

NBER WORKING PAPER SERIES

INTELLECTUAL CAPITAL AND THE
BIRTH OF U.S. BIOTECHNOLOGY
ENTERPRISES

Lynne G. Zucker
Michael R. Darby
Marilynn B. Brewer

Working Paper No. 4653

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 1994

This research has been supported by grants from 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 on earlier drafts from David Butz, Harold Demsetz, Robert Drazin, Martin Feldstein, Zvi Griliches, Adam Jaffe, Benjamin Klein, Josh Lerner, Gary Pisano, Jeff Rosensweig, L.G. Thomas, and others. We are indebted to a remarkably talented team of post-doctoral fellows Zhong Deng, Julia Liebeskind, and Yu Sheng Peng and research assistants Paul J. Alapat, Jeff Armstrong, Lynda J. Kim, Amalya Oliver, Alan Paul, and Maximo Torero. 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.

INTELLECTUAL CAPITAL AND THE
BIRTH OF U.S. BIOTECHNOLOGY
ENTERPRISES

ABSTRACT

We examine the relationship between the intellectual capital of scientists making frontier discoveries, the presence of great university bioscience programs, the presence of venture capital firms, other economic variables, and the founding of U.S. biotechnology enterprises during 1976-1989. Using a linked cross-section/time-series panel data set, we find that the timing and location of the birth of biotech enterprises is determined primarily by intellectual capital measures, particularly the local number of highly productive "star" scientists actively publishing genetic sequence discoveries. Great universities are likely to grow and recruit star scientists, but their effect is separable from the universities. When the intellectual capital measures are included in our poisson regressions, the number of venture capital firms in an area reduces the probability of foundings. At least early in the process, star scientists appear to be the scarce, immobile factors of production. Our focus on intellectual capital is related to knowledge spillovers, but in this case "natural excludability" permits capture of supranormal returns by scientists. Given this reward structure technology transfer was vigorous without any special intermediating structures. We believe biotechnology may be prototypical of the birth patterns in other innovative industries.

Lynne G. Zucker
Department of Sociology
University of California, Los Angeles
Los Angeles, CA 90024-1551

Michael R. Darby
Anderson Graduate School of
Management
University of California, Los Angeles
Los Angeles, CA 90024-1481
and NBER

Marilynn B. Brewer
Department of Psychology
Ohio State University
1885 Neil Avenue
Columbus, OH 43210-1222

The American biotechnology industry has developed rapidly from nonexistent to about seven hundred active firms in less than two decades. We believe that the birth of this industry is essentially intertwined with the development of the underlying science in a way which has not been previously considered for this industry or any other in either the economics or the organizations literature. Indeed we present here strong evidence that the timing and location of new biotech firms (NBFs) and new biotech subunits of existing firms (NBSs) are primarily explained by the presence at a particular time and place of scientists who are actively contributing to the basic science as represented by publications in major academic journals.

The novelty of our approach lies first in treating the "intellectual capital" of these scientists, which we shall discuss at length below, as the crucial resource for founding new biotech enterprises (NBEs = NBFs + NBSs). In the mainstream economic literature, where the timing and location of firm births is related at all to localized resources, it is normally either at the country level as in the doctrine of comparative advantage or in terms of natural resources as in the discovery of gold deposits. Comparable to measures for traditional economic inputs, we develop detailed quantitative measures of intellectual capital relevant to biotechnology. Here we show that localized intellectual-capital resources were key for the development of at least the biotech industry. This quantitative approach differs methodologically as well as conceptually from the few case studies that have attempted to link universities to new commercial enterprises.¹

Our focus on intellectual capital is related to but goes beyond a related body of work

which focusses on knowledge spillovers. Zvi Griliches (1992) has recently done a masterly survey of that body of work and its potential importance in understanding continued technical progress. We cannot compete with his effort, but we do want to focus on few recent papers which we believe are related to and (implicitly) supportive of the approach we take here. Adam Jaffe (1989) and Jaffe, Manuel Trajtenberg, and Rebecca Henderson (1992) have exploited patent data to show that there is a very significant, geographically limited connection between university and other research and the application of that research to patentable innovations.² Edwin Mansfield (1992, p. 17) confirmed in a survey of 70 corporations that a "substantial portion of industrial innovations in high technology industries ... have been based directly on recent academic research, although in many cases the invention itself did not stem from the universities." Furthermore, he finds that "the extent to which a university is credited by firms ... with making major contributions to these firms' innovations tends to be directly related to the quality of the university's faculty in the relevant department (according to the NAS ratings), to the size of its R and D expenditures in relevant fields, and to the proportion of the industry's members located nearby." The idea of intellectual capital provides a specific channel through which knowledge spillovers were embodied and returns to them captured in the case of biotechnology. Michael Darby (1984, 1993) has emphasized the importance of technological advances embodied in the quality of the labor force, but intellectual capital will be seen to differ from that concept too because it involves supranormal returns to education.

A second innovation in our approach is the use of panel data to analyze the development of the industry: The entire United States is divided into the 183 functional

economic areas (BEA areas) as defined by the U.S. Bureau of Economic Analysis. Much of our data is available and coded at the zip-code or county level by year and then aggregated to the BEA level. While we exploit the finer geographic detail in some analyses, our primary focus in this paper is explaining the number of firms born in each BEA area each year from the beginning 1976 through 1989. The use of panel data permits us not only to better answer questions about the sources of births of biotech firms, but also to address some fundamental questions of educational and scientific policy and regional development.

The next section presents our conception of intellectual capital and describes how we implemented it empirically for biotechnology. Other factors which may also influence the founding of biotech firms are introduced in Section II. Our empirical results on firm births are reported in Section III. We conclude and summarize the main body of the paper in Section IV. A Data Appendix is included to describe the extensive work required to prepare the data set for this paper.

I. Intellectual Capital and Natural Excludability

In this section we first define intellectual capital as distinct from intellectual property, information, and human capital. We then describe our empirical implementation of this concept to develop measures of intellectual capital applicable to biotechnology. In particular we identify a group of "star" scientists who are endowed at particular times with intellectual capital and their collaborators who may also become endowed with intellectual capital. We then describe the process by which intellectual capital was created and diffused in the 1970s and 1980s and present substantial evidence that our implementation of the intellectual capital concept has predictive value for explaining the formation of NBEs.

I.A. The Concept of Intellectual Capital

In economics information is often treated as a public good -- freely available to an unlimited number of potential users at a zero marginal cost. In order to provide incentives for the creation of information, and particularly for commercially valuable scientific research, a system of property rights -- patents, trade secrets, and the like -- has been created so that the creator of information can control the use of his or her intellectual property. The international protection of intellectual property is of such importance as to be a key element in the current Uruguay round of the GATT negotiations.

A complementary literature recognizes that some information requires an investment of considerable time and effort to master. The human capital developed by this investment is seen as earning a normal return on the cost of the investment, both direct costs and

foregone earnings. The average level of human capital embodied in a country's labor force should be an important factor in determining whether the nation's production is concentrated in high-wage, high-skill industries or lower-wage, lower-skilled industries. However, we do not expect a produced factor of production like human capital to be a restrictive endowment such that particular kinds of firms arise where it is present and are absent otherwise.

As we use the term, intellectual capital incorporates elements of both intellectual property and human capital but is distinct from both. A person possesses intellectual capital if he or she embodies a specialized body of knowledge which enables the individual to earn supranormal returns on the cost of obtaining that knowledge. How would this come about? Consider first the discoverers of a radically new method for doing commercially valuable research. Doubtless their discovery will lead to the creation of intellectual property as for example did the 1973 discovery by Stanley Cohen and Herbert Boyer of the basic technique for recombinant DNA (rDNA) which ultimately led to today's biotechnology industry.³ However, in frontier science it is sometimes the case that the methodology actually required to do the science has to be learned first-hand in the laboratories in which it is done.⁴ One may need to license relevant patents, but scientists who have the ability to actually do the science will initially be able to earn extraordinary returns. Some scientists will choose to realize those returns by using their intellectual capital to create commercially valuable products. Others may pursue more academic rewards, such as career acceleration, sometimes the only path to realize intellectual-capital returns.

Techniques which must be learned first-hand through collaboration or apprenticeship

may be described as possessing "natural excludability" when they are first discovered. Over time as the techniques become widely diffused and their use represents routine rather than frontier science, all major graduate schools come to have the capacity to teach the techniques as part of the normal educational process of creating human capital. Thus our intellectual capital has transient value until the "quasi-rents" to those who embody the specialized body of knowledge is driven down to a normal return. We abstract some evidence below from Lynne Zucker, Marilynn Brewer, and Darby (1994) which is consistent with scientists taking informal actions, particularly when the information is most valuable, to slow its diffusion.⁵

Intellectual capital is by its nature a transient property of disequilibria. It arises because few people have mastered certain valuable techniques relative to the number that will master those techniques and earn normal returns on that mastery when full equilibrium is achieved. It is true that the supranormal returns during this period are analogous to lottery winnings and do not imply that supranormal returns -- given the considerable talents and human capital of these individuals -- are expected *ex ante*. Rather our point is that the winners who make certain frontier discoveries or very early learn their techniques will command enormous wages in the early days of commercialization and the location and timing of firm formation will be determined by where and when they are active.⁶ While this is a transient condition of disequilibrium for the labor market, industry creation is itself a coincident disequilibrium process.

Economists and organizational theorists are not likely so much to object to the theoretical construct of intellectual capital presented here as to doubt that it could be of

both sufficient value and persistence to play an important role in the development of a major industry. Furthermore, Karen Seashore Lewis, Blumenthal, Michael Gluck, and Michael Stoto (1989) argue that scientific entrepreneurship leading to commercialization is not distributed evenly across universities so that whether or not a given scientist is likely to start a firm depends upon where he or she is located. Both of these objections to focussing on the scientists who embody intellectual capital come down to empirical questions which we believe are answered convincingly in the rest of this paper for at least the biotech industry.

We hope that our work on the role of intellectual capital in the formation of biotechnology firms will ultimately help us to understand the creation of firms in other very highly technological industries. Harold Demsetz (1988) has argued that the costs of mastering specialized bodies of knowledge play a central role in delimiting the boundaries of the firm. Here we see that at least in one case the very existence of a firm requires access to the very frontier of a science. Whether there are other industries in which intellectual capital plays an equally important role remains the subject of future research.

I.B. Measuring Intellectual Capital in Biotechnology

Broadly enough defined, biotechnology has been used as long as people have baked bread and drank wine. Cross breeding of animals and growing penicillin are other examples of biotechnology. Since the mid-1970s the biotechnology industry has been driven by advances in the biological sciences, particularly in fields such as genetic engineering and cell hybridization. Accordingly today biotechnology refers principally to the application of

genetic engineering based upon taking a gene from one organism and implanting it in another (rDNA) and production of the outcome of this process. The other basic technology is cell fusion in which lymphocytes are fused with myeloma cells to create rapidly proliferating antibody-producing cells.⁷

A very important measure of research success is the discovery of nucleotide sequences that determine the characteristics of proteins and other molecules. In the earlier stages of the project, GenBank was used to identify all articles reporting gene sequence discoveries up to 1990.⁸ The rDNA technology devolved from high to routine science in the late 1980s: for example, a gene sequence discovery was no longer the basis for a Ph.D. dissertation at a major university by the end of the decade. So 1990 is a good year to end the literature base for identifying scientists possessing intellectual capital.

First, as detailed in Data Appendix A.1, we identified 337 leading researchers whom we termed "stars" on the basis of the number of genetic sequence discoveries reported up to 1990 for which they were an author and the number of such articles. These 337 stars were listed as authors on 4,315 distinct articles in major journals. These articles were hand collected (see Data Appendix A.2) to identify and locate institutional affiliations at the time of publication for each of our stars and their coauthors.

Because research discoveries frequently occurred through teams, another 7,718 scientists who were coauthors with a star but who had not themselves met the star criteria were labeled "collaborators." The left side of Table 1 provides some characteristics of those 4,196 stars and collaborators who were located in the United States at the time of their first publication. The average citation data (see Data Appendix A.3) indicate that scientists in

firms are not necessarily inferior to their academic colleagues. In the absence of tenure pressures, it would not be surprising to see fewer "minor" publications produced by firm scientists. Nonetheless, it is surprising that the total number of citations for all gene-sequence-reporting publications should be higher for scientists located in firms than those located in universities or in research institutes and hospitals. While the firm-born star citations are dominated by a single individual, there are 390 firm-born collaborators who similarly dominate their university and institution colleagues in citation rate.

Discussions with scientists in the field indicate that simple co-authorship may not indicate that the individual in question had acquired substantial mastery of the rDNA technology. Accordingly, we developed a screen to focus on those scientists who are actively working -- presumably with mastery of the then-current state of the art -- at a particular time: a star or collaborator is "active" in any year in which he or she has published three or more articles in the three-year period ending with that year. This is a substantial screen, only 134 of the 213 U.S.-born stars are ever active in the U.S. while only 12.5 percent (498 out of 3983) U.S.-born collaborators are ever active in the U.S.⁹ Note that some of the stars and collaborators "born" in the United States may have returned to their native countries after training and published sufficiently to be "active" there. Restricting our attention to these ever-active individuals does not eliminate the higher citation rate for scientists "born" in firms: While we might not put much weight on the single ever-active firm-born star, the 46 ever-active firm-born collaborators are as productive of citations as our ever-active university-born stars!

Intellectual capital is rather concentrated in location. In the United States, only 263

distinct organizations were given at any time as affiliations of any of the stars or collaborators in our articles data set. These organizations comprised 141 universities, 74 research institutes and hospitals, and 48 firms. Of course, scientists may have secondary affiliations such as service on scientific advisory boards with a number of other organizations which would not be reflected in our data set.

There is some ambiguity as to whether intellectual capital is best conceptualized simply as mastery of a particular specialized body of knowledge or whether it also involves the individual's capacity to utilize that knowledge. Obviously, anyone who meets our active criterion has some substantial ability to apply the rDNA technology. Nonetheless, we plan to use our citation data in future research to test whether it is enough to know how the technology is applied or instead whether the interaction of "genius" with know-how is the key determinant to where and when biotech firms will be born.

I.C. The Diffusion of Intellectual Capital among Stars and Their Collaborators

After the 1973 discovery, rDNA technology diffused among the stars and their collaborators primarily by bringing the new people in to work in labs where rDNA technology was being utilized. We believe that doing so was consistent with optimizing behavior by the individuals endowed with the intellectual capital. We plan to report on that process in detail separately, but here will simply indicate how the diffusion occurred.

One way to see how the diffusion process occurred is to see how few articles even our stars wrote reporting sequence discoveries each year from 1967 until 1977 when the process began to take off (Figure 1). In the early years, the technique appears to have been

at least informally bounded by excluding coauthors from other institutions (Figure 2); as time goes on and the know-how declines in value, it becomes more likely that research teams will include scientists from other institutions. The logit regression in Table 2 both (a) confirms the declining probability over time of publishing with same-institution authors even when other factors are held constant and (b) indicates that researchers from firms are especially likely to write with coauthors from their firm in comparison with other authors.¹⁰ Similarly, we see in Figure 3 that the number of coauthors rose over time as value of know-how fell.

Restricting our focus henceforth to only scientists and firms located in the U.S., we plot in Figure 4 the cumulative number of stars, collaborators, and new biotech enterprises with location and birth date of scientists defined according to their first publication and founding dates for firms present in the United States as of 1990.¹¹ We can see that the number of stars began to grow rapidly some years before either collaborators or firms.¹² The map (Figure 5) illustrates that firms seemed to follow stars not only in time but also in place: The location of biotech enterprises in 1990 is well predicted by the areas in which any of our stars first became "active" in the period 1976-89. We will expand on these simple correlations below after including other variables which might also be important, but the basic message will remain. When and where star scientists are active, firms are likely to be born.

I.D. Universities and the Production of Intellectual Capital

It is hardly novel that important discoveries are disproportionately made at great

universities and diffused there from the original discoverers to colleagues and students. This is true for biotechnology as well. In the United States, this incubator role of the university is clearly indicated in Figure 6 where initially all the active stars and their active collaborators are to be found in universities, while by the end of the period this share falls below one half for both stars and their collaborators.

University faculties and administrators make fundamental investment decisions -- to grow a certain area (e.g., biochemistry) or to "go after" a top scientist -- and these decisions at least early in the discovery process have major impacts on the direction and development of science, including discovery rates. The faculties of great universities make these decisions better because they are selected and rewarded on the basis of their individual records of demonstrated productivity and the scientific intuition or "taste" that guides their research and teaching as to what is really important.

Although we have been able to measure intellectual capital in the rDNA field more directly by looking at research output, we expect that the presence of research universities, and especially great research universities, should also be separately entered into the equation explaining firm births for two reasons: First, to the extent that our intellectual capital measures imperfectly capture all the scientists with the requisite rDNA know-how, the university variables should improve the explanatory power of the equation along standard specification analysis lines. Second, the biotech industry has substantial components which do not require knowledge of gene splicing, e.g., growing proteins which have been created by others or producing bases used by the gene splicers. These ancillary technologies also require different forms of intellectual capital which we measure indirectly

by the universities capable of producing these other forms. Obviously, it is an empirical question as to whether our empirical measures of rDNA intellectual capital will continue to enter in the presence of university variables or whether the university variables will have any independent explanatory power in the presence of our measures.

For empirical purposes, we define biochemistry, cellular/molecular biology, microbiology as "biotech-relevant." With reference to just these biotech-relevant departments, we used two different types of measures for research universities: QUAL1, which is the number of universities in a BEA with one or more most highly rated programs in Jones et al. and FEDGRANT which is the total number of faculty supported by 1979-80 federal grants to all universities in each BEA for biotech-relevant research.¹³ (See Data Appendix A.4 for details.)

II. Other Factors Affecting Firm Births

The basic optimizing rule for firm birth is that a potential founder will start a firm if and only if the net present value of the discounted expected cash flows is positive. In fact, we could in principle expand on that simple insight following Boyan Jovanovic (1982) to allow for an entrepreneur's choosing an optimal time to start a firm given a range of times for which there are positive discounted expected cash flows. Given the limitations of the available data, we do not believe very refined models are likely to be productive. Instead we pursue a simple "at risk" modelling approach which views the following agents in a BEA area as potentially starting a NBE: scientists with intellectual capital located in the area, venture capital firms located in the area, pharmaceutical firms operating in the area, and other economic agents active in the area. Thus at any time t the instantaneous probability p_i of a firm founding in area i is determined by the numbers of these agents in the area, and their respective probabilities of founding an NBE. Over finite periods, like years, the number of firms actually founded will follow a poisson process.¹⁴ In this section we will discuss and motivate measures of each of the remaining agents and factors affecting their probability of firm founding. The next section will report on the empirical results.

II.A. Venture Capital Firms

The development of the U.S. biotech industry has largely been financed in the initial stages by venture capital firms. These firms serve as intermediaries putting together private financing for new and young firms. In doing so, the venture capital firms lower the cost of

financial capital to the new firms during the lengthy period of research before there is either product or intellectual property with significant market value.¹⁵ These firms monitor their investments for themselves and their partners, and in doing so provide management expertise to their client firms (see Lerner 1993b).

We believe that the existence of the venture capital industry in America has had a significant effect on the development of the biotech industry. In the U.S. 74 percent of NBEs were started as independent firms compared to only 5 percent in Japan where there is no significant venture capital industry. The difference, as we see it, is that in the U.S. there is the choice of starting a new firm or starting a new subunit within an existing firm while the latter is effectively the only choice in Japan. Since where there is the option, new firms are the dominant form of business, we conclude that this difference may be a significant constraint on Japanese biotech firms. In principle, the existence of venture-capital financing in the U.S. may have permitted NBEs to have been founded sooner in the U.S. and possibly closer to the intellectual capital resources than would otherwise be the case. This systemic argument is the main one for an important role for venture capitalists.

Japanese participants in the biotech industry believe that the interaction between the lack of venture capital firms willing to finance start-ups and the effective prohibition of initial public offerings without a track record of substantial profits and revenues makes it impossible for Japanese scientists to start firms as is done in the U.S. and thus to capture fully returns to their intellectual capital (see Zucker and Darby 1993). It should be noted that there are alternative explanations for the predominance of firms in the U.S., such as our predilection for tort litigation so that existing firms may have greater reason to avoid

risky new technologies and their associated product liabilities.

Quite another argument is frequently made: that venture capitalists are the key catalysts for the birth of NBFs and that in areas where they are absent no firms will be formed while where they are present we will see NBFs formed.¹⁶ A weaker form of the same argument is that there is not an efficient national venture capital market; so the cost of financial capital is much higher to start-up firms in areas where venture capitalists are absent or not subject to vigorous competition. Because of the hands-on role of the venture capitalist, it is argued, they are not likely to fund operations more than several hours from their offices. The higher is the cost of financial capital in a particular area, the less likely is it that a potential entrepreneur will compute a positive discounted expected cash flow from founding a new biotech firm.

Since the venture capital industry existed during the entire time the biotech industry was developing we cannot in this paper test the systemic hypothesis for its importance. We plan to do so, however, in future cross-country comparative work. We can and do test here whether the number of venture capital firms in a BEA area $VENTCAP_i$ has a significant positive effect on the rate at which NBFs are founded. If it does, it is evidence for the absence of an efficient national venture capital market. If it does not, then we can conclude that there is an efficient national market and that venture capital funds will flow to firms of good potential regardless of where they are founded. By way of analogy, we might ask whether great universities need to be founded in areas rich in good book stores or whether such book stores will be induced wherever there are great universities. Stretching the analogy further, the systemic argument is about whether it is a significant impediment to the

development of universities if private book stores are not permitted and the universities must found their own.

A caveat is in order at this point. Just as bookstores spring up when a university opens, so too venture capital firms may arise in areas where there is new business to be done. Hence, we should either restrict ourselves to looking at those venture capital firms which existed before the birth of the biotech era or else use a simultaneous equation method. The latter is problematic in the case of poisson-process estimation. As explained in Data Appendix A.7, we have chosen to try to limit the simultaneity problem by freezing our data on venture capital firms after 1981 when the main burst of biotech births begin.¹⁷

II.B. Other Economic Agents

We would like to identify other economic units at risk for starting NBEs. We believe that pharmaceutical firms are the remaining agents most likely to start NBEs in the United States. To date, primarily because of disclosure restrictions on government statistical agencies, we have not found an acceptable measure of the presence of pharmaceutical firms in a BEA area. We hope to add such a PHARM_{*t*} variable in future research.¹⁸

Having accounted for individuals possessing intellectual capital and venture capital firms -- and unable to measure pharmaceutical firms -- it is difficult to identify other potential firm founders, but we also include total employment EMP_{*t*} and average earnings per job EJOB_{*t*} as measures of other at-risk units not accounted for by our prime categories.

II.C. Carrying Capacity and Population Ecology

In Section I, we related measures of intellectual capital to NBE_{it} , the number of new biotech enterprises in particular BEA areas and years. The primary purpose of this paper is to explain the birth of those enterprises or $BNBE_{it}$. The measurement of these and related concepts is detailed in Data Appendix A.6.

Once the commercial usefulness of biotechnology is appreciated, the net present value of a prospective NBE should be a decreasing function of the number of enterprises in existence when the period begins. Since the competition in the output market is at the national level, it is appropriate that we use the national total of enterprises in the industry. Hence, we would expect the lagged stock of new biotech enterprises $NBE_{i,t-1}$ to have a negative effect on the birth of NBEs in year t . Note that the national aggregate is indicated by the absence of the subscript i .

Although we can be confident that at the national level the effect of the lagged stock should be negative reflecting competition in the product market, there also may be a separate, ambiguously signed effect on the birth of NBEs in area i from the lagged stock of new biotech enterprises in that particular BEA area ($NBE_{i,t-1}$). We have not found previous analyses that have considered the two effects, national and local, separately.¹⁹ For the local effect, on the one hand, the more firms there are locally the more local competition exists for inputs. But in the spirit of Adam Smith's dictum that "specialization is limited by the extent of the market" or the geographer's concept of economic agglomeration, the existence of a local biotech industry may lower the cost of goods and services to the firm. Therefore, whether $NBE_{i,t-1}$ has a positive, zero, or negative effect is strictly an empirical question.

Endogenous population dynamics are the cornerstone of population ecology models of birth. Recent empirical studies rely, sometimes exclusively, on variables that describe the population at an earlier time point (for an excellent review, see Singh and Lumsden 1990). For example, in models of labor union birth at the national level the authors include the number of national labor unions in the prior year and its square, the number of births in the prior year and its square, plus period effects (Hannan and Freeman 1989, pp. 201-224; Hannan and Freeman 1987). Other authors have introduced some other variables, but the focus has remained on lagged births and stock of firms.²⁰

On our view, it is far better to explain the forces determining births in a particular area and year than assume that the forces are pretty much like last year as is implicit in the use of the lagged dependent variable (see Zucker 1989). Below we do run some models with the lagged dependent variable for comparison purposes, but we believe that their inclusion is both theoretically and econometrically suspect.

II.D. Other Factors

Two other factors are considered as determinants of the birth of new biotech enterprises:²¹ the S&P500 earnings-price ratio $E/PRATIO_t$ and a time trend $YEAR_t$. The S&P500 earnings-price ratio is a natural measure of the all-equity cost of capital in the economy and hence should enter negatively as a determinant of births. In some regressions we include a time trend not so much as an explanatory variable as a check on the significance of omitted variables which vary with time. Note that these variables are national observations and thus cannot enter when we do purely cross-section analyses.

In earlier studies such as Mansfield (1962) industry profitability, capital requirements, and minimum efficient firm size were important determinants of entry and exit.²² We do attempt below through period interactions to capture the declining profitability to scientists of starting a firm as the technology diffuses, but firm profitability as conventionally measured is essentially nonexistent throughout this period even as entrepreneurs made fortunes taking their "unprofitable" firms public. We have no way to measure capital requirements or minimum efficient firm size but see no obvious reason for there to have been significant change in either of these parameters.

III. Empirical Results

A formal model is premature if not presumptuous at this point. In what follows we generally enter the variables identified in Sections I and II into poisson or OLS regressions on $BNBE_{it}$, the birth on NBEs in area i in year(s) t . These regressions are estimated using the LIMDEP package.²³ As discussed in Jerry Hausman, Bronwyn H. Hall, and Griliches (1984), the poisson process is the most appropriate statistical model for count data such as we are dealing with here. The poisson process assumes that births occur in a BEA area with a probability $LAMBDA_{it}$ per unit time. The logarithm of $LAMBDA_{it}$ is a linear function of the explanatory variables included in the regression. This process is consistent with the dependent variable which is a non-negative integer with significant mass at zero. OLS regressions are not really suitable for the structure which we are estimating, but are included at times for comparison purposes.

Active stars and their collaborators are denoted by $ACSTAR_{it}$ and $ACCOLL_{it}$, respectively. Squared values of these variables are $ACSTAR_{it}^2$ and $ACCOLL_{it}^2$, respectively.

We also ran the same regressions for NBFs and NBSs defined both exclusive and inclusive of the arguable case of joint ventures. The results were generally very similar for the subcomponent regressions and accordingly are not reported or discussed here except in Table 8 in connection with discussion of the venture capital variable.

III.A. The Long-Run Model

Because of concerns about possible simultaneity biases once the industry became a significant economic force, we begin our empirical discussion with models which relate the number of firms in each BEA area at the beginning of 1990 to the distribution of intellectual capital and other variables as of 1980. These results provide something of an acid test of our approach.

In Table 3, we present cross-section poisson regression across the 183 BEA areas explaining $NBE_{i,90}$, the number NBEs at the beginning of 1990 when our data set ends. As explanatory variables for the logarithm of the poisson parameter we use: $ACSTAR_{i,76-80}$, $ACCOLL_{i,76-80}$, $QUAL1_i$, $FEDGRANT_i$, $VENTCAP_{i,80}$, $EMP_{i,80}$, $EJOB_{i,80}$, and $NBE_{i,80}$, where the variables with 76-80 subscripts are defined as:

$ACSTAR_{i,76-80}$ the cumulative number of stars "active" at any time during the years 1976-1980 in area i

$ACCOLL_{i,76-80}$ the cumulative number of their collaborators "active" at any time during the years 1976-1980 in area i

The other variables are values for or about 1980. The earnings-price ratio had to be dropped for these analyses because it is available only nationally over time.

The first column of results restrains the analysis to only stars and collaborators, while the second column adds in the other intellectual capital variables. The third column adds in the number of venture capital firms, the fourth adds employment and average earnings by area, and the fifth adds those births which occur during the very early years of the industry. Since those births also appear in the variable to be explained by the regression,

Table 4 reports the same regression with the dependent variable $NBE_{1981-90}$ instead of NBE_{1990} , where the former variable excludes those firms alive at the beginning of 1990 which were born before 1981. The results are essentially identical which reflects the fact that bulk of new biotechnology enterprises were founded after 1980.

In the fourth column (and generally before, where applicable) the four intellectual capital variables all enter significantly. The coefficient on collaborators is negative, however. The effect of collaborators is unstable in other regressions both reported below and unreported. We believe that this may reflect the fact that training collaborators, a useful and rewarded thing particularly for the academic stars, may take more of the stars' energy than it is worth if firm birth were the only criteria. We shall see below that this is particularly the case for the earlier years of the industry.

In the fifth column, the only significant effects are (a) to lower the overall size of coefficients as generally occurs with lagged dependent variables,²⁴ and (b) to eliminate any significant effect for the number of great university programs and the number of venture capital firms in the area. As to the latter, it appears that the proto-industry's geographical distribution in 1980 is so highly correlated with $QUAL1$, and $VENTCAP_{1980}$ that the latter variable has no significant independent effect. We should note that the properties of poisson regressions in the presence of lagged dependent variables are not well worked out and these results should be approached with caution.

The significant negative sign of the $VENTCAP_{1980}$ variable in columns 3 and 4 is the most surprising result in Tables 3 and 4. As we discussed above, there is a significant body of literature which argues that the presence or absence of local venture capital firms is the

key determinant of where the biotech industry developed. We accordingly expected a positive effect if this literature is correct or an insignificant effect if the venture capital market is "efficient" in the sense that venture capital is allocated to its highest valued uses regardless of the location of venture capital firms. As of this time, we have come up with only one substantive and one statistical explanation for why venture capital firms would have a negative effect: Where venture capital firms are numerous, firms may tend to be larger and better capitalized using up more stars than in those areas where venture capital is scarce.²⁵ Statistically, columns 5 of Tables 3 and 4 indicate that multicollinearity between $QUAL1_t$ and $VENTCAP_{t,80}$ may be affecting the estimates. In future research, we shall examine aspects of firm development other than birth which should shed more light on this apparent anomaly.

If the $VENTCAP_{t,80}$ variable is entered with only the economic variables (that is, with the intellectual-capital variables excluded) then a significant positive coefficient appears:²⁶

$$\log LAMBDA_{it} = -2.056 + 0.014 VENTCAP_{t,80} + 25E-8 EMP_{t,80} + 15E-5 EJOB_{t,80}$$

(.2712)
(.0040)
(37E-9)
(13E-6)

Dep. Var. $NBE_{t,80}$ Log-likelihood = -864.0

$$\log LAMBDA_{it} = -2.309 + 0.015 VENTCAP_{t,80} + 24E-8 EMP_{t,80} + 15E-5 EJOB_{t,80}$$

(.2985)
(.0044)
(41E-9)
(15E-6)

Dep. Var. $NBE_{t,81-90}$ Log-likelihood = -754.3

In Section IV.B below, we report very similar results in a cross-section/time-series context. We also note that in the OLS regressions in Tables 5 and 6 $VENTCAP_{t,80}$ appears to be positive. Thus, it is certainly easy to see why the evidence for the importance of venture

capital may have appeared stronger in previous work than seems warranted based on fuller models and more appropriate estimation techniques.

Turning to Tables 5 and 6 briefly, we note again that we believe that the poisson regressions are more appropriate than the OLS regressions presented here. Nonetheless, the results are qualitatively fairly similar except for the venture capital variable. The explanatory power of these regressions is also impressive judging from the adjusted R^2 values.

In conclusion, the intellectual capital variables play a strong role in determining where the U.S. biotech industry developed during the 1980s. This strong role is most robust for the variables which are related to particular individuals (counts of stars, their collaborators, and scientists supported by federal grants) in comparison with the reputation of the university programs with which most of these individuals are affiliated. The strong role of venture capital variable reported previously is not supported. Indeed, if anything the data tell us that there were fewer firms founded, other things equal, where there more venture capital firms.

III.B. The Short-Run Model

Our data set permits us to examine the panel of 183 BEA areas for each of the years 1976 through 1989. Table 7 reports poisson regressions for this entire panel.

The first column reports the results using only the counts of stars and their collaborators active each year in each BEA area. Examination of the data suggested that these effects -- particularly for stars -- were nonlinear so we add squared values in the

second column. As the number of stars increases, their marginal contribution diminishes eventually passing through zero.

We believed that these nonlinearities might reflect the declining value over time of the intellectual capital as we have measured it. Basically, as the knowledge diffuses we observe both more and more stars and less and less payoff to any one of them if he or she were to start a firm. There is evidence in the data such that stars are less likely to result in birth of NBEs after 1985 than before. This is illustrated in the third column of Table 7 where the prefix 8689 means that the variable has been multiplied by a dummy equal to 1 during 1986-1989 and 0 otherwise. The effect of stars virtually disappears while there is some more positive effect from their collaborators who may provide labor to new firms.²⁷ In the same regression, we see that university quality and federal support are also significant measures of intellectual capital relevant to firm founding.

The fourth column repeats the results of Section III.A: The economic variables enter significantly with the expected sign if the intellectual capital variables are omitted from the regression. However, unlike the previous long-run case, we can now enter the earnings-price ratio. Here this variable enters with the correct sign and insignificantly.

The final column of Table 7 presents the combined model incorporating the intellectual capital and other variables. The results for the intellectual capital measures are robust while the sign of the venture capital variable turns significantly negative and the employment variable turns insignificantly negative. Thus, taken as a whole the results summarized in Table 7 support the strong role of intellectual capital variables in determining the development of the American biotech industry.

The role of the economic variables, particularly $VENTCAP_{it}$, is explored further in Table 8. This table presents representative results for births in the NBF and NBS subcomponents of NBE. We see in columns 1 and 2 that if only the economic variables are introduced we get all the expected signs at appropriate significance including a nice result that $VENTCAP_{it}$ has a significantly positive effect on the birth of new firms but an insignificant effect on the birth of subunits of existing firms which would not normally be financed by venture capital firms. The combined model is reported in columns 3, 4, and 5 for births of firms, subunits, and enterprises, respectively.²⁸ Again, in the presence of intellectual capital the nice simple economic story does not hold up: the coefficients of $VENTCAP_{it}$ and EMP_{it} turn negative, significantly so in the case of venture capital.

To complete the short-run model, we next examine the dynamic nature of the birth process. To do so, we need to explore the effects of the stocks of NBEs both locally and nationally (i.e., NBE_{it-1} and NBE_{t-1}), the presence of a time trend $YEAR_t$ ($YEAR_{1976} = 1$, $YEAR_{1977} = 2$, etc.), and the effect of the statistically problematic lagged dependent variable $BNBE_{it-1}$. Collectively, these are the main variables that have been emphasized in the population ecology models of firm birth. In Table 9 we see that like other extrapolative forecasting methods, a combination of the dynamic variables -- with or without the lagged dependent variable -- yields reasonable models. In column 4 or 5, we interpret the positive coefficient on NBE_{it-1} as reflecting local agglomeration effect and the negative coefficient on NBE_{t-1} as due to the depressing effect on profitability as the national total of biotech firms rises. However, we believe that these models are misleading by themselves without including the structural variables which explain much of the lagged values.

In Table 10, we add these four variables to first the intellectual capital variables, then to the other economic variables, and finally to the full fundamentals model. In the fourth column we drop the statistically insignificant but problematic lagged dependent variable. Comparing the log-likelihoods of column 5, Table 7, columns 4 or 5 of Table 9, and columns 3 or 4 of Table 10, we see that the dynamic variables add some explanatory power to the basic model, but do not alter the qualitative description of the structural variables: intellectual capital as measured by the star variables, QUAL1, and FEDGRANT remain the dominant forces with supporting roles for the other economic and dynamic variables. VENTCAP_{it} and EMP_{it} continue to display significant coefficients of the wrong sign while EJOB_{it} and E/PRATIO_{it} enter correctly and significantly. The local lagged stock of firms NBE_{it-1} and the corresponding national stock NBE_{t-1} retain significance consistent with both local agglomeration effects and a national carrying capacity for the industry. Of course, we cannot distinguish between agglomeration or other "real" explanations for the significantly positive coefficient on NBE_{it-1} and omitted or mismeasured variable interpretations. The fourth column of Table 10 reflects our preferred model.

We conclude by observing that by and large the dynamic behavior of births of new biotech enterprises reflects the dynamic behavior of the underlying fundamental variables, principally the intellectual capital variables. The lagged local stock of firms, the lagged national stock, and a time trend play a significant but supporting role.

IV. Conclusions and Summary

The American biotechnology industry which was essentially nonexistent in 1975 grew to seven hundred active firms over the next 15 years. This industry is a testament to the value of basic scientific research. In this paper, we show the tight connection between the intellectual capital created by frontier research and the founding of firms in the industry. At least for this high-tech industry, the growth and location of intellectual capital was the principal determinant of the growth and location of the industry itself. Venture capital firms, which appear to be significant determinants when intellectual capital variables are excluded from the regressions, are at best insignificant as determinants of firm births and generally exhibit anomalously significantly negative coefficients. This may reflect a tendency toward fewer, larger firms in areas rich in venture capital firms, but that remains conjecture until further research is done. The crucial importance of lagged variables and time trends which have been popularized by the population ecologists also does not survive in the presence of structural variables.

We conclude that the growth and diffusion of intellectual capital was the main determinant of where and when the American biotechnology industry developed. Intellectual capital tended to flourish around great universities, but the existence of outstanding scientists measured in terms of research productivity was key. In future research, we hope to resolve some of the puzzles we have raised and to extend the analysis to other countries and industries.

Appendix A: Data

In this paper we used the data sets developed for the Project on "Intellectual Capital, Technology Transfer, and the Organization of Leading-Edge Industries: The Case of Biotechnology," Lynne G. Zucker, Marilyn B. Brewer, and Michael R. Darby, Principal Investigators. These data will be archived upon completion of the project in the Data Archives at the UCLA Institute for Social Science Research. The project has been housed at and supported in part by ISSR.

The data sets contain a variety of linkages, but are derived from distinct sources and are most easily discussed sequentially. Before doing so, however, it is important to understand a few salient features of the data sets.

First the data generally are in the form of panels: cross sections observed annually from 1975 or earlier through 1989 or later. The cross-sections are generally defined geographically. For this purpose, geocoding was originally done at the finest level possible of the zip-code, county, BEA- area, and national levels of geography. For the analyses reported in this paper, all data for finer levels of geography have been aggregated to the 183 BEA-defined functional economic areas in the United States. Every U.S. zip code and county is assigned to one of these areas.

Generally, natural geography exists only for institutions (universities, research institutes and hospitals, biotechnology firms, and venture capital firms), but a variety of economic data is collected at or can be aggregated to BEA level as well. Individual scientists are linked to locations through the institutional affiliations reported in their

publications in the article data set. The citations were also collected for articles if and only if they appeared in the article data set; so scientists are credited with citations only insofar as they are to the 4,315 articles reporting gene sequence discoveries and published in major journals.

A.1. Scientist Data Set

As mentioned above, in the earlier stages of the project, GenBank was used to determine for each scientist the total number of genetic sequence discoveries reported in all articles reporting gene sequence discoveries up to 1990 for which that scientist was listed as an author.²⁹ On the basis of these total number counts, we identified 315 leading researchers whom we termed "stars." We used a cutoff of at least 41 sequence discoveries. However, total gene sequence discoveries is an imperfect measure of scientific output: many technological breakthroughs in the 1970s and 1980s made gene sequencing more efficient, and the discovery of sequences for some types of genes is either more difficult or more important. Accordingly, another 22 stars were identified based on frequency of publication of articles reporting gene sequences bringing the total to 337 stars.

Because research discoveries frequently occurred through teams, scientists who were coauthors with a star but who had not themselves met the star criteria were labeled "collaborators." In terms of distinct individuals, we captured data for 7,718 collaborators in addition to the 337 stars, or a total of 8,055 individuals worldwide. Each star and collaborator was assigned a unique scientist ID number for use in linking to the other data files.

As discussed in Data Appendix A.2 below, we have the affiliations listed by each of these scientists on 4,315 journal articles in our article data set. We use the institutional affiliations to locate scientists: Generally a scientist is located for our analysis at the institution given in his or her last publication; however, scientists are said to be "born" in the year and at the institution of their first publication.

As discussed in the text, a star or collaborator is defined as "active" in any year in which he or she has published three or more articles in the three-year period ending with that year. The variables $ACSTAR_t$ and $ACCOLL_t$ are counts of the number of active stars or collaborators by year in each BEA area.

A.2. Article Data Set

Our article data set consists of all 4,315 articles in major journals listed in GenBank as reporting genetic sequence discoveries for which one or more of our 337 stars were listed as authors.³⁰ All of these articles were assigned unique article ID numbers and collected by hand. For each article, scientist ID numbers are used to identify the order of authorship and the institutional affiliation and location for each author on each article. This hand coding was necessary because, under the authorship traditions for these fields, the head of the laboratory who is often the most prestigious author frequently appears last. Our stars, for example, were first authors on 18.3 percent of the articles and last authors on 69.1 percent of the 4,285 articles remaining after excluding the 30 sole-authored articles.³¹ Unfortunately, only first author affiliations are available in machine-readable sources.

The resulting authorship data file contains 20,669 observations, approximately 5

authors for each of the 4,315 published articles. Each authorship observation gives the article ID number, the order of authorship, the scientist ID number of one of our stars and collaborators, and an institutional ID number for the authors affiliation which links him or her to a particular institution with a known zip code as of the publication date of the article.

A.3. Citation Data Set

We have collected data for 1982, 1987, and 1992, on the total number of citations to each of our 4,315 published articles listed in the Institute for Scientific Information's Science Citation Index. These citation counts are linked to the article and authorship data set by the article ID number.

A.4. University Data Set

Our university data set consists 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, 1983-84.³² 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. All of the following additional variables are based on data in the National Academy of Sciences study by Lyle Jones, Gardner Lindzey, and Porter Coggeshall (1982).

$FEDGRANT_i$ is the total number of scientists in all biotech-relevant departments in

BEA area is supported by 1979-80 grants from the NSF, NIH and the Alcohol, Drug Abuse and Mental Health Administration (ADAMHA).

R&DEXP_i is the sum in thousands of dollars of total reported university expenditures for research and development in the biotech-relevant departments over all the universities listed in Jones et al. which are located in the BEA area. We note that reporting practices for these expenditures seemed inconsistent across universities, perhaps because they do not keep accounts as would have been required to accurately answer the survey.

We define university quality level based on the scholarly quality rating in the reputational survey in Jones et al. Reputational ratings were 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. The number of universities in a BEA with one or more most highly rated programs (rated above 4) is our variable QUAL1.³³ Similarly, QUAL2, QUAL3, and QUAL4 represent the numbers of universities with rated programs of above 3 through 4, of above 2 through 3, and of 2 and below, respectively. That is, QUAL1 is the count of the very best universities in an area while QUAL4 is a count of those of the lowest quality to be rated at all.

A.5. Research Institute and Hospitals Data Set

For those U.S. research institutions and hospitals listed as affiliations in the article data set, we assigned an institutional ID number and an institute/hospital flag, and obtained

an address including a zip code as required for geocoding. No additional information has been collected on these institutions.

A.6. Biotechnology Firm Data Set

The starting point for our firm data set covered the industry as of April 1990 and was purchased from the North Carolina Biotechnology Center (1991), a private firm which tracks the industry. This data set identified 1075 firms; some of which were duplicates or foreign and others of which had died or merged. Further, there were a significant number of missing firms which had died earlier. For these reasons, an intensive effort was made to supplement the NCBC data with information from *Bioscan* (1989-1993) and an industry data set provided by a firm in the industry which was also the ancestor of the *Bioscan* data set (Pre-Bioscan 1988). Each of the firms was assigned an institutional ID number and an enterprise flag.

We combined these three sources to identify 751 distinct U.S. firms for which we could determine a zip code and a date of founding (or entry into biotechnology for subunits of preexisting firms). Of these, based on the financial information in the NCBC data supplemented where necessary and possible by our other sources, 511 are classified as NBFs, 150 as NBSs, and another 90 could not be classified clearly into either subcategory but were retained as NBEs. Of the 90 "others," 18 were in the problematic category of joint ventures and the remaining had insufficient data to classify as independent firm or subunit. We also obtained for 52 of these 751 firms the date of their exit by death or merger.

Based on these data, we have developed a continuous series on the number of active

new biotech enterprises NBE_{it} (and its subcomponents NBF_{it} and NBS_{it}) and their births ($BNBE_{it}$, $BNBF_{it}$, and $BNBS_{it}$), all series being by year and BEA area.³⁴ The national stock of new biotech enterprises NBE_t is simply the annual aggregate of the BEA area values.

A.7. Venture-Capital Firm Data Set

We created a venture-capital firm data set by extracting from the Stanley Pratt (1982) directory the name, type, location, year of founding, and interest in funding biotech firms. This information was extracted for all venture capital which were legally permitted to finance start-ups. This latter requirement eliminated a number of firms which are chartered under government programs targeted at small and minority businesses.

We developed two alternative measures of the availability of venture capital: $VENTCAP_{it}$ which is the total number of eligible venture capital firms in BEA area i in year t and $VCBIO_{it}$ which counts only the subset of those firms indicating a specific interest in their Pratt listing in funding biotech firms. This approach accounts for the founding date of firms appearing in the 1982 Pratt directory, but not for those firms that may have either entered thereafter or existed in earlier years but exited before the Directory was compiled.

A.8. Economic Data Set

This data set consist of three variables EMP_{it} , $EJOB_{it}$, and POP_{it} observed at the BEA area and year level, one variable $SIC2830$, which varies across BEA areas but not over time, and three national variables $UNEMP_t$, $E/PRATIO_t$, and $YEAR_t$ which vary only with the

year.

Total employment EMP_{it} , average earnings per job $EJOB_{it}$, and total population POP_{it} are all reported by the BEA at the BEA area level from county level data in U.S. Department of Commerce (1992):

EMP_{it} Table K, line 010

$EJOB_{it}$ Table V, line 290 (wage & salary disbursements, other labor income, and proprietors income per job) deflated by the implicit price deflator for personal consumption expenditures³⁵

POP_{it} Table B, line 020

$SIC2830_{it}$ is the number in the BEA in 1980 of establishments producing goods with SIC code 2830 (Drugs, Total; Number of Establishments). This variable is aggregated to the BEA area level from county level data in U.S. Department of Commerce, Bureau of the Census, (1982).

The total unemployment rate $UNEMP_{it}$ and the S&P500 earnings-price ratio $E/PRATIO_{it}$ were taken from CITIBASE (1993), series LHUEM and FSEXP, respectively.

The time trend $YEAR_{it}$ is 1 in 1976, 2 in 1977, etc.

References

- Bania, Neil, Randall Eberts, and Michael Fogarty, "Universities and the Startup of New Companies: Can We Generalize from Route 128 and Silicon Valley," *Review of Economics and Statistics*, November 1993, in press.
- Beesley, Michael, "The Birth and Death of Industrial Establishments: Experience in the West Midlands Conurbation," *Journal of Industrial Economics*, October 1955, 4: 45-61.
- Bilofsky, H.S., and C. Burks, "The GenBank (R) Genetic Sequence Data Bank," *Nucleic Acids Research*, 1988, 16: 1861-1864.
- Blumenthal, David, "Academic-Industry Relationships in the Life Sciences: Extent, Consequences, and Management," *Journal of the American Medical Association*, December 16, 1992, 268: 3344-3349.
- Carroll, Glenn R., *Publish and Perish: The Organizational Ecology of Newspaper Industries*, Greenwich, CT: JAI, 1987.
- Carroll, Glenn R., and Yangchung Paul Huo, "Organizational Task and Industrial Environments in Ecological Perspective: Findings from the Local Newspaper Industry," *American Journal of Sociology*, January 1986, 91: 838-873.
- Cohen, Stanley, A. Chang, Herbert Boyer, and R. Helling, "Construction of Biologically Functional Bacterial Plasmids *in vitro*," *Proceedings of the National Academy of Sciences*, 1973, 70: 3240-3244.
- Darby, Michael R., "The U.S. Productivity Slowdown: A Case of Statistical Myopia," *American Economic Review*, June 1984, 74: 301-322.
- Darby, Michael R., "Causes of Declining Growth," in *Policies for Long-Run Economic Growth*, proceedings of A Symposium Sponsored by The Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, August 27-29, 1992, Kansas City: Federal Reserve Bank of Kansas City, 1993.
- Delacroix, J., and Glenn R. Carroll, "Organizational Foundings: An Ecological Study of the Newspaper Industries of Argentina and Ireland," *Administrative Science Quarterly*, 1983, 28: 274-291.
- Delacroix, J., and M. E. Solt, "Niche Formation and Foundings in the California Wine Industry," in Glenn R. Carroll, ed., *Ecological Models of Organizations*, Cambridge, MA: Ballinger, 1988.

- Demsetz, Harold, "The Theory of the Firm Revisited," *Journal of Law, Economics, and Organization*, Spring 1988, 4: 141-161.
- Dorfman, Nancy S., "Route 128: The Development of a Regional High Technology Economy," in David Lampe, ed., *The Massachusetts Miracle: High Technology and Economic Revitalization*, Cambridge, MA: MIT Press, 1988.
- Dunne, Timothy, Mark J. Roberts, and Larry Samuelson, "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries," Pennsylvania State University, working paper, December 1987.
- Evans, David S., "Tests of Alternative Theories of Firm Growth," *Journal of Political Economy*, August 1987, 95: 657-674.
- Greene, William H., *LIMDEP: User's Manual and Reference Guide, Version 6.0*, Bellport, NY: Econometric Software, Inc., 1992.
- Griliches, Zvi, "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, December 1990, 28:1661-1707.
- Griliches, Zvi, "The Search for R&D Spillovers," *Scandinavian Journal of Economics*, 1992 Supplement, 94: 29-47.
- Hannan, Michael T., "A Model of Competitive and Institutional Processes in Organizational Ecology," Technical Report 86-13, Department of Sociology, Cornell University, 1986.
- Hannan, Michael T., and Glenn R. Carroll, *Dynamics of Organizational Populations: Density, Legitimation, and Competition*, New York: Oxford University Press, 1992.
- Hannan, Michael T., and John Freeman, "The Ecology of Organizational Founding: American Labor Unions, 1836-1985," *American Journal of Sociology*, 1987, 92: 910-943.
- Hannan, Michael T., and John Freeman, *Organizational Ecology*, Cambridge, MA: Harvard University Press, 1989.
- Hause, John C., and Gunnar Du Rietz, "Entry, Industry Growth, and the Microdynamics of Industry Supply," *Journal of Political Economy*, August 1984, 92: 733-757.
- Hausman, Jerry, Bronwyn H. Hall, and Zvi Griliches, "Econometric Models for Count Data with an Application to the Patents-R&D Relationship," *Econometrica*, July 1984, 52: 909-938.

- Jaffe, Adam B., "Real Effects of Academic Research," *American Economic Review*, December 1989, 79: 957-970.
- Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," NBER Working Paper No. 3993, February 1992.
- Jovanovic, Boyan, "Selection and the Evolution of Industry," *Econometrica*, May 1982, 50: 649-670.
- Kenney, Martin, *Biotechnology: The University-Industrial Complex*, New Haven: Yale University Press, 1986.
- Lerner, Joshua, "Venture Capitalists and the Decision to Go Public," working paper, Harvard Business School, August 1993. (1993a)
- Lerner, Joshua, "Venture Capitalists and the Oversight of Private Firms," working paper, Harvard Business School, August 1993. (1993b)
- Louis, Karen Seashore, David Blumenthal, Michael E. Gluck, and Michael A. Stoto, "Entrepreneurs in Academe: An Exploration of Behaviors among Life Scientists," *Administrative Science Quarterly*, March 1989, 34: 110-131.
- Mansfield, Edwin, "Entry, Gibrat's Law, Innovation, and the Growth of Firms," *American Economic Review*, December 1962, 52: 1023-1051.
- Mansfield, Edwin, "Academic Research Underlying Industrial Innovations: Sources and Characteristics," working paper, University of Pennsylvania, no date/circa 1992.
- Pisano, Gary P., "Using Equity Participation to Support Exchange: Evidence from the Biotechnology Industry," *Journal of Law, Economics, and Organization*, Spring 1989, 5: 109-126.
- Pisano, Gary P., "The R&D Boundaries of the Firm: An Empirical Analysis," *Administrative Science Quarterly*, March 1990, 35: 153-176.
- Pisano, Gary P., and Paul Y. Mang, "Collaborative Product Development and the Market for Knowhow: Strategies and Structures in the Biotechnology Industry," in Richard Rosenbloom and Robert Burgelman, eds., *Research on Technological Innovation Management*, vol. 5, Greenwich, CN: JAI Press, in press 1994.
- Singh, Jitendra V., and Charles J. Lumsden, "Theory and Research in Organizational Ecology," *Annual Review of Sociology*, 1990, 16: 161-195.

- Singh, Jitendra V., D.J. Tucker, and A.G. Meinhard, "Institutional Change and Ecological Dynamics," in Walter W. Powell and Paul J. DiMaggio, eds., *The New Institutionalism in Organizational Analysis*, Chicago: University of Chicago Press, 1991.
- Smilor, Raymond W., George Kozmetsky, and David V. Gibson, *Creating the Technopolis: Linking Technology, Commercialization, and Economic Development*, Cambridge, MA: Ballinger Publishing Co., 1988.
- Teece, David J., "Profiting from Technological Innovation: Implications for Integration, Collaboration, Licensing, and Public Policy," *Research Policy*, 1986, 15: 285-305.
- Trajtenberg, Manuel, Rebecca Henderson, and Adam Jaffe, "Ivory Tower versus Corporate Lab: An Empirical Study of Basic Research and Appropriability," NBER Working Paper No. 4146, August 1992.
- Tucker, D.J., Jitendra V. Singh, and A.G. Meinhard, "Organizational Form, Population Dynamics, and Institutional Change: A Study of Founding Patterns of Voluntary Organizations," *Academy of Management Journal*, 1990, 33: 151-178.
- Zucker, Lynne G., "Combining Institutional Theory and Population Ecology: No Legitimacy, No History," *American Sociological Review*, 1989, 54: 542-545.
- Zucker, Lynne G., Marilyn B. Brewer, and Michael R. Darby, "Information Dilemmas and Collaboration Structure in Biotechnology: Interpersonal and Organizational Boundaries as Trust Production," in Barry Markovsky, Jodi O'Brien, and Karen Heimer, eds., *Advances in Group Processes*, Volume 11. Greenwich, Connecticut: JAI Press, in press 1994.
- Zucker, Lynne G., Marilyn B. Brewer, Amalya Oliver, and Julia Liebeskind, "Basic Science as Intellectual Capital in Firms: Information Dilemmas in rDNA Biotechnology Research," working paper, UCLA Institute for Social Science Research, 1993.
- Zucker, Lynne G., and Michael R. Darby, "Perceptions on the Organization of Biotechnology Science and Its Commercialization in Japan," working paper, UCLA Institute for Social Science Research, December 1993.

Data Sources

Bioscan, Volume 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, Lyle V., Gardner Lindzey, and Porter E. Coggeshall, eds., *An Assessment of Research-Doctorate Programs in the United States: Biological Sciences*, Washington: National Academy Press, 1982.
- 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, Stanley E., *Guide to Venture Capital Sources*, 6th ed., Englewood Cliffs, NJ: Prentice-Hall, Inc., 1982.
- U.S. Department of Commerce, Bureau of the Census, *County Business Patterns, 1980: U.S. Summary, State, and County Data*, machine readable data base, Washington: Bureau of the Census, 1982. [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, *National Income and Product Accounts of the United States, Volume 2, 1959-88*, Washington: U.S. Government Printing Office, 1992. (1992a)
- U.S. Department of Commerce, Economics and Statistics Administration, Bureau of Economic Analysis, *Regional Economic Information System, Version 1.3*, CD-ROM, machine readable data base, Washington: Bureau of Economic Analysis, May 5, 1992. (1992b)
- 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.

Footnotes

1. Nancy Dorfman (1988) and Raymond Smilor, George Kozmetsky, and David Gibson (1988) report recent case studies relating industrial development to M.I.T. and the University of Texas at Austin, respectively. Neil Bania, Randall Eberts, and Michael Fogarty (1993) in an exploratory regression study of firm founding in six industries (including two high-tech industries but not including biotechnology) find that university research expenditures produced higher local foundings only for the Electrical and Electronic Equipment industry.
2. For a discussion of the potentials and pitfalls of use of patent statistics, see Griliches (1990). Trajtenberg, Henderson, and Jaffe (1992) present empirical evidence based on patterns of patent citations showing that university research is indeed more "basic" than corporate research efforts.
3. Cohen, Chang, Boyer, and Helling (1973).
4. High temperature superconductivity was certainly a Kuhnian paradigm shift of dramatic magnitude comparable to those underlying biotechnology, but we do not believe any comparable element of intellectual capital was created there. The difference is that the technique of research for these superconductors was widely known and easily replicable by any competent chemist. Thus, the publication of the discovery of a ceramic superconductor embodying rare earths broke the bounds of the existing science and set off an international search for superconductors which worked at yet higher temperatures or which possessed more usable mechanical properties.

5. There is a tension between the academic norms of openness and free publication and the creation of intellectual property. David Blumenthal (1992) and many Japanese academics (see Zucker and Darby 1993) view these ends as antithetical. Natural excludability goes part way toward resolving such conflicts since one can publish without really giving away the commercially valuable information.

A different philosophy underlies much recent federal legislation and regulation: If commercially valuable discoveries are made available to everyone, then no single firm will have the incentive in many cases to invest the large additional sums to transform a discovery into a commercially valuable product. Thus most federally funded research now grants patent rights to the discoverer so that the incentives are aligned to promote and commercialize a discovery to the benefit of society. Since this is known at the time of contracting, there is no ethical issue of private benefit from publicly funded research as the federal government pays less in view of its assignment of rights to the discoverer.

6. Because of very obvious principal-agent problems, these enormous wages are generally seen to be taken in the form of equity in the firms built around the scientists. For those who became scientific entrepreneurs, the rewards for successful application of their intellectual capital easily ran up into the hundreds of millions of dollars.

7. This technology is also identified by the terms monoclonal antibodies, MABs, or hybridomas.

8. See Zucker, Brewer, Oliver, and Liebeskind (1993). The Genbank data set and our methods of identifying stars and their collaborators is discussed in more detail below in the Data Appendix. As will be obvious, much of the time since 1990 has been spent in

developing reasonable measures of intellectual capital and in collecting and coding data necessary to locate the authors of the discoveries reported in the articles in question and to trace the diffusion process.

9. Because of the way the data set was constructed, the screen is in fact more severe for collaborators than for stars: If a collaborator wrote any articles reporting gene sequence discoveries which were not coauthored with any of our stars, they would not appear in our set of 4,315 articles and hence not count toward active status.

10. In this paper we are limiting our consideration of collaborations to those involving joint publications. David Teece (1986), Gary Pisano (1989,1990), and Pisano and Mang (1994) analyze a variety of forms of collaboration within the biotechnology industry. We intend to extend our work along these lines in future research.

11. In order to achieve scales conformable to the stars, the collaborators are measured in units of 20 people and the NBEs are measured in units of 4 firms.

12. The comparison is somewhat misleading since prior to the discoveries of the early 1970s there was no apparent commercial application of rDNA and we accordingly date the first biotech firms established in response to the scientific revolution in 1976. Even if we look at the firms established between 1967 and 1975 which ultimately adopted or established subunits using biotechnology, there were only 89 firms founded in these nine years with no more than 16 founded in any one year.

13. Besides the counts of universities with one or more great biotech relevant departments, we also had counts of the number of universities with only lower ranking departments, QUAL2, QUAL3, and QUAL4. We also tried a third quality variable R&DEXP which is

the sum in thousands of dollars of total reported university expenditures for research and development in the biotech-relevant departments in the BEA area.

We found that because of high multicollinearity among the various university variables it is generally necessary to restrict the analysis to two of these variables. QUAL1 and either FEDGRANT or R&DEXP seemed to capture nearly all the explanatory power in the data. We preferred the FEDGRANT data supplied from the granting agencies to the raw data supplied by the universities which was the basis of R&DEXP; reporting practices for the latter seemed inconsistent across universities, perhaps because they do not keep accounts as would have been required to accurately answer the survey.

The eighteen universities with sufficiently high ratings in biotech-relevant departments to count in the QUAL1 variable were: Brandeis University, California Institute of Technology, Columbia University, Cornell University, Duke University, Harvard University, Johns Hopkins University, Massachusetts Institute of Technology, Rockefeller University, Stanford University, University of California-Berkeley, University of California-Los Angeles, University of California-San Diego, University of California-San Francisco, University of Colorado at Denver, University of Washington (Seattle), University of Wisconsin-Madison, Yale University.

14. We are indebted to Zvi Griliches for suggesting estimation using the poisson process.
15. Lerner (1993a) shows that venture capitalists are effective at maximizing the initial public offering value of biotech firms which is one way of lowering the cost of capital to these firms.
16. Few advocates of venture capital would go as far as Martin Kenney (1986, p.133) to say

"Biotechnology has emerged as an industry largely because of one economic institution: venture capital."

17. We also tried a variable $VCBIO_{it}$ which counts only the subset of those firms indicating a specific interest in their Pratt listing in funding biotech firms. We found that $VENTCAP_{it}$ and $VCBIO_{it}$ yielded substantially identical results and we report only the $VENTCAP_{it}$ results in this paper.

Our (marginal) preference for the $VENTCAP$ rather than $VCBIO$ measure may need some explanation. Although the latter is more appealing on the surface, appearances may be misleading. Firms that funded biotech firms before responding to the questionnaire may not then be particularly looking for more such firms. Furthermore, we note that some venture capital firms expressing a desire to fund biotech firms in areas with no biotech activity may be expressing a forlorn hope for the profitable opportunities which had been available to their peers in other areas.

18. In empirical work not reported below, we have measured their presence in a BEA area in two alternative ways: $SIC2830_{it}$ is the number in 1980 of establishments producing goods with SIC code 2830 in the BEA. $PRE67BE_{it}$ is the number of enterprises founded before 1967 engaged in biotechnology in 1990. Each measure proved less than ideal.

$SIC2830_{it}$ had to be defined based on an SIC code which included biological products, medicinals, and pharmaceuticals because confidentiality rules limit the detail which the Census Bureau reports in the county level data from which we constructed this variable. $PRE67BE_{it}$ is more precisely focussed, but may merely account for the birth of the NBSs represented in this variable. After trying both in a variety of regressions, we concluded that

the results with these variables were too unstable and problematic to warrant reporting.

19. Care should generally be given to the issue of the extent of the market appropriate for analysis of competitive pressures on firm foundings. As in this case, there may be national competition in the output market and significant local competition in the input market. A number of authors use only one lagged number of firms variable without apparent regard for whether local, regional, national or several levels of competition might be significant. For example, Michael Hannan and Glenn Carroll (1992, p. 85) use either local or national counts on numbers of newspapers without apparent concern for whether one, the other, or both are appropriate definitions of significant newspaper markets.

20. See, for example, Delacroix and Carroll (1983), Carroll and Huo (1986), Carroll (1987), Delacroix and Solt (1988), Tucker, Singh, and Meinhard (1990) and Singh, Tucker, and Meinhard (1991).

21. We also planned to include the national unemployment rate UNEMP, which is a generally robust measure of the cyclical state of the economy. Higher values of UNEMP were expected to discourage births. However, when this variable is included it always enters positively, higher unemployment encourages births. We are indebted to Zvi Griliches for the observation that this is an artifact of the high unemployment rates during the early 1980s when the time was ripe for starting biotech firms. We judge the estimated effect to be spurious and dropped the variable from the regressions reported here, but do not believe that its inclusion would have qualitatively affected any of the results.

22. See also Beesley (1955), Hause and Du Rietz (1984), Dunne, Roberts, and Samuelson (1987), and Evans (1987).

23. See Greene (1992, pp. 539-549) for a discussion of this procedure.
24. This can be understood in the simple case of $y_t = ax_t + by_{t-1}$, by noting that the long-run value of y_t is $ax_t/(1 - b)$.
25. This hypothesis is problematic too, however. In Table 8 below we present evidence that, in the presence of intellectual capital variables, the coefficient on $VENTCAP_{it}$ is negative in predicting the births of NBSs as well as NBEs and NBFs. Since venture capital firms are not a major source of financing subunits of incumbent firms, it is hard to see how their presence could affect the scale of those units.
26. All coefficients in these poisson regressions are significant at the .001 level.
27. We examined also interactions with dummy variables for 1976-1980 and with a time trend. Since the 1976-1980 dummies did not enter significantly, we believe that the reported form more accurately reflect the time or diffusion dependence than a negative trend throughout the period.
28. Column 5 reproduces for ready comparison column 5 from the previous table.
29. 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.
30. A small number of unpublished papers and articles appearing in proceedings volumes and obscure journals were excluded to permit the hand coding detailed below.
31. This positional tradition holds across national boundaries: As a percentage of articles co-authored by their fellow nationals, American stars are 16.4 percent of first authors and 71.2 percent of last authors, compared to 21.2 percent and 63.1 percent, respectively, for

Japanese, and 19.7 percent and 69.2 percent for other nationalities.

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

33. 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.

34. Where a new enterprise enters the data set due to the merger of a NBF and another firm, we count it for the purposes of this paper as a continuation of the original NBF and not a new birth (the older NBF if two are involved). If enterprises already in the data set merge and one continues with the other(s) absorbed, the enterprise is counted as the continuing enterprise and not a new birth.

35. The annual data for the implicit price deflator for personal consumption expenditures were taken from U.S. Department of Commerce (1992a, p. 247, line 16) as updated in the July 1992 *Survey of Current Business*, (p. 92, line 16).

Table 1
Distribution of U.S.-Born Stars
and U.S.-Born Collaborators

Organization Type	<u>Full Data Set^a</u>		<u>Ever Active in U.S.^b</u>	
	Number of Scientists	Citations ^c /Scientist/yr.	Number of Scientists	Citations ^c /Scientist/yr.
<u>Stars:</u>				
University	163	96.9	108	126.9
Institute	44	72.2	25	111.9
Firm	6	144.8	1	822.0
Dual	0	n/a	0	n/a
Total	213		134	
<u>Collaborators:</u>				
University	2887	11.8	367	30.8
Institute	703	16.1	85	40.4
Firm	390	38.0	46	132.4
Dual	3	10.8	0	n/a
Total	3983		498	

Notes:

- a. "U.S.-Born" scientists are defined by the location of the affiliation listed on their first publication.
- b. Ever active in the U.S. means that in at least one three year period beginning 1974 or later and ending 1989 or earlier, the scientist was listed on at least three articles appearing in our data set of 4,315 articles which reported gene sequence discoveries and were published in major journals and that the affiliation listed in the last of the three articles was located in the United States.
- c. Citation counts are for 1987 and 1992 for all articles in our data set (whenever published) for which the individual was listed as an author.

Table 2
Logit Regression
Probability that a Pair
of Coauthors Are from Same Institution
1967-1990

Intercept	6.778**	[.393]
Year	-0.063**	[.005]
Both Foreign	0.714**	[.031]
No. Coauthors	-0.162**	[.005]
Both in a Firm	3.356**	[.203]
Both Stars	0.147**	[.052]

Chi-square for -2 log likelihood (5 d.f.) = 2219.5

Concordance of predicted probabilities & observations:
 Concordant 68.6%, Discordant 30.9%, Tied 0.5%

Notes:

Standard errors in square brackets.

** Parameter significant at the 0.01 level.

Table 3
Poisson Regressions on Stock of
New Biotech Enterprises at the beginning of 1990

Constant	0.927*** (.0467)	0.592*** (.0545)	0.379*** (.0619)	-4.981*** (.4253)	-3.666*** (.5107)
ACSTAR _{1,76-80}	0.574*** (.0499)	0.228*** (.0397)	0.505*** (.0403)	0.378*** (.0417)	0.303*** (.0428)
ACCOLL _{1,76-80}	-0.084*** (.0196)	-0.042 (.0255)	-0.142*** (.0215)	-0.154*** (.0208)	-0.193*** (.0200)
QUAL ₁		0.606*** (.1619)	1.010*** (.1295)	1.205*** (.1220)	-0.209 (.1516)
FEDGRANT ₁		65E-4*** (89E-5)	96E-4*** (73E-5)	69E-4*** (80E-5)	64E-4*** (85E-5)
VENTCAP _{1,80}			-0.036*** (.0029)	-0.039*** (.0036)	-56E-4 (.0049)
EMP _{1,80}				-12E-8*** (33E-9)	-13E-8*** (34E-9)
EJOB _{1,80}				28E-5*** (21E-6)	20E-5*** (27E-6)
BNBE _{1,76-80}					0.375*** (.0287)
Log-likelihood	-826.8	-583.5	-516.2	-438.2	-339.5

Note: Standard errors are in parentheses below coefficients.
Probability t>x: * < .05, ** < .01, *** < .001

Table 4
Poisson Regressions on Stock of New Biotech
Enterprises born 1981-89 at the beginning of 1990

Constant	0.709*** (.0522)	0.389*** (.0601)	0.162* (.0689)	-5.381*** (.4678)	-4.183*** (.5541)
ACSTAR _{i,76-80}	0.650*** (.0530)	0.267*** (.0437)	0.542*** (.0439)	0.411*** (.0454)	0.339*** (.0468)
ACCOLL _{i,76-80}	-0.113*** (.0208)	-0.062* (.0279)	-0.159*** (.0233)	-0.171*** (.0225)	-0.204*** (.0218)
QUAL _i		0.660*** (.1775)	1.040*** (.1411)	1.243*** (.1327)	-0.081 (.1689)
FEDGRANT _i		62E-4*** (97E-5)	96E-4*** (80E-5)	70E-4*** (87E-5)	67E-4*** (92E-5)
VENTCAP _{i,80}			-0.036*** (.0032)	-0.038*** (.0040)	-79E-4 (.0054)
EMP _{i,80}				-13E-8*** (37E-9)	-14E-8*** (37E-9)
EJOB _{i,80}				29E-5*** (23E-6)	22E-5*** (29E-6)
BNBE _{i,76-80}					0.345*** (.0313)
Log-likelihood	-712.4	-514.2	-456.5	-388.3	-317.6

Note: Standard errors are in parentheses below coefficients.
Probability t>x: * < .05, ** < .01, *** < .001

Table 5
 OLS Regressions on Stock of
 New Biotech Enterprises at the beginning of 1990

Constant	1.945*** (.4960)	0.751 (.3994)	0.799* (.4004)	-6.703 (3.906)	-3.303 (2.246)
ACSTAR _{1,76-80}	4.135*** (1.145)	2.797** (0.887)	2.583** (.9010)	2.442** (.8860)	3.595*** (.5114)
ACCOLL _{1,76-80}	1.448** (.4849)	0.018 (.4469)	0.030 (.4463)	0.081 (.4338)	-1.782*** (.2673)
QUAL ₁		10.30*** (2.150)	9.758*** (2.187)	10.28*** (2.132)	2.498 (1.289)
FEDGRANT ₁		0.070*** (.0124)	0.064*** (.0132)	0.037* (.0152)	0.014 (.0088)
VENTCAP _{1,80}			0.106 (.0833)	-0.050 (.1029)	0.223*** (.0652)
EMP _{1,80}				14E-7* (70E-8)	-59E-8 (42E-8)
EJOB _{1,80}				38E-5 (22E-5)	19E-5 (13E-5)
BNBE _{1,76-80}					4.562*** (.2408)
Adjusted R ²	.6347	.7902	.7910	.8030	.9353

Note: Standard errors are in parentheses below coefficients.
 Probability t>x: * < .05, ** < .01, *** < .001

Table 6
 OLS Regressions on Stock of New Biotech
 Enterprises born 1981-89 at the beginning of 1990

Constant	1.559*** (.4223)	0.554 (.3425)	0.609 (.3421)	-5.958 (3.357)	-3.297 (2.243)
ACSTAR _{1,76-80}	4.211*** (0.975)	3.085*** (.7604)	2.838*** (.7697)	2.699*** (.7614)	3.601*** (.5108)
ACCOLL _{1,76-80}	0.859* (.4129)	-0.351 (.3833)	-0.336 (.3813)	-0.301 (.3728)	-1.759*** (.2670)
QUAL1 ₁		8.716*** (1.844)	8.096*** (1.868)	8.528*** (1.832)	2.440 (1.287)
FEDGRANT ₁		0.059*** (.0106)	0.052*** (.0113)	0.031** (.0131)	0.014 (.0088)
VENTCAP _{1,80}			0.123 (.0712)	0.011 (.0884)	0.225*** (.0606)
EMP _{1,80}				97E-8 (60E-8)	-59E-8* (42E-8)
EJOB _{1,80}				34E-5 (19E-5)	19E-5 (13E-5)
BNBE _{1,76-80}					3.571*** (.2405)
Adjusted R ²	.6214	.7794	.7819	.7920	.9077

Note: Standard errors are in parentheses below coefficients.
 Probability t>x: * < .05, ** < .01, *** < .001

Table 7
Poisson Regressions of Annual Births of
New Biotech Enterprises, 1976-89

Constant	-1.572*** (.0429)	-1.871*** (.0501)	-2.030*** (.0533)	-4.262*** (.2662)	-4.274*** (.3475)
ACSTAR _{it}	0.160*** (.0145)	0.525*** (.0394)	0.337*** (.0526)		0.394*** (.0520)
ACCOLL _{it}	0.039*** (.0089)	0.069** (.0255)	0.0342 (.0322)		-47E-4 (.0328)
ACSTARSQ _{it}		-0.022*** (.0018)	-0.017*** (.0030)		-0.016*** (.0028)
ACCOLLSQ _{it}		-78E-5 (64E-5)	44E-5 (88E-5)		12E-4 (90E-5)
8689STAR _{it}			-0.269*** (.0769)		-0.360*** (.0750)
8689COLL _{it}			0.146** (.0485)		0.168*** (.0472)
8689STARSQ _{it}			89E-4* (.0040)		0.012** (.0037)
8689COLLSQ _{it}			-20E-4 (.0012)		-27E-4* (.0012)
QUAL1 _{it}			0.213** (.0725)		0.254*** (.0696)
FEDGRANT _{it}			67E-4*** (57E-5)		91E-4*** (81E-5)
VENTCAP _{it}				0.017*** (.0042)	-0.028*** (.0041)
EMP _{it}				18E-8*** (34E-9)	-13E-9 (37E-9)
EJOB _{it}				15E-5*** (11E-6)	12E-5*** (16E-6)
E/PRATIO _t				-0.026 (.0152)	-0.013 (.0193)
Log-likelihood	-1674.9	-1454.1	-1338.3	-1666.4	-1280.1

Note: Standard errors are in parentheses below coefficients.
Probability t>x: * < .05, ** < .01, *** < .001

Table 8
Poisson Regressions of Annual Births of
New Biotech Firms, Subunits, and Enterprises, 1976-89

Dep. Var.	BNBF _{it}	BNBS _{it}	BNBF _{it}	BNBS _{it}	BNBE _{it}
Constant	-4.726*** (.3109)	-5.603*** (.5624)	-4.940*** (.4158)	-5.576*** (.7367)	-4.274*** (.3475)
ACSTAR _{it}			0.420*** (.0631)	0.330** (.1096)	0.394*** (.0520)
ACCOLL _{it}			-18E-5 (.0401)	0.019 (.0677)	-47E-4 (.0328)
ACSTARSQ _{it}			-0.017*** (.0033)	-0.016** (.0059)	-0.016*** (.0028)
ACCOLLSQ _{it}			99E-5 (.0011)	14E-4 (19E-4)	12E-4 (90E-5)
8689STAR _{it}			-0.274** (.0880)	-0.541*** (.1637)	-0.360*** (.0750)
8689COLL _{it}			0.122** (.0563)	0.255** (.0980)	0.168*** (.0472)
8689STARSQ _{it}			87E-4* (.0044)	0.019* (.0080)	0.012** (.0037)
8689COLLSQ _{it}			-18E-4 (.0014)	-44E-4 (.0024)	-27E-4* (.0012)
QUAL1 _{it}			0.270** (.0831)	0.279 (.1468)	0.254*** (.0696)
FEDGRANT _{it}			97E-4*** (95E-5)	92E-4*** (.0018)	91E-4*** (81E-5)
VENTCAP _{it}	0.023*** (.0053)	86E-4 (80E-4)	-0.028*** (.0050)	-0.029*** (.0084)	-0.028*** (.0041)
EMP _{it}	13E-8*** (43E-9)	26E-8*** (65E-9)	-71E-9 (44E-9)	44E-9 (75E-9)	-13E-9 (37E-9)
EJOB _{it}	16E-5*** (13E-6)	15E-5*** (24E-6)	13E-5*** (19E-6)	13E-5*** (34E-6)	12E-5*** (16E-6)
E/PRATIO _t	-0.036* (.0185)	-0.065* (.0326)	-0.018 (.0237)	-0.067 (.0415)	-0.013 (.0193)
Log-likelihood	-1265.1	-529.0	-962.8	-434.2	-1280.1

Note: Standard errors are in parentheses below coefficients.
Probability t>x: * < .05, ** < .01, *** < .001
For this table NBSs are defined inclusive of joint ventures.

Table 9
Poisson Regressions of Annual Births of
New Biotech Enterprises, 1976-89

Constant	-1.583*** (.0429)	-1.410*** (.0606)	-1.643*** (.0848)	-3.805*** (.2527)	-3.787*** (.2622)
NBE _{i,t-1}	0.065*** (.0015)			0.078*** (.0021)	0.053*** (.0029)
NBE _{t-1}		61E-5*** (15E-5)		-0.015*** (.0011)	-0.013*** (.0012)
YEAR _t			0.052*** (.0092)	0.826*** (.0712)	0.763*** (.0740)
BNBE _{i,t-1}					0.178*** (.0115)
Log-likelihood	-1713.4	-2181.9	-2173.2	-1599.9	-1516.2

Note: Standard errors are in parentheses below coefficients.
Probability t>x: * < .05, ** < .01, *** < .001

Table 10
Poisson Regressions of Annual Births of NBEs, 1976-89

Constant	-3.823*** (.2635)	-4.415*** (.5020)	-5.593** (.5395)	-5.613*** (1.788)
ACSTAR _{it}	0.247*** (.0578)		0.225*** (.0561)	0.255*** (.0538)
ACCOLL _{it}	-30E-5 (.0329)		45E-4 (.0330)	-69E-4 (.0325)
ACSTARSQ _{it}	-0.018*** (.0034)		-0.014*** (.0032)	-0.016*** (.0029)
ACCOLLSQ _{it}	11E-4 (93E-5)		45E-5 (93E-5)	10E-4 (88E-5)
8689STAR _{it}	-0.351*** (.0827)		-0.273*** (.0812)	-0.328*** (.0747)
8689COLL _{it}	0.107* (.0487)		0.065 (.0501)	0.088 (.0480)
8689STARSQ _{it}	0.013** (.0043)		86E-4* (.0041)	0.011** (.0038)
8689COLLSQ _{it}	-11E-4 (.0012)		-35E-5 (.0013)	-11E-4 (.0012)
QUAL _{it}	0.284*** (.0814)		0.293*** (.0817)	0.289*** (.0813)
FEDGRANT _{it}	55E-4*** (60E-5)		85E-4*** (80E-5)	86E-4*** (79E-5)
VENTCAP _{it}		-0.018*** (.0034)	-0.018*** (.0044)	-0.016*** (.0043)
EMP _{it}		26E-8*** (26E-9)	-11E-8** (36E-9)	-12E-8E** (36E-9)
EJOB _{it}		12E-5*** (16E-6)	15E-5*** (16E-6)	15E-5*** (16E-6)
E/PRATIO _t		-0.196*** (.0283)	-0.149*** (.0300)	-0.144*** (.0298)
NBE _{1,t-1}	0.060*** (.0069)	0.037*** (.0038)	0.046*** (.0075)	0.052*** (.0068)
NBE _{t-1}	-0.010*** (.0012)	-0.018*** (.0013)	-0.015*** (.0014)	-0.015*** (.0014)
YEAR _{it}	0.617*** (.0756)	0.939*** (.0748)	0.805*** (.0776)	0.802*** (.0774)
BNBE _{1,t-1}	-66E-4 (.0181)	0.169*** (.0136)	0.034 (.0190)	
Log-likelihood	-1262.8	-1368.8	-1197.0	-1198.5

Figure 1. Total Number of Articles Reporting Genetic Sequence Discoveries Authored by Stars (1967-1989)

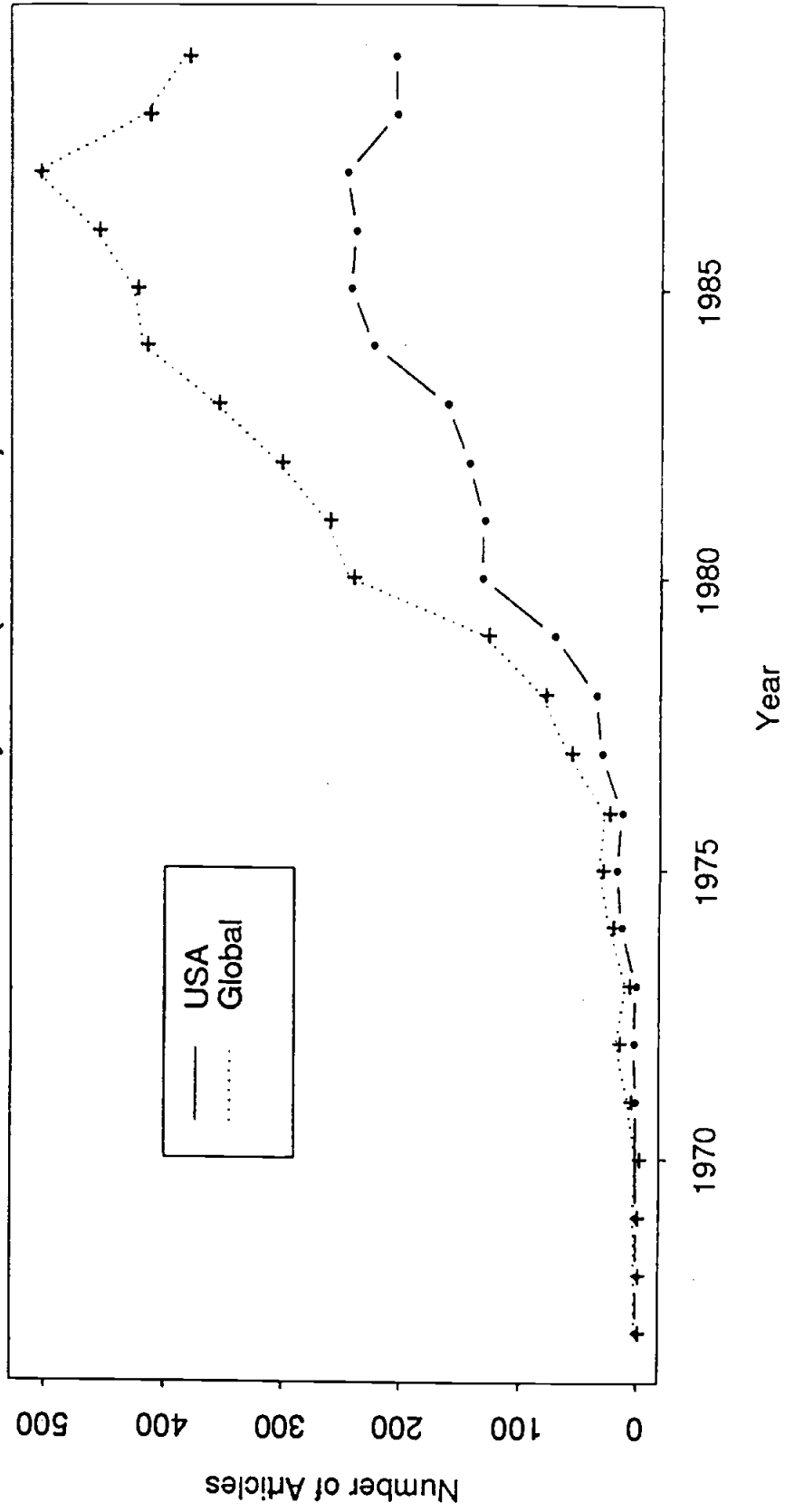


Figure 2. Percentage of Coauthors from the Same Institution as Stars (1967-1989)

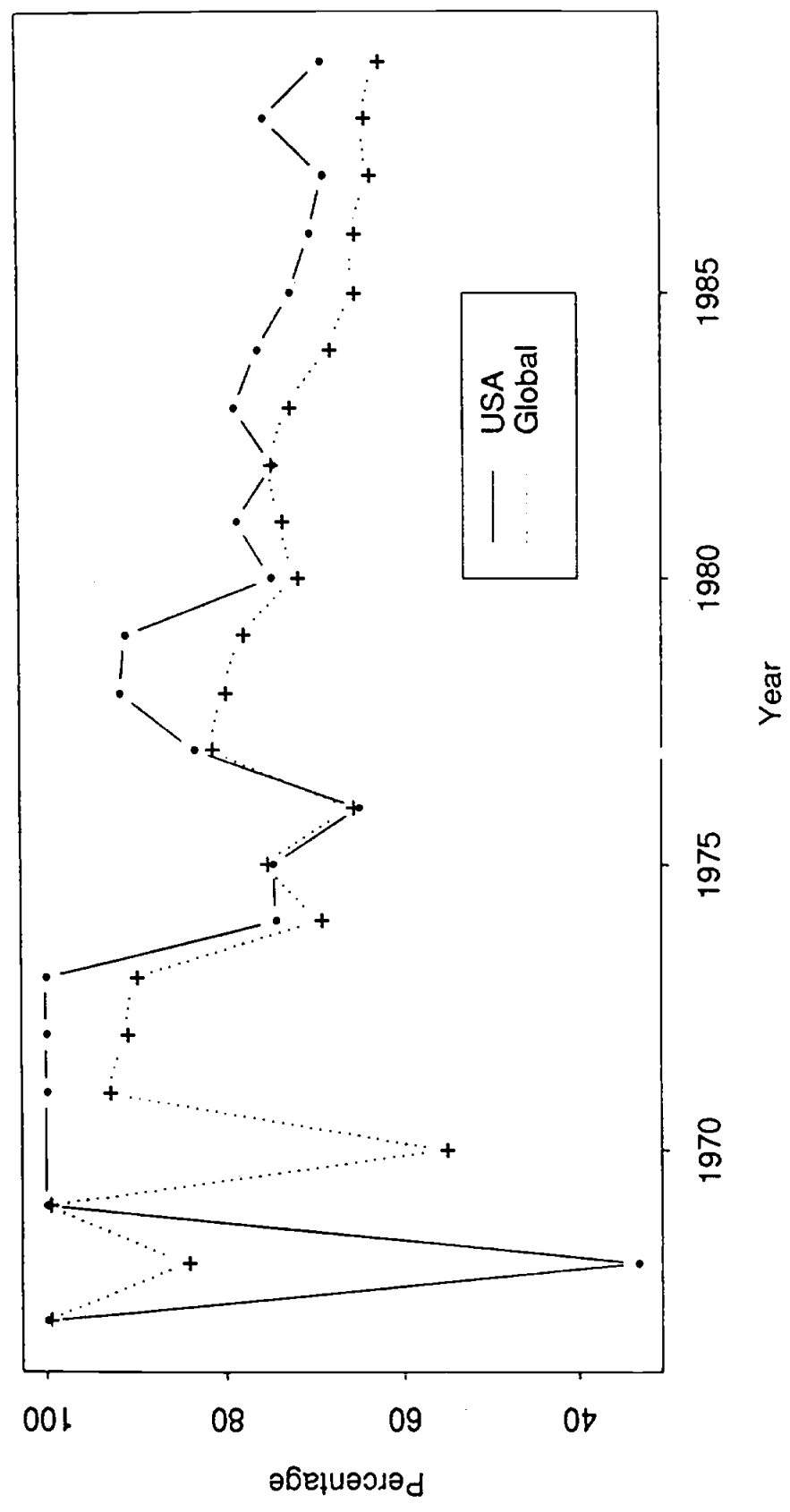


Figure 3. Average Number of Authors per Article Published by Stars (1967-1989)

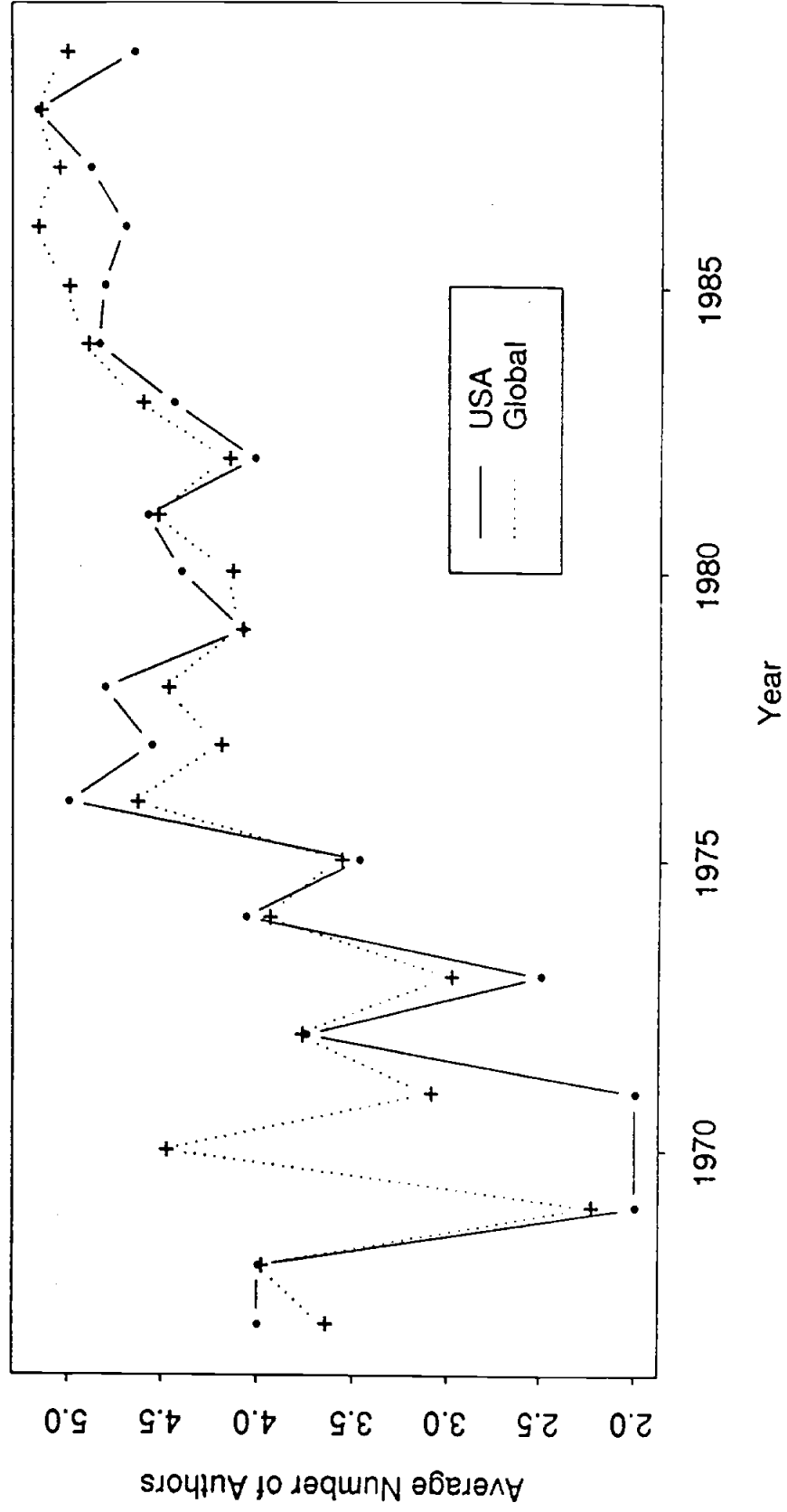


Figure 4. Cumulative Number of Stars, Collaborators, and New Biotech Enterprises in the U.S. (1967-1989)

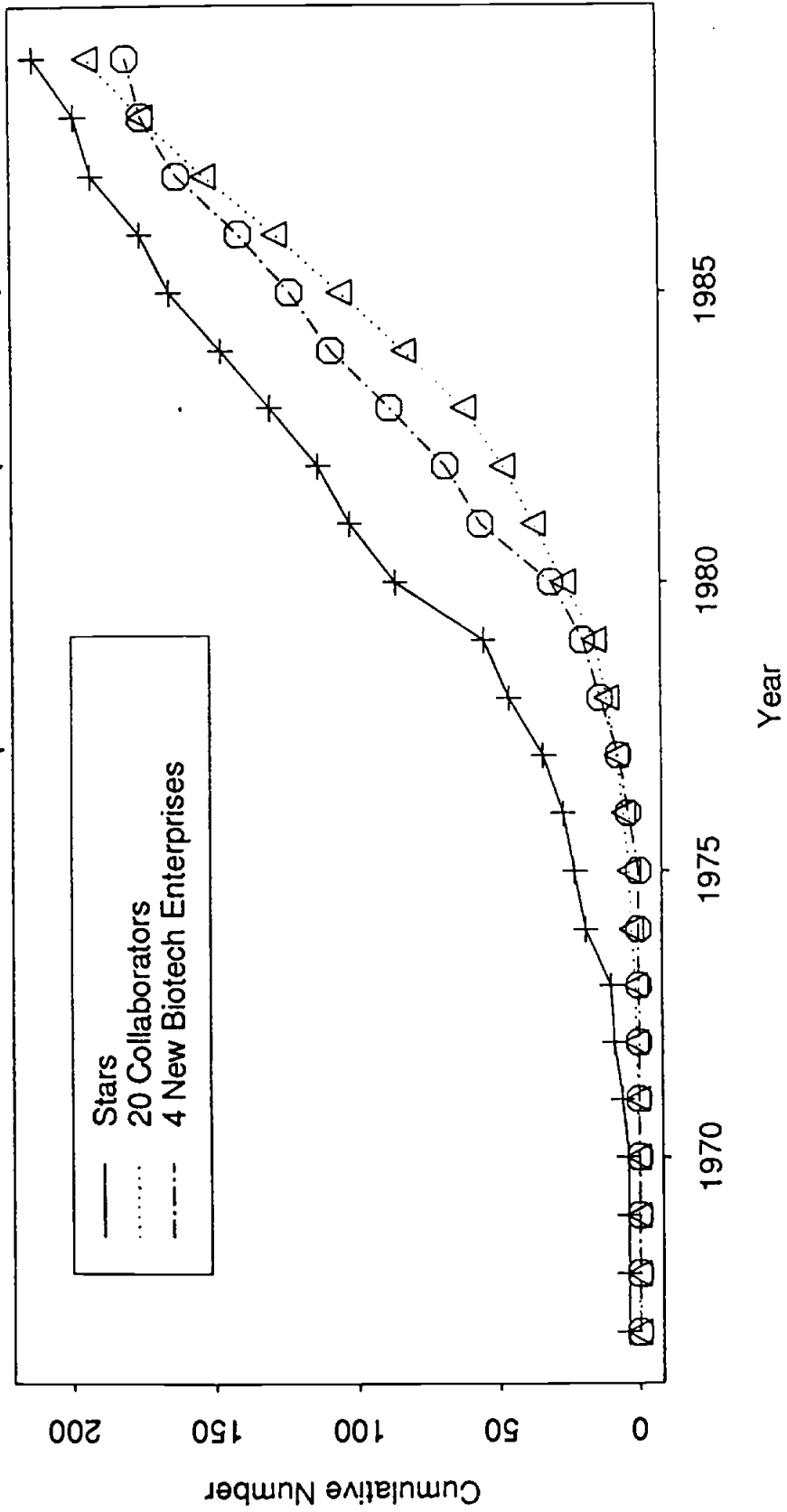


Figure 5: Active Stars and New Biotechnology Enterprises as of 1990

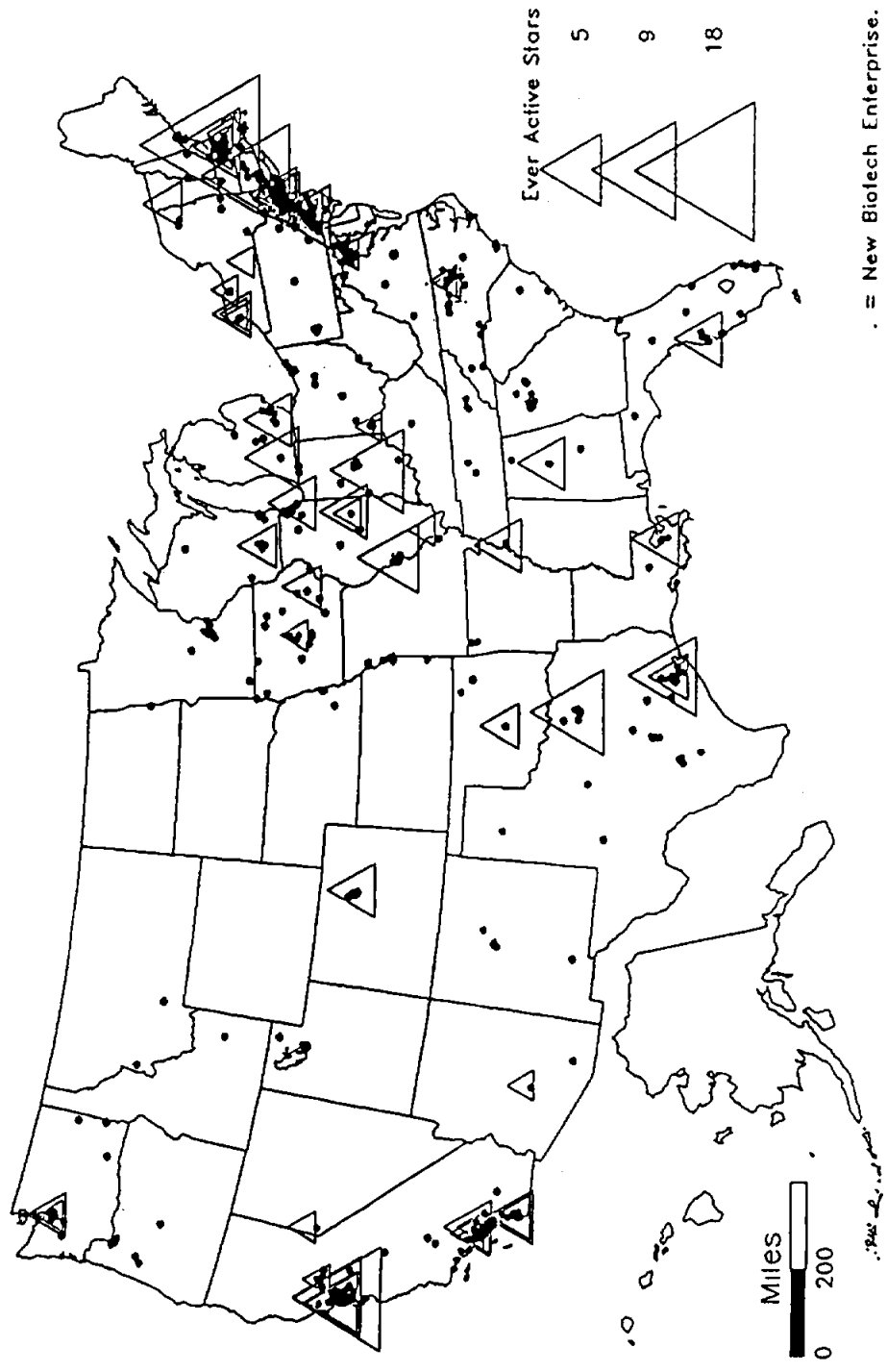


Figure 6 Percentage of Active Stars and Collaborators in Universities, U.S. only (1967-1989)

