Labor Mobility from Academe to Commerce

Lynne G. Zucker, *University of California, Los Angeles, and National Bureau of Economic Research*

Michael R. Darby, *University of California, Los Angeles, and National Bureau of Economic Research*

Maximo Torero, *Grupo de Análisis para el Desarrollo (GRADE), and University of California, Los Angeles*

Breakthroughs with natural excludability are transferred to industry by top academic scientists (stars) working in or with firms. Movement to firms depends on scientists' quality, moving costs, and reservation wage. Scientists' quality, moving costs, trial frequency, interfering academic offers, and productivity of stars already in firms determine

This research has been supported by grants from the University of California Systemwide Biotechnology Research and Education Program, the University of California President's Industry-University Cooperative Research Program, the Alfred P. Sloan Foundation through the National Bureau of Economic Research (NBER) Research Program on Industrial Technology and Productivity, and the National Science Foundation (SES 9012925). This article 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 postdoctoral 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 Jeff Armstrong and Rick Flyer substantially improved this article. This article is a part of the 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.

[*Journal of Labor Economics,* 2002, vol. 20, no. 3]

0734-306X/2002/2003-0007$10.00

reservation wage. In group-duration analysis for biotechnology, stars move to firms faster as their quality, human focus, and outside coauthorships increase; local firms and productivity of local stars in firms increase; and top local universities decrease. Stars move to firms full or part time similarly, but significance drops for rarer full-time moves.

> Technology transfer is the movement of ideas in people. (Donald Kennedy, Stanford University, March 18, 1994)

I. Introduction

In breakthrough discoveries where scientific productivity becomes relevant to commercialization, the labor of the most productive scientists is the main resource around which firms are built or transformed. We can think of the top scientists as the "seeds" around which crystals form. The scientist is the key resource, but he or she must attract other resources to augment his or her own research productivity. Further, when that research is highly commercializable, stars similarly augment the productivity of a firm when they move their labor from universities and research institutes to firms. But how do these top academic scientists become involved in commercializing their discoveries?

We model labor mobility as a function of the 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 (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 involved 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 1996*a,* 1997), but many of the firms are newly created around these "star" scientists, who often become residual owners as well as employees (Zucker, Darby, and Brewer 1998).

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 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.

Our analysis is organized as follows. In Section II, we discuss the issues that need to be taken into account to be able to understand labor mobility as a technology transfer process in biotechnology. In Section III, we develop a theoretical model to explain the decision process of a star scientist when moving to a firm, and in Section IV we detail the econometric techniques used to estimate our model. The main results are then explained in Section V. Finally, in Section VI, we conclude by discussing our identification of an important neglected set of processes that allow retention of at least part of the value of knowledge by a discoverer and explaining how this knowledge is transferred through labor mobility to the firms.

II. Labor Mobility as Technology Transfer

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; therefore, these scientists become the main resource around which firms are built or transformed (Zucker, Darby, and Armstrong 1998; Zucker, Darby, and Brewer 1998). "Star" scientists are, therefore, important in the process of technology transfer because of the value of their knowledge to the success of firms.[1]

[1] In related research, we have found that the one variable that has a significantly positive effect on all measures of firm success is the count of "linked" articles authored by stars with firm employees. These linked stars are most often local academic scientists-entrepreneurs who possess a significant equity or founding interest in the firm. The number of such articles serves as an indicator of the depth of the star's involvement with the firm's research effort. Just two such linked articles results in about one more product in development, about one more product on the market, and about 345 more person employment growth from 1989 to 1994. For five such articles, the impact was 4.7 more products in development, 3.5 more products on the market, and about 860 more employees (Zucker and Darby 1996*b*; Zucker, Darby, and Armstrong 1998). These results generalize. In Japan, articles written with biotech stars have been shown to determine which firms are most

A. Barriers to Information Flow: Stars' Knowledge Advantage

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. 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.[2]

Coauthoring, which implies bench-level collaboration, provides our measure of tacitness: the degree of tacitness is high when most new authors in an area of research are publishing with at least one old author, defined as one who has previously published in the same field, and low if most new entrants to the field can do the research either by himself or herself or with all new authors.[3] For biotechnology, the field is well defined by inclusion in GenBank, which is a worldwide directory of all articles reporting newly discovered genetic sequences.

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 is also shown in figure 1, our tacitness measure declines more slowly than scarcity. In fact, new scientists continue to enter throughout the 1969–92 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% of entry from 1969 through 1992.[4] Excluding sole-

successful (Zucker and Darby 2001). In the United States, Torero (1998) finds star scientists' involvement explains semiconductor firms' success.

[2] Exceptions typically include the handful of scientists working in the same 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 be only to the graduate students and postdocs working in the same lab as the discovering scientists.

[3] 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.

[4] 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% of total entry.

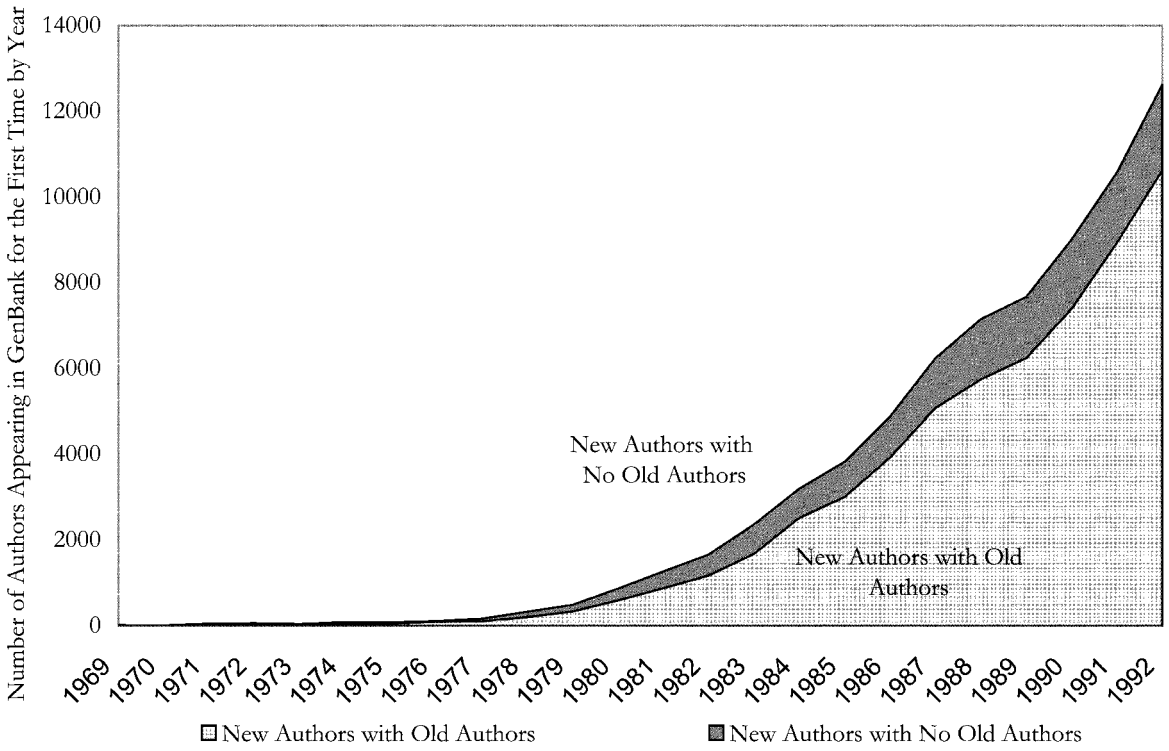

FIG. 1.—New authors in GenBank by whether or not they enter by coauthoring with any authors who have published previously

authored articles, which may be dissertations for new authors and review articles by established authors, new authors write exclusively with other new authors 36% less frequently than old authors write exclusively with other old authors.[5] The overall significance of these differences was confirmed in a loglinear analysis ($\chi^2 = 1{,}265.45$; $G^2 = 1{,}202.83$; $p < .0001$) for both values.

B. Transmission of Information: Stars' Velocity of Movement to Firms

Is there an explanation of how quickly a star scientist will move? We define velocity of movement as the inverse of the expected number of years that a star scientist will take to move to a firm. We see three possible interpretations of the velocity of movement of scientists to firms in biotechnology. The first explanation is based on existence of a market imperfection, such as a monopoly or cartel. For example, existing pharmaceutical firms might collude to block the new technology as threatening their market-sharing arrangements. In this case, the velocity of movement depends on how long it takes the market to break the cartel. The second explanation is a pure diffusion of information problem in which what is critical is how long it takes for the companies to get pure information about the quality and capacity of the respective scientists (see Conlin 1999). Therefore, the fact that the firms have to make a decision under uncertainty is central.

Finally, the third explanation is a pure diffusion of knowledge problem in which what is critical is how long it takes for the scientist to realize the commercial value of their discoveries. In this case, the importance of the scientist's work (measured through citations), the type of research they do (human related or not), as well as their relationship with other scientists with commercial experience, is crucial.

In all of these instances, analyzing the velocity of movement is interesting. In the first explanation, we can find how quickly the market imperfection is solved. In the second explanation, velocity will depend on how long it takes to get pure information and the extent to which the companies are risk adverse. Finally, in the third explanation, we can find out how fast the scientists realize the value of their research, depending on characteristics of their research, their research ties, and their personal risk aversion (determining how quickly they will be willing to move out from the academic environment).

Specifically, for the purposes of this article and given the nature of the

[5] Sole-authored articles account for only 6.5% of the authorships of new authors and 7.8% of the authorships of old authors over this period. It is interesting that new sole authors became 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).

sector with which we are working, the second and third alternatives are the most empirically relevant. Further, the measures used for productivity are generally considered good indicators of the quality of the scientists; moreover, companies actually use publications and citations as measures of a scientist's quality (this has been confirmed by speaking with top R&D executives at a number of major biotechnology companies). Therefore, at least in science-driven industries such as biotechnology, firms have a way to measure (or at least an excellent indicator of) the quality of the scientists and can thus make offers to the best of the star scientists, directly influencing the speed at which these scientists move. Of course, there are other variables that are important in the model, such as development of human-related sequences and the network size of the scientists (which increases the probability that an individual scientist's work will be recognized).

Now, as mentioned previously, there are three possible alternative movements of stars to firms: (*a*) a scientist completely moves to an existing firm (affiliated), (*b*) a scientist partially moves to an existing firm (local linked or external link), and (*c*) a scientist builds a new firm working there full time (affiliated) or part time (linked). In the first two cases, the firms will be the ones making the offers according to the quality of the scientist. However, if the firms are not able to understand the quality of the scientist's work, the scientist could realize the commercial value of his invention and, therefore, initiate a startup.[6] On the other hand, the time to move will be affected, but to a smaller degree, by a partial adjustment process until the market clarifies the price, though given the scarcity and tacitness of this new technology, the latter will not be the main variable behind the analysis. 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.

III. The Model

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.[7] 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

[6] Because the biotechnology revolution involved much different scientific skills than those used at the then-incumbent pharmaceutical firms, the importance of the breakthrough was not initially recognized by the bulk of the incumbent firms (Zucker and Darby 1997). However, successful adoption of biotechnology was ultimately essential to their survival (Darby and Zucker 1999).

[7] This is equivalent to the velocity of movement defined above.

marginal product.[8] 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 (see the list of variables in table 1; details are provided in app. A). Increases in each of those variables would increase the expected value of a star scientist to any given firm:

$$\Pr(w_0 < z;\ \mathbf{Q}) = G(z;\ \mathbf{Q}), \quad \mathbf{G}_{\mathbf{Q}}(z;\ \mathbf{Q}) < 0, \tag{1}$$

where $\mathbf{G}_{\mathbf{Q}}(z;\ \mathbf{Q})$ 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 that exceeds any given value generally increases with the characteristics in $\mathbf{Q}$.[9]

Four distinct types of potential alternate (or joint) employers typically employ star bioscientists moving from a university (in whole or part): 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, as indicated by articles or patents, 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 we distinguish between full- and part-time work with firms.

Ideally, the reservation wage for leaving the current employer would be

[8] 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 that reflects the variation in expected values of marginal product across employers. The location of this distribution should vary across individuals according to their characteristics that 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: $\Pr(w_0 < z;\ X) = G(z;\ X), GX(z;\ X) < 0$. Topel and Ward note that experience increases wage offers if the last inequality in the previous equation is strict, but observed wages will increase with experience due to search even if expected productivity is independent of experience $[GX(z;\ X) = 0]$.

[9] In principle, the labor market experience X, mentioned by Topel and Ward, might be an element of $\mathbf{Q}$, but we see below that X is generally insignificant as a predictor of mobility, with the sole exception of the first year the scientist publishes in GenBank, an experience variable that is both highly specific and very relevant to the firm (in analyses not reported here, the other experience variables remain nonsignificant even when the "first year" variable is removed from the equation).

Table 1
Descriptive Statistics

Variables	Mean	SD	Minimum	Maximum
Individual characteristics:				
Gender of star scientist ($M = 1, F = 0$)	.96	.20	0	1
Age of star scientist	54.22	10.79	38	91
Age squared	3,055.61	1,245.95	1,444	8,281
Marital dummy ($M = 1, 0$ otherwise)	.77	.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 prior articles in GenBank	9.35	10.22	1	55
Total citations to prior articles in GenBank	126.38	185.31	0	953
Nobel Prize dummy (yes = 1, no = 0)	.04	.19	0	1
Tenure dummy (yes = 1, no = 0)	.88	.33	0	1
Characteristics of university or research institute:				
University top quality dummy (yes = 1, no = 0)	.49	.50	0	1
University average reputation	3.92	.67	1.7	4.93
MIT or Harvard University dummy (yes = 1, no = 0)	.15	.35	0	1
Stanford or University of California, San Francisco, dummy (yes = 1)	.12	.32	0	1
National Cancer Institute dummy (yes = 1, no = 0)	.05	.22	0	1
Index of wages	.86	.18	.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
Proportionate change in annual citation rate of other stars in same region while tied to firm	.19	.62	−1	2.22
Proportionate change in annual citation rate of stars in different regions while tied to firm	.15	.55	−.73	3.33
Indicators of size of social networks:				
Proportion coauthors from different institutions	.31	.22	0	1
Number of times star changes university or research institute	2.35	1.28	1	7
Dependent variable:				
Star scientist movement to firm (1 = period star is first affiliated or linked, 0 otherwise)	.39	.49	0	1

$N = 248$

an increasing function of the wage earned there, but we do not have data on individual wages, and a wage index for the scientist's university proved inadequate empirically (this is discussed at table 4 below). So we assume that current wages are also a function of **Q** but that egalitarian pressures within the university as well as the potentially greater returns to commercial applications of the star's intellectual human capital imply that higher values of any of the elements of **Q** shift the location of the G function by more than the reservation value R. Thus, the probability of an acceptable offer from any of the six types of potential employment (indexed by i) increases in **Q** also:

$$\Pr(w_0 < R;\ \mathbf{Q}, i) = G(R;\ \mathbf{Q}, i), \quad G_Q(z;\ \mathbf{Q}, i) < 0, \tag{2}$$

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 job, and 6 for external university job.

This assumption is more obvious for movements to firms ($i = 1, 2, 3,$ or 4) than for universities ($i = 5$ or 6), but academics frequently note that much greater weight is placed on externally visible research productivity in external hiring than in promoting from within. In any case, our principal concern in this article is explaining embodied technology transfer from universities to firms, so movements to other universities enter only as potential temporary interference with that process.

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

$$\lambda(t;\ \mathbf{Q}, \mathbf{H}) = \sum_{i=1}^{4} \lambda_i(t;\ \mathbf{Q}_{t-1}, \mathbf{H}_{t-1}), \tag{3}$$

where **Q,** as before, is our vector of externally observed measures of intellectual human capital and **H** represents other factors affecting the hazard rate. Both **Q** and **H** are observed in the previous period to eliminate potential endogeneity problems. 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 (2) 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 that might belong in **H** are those that increase the rate at which individual employer-scientist matches are considered per unit of time. We, again, refer to table 1. Other things equal, we expect that the lower 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; we therefore include the lagged 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 interuniversity movements.[10] 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 1 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 that may reduce the reservation wage for firm employers is favorable working conditions in the form of increased scientific productivity. 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.[11] 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 observed increase in productivity by other star scientists who have previously moved to firms: (*a*) the proportionate change in the annual citation rate for other local stars during years through the prior period comparing citations to their articles written during to articles written prior to ties to a firm and (*b*) the proportionate change in the annual citation rate for tied stars in other regions during years through the prior period comparing citations to their articles written during to articles written prior to ties to

[10] 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.

[11] 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 1998). 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.

a firm.[12] We expect that a star scientist will be more aware of the scientific payoff experienced by stars in his or her own region than elsewhere in the country, but that is an empirical question. Again, these variables are listed in table 1 under regional variables. Region is defined in this article by the functional economic areas in U.S. Department of Commerce, Economics, and Statistics Administration, Bureau of Economic Analysis (1992).

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 Armstrong 1998; Zucker, Darby, and Brewer 1998). That is, what appears to be employment may in fact be entrepreneurship. We expect that characteristics that predict a high marginal product to potential employers will also be attractive to potential investors, so the analysis is not greatly affected by whether the scientist is searching for an employer or venture capital. Indeed, prospectuses for initial public offerings of new biotechnology firms frequently list precisely the sorts of qualifications in **Q** for key associated scientists. Since there is a significantly positive agglomeration effect, as 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 also be positive here.

We use group duration analysis 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 velocity and probability that a star will become employed by or collaborate with a firm and that these effects will dominate traditional measures such as experience. The duration process captures the timing and motivational aspects of the move decision. Finally, we use the multinomial choice model to analyze the type of move made, comparing linked versus affiliated. Before turning to the estimation, in the next section we first introduce the variables not yet discussed and then provide a brief explication of the methodology. Details of the econometric modeling may be found in appendix B.

IV. Econometric Methods

We estimate the model using group-duration techniques described in this section. Basically, the time that a scientist works in academe without firm involvement is described as a sequence of discrete periods. Each scientist is observed for each of these discrete periods until he or she

[12] The scientists are assumed to directly observe the quality and quantity change in the articles published by those stars at the time of publication, with citations the econometrician's retrospective proxy for the observed scientific productivity increase.

moves to a firm. The estimation basically computes a conditional logit for each period conditional upon not having moved previously.

The expected amount of time the scientist stays in universities without moving to a firm differs because each scientist has an individual time-varying vector **Q** of indicators of expected value of marginal product (whether employed by a top-quality university, tenured there, the quality and quantity of articles published previously, and the number of human genetic sequences discovered) and faces different local economic areas that alter each scientist's probability of receiving an offer from a firm greater than his or her 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.

Most individual variables are defined in the period prior to each observation, including the total number of articles and citations to these articles, the number of human sequences, the percentage of collaborators from outside his or her organization, and proportionate change in annual citation rate for other stars in the same or other region while (up through the prior period) affiliated or linked to a firm.[13] The number of local biotechnology firms and variables that describe the university or research institute (top-quality university, current wage index, location in specific key university clusters, Stanford/University of California, San Francisco, or MIT/Harvard, and location at the National Cancer Institute) are updated each time the scientist changes university or institute without moving to a firm.

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 of moves are based on articles published by year and so time is measured (grouped) at intervals, available discretely at the level of the year. These models have the advantage of being derived from discrete time and therefore more directly conform to the discrete nature of the data. 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 app. A and the reference list for examples) or the first date of publishing in GenBank, and the first article in GenBank that shows him or her affiliated or linked to a firm through coauthorship. See appendix B for a derivation of the group duration procedure that we use in our main analyses.

[13] Citations are all past and future citations to past articles as the econometrician's proxy for article quality, which is assumed to be directly observable by scientists and firms.

As described in appendix B, group-duration information can be summarized as a sequence of binary outcomes (exit or survive in each successive period), allowing us to apply a logistic function that 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.

Finally, we selected multinomial logit to explore the determinants of selecting different relationships to a firm. Each star scientist is assumed to have preferences defined over a set of alternatives: untied, affiliated with a firm, or linked to a firm. Since this technique is more commonly used, we do not go into further detail on it.

V. Empirical Results for the Group Duration Model

The results reported in tables 2 and 3 are generally supportive of the suppositions contained in our mobility model. Table 2 reports the standard coefficient estimates, and table 3 reports the corresponding partial derivatives of probabilities with respect to the vector of characteristics (marginal effects).[14] 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 the later the year of entry, the less probable that the star becomes affiliated with 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.[15] Even more, using the marginal effects, it can be shown that an increase in 10 citations will result in an increase in 1% of the probability of moving to a firm. The insignificant coefficient on the quantity of articles suggests that firms do not 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

[14] As is well known, in the logit model, the coefficients do not indicate the increase in the probability of the event occurring, given a one-unit increase in the corresponding independent variable; rather, they reflect the effect of a unit change in an independent variable on $\ln (P_i/(1 - P_i)$, the log of the odds ratio. The marginal effects are computed at the means of the explanatory variables. In the case of dummy variables, the marginal effect is computed as the difference in the probability function when it is evaluated at the values 1 and 0.

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

Table 2
Duration Model of Mobility to Firms of Star Scientists in the United States: Coefficient Estimates

Variables	Model A	Model B	Model C	Model D
Constant	110.123	96.232	216.738**	139.960*
	(57.840)	(78.432)	(87.552)	(74.366)
Gender of star scientist	.220	.425		
	.728)	(.857)		
Marital dummy	−.068	.479		
	(.432)	(.515)		
Age	.157	.042		
	.125)	(.151)		
Age squared	−.001	−.000		
	(.001)	(.001)		
Number of children	.068	.017		
	(.147)	(.166)		
First year star publishes in GenBank	−.058*	−.050	−.109**	−.072
	(.029)	(.039)	(.044)	(.038)
Nobel Prize dummy		.662	.624	
		(.877)	(.922)	
Tenure dummy		−.639	−.453	
		(.506)	(.472)	
Total citations to prior articles in GenBank		.005***	.004***	.004***
		(.001)	(.001)	(.001)
Total number of prior articles in GenBank		−.017	−.023	−.005
		(.020)	(.022)	(.020)
University top quality dummy		.463	.712	.387
		(.522)	(.534)	(.472)
University average reputation			−.415	
			(.311)	
Stanford or University of California, San Francisco, dummy		.428		
		(.726)		
MIT or Harvard University dummy		.912		
		(.694)		
National Cancer Institute dummy		.934		
		(.738)		
Number of human genetic sequences		.095*	.090**	.087*
		(.040)	(.040)	(.039)
New biotech enterprises in region		.021*	.025**	.023**
		(.010)	(.010)	(.009)
Top-quality universities in region		−.921**	−.677**	−.689**
		(.327)	(.247)	(.242)
Proportion coauthors from different institutions			1.657*	1.511*
			(.708)	(.697)
Number of times star changes university or research institution			−.262	
			(.158)	
Proportionate change in annual citation rate of other stars in same region while tied to firm		.735**	.776**	.732**
		(.277)	(.272)	(.266)
Proportionate change in annual citation rate of stars in different regions while tied to firm		−.239	−.171	−.044
		(.313)	(.312)	(.292)
Log likelihood	−160.676	−133.270	−131.190	−134.190
Restricted log likelihood	−165.523	−165.523	−165.523	−165.523
Wald test	9.09	43.32***	46.28***	42.58***

NOTE.—Standard errors are in parentheses.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 3
Duration Model of Mobility to Firms of Star Scientists in the United States: Marginal Effects

Variables	Model A	Model B	Model C	Model D
Constant	26.013	22.464	50.426**	32.724*
	(13.660)	(18.279)	(20.226)	(17.285)
Gender of star scientist	.052	.099		
	(.172)	(.200)		
Marital dummy	−.016	.112		
	(.102)	(.120)		
Age	.037	.010		
	(.029)	(.035)		
Age squared	−.000	−.000		
	(.000)	(.000)		
Number of children	.016	.004		
	(.035)	(.039)		
First year star publishes in GenBank	−.014*	−.012	−.025**	−.017
	(.007)	(.009)	(.010)	(.009)
Nobel Prize dummy		.155	.145	
		(.205)	(.215)	
Tenure dummy		−.149	−.105	
		(.118)	(.110)	
Total citations to prior articles in GenBank		.001***	.001***	.001***
		(.000)	(.000)	(.000)
Total number of prior articles in GenBank		−.004	−.005	−.001
		(.005)	(.005)	(.005)
University top quality dummy		.108	.166	.090
		(.121)	(.124)	(.110)
University average reputation			−.097	
			(.072)	
Stanford or University of California, San Francisco dummy		.100		
		(.169)		
MIT or Harvard University dummy		.213		
		(.161)		
National Cancer Institute dummy		.218		
		(.172)		
Number of human genetic sequences		.022*	.021**	.020*
		(.009)	(.009)	(.009)
New biotech enterprises in region		.005*	.006**	.005**
		(.002)	(.002)	(.002)
Top-quality universities in region		−.215**	−.158**	−.161**
		(.075)	(.057)	(.056)
Proportion coauthors from different institutions			.385*	.353*
			(.163)	(.162)
Number of times star changes university or research institute			−.061	
			(.037)	
Proportionate change in annual citation rate of other stars in same region while tied to firm		.172**	.180**	.171**
		(.064)	(.063)	(.062)
Proportionate change in annual citation rate of stars in different regions while tied to firm		−.056	−.040	−.010
		(.073)	(.072)	(.068)

NOTE.—Standard errors are in parentheses.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

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 discovered, 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. Discovering one additional genetic sequence will increase by 2% the probability of a star scientist's moving to a 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 have opposite effects 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. Examining table 3, we see that a unit increase in the number of top-quality universities has a bigger interfering effect than a unit increase in the number of firms has an attracting effect. However, there are many more firms than top-quality universities, so that the firm effect is dominant even where top-quality universities are present. Figure 2 shows the total number of stars, the numbers of stars that move at least some of their labor effort to a firm, and the number of firms (rescaled by dividing by 4) in each of the top-20 biotechnology regions (those with four or more stars or 10 or more firms). These 20 regions account for 88%, 85%, and 78%, respectively, of these totals for all 183 U.S. regions.

Econometrically, this relationship could be argued to be because of possible endogeneity of the scientists' original location, that is, scientists who anticipate moving look for those characteristics such as many new biotechnology firms in the region rather than regional characteristics increasing the probability of moving. Although this could be a possibility, the fact that we are analyzing the industry at its beginning and that we are looking to the behavior of the initial stock of scientists who develop the major innovations in biotechnology and even gave birth to some of the new biotechnology enterprises (both founding new entrants and bringing new technology to incumbent firms) reduces the likelihood of significant reverse causality. Moreover, when we follow the mobility of scientists moving between universities to see if they move to universities in regions where there are more biotechnology firms, the result is the opposite: only 21.8% of scientists moving to another university move to

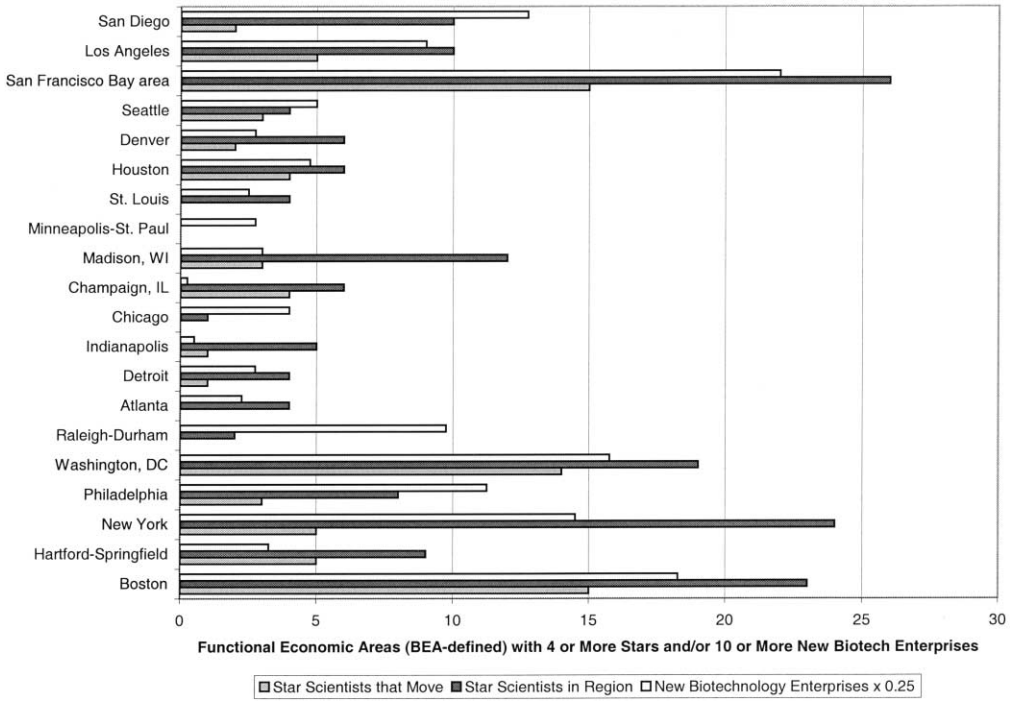

FIG. 2.—Mobility from academe to commerce of U.S. biotechnology star scientists by region as of 1990

a region where there were more biotech firms, while 67.2% move to a university in a region where there were a smaller number of biotech firms relative to the Bureau of Economic Analysis (BEA) region from which they came.[16]

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

Overall, tables 2 and 3 provide 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 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 wage index generally entered with the right sign, but it never entered significantly at the .05 level. Table 4 presents these results.

Table 5 reports the estimation of a multinomial logit model that examines the choice of staying untied to a firm or becoming affiliated with or linked to a firm. When we analyze only the coefficients (the first two columns) relative to being untied, 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

[16] The remaining 10.9% are regions with the same number of new biotechnology enterprises.

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

	Coefficients (Standard Errors)			Marginal Effects (Standard Errors)		
Variables	Model A	Model B	Model C	Model A	Model B	Model C
Constant	.627	200.219*	121.559	.149	46.483*	28.351
	(.664)	(88.244)	(74.554)	(.158)	(20.357)	(17.301)
Index of wages at university or research institute	−1.266	−1.848	−1.685	−.300	−.429	−.393
	(.763)	(1.055)	(1.029)	(.180)	(.244)	(.239)
First year star publishes in GenBank		−.100*	−.062		−.023*	−.014
		(.044)	(.038)		(.010)	(.009)
Nobel Prize dummy		.682			.158	
		(.956)			(.222)	
Tenure dummy		−.330			−.077	
		(.477)			(.111)	
Total citations to prior articles in GenBank		.005***	.004***		.001***	.001***
		(.001)	(.001)		(.000)	(.000)
Total number of prior articles in GenBank		−.020	.000		−.005	.000
		(.023)	(.020)		(.005)	(.005)
University top quality dummy		.653	.305		.152	.171
		(.545)	(.485)		(.126)	(.123)
University average reputation		−.480			−.111	
		(.318)			(.074)	
Number of human genetic sequences		.089*	.085*		.021*	.020*
		(.039)	(.038)		(.009)	(.009)

New biotech enterprises in region		.020	.018		.005	.004
		(.010)	(.010)		(.002)	(.002)
Top-quality universities in region		−.499	−.540*		−.116	−.126*
		(.269)	(.261)		(.062)	(.061)
Proportion coauthors from different institutions		1.734*	1.568*		.402*	.366*
		(.716)	(.701)		(.165)	(.163)
Number of times star changes university or research institute		−.275			−.064	
		(.159)			(.037)	
Proportionate change in annual citation rate of other stars in same region while tied to firm		.846**	.797**		.196**	.186**
		(.276)	(.271)		(.064)	(.063)
Proportionate change in annual citation rate of stars in different regions while tied to firm		−.205	−.067		−.048	−.016
		(.314)	(.293)		(.073)	(.068)
Log likelihood	-164.088	−129.592	−133.040	N.A.	N.A.	N.A.
Restricted log likelihood	−165.523	−165.523	−165.523	N.A.	N.A.	N.A.
Wald test	2.760	46.600***	42.820***	N.A.	N.A.	N.A.

NOTE.—N.A. = not available.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

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

	Coefficients (Standard Errors)		Marginal Effects (Standard Errors)		
Variables	Affiliated	Linked	Untied	Affiliated	Linked
Constant	518.518***	219.449**	−32.583*	11.580	21.003
	(159.865)	(90.533)	(15.129)	(6.094)	(16.390)
First year star publishes in GenBank	−.264***	−.112**	.017**	−.006	−.011
	(.081)	(.046)	(.008)	(.003)	(.008)
Total citations to prior articles in GenBank	.022***	.020***	−.003***	.000	.003***
	(.005)	(.005)	(.000)	(.000)	(.000)
Total number of prior articles in GenBank	−.102*	−.011	.002	−.003	.001
	(.053)	(.028)	(.004)	(.002)	(.004)
University top quality dummy	1.080	.992	−.140	.009	.131
	(1.102)	(.546)	(.083)	(.035)	(.089)
Number of human genetic sequences	.165	.330***	−.045**	−.004	.049**
	(.284)	(.101)	(.017)	(.009)	(.018)
New biotech enterprises in region	.046*	.021*	−.003	.001	.002
	(.022)	(.011)	(.002)	(.001)	(.002)
Top-quality universities in region	−.748	−.664*	.094*	−.007	−.087
	(.537)	(.277)	(.046)	(.017)	(.049)
Proportion coauthors from different institutions	1.146	1.863*	−.257*	−.013	.271
	(1.819)	(.776)	(.131)	(.060)	(.141)
Proportionate change in annual citation rate of other stars in same region while tied to firm	.919	.228	−.036	.025	.011
	(.598)	(.311)	(.045)	(.020)	(.050)
Proportionate change in annual citation rate of stars in different regions while tied to firm	−1.093	−.187	.032	−.032	.000
	(1.142)	(.331)	(.048)	(.037)	(.057)
Log likelihood	−136.523				
Restricted log likelihood	−216.000				

* $p < .05$.
** $p < .01$.
*** $p < .001$.

the star publishes in GenBank. Linked stars show a very similar pattern of significant variables to the overall results reported in table 4 and discussed above, except that the proportionate change in citation rate for tied local stars loses its significance.

The last three columns of table 5 report the corresponding marginal effects of each of the choices (untied, affiliated, or linked). These numbers tell us the effect of each variable on each choice relative to all the other choices and not to just being untied. These results suggest that there is an ordering of labor arrangements for these top scientists: those with the fewest citations, the least work on human sequences, the fewest coauthors at other institutions, and the least experience are most likely to be exclusively in academe. Their opposites—the elite of this elite group—can simultaneously work in the university and a firm. Full-time work in firms seems to have an intermediate position—attractive to stars who are good enough to be attractive to firms but not quite good enough to write their own ticket.

VI. Summary and Implications

We have shown across a series of analyses that star scientists move more quickly from academe to commercial involvement if they have higher-quality intellectual human capital (here measured in terms of number of citations to genetic-sequence discovery articles) and if that capital is more relevant to firms commercializing biotechnology (i.e., amplified by discovery of human genetic sequences). We have also demonstrated strong effects of the opportunities available in the star's 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 the 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 the productivity of other stars in their region who have previously moved to firms; when these other stars' annual citation rates increase, 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 them but with the smaller number of 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 1996*a*). The value of the information is a key determinant. We examine value in two principal ways. In this article, we operationalize an important new measure of the degree of tacit knowledge, resting on a coauthorship measure we developed to examine labor mobility of star scientists to firms: even as scarcity of the knowledge may be declining, tacitness may not be—or at least not as fast.[17] 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% 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 series of related papers on the effects of stars on the success of new biotechnology enterprises, and we 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 United States has strongly moved to firms and has done so in very concentrated, localized areas, as illustrated in figure 2.

We 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

[17] We build here on a novel empirical measure we developed in earlier research: "copublishing," that is, 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 Sec. III), effects of organizational boundaries as information envelopes slowing diffusion of scientific knowledge, and size and geography of scientific networks used by firms (Zucker, Brewer, Oliver, and Liebeskind 1993; Zucker, Darby, and Armstrong 1994, 1998; Liebeskind, Oliver, Zucker, and Brewer 1996; Zucker and Darby 1996*a,* 1996*b*; Zucker et al. 1996). The validity of our copublishing 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 et al. [1993]; Zucker and Darby [1997]).

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."[18] Many universities do not place any restrictions on a professor's outside employment, while universities with rules typically allow 40% (e.g., 1 week day plus 1 weekend day) of faculty time to be spent on outside consulting. One study of academics found that 20%–25% of faculty income was earned outside the university (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 1999).

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 have 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.

[18] Most research on part-time work 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 does not necessarily mean marginal, either in terms of the amount of effort or 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 1996*b*). 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 nonacademic arenas (e.g., increasing fund raising success among entrepreneurs).

Appendix A

Data Appendix

A. Star Scientists Database

Given the fundamental role of recombinant DNA (genetic engineering) 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 genetic sequence discoveries up to 1990 (see Zucker, Darby, and Brewer 1994). Worldwide, 327 leading researchers (the "stars") were identified on the basis, up to 1990, of the number of genetic sequence discoveries and articles reporting them for which they were an author. 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 because available machine-readable databases give only the location of the first (or corresponding) author who, given the authorship conventions of the field, is rarely a star scientist.

B. Identifying Employment Relationships with Firms: Dependent Variable

In this article, stars may be affiliated, that is, working for firms (measured as listing the firm as affiliation on the article or genetic sequence patent), or stars may be linked, that is, working with firms while maintaining their primary affiliation with a university or research institute (measured as coauthoring 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 stars 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). When a star scientist becomes for the first time affiliated or linked, we will call him a mover from a university or research institute to a firm.

C. Other Databases Used

Individual Characteristics of the Stars

Using the names of the 327 stars, detailed bibliographic information was collected from five major sources: *American Men and Women of Science* (1971–94); *Biotechnology Research Directory 4000 Faculty Profiles* (North Carolina Biotechnology Center 1991); *Who's Who of Nobel Prize Winners, 1901–90* (1991); National Academy of Sciences, *Organization*

and Members 1993 (1993); and the *1990 Directory* of the American Association for the Advancement of Science (1990). We also filled in some missing data for particular stars from *Who's Who of British Scientists, 1980/81* (1980), *Who's Who in Science in Europe* (1991), and *Who's Who in Biotechnology* (1986). Finally, annual salary data were collected for associate and full professors from most U.S. research universities from the *American Association of University Professors Bulletin* (1979–93) and *Academe* (1979–93), and for the handful of institute affiliations from telephone interviews with the respective institutes.

Quality Characteristics of the Stars

To measure the quality of the star scientists we use four measures. The first one, number of prior articles in the GenBank, was obtained by counting how many articles the star scientists have in the GenBank prior to each discrete-period observation (GenBank 1990, 1994). Second, we have collected data for 1982, 1987, and 1992, on the total number of citations to each of our 4,061 published articles listed in the Institute for Scientific Information's *Science Citation Index* (1982, 1987, 1992). These citation counts are linked to the article and authorship data set by the article ID number, and therefore we were able to get a count of citations of all articles written prior to each observation. Third, we use the *Who's Who of Nobel Prize Winners* (1991) to identify which of our star scientists had received a Nobel Prize. Finally, from university directories and from bibliographic directories, we identified if the scientists were tenured or not in their respective institutions.

Indicator of Commercial Potential

The GenBank, in addition to the article information, also classifies into 13 different types each of the genetic sequences discoveries reported in each article. This article uses the total number of human-related genetic sequences (sequence types 1 and 4) as a proxy measure of commercial potential of the discoveries of the star scientist.

Characteristics of Universities and Research Institutes

Our university data set consists of all U.S. institutions listed as granting the Ph.D. degree in any field in the Higher Education General Information Survey (HEGIS), Institutional Characteristics, 1983–84 (U.S. Department of Education, National Center for Education Statistics, 1985). Each university is assigned an institutional ID number and a university flag and is located by zip code based on the HEGIS address file.

Additional information was collected for those universities granting the Ph.D. degree in biochemistry, cellular/molecular biology, or microbiology, which we define as "biotech-relevant" fields. All of the following additional variables are based on data in the National Academy of Sciences study by Jones, Lindzey, and Coggeshall (1982).

We define university quality level based on the scholarly quality rating in the reputational survey in Jones et al. (1982). Reputational ratings were

based on responses from approximately 15% of the faculty in the fields studied. Since we were interested in identifying the very best programs, we considered only the highest rated of the biochemistry, cellular/molecular biology, or microbiology programs offered by a particular university. The number of universities in a region with one or more most highly rated programs (rated above 4) is our variable top-quality university.

In addition, for those U.S. research institutions and hospitals listed as affiliations in the article data set, we assigned an institutional ID number and an institute/hospital flag, and we obtained an address, including a zip code, as required for geocoding. We collected, by phone survey, salary indices for those research institutions and hospitals.

Biotechnology Firm Data Set

The starting point for our firm data set covered the industry as of April 1990 and was purchased from the North Carolina Biotechnology Center (NCBC; 1992), a private firm that tracks the industry. This data set identified 1,075 firms, some of which were duplicates or foreign and others of which had died or merged. Further, there were a significant number of missing firms that had died earlier. For these reasons, an intensive effort was made to supplement the NCBC data with information from *Bioscan* (1989–93) and an industry data set provided by a firm in the industry, which was also the ancestor of the *Bioscan* data set (Cetus Corporation 1988). Each of the firms was assigned an institutional ID number and an enterprise flag.

We combined these three sources to identify 751 distinct U.S. firms for which we could determine a zip code and a date of founding (or entry into biotechnology for subunits of preexisting firms). Based on these data, we have developed a continuous series on the number of active new biotech enterprises by year and region.

Indicators of Size of Social Networks

Since our sample is selected based on star scientists and sequence-reporting articles and since collaborators appear in our sample only if one of our star scientists is an author, we are able to define the collaboration as all possible pairs of coauthors that include at least one star. This measure allowed us to identify the proportion of coauthors a star has from different institutions. Since we have the institutional affiliation for each author in each of his or her genetic-sequence-discovery articles, we can identify when a star moves from one university or research institute to another prior to their movement to a firm.

Appendix B

Econometric Modeling

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; Ryu 1994).

First, divide the interval between the beginning of the measurement period, $T = 0$, to the time of the measurement, $T = t_i$, into j exhaustive nonoverlapping intervals, $a_0 < a_1 < \ldots < a_{j-1} < a_j$. 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)

$$A_j = [a_{j-1}, a_j), j = 1, \ldots, ra_0 = 0, a_r = \infty, \tag{B1}$$

and the failure times in A_j are recoded as $\mathbf{t_i}$.

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_j equal the probability of surviving the jth interval given that an individual has survived up to the $(j-1)$th interval (conditional probability).

Therefore,

$$\alpha_j = \Pr(T > a_j \setminus T > a_{j-1}) = \frac{\Pr(T > a_j)}{\Pr(T > a_{j-1})} = \frac{S(a_j)}{S(a_{j-1})}. \tag{B2}$$

However,

$$S(\alpha_j) = e^{\int_0^{a_j} h(u)du.} \tag{B3}$$

Therefore,

$$\alpha_j = e^{\int_{a_{j-1}}^{a_j} h(u)du.} \tag{B4}$$

Then the probability of observing a failure at time t_i on an individual with regression vector $\mathbf{x}$ is

$$[1 - \alpha_j^{\exp(x,\beta)}] \prod_{j=1}^{j-1} \alpha_j^{\exp(x,\beta)}, \tag{B5}$$

where the probability of surviving to the beginning of A_j is

$$P(t_i, x) = \prod_{j=1}^{j-1} \alpha_j^{\exp(x,\beta)}. \tag{B6}$$

Given that the group duration information is a sequence of binary observations, we can apply a logistic function for computational simplicity; therefore, we will have

$$\alpha_j = e^{-exp(x\beta+\gamma j)} \text{ where } \gamma_j = \int_{Aj} h_o(t)dt, \tag{B7}$$

and $h_o(t)$ is the baseline hazard at period t.

This model is easy to estimate using either an ordered-probit or or-

dered-logit approach. In some exploratory research of Han and Hausman (1990), the estimates of the ordered-logit and ordered-probit models were very similar except in the extreme left tail. Accordingly, we select the ordered-logit model because of the simplicity of its calculation.

References

Academe. Vols. 65–79. Washington, DC: American Association of University Professors, 1979–93.

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 of University Professors Bulletin. Vols. 55–64. Washington, DC: American Association of University Professors, December 1969–December 1978.

American Men and Women of Science, 1st–19th eds. New York: Jacques Cattell/R. R. Bowker, 1971–94.

Bioscan. Vols. 2–8. Phoenix: Oryx, 1988–94.

Cetus Corporation. "Biotechnology Company Data Base." Printout of predecessor source for *Bioscan.* Emeryville, CA: Cetus Corp., 1988.

Conlin, Michael. "Empirical Test of a Separating Equilibrium in National Football League Contract Negotiations." *Rand Journal of Economics* 30, no. 2 (Summer 1999): 289–304.

Darby, Michael R., and Zucker, Lynne G. "Local Academic Science Driving Organizational Change: The Adoption of Biotechnology by Japanese Firms." Working Paper no. 7248. Cambridge, MA: National Bureau of Economic Research, 1999.

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.

Han, Aaron, and Hausman, Jerry A. "Flexible Parametric Estimation of Duration and Competing Risk Models." *Journal of Applied Econometrics* 5, no. 1 (January–March 1990): 1–28.

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.

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

Liebeskind, Julia Porter; Oliver, Amalya Lumerman; Zucker, Lynne G.; and Brewer, Marilyn B. "Social Networks, Learning, and Flexibility: Sourcing Scientific Knowledge in New Biotechnology Firms." *Organization Science* 7, no. 4 (July–August 1996): 428–43.

National Academy of Sciences, National Academy of Engineering, In-

stitute of Medicine, and National Research Council. *Organization and Members 1993.* Washington, DC: National Academy Press, 1993.
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.
North Carolina Biotechnology Center, Biotechnology Information Division. *Biotechnology Research Directory: 4000 Faculty Profiles.* Washington, DC: Bureau of National Affairs, 1991.
Prentice, Ross L., and Gloeckler, Lynn A. "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data." *Biometrics* 34 (March 1978): 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* 89 (December 1994): 1386–97.
Spence, Michael. "Job Market Signalling." *Quarterly Journal of Economics* 87, no. 3 (August 1973): 355–75.
———. *Market Signalling: Informational Transfer in Hiring and Related Screening Processes.* Cambridge, MA: Harvard University Press, 1974.
Stigler, George J. *Employment and Compensation in Education.* Occasional Paper no. 33. New York: 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* 8, pt. A (1986): 199–233.
Topel, Robert H., and Ward, Michael P. "Job Mobility and the Careers of Young Men." *Quarterly Journal of Economics* 107, no. 2 (May 1992): 439–79.
Torero, Maximo. "Essays on Diffusion of Technical Change." Ph.D. dissertation, University of California, Los Angeles, Department of Economics, 1998.
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, DC: 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.* U.S. Department of Education, National Center for Education Statistics. 1985. Machine readable data base, ICPSR 8291. Ann Arbor, MI: Inter-University Consortium for Political and Social Research, circa 1985.
Who's Who in Biotechnology, 5th ed. (North American ed.). Woodbridge, CT: Research Publications, 1986.
Who's Who in Science in Europe, 7th ed. Harlowe, Essex, United Kingdom: Longman Group, U.K., 1991.
Who's Who of British Scientists, 1980-81. Dorking, Surrey, United Kingdom: Simon Books.
Who's Who of Nobel Prize Winners. 1901–90. 2d ed. Phoenix: Oryx, 1991.

Zucker, Lynne G.; Brewer, Marilynn B.; Oliver, Amalya; and Liebeskind, Julia. "Basic Science as Intellectual Capital in Firms: Information Dilemmas in rDNA Biotechnology Research." Working paper. Los Angeles: University of California, Los Angeles, Institute for Social Science Research, 1993.

Zucker, Lynne G., and Darby, Michael. "Costly Information: Firm Transformation, Exit, or Persistent Failure." *American Behavioral Scientist* 39, no. 8 (August 1996): 959–74. (*a*)

———. "Star Scientists and Institutional Transformation: Patterns of Invention and Innovation in the Formation of the Biotechnology Industry." *Proceedings of the National Academy of Sciences* 93, no. 23 (November 1996): 12709–16. (*b*)

———. "Present at the Revolution: Transformation of Technical Identity for a Large Incumbent Pharmaceutical Firm after the Biotechnological Breakthrough." *Research Policy* 26, nos. 4–5 (December 1997): 429–46.

———. "Capturing Technological Opportunity via Japan's Star Scientists: Evidence from Japanese Firms' Biotech Patents and Products." *Journal of Technology Transfer* 26, nos. 1–2 (January 2001): 37–58.

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

———. "Geographically Localized Knowledge: Spillovers or Markets?" *Economic Inquiry* 36, no. 1 (January 1998): 65–86.

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

———. "Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises." *American Economic Review* 88, no. 1 (March 1998): 290–306.

Zucker, Lynne G.; Darby, Michael R.; Brewer, Marilynn B.; and Peng, Yusheng. "Collaboration Structure and Information Dilemmas in Biotechnology: Organizational Boundaries as Trust Production." In *Trust in Organizations,* edited by Roderick M. Kramer and Tom R. Tyler, pp. 90–113. Thousand Oaks, CA: Sage, 1996.