NBER WORKING PAPER SERIES

LABOR MOBILITY FROM ACADEME TO COMMERCE

Lynne G. Zucker
Michael R. Darby
Maximo Torero

Working Paper 6050

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 1997

This research has been supported by grants from the Alfred P. Sloan Foundation through the National Bureau of Economic Research (NBER) Project on Industrial Technology and Productivity, the National Science Foundation (SES 9012925), and the University of California Systemwide Biotechnology Research and Education Program. This paper is a revised version of papers presented at the annual meetings of the Western Economic Association, San Francisco, CA, June 30, 1996. We are indebted to a remarkably talented team of post-doctoral fellows Jeff Armstrong, Zhong Deng, Julia Liebeskind, and Yusheng Peng and research assistants Yui Suzuki, Akio Tagawa, and Alan Wang, with special assistance in coding biographical data from Mavis Wu and Cheri Barba. Comments from Armstrong substantially improved the paper. This paper is part of NBER's research program in Productivity. Any opinions expressed are those of the authors and not those of their employers, funders, or the National Bureau of Economic Research.

Labor Mobility from Academe to Commerce
Lynne G. Zucker, Michael R. Darby and Maximo Torero
NBER Working Paper No. 6050
May 1997
JEL Nos. J60, J44, O31
Productivity

ABSTRACT

Following a breakthrough discovery, scientific knowledge with natural excludability may be best transferred to industry by the labor mobility of top scientists from universities and research institutes to firms. We model labor mobility as a function of scientist's quality (as measured by scientific citations) and his or her reservation wage. The reservation wage is determined by labor quality and the cost of moving, and also depends on the trial frequency (number of potential firm employers), potential interfering offers from universities, and experienced increase in productivity of top scientists already in firms (reducing reservation values). Applying our model to bioscience and related biotechnology industries, we find broad support in a group duration analysis. The time a star scientist remains in a university before moving to a firm is significantly: decreased as the quality of the bioscientist and as his or her focus on human genetics increases; decreased as the expected frequency of offers increases with increases in local firms commercializing the technology and the percentage of ties to scientists outside the bioscientist's organization; decreased by experienced increase in productivity by other nearby star scientists who have already moved to firms. Only the number of top quality universities in the local area, via interfering university moves, increases the time a star scientist remains in a university before moving to a firm. We find some evidence of heterogeneity when we decompose the sample of bioscientists by their destination status, finding only quality remains significant across both *affiliated* scientists (full-time employment in a firm) and *linked* scientists (part-time employment), with all variables that are significant in the duration model also entering for linked scientists.

Lynne G. Zucker
Department of Sociology
University of California
Los Angeles, CA 90095-1551
and NBER
zucker@soc.sscnet.ucla.edu

Michael R. Darby
Anderson Graduate School of Management
University of California
Los Angeles, CA 90095-1481
and NBER
michael.r.darby@anderson.ucla.edu

Maximo Torero
Department of Economics
University of California
Box 951477
Los Angeles, CA 90095-1477

Labor Mobility from Academe to Commerce

"Technology transfer is the movement of ideas in people."

- Donald Kennedy, Stanford University, March 18, 1994

Following a breakthrough discovery, scientific knowledge with natural excludability may be best transferred to industry by the labor mobility of top scientists from universities and research institutes to firms. We model labor mobility as a function of scientist's quality (as measured by scientific citations) and his or her reservation wage. Labor mobility generally is based on visible or easily obtainable signals of underlying labor quality, such as education (A. Michael Spence 1973; 1974); labor mobility of top scientists is no different, but the signals typically contain more differentiated information concerning current levels of output and more evaluative information on the quality of that output. Returns to detailed monitoring of the quantity and quality of scientists' performance are sufficiently high to employers to offset the costs involved.

In order to gain access to the knowledge of discovering scientists, firms in related areas of technology employ them. In biotechnology, the discovering scientists were initially employed by universities and research institutes; we are concerned with explaining the mobility processes involving in moving at least part of their labor effort to specific firms. Some of these firms are incumbent firms which adopt the new technology (see Zucker and Darby 1997, 1996a), but most of the firms are newly created around these "star" scientists, who often become residual owners as well as employees (Zucker, Darby, and Brewer 1997).

We investigate two somewhat different sources of labor mobility: the "classic" labor

mobility of changing employer from a university or research institute to a firm ("affiliated scientists"), and the empirically more common labor mobility we observe when academic or research institute scientists collaborate on joint research projects and/or patenting with a firm ("linked scientists"). Both kinds of mobility generally involve working at the bench science level with firm scientists. Some of the linked bioscientists retain their full university positions, but others have opted for adjunct or other titles that involve less active day-to-day participation, while still retaining their academic positions and identifying their affiliation as the university on their publications.

Labor Mobility As Technology Transfer

Labor quality is variable among scientists working on recombinant DNA; some scientists are very productive while others are not. In breakthrough discoveries scientific productivity becomes relevant to commercialization, and hence the labor of the most productive scientists is the main resource around which firms are built or transformed. We can think of these scientists as "seed crystals" forming the center of a set of resources that they actively seek to augment their own research productivity, and when that research is highly commercializable, similarly augment productivity of a firm.

"Star" scientists are important in the process of technology transfer because of the value of their knowledge to success of firms. In related research, we have found that for an average firm involved in biotechnology, two genetic sequence articles co-authored by an academic star scientist and a firm's scientists result in about 1 more product in development, 1 more product on the market, and 344 more employees; for five articles these numbers are 5, 3.5, and 860

respectively (Zucker, Darby, and Armstrong 1997a; Zucker and Darby 1996b).[1]

Identifying Star Scientists

The breakthrough discovery by Stanley Cohen and Herbert Boyer of the basic technique for recombinant DNA is the foundation both of a burst of related scientific innovation in the biosciences and of commercial biotechnology (reported in Cohen, Chang, Boyer, and Helling 1973). While other discoveries and techniques have become important in biotechnology, the core technology is the application of genetic engineering based upon the Cohen-Boyer breakthrough of taking a gene from one organism and implanting it in another.[2]

A very important measure of research success is the discovery of nucleotide sequences that determine the characteristics of proteins and other molecules; these sequences and the articles that report them are cataloged in an international scientific data base, GenBank (1990; 1994). In this paper, we analyze gene sequence articles up to 1990 by at least one U.S. scientist (with one exception that includes international articles through 1992). GenBank assigns to each article one or more "primary accession numbers" to identify each genetic sequence.

Based on these accession numbers, we identified a set of 327 star scientists, 305 with more than 40 genetic sequences through April 1990 and 22 with 20 or more articles (with at least 20 primary accession numbers) to include difficult discoveries that may report fewer sequences per article on average.[3] Of these stars, 207 published with an affiliation to an institution in the United States at least once. Affiliation, including both institution and address (country), are not included in GenBank but were hand coded from 4,061 articles authored by stars.[4] When we examine the full GenBank in the next section, then, we do not have information on location/affiliation of most authors.

Barriers to Information Flow: Stars' Knowledge Advantage

Labor mobility of discovering scientists becomes important in technology transfer when a new discovery has both high commercial value and a combination of scarcity and tacitness that defines *natural excludability*, the degree to which there is a barrier to the flow of the valuable knowledge from the discoverers to other scientists. Those with the most information about breakthrough discoveries are the scientists actually making them, so there is initial scarcity. To the extent that the knowledge is both scarce and tacit, it constitutes intellectual human capital retained by the discovering scientists (Zucker, Darby, and Brewer 1997 and 1994, Appendix).

Scarcity of the new knowledge is reflected in classic diffusion, beginning with just a handful of discoverers and growing at a pace that reflects both the value of the knowledge, where high value discoveries will diffuse more widely and rapidly than those with low value, and its tacitness.[5] When the value is high, as in biotechnology, other scientists are motivated to learn the new knowledge; however when tacitness is high, these other scientists are limited in their ability to learn it depending on the relative scarcity of those who already know it since scientists desiring to enter the new area of research may need to have hands-on experience at the bench before they are able to do so.[6] Coauthoring, which implies bench level collaboration, provides our measure of tacitness: Degree of tacitness is high when most new authors are publishing with at least one old author defined as those who have published before in GenBank, and low if most new entrants to GenBank can do the research either by him/herself or with all new authors.

Figure 1 illustrates the initial scarcity of the new knowledge, and the overall drop in scarcity as new scientists increasingly publish in GenBank, enlarging the pool over time of scientists who continue to publish on genetic sequence research. As also shown in Figure 1, our

tacitness measure declines more slowly than scarcity. In fact, new scientists continue to enter throughout the 1969 through 1992 period predominantly by publishing with old, experienced scientists who have previously published in GenBank, and thus demonstrably know the relevant techniques, with this mode accounting for 81 percent of entry from 1969 through 1992.[7] Excluding sole-authored articles, which may be dissertations for new authors and review articles by established authors, new authors write exclusively with other new authors 36 percent less frequently than old authors write exclusively with other old authors.[8] The overall significance of these differences was confirmed in a loglinear analysis (chi-square $= 1265.45$; G-squared $= 1202.83$; $p > .0001$ for both values).

Star Productivity and Labor Mobility to Firms

We have already reviewed evidence identifying the productivity effects of stars coauthoring with firm scientists--products in development, on the market, and employment growth--on firm success, explaining why firms are motivated to employ stars. But why are stars motivated to work with firms? Given the enormous commercial value of their initially scarce knowledge, financial incentives are obvious, but we believe that there are also important scientific productivity effects, and again focus specifically on the U.S.

Briefly reviewing the kinds of employment relations with firms, stars may be **affiliated**, that is working for firms (measured as listing the firm as affiliation on the article), or stars may be **linked**, that is working with firms while maintaining their primary affiliation with a university or research institute (measured as coathoring with firm scientists while simultaneously listing the university or research institute as their primary affiliation or assigning a patent on issuance to a firm rather than to their university or research institute) While affiliated by definition work in

the same region as the firm, linked scientists may coauthor either with firms in their region (**local link**) or with firms outside their region (**external link**, or link to different region). We define region here as one of 183 functional economic areas in the U.S. as defined by the Bureau of Economic Analysis (U.S. Department of Commerce 1992).

To measure scientific productivity, we here examine both the number of articles published in GenBank and the citations to those articles in 1987 and 1992. We find generally that scientists who became affiliated or linked to firms increase their productivity during the time they are working with firm scientists, especially in citations/quality, and that for the most part this higher productivity is not due to a selection effect, but rather to experience during the time the scientist was working with the firm, generally confirming our earlier results (Zucker and Darby 1996b).

In Table 1 we present the mean number of GenBank articles published pre/during/post working with the firm for each U.S. star scientist who is ever linked or affiliated with a firm, and compare productivity across these periods, and also with stars who are never linked or affiliated. Productivity effects are estimated for affiliated, local linked, and external stars. Underlying each of these comparisons is a series of regression analyses that because of their complexity will be summarized here. Both the total output of articles and the rate of articles per year show generally similar effects: all tied (affiliated or linked) scientists have significantly higher publications pre, during, and after the firm, but only for locally linked stars is there clear evidence that the rate of publication increases significantly during the firm, compared to before and after.

Turning now to the quality indicator, citations, affiliated stars show a significantly higher rate of citations during the firm than before or after; local linked show the same tendency, but the increase is not significant (though it reaches the $p = .08$ level). For external links, the

increase during linkage with the firm does reach significance, but there is also evidence of a similarly high rate of citation to their pre-firm articles, suggesting a selection effect. The increased quality of the articles is probably due to increased resources and higher standards of reliability at firms relative to universities, based on interviews with firm-tied stars.

There are many specific gains to trade between star scientists and firms that make it attractive to both to move at least part of the labor effort of stars to firms that are working to commercialize their discoveries. Figure 2 uses a dark-lined triangle to indicate the stars that move at least some of their labor effort to a firm, with a lighter-edged triangle indicating the total number of stars. Firms are indicated by a burred star. It provides a graphic illustration of the amount of labor mobility of stars, and shows that the locations of mobile stars and firms are both concentrated and highly correlated geographically. In the next section, we develop a model of this mobility process, and then use it to motivate estimation of mobility in the following sections.

The Model:

In the job mobility literature, experience and employer-specific human capital are the key variables explaining movement. Specific human capital is not significant for the current case of star bioscientists and will not be discussed further (but it is detailed under individual characteristics in Table 2). In analyzing job mobility between employers, Topel (1986) and Topel and Ward (1992) assume that wage offers from potential employers are generated by a known offer distribution which reflects the variation in expected values of marginal product across employers. The location of this distribution should vary across individuals according to their characteristics which indicate differences in productivity to potential employers. Topel and Ward

(1992) abstract from individual differences and assume that the location of the external wage offer distribution depends on an individual's cumulative labor market experience X:

$$(1) \qquad prob(w_0 < z;X) = G(z;X), \qquad G_X(z;X) \leq 0$$

Topel and Ward note that experience increases wage offers if the last inequality in (1) is strict, but observed wages will increase with experience due to search even if expected productivity is independent of experience [$G_X(z;X) = 0$].

For the star scientists we are considering, it is possible for both potential employers and econometricians to readily measure a vector Q of indicators of expected value of marginal product. Elements of this vector would include whether the scientist is employed by a top-quality university, is tenured there, the quantity and quality of articles published previously (quality is observed directly by the firm but proxied for us by the number of citations per article), and whether the scientist's work has concentrated on human genetic sequences (refer again to Table 2, quality characteristics of star and commercial potential). Increases in each of those variables would increase the expected value of a star scientist to any given firm.

$$(2) \qquad prob(w_0 < z;Q) = G(z;Q), \qquad G_Q(z;Q) \leq 0$$

where $G_Q(z;Q)$ is the vector consisting of partial derivatives for the continuous variables and partial differences for the categorical variables (i.e., top-quality university, tenured). Hence the probability of receiving an offer from some firm which exceeds any given value generally

increases with the characteristics in Q. In principle, experience X might be an element of Q, but we see below that in the presence of more direct productivity measures X is insignificant as a predictor of mobility.

Star bioscientists moving from a university (in whole or part) are typically employed by four distinct types of potential alternate (or joint) employers: local firms, external (out-of-local-region) firms, local universities, and external universities. Movement to another university or firm or even part-time collaboration (linkage) to a firm generally involves a major time investment for the scientist and occurs infrequently; so, for practical purposes, we can assume that only one such move is possible in any given period. We model the scientist as acquiring a new employer if an offer exceeding the type-specific (see below) reservation value is made by any of each of the four types of employers and distinguish between full- and part-time work with firms.

We assume that because of egalitarian pressures within the university as well as the potentially greater returns to commercial applications of the star's intellectual human capital that higher values of any of the elements of Q shifts the location of the G function by more than the reservation value R so that the probability of an acceptable offer from any of the four types of potential employers (indexed by i) increases in Q also:

(3) $prob(w_o < R;Q,i) = G(z;Q,i), \qquad G_Q(z;Q,i) \leq 0$

where i *= 1 for full-time local firm job, 2 for full-time external firm job, 3 for part-time local firm link, 4 for part-time external firm link, 5 for other local university jobs, and 6 for external university jobs*

This assumption is more obvious for movements to firms (i = 1-4) than for universities (i = 5,6), but academics frequently note that much greater weight is placed on externally visible research productivity in hiring from the outside than in promoting from within. In any case, our principal concern in this paper is explaining embodied technology transfer from universities to firms, so movements to other universities enter only as potential temporary interference with that process.

Specifically, we want to explain the probability per unit of time that a scientist will become involved in commercial applications of biotechnology full or part-time with either a local or external firm. The overall hazard function therefore can be written as the sum of the firm-type-specific hazard functions (Kalbfleisch and Prentice 1980, p. 167):

$$(4) \qquad \lambda(t;Q,H) = \sum_{i=1}^{4} \lambda_i(t;Q,H)$$

where Q as before is our vector of externally observed measures of intellectual human capital and H represents other factors affecting the hazard rate. The additive form of the hazard function implies that we can group relevant subsets for empirical purposes such as full-time versus part-time or local versus external employers.

Since equation (3) describes the conditions under which a single trial will result in an offer greater than the reservation value for that type of firm, prominent candidates for variables which might belong in H are those which increase the rate at which individual employer-scientist matches are considered per unit of time. We again refer back to Table 2. Other things equal, we expect that the cost of moving residence and family (and those of research lab teams)--or the

cost of travel for part-time work--gives a lower reservation value and hence higher probability per trial for local employers. However, the number of local trials is limited by the extent of the market; so we include the number of new biotechnology enterprises in the same region as the scientist's university and expect that variable to increase the probability of initiating (local) commercial ties. Similarly, a higher number of top-quality universities in the same region should reduce the probability of initiating commercial ties by increasing the probability of interfering inter-university movements.[9] External employers are numerous relative to the feasible number of trials for a scientist over any short number of years, but to the extent that the scientist has a higher fraction of his or her coauthors at organizations elsewhere we anticipate that the frequency with which alternative employment opportunities can be explored is increased per unit time. This variable appears in the variable list in Table 2 under size of social networks. Changing employers among universities or research institutions may play a similar role in increasing the probability of receiving information about alternative employment opportunities.

One major factor which may reduce the reservation wage for firm employers is the experience that star scientists saw their productivity maintained or increased in quantity of publications when they became employed by or collaborated with firms and dramatically increased in quality in terms of citations per article while thus tied to firms, especially for affiliated stars, as we discussed above in connection with Table 1 (see also Zucker and Darby 1996b).[10] We assume that this symbiotic effect on personal productivity and hence scientific prestige and expected future earnings was not expected by scientists until it was observed; so we include in H two measures of experienced increase in productivity by other star scientists who have previously moved to firms: The average change in number of citations by stars during the

time they are affiliated with a firm, using the measure in the prior period to predict mobility in the current period, and separating the experience of stars in the same region from that of stars in other regions since the former may be better known. Again, these variables can be found in Table 2 under regional variables.

It should be noted that the star scientists frequently play a key role in the founding of the firms with which they become affiliated or linked (Zucker, Darby, and Brewer 1997 and 1994 Appendix; Zucker, Darby, and Armstrong 1997a). That is, what appears to be employment may in fact be entrepreneurship. We expect that characteristics which predict a high marginal product to potential employers will also be attractive to potential investors, so the analysis is not greatly affected whether the scientist is searching for an employer or venture capital. Indeed, prospecti for initial public offerings of new biotech firms frequently list precisely these sorts of qualifications in Q for key associated scientists. Since there is a significantly positive agglomeration effect reported by Zucker, Darby, and Brewer (1994), a star should find it easier to start a new firm where there are more firms already, so the sign of total new biotechnology enterprises in the region should be positive here too.

In the next section, we use group duration analysis of proportional hazard models to test the hypothesis that our measures Q of scientific quality and H of factors affecting trial frequency, reservation values, and interfering university offers have the predicted effects on the probability that a star will become employed by or a collaborator of a firm and that these effects will dominate traditional measures such as experience. We find that early in a star's university career, there is a very low probability of affiliating or linking to a firm, although this is higher for stars located in regions and times with more numerous new biotechnology enterprises. The hazard rate

increases over a star's early publishing career in GenBank (measured as years since first publication appeared in GenBank), with an estimated peak around the tenth year. Only one star started publishing with a firm, that is, became affiliated with a firm as his first job.

In the next section we first introduce the variables not yet discussed and then provide a brief explication of the methodology, with details in the Appendix. Initially a grouped duration model is applied to determine what could be an appropriate assumption for the base hazard function and finally a correction for heterogeneity is carried out.

Methods

Data Sources

We have now reviewed most of the variables listed in Table 2. Most of our variables are drawn from GenBank and the articles published by the star scientists that are catalogued in GenBank (for affiliation, location, and nature of coauthor relations from collaborating with firm scientists to size of social network), and the genetic sequences that are given primary accession numbers in GenBank.

There are five major sources of additional data: (1) Institute for Scientific Information's *Science Citation Index* (1987, 1992) on the total number of citations to each of our 4,061 published articles for each of the indicated two years;[11] (2) Listing and location of all U.S. universities provided by Higher Education General Information Survey (HEGIS), Institutional Characteristics, 1983-84 (U.S. Department of Education, National Center for Education Statistics, 1985) and National Research Council's rating survey of research university departments (Jones, Lindzey, and Coggeshall 1982), where we use the presence of at least one "biotech-relevant"

department (biochemistry, cellular/molecular biology, and microbiology) with scholarly-quality reputational ratings of 4.0 or higher in the 1982 to define top quality universities of which there are 20 in the United States;[12] (3) *Bioscan* (1988-1994) and Cetus Corp. (1986), coupled with North Carolina Biotechnology Center's *U.S. Companies Database* (1992), on the total number of new biotechnology enterprises by location in the United States; (4) Detailed bibliographic information from five major sources listed in Data Sources following the references (*American Men and Women of Science*, *Biotechnology Research Directory 4000 Faculty Profiles*, *Who's Who of Nobel Prize Winners, 1901-1900*, National Academy of Sciences *Organization and Members 1993*, and the *1990 Directory* of the American Association for the Advancement of Science);[13] and (5) Annual salary data for associate and full professors from most U.S. research universities from *The American Association of University Professors Bulletin* and *Academe* from 1970 through 1989, and for the handful of institute affiliations from telephone interviews with the respective institutes.

Group Duration Models

The expected amount of time the scientist stays in universities without moving to a firm differs because each scientist has a different vector **Q** of indicators of expected value of marginal product (whether employed by a top quality university, is tenured there, the quality and quantity of articles published previously, and number of human genetic sequences) and faces different local economic areas that alter each scientist's reservation wage (number of new firms, number of top quality universities). All of these sources of differences are represented by a regressor vector **x**, for each scientist. This regressor vector may have elements that only change across individuals and stay constant during the time of the duration (most of our variables are time

invariant covariates) or might change through the time of the duration (time varying covariates are total citations, total publications, and number of new biotechnology firms).

Time invariant covariates fall into two classes: variables that are defined **prior** to the move, including the number of human sequences and the percent collaborators from outside his or her organization, and variables that describe the university or research institute that are updated each time the scientist changes university or institute (top quality university, location in specific key university clusters, Stanford/UC San Francisco or MIT/Harvard, and location at the National Cancer Institute).

We follow in the footsteps of initial uses of duration models as reported in Lancaster (1979), Nickell (1979) and Petersen (1986a, 1986b). These duration models were selected based on concern with unobserved heterogeneity, while statisticians have instead emphasized the elegance of the use of semiparametric models which do not require parametric specification of the baseline hazard, most often preferring the proportional hazard-partial likelihood specification (PHM, see Cox 1972, 1975, Cox and Oakes 1984, Kalbfleisch and Prentice 1980) . The semiparametric models have not received much use in economics, despite their elegance, for three reasons identified by Han and Hausmann (1990): (1) It is a continuous time specification while most duration data in econometrics are discrete, e.g. determining duration based on publication date of the first article using a firm address for location or writing with firm coauthors; (2) Ad hoc procedures used to treat tied failure times within the partial likelihood framework are cumbersome in the presence of many ties (many simultaneous failures), as found in our data where the majority of scientists exit the universities in their fourth year; (3) Unobservable heterogeneity cannot be included without the presence of multiple integrals of the same order as

the number of individuals in the risk set, which makes estimation difficult, if not impossible.

We elect to use the grouped data version of the proportional hazard model to develop computationally feasible estimators of the relative risk function and the corresponding survivor function in the presence of many tied failure times. Specifically, we apply the technique of group duration analysis developed and used by Prentice and Gloeckler (1978) and Ryu (1994), given that our observations are the articles published by year and so time is measured (grouped) at intervals, available discretely at the level of the year. The spell **T**, number of years, is the difference in years between either the date each star scientist entered a university, as recorded in one of the biographical directories (see Data Sources), or the first date of publishing in GanBank, and the first article in GenBank that shows him/her affiliated or linked to a firm through coauthorship. See the Appendix for a derivation of group duration analysis that we use in our main analyses.

Given then that group duration information is a sequence of binary information we can apply a logistic function which is inherently easier to compute, selecting from ordered probit and ordered logit models, as suggested by Han and Hausman (1990). In some exploratory research (Han and Hausman 1990), the estimates of the ordered logit and ordered probit models are very similar except in the extreme left tail. Given these small differences, we selected the ordered logit model because of the simplicity of its calculation.

Multinomial Logit

We selected multinomial logit to explore the effects of selecting different relationships to a firm. Each star scientist is assumed to have preferences defined over a set of alternatives: Affiliated or linked to a firm. Since this technique is more commonly used, we do not go into

further detail on it.

Empirical Results for the Group Duration Model

The results reported in Table 3 are generally supportive of the suppositions contained in our mobility model. Standard individual characteristic variables generally fail to reach significance, though they are generally in the expected direction. The one exception is the "first year that the star publishes in GenBank" which is experience of a very special sort; the negative sign indicates that later the year of entry, the less probable that the star becomes affiliated or linked to a firm.

Of the quality variables only the number of citations enters significantly. The larger the number of citations, the more likely the star will be to move out of the university. The insignificant coefficient on the quantity of articles suggests that firms don't distinguish between a scientist with a few highly cited articles and another with many lesser cited articles as long as total citations are the same. (In an analysis not reported here, we find that the number of articles enters significantly to increase the probability of moving to a firm if number of citations is removed from the equation, but the overall fit declines.) Receipt of tenure or the Nobel Prize appears to raise the reservation wage as much as the offer distribution with no net effect on mobility.

Our indicator that the scientist's work has more immediate commercial potential, the number of human genetic sequences, enters significantly robustly across the different specifications, increasing the "death" rate or rate of labor mobility from the university or research institute to the firm. In contrast, none of the characteristics of the university or research institute

currently employing the star are ever significant (the index of wages is considered below).

The count of new biotechnology enterprises and top quality universities in the region are both significant, but as expected they act in opposite directions on the probability of moving to a firm. As the number of firms grows larger, so does the probability of a star scientist becoming tied to a firm; as the number of top quality universities grows larger, the probability of a star becoming tied to a firm declines.

The proportion of a star's coauthors that are from different institutions increase the probability of moving to a firm, as we would expect based on increasing information about potential opportunities. Another sort of information, about the quality of the experience other stars' have had working with firms, entered significantly: the larger the increase in citations to local stars who became involved with firms, the more likely is a star to become involved. However, the citation experience of stars outside the region has no significant effect.

Overall, Table 3 provides strong support for our conjectures. While many of the variables are not significant, key variables measuring quality and commercial potential of the intellectual human capital significantly increase the probability of moving to a firm, as do various measures of increasing information about opportunities (social network) and about scientific productivity gains to working with firms.

We now consider briefly the issue of wages earned in the university, under the hypothesis that higher university wages would increase the reservation wage and hence the time it takes scientists to move to a firm. Unfortunately, we were not able to obtain the actual salary paid to each star scientist while he/she was in the university, so we constructed a proxy index of wages by dividing the specific wage in the university or institute employing the star scientists over the

average of the wages for all the universities and institutes in the relevant year This index of wages never entered significantly, but was in the right direction and approaching significance when entered with only the first year star publishes in GenBank. Table 4 presents these results.

Table 5 reports the estimation of a multinomial logit model which examines the choice of becoming affiliated with or linked to a firm. The parameter estimates indicate that very similar processes are involved in the decision to become wholly or partially involved with a firm, although fewer of the coefficients are significant for affiliated stars, apparently because of the relatively fewer observations for affiliated stars. For affiliated stars, the quality of the star scientist is the most important variable affecting the probability of a move to a firm, though the number of articles is negative indicating a premium for earning total citations in fewer more highly cited articles. New biotech enterprises is significant in the expected direction, as is the first year that the star publishes in GenBank. Linked stars show a very similar pattern of significant variables to the overall results reported in Table 3 and discussed above, except that the average change in citations of other local stars lose their significance.

Summary and Implications

We have shown across a series of analyses that star scientists with high quality intellectual human capital--here measured in terms of number of citations to genetic-sequence-discovery articles--that is relevant to firms commercializing biotechnology (i.e.,amplified by discovery of human genetic sequences) leads to moving at least some of their labor from universities to firms earlier in the process (after a shorter duration in the university). We have also demonstrated strong effects of the opportunities available in the stars' own region: stars have a higher

probability of moving to a firm when there are more biotech enterprises in their region, and a lower probability of moving to a firm when there are more top quality universities in their region, a competing influence. The size of stars' networks outside of the university also increased the likelihood of their leaving the university after a shorter duration. Stars also seem to be paying attention to changes in productivity of other stars in their region who have previously moved to firms: when these other stars' citations increase on average, the probability of moving to a firm after a shorter duration increases. Our relatively weak measure of wages did not have a significant impact, but it is not clear whether measurement was the problem or the astronomically higher wages (especially if full or partial ownership of the firm is included) on the other side of the equation. The multinomial logit results for the choice of becoming affiliated or linked to a firm show a generally similar pattern of results as the pooled analyses, with linked scientists close to matching but with the smaller affiliated stars having fewer significant explanatory variables.

Overall, the empirical analysis provided strong support for the model we developed. We hypothesized that the very valuable intellectual human capital would serve as the basis for mobility, not the much less precise measures of experience and firm-specific experience that are typically used in these models. When it is worth investing in costly information, both the individuals and the organizations involved will invest in collecting and using it (Zucker and Darby 1996a). The value of the information is a key determinant. We examine value in two principal ways. In this paper, we operationalize an important new measure of the degree of tacit knowledge, resting on a coathorship measure we developed to examine labor mobility of star scientists to firms:[14] Even as scarcity of the knowledge may be declining, tacitness may not be--or at least not as fast. Throughout the period in which we are examining star scientist mobility,

most new authors entered GenBank by publishing with at least one old author (81 percent of the entry from 1969 through 1992). While there are competing explanations for this finding, none are as parsimonious as the high and only gradually declining tacitness of the knowledge, which provides natural excludability or a natural barrier to the entry of new scientists and hence returns to those who hold the tacit knowledge.

We also measure value in a related paper on the effects of stars on the success of new biotechnology enterprises, and find that university star scientists who actually work with firm scientists have a strong positive effect on products in development, products on the market, and employment growth. Due to both of these sources of value, the labor of star scientists in the U.S. has strongly moved to firms and has done so in very concentrated, localized areas, as illustrated in Figure 2 above.

Finally, we wish to conclude with the observation that scientists and the universities, research institutes, and high technology firms that they work in are recurrently faced with knowledge discontinuities that require some kind of technology transfer mechanism. There are thus incentives for them to construct structures--or to be "born" with structures--that lower the costs of new knowledge acquisition: Both affiliation and link to firms fit well within the structure of a "normal" academic career. For scientists, moving part of their labor effort outside the university is common, and is concentrated in the high quality end of the faculty distribution, certainly not "marginal."[15] Many universities do not place any restrictions on professor's outside employment, while universities with rules typically allow 40 percent of faculty time to be spent on outside consulting. One study of academics found that 20 to 25 percent of faculty income was earned outside the university (George G. Stigler 1950, pp. 42, 60). High technology firms

routinely employ the very top scientists across a wide variety of positions, from heads of scientific teams to members of scientific advisory boards, some full time and some traditionally part-time. Even in countries with substantial barriers to collaboration across university boundaries, firms and entrepreneurial academic scientists find "work-arounds" such as bringing firm scientists into the university labs along with a "stipend" from the firm to cover laboratory materials, as is routine in the national universities in Japan (Darby and Zucker 1996).

We have uncovered an important and neglected set of processes that allow retention of knowledge by its discoverer and incorporation of that knowledge--at least for some period of time--into the intellectual human capital of the discoverer. When this knowledge is valuable, there will be high demand for those who retain it and structures that allow technology transfer between the discoverers and those who wish to use it in science or commerce will develop, even around significant institutional barriers. We have examined the employment relation of star scientists through affiliation and linkage to firms as one structural mechanism that facilitates technology transfer from universities and research institutes to firms.

APPENDIX

We use the grouped data version of the proportional hazard model in an attempt to develop computationally feasible estimators of the relative risk function and the corresponding survivor function in the presence of many tied failure times (Prentice and Gloeckler 1978 and Ryu 1994).

First divide the interval between the beginning of the measurement period, **T**=0, to the time of the measurement, **T**=ti, into j exhaustive non overlapping intervals, $a_0 < a_1 < \ldots < a_{j-1} < a_j$ and the covariates will be assumed to stay constant within each of the j intervals and may change from one interval to the next. Given that the observations are the articles published by year and so T is available discretely and only up to years , the best technique to consider is group duration.

The main idea is that there is an observation scheme grouped into intervals (years in our case):

(1) $A_i = [\, a_{i-1}, a_i \,), \; i = 1, \ldots\ldots r$
$with \; a_o = 0 \,, \; a_r = \infty$

and the failure times in Ai are recoded as **ti**.

Now for each interval we observe (X_i, T_i) where X_i refers to the characteristics of the individuals and T_i refers to the duration. Let α_i = the probability of surviving the ith interval given that an individual has survived up to the (i-1)th interval (conditional probability). Therefore:

(2) $$\alpha_j = e^{-\int_{aj-1}^{aj} h(t)\,dt}$$

(3) $$So: \quad \alpha_j = \Pr(T > a_j \mid T > a_{j-1}) = \frac{\Pr(T > a_j)}{\Pr(T > a_{j-1})} = \frac{S(a_j)}{S(a_{j-1})}$$

(4) $$but: \; S(a_j) = e^{-\int_{0}^{aj} h(u)\,du}$$

(5) $$Therefore: \; \alpha_j = e^{-\int_{aj-1}^{aj} h(u)\,du}$$

Then the probability of observing a failure at time ti on an individual with regression vector x is:

$$(6) \qquad [1 - \alpha_i^{\exp(x,\beta)}] \prod_{j=1}^{j-1} \alpha_j^{\exp(x,\beta)}$$

where the probability of surviving to the beginning of Ai is:

$$(7) \qquad P(ti, x) = \prod_{j=1}^{i-1} \alpha_j^{\exp(x,\beta)}$$

Given then that group duration information is a sequence of binary information we can apply a logistic function given that it is easier for computation:

$$(8) \quad \alpha_j = e^{-e^{x\beta + \gamma j}} \qquad where: \gamma_j = \log \int_{Aj} h_o(t)\, dt$$

This model is easy to estimate using either an ordered probit or ordered logit approach. In some exploratory research by Han and Hausman (1990), the estimates of the ordered logit and ordered probit models are very similar except in the extreme left tail. Accordingly, we select the ordered logit model because of the simplicity of its calculation.

REFERENCES

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.

Cox, D. R., "Regression Models and Life Tables," *Journal of the Royal Statistical Society*, Series B, 1972, 34: 187-220.

Cox, D. R., "Partial Likelihood," *Biometrika*, August 1975, 62(2): 269-276.

Cox, D. and Oakes, R, *Analysis of Survival Data*, London, U.K.: Chapman and Hall, London, 1984.

Darby, Michael R., and Lynne G. Zucker, "Star Scientists, Institutions, and the Entry of Japanese Biotechnology Enterprises," National Bureau of Economic Research Working Paper No. 5795, October 1996.

Han, Aaron, and Jerry A. Hausman, "Flexible Parametric Estimation of Duration and Competing Risk Models," *Journal of Applied Econometrics*, January-March 1990, 5(1): 1-28.

Kalbfleisch, John D., and Ross L. Prentice, *The Statistical Analysis of Failure Time Data*, New York, NY: John Wiley and Sons, 1980.

Lancaster, Tony, "Econometric Methods for the Duration of Unemployment," *Econometrica*, July 1979, 47(4): 939-56.

Liebeskind, Julia Porter, Amalya Lumerman Oliver, Lynne G. Zucker, and Marilynn B. Brewer, "Social Networks, Learning, and Flexibility: Sourcing Scientific Knowledge in New Biotechnology Firms," *Organization Science*, July/August 1996, 7(4): 428-443.

Nickell, Stephen, J, "Estimating the Probability of Leaving Unemployment," *Econometrica*, September 1979, 47(5): 1249-66.

Petersen, Trond, "Fitting Parametric Survival Models with Time-Dependent Covariates," *Journal of the Royal Statistical Society*, Series C, 1986, 35(3): 281-288. (1986a)

Petersen, Trond, "Estimating Fully Parametric Hazard Rate Models with Time-Dependent Covariates," *Sociological Methods and Research*, 1986, 14: 219-246. (1986b)

Prentice, R. L., and L. A. Gloeckler, "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data," *Biometrics*, March 1978, 34: 57-67.

Ryu, Keunkwan, "Group Duration Analysis of the Proportional Hazard Model: Minimum Chi-square Estimation and Specification Tests," *Journal of the American Statistics Association*, December 1994, 89(428): 1386-1397.

Sindelar, Robert D., "Overview/Preview of Current and Future Recombinant DNA-Produced Pharmaceuticals," *Drug Topics*, April 20, 1992, Supplement, pp. 3-16.

Sindelar, Robert D., "The Pharmacy of the Future," *Drug Topics*, May 21, 1993, 137(9): 66-84.

Spence, Michael, "Job Market Signalling." *Quarterly Journal of Economics*, August 1973, 87(3): 355-375.

Spence, A. Michael, *Market Signalling: Informational Transfer in Hiring and Related Screening Processes*, Cambridge, MA: Harvard University Press, 1974.

Stigler, George J., *Employment and Compensation in Education*, National Bureau of Economic Research Ocassional Paper 33, Nee York, NY: National Bureau of Economic Research, 1950.

Topel, Robert H., "Job Mobility, Search, and Earnings Growth: A Reinterpretation of Human Capital Earnings Functions," *Research in Labor Economics*, 1986, 8(Part A): 199-233.

Topel, Robert H., and Michael P. Ward, "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics*, May 1992, 107(2): 439-479.

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., and Michael R. Darby, "Costly Information: Firm Transformation, Exit, or Persistent Failure," *American Behavioral Scientist*, August 1996, 39(8): 959-974. (1996a)

Zucker, Lynne G., and Michael R. Darby, "Star Scientists and Institutional Transformation: Patterns of Invention and Innovation in the Formation of the Biotechnology Industry," *Proceedings of the National Academy of Sciences*, November 12, 1996, 93(23): 12709-12716. (1996b)

Zucker, Lynne G., and Michael R. Darby, "Present at the Revolution: Transformation of Technical Identity for a Large Incumbent Pharmaceutical Firm after the Biotechnological Breakthrough," *Research Policy*, 1997 in press.

Zucker, Lynne G., Michael R. Darby, and Jeff Armstrong, "Intellectual Capital and the Firm: The Technology of Geographically Localized Knowledge Spillovers." National Bureau of Economic Research Working Paper No. 4946, December 1994.

Zucker, Lynne G., Michael R. Darby, and Jeff Armstrong, "Geographically Localized Knowledge: Spillovers or Markets?", *Economic Inquiry*, 1997 in press. (1997a)

Zucker, Lynne G., Michael R. Darby, and Jeff Armstrong, "Sources of Superior Performing Firms: Evidence from the U.S. Biotechnology Industry," working paper, UCLA Institute for Social Science Research, February 1997. (1997b)

Zucker, Lynne G.; Darby, Michael R. and Brewer, Marilynn B. "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*, September 1997, 87(4): in press.

Zucker, Lynne G., Michael R. Darby, Marilynn B. Brewer, and Yusheng Peng, "Collaboration Structure and Information Dilemmas in Biotechnology: Organizational Boundaries as Trust Production," in Roderick M. Kramer and Tom R. Tyler, eds., *Trust in Organizations*, Thousand Oaks, CA: Sage, 1996. [Pp. 90-113]

DATA SOURCES

Academe, Volumes 65-79, 1979-1993.

American Association for the Advancement of Science, *1990 Directory: Consortium of Affiliates for International Programs*, Washington, DC: American Association for the Advancement of Science, 1990.

American Association for the Advancement of Science, *AAAS Handbook 1995/1996: Officers, Organization, Activities*, Washington, DC: American Association for the Advancement of Science, 1995.

American Association of University Professors Bulletin, Volumes 55-64, December 1969-December 1978.

American Men and Women of Science, 1st-19th editions, New York, NY: Jacques Cattell Press/R. R. Bowker Company, 1971-1994.

Bioscan, Volumes 2-8, 1988-1994.

Cetus Corp., "Biotechnology Company Data Base," printout of predecessor source for *Bioscan*, Emeryville, CA: Cetus Corp., 1986.

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

GenBank, Release 81.0, machine readable data base, Bethesda, MD: National Center for Biotechnology Information, February 15, 1994.

Institute for Scientific Information, *Science Citation Index*, ISI Compact Disc Editions, machine readable data base, Philadelphia: Institute for Scientific Information, 1982, 1987, 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.

National Academy of Sciences, National Academy of Engineering, Institute of Medicine, and National Research Council, *Organization and Members 1993*, Washington DC: National Academy Press, 1993.

North Carolina Biotechnology Center, Biotechnology Information Division, *Biotechnology Research Directory: 4000 Faculty Profiles*, Washington, DC: Bureau of National Affairs, Inc., 1991.

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.

Schlessinger, Bernard S., and June H. Schlessinger, eds., *The Who's Who of Nobel Prize Winners, 1901-1990*, 2nd ed., Phoenix, AZ: Oryx Press, 1991.

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 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. For about 15 years after the breakthrough, only a handful of firms have positive profits from using the new technology in part because of the industry context (FDA approval requires 11 to 12 years) and in part because of the very high research and development costs. Thus, it is not possible to use firm financial data or financial market value as measures of productivity.

2. The other basic technology is cell fusion (also termed monoclonal antibodies, MABs, or hybridomas) in which lymphocytes are fused with myeloma cells to create rapidly proliferating antibody-producing cells (see Sindelar 1992 and 1993 for more detail).

3. These 327 stars were only 3/4 of one percent of the authors in GenBank but accounted for 17.3 percent of the published articles, almost 22 times as many articles as the average scientist. The Genbank data set, methods of identifying stars, and productivity of the stars are discussed in more detail in the Data Appendix included in Zucker, Darby, and Brewer (1994) and in Zucker and Darby 1996b).

4. The Science Citation Index lists up to six of the affiliations listed on the paper but only links the corresponding author to a particular affiliation. Thus, only first- and/or corresponding-author affiliations are available in machine-readable sources and bioscience papers frequently list the head of the lab last. As might be expected, our stars, excluding sole authored articles, were last authors on over 69 percent of the articles, where GenBank articles have on average about 4.8 authors per article.

5. Comparing different scientific breakthroughs to determine the initial starting size of the discoverers, the degree to which learning by doing is involved (coauthoring with "old" scientists as the predominant mode of entry), and the relative rates of "diffusion" is an important next step. For example, a much less tacit process appears to operate in the case of high-temperature

superconductors where the know-how was widespread prior to the breakthrough experiment that demonstrated that ceramics incorporating rare earths can work as superconductors at economically interesting temperatures.

6. Exceptions typically include the handful of scientists working in the very narrow specialized area as the discovering scientists. At the extreme, when initial scarcity and tacitness are very high, transmission of the new knowledge will only be to the graduate students and postdocs working in the same lab as the discovering scientists.

7. Reports of publications for 1993 were incomplete in February 1994 so that year has been excluded from the figure and these calculations. In the incomplete reports for 1993, entry with old authors amounted to 83 percent of total entry.

8. Sole-authored articles account for only 6.5 percent of the authorships of new authors and 7.8 percent of the authorships of old authors over this period. Interestingly, new sole authors become more frequent later in the period as the value of the tacit knowledge declined as it became more widespread (see also, Zucker, Darby, Brewer, and Peng 1996).

9. While star scientists occasionally accept an extraordinary offer from universities below top-quality rating, we believe a count of top-quality universities is an adequate measure of the local university market.

10. In addition, local-linked star scientists generally have significantly greater impact on the firm's success than do scientists from other areas (Zucker, Darby, and Armstrong 1997a and b). Thus local firms should have a higher expected offer for part-time linkage than external firms, reinforcing the higher probability that an offer of linkage by a local firm will exceed the reservation value.

11. We also use citations for 1982 in computing experienced change in citations during firm ties, but exclude it from the main analysis because of the small number of articles and stars with significant 1982 citations.

12. The twenty universities 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 Chicago, University of Colorado at Denver, University of Pennsylvania, University of Washington (Seattle), University of Wisconsin-Madison, Yale University.

13. We filled in some missing data for particular stars from *Who's Who of British Scientists, 1980/81*, *Who's Who in Science in Europe*, and *Who's Who in Biotechnology*.

14. We build here on a novel empirical measure we developed in earlier research: "co-publishing," examining all scientists who publish together, to measure who the stars are working with at the bench science level and which organizations are involved in the collaboration (by obtaining the organizational affiliation of all scientists). We have previously used our measure to examine reciprocal productivity effects of star scientists working with scientists in firms (see our discussion of these results in Section III below), effects of organizational boundaries as information envelopes slowing diffusion of scientific knowledge, and size and geography of scientific networks used by firms (Zucker, Darby, and Armstrong 1994 and 1997; Zucker and Darby 1996; Zucker, Darby, Brewer, and Peng 1996; Liebeskind, Oliver, Zucker, and Brewer 1996; Zucker, Brewer, Oliver, and Liebeskind 1993). The validity of our co-publishing indicator for the existence of contractual or ownership relationships with firms has been confirmed through

extensive interviews conducted with university scientists and administrators, and with firm scientists, CEOs, and corporate board members (for U.S. examples, see Zucker and Darby 1997; Zucker, Brewer, Oliver, and Liebeskind 1993).

15. Most of research on part-time work and/or multiple jobs focuses on low skill, low wage employment and "moonlighting." The common, and perhaps even typical, pattern of top academic scientists routinely and recurrently moving a significant part of their labor outside the university to another organization, sometimes created by them, has received much less empirical attention. Labor effort can be quite mobile. Part-time doesn't necessarily mean marginal, either in terms of the amount of effort nor in terms of the effects of that effort on productivity, here of both the firm and the scientist. Part-time "consulting" or control of an outside business often involves substantial labor effort; at least in our research on biotechnology, we find strong positive effects of that effort on productivity of both the firm and the scientist (Zucker and Darby 1996b). Possible benefits to the university include paying lower wages than would otherwise be necessary, receiving acclaim for the net productivity of the scientist (including the--sometimes higher--productivity achieved through outside employment), and increased visibility of the university in non-academic arenas (increasing fund raising success among entrepreneurs, for example).

Table 1

Mean Numbers of Articles and Citations by Commercial Ties of U.S. Star Scientists

	Mean number of articles per year			
Type of stars	pre-firm	during firm	after firm	never in firm
Affiliated stars	2.080	2.481	3.500	
Local linked stars	2.297	3.322	1.789	
External linked stars	1.978	2.452	2.077	
Never linked or affiliated stars				1.611

	Mean number of citations per article			
Type of stars	pre-firm	during firm	after firm	never in firm
Affiliated stars	14.495	29.738	6.868	
Local linked stars	13.723	16.672	8.145	
External linked stars	16.914	17.335	6.830	
Never linked or affiliated stars				11.710

Note: Mean citations are the ratio of total citations in the *Science Citation Index* for 1987 and 1992 summed over all the genetic-sequence-discovery articles (up to April 1990) in Genbank (1990) authored or coauthored by each of the stars in the cell to the total numbers of those article summed over all the stars in the cell.

Table 2
Descriptive Statistics

Variables	Mean	S.E.	Min.	Max.
INDIVIDUAL CHARACTERISTICS				
Gender of star scientist (M=1, F=0)	0.96	0.20	0	1
Age of star scientist	54.22	10.79	38	91
Age squared	3055.61	1245.95	1444	8281
Marital dummy (M=1, otherwise 0)	0.77	0.42	0	1
First year star publishes in GenBank	1980.51	4.94	1967	1989
Number of children	1.69	1.21	0	7
QUALITY CHARACTERISTICS OF STAR				
Total number of articles in GenBank	9.35	10.22	1	55
Total citations to articles in GenBank	126.38	185.31	0	953
Nobel prize dummy (yes=1, no=0)	0.04	0.19	0	1
Tenure dummy (yes=1, no=0)	0.88	0.33	0	1
CHARACTERISTICS OF UNIVERSITY OR RI				
University *top quality dummy* (yes=1, no=0)	0.49	0.50	0	1
University average reputation	3.915	0.67	1.7	4.93
MIT or Harvard University dummy (yes=1, no=0)	0.15	0.35	0	1
Stanford or UC-San Francisco dummy (yes=1)	0.12	0.32	0	1
National Cancer Institute dummy (yes=1, no=0)	0.05	0.22	0	1
Index of wages	0.86	0.18	0.42	1.71
INDICATOR OF COMMERCIAL POTENTIAL				
Number of human genetic sequences (Human = sequence type 1 or type 4)	1.54	5.30	0	49
REGIONAL VARIABLES				
New biotech enterprises in region (count)	26.62	23.51	0	82
Top quality universities in region (count)	1.44	1.18	0	3
Average change in citations of other stars in *same* region while affiliated or link to firm	0.19	0.62	-1	2.22
Average change in citations of other stars in *different* region while affiliated or linked to firm	0.15	0.55	-0.73	3.33
INDICATORS OF SIZE OF SOCIAL NETWORKS				
Proportion coauthors from different institutions	0.31	0.22	0	1
Number of times star changes univ. or res. inst.	2.35	1.28	1	7
DEPENDENT VARIABLE				
Star scientist movement to firm (1 = period star is first affiliated or linked, 0 otherwise)	0.39	0.49	0	1

N =248

Table 3
Duration Model of Mobility to Firms of Star Scientists in the U.S.

Variables	Coefficients (standard errors)			
	model a	model b	model c	model d
Constant	110.120	96.232	216.740 **	139.960*
	(57.840)	(78.432)	(0.013)	(74.366)
Gender of star scientist	0.220	0.425		
	(0.730)	(0.857)		
Marital dummy	-0.068	0.479		
	(0.432)	(0.515)		
Age	0.157	0.042		
	(0.125)	(0.151)		
Age squared	-0.001	-0.000		
	(0.001)	(0.001)		
Number of children	0.068	0.017		
	(0.147)	(0.167)		
First year star publishes in GenBank	-0.058 *	-0.049	-0.108 **	-0.072
	(0.028)	(0.039)	(0.044)	(0.038)
Nobel prize dummy		0.662	0.624	
		(0.877)	(0.921)	
Tenure dummy		-0.639	-0.453	
		(0.506)	(0.472)	
Total citations to articles in GenBank		0.005 ***	0.004 ***	0.004 ***
		(0.001)	(0.001)	(0.001)
Total number of articles in GenBank		-0.017	-0.022	-0.005
		(0.020)	(0.021)	(0.019)
University top quality dummy		0.463	0.711	0.387
		(0.521)	(0.534)	(0.472)
University average reputation			-0.415	
			(0.311)	
Stanford or UC-San Francisco dummy		0.428		
		(0.726)		
MIT or Harvard University dummy		0.912		
		(0.694)		
National Cancer Institute dummy		0.934		
		(0.738)		
Number of human genetic sequences		0.095 *	0.090 *	0.087 *
		(0.040)	(0.039)	(0.039)
New biotech enterprises in region		0.214 *	0.025 **	0.023 **
		(0.010)	(0.009)	(0.009)
Top quality universities in region		-0.921 **	-0.677 **	- 0.689 **
		(0.327)	(0.247)	(0.242)
Proportion coauthors from different institutions			1.657 *	1.518 *
			(0.708)	(0.697)
Number of times star changes univ. or res. inst.			-0.262	
			(0.158)	
Average change in citations of other stars in same region while affiliated or linked to firm		0.735 **	0.776 **	0.731 **
		(0.277)	(0.272)	(0.265)
Average change in citations of other stars in different region while affiliated or linked to firm		-0.238	-0.171	-0.044
		(0.313)	(0.311)	(0.292)
Log Likelihood	-160.676	-133.269	-131.189	-134.441
Restricted Log Likelihood.	-165.523	-165.523	-165.523	-165.523

Significance levels: *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001

Table 4
Duration Model of Mobility to Firms of Star Scientists in the U.S.
Including Index of Wages at Current University or Research Institute

Variables	Coefficients (standard errors)		
	Model a	Model b	Model c
Constant	0.627	200.220*	121.560
	(0.664)	(88.244)	(74.554)
Index of wages at university or research institute	-1.266	-1.848	-1.685
	(0.763)	(1.055)	(1.029)
First year star publishes in GenBank	-0.100*	-0.062	
		(0.044)	(0.038)
Nobel prize dummy		0.682	
		(0.956)	
Tenure dummy		-0.330	
		(0.477)	
Total citations to articles in GenBank		0.005**	0.004***
		(0.001)	(0.001)
Total number of articles in GenBank		-0.020	0.000
		(0.023)	(0.020)
University top quality dummy		0.653	0.305
		(0.545)	(0.485)
University average reputation		-0.480	
		(0.318)	
Number of human genetic sequences		0.089*	0.085*
		(0.039)	(0.038)
New biotech enterprises in region		0.020*	0.018
		(0.010)	(0.010)
Top quality universities in region		-0.499	-0.540*
		(0.269)	(0.261)
Proportion coauthors from different institutions		1.734*	1.568*
		(0.716)	(0.701)
Number of times star changes university or res. inst.		-0.275	
		(0.159)	
Average change in citations of other stars in same region while affiliated or linked to firm		0.846**	0.797**
		(0.276)	(0.271)
Average change in citations of other stars in different region while affiliated or linked to firm		-0.205	-0.067
		(0.314)	(0.293)
Log Likelihood	-164.088	-129.592	-133.040
Restricted Log Likelihood.	-165.523	-165.523	-165.523

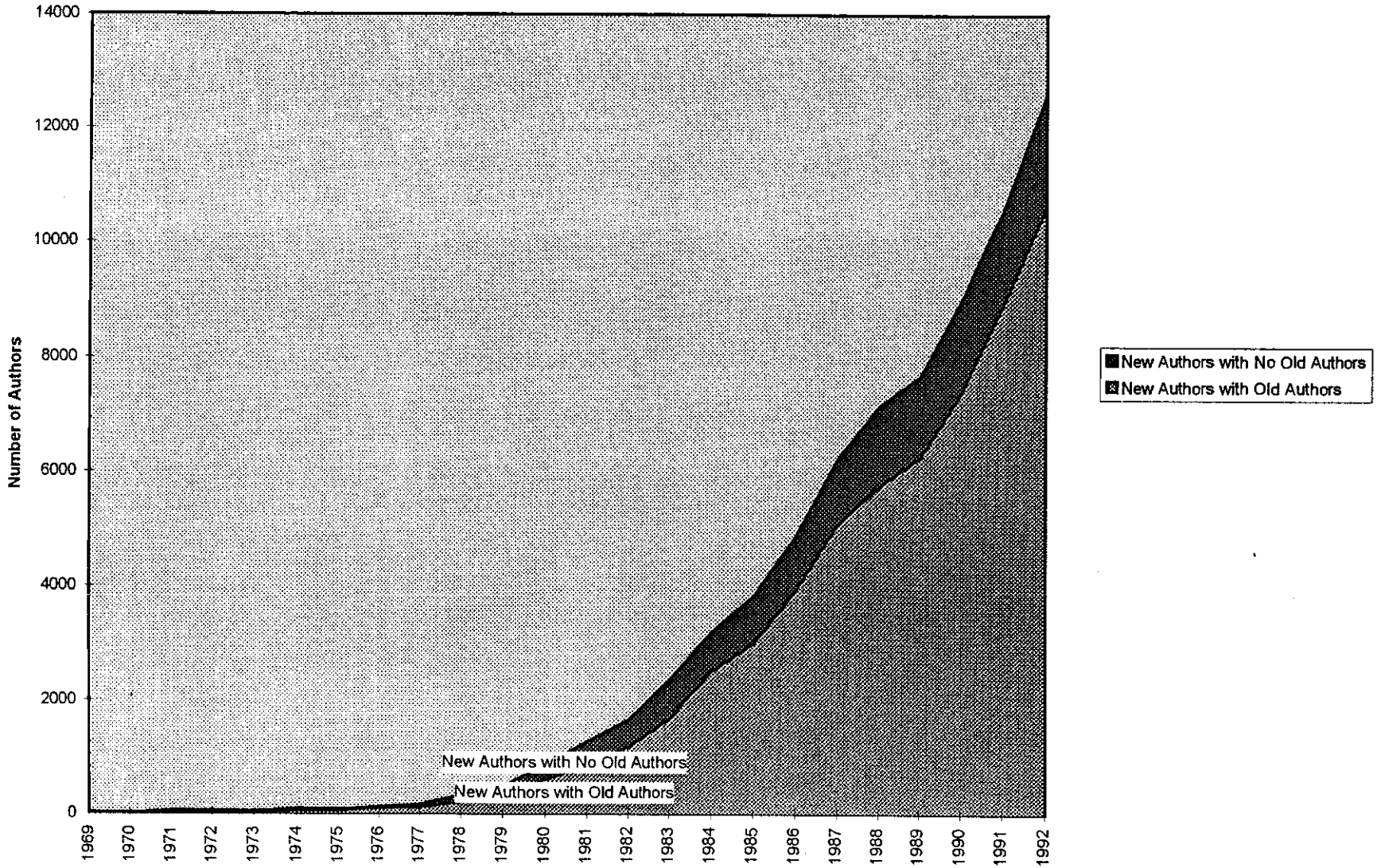
Significance levels: *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001

Table 5
Multinomial Logit Model of Choice of Becoming Affiliated with or Linked to a Firm

Variables	Coefficients (standard errors)	
	Affiliated	Linked
Constant	518.520***	219.450**
	(159.870)	(90.533)
First year star publishes in GenBank	-0.264***	-0.112**
	(0.081)	(0.046)
Total citations to articles in GenBank	0.022***	0.020***
	(0.005)	(0.005)
Total number of articles in GenBank	-0.102*	-0.011
	(0.053)	(0.028)
University top quality dummy	1.080	0.992
	(1.102)	(0.546)
Number of human genetic sequences	0.165	0.330***
	(0.284)	(0.101)
New biotech enterprises in region	0.046*	0.021*
	(0.022)	(0.011)
Top quality universities in region	-0.748	-0.664**
	(0.537)	(0.277)
Proportion coauthors from different institutions	1.146	1.863*
	(1.819)	(0.776)
Average change in citations of other stars in same region while affiliated or linked to firm	0.919	0.228
	(0.598)	(0.311)
Average change in citations of other stars in different region while affiliated or linked to firm	-1.093	-0.187
	(1.142)	(0.331)
Log Likelihood	-136.5233	
Restricted Log Likelihood	-216.0000	

Significance levels: *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001

Figure 1
New Authors in GenBank by Whether or Not They Enter by Coauthoring with Old Authors



Source: Calculations of the authors based upon GenBank (1994).

Figure 2
Mobility from Academe to Commerce of U.S. Biotechnology Star Scientists by Region as of 1990