, significant deviation from the norms requiring biologic materials necessary for replication to be made available to other scientists were acknowledged by a majority of ad hoc committee members reviewing the publication policy of the *Journal of Biological Chemistry* (Dr. Donald Brown, Department of Embryology, Carnegie Institution of Washington; Dr. I.S. Johnson, Vice President, Lilly Research Laboratories, Eli Lilly; Dr. Daniel Nathans, Department of Microbiology, The Johns Hopkins University School of Medicine; Dr. Jesse C. Rabinowitz, Department of Biochemistry, University of California, Berkeley; as identified in Eisenberg 1987: 201-202, footnotes 132, 137, and 140).

5. Shapiro (1987) has also outlined some principals of formal structure that increase trust, but her argument is not process-focused and doesn't relate directly to the extension of organization theory proposed in our paper.

6. Explicit agreements, such as joint ventures between organizations, may produce similar trust-enhancing results. We will estimate these equations in later work.

7. It is important to draw a parallel here to economic activity. It is empirically as ubiquitous, but its production is seen as highly variable; when the incentives are not strong, as in the command economies, economic action will be depressed and economic structure will be poorly developed. We need to rethink institutional structure and social action in these terms, and consider developing an explicit demand framework. We do already have the concept of "resource mobilization," developed in social movement research, that recognizes the need for resources to work for social change, effectively creating a demand for institutions of a particular

sort (McCarthy and Zald 1977). The concept of "social capital" may also serve as a basis for revising sociological theory if it is redeveloped along the lines suggested here (see Coleman 1986).

8. By the cost of learning the information, we specifically mean the eventual cost of mastering the information as part of a normal education and/or on-the-job training program of an entering scientist in the field. It is this cost, and not the cost of actually making the discovery that will ultimately determine the returns to the information when equilibrium is reached.

9. In the limit, where the discovery can be easily incorporated into the human capital of any competent scientist, the discoverer(s) cannot earn any personal returns -- as opposed to returns to intellectual property such as patents or trade secrets (see below). In the case of biotechnology, it may be empirically difficult to separate intellectual capital from the conceptually distinct value of cell cultures created and controlled by a scientist who used his or her nonpublic information to create the cell culture.

Besides natural excludability, the value of intellectual capital depends on the novelty and value of the discovery. For example, many discoveries have little or no value--whether monetary or nonmonetary. They are viewed as a dead-ends, of interest only within a few collaborations. The techniques and research program crawl along, diffusing slowly if at all.

10. We used GenBank Release 65.0 which combines data from DNA Data Bank of Japan (Mishima, Japan), EMBL Data Library (Heidelberg, Germany), and GenBank (Los Alamos, New Mexico). See H.S. Bilofsky and C. Burks (1988) for a description of the GenBank.

11. We excluded unpublished manuscripts and papers published in proceedings and in obscure journals in order to obtain consistent information on institutional affiliation and location.

12. Gene splicing machines were initially extremely unreliable. By 1983, the machines had improved in reliability, but still not sufficiently for wide adoption (Gebhart 1983). The high cost

remained an additional barrier to adoption until the late 1980s.

13. See U.S. Department of Education, National Center for Education Statistics (1985).

14. The respondents were asked to rate programs using the following scale: 5 for distinguished, 4 strong, 3 good, 2 adequate, 1 marginal, and 0 not sufficient for doctoral education. The reported scores are the averages among respondents.

15. In informal discussion, Ed Penhoet emphasized the role of diffusion in increasing the number of coauthors over time. As an early innovator in the area, both as a founder and the current CEO of Chiron, he has deep insight into the development of the science over the full time period of our study. We will explore this competing hypothesis more adequately in a companion paper that develops a model of the diffusion process underlying rDNA research.

16. We are indebted to Kerry Knight for this point.

Table X.1: Variable Definitions

Value of Intellectual Capital

BOTHSTAR _i	1 if both scientists in coauthor pair i are stars; 0 otherwise.
BOTHTQU _i	1 if both scientists in coauthor pair i are from top quality universities; 0 otherwise.
BOTHFIRM _i	1 if both scientists in coauthor pair i are from firms; 0 otherwise.
BSTARCLB _j	For each star j, total number of distinct other stars listed as coauthors across all articles.
TQCOLNET _j	For each star j, total number of distinct coauthors who list affiliation as top quality universities across all articles.
FIRMLAST _j	1 if last affiliation of star j was a firm; 0 otherwise.
STAR _{kt}	Number of stars affiliated with an organization in BEA area k in year t.

Location: Organizational and Geographic

SAMEORG _i	1 if both scientists in coauthor pair i list affiliation to the same organization; 0 otherwise.
SAMEOCLB _j	For each star j, total number of coauthors that list affiliation to the same organization.
DIFFOCLB _j	For each star j, total number of coauthors that list affiliation to different organizations.
DIFFBEA _{kt}	Number of coauthor pairs who list affiliations in BEA area k and a different BEA area in year t.
MOBILITY _j	Number of moves by each star j from one organization to another.

Diffusion Measures

CLBNET _j	For each star j, total number of distinct coauthors across all articles.
NCOLLAB _{kt}	Number of collaborators who are not also stars born in BEA area k in year t.
NCOLLAB1 _{kt}	Lagged one year, number of collaborators, stars excluded, born in BEA area k in year t.

Control Variables

CLBSIZE _i	Number of authors of the article from which coauthor pair i is drawn.
TOTPUBS _j	Total number of articles by each star j.
FIRSTYR _j	Year of entry; calendar year of first article by star j; first year is 1967.
YEAR _i	Calendar year in which the article with coauthor pair i was published; first year is 1967.
YEARSQ _i	Calendar year squared; YEAR _i ² .

Table X.2
Poisson Regression on the Star's Collaboration Network,
Distinct Coauthors Across All Articles, U.S. Only
1967-1990

Constant	21.897***	[6.351]
SAMEOCLB	0.004***	[0.0005]
DIFFOCLB	0.008***	[.0004]
BSTARCLB	-0.009**	[.0043]
TQCOLNET	0.005***	[.0007]
FIRMLAST	0.148***	[.0476]
MOBILITY	0.071***	[.0095]
FIRSTYR	-0.010***	[.0032]
TOTPUBS	0.015***	[.0022]

Log Likelihood = -899.706

Notes:

Standard errors in square brackets.

* Parameter significant at the 0.05 level.

** Parameter significant at the 0.01 level.

*** Parameter significant at the 0.001 level.

Table X.3
 Number of Collaboration Pairs Within the Same Organization and Between Different Organizations:
 Universities, Research Institutes/Hospitals, and Firms, 1967-1990

	<u>University</u>		<u>Research Institute</u>		<u>Firm</u>	
	Same	Different	Same	Different	Same	Different
University	2747	771		420		302
Research Institute		420	532	141		111
Firm		302		111	346	17
Totals	2747	1493	532	672	346	330

Notes:

Location based on affiliation given in the first article published by each pair of scientists.

Excludes scientists listing affiliation with more than one institution (N = 16).

Table X.4
 Logit Regression
 Probability that a Pair
 of Coauthors Are from Same U.S. Organization
 1967-1990

Intercept	1.420**	[.597]
YEAR	-0.001	[.007]
CLBSIZE	-0.249***	[.009]
BOTHFIRM	4.707***	[.246]
BOTHSTARS	0.496***	[.088]
BOTHTQU	2.494***	[.068]

Chi-square for -2 log likelihood (5 d.f.) = 3571.36***

Concordance of predicted probabilities & observations:
 Concordant 80.7%, Discordant 16.2%, Tied 3.1%

Notes:

For each coauthor pair, at least one author must be from the U.S.

Standard errors in square brackets.

Results for predicting different organization are opposite signs, but same magnitude for coefficients.

* Parameter significant at the 0.05 level.

** Parameter significant at the 0.01 level.

*** Parameter significant at the 0.001 level.

Table X.5
Average Number of Citations to Genetic Sequence Articles
By Location of Collaboration Pair, 1967-1990

	<u>University</u>		<u>Research Institute</u>		<u>Firm</u>	
	Same	Different	Same	Different	Same	Different
University	17.64	29.01		22.00		49.53
Research Institute		22.00	18.03	26.45		95.60
Firm		48.53		95.60	69.57	64.18

Notes:

Citation counts are for 1982, 1987, and 1992.

Average number of citations is determined by dividing per scientist.

Excludes scientists listing affiliation with more than one institution (N = 16).

Table X.6
Poisson Regressions:
Determinants of Number of Collaborators Born
By BEA Area in the U.S. and Year, 1967-1989

Dependent variable: $NCOLLAB_{kt}$

Independent Variables	Coefficients and Standard Errors	
	Model 1	Model 2
CONSTANT	-26.793* (12.500)	11.429 (12.840)
SAMEORG	-0.012*** (0.002)	-0.003 (0.002)
DIFFBEA	0.005 (0.005)	0.039*** (0.005)
STAR	0.134*** (0.006)	0.121*** (0.006)
YEAR	0.567 (0.301)	-0.375 (0.309)
YEAR ²	-0.003 (0.002)	0.003 (0.002)
NCOLLAB1		-0.032*** (0.002)
Log Likelihood	-3058.96	-2952.64

Notes:

For each coauthor pair, both authors are from the U.S.

Each BEA gets 1/2 birth credit for across BEA collaboration pairs.

Standard errors are in parentheses.

Results for predicting different organization are opposite signs, but same magnitude for coefficients.

* Parameter significant at the 0.05 level.

** Parameter significant at the 0.01 level.

*** Parameter significant at the 0.001 level.

Figure X.1
Total Number of Articles Reporting Sequence Discoveries
Authored by Stars: Total and U.S. only, 1967-1989

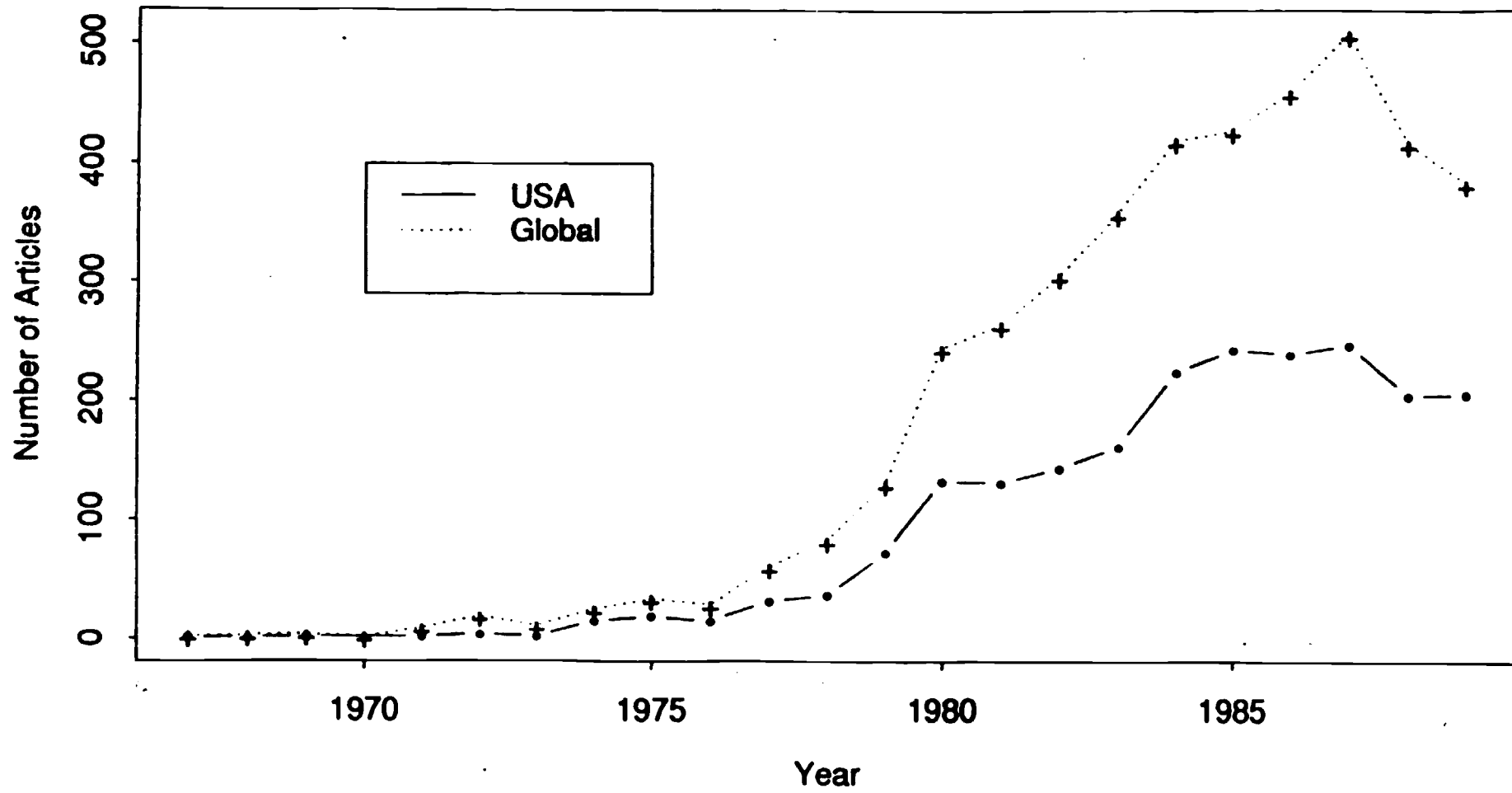


Figure X.2
Average Number of Authors per Article Published by Stars, 1967-1989

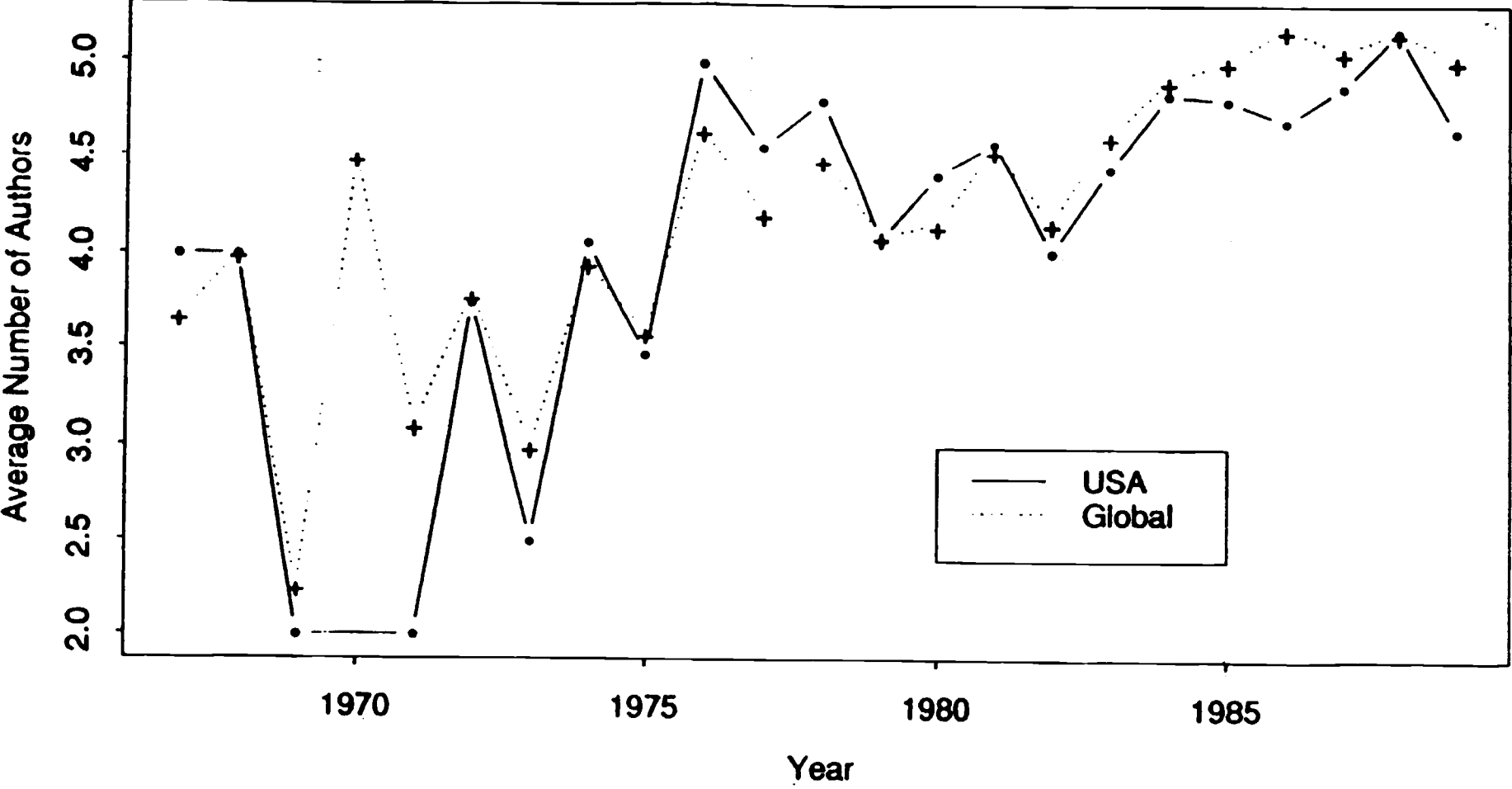


Figure X.3
Percentage of Coauthors from the Same Institution as the Star, 1967-1989

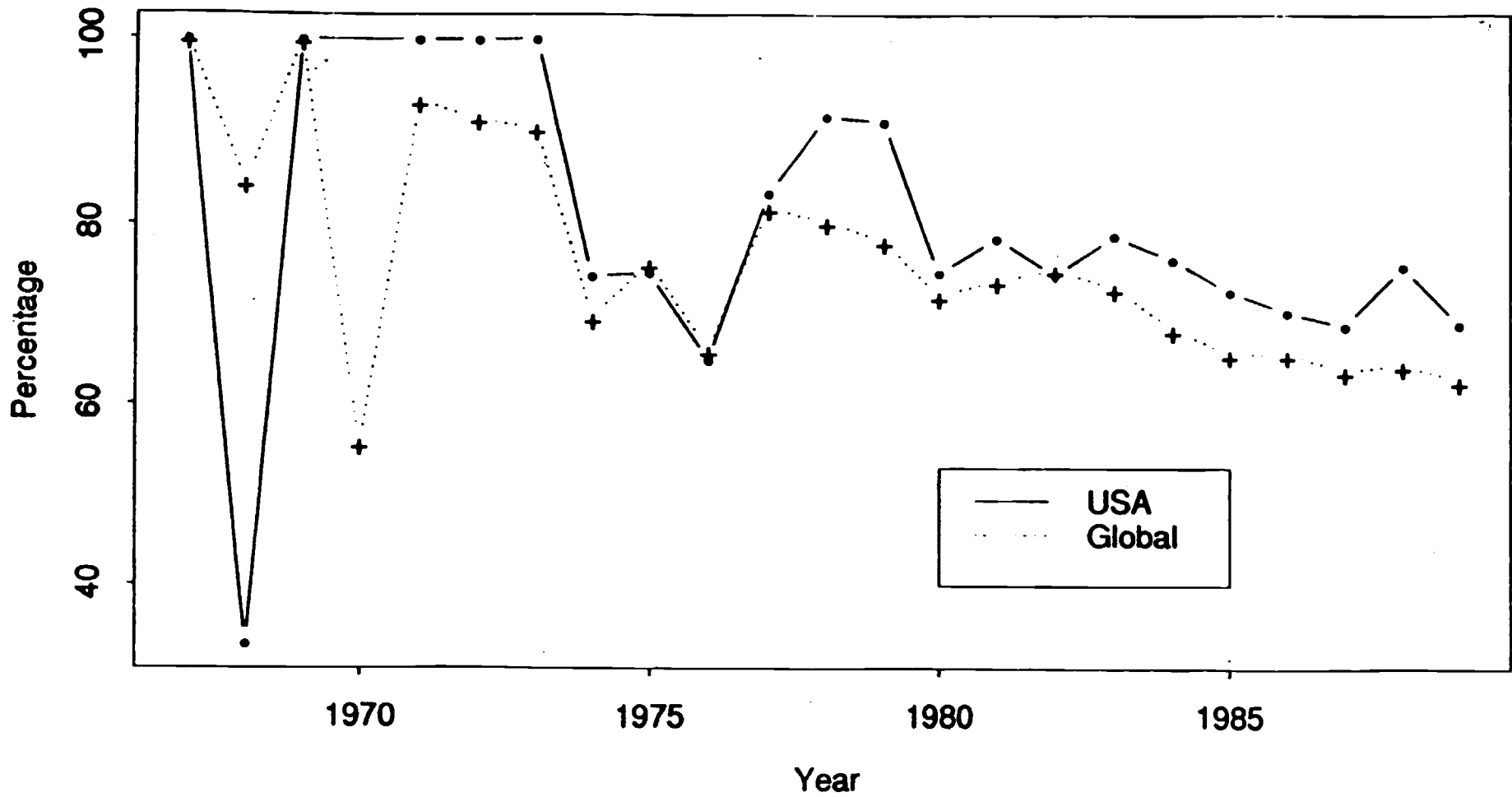


Figure X.4
Cumulative Geographic Agglomeration and Distribution of Intellectual Capital: Stars and Collaborators in the U.S., 1980 and 1990

