GOING PUBLIC WHEN YOU CAN IN BIOTECHNOLOGY

Michael R. Darby
Lynne G. Zucker

Going Public When You Can in Biotechnology
Michael R. Darby and Lynne G. Zucker
NBER Working Paper No. 8954
May 2002
JEL No. O31, G32, L21, L65, M13

ABSTRACT

Scientist-entrepreneurs prominent in biotech and other high-technology industries view going public not as a cost-effective source of capital but as a cross between selling a now-proven innovation and winning a lottery. Unlike most empirical IPO analyses confined to those firms that go public, we study substantially all the non-public biotech firms founded up through 1989. The probability that one of these firms goes public in any given year increases with the quality of the firm's science base (use of recombinant DNA technology, number of articles by star scientists as or with firm employees, number of biotech patents), the percentage of eligible firms going public the year the firm was founded as a strategy indicator, recent biotech returns as an indicator of a hot market, and whether or how many rounds of venture capital has been obtained. The same key factors increase the expected proceeds raised from IPOs, but the quality of the firm's science base plays a more dominant role. All firms going public try to look like the next Genentech, but only those with the strong science base necessary for success attract large investments.

Michael R. Darby
Cordner Professor of Money & Financial Markets
Depts. of Management, Economics & Policy Studies
Anderson Graduate School of Management
University of California, Los Angeles
Los Angeles, CA 90095-1481
and NBER
darby@ucla.edu

Lynne G. Zucker
Professor of Sociology &
Director, Center for International Science,
  Technology, and Cultural Policy
School of Public Policy & Social Research
University of California, Los Angeles
Los Angeles, CA 90095-1551
and NBER
zucker@ucla.edu

Going Public When You Can in Biotechnology

by Michael R. Darby and Lynne G. Zucker

> *Finance Professor:* What were the factors which made it optimal to
> raise money through an initial public offering at this time?
>
> *New Biotech Millionaire:* Because I could.

This joke sums up a disconnect between the academic literature on the one hand and both Main and Wall Streets on the other. The scientist-entrepreneurs prominent in the biotech and other high-technology industries view going public not so much as a sometimes reasonably priced alternative source of working capital as a cross between a well-earned sale of a now-proven innovation embodied in a firm and the winning of a lottery.

The use of modern biotechnology mushroomed from nonexistent in 1975 to over 700 firms by 1990 in the United States alone. Of those firms, some 512 were new entrants, most of which were attempting both to successfully apply the new bioscience breakthroughs to create commercially valuable products and to go public insuring the survival of the firm and wealthy founders and early investors. By 1992, 162 of these firms had traded publicly including 156 (30.5 percent of the 512 entrants) which made an initial public offering (IPO).[1] These firms are an attractive natural experiment for understanding the IPO process.

This paper is unique in beginning with a universe of non-public firms and analyzing the determinants of the probability a firm will go public in a given year and, if it does, the amount of capital which it raises in the IPO. Deeds, Decarolis, and Coombs (1997) and Stephan and Everhart (1998) have previously examined the amount raised in biotech IPOs for those firms that go public during their time frame, but we are able to add to and improve on their analyses in several ways.

First, because of our rich data resources from the larger biotechnology project, we are to examine all IPOs occurring over a much longer period than could either of those research teams. Second, because we have the universe of private firms, we are able to apply the Tobit technique to properly estimate the funding amounts conditional on the probability that a firm goes public.

Finally, and most important, we have much more information on the scientists involved in firms and are able specifically to identify publications with star scientists that are joint with the firms. These joint publications identify the labor effort that has been moved over to the firm by top scientists, those scientists who have made breakthrough discoveries in genetic sequencing.[2] The importance of this labor effort has been documented in a series of substantive analyses, predicting a wide variety of performance measures from patents and products to employment growth. In our most recent analyses, we examine publications joint with firm scientists by all scientists in the relevant science areas and by the stars (Zucker, Darby, and Armstrong 2002). We find that stars' joint publications provide a significant increment above the significant positive effect of all scientists that are co-publishing with firm scientists on a variety of performance measures, especially strong and consistent in panel analyses.

Our story, then, is that access to tacit knowledge held by the discovering scientists has a strong and consistent impact on firms' performance. Such tacit knowledge is best communicated from discovering to firm scientists by working jointly at the lab bench. Learning-by-doing-with limits the speed of diffusion of the knowledge, conferring natural excludability. We find traces of this process and its strong impact in both biotechnology and semiconductors (on semiconductors, see Torero, Darby, and Zucker 2001). Number of articles or citations measures of their quality fail to capture the process of transmitting tacit knowledge held by the discovering scientists that our co-publication measured has been designed to uncover.

We argue that the strong positive effects of access to this tacit knowledge on firm performance alters the financial prospects of the firm, making it easier to find capital and to obtain it in larger amounts. This is true of early funding by venture capitalists and of funding through an IPO. Since firms with collaborating relationships with star scientists are more likely to go public, censoring firms that do not go public means that the total effects of stars on financial success of firms are seriously underestimated. Here, we investigate the full process, from birth through IPO, of a substantial portion of those dedicated biotech firms that ever go public.

I. Hypotheses

Biotechnology is characteristic of many industries born based on scientific and/or technological breakthroughs that make it possible to do what was before impossible or impossibly uneconomical. Literally hundreds if not thousands of firms entered each time to exploit such breakthroughs as the internal combustion engine, the integrated circuit, biotechnology, and, most recently, the internet. As the industries mature and rapidly grow, a relatively small number of firms grow rapidly while most stagnate, shrink, or exit through merger or failure.

The entrant biotech firm must win or at least place in repeated rounds of beauty contests if it is to attract sufficient funding from angel investors, venture capitalists, and ultimately public investors to fund the long years of research and clinical trials which typically lie between brilliant idea and marketed, revenue-producing product. The firms that will be most successful in passing through rounds of financing and simultaneously achieving sufficient R&D success to be an attractive contestant will be those with the deepest scientific base.

For private firms, we do not have the rich data sets available to students of public firms, but our previous work has identified some key variables which serve as useful indicators of a high-quality science base and of interim progress toward ultimate profitability. As discussed above, the single most powerful indicator of ultimate success is the active working involvement of the star scientists who – although small in number -- are responsible for a large fraction of the most important discoveries. When these scientists become involved in commercializing their discoveries, they frequently have the insight and scientific taste to identify the sweet spot where scientific possibility and economic payoff are combined. A simple indicator of the level of

technological sophistication of a biotech firm in the 1980s was whether or not it was actively using the recombinant DNA (rDNA) technology, also referred to as genetic engineering. The presence of nearby top-quality universities appears to be valuable also, although it is not clear whether this represents the presence of stars unidentified by our empirical definition, geographically localized knowledge spillovers, or the pecuniary externalities inherent in subsidized training of labor and geographic agglomerations generally.

We usually use patents granted as an indicator of R&D productivity although it is often remarked that it may be a better indicator of past R&D expenditures than production of valuable innovation (Griliches 1990). Since we do not have a history of R&D expenditure data available for these private firms, this collinearity may be a blessing despite implying some difficulty iin interpretation.

Unlike these proven scientific indicators, we also try in some regressions a count of the number of SBIR grants received by the firm.[3] The number of SBIR grants received might be expected to reduce the probability of remaining private either because the resources permit creation of more intellectual property or serve as a government certification that the firm is doing good science (the "halo effect") or for both reasons. Venture capitalists sometimes see this and similar government programs as competitive, so including venture-capital and SBIR funding in the same regression permits us to assess whether these funding sources are equally valued in the financial markets.

We also expect that the firm's history of receiving venture capital investments will have a positive impact on the probability of the firm going public. First of all, venture capitalists are investors themselves and their funding connotes their expectation that they will be able to exit after a few years via a successful public offering. Their investment and due diligence also

provide a behavioral signal upon which uninformed investors can cascade (Banerjee 1992, Bikhchandani, Hirshleifer, and Welch 1992, and Welch 1992).

As is well known there are waves – or "windows" – in which financing of particular technologies is in vogue and IPOs are particularly favorable for firms that are "ready" to go in terms of a technology at least a reasonably proven in the laboratory. We thus expect firms more likely to go public in periods following rapid increases in biotech stock prices.

We include a final variable to help distinguish between two types of firms: those that are founded in hopes of becoming a major player by proceeding trough rounds of private financing to the IPO and beyond and those that in fact are intended by the proprietor to remain indefinitely a small or medium sized firm with a narrow scope and private ownership. The only variable we have found which distinguishes at all effectively between firm types is the fraction of private biotech firms going public in the year of founding. Our fieldwork suggests that number of scientists founded eventually successful firms in order to imitate if not outdo the commercial success of the suddenly rich rival down the university corridor.

IPOs for high-technology firms are things of remarkable beauty: Investors provide large amounts of money to firms with few employees and relatively few assets beyond some lab equipment and the brains of the personnel represented by accumulated losses which have proven – more or less – the plausibility of the firm's nascent technology succeeding in the market. Investors in these firms seek to diversify their portfolios in the hope that 2 or 3 winners will much more than compensate for the inevitable 40 or 50 losers. Not surprisingly, a significant minority of IPOs involve firms which are more or less fraudulent. We hypothesize that investors will foresee this possibility and provide more funding to firms in which star scientists are actually involved in bench

level science for the firm (indicated for us by joint publications) than merely lending their name to a scientific advisory board.

Aghion and Tirole (1994) provide a more general analysis of why investors would place a higher value on firms in which the research principals are deeply involved. Deeds, Decarolis and Coombs (1997) give a signalling interpretation to their finding that proceeds increase with the number of citations to publications authored by the full-time executives and employees of the firm. Stephan and Everhart (1998) confirm for two years of IPOs the value of highly cited scientists associated with the firm. We have emphasized the importance of having the best scientists to the real productivity of the firm's research and development efforts (Zucker, Darby, and Armstrong 1998, and Darby, Liu, and Zucker 1990).

Without necessity to choose among these motivations here, we hypothesize that the amount of IPO proceeds raised if and when the firm goes public will be greater the deeper the science base of the firm.

We also expect that the firm's history of receiving venture capital investments and p"hot market" conditions indicated by prior high biotech returns will have positive impacts on the amount the firm is able to raise for very much the same reason that each factor increases the probability that the firm goes public.

II. Data

Estimating the probability that private firms go public is inherently data constrained since by definition they are under no compulsion to make public disclosures in the systematic and detailed ways that public firms must. However, a surprising amount of information can be gained by coding the data contained in commercial directories which serve to attract customers and by matching to this data information on recipients of venture capital, assignees of patents, and affiliations of authors of scientific papers covering areas closely related to the firms' technologies. Since 1988 we have been building a matched, linked, and cleaned data base for examining the interaction among firms, universities, research institutes, and scientists involved in biotechnology. Detailed description of the data used in this paper is reported in the Data Appendix. Here we outline the sources and definition of the main variables as a way of introduction to the empirical results and temptation to delve into the data appendix.

The principle themes of our research are: (1) Biotechnology -- like many new scientific breakthroughs with commercial value – was characterized by tacit knowledge transferable only by working at the bench level with scientists who have mastered its practice – often as the original discoverers but sometimes through sophisticated reverse engineering. Firms that wanted to practice the key technology of genetic engineering had to learn its techniques from someone who already knew them. (2) The very best "star" scientists were the most effective engines of transfer and guides to commercial application of the new methods.

The scientists developed a central data depositary GenBank for reporting genetic-sequence discoveries and reporting was effectively enforced worldwide by requirements at all the major journals that appropriate GenBank accession codes be reported prior to publication for any article

involving such discoveries. We used this GenBank to identify the 307 most productive researchers in the world defined as those who had discovered over 40 sequences by April 1990. At the insistence of bioscientist critics, we added another twenty stars on the basis of publishing twenty or more articles reporting genetic sequence discoveries.[4] We hand collected the stars articles and identified all the authors and their affiliations on each article. This is the source of our count of "tied" articles written by a star as or with a firm's employees, a variable which has proved a powerful indicator of star involvement and hence future firm success.

We similarly compiled a universe of 752 biotech firms that had operated at some time through 1990 and matched that list to assignee names in the patent records and to information on the geographic region(s) in which the firms operated. We also coded whether the firms reported using the recombinant DNA technology (rDNA) in the industry directories or were licensed by Stanford and the University of California to do so.

We obtained data on the date of IPOs and the proceeds raised using *Bioscan* (1989-1997), its precursor Cetus Corp. (1988), the *IPO Reporter*, and the Securities Data Company (1998a) *Global New Issues* on-line electronic data base. *Moody's* manual and purchased copies of tIPO prospectus were used to resolve a few instances of conflict among these sources.

We constructed a biotech equity index based on a portfolio consisting of all publicly traded biotech-using firms (including incumbent adopters) in our biotech firm data set. Whether and how many rounds of venture capital were received was obtained by use of the Securities Data Company (1998b) *VentureXpert* on-line electronic database.

III. Empirical Results

In this section we first report estimates on the probability per period that a private biotech firm will go public and assess the cumulative impact of the firm's science base on this probability. We then estimate Tobit regressions explaining the amount raised given that the firm goes public. The effects of the firm's science base are even stronger for proceeds than for duration to IPO. Finally, we examine what appears to be a reciprocal causal relationship between the firm's science base and its receipts of venture capital funding.

III.A. Duration to IPO Results

In Table 2 we report estimates of standard Weibull survival models explaining the duration from the founding of the firm to the firm's initial public offering (if any). In the (unfortunate) terminology used for these models positive coefficients indicate greater probabilities of surviving as a private firm so negative coefficients indicate a greater probability of exiting – that is, making an initial public offering.

In column (a) of Table 2 we present our simplest model in which the duration to IPO is predicted by six variables: three indicators of the firm's science base, the percentage of biotech firms going public the year the firm was founded, and two indicators of whether there is a "hot" biotech market. The three variables describing the firm's science base are the number of articles written to date by star scientists as or with a firm employee, the number of eventually granted patents applied for to date, and whether or not the firm uses rDNA technology. Despite some multicollinearity, each of these variables is significant at the 5 percent or better level, and the chi-squared test at the bottom of the table indicates that we can reject the joint hypothesis that all the science-base coefficients in this model (or any of the models in Table 2) are zero at better than the

0.001 level of confidence.   The percentage of biotech firms going public in the year the firm was founded is also very significant, indicating that a high IPO rate leads to imitative entry of firms pursuing the same strategy. Finally, whether the biotech market has experienced high returns in the previous year also very significantly reduces the probability that a private firm will remain private. Thus these results all closely correspond to our basic hypotheses about what drives the process of high-tech firms going public.

In models (b) and (c) we consider the issue of whether in fact the science base is important only because it attracts venture capital support or whether it works separately in terms of pure productivity of producing new, valuable intellectual property. After some experimentation, we found that the best available indicators of venture capital support were simply the dummy variable for ever having such support or the count of the number of rounds of support received. All the indicators are highly correlated, and we suspect that omissions in the SDC survey coverage introduce significant measurement error into the alternative venture-capital-funds-received variables which we tried.   We introduce the dummy and cumulative rounds separately because multicollinearity becomes severe when both are in the same equation. In any case, we get similar results: The venture capital variables are indeed highly enough correlated with the science-base variables to make it difficult to measure their separate effects robustly, but – based on the chi-squared test – we see that both science-base variables as a group and venture capital support increase the likelihood of going public. The results for the strategy and market indicators are qualitatively the same for all the models reported in Table 2.

In models (d) and (e) we report the results from experimentation with other science base indicators which we have used in our other work on biotechnology. The number of SBIR grants received might be expected to reduce the probability of remaining private either because the

resources permit creation of more intellectual property or serve as a government certification that the firm is doing good science (the "halo effect") or for both reasons.[5] Similarly the number of nearby universities with top-ranked biotech-relevant doctoral programs serves as an indicator of geograpically localized knowledge spillovers and/or favorable labor cost conditions.[6] Again, the science-base coefficients are all negative as expected. However, although as a group they are highly significant in explaining the probability of going public, the individual coefficients are not robustly statistically significant except for whether or not the firm does genetic engineering (uses rDNA technology).

One way to interpret the survival models is to simulate them using different assumptions for the values of the determining variables. This is done in Table 3 for model (e) of Table 2. Table 3 reports the predicted number of years from firm founding that it would take for a cohort of firms with given characteristics to reach various percentages of firms having gone public.[7] The base case assumes for each year of the simulation that the surviving firms have the mean values for firms their age of each of the variables in model (e). Case (c) sets all the science base variables at 0 instead of their mean values and reruns the simulation. The difference between case (a) and case (c) indicates the estimated reduction in duration to IPO due to the combined effects of all the firms' science-base variables. Case (b) is like case (a) except it sets at 0 all the science base variables other than the number of Articles to Date by Stars with Firm ("ties") and the indicator for Use of rDNA Technology ("rtech"). Comparing all three cases, we see that overall the science-base variables reduced the time to IPO by a little over a third and a little more than half of this reduction is attributable to the number of articles the firms have authored with stars and to the use of genetic engineering. Figure 1 plots the simulated effects in Table 3, illustrating

graphically the economically significant effect of the firm's science base holding constant the market-condition variables and the firm's strategy and receipts of venture capital.

Taken as a whole, Table 2 implies that market-conditions and the firm's science base, strategy, and receipts of venture capital all play significant roles in determining the probability per year that a particular private biotech firm will go public. We cannot robustly characterize which science-base indicators are the most important, but the use of genetic engineering was always significant in determining which firms were able to go IPO when the market conditions were favorable.

III.B. Results on Proceeds from the IPO

Going public is much like the consumer buying a car, there is not only the of whether it is done at all, but also, if it is done, how much money is involved. Accordingly, we turn to estimating the funds raised in those IPOs that do occur using the Tobit technique. The results are reported in Table 4. Although the reported coefficient estimates do not have the easy interpretation of OLS regressions, significant positive coefficients do mean that the variable results in a predicted increase in IPO proceeds if an IPO occurs.

Models (a), (b), and (c) estimate the Tobit regression with the core science base variables (articles by stars with or as firm employees, patents granted applied for by this date, and use of rDNA technology), two-year lagged biotech returns as our market indicator, and no, dummy, or round count venture-capital variables.[8] Unlike the survival models, we obtain distinct highly significant, positive coefficients for each of these variables in explaining how much money the firm can raise. The market indicator and venture capital variables also have robustly significant, positive coefficients in Table 4.

In models (d) and (e) we add the SBIR and local top-quality universities variables, but only the latter has significant, positive coefficients. It appears that investors in initial public offerings are much more impressed by investments by other private investors than by government bureaucrats allocating SBIR funds.

Table 5 computes the partial derivatives of expected IPO proceeds with respect to each of the variables in the Table 4 models assuming in each case that all the variables are at their sample-mean values. Each article written by a star as or with a firm employee increases IPO proceeds by from $0.9 to $1.3 million depending on model specification. The effect on the value of the firm will typically be a significant multiple of these derivative amounts since only a fraction of the equity value is sold at IPO. Each patent is worth between $1 and $2 million as is the use of rDNA technology. A nearby top-quality university adds from $0.5 to $0.7 million to IPO proceeds. Receiving venture capital is worth about $3.3 million in IPO proceeds as a yes or no matter or about $0.6 million per round of venture capital received. Higher prior returns also increase current IPO proceeds.

As with the probability of going public, the firm's science base, market-conditions, and receipts of venture capital all play distinct and significant roles in determining the amount of money raised by the firms that indeed do go public. In this case, we can robustly characterize the impact of four distinct indicators of the science base, with only SBIR grants failing to make a significant positive contribution.. Also as with the probability of going public, the use of genetic engineering an prior high biotech returns always significantly increase expected IPO proceeds.

III.C. Granger Causality Analysis of the Firm's Science Base and Venture Capital Funding

Table 6 reports the results of Granger causality tests to see whether venture-capital funding enables the firm to engage in more joint research with star scientists or whether more joint research with star scientists enables the firm to get venture-capital funding. As is often the case with these experiments the results are somewhat ambiguous, but we believe they are most consistent with the view that the variables are mutually reinforcing in a virtuous circle in which more of each increases the other.

The particular difficulty in using the Granger methodology here is that we do not have a clear dating of when the work which culminates in an article published on a given date was actually done. The firm would be sure that potential venture-capital funders were aware of star involvement in the firm while it was actually in process as opposed to completed, submitted, revised, and eventually published. Table 6 reports Granger causality tests assuming that the work was done precisely two years before the article was published. In this case we can reject both the hypothesis that star articles does not cause venture-capital funding and the hypothesis that venture-capital funding does not cause star articles, implying reciprocal causality. If one believes that the R&D work plus publication must have significantly exceeded two years, then the hypothesis that venture-capital funding does not cause star articles becomes tenable. If, on the other hand, one believes that the R&D work plus publication takes only one year, then the hypothesis that star articles do not cause venture-capital funding becomes tenable.

We see reciprocal causality as the most reasonable interpretation of the results. On the other hand, based on extensive interviews with the scientists involved, we find it much more plausible that the work plus publication process significantly exceeds two years than that it is significantly less. Thus, there is stronger evidence for star articles to venture funding than for the reverse.

15

IV. Conclusions

This paper starts from a universe of private biotech firms and estimates survival models to explain the probability each year that a particular firm will go public. We find that a stronger science base, venture capital funding, and a hot market all increase the probability that the firm will go public. Our most surprising finding on duration to IPO was the power of imitation -- or envy: Firms founded in years with exceptionally high IPO rates are themselves significantly more likely to eventually go public.

We also find that a stronger science base, venture capital funding, and a prior hot market all increase the expected amount of money that the firm will raise if it does go public. In the case of proceeds raised, we can estimate strongly significant and distinct positive contributions for four of our five science-base indicators – the only exception being the number of SBIR grants received. Each article written by a star scientist as or with a firm employee increases IPO proceeds by around $1 million and each patent by $1-2 million. Each nearby top-quality university or round of venture capital adds about $0.6 million to IPO proceeds. Thus star articles are worth nearly as much as a patent in terms of firm value and about twice as much as a round of venture-capital financing. As to the chicken-egg question, the evidence is most consistent with mutual (i.e., two way) star-articles and venture-capital causation, but if we had to choose one-way causation our reading would be that star involvement in a firm is more likely to bring venture-capital financing than is venture-capital financing to attract star involvement.

We conclude that available data on the financing and science base of private biotech firms is sufficiently rich to permit meaningful estimation of both duration and proceeds models for IPOs. It remains for future research to apply the same methods to other high-tech industries.

Data Appendix

In this paper we use Zucker-Darby data sets developed for the Project on "Intellectual Capital, Technology Transfer and the Organization of Leading-Edge Industries: The Case of Biotechnology." The basic data sets developed for this project are described in Zucker, Darby, and Brewer (1994, 1998); details of how we defined star ties to firms through co-authorship on scientific articles—either by affiliation to the firm or by linkage via co-authorship with firm employees—can be found in Zucker, Darby, and Armstrong (1998). These data have been or will be archived upon completion of the project in the Data Archives at the UCLA Institute for Social Science Research (ISSR).

A.1. Firms "At Risk" for an IPO

The firms at risk for going public are the 512 NBFs, among the 752 identified U.S. biotech enterprises. The remainder of the firms are either incumbents (150) or could not be classified clearly into either subcategory (90).

Each NBF is geo-coded as being located in one of the 183 functional economic areas (BEA areas) as defined by the U.S. Bureau of Economic Analysis.

A.2. Measures of Science and Invention

We use five measures of "science-ness" of the firm: cumulative number of articles with affiliated and tied star scientists, cumulative number of patents granted (using the application date), use of rDNA technology by the firm, cumulative number of SBIR (Small Business Innovation

Research) grants awarded to the firm, and the number of local—in the same BEA area—top-quality universities.

Star Scientists: Zucker, Darby, and Brewer (1994; 1998) demonstrate the key role played in determining where and when NBEs are founded by intellectual capital, in particular by where and when leading frontier scientists ("stars") are actively publishing. Given the fundamental role of rDNA in modern biotechnology, a very important measure of research success in the basic science 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.[9] Worldwide 327 leading researchers (the "**stars**") were identified on the basis of the number of genetic sequence discoveries (over 40) and articles reporting them (20 or more) up to April 1990 for which they were an author. Of these, 107 published at some time in the U.S.

These 327 stars were listed as authors on 4,061 distinct articles in major journals. These articles were hand collected and used to identify and locate institutional affiliations at the time of publication for each of our stars and their coauthors who were either other stars or "collaborators" (6,082 scientists worldwide). This hand coding was necessary at that time because available machine-readable data bases gave only the location of the first author who, given the authorship conventions of the field, is rarely a star scientist.

In an extension of that work (Zucker, Darby and Armstrong 1998), we introduced the concept of a scientist become "tied" to the firm. We weight each star by the number of articles each has published with firm scientists. Tied scientists are a combination of those that list the firm as an affiliation (we call these elsewhere "affiliated scientists") and those that list a scientist from that firm as a co-author ("linked scientists").

Patent Data: We obtained the patent data for each biotech firm from the U.S. Department of Commerce (1993) CD-ROM *Patent Technology Set: Genetic Engineering* produced by the Office of Electronic Information Products and Services of the U.S. Patent and Trademark Office. All patents were included in our analyses: class 935 (all subclasses) or class 435 (subclass 172.3 only) of the U.S. Patent Classification System.   Class 935 is entitled Genetic Engineering: Recombinant DNA Technology, Hybrid or Fused Cell Technology and Related Manipulations of Nucleic Acids. It is a collection of patents which disclose: (a) processes of altering the genetic structure of microorganisms, (b) genes and methods of modifying genes and their expression, (c) vectors and methods of modifying vectors, (d) methods of introducing DNA into a cell, (e) microorganisms per se which have had their genetic sequences altered by recombinant DNA techniques or by cell fusion or by uptake of DNA, (f) testing, (g) separation techniques (h) apparatus, (i) methods of use of vectors or of the genetically engineered microorganisms, and (j) methods of gene therapy or genetic modification of living organisms. Class 435, subclass 172.3, includes patents which disclose methods and materials of recombinant DNA technology. Thus, these patent classes cover the core inventions of modern biotechnology.

The Patent Technology Sets include patents which issued through January 26, 1993, and which were included in above classifications as of June 29, 1993. We counted the number of patents granted to the firm using year of application date through 1991. We used year of application because the work underlying the patent had by then been done and would be available for disclosure -- and normally would be disclosed -- to the financial markets in the event of an IPO. We also calculated the total patents applied for by each firm from 1976 to 1991 which were later granted.

rDNA Technology: We also coded (1 yes, 0 no) whether the firms reported using the recombinant DNA technology (rDNA) in BioScan, the leading commercial directory of biotech-

producing firms, or were licensed by Stanford and the University of California to do so. This rDNA technology is a marker of a high-science firm.

SBIR Awards: Small Business Innovation Research grants are a set-aside program of all Federal agencies that make any research grants.

Top Quality Universities: There are 20 U.S. universities with biotech-relevant departments receiving the highest ratings (rated above 4 on a scale of 5) on overall scholarly quality as reported in the 1982 National Academy of Sciences' reputational survey of doctoral programs (Jones, Lindzey and Coggeshall 1982).[10] Biotech-relevant departments included biochemistry, cellular/molecular biology, and/or microbiology.

A.3. IPO Variables

In our study, we define IPO proceeds as the capital raised by company in the IPO before transaction fees and including known over-allotment. There are four main sources from which we obtained biotech IPO dates and proceeds data: (1) *BioScan* (1989-1997); (2) BioScan's precursor Cetus Corp. (1988); (3) *IPO Reporter*, a weekly publication: and (4) *Global New Issues,* an on-line electronic data base produced by Securities Data Company (1998a) with data on over 72,000 U.S. IPOs since 1970.

In most cases, the dates and proceeds from different sources matched. There are several cases in which we did not have matched IPO data: (1) Data missing in one or more of our sources generally could be replaced with data from Global New Issues that was the most comprehensive in coverage; (2) Different data was occasionally reported in different sources, but generally we could use data from Global New Issues to determine which values were most accurate. In most cases, differences could be identified as due to rounding errors, transaction fees, over-allotments, or

foreign proceeds (which we excluded). For those cases where we could not reconcile the differences by referring to our four primary sources, we checked the *Moody's* manual. Generally, *Moody's* would verify one of the data values we had. In three cases, we had to purchase copies of the IPO prospectus in question to obtain cleaned and verified IPO data.

In our search, we found that 162 of the 512 biotech entrants began public trading during the 1976-1992 period. We dropped 6 of these companies that became publicly traded without an IPO through spin-offs or merger of a private company with one already trading publicly. We ended up with 156 biotech IPOs between 1976 and 1992. For each of these cases, we obtained the exact IPO date and the proceeds raised.

We counted the number of biotech IPOs in each year from 1976 to 1992 at the U.S. level and then subdivided the total into 183 regional observations for each of the functional economic areas ("BEA areas" for short) into which the U.S. is divided by the Bureau of Economic Analysis in U.S. Department of Commerce (1992). We also calculated the total proceeds raised in IPOs in each year at both levels. In our study, we also examined use of the cumulative number of biotech IPOs and the cumulative value of biotech IPO proceeds as control variables.

For all IPO value variables, we calculated their deflated values in terms of 1983 U.S. dollars by using the Consumer Price Index for All Urban Consumers as reported in the *Economic Report of the President* (February 1995). We also considered the equivalent values of proceeds in 1989 where proceeds from years before 1989 grow according to compound interest using a short-term interest rate. The short-term rate we use in these calculations is the U.S. average annual 3 month certificate of deposit rate as listed in the International Monetary Fund's *International Financial Statistics* for October 1994.

A.4. Biotech Equity Index and Returns

We define a biotech equity index and use its returns to measure the overall performance of biotech stocks. We considered two different portfolios constructed to calculate monthly weighted returns from January 1975 to December 1995: (1) Our basic biotech equity index is defined based on a portfolio consisting of all publicly traded biotech-using firms (including incumbent adopters) in our larger project biotech firm data set. (2) We defined an alternative portfolio consisting of all publicly traded firms in the biotech-relevant Standard Industrial Classifications 2830 through 2836, inclusive. We retrieved return data from CRSP and calculated the monthly weighted returns for each portfolio. Based on these monthly returns, we calculated annualized returns for each portfolio. In this paper, we use the annualized weighted returns of the first portfolio since the second portfolio included numerous non-biotech using firms.

A.5. Venture Capital Investment Variables

We use a number of venture capital investment variables based on data available in the *VentureXpert* on-line electronic database (maintained by the Securities Data Company 1998b). The database contains detailed information on the date, stage, and amount for each round of funding for each firm. We drop "Bridge", "Bridge Loan", "Open Market Purchase", "Other Spec Situation", "Secondary Purchase", and "Turnaround" stage funding. We drop observations where there are missing values on date or on the amount of venture capital investment.

Based on these data, we calculated the rounds of venture capital and amount of venture capital received by each firm in each year. The latter value is converted into 1984 US dollars. For both the number of rounds and the amount of venture capital, we calculated the cumulative values up through the current year. Additionally, we constructed a categorical variable (1 yes, 0

no) indicating whether each of the firms has ever been reported as receiving venture capital. This variable equals 1 if the value of the cumulative rounds variable is nonzero.

REFERENCES

Aghion, Philippe, and Jean Tirole, "The Management of Innovation," *The Quarterly Journal of Economics*, November 1994, 109(4): 1185-1209.

Banerjee, Abhijit V., "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, August 1992, 107(3): 797-818 .

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades, *Journal of Political Economy*, October 1992, 100(5): 992-1026.

Bioscan, Volumes 3-11, 1989-1997.

Cetus Corporation, "Biotechnology Company Data Base," [also known as Pre-Bioscan], predecessor source for *Bioscan*, provided in printout form by the former C.E.O. of Cetus Corp., 1988.

Darby, Michael R., Qiao Liu, and Lynne G. Zucker, "Stakes and Stars: The Effect of Intellectual Human Capital on the Level and Variability of High-Tech Firms' Market Values," National Bureau of Economic Research Working Paper No. 7201, June 1999.

Deeds, David L., Dona Decarolis, and Joseph E. Coombs, "The Impact of Firm-Specific Capabilities on the Amount of Capital Raised in an Initial Public Offering, *Journal of Business Venturing*, 1997, 12: 31-46.

Griliches, Zvi, "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, December 1990, 28:1661-1707.

IPO Reporter, various issues, 1976-1992.

Jones, Lyle V., Lindzey, Gardner and Coggeshall, Porter E., eds., *An Assessment of Research-Doctorate Programs in the United States: Biological Sciences*, Washington, DC: National Academy Press, 1982.

Securities Data Company, *Global New Issues*, on-line electronic data base, New York, NY: Securities Data Company, as of December 1998. (1998a)

Securities Data Company, *VentureXpert*, on-line electronic data base, New York, NY: Securities Data Company, as of December 1998. (1998b)

Stephan, Paula E., and Everhart, Steve, "The Changing Rewards to Science: The Case of Biotechnology," *Small Business Economics*, March 1998, 10(2): 141-151.

Torero, Máximo, Michael R. Darby, and Lynne G. Zucker, "The Importance of Intellectual Capital in the Birth of the Semiconductor Industry," UCLA Working Paper, 2001.

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.

U.S. Department of Commerce, Patent and Trademark Office, *Patent Technology Set: Genetic Engineering*, CD-ROM, machine readable data base, Washington, DC: U.S. Department of Commerce, Office of Information Systems, June 29, 1993.

Welch, Ivo, Sequential Sales, Learning, and Cascades," *Journal of Finance*, June 1992, 47(2): 695-732

Zucker, Lynne G., Marilynn 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., Michael R. Darby, and Jeff Armstrong, "Geographically Localized Knowledge: Spillovers or Markets?", *Economic Inquiry*, January 1998, <u>36</u>(1): 65-86.

Zucker, Lynne G., Michael R. Darby, and Jeff Armstrong, "Commercializing Knowledge: University Science, Knowledge Capture, and Firm Performance in Biotechnology," *Management Science*, January 2002, <u>48</u>(1): 138-153.

Zucker, Lynne G., Michael R. Darby, and Marilynn B. Brewer, "Intellectual Capital and the Birth of U.S. Biotechnology Enterprises," National Bureau of Economic Research Working Paper No. 4653, February 1994.

Zucker, Lynne G., Michael R. Darby, and Marilynn B. Brewer, "Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises," *American Economic Review*, March 1998, <u>88</u>(1): 290-306.

Zucker, Lynne G., Michael R. Darby, and Máximo Torero, "Labor Mobility from Academe to Commerce," *Journal of Labor Economics*, July 2002 in press.

FOOTNOTES

[1] The other 6 firms became publicly traded through merger with a public company or spinoff.

[2] Factors that predict stars working with firms are identified in Zucker, Darby, and Torero (2002).

[3] The Small Business Innovation Research (SBIR) sets aside for small-business contracts and grants a portion of all federal research funding programs and a number of the private biotech firms obtained research funding from this source.

[4] The scientists who criticized our simple criterion observed that not all sequences are equally difficult and some earned an article for only one sequence while others wer reported jointly.

[5] The Small Business Innovation Research (SBIR) sets aside for small-business contracts and grants a portion of all federal research funding programs and a number of the private biotech firms obtained research funding from this source.

[6] Top-quality universities are defined as those with a quality ranking of 4.0 or above (on a scale of 1 to 5) for one or more of the biotech-relevant doctoral programs in biochemistry, cellular/molecular biology, and microbiology in the 1981 National Research Council survey (Jones, Lindzey and Coggeshall 1982). Twenty U.S. universities were so defined as top quality. Nearby is defined as in the same BEA area as the firm. Zucker, Darby, and Brewer (1998) found that top-quality universities had a significant positive effect on entry of biotech firms over and above that of the number of local actively publishing star scientists.

[7] This table would look like a conventional survival table if we subtract the percentage public from 100 percent to get the percentage "surviving" as private firms.

[8] The models were unstable when one-year and two-year lagged biotech returns were included simultaneously and those with two-year returns generally had a higher log-likelihood. We did not

include the strategy variable in the proceeds models because it was used to identify firms more likely to be pursuing an IPO strategy but would not affect the proceeds of those firms which actually do go public.

[9.] See Zucker, Brewer, Oliver, and Liebeskind (1993). The rDNA technology devolved from high to routine science in the late 1980s; so 1990 was a good year to end the literature base for identifying scientists possessing intellectual capital.

[10] Reputational ratings were based on responses from approximately 15 percent of the faculty in the fields studied.

Table 1
Descriptive Statistics

Variable	Mean	Std. Dev.	Minimum	Maximum
Use of rDNA Technology by Firm	0.4161	0.4930	0	1
Articles to Date by Stars with Firm	0.0103	0.2272	0	9
Patents Granted with Application to Date	0.0272	0.2487	0	8
Number of SBIR Grants to Date	0.0313	0.2110	0	5
Local Top-quality Universities	1.2280	1.1343	0	3
Biotech Returns Two Years Prior	0.2145	0.3504	-1.6541	1.0331
Biotech Returns One Year Prior	0.3709	0.6243	-1.6541	2.0451
Firm Has Received Venture Capital	0.2191	0.4137	0	1
Rounds of Venture Capital Received	0.6269	1.5515	0	13
IPO Proceeds Raised by Firm	2.1761	6.8804	0	7997
IPOs/Private NBFs in Year Firm Born	0.0640	0.0563	0	0.1667

Number of Observations = 3675

Table 2
Estimates of Weibull Survival Model of Duration from Founding to Initial Public Offering

	(a)	(b)	(c)	(d)	(e)
Constant	3.760***	4.111***	4.093***	3.963***	-4.170***
	(0.235)	(0.315)	(0.349)	(0.280)	(0.376)
Firm's Science Base					
Articles to Date by Stars with Firm	-0.207*	-0.161†	-0.235	-0.195*	-0.236
	(0.086)	(0.098)	(0.187)	(0.093)	(0.206)
Patents Granted with Application to Date	-0.266*	-0.167	-0.268	-0.208†	-0.225
	(0.117)	(0.162)	(0.172)	(0.119)	(0.176)
Use of rDNA Technology by Firm	-0.519***	-0.271†	-0.379*	-0.494**	-0.365*
	(0.153)	(0.145)	(0.169)	(0.153)	(0.170)
Number of SBIR Grants to Date	-	-	-	-0.302†	-0.328
				(0.171)	(0.221)
Local Top-Quality Universities	-	-	-	-0.143*	-0.052
				(0.066)	(0.072)
Firm Strategy Indicator					
IPOs/Private NBFs in Year Firm Born	-4.347***	-4.050***	-4.304**	-4.392***	-4.248**
	(1.192)	(1.210)	(1.365)	(1.204)	(1.376)
Market Indicators					
Biotech Returns One Year Prior	-0.339**	-0.376**	-0.393**	-0.337**	-0.396**
	(0.121)	(0.132)	(0.146)	(0.122)	(0.149)
Biotech Returns Two Years Prior	-0.248	-0.240	-0.325	-0.258	-0.348
	(0.192)	(0.203)	(0.232)	(0.196)	(0.237)
Venture Capital					
Firm Has Rec'd Venture Capital	-	-1.205***	-	-	-
		(0.247)			
Rounds of Venture Capital Received	-	-	-0.209***	-	-0.204***
			(0.056)		(0.056)
Sigma	0.774***	0.779***	0.883***	0.778***	0.886***
	(0.089)	(0.097)	(0.135)	(0.091)	(0.137)
Log-likelihood	-607.6	-566.5	-579.5	-602.9	-577.9
Chi-squared test [d.f.]	39.56 [3]***	15.56 [3]***	24.48 [3]***	49.98 [5]***	27.68 [5]***

Notes: Standard errors in parentheses. N = 3675.

Prob($|t| \geq x$) or Prob($|\chi^2| \geq x$): $^\dagger < 0.1$, $* < 0.05$, $** < 0.01$, $*** < 0.001$

The Chi-squared test is for the null hypothesis that the coefficients of the firm science-base variables all = 0. The degrees of freedom for the statistic are given in square brackets.

Table 3
Percentage of Private Biotech Firms Going Public by Age and Science Base

---

This table presents the predicted number of years from founding for various percentages of a cohort of entering private firms to have gone public based on model (e) in Table 2. Model (e) is simulated for three different cohorts: In the case labeled (a), the firms have the mean values each year of all the firms of a given age in our sample. Case (b) is the same as (a) except that the values of the science variables other than the number of Articles to Date by Stars with Firm ("ties") and the indicator for Use of rDNA Technology ("rtech") are set to 0. Case (c) is the same as (a) except that the values of *all* the science variables are set to 0.

Percentage of Public Firms	Base Case All Variables = Sample Means (a)	Case with Science-Base Variables = 0 Except ties and retch (b)	Case with All Science-Base Variables = 0 (c)
5%	1.43	1.77	2.25
10%	2.49	3.10	3.95
20%	4.47	5.56	7.07
30%	6.44	8.00	10.18
40%	8.51	10.58	13.46
50%	10.79	13.41	17.07
60%	13.41	16.67	21.21
70%	16.58	20.61	25.93
80%	20.78	25.83	32.86
90%	27.45	34.12	43.42

Table 4
Estimates of Tobit Regressions for IPO Proceeds

	(a)	(b)	(c)	(d)	(e)
Constant	-23.026***	-25.017***	-24.909***	-27.610***	-28.044***
	(1.034)	(1.017)	(0.349)	(1.220)	(1.215)
Firm's Science Base					
Articles to Date by	6.806***	5.765***	6.549***	6.001***	5.951***
Stars with Firm	(1.440)	(1.235)	(1.382)	(1.399)	(1.356)
Patents Granted with	10.970***	6.718***	8.548***	9.945***	8.158***
Application to Date	(1.348)	(1.156)	(1.295)	(1.317)	(1.279)
Use of rDNA Tech-	10.570***	6.105***	8.506***	10.232***	8.428***
nology by Firm	(0.915)	(0.817)	(0.893)	(0.895)	(0.880)
Number of SBIR	-	-	-	0.393	-1.036
Grants to Date				(1.838)	(1.853)
Local Top-Quality	-	-	-	3.917***	2.937***
Universities				(0.383)	(0.377)
Market Indicator					
Biotech Returns Two	6.390***	6.649***	7.890***	6.376***	7.682***
Years Prior	(1.232)	(1.109)	(1.213)	(1.208)	(1.196)
Venture Capital					
Firm Has Rec'd	-	20.711***	-	-	-
Venture Capital		(0.932)			
Rounds of Venture	-	-	3.612***	-	3.235***
Capital Received			(0.244)		(0.241)
Sigma	18.996***	16.247***	18.214***	18.410***	17.832***
	(0.566)	(0.468)	(0.537)	(0.545)	(0.524)
Log-likelihood fn.	-4274.0	-3972.1	-4152.0	-4219.3	-4121.2

Standard errors in parentheses. N = 3675.
Prob($|t| \geq x$): $^{\dagger} < 0.1$, $* < 0.05$, $** < 0.01$, $*** < 0.001$

Table 5
Marginal Effects of Explanatory Variables on IPO Proceeds

The table presents the partial derivatives of expected value with respect to the vector of explanatory variables. They are computed at the means of the variables based on the corresponding model estimates reported in Table 4.

	(a)	(b)	(c)	(d)	(e)
Firm's Science Base					
Articles to Date by Stars with Firm	1.2727*** (0.2725)	0.9127*** (0.1991)	1.1379*** (10.2436)	1.0822*** (0.2553)	1.0127*** (0.2340)
Patents Granted with Application to Date	2.0515*** (0.2577)	1.0634*** (0.1874)	1.4853*** (0.2298)	1.7934*** (0.2433)	1.3882*** (0.2224)
Use of rDNA Technology by Firm	1.9766*** (0.1674)	0.9664*** (0.1288)	1.4779*** (0.1530)	1.8452*** (0.1583)	1.4344*** (0.1478)
Number of SBIR Grants to Date	-	-	-	0.0709 (0.3314)	-0.1764 (0.3154)
Local Top-Quality Universities	-	-	-	0.7064*** (0.0681)	0.4998*** (0.0639)
Market Indicator					
Biotech Returns Two Years Prior	1.1950*** (0.2292)	1.0526*** (0.1747)	1.3709*** (0.2091)	1.1498*** (0.2169)	1.3073*** (0.2020)
Venture Capital					
Firm Has Rec'd Venture Capital	-	3.2787*** (0.1590)	-	-	-
Rounds of Venture Capital Received	-	-	0.6275*** (0.0428)	-	0.5506*** (0.0416)

Standard errors in parens. N = 3675.
Prob($|t| \geq x$): [†] < 0.1, * < 0.05, ** < 0.01, *** < 0.001

Table 6
Causality Analysis between Science Variables and Venture Capital Variables
(Poisson Regressions)

	(a)	(b)	(c)	(d)
Dependant Variable	Articles to date by stars first tied to firm[1]	Article to date by stars first tied to firs	Rounds of venture capital	Rounds of venture capital
Constant	-4.501*** (0.173)	-4.791*** (0.202)	-2.190*** (0.055)	-2.197*** (0.055)
Article by stars one year prior	2.345*** (0.219)	2.244*** (0.218)	–	0.636* (0.251)
Rounds of venture capital one year prior	–	0.972*** (0.194)	1.054*** (0.050)	1.054*** (0.050)
Log-likelihood fn.	-214.4	-205.9	-1255.6	-1253.4

Standard errors in parentheses. N = 2989.
Prob($|t| \geq x$): $^{\dagger} < 0.1$, * < 0.05, ** < 0.01, *** < 0.001

Chi-squared tests:

H₀: Venture Capital variable does not cause Star Article
$\chi^2(1) = 17.0$***

H₀: Star Article does not cause Venture Capital
$\chi^2(1) = 4.4$*

Note:
[1] We assume it takes two years for an article to be finished and published (i.e., if an article is published in year t, we assume its authors were first tied to the firm in year t-2).

Figure 1
Percentage of Private Biotech Firms Going Public by Age and Science Base