

NBER WORKING PAPER SERIES

MOVEMENT OF STAR SCIENTISTS AND ENGINEERS AND HIGH-TECH FIRM
ENTRY

Lynne G. Zucker
Michael R. Darby

Working Paper 12172
<http://www.nber.org/papers/w12172>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2006

This research has been supported by grants from the National Science Foundation (grants SES-0830983 and SES-1158907) and the Ewing Marion Kauffman Foundation (grants 2008-0028 and 2008-0031). We are indebted to our research team members Emre Uyar, Jason Fong, Robert Liu, Jade Yu-Chi Lo, Tim Loon, Hongyan Ma, Amarita Natt, and Yong Yang. The paper was improved by insightful comments from several discussants and anonymous referees. Certain data included herein are derived from the Science Citation Index Expanded, Social Sciences Citation Index, Arts & Humanities Citation Index, High Impact Papers, and ISI Highly Cited of the Institute for Scientific Information®, Inc. (ISI®), Philadelphia, Pennsylvania, USA: © Copyright Institute for Scientific Information®, Inc. 2005, 2006. All rights reserved. Certain data included herein are derived from the Connecting Outcome Measures of Entrepreneurship, Technology, and Science (COMETS) database and the associated COMETSbeta and COMETSandSTARS databases © Lynne G. Zucker and Michael R. Darby. All rights reserved. This paper 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 or the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2006 by Lynne G. Zucker and Michael R. Darby. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Movement of Star Scientists and Engineers and High-Tech Firm Entry
Lynne G. Zucker and Michael R. Darby
NBER Working Paper No. 12172
April 2006, Revised June 2014
JEL No. O31,J61,033,J44,M13

ABSTRACT

This paper analyzes the effects of top nanoscale scientists on industry entry in the comparative context of 5 major areas of science and technology, extending the concept of star scientist to all areas of science and technology. The results for nanotechnology are replicated using the comprehensive list of firms from NanoBank.org that provide an alternative industry entry measure not available for other high-tech fields. We follow careers 1981-2004 for 5,401 stars as identified in ISIHighlyCited.comSM, using their publication history to locate them each year. The number of stars in a U.S. region or in one of the top-25 science and technology countries generally has a consistently significant and quantitatively large positive effect on the probability of firm entry in the same area of science and technology. Other measures of academic knowledge stocks have weaker and less consistent effects. Thus the stars themselves rather than their potentially disembodied discoveries play a key role in the formation or transformation of high-tech industries. We identify separate economic geography effects in poisson regressions for the 179 BEA-defined U.S. regions, but not for the 25 countries analysis. Stars become more concentrated over time, moving from areas with relatively few peers to those with many in their discipline. A counter-flow operating on the U.S. versus the other 24 countries is the tendency of foreign-born American stars to return to their homeland when it develops sufficient strength in their area of science and technology. In contrast high impact articles and university articles and patents all tend to diffuse, becoming more equally distributed over time.

Lynne G. Zucker
Departments of Sociology & Public Policy
UCLA
Box 951551
Los Angeles, CA 90095-1551
and NBER
zucker@ucla.edu

Michael R. Darby
John E. Anderson Graduate School of Management
University of California, Los Angeles
110 Westwood Plaza, Box 951481
Los Angeles, CA 90095-1481
and NBER
michael.r.darby@anderson.ucla.edu

Movement of Star Scientists and Engineers and High-Tech Firm Entry*

Lynne G. Zucker and Michael R. Darby

Harberger (1998) in his A.E.A. presidential address reports that nearly all total factor productivity (TFP) growth at any given time is concentrated in 10-20 percent of the firms in only 10-20 percent of the industries. That is, technological progress is not spread like manna from heaven smoothly across firms and industries a la Solow but concentrated in a few thousand or even hundreds of firms making very rapid, metamorphic progress (Darby and Zucker 2003). In a series of articles, we have provided evidence for rapidly advancing science and technology areas such as biotechnology (where hedonic TFP measures are usually not feasible due to radical product innovation), that individual “star” scientists making major discoveries play an important role in determining where and when new or previously existing firms begin using the new technologies and which firms are most successful (Zucker and Darby 1996, 2001, 2006a, 2006b; Zucker Darby and Brewer 1998; Zucker, Darby, and Armstrong 1998, 2002; Darby and Zucker 2001, 2005). These results suggests that growth beyond that due to standard measures of investment in human and physical capital may be due to the purposive actions of a relatively few people embodying rare intelligence, energy, and drive. In the words of William James (1911): “Geniuses are ferments; and when they come together, as they have done in certain lands at certain times, the whole population seems to share in the higher energy which they awaken.”

Nanotechnology, driven by breakthroughs in nanoscale science and engineering, is – in our view – well along in the early metamorphic progress phase of industrial evolution. Centered on utilizing properties occurring at the atomic and sub-atomic scale, nanotechnology has application to a particularly wide range of industries. In this paper we analyze the effects of top nanoscale scientists on industry entry in the comparative context of 5 major areas of science and technology (S&T areas) which have been used by us and others for tracing the process of knowledge creation and its flow to new commercial applications: biology/chemistry/medicine; computing, information and communications; semiconductors; other science; and other engineering. We do so by extracting specific nano articles and patents from the S&T areas where they would otherwise be included so that the processes surrounding these new technologies are not swamped by the broader contexts in which they occur. In developing NanoBank.org,

This research has been supported by grants from the National Science Foundation (grants SES-0830983 and SES-1158907) and the Ewing Marion Kauffman Foundation (grants 2008-0028 and 2008-0031). We are indebted to our research team members Emre Uyar, Jason Fong, Robert Liu, Jade Yu-Chi Lo, Tim Loon, Hongyan Ma, Amarita Natt, and Yong Yang. The paper was improved by insightful comments from several discussants and anonymous referees. Certain data included herein are derived from the *Science Citation Index Expanded*, *Social Sciences Citation Index*, *Arts & Humanities Citation Index*, *High Impact Papers*, and *ISI Highly Cited* of the Institute for Scientific Information®, Inc. (ISI®), Philadelphia, Pennsylvania, USA: © Copyright Institute for Scientific Information®, Inc. 2005, 2006. All rights reserved. Certain data included herein are derived from the Connecting Outcome Measures of Entrepreneurship, Technology, and Science ([COMETS](#)) database and the associated [COMETSbeta](#) and COMETSandSTARS databases © Lynne G. Zucker and Michael R. Darby. All rights reserved. This paper 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 or the National Bureau of Economic Research.

we built a very comprehensive list of firms that have entered into nanotechnology with entry dates based on firm creation for new nano-focused firms and stated entry or first evidence of activity for pre-existing firms previously using other technologies only (see Section II.1.3 for details). Substitution of these alternative entry dates did not materially alter the results based on entry inferred from first publication and/or patenting activity in the application to all high-tech firms, confirming the results for nanotechnology and validating the method for the other S&T areas.

In addition to treating nanotechnology as a distinct S&T area, this paper expands on our previous work in four other major ways: First, we focus on the question of whether or not the star scientist has an independent role separate from the discoveries he or she makes. Second, we expand our coverage of science base to all areas covered by the Science Citation Index and our coverage of firms expands to all those for which U.S. patents have been assigned on issue or with which authors of articles are affiliated in the Science Citation Index, 1981-2004. Third, we provide evidence on 25 countries which do most of the world's scientific research and commercial innovation instead of concentrating on one or two countries.¹ Fourth, we validate ISIHighlyCitedSM as an empirically useful means of identifying scientific stars.² While we focus on high-tech industries and star scientists, we conjecture that they are merely the most easily identifiable representatives of a larger class of creative geniuses – star innovators – who normally drive metamorphic progress wherever it occurs, be it Genentech, Intel, McDonalds, or WalMart. The scientific stars have proven their drive, energy, and brilliance in their published scientific work; we find empirically here that they indeed often choose to apply these qualities in commercialization of their discoveries – discoveries often characterized initially by extensive tacit knowledge and even controversy over reproducibility which gives these discoverers a period of natural excludability of potential competitors.

We show here that star scientists do have a statistically and substantially significant impact on firm entry even after accounting for such measures of local knowledge stock as high-impact (highly cited) articles except those with firm authors, all publications by university authors except those with firm co-authors, and U.S. patents assigned at issue to universities. Since the embodied knowledge, insight, taste, and energy of the stars plays a role separate from their potentially disembodied discoveries, this evidence strengthens the case for the importance of these extraordinary individuals for the economic development of regions and nations. These statements apply equally to nanotechnology and the five S&T areas that have been previously considered. Furthermore, while the distribution of knowledge stock measures shows a slight tendency – if anything – toward more evenness over time, the distribution of star scientists in many of the cases becomes significantly more concentrated in the leading centers over time. Counts of non-university patents show an inconsistent pattern of dependence on the local presence of star scientists.

Section I lays out the analytical approach and hypotheses to be examined. We discuss the data set and estimation methodology in Section II. Our empirical results are reported in the next section. We summarize the results and draw our conclusions in Section IV.

¹ The 25 countries are: Australia, Austria, Belgium, Brazil, Canada, China, Denmark, Finland, France, Israel, India, Italy, Japan, Germany, the Netherlands, Norway, Poland, South Korea, Spain, Sweden, Switzerland, Taiwan, the United Kingdom, the United States, and the USSR & Russia counted as the same country.

² Was renamed [Highly Cited Research](#) and discontinued as a stand-alone resource by Thomson Reuters in 2011. The combination of their [ResearcherID](#) and [Essential Science Indicators®](#) provide is their suggested replacement.

I. Analytical Approach and Hypotheses

As in our prior work we assume that the probability λ that a firm will begin to use a given type of new technology through birth or change in focus in a particular country or region is small for an arbitrarily short period of time, so that entry occurs randomly over time in accordance with the poisson or related processes. While the poisson process is frequently used to characterize the distribution of failures – such as light bulbs burning out – it is useful for characterizing countable events of a positive nature as well. The probability λ is assumed to vary across regions and years according to $\log \lambda = x\beta$ where x is a row vector of the explanatory variables and β is a parameter vector to be estimated.³ In our previous work the significant explanatory variables have been primarily counts of star scientists working in the region, other measures of the knowledge base in the region and measures of the economic geography (employment and average wage per job as a proxy for education level of the local labor force).

Zucker and Darby (1996) and Zucker, Darby, and Brewer (1998) introduced the concept of biotechnology stars based upon productivity measured by the number of articles written through 1990 which reported a genetic-sequence discovery. These biotech stars were the top 0.7 percent of all the authors of such discoveries and accounted for 17.3 percent of the total number of discoveries. Direct involvement of these stars with specific firms proved to be a major factor in determining which firms were ultimately major winners in biotechnology (Zucker, Darby, and Armstrong 1998, 2002; Zucker and Darby 2001). In this paper we operationalize the concept of star scientist and engineers as those 5,401 very productive scientists and engineers selected by ISI and profiled in ISIHighlyCited.comSM across the range of science and engineering topics covered in the Science Citation Index. By including the number of these stars active in a region and year as an explanatory variable, we specifically investigate whether these extraordinary individuals play an independent role in promoting the entry of firms into their area of science and technology when their discoveries are accounted for in measures of the local knowledge stocks of high impact articles, all university articles, and university patenting. Based on our biotechnology work, we hypothesize that they do have a separate positive impact on λ , but acknowledge controversy as to how far beyond biotechnology and other high-science-driven areas that effect will be present.

We hypothesize that a very similar process explains commercial development in the form of non-university patenting, although the corresponding λ would surely be of larger magnitude per unit of time. Since stars mostly work in the university (even stars affiliated with firms most often have a primary appointment with a university), we hypothesize that their effect on non-university patenting is likely to be weaker but still present. Moreover, we believe that it is interesting to quantify the effects of the academic knowledge stocks more generally on regional patenting, given the gradual shift in the search for the basis of high-tech-industry innovations toward a “knowledge economy” explanation [recent examples include Baba, Shichijo, and Sedita (2009), Elfenbein, Hamilton, and Zenger (2010), and Jiang, Tan, and Thursby (2011)].

³ If $\lambda = X\beta + \varepsilon$ (i.e., has a disturbance term ε) and if so the distribution of ε affects the estimation methodology used as discussed in Section II.B below.

II. Empirical Methodology

Our empirical analysis focuses on entry of firms and non-university patenting over time and by U.S. regions or by countries. The data bases for this study have been substantially enlarged in both size and coverage from those used in any other study of which we are aware. Section II.1 describes the data used in the empirical analysis. Section II.2 summarizes the standard estimation methodology.

II.1. The Data

The primary source databases for the analysis are the complete, continuously updated and parsed U.S. Patent database of the Zucker-Darby Knowledge, Innovation, and Growth Project and the *Science Citation Index Expanded*, *Social Sciences Citation Index*, *Arts & Humanities Citation Index*, *High Impact Papers*, and *ISI Highly Cited* of the Institute for Scientific Information[®], Inc. (ISI[®], 2005, 2006). Our patent data cover the 3,891,720 U.S. patents granted by USPTO from 1976 to 2005. The ISI database contains more than 24,250,000 records from over 8700 peer-reviewed scientific journals. All of this data is integrated into the Connecting Outcome Measures of Entrepreneurship, Technology, and Science ([COMETS](#)) database and the associated [COMETSbeta](#) and COMETSandSTARS databases constructed and still being extended in scope and years as a research community resource by the authors and our team with support from the National Science Foundation and the Ewing Marion Kauffman Foundation (Zucker, Darby and Fong 2014). Other sources are noted where relevant.

Although our data cover all countries, computational considerations led us to limit our analysis to the 25 top science and engineering countries defined as all countries that accounted for at least 0.5% of all ISI articles or at least 0.1% of all U.S. patents granted, 1976-2004, or both, with articles prorated by authors' addresses and patents prorated by inventors' addresses. These "top-25 science and technology (S&T) countries" are: Australia, Austria, Belgium, Brazil, Canada, China, Denmark, Finland, France, Israel, India, Italy, Japan, Germany, the Netherlands, Norway, Poland, South Korea, Spain, Sweden, Switzerland, Taiwan, the United Kingdom, the United States, and (counted as the same country) the USSR and Russia. These 25 countries account for 92.8% of all ISI articles and 99.2% of U.S. patents.

These data are used to create two analysis data sets containing data from 1981 through 2004 for each of the 179 U.S. regions and also for each of 25 top science and engineering countries (including the U.S.). These longitudinal (panel) data sets consist of 179 regions x 24 years = 4,296 observations and 25 countries x 24 years = 600 observations, respectively. Some analysis is done with the U.S. deleted from the country data set, leaving 576 observations. One year of observations is lost in regressions run with location-specific effects and all independent variables lagged by one year. The U.S. Bureau of Economic Analysis defines the 179 regions as functional economic areas such that each U.S. county is assigned to a region which includes the major metropolitan center for which commuting, shopping, and newspaper readership predominates (Johnson and Kort 2004). The variables contained in the data sets and their summary statistics are listed in Table 1. Their construction is described immediately below.

The variables in each data set are categorized into six science and technology areas: Biology, Chemistry & Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Each organization name appearing in the assignee-at-issue field of a patent is categorized as a firm, a

university, or put in a miscellaneous other category which includes governmental organizations and research institutes. If a patent has n assignees, where $n \geq 1$, each assignee is assigned $1/n$ for that patent. Organization names appearing in research address or reprint address fields of ISI articles are similarly grouped. Articles that have reported affiliations in the research address field are assigned to those organizations, with each organization getting $1/m$ credit for that article, where m is the number of research addresses reported. Those without a research address are assigned to the organization reported in the reprint address, provided that one exists.

Table 1. Summary Statistics for Variables Used in Main Empirical Analysis, 1981-2004

Variables	U.S. Regions					Top-25 Sci. & Tech Countries					Top-24 non-U.S S&T Countries				
	N	Mean	S.D.	Min	Max	N	Mean	S.D.	Min	Max	N	Mean	S.D.	Min	Max
Entry of Firms into:															
Biology/Chemistry/Medicine	4296	7.2	15.9	0	165	600	45.8	149.4	0	1067	576	21.8	24.9	0	125
Computing/Information Technology	4296	3.2	9.8	0	203	600	20.7	80.5	0	738	576	7.3	10.3	0	56
Nanotechnology	4296	1.9	6.2	0	87	600	12.5	53.5	0	483	576	5.4	9.7	0	71
Semiconductors	4296	2.7	7.3	0	90	600	24.5	80.7	0	657	576	9.3	12.2	0	73
Other Sciences	4296	3.6	8.3	0	104	600	24.5	89.0	0	609	576	9.5	13.5	0	66
Other Engineering	4296	12.8	24.7	0	248	600	72.8	287.0	0	2107	576	29.6	42.3	0	207
Non-university patenting:															
Biology/Chemistry/Medicine	4296	30.6	147.7	0	1986.4	600	729.8	2081.8	0	14783.0	576	344.7	663.8	0	4079.3
Computing/Information Technology	4296	20.4	165.2	0	4130.3	600	446.5	1709.5	0	14101.2	576	214.9	828.1	0	9188.5
Nanotechnology	4296	9.8	73.7	0	2089.8	600	247.8	966.5	0	8245.5	576	141.2	625.5	0	7967.5
Semiconductors	4296	5.4	46.7	0	1126.7	600	162.1	544.0	0	4453.6	576	96.1	347.4	0	3039.5
Other Sciences	4296	11.1	50.4	0	946.4	600	311.7	881.2	0	6453.9	576	164.3	446.2	0	4102.3
Other Engineering	4296	71.9	262.1	0	3422.7	600	2286.5	6261.7	0	41986.1	576	1241.1	3133.4	0	23161.3
Star Scientists & Engineers Active															
Biology/Chemistry/Medicine	4296	11.3	33.8	0	256	600	126.2	395.8	0	2385	576	46.8	59.3	0	300
Computing/Information Technology	4296	2.5	9.2	0	96	600	25.4	87.3	0	516	576	7.9	9.4	0	41
Nanotechnology	4296	0.3	1.6	0	22	600	3.9	14.6	0	116	576	1.7	3.4	0	20
Semiconductors	4296	2.0	7.0	0	79	600	20.7	70.5	0	472	576	6.6	7.8	0	41
Other Sciences	4296	1.6	6.3	0	62	600	17.0	57.7	0	361	576	5.6	8.5	0	42
Other Engineering	4296	0.8	3.2	0	35	600	8.6	29.3	0	200	576	2.7	4.3	0	25
High Impact Articles Knowledge Stock															
Biology/Chemistry/Medicine	4296	13.1	55.0	0	518.9	600	277.3	834.2	0	4,911.6	576	115.0	156.2	0	870.3
Computing/Information Technology	4296	2.5	11.5	0	123.5	600	58.4	162.4	0	1,194.3	576	27.0	34.6	0	248.2
Nanotechnology	4296	0.5	2.9	0	65.9	600	15.1	53.7	0	513.1	576	7.6	14.4	0	96.9
Semiconductors	4296	2.0	9.7	0	115.4	600	53.4	146.1	0	907.9	576	25.3	32.0	0	137.1
Other Sciences	4296	5.0	25.1	0	246.0	600	96.7	324.2	0	1,994.1	576	34.7	52.6	0	269.4
Other Engineering	4296	1.3	7.3	0	75.2	600	30.4	91.2	0	528.3	576	12.6	18.7	0	103.4
University Articles Knowledge Stock															
Biology/Chemistry/Medicine	4296	1485.0	5068.2	0	61,361.2	600	46272.6	108860.9	64.0	729,673.6	576	26210.2	31255.5	64.0	159,973.2
Computing/Information Technology	4296	90.9	317.1	0	3,803.9	600	3501.5	7268.2	6.0	49,950.6	576	2207.5	2476.7	6.0	12,449.2
Nanotechnology	4296	20.0	94.2	0	1,509.7	600	953.9	2557.6	0	24,685.3	576	654.0	1338.9	0	9,289.5
Semiconductors	4296	159.7	577.8	0	6,030.3	600	7438.5	13386.2	24.4	81,232.2	576	5198.3	6455.8	24.4	36,102.0
Other Sciences	4296	99.7	374.2	0	4,161.1	600	3245.2	7703.8	3	48,111.6	576	1816.8	2207.4	3	11,874.8
Other Engineering	4296	44.0	158.6	0	1,691.6	600	1534.7	3622.9	1	23,672.5	576	875.5	1166.2	1	7,241.4
University Patents Knowledge Stock															
Biology/Chemistry/Medicine	4296	8.7	40.4	0	564.3	600	134.7	672.2	0	5259.8	576	17.4	37.4	0	262.0
Computing/Information Technology	4296	0.7	3.9	0	66.6	600	11.3	61.0	0	576.3	576	1.3	3.6	0	35.5
Nanotechnology	4296	3.3	21.7	0	415.8	600	53.5	334.6	0	3599.1	576	6.3	18.0	0	237.4
Semiconductors	4296	0.3	2.4	0	44.5	600	5.8	28.5	0	294.8	576	0.8	4.4	0	92.8
Other Sciences	4296	2.2	12.3	0	224.4	600	34.8	163.4	0	1281.0	576	4.7	10.4	0	58.9
Other Engineering	4296	3.4	15.3	0	255.8	600	62.1	291.6	0	2330.0	576	9.5	20.0	0	129.1
Total Employment in Region/Country	4296	0.8	1.4	0.0	12.9	600	47.3	123.3	1.0382	752	576	44.2	124.9	1.0	752.0
Average Wage per Job in Region	4296	26.2	4.1	18.3	51.0	—	—	—	—	—	—	—	—	—	—

Notes: 1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. U.S. regions are the 179 functional economic areas defined by the U.S. Bureau of Economic Analysis (Johnson and Kort 2004).

For purposes of locating observations, each valid U.S. address in these fields is assigned to a county and the corresponding region using the Federal Information Processing Standard (FIPS55) database maintained by the U.S. Geological Survey. (<http://geonames.usgs.gov/fips55.html>). Foreign addresses are grouped based on the country of origin.⁴

II.1.1. Science and Technology Areas

⁴ Countries are defined as in the international standard ISO 3166 – 1 (see ISO Country Info tab in the COMETS codebook at <http://www1.kauffman.org/Comets/Codebook-2.aspx> for details on coding merged countries (e.g., Germany) and split countries (e.g., Russia).

Tushman and Anderson (1986) emphasize the stability in the science and technology base of a given firm so that it is a major and perilous event to enter a new area of technology comparable to birth of a start-up firm with its own science and technology base. Mansfield (1995) focuses on the ties between particular industries and academic disciplines. Darby and Zucker (1999) attempt to capture these insights in a set of seven area clusters which can be used to compare activity in journal articles (Institute for Scientific Information 1981-1997), university doctoral programs (National Research Council 1995), and patents (Zucker and Darby 1999a).

These clusters are used here with two exceptions: First, the humanities and social sciences are dropped for this study because they have little specific applicability to particular high technology industries. Second, we have developed a public digital library NanoBank.org for the emergent, highly interdisciplinary nanotechnologies which utilize the unique properties that occur at the atomic and sub-atomic level (Zucker and Darby 2007, Zucker, Darby and Fong 2011). We subtract those articles and patents identified for NanoBank.org from the area in which they would have been previously classified. Those nanotechnology patents are identified as the union of a standard Boolean search of titles, abstracts, and patent descriptions using nano-specific terms (keywords from official websites), a list of authority documents from programs targeted to nanotechnology, and an iterative probabilistic method which scores words and phrases according to their relative frequency of appearance in a learning set of expert-identified nano-articles and articles and patents generally (Ma, Furner, Zucker, and Darby 2006; Zucker, Darby, Ma, Furner, and Liu 2007).⁵ Zucker and Darby (2008) and Table B8 in the electronic-only Appendix B detail the Web of Science subject category codes, International Patent Classes, and National Research Council doctoral program names corresponding to each of these five science and technology areas other than nanotechnology.

II.1.2. Star Scientists and Engineers

Zucker and Darby 1996 and Zucker, Darby, and Brewer (1998) introduced the concept of biotechnology stars based upon productivity measured by the number of articles written through 1990 which reported a genetic-sequence discovery. That concept has been generalized to other cases of obvious linkage between development in science and in high technology, such as nanotechnology (Darby and Zucker 2005) and stem cells (Zucker and Darby 2006a). We here apply the star scientist concept to all fields of science and engineering and test the hypothesis that locally active star scientists generally exert an independent positive effect on commercial development (here, firm entry and patenting) in related high-technology industries.

Star scientists previously have been empirically identified by field specific definitions which pick out a highly productive group comprising on the order of 0.5-1.0 percent of all scientists authoring any papers in the field of interest. The *ISIHighlyCited.com* website offers a database of the top 250 individual researchers in terms of 20-plus-year-window citation counts in each of 21 subject fields – 19 of which are science and engineering fields included in this study. Because individuals are not dropped when periodic updates occur, our stars include all 5,401 individual scientists and engineers listed by ISI in 2005 when the analysis data set was constructed, 13.7 percent higher than the nominal 250 most highly cited authors in each field. Information for each highly cited author includes (potentially) full listings of publications

⁵ Details are included in the downloadable codebook at <http://www.nanobank.org/downloads.php>.

and links where available to the full bibliographic information indexed in the *ISI Web of Science*[®]. Since data quality depends in part on the cooperation of the highly cited authors, data quality is as variable as our colleagues. Nonetheless, the database seems to offer a comprehensive list of top researchers across the breadth of science and engineering and the research reported here largely validates its usefulness as an empirical tool for identifying star scientists. The proportion of “stars” identified here is somewhat larger than for the 327 biotech stars identified by gene-sequence discoveries and about a quarter as much as the 10,349 (5 percent of the total) life-science “superstars” identified by Azoulay, Graff Zivin, and Wang (2010). We present some evidence in Appendix A favoring a more select subset of the ISIHighlyCited as a star definition.

Altogether we thus identify 5,401 star scientists, one or more of who are credited with authorship of some 520,839 articles that appear in the *ISI Web of Science*[®] database. If we count articles each time a star scientist appears, there are 571,068 article authorships. For a sizeable minority (2,042 or 37.8%) of these stars, exactly 10 ISI articles are listed as their full publication list – apparently representing those listed on an NSF or NIH abbreviated vita or supplied by ISI itself.⁶ On completion of our STAR person-matching project for all ISI article authors and U.S. patent inventors since 1981, we will be able to add to that article count, but we have not done so at this point. The articles are used to identify where the stars are active based on those 299,583 cases (52.5% of the star authorships) where their affiliation is unambiguous because they are the corresponding author, the sole author, or there is only 1 listed corresponding or research address for a journal that reports multiple addresses on other articles in the same year.⁷

We have used these addresses to identify each U.S. region or non-U.S. country in which these star scientists were active 1981-2004. We use each article for which a star’s location can be definitely determined as fixing a location as of that particular year. We then connect these locations to create continuous location histories for each star: We code the stars as active in a region from two years before their first publication there (based on research and publication lags in a 40-star curriculum vitae study) until they move to another location. During transitional phases they are coded as active for up to two years in both locations. Stars who maintain long-term affiliations in multiple countries also are coded as active in each location. For stars with sparse publications in the ISIHighlyCited list, this procedure may induce a greater lag in our recognizing moves than those with numerous publications each year. A referee pointed out that it will also induce a somewhat longer average lag in recording moves for authors who participate disproportionately in multi-institution collaborations. It seems to us that any noise added to the measurement of movements will bias estimated coefficients on star counts toward zero, making it harder to obtain statistically significant results for our central hypothesis. Perfect measurement would be best, but biases against obtaining the hypothesized effects at least provide an acid test.

⁶ Another 3.2% of the stars are listed as having 5-9 ISI-indexed articles, presumably having cited works in press or conference volumes on the NSF/NIH vita. We dropped entirely from the analysis those 30 highly cited authors out of 5,431 for which there are no ISI-indexed articles listed.

⁷ ISI article data do not distinguish which address (normally an organization) goes with which author except for a possible single author designated corresponding author who then matches (at least) to the corresponding address. The cases indicated in the text are those for which the star scientist can be definitively located with an address. The 299,583 authorships corresponded to 276,182 different articles, with the difference (23,401) all accounted for by multiple star authors with a single address.

The author is assigned to each of one or more science and technology areas in those years when that area is reflected in the article keywords (author keywords and ISI Keywords Plus) appearing in all of their publications (whether we have an assured location or not). This potential double-counting of both areas and locations is not a major concern in practice: the average number of stars per year across all countries, areas, and years is only 5,038 or 33.3% more than the 3,780 average number of unique individuals per year. Nonetheless, we believe that it more accurately captures the ability of these extraordinary individuals to catalyze the founding of a firm or entry of an existing firm into a new technology area.

II.1.3. Firm Entry into a New Area of Technology

Darby and Zucker (2005) have shown that the first time in which a firm publishes an article in an area is a good indicator of entry into high-technology industries. We generalize that approach here by counting as entry the first-to-appear of publications with a firm-affiliated author or patents assigned to the firm at the time the patent is granted (assignee at issue). Appearance is dated by article publication or patent application date. For the country data set, a particular firm can enter each time it first publishes or patents in a given technology area in a given country. Firm location corresponds to the address given on the article or the region or non-U.S. country of each inventor on a patent (for over 89% of patents there is only one region or non-U.S. country). For the U.S. data set, a particular firm can enter each time it first publishes or patents in a given technology area in a given region. Thus, entries by a firm in a given area and U.S. region are not counted in the country data set after the first time the firm enters that technology area in any region in the U.S. In practice, the vast majority of firms publish and/or patent in only a single area and single country or region as defined by the addresses of authors or inventors.

It is important to emphasize that having used articles with firm-affiliated authors and patents with firm assignees to define our primary dependent variable, such articles and patents cannot be and are not used in the construction of any of the independent variables for the empirical analysis.

Even if they are also a principal in a firm, star scientists generally give their affiliation as a university or research institute. However, we have found star collaborations with firm employees to be particularly powerful indicators of future firm success. There were 38,100 articles authored by stars either with or as firm employees – 7.3% of all stars’ articles. However, 3,291 stars – 59.6% of all stars – ever had any articles as or with firm employees.

When it comes to establishing property rights to their inventions, however, 28,827 or 63.3% of stars’ patents have a firm listed as or among the assignees at issue out of their 45,542 total U.S. patents.⁸ In contrast, only 5,040 or 11.1% have a university listed as or among the assignees at issue. The number of stars listed as inventors on one or more patents was 2,771 of which 1,976 (71.3% of those with any patents) had one or more patents with a firm listed as or among the assignees at issue. However, 1,354 of the stars with firm-assigned patents also had articles linked to firms, so patents identify only 622 additional stars with some firm involvement.

⁸ Since we cannot rely on a definitive list of articles from ISI, we performed a name match based on exact match of both the last name and the first name and to the extent they are available middle name or initial. This method probably results in an overestimate of the number of star patents with false positives outweighing missed matches due to misspellings.

A total of 3,913 stars (72% of all stars) have been identified to have some relationship to a firm at some point in their career, which may range from arms length patent transfer or collaboration with an employee through consulting, advisory or principal role to employment and/or ownership. Of this total, 84.1% can be identified by co-publication, 50.5% by patent assignment, with a 34.6% overlap identified by both ($84.1\% + 50.5\% - 34.6\% = 100.0\%$).

For one S&T area, nanotechnology, it is possible to compare our measure of firm entry based on publication and patenting with a measure based upon archival data. For NanoBank.org we developed a nearly comprehensive set of all firms involved in nanotechnology. The main sources for firm data are two industry directories – one produced by the leading industrial magazine and the other aimed at potential investors in nanotechnology. These firm names were supplemented by a web search, lists of conference participants, and by publishing and patenting in nanotechnology. For comparison purposes, we have eliminated all firms for which publishing and patenting was the sole source. Entry dates are generally readily available for new entrants (founded after 1984) for which founding date can be taken as the entry date into nanotechnology. For incumbent firms (founded before 1985) it is very difficult to obtain any dates for when the firm began doing nanotechnology. With the exception of 3 incumbent firms which provide entry dates, we assign incumbent entry dates by random draw from the known entry dates in the same U.S. region for a U.S. firm for which there are known regional entry dates, and otherwise drawn from the known entry dates in the same country. This procedure is designed to utilize the where information in incumbent entry while relying (with 3 exceptions) on the entrants for the when information.

II.1.4. Non-University Patenting

The second dependent variable – non-university patenting – measures an aspect of the development of commercial technology by region or country. We use non-university patenting (i.e., patents with no university as any of the assignees at issue) rather than firm patents because the bulk of those patents not assigned to identified firms or universities appear to be commercial in nature although we have not identified the assignee as a particular firm. These patents are geolocated by the inventors' addresses as a more reliable indicator of where the research was done than the assignee address which is often the firm's headquarters. Where there are $n > 1$ inventors, the patent is counted as $1/n$ for the location of each inventor's address.

II.1.5. Knowledge Stocks

We used three separate sources to develop measures of the non-firm science base by region (or country) and year: university articles, university patents, and high impact articles. In constructing these measures we first delete all articles for which a firm is included on any of the reprint and research addresses (articles) or as an assignee at issue since those articles or patents could have been used to define entry and introduce a subtle bias into the analysis. High impact articles are those in the *High Impact Papers* database of the ISI cited above. University articles and patents are those with a university (but no firm) named either on any of the reprint and research addresses or as an assignee at issue. High-impact articles exclude only those with firm authors, but are nearly all also included in the university articles file. Similarly, almost all articles by stars and most of their patents (especially so prior to founding of a firm in which they are a principal) will appear in these input series to the knowledge stocks.

With firm observations excluded, all the articles and patents used in these variables can belong to either universities or “other” organizations (such as national labs or research institutes). In case of multiple organizational affiliations for an observation (due to the presence of multiple research addresses for articles or multiple assignees for patents), each organization is credited with the corresponding fractional amount for that observation. For example, if a patent (article) has listed 1 university and 1 national lab as assignees (research addresses), each is credited $\frac{1}{2}$, and so only 0.5 is added to the university patent (article) total for the corresponding total.

Knowledge stocks are measured as conventional (see Griliches 1990) in the economics of science and technology literature as a perpetual inventory with depreciation rate $\delta = 0.20$:

$$K_{i,t} = I_{i,t} + (1 - \delta)K_{i,t-1} \quad (1)$$

where $K_{i,t}$ is the knowledge stock of type i (denoting science and technology area and region or country) at time t and $I_{i,t}$ is the input series for this knowledge stock – alternatively counts by region/country and science and technology area of (non-firm) university articles, university patents, and high impact articles.

While creating the input series counts for each of these measures, we determine the articles or patents in each science and technology area. (If an article or patent that can be considered belonging to more than one area, each area is credited a fraction.⁹) These science and technology area counts are then allocated to U.S. regions and/or to countries with each research address or assignee address receiving equal credit.¹⁰ For example, if an article had seven authors and listed two British addresses and one French address, Britain would get two thirds of the article’s credit and France one-third since we cannot assign each of the seven authors to any particular research address. After creating the basic counts for each year by area and region or country, we use formula (1) to accumulate them year by year with a 20% depreciation rate to create the knowledge stocks by science and technology area, region or country, and year for each of the two analysis data sets (U.S. regions and top-25 science and technology countries).

II.1.6. Other Variables

The employment and average wage-per-job data for the U.S. regional data set were downloaded from the BEA website (<http://www.bea.gov/bea/regional/reis/>) and the wages were deflated to thousands of 2000 dollars per year using the BEA’s chain-type price index for consumer expenditures. The employment data for the 25-country data set were obtained from IMF (<http://ifs.apdi.net/imf/>) with the exception of Taiwan data which were downloaded from

⁹ Each International Patent Classification code and each ISI Web of Science category code has been associated with one major science and technology area. Since patents can report more than one IPC code (and journals more than one ISI category), we can have observations associated with more than one major science and technology area. If a patent (or article) has n IPC codes (or n ISI categories), each science and technology area is credited with $1/n$ for that patent (article) for each IPC code (ISI category) on it which falls in the area.

¹⁰ If an article has research address listed they include the reprint address and so that is not counted again. If there is only a reprint address, full credit for the article goes to that location. Since this assignment is made by address, the sum of the U.S. regional assignments in principle equals the number assigned to it in the 25-country data set. The only differences arise because of a few U.S. observations which were omitted in the U.S.-only data set because we were unable to assign the partial or garbled address to a region.

http://2k3dmz2.moea.gov.tw/gnweb/english/e_main.aspx?Page=D. Missing observations were interpolated by linear regressions.

II.2. Estimation Method

There is some controversy among practitioners as to the best method to estimate count models with a poisson-like structure. The mean and variance of the poisson distribution both equal the single parameter λ . However, overdispersion (variance $>$ mean) will be observed if there is unobserved heterogeneity across observations. This is frequently dealt with by assuming that the parameter λ is distributed according to

$$\log \lambda = x\beta + \varepsilon \quad (2)$$

where the disturbance term ε is distributed as a gamma distribution. Kennedy (1998, pp. 247-248) notes that this “leads to a negative binomial distribution for the number of occurrences, with mean λ and variance $\lambda + \alpha^{-1}\lambda^2$ where α is the common parameter of the gamma distribution.” Estimation by negative binomial if poisson fails a pretest for overdispersion is a frequent recourse for dealing with potential overdispersion.

We have continually avoided this practice in our own work because (a) pretesting invalidates the nominal statistical significance levels computed for the negative binomial (or any) estimator and (b) if the binomial is inappropriate (i.e., ε is not gamma-distributed) the estimated coefficients will be biased with the negative binomial method while these coefficients are estimated without bias using the poisson method even if the negative binomial method is appropriate. A similar critique applies to the use of the zero-inflated-poisson estimator which imposes the restriction of literally 0 probability of entry on some regions or countries. We are persuaded by Wooldridge (1991) that the better way to deal with possible overdispersion (and underdispersion which also occurs) is to eschew pre-testing for model/estimator selection and instead separately estimate standard errors for the poisson coefficients which are also unbiased across a range of plausible models. In the past we have used Wooldridge’s regression based method which works but requires writing your own subroutine. Our estimates were calculated by the Stata 9.0 statistical package which includes robust standard errors as an option for poisson estimation; this option solves the problems discussed by Wooldridge and others.

III. Empirical Results

This section discusses our empirical results for both firm entry (III.1) and non-university patenting (III.2). We use the Stata 9.0 statistical package for poisson estimation with robust standard errors for all the estimates presented in these subsections for the reasons just discussed. The third subsection examines whether the major determinants in these regressions are becoming more diffuse or more concentrated over time.

III.1. Entry into New Technology Areas

Table 1 above provides summary statistics for the variables used in the main analyses. Our empirical results for entry of firms into new (to them) science and technology areas are reported in Table

2 for the U.S. regions data set, Table 3 for the top-25 science and technology countries data set, and Table 4 for the top-24 non-U.S. science and technology countries data set.

**Table 2. Firm Entry into New Technologies – Poisson Regressions
U.S. Functional Economic Regions, 1981-2004**

Explanatory Variables	Science and Technology Areas of Firm Entry					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Region/ Country in Same S&T Area as Entry	0.0077*** (0.0008)	0.0088 (0.0058)	0.0459*** (0.0139)	0.0045 (0.0059)	0.0173*** (0.0031)	0.0355*** (0.0062)
High Impact Articles Knowledge Stock in Same S&T Area as Entry	-0.0012* (0.0006)	0.0292*** (0.0033)	0.0456* (0.0224)	0.0402*** (0.0055)	0.0037*** (0.0011)	-0.0295*** (0.0070)
University Articles Knowledge Stock in Same S&T Area as Entry	0.0000* (0.0000)	-0.0014*** (0.0002)	-0.0012 (0.0008)	-0.0010*** (0.0001)	-0.0006*** (0.0001)	0.0016*** (0.0004)
University Patents Knowledge Stock in Same S&T Area as Entry	-0.0048*** (0.0009)	-0.0075 (0.0078)	-0.0056 (0.0050)	0.0376*** (0.0094)	0.0048** (0.0021)	-0.0090** (0.0032)
Total Employment in Region/Country (millions of persons)	0.0961*** (0.0186)	0.1564*** (0.0261)	0.0971*** (0.0028)	0.2391*** (0.0341)	0.1921*** (0.0191)	0.1434*** (0.0171)
Average Wage per Job in Region (thousands of 2000 dollars)	0.1287*** (0.0072)	0.1783*** (0.0094)	0.1990*** (0.0098)	0.1318*** (0.0088)	0.1161*** (0.0083)	0.1213*** (0.0065)
Constant	-1.8895*** (0.1872)	-4.1122*** (0.2459)	-5.1264*** (0.2627)	-3.0198*** (0.2278)	-2.2201*** (0.2225)	-0.9775*** (0.1713)
Dummy = 1 in 2002, else 0	-0.8422*** (0.1117)	-0.9257*** (0.1033)	-0.7564*** (0.1291)	-0.7297 (0.1170)	-0.9511*** (0.1399)	-1.1007*** (0.1127)
Dummy = 1 in 2003, else 0	-1.2904*** (0.11512)	-1.6108*** (0.1258)	-1.4691*** (0.1323)	-0.9290*** (0.1189)	-1.2570*** (0.1202)	-1.7382*** (0.1202)
Dummy = 1 in 2004, else 0	-2.0501*** (0.1377)	-2.3881*** (0.1341)	-2.6671*** (0.1747)	-1.5562*** (0.1381)	-2.2833*** (0.1407)	-3.1179*** (0.1432)
Pseudo R ²	0.6181	0.5802	0.5278	0.5266	0.5275	0.5822

Notes: Robust standard errors in parentheses below coefficient estimates. N = 4296. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001
1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

First, however, it is important to emphasize the major result: the number of star scientists and engineers active in a region or country has positive (with one exception) and generally significant effects on the probability of a firm entering in all six science and technology areas. These results are even stronger – as reported in Appendix A – if the analysis is restricted to the first third of the stars who met a higher ISI hurdle than required of later selectees.¹¹ Although we present the estimates for all S&T areas lumped together, we do not discuss them as their magnitudes reflect the mix of areas in a nation or region.

Focusing now on Table 2, we note that only for the computing and semiconductors area is the coefficient on stars not statistically significant. This is consistent with the views expressed by some industry observers that many or most of the most important advances in these areas is made in industry rather than by university faculty who are more prone to publication in ISI-indexed journals. However, note that in the international regressions (Tables 3 and 4) computing is highly significant and

¹¹ Since we know this only by an accident of timing of a required early draft for a conference and want to avoid data mining leading to misleading significance levels, we do not make these the focus results. We do, however, consider setting the right threshold for stardom to be a prime issue for future research.

semiconductors is significant at the 5 percent level when the U.S. is included in the sample. An important issue for future research is to model which areas of science will have the greatest impact on entry and success of high-tech firms. The three knowledge stock variables measuring university science base including the bulk of the stars' actual discoveries are not even consistently positive in their impact with the simple counts of stars included.¹² This seems to say that top scientists and engineers are the ferment driving the formation and transformation of high-tech firms, not their inventions separate from themselves. The economic geography variables seem to be well behaved, with both the size of the region and the average education level (captured here by average wage) having significant positive effects on the entry of firms into new areas.¹³ The dummy variables were included to capture the fact that some of the entries which will be eventually observed using our methodology cannot be seen yet since many applied for patents from 2002-2004 were still pending when our data set was created and even some articles to appear in journals with 2004 cover dates had yet to be published. These dummies had the expected negative sign and increasing size indicating increasing truncation effects.

**Table 3. Firm Entry into New Technologies – Poisson Regressions
Top-25 Science & Technology Countries, 1981-2004**

Explanatory Variables	Science and Technology Areas of Firm Entry					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Region/ Country in Same S&T Area as Entry	0.0035*** (0.0001)	0.0140*** (0.0025)	0.0651*** (0.0049)	0.0133* (0.0055)	0.0192*** (0.0038)	0.0254** (0.0081)
High Impact Articles Knowledge Stock in Same S&T Area as Entry	-0.0019*** (0.0001)	-0.0043** (0.0014)	-0.0008 (0.0045)	-0.0031 (0.0027)	-0.0036*** (0.0007)	0.0024 (0.0023)
University Articles Knowledge Stock in Same S&T Area as Entry	0.0000*** (0.0000)	0.0001^ (0.0000)	0.0002*** (0.0001)	0.0000 (0.0000)	0.0002*** (0.0001)	0.0001 (0.0001)
University Patents Knowledge Stock in Same S&T Area as Entry	-0.0010*** (0.0001)	-0.0023^ (0.0013)	-0.0022*** (0.0005)	-0.0034 (0.0027)	-0.0046*** (0.0006)	-0.0019** (0.0006)
Total Employment in Region/Country (millions of persons)	-0.0017*** (0.0003)	-0.0002 (0.0003)	0.0004 (0.0004)	-0.0001 (0.0004)	-0.0004 (0.0003)	-0.0001 (0.0004)
Constant	2.5454*** (0.0532)	1.7118*** (0.0728)	1.3053*** (0.1304)	2.1689*** (0.1192)	1.7140*** (0.1169)	2.9924*** (0.1118)
Dummy = 1 in 2002, else 0	-0.1958^ (0.1113)	-0.2609 (0.1904)	-0.5906*** (0.1097)	0.2894 (0.1890)	-0.3826* (0.1820)	-0.4882^ (0.2586)
Dummy = 1 in 2003, else 0	-0.3124^ (0.1690)	0.4756^ (0.2585)	-1.2523*** (0.2036)	0.3956^ (0.2110)	-0.9073*** (0.2691)	-0.9040*** (0.2120)
Dummy = 1 in 2004, else 0	-2.1357*** (0.5411)	-0.7148** (0.2402)	-1.6237** (0.5712)	-0.1660 (0.2171)	-2.3861*** (0.5524)	-1.5355*** (0.2039)
Pseudo R ²	0.8889	0.8557	0.7737	0.7720	0.8597	0.8589

Notes: Robust standard errors in parentheses below coefficient estimates. N = 600. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

¹² In results not reported here, we checked that these variables do have their usual positive impact if stars are excluded from the regressions. If, on the other hand, these variables are excluded then the estimated impact of the stars generally increases. This is to be expected given the moderately high correlations among these explanatory variables reported in the electronic-only Appendix B, Table B7.

¹³ An anonymous referee has pointed out that this coefficient is biased to the extent that labor supply and demand shocks induce correlation of average wages with the error term. Fortunately, the variable is included only in the U.S. regional regressions and does not seem to affect variables of main interest qualitatively. In future research we plan to replace the wage variable with a suitable direct measure of skill levels in the region or country.

The results for the 25-country regressions in Table 3 are similar to Table 2 with the notable exception that the university articles stock does appear to have a consistent positive and mostly significant effect on firm entry. Economic geography loses its consistent effects, perhaps because of our inability to find a commensurable average wage or education level variable and perhaps also because of the very large size of China and India. The 2002-2004 dummy variables follow the same general pattern except for computing/IT and semiconductors where there are two borderline significant positive values.

**Table 4. Firm Entry into New Technologies – Poisson Regressions
Top-24 Non-U.S. Science & Technology Countries, 1981-2004**

Explanatory Variables	Science and Technology Areas of Firm Entry					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Region/ Country in Same S&T Area as Entry	0.0091*** (0.0014)	0.0349*** (0.0076)	0.0434^ (0.0227)	0.0046 (0.0099)	-0.0096 (0.0139)	0.0721*** (0.0186)
High Impact Articles Knowledge Stock in Same S&T Area as Entry	-0.0014* (0.0006)	-0.0106** (0.0039)	0.0305*** (0.0089)	0.0174*** (0.0027)	0.0042 (0.0026)	0.0028 (0.0073)
University Articles Knowledge Stock in Same S&T Area as Entry	0.0000*** (0.0000)	0.0003*** (0.0000)	-0.0000 (0.0001)	0.0000* (0.0000)	0.0002*** (0.0001)	0.0001 (0.0001)
University Patents Knowledge Stock in Same S&T Area as Entry	-0.0001 (0.0009)	-0.0044 (0.0099)	0.0136*** (0.0021)	-0.0068 (0.0076)	-0.0131^ (0.0078)	0.0068** (0.0027)
Total Employment in Region/Country (millions of persons)	-0.0009*** (0.0002)	-0.0010*** (0.0002)	-0.0003 (0.0003)	-0.0004 (0.0003)	-0.0007** (0.0003)	-0.0003 (0.0003)
Constant	2.3401*** (0.0423)	1.0167*** (0.0675)	1.1450*** (0.0604)	1.3694*** (0.0569)	1.5851*** (0.0610)	2.8621*** (0.0565)
Dummy = 1 in 2002, else 0	-0.1984 (0.1570)	-0.5293*** (0.1503)	-0.8945*** (0.2464)	0.1950^ (0.1031)	-0.0923 (0.1486)	-0.7844** (0.2584)
Dummy = 1 in 2003, else 0	-0.2997^ (0.1644)	-0.2219 (0.2099)	-1.5045*** (0.2427)	-0.0479 (0.1450)	-0.5350*** (0.1679)	-1.1910*** (0.2516)
Dummy = 1 in 2004, else 0	-0.8248*** (0.1884)	-1.2219 (0.2628)	-1.7750*** (0.3071)	-0.1240 (0.1136)	-0.9641*** (0.2186)	-2.0436*** (0.2513)
Pseudo R ²	0.5905	0.4685	0.4421	0.5009	0.4445	0.4805

Notes: Robust standard errors in parentheses below coefficient estimates. N = 576. Significance levels: ^ 0.10, * 0.05, ** 0.01, *** 0.001
1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

We were concerned that the U.S. might dominate the country regressions since it is such a scientific and commercial giant: In Table 1 the standard deviations of the data for the 25-country data set are typically an order of magnitude higher than those for the 24-country data set excluding the U.S. With the U.S. excluded from the sample in the regressions reported in Table 4, the results for the star counts are indeed somewhat less strong with three areas showing highly significant impacts and a fourth with a large coefficient but only borderline significance. Semiconductors and the catch-all Other Sciences area are statistically insignificant and in the latter case the only observed negative value for the star counts in these three tables.

Appendix A reports matched regressions where the empirical definition of stars is restricted to the first third (1,838 or 34%) of the ISI Highly Cited, who on average met a somewhat higher hurdle than subsequent highly cited authors. Star scientists have a significant, positive effect in every industry when the models of Tables 2 and 3 are rerun with these data, suggesting that stars might be better defined by

starting with the ISI Highly Cited data and then imposing further screens. We will explore that hypothesis in future research.

If a significant number of stars were moving to new locations to work at or with new firms started by other scientists or entrepreneurs, it is possible that they would provide the firm with its first article or patent, creating a spurious entry date corresponding with their observed arrival in a new location. The electronic-only Appendix A demonstrates that eliminating all firms for which entry is identified by an article or patent on which a star is an author or inventor yields results which are very similar to those in Tables 2-4.

Table 5. Firm Entry into Nanotechnology – Comparing Archival and Article/Patent Measures of Firm Entry
U.S. Functional Economic Regions, Top-25 Science & Technology Countries, and
Top-24 Non-U.S. Science & Technology Countries – Poisson Regressions – 1981-2004

	U.S. Regions		Top-25 S&T Countries		Top-24 Non-US S&T Countries	
	Archival Firms	Articles/Patents	Archival Firms	Articles/Patents	Archival Firms	Articles/Patents
Star Scientists & Engineers Active in Region/ Country in Same S&T Area as Entry	0.0719*** (0.0152)	0.0459*** (0.0139)	0.0447*** (0.0049)	0.0651*** (0.0049)	0.0452* (0.0222)	0.0434^ (0.0227)
High Impact Articles Knowledge Stock in Same S&T Area as Entry	0.0840*** (0.0177)	0.0456* (0.0224)	0.0056* (0.0027)	-0.0008 (0.0045)	0.0432*** (0.0089)	0.0305*** (0.0089)
University Articles Knowledge Stock in Same S&T Area as Entry	-0.0023*** (0.0006)	-0.0012 (0.0008)	0.0001* (0.0001)	0.0002*** (0.0001)	-0.0002^ (0.0001)	-0.0000 (0.0001)
University Patents Knowledge Stock in Same S&T Area as Entry	-0.0077** (0.0036)	-0.0056 (0.0050)	-0.0019*** (0.0004)	-0.0022*** (0.0005)	0.0137*** (0.0036)	0.0136*** (0.0021)
Total Employment in Region/Country (millions of persons)	0.0908** (0.0299)	0.0971*** (0.0028)	0.0004 (0.0004)	0.0004 (0.0004)	0.0006 (0.0006)	-0.0003 (0.0003)
Average Wage per Job in Region (thousands of 2000 dollars)	0.2273*** (0.0128)	0.1990*** (0.0098)	-	-	-	-
Constant	-8.2485*** (0.3593)	-5.1264*** (0.2627)	-0.1238 (0.1449)	1.3053*** (0.1304)	-0.8266*** (0.1000)	1.1450*** (0.0604)
Dummy = 1 in 2002, else 0	-0.4405** (0.1558)	-0.7564*** (0.1291)	-0.3912* (0.1871)	-0.5906*** (0.1097)	-0.3508 (0.3345)	-0.8945*** (0.2464)
Dummy = 1 in 2003, else 0	-0.5782** (0.1927)	-1.4691*** (0.1323)	-0.6617** (0.2415)	-1.2523*** (0.2036)	-0.5746 (0.3640)	-1.5045*** (0.2427)
Dummy = 1 in 2004, else 0	-1.2253*** (0.1825)	-2.6671*** (0.1747)	-0.5207^ (0.3150)	-1.6237** (0.5712)	-0.6116 (0.4000)	-1.7750*** (0.3071)
Pseudo R ²	0.4906	0.5278	0.6788	0.7737	0.2898	0.4421
N	4296	4296	600	600	576	576

Notes: Robust standard errors in parentheses below coefficient estimates. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001

1. The Archival Firms measure of firm entry is based on firms identified in two major nanotech firm directories supplemented by a web-search utilizing firm references found in other firms. Founding dates are used as entry dates for firms founded 1985 or later. For incumbent firms (those founded before 1985) only 3 entry dates into nanotechnology were obtained from firm websites and press releases. For other incumbent firms, entry dates were drawn from the distribution of known entry dates in the same U.S. region (if applicable) or otherwise from the distribution of known entry dates in the same country; thus incumbent firms (with 3 exceptions) provide information on the where but not the when of entry on the archival measure. The Articles/Patents measure of firm entry is that used in Tables 2-4 from which the so-titled regressions are extracted.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

For nanotechnology we can compare the results in Tables 2-4 with results based on the NanoBank.org data on firms involved in nanotechnology based on archival sources. Table 5 estimates entry based on these archival data and repeats for comparison the corresponding regressions based on our

basic article publication and patenting measure of firm entry. The results are very similar and in every case the independent effect of the star scientists on entry are either as significant as or more significant with the archival measure than with the article/patents measure. The pseudo R^2 is generally lower for the archival than our base measure, and this probably reflects the need for random assignment of entry dates for most incumbent firms. We interpret these results as very supportive of our base measure of entry.

Table 6. Firm Entry into New Technologies – Poisson Regressions with Location-Specific Fixed Effects and Explanatory Variables Lagged One Year U.S. Functional Economic Regions, 1982-2004

Explanatory Variables	Science and Technology Areas of Firm Entry					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Region/ Country in Same S&T Area as Entry	0.0088*** (0.0015)	0.0117* (0.006)	0.0712*** (0.0167)	0.0034 (0.0041)	0.0051 (0.005)	0.0105** (0.0038)
High Impact Articles Knowledge Stock in Same S&T Area as Entry (hundreds)	-0.1107^ (0.058)	-0.6360^ (0.3476)	-2.7154* (1.0847)	-0.7102*** (0.2164)	0.1022 (0.1006)	0.2125 (0.3004)
University Articles Knowledge Stock in Same S&T Area as Entry (hundreds)	0.0018* (0.0009)	0.0534*** (0.0132)	0.1678*** (0.0419)	0.0068^ (0.0039)	0.0094 (0.0085)	0.0788*** (0.0187)
University Patents Knowledge Stock in Same S&T Area as Entry (hundreds)	-0.1186** (0.0422)	-0.8275^ (0.455)	-0.6574*** (0.1443)	0.1467 (0.4744)	-0.3281* (0.1302)	-0.0909 (0.0838)
Total Employment in Region/Country (millions of persons)	0.0400 (0.0614)	0.1154 (0.0984)	0.4899*** (0.1362)	-0.1440* (0.057)	-0.0511 (0.0491)	-0.1652* (0.0732)
Average Wage per Job in Region (thousands of 2000 dollars)	0.0031 (0.0078)	0.0039 (0.0139)	0.0845*** (0.0163)	0.0253*** (0.0076)	0.0177* (0.0076)	-0.0146* (0.0075)
Constant	0.838*** (0.2225)	-0.3242 (0.4279)	-3.9565*** (0.5387)	-1.5924*** (0.3779)	-0.4671 (0.2856)	2.5526*** (0.1986)
Dummy = 1 in 2003, else 0	-0.7089*** (0.0601)	-1.0758*** (0.1187)	-1.2996*** (0.1165)	-0.2969*** (0.078)	-0.5961*** (0.0685)	-1.0621*** (0.0688)
Dummy = 1 in 2004, else 0	-1.3461*** (0.0748)	-1.8147*** (0.119)	-2.7430*** (0.1978)	-0.8920*** (0.102)	-1.5510*** (0.094)	-2.2527*** (0.0763)
Pseudo R^2	0.7910	0.7614	0.7007	0.7396	0.7381	0.8202

Notes: Robust standard errors in parentheses below coefficient estimates. N = 4117. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

Some readers may have concerns as to whether we have demonstrated the direction of causality. Unfortunately we have been unable to come up with any instruments to identify the effect of the stars along the lines served by early and unexpected death of stars for the analysis of Azoulay, Graff Zivin, and Wang (2010). Instead, following the suggestions of workshop participants and an anonymous referee, we re-estimated the basic models with location fixed effects and also lagged the explanatory variables by one year. These equations rely on the changes over time in each region or country as opposed to differences across regions or countries. Thus, any unmeasured region/country specific factor which might make a given location both more attractive to stars and a better place to start firms will be attributed to the location-specific dummy and not to the count of stars. By using lagged variables, we further reduce the plausibility of any reverse causation in which the entry of firms causes the movement of stars. Tables 6, 7, and 8 report the results of these estimations. The results are qualitatively the same or stronger as compared to our basic results in Tables 2-4: In every case at least two-thirds of the science and technology areas show coefficients on active stars which are significant at the 0.01 level or better. The

evidence of this star effect is weak only for semiconductors and the catch-all other science areas. As expected based on our field interviews with star scientists, the evidence does not support the hypothesis that they move their location to join or work with a new firm created by others.

Table 7. Firm Entry into New Technologies – Poisson Regressions with Location-Specific Fixed Effects and Explanatory Variables Lagged One Year Top-25 Science & Technology Countries, 1982-2004

Explanatory Variables	Science and Technology Areas of Firm Entry					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Region/ Country in Same S&T Area as Entry	0.0026*** (0.0003)	0.0040 (0.0035)	0.0205*** (0.0032)	0.0029*** (0.0008)	0.0048* (0.002)	0.0153*** (0.0024)
High Impact Articles Knowledge Stock in Same S&T Area as Entry (hundreds)	-0.0856*** (0.0149)	-0.4093*** (0.0647)	-0.4796 (0.4)	-0.3491*** (0.039)	-0.0341 (0.0295)	-0.1938 (0.1382)
University Articles Knowledge Stock in Same S&T Area as Entry (hundreds)	0.0006*** (0.0001)	0.0098*** (0.0024)	0.0119** (0.0045)	0.0036*** (0.0005)	0.0030^ (0.0017)	0.0102* (0.0043)
University Patents Knowledge Stock in Same S&T Area as Entry (hundreds)	-0.0514*** (0.0081)	-0.4521*** (0.1059)	-0.0905* (0.0375)	-0.8982*** (0.1058)	-0.1692*** (0.037)	-0.1771*** (0.026)
Total Employment in Region/Country (millions of persons)	0.0082*** (0.0013)	0.0086* (0.0037)	0.0626*** (0.0122)	0.0068*** (0.0021)	0.0127*** (0.003)	0.0138*** (0.0031)
Constant	1.6361*** (0.1428)	1.7623*** (0.1866)	0.6512* (0.2547)	1.5485*** (0.2018)	1.4920*** (0.2152)	3.1938*** (0.1458)
Dummy = 1 in 2003, else 0	-0.2974** (0.0976)	-0.4810** (0.1677)	-0.6708*** (0.1041)	0.1068 (0.0845)	-0.5894*** (0.1313)	-0.9452*** (0.11)
Dummy = 1 in 2004, else 0	-0.8118*** (0.1292)	-0.4427* (0.2206)	-1.4874*** (0.1857)	-0.1297 (0.1055)	-1.1109*** (0.1676)	-1.7574*** (0.1292)
Pseudo R ²	0.9441	0.9233	0.8984	0.9316	0.9410	0.9652

Notes: Robust standard errors in parentheses below coefficient estimates. N = 575. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001
1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

Taken as a whole, the results show that the physical presence of star scientists and engineers plays an important role in the formation and transformation of high-tech firms. These results have held up despite subjecting the data to reanalyses proposed by a number of independent discussants and referees. There is some evidence, especially in the country regressions, that academic discoveries (or unmeasured discoverers) play an independent role as well. The economic geography variables work well in the U.S. regional context, but do not translate well in the international context.

III.2. Non-university Patenting

Our attempt to explain non-university patenting using stars, academic knowledge stocks, and economic geography were not so successful as reported in the electronic-only Appendix B in Tables B5-B10 spanning both the full and restricted sets of stars. The university articles knowledge stock does show consistent positive effects. Education levels as measured by average wages are also consistently positive for the U.S. regional results. Our problem may be that the three knowledge stocks are correlated not only among themselves but with the star counts (see electronic-only Appendix B, Table B7). Apparently, any positive effect of the local academic knowledge stock or stars is much weaker than is the case for firm entry into new science and technology areas.

While the results for patenting are not commensurate with the effort that went into the analysis, they do illustrate one important point: There is nothing inherent in the empirical methodology that artificially creates significant coefficients.

Table 8. Firm Entry into New Technologies – Poisson Regressions with Location-Specific Fixed Effects and Explanatory Variables Lagged One Year Top-24 Non-U.S. Science & Technology Countries, 1982-2004

Explanatory Variables	Science and Technology Areas of Firm Entry					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Region/ Country in Same S&T Area as Entry	0.0062*** (0.0012)	0.0413*** (0.0089)	0.0536** (0.0176)	0.0040 (0.0055)	0.0068 (0.0064)	0.0647*** (0.0112)
High Impact Articles Knowledge Stock in Same S&T Area as Entry (hundreds)	0.1259 (0.0831)	-0.8028^ (0.4236)	4.3287*** (0.6711)	-0.2911 (0.2114)	0.0600 (0.2007)	0.4702 (0.3278)
University Articles Knowledge Stock in Same S&T Area as Entry (hundreds)	0.0000 (0.0002)	0.0269*** (0.004)	-0.0417*** (0.0048)	0.0037*** (0.0007)	0.0058 (0.0051)	-0.0057 (0.0056)
University Patents Knowledge Stock in Same S&T Area as Entry (hundreds)	-0.0519 (0.0652)	-4.9914*** (1.1518)	0.5993* (0.2452)	-2.5004^ (1.3833)	-0.3692 (0.3549)	-0.4310* (0.2177)
Total Employment in Region/Country (millions of persons)	0.0120*** (0.002)	0.0159*** (0.0036)	0.0406** (0.0149)	0.0131*** (0.0025)	0.0157*** (0.0033)	0.0215*** (0.0031)
Constant	1.6093*** (0.1439)	1.4920*** (0.1789)	0.9118*** (0.2768)	1.5156*** (0.2042)	1.4537*** (0.2135)	3.2031*** (0.1498)
Dummy = 1 in 2003, else 0	-0.2561* (0.1091)	-0.4428*** (0.111)	-0.8486*** (0.169)	0.1355 (0.1117)	-0.4381** (0.1672)	-0.7956*** (0.1227)
Dummy = 1 in 2004, else 0	-0.7476*** (0.1299)	-0.6199* (0.2909)	-1.1978*** (0.235)	-0.1977^ (0.1171)	-0.9358*** (0.2063)	-1.6955*** (0.1374)
Pseudo R ²	0.7206	0.6301	0.6275	0.6538	0.6953	0.8194

Notes: Robust standard errors in parentheses below coefficient estimates. N=552. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001
1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

III.3. Concentration or Diffusion of Stars and Knowledge

Given the clear importance of stars for firm entry and – in the country regressions – also the university articles knowledge stock, it is interesting to know how these change over time. Table 9 reports the correlation coefficients between the growth rates and levels of stars and each of the three knowledge stock variables by science and technology area. If this correlation is significantly positive, it means that those regions with above average levels of the variable are also growing faster in percentage terms implying increasing concentration over time: the rich get richer. On the other hand, a significant negative value of this correlation coefficient means below average regions or countries are growing more rapidly and above average ones less rapidly implying diffusion or a tendency toward equalization over time.

Stars show a clear tendency toward concentration by area in the U.S. regional data set with the effect weakest for the Biology, Chemistry, and Medicine area. (Note that aggregating the data into all scientists and engineers masks the pattern apparent when examined by S&T areas.) For the 24 non-U.S. countries, three of the S&T areas have insignificant correlations, two have strongly positive correlations, and 1 has a borderline significantly positive correlation. The pattern of stars moving to where there are more stars in the same S&T area disappears entirely in the 25 country data set where no correlations are significant and they are positive for only 4 of the six fields.

Table 9. Correlation Coefficients for the Levels and Growth Rates of Star Scientists & Engineers and Knowledge Stocks

	Correlation Coefficients of Level and Growth Rate across Years and Regions/Countries by S&T Field						
	All Sci. & Eng.	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
<u>US Regions</u>							
Star Scientists & Eng.	0.02	0.04 [^]	0.08**	0.19***	0.07**	0.11***	0.19***
High Impact Articles	-0.03	-0.02	-0.04	-0.03	-0.03	-0.03	-0.04
University Articles	-0.07**	-0.07**	-0.08**	-0.09**	-0.08**	-0.04 [^]	-0.04 [^]
University Patents	-0.03	-0.03	0	-0.02	-0.02	-0.03	-0.03
<u>Top-25 S&T Countries</u>							
Star Scientists & Eng.	-0.03	-0.05	0.04	-0.01	0.00	0.02	0.01
High Impact Articles	-0.09*	-0.08 [^]	-0.06	-0.08	-0.05	-0.05	-0.05
University Articles	-0.16***	-0.16***	-0.18***	-0.18***	-0.21***	-0.08*	-0.16***
University Patents	-0.03	-0.03	-0.02	-0.04	-0.02	-0.04	-0.01
<u>Top-24 Non-US S&T Countries</u>							
Star Scientists & Eng.	-0.07 [^]	-0.06	0.04	0.20***	0.08 [^]	0.08	0.17**
High Impact Articles	-0.18***	-0.15***	-0.01	-0.13*	-0.09*	-0.11*	-0.09 [^]
University Articles	-0.29***	-0.29***	-0.29***	-0.23***	-0.27***	-0.15***	-0.26***
University Patents	-0.03	-0.05	-0.01	0.02	0.01	-0.03	-0.02

Notes: Significance levels: [^] 0.10, * 0.05, ** 0.01, ***0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles and patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.

Explaining these results is an interesting agenda for future research. We conjecture based on anecdotal evidence – and supported by Kim’s (2010) case study for Korea – that the results reflect two contradictory motivations for movement of stars. First, stars tend to move to where there are more other stars – for example from lower to higher ranked universities – as well as to where there are greater commercial opportunities (Zucker and Darby (1999b)).¹⁴ Overlaying this pattern during the last quarter century, however, are movements of many U.S. trained foreign students who build successful careers in American academe, perhaps moving from lower to higher ranked U.S. universities but choosing to return home when their native countries develop sufficient strength in their disciplines to both seek them out and to be attractive (Saxenian 2005). This countercurrent could cancel the tendency toward concentration when the U.S. is in the country data set. Since this effect is present to a somewhat similar degree in all American universities, the reverse brain drain of expatriate stars affects the average growth rate of stars in the U.S. without weakening the positive correlation across countries. The intermediate results for the non-U.S. data set may well reflect a similar pattern occurring in some other countries which are also centers of doctoral education.

For the knowledge stocks in the U.S regions and 25 countries data sets, there is a general tendency toward diffusion or equalization of knowledge stocks, with significant negative correlations appearing most consistently for university articles – possibly the diffusion of the “publish or perish” standard from elite American universities to others both domestically and internationally. Domestic diffusion may have kept up the overall growth rate of American knowledge stocks, since the pattern of diffusion of university articles and even high impact articles is much stronger among the 24 non-U.S. countries.

¹⁴ A seemingly related paper, Zucker, Darby, and Torero (2002), focused not on inter-regional or international mobility of biotech stars, but mobility from purely academic to either affiliated or linked (co-authoring) with firm scientists.

IV. Conclusions

We have shown that it is possible to generalize the definition of star scientists and engineers far beyond its biotechnology origins. We have also seen that doing so offers convincing evidence that these extraordinary people play a key role in the formation and transformation of high-tech firms. We have seen that while there is a general tendency toward diffusion of academic knowledge stocks, the movement of star scientists and engineers is toward concentration. One offsetting trend may be the tendency for foreign-born American academics to return home when their country becomes strong in their disciplines.

Appendix A

In the first version of this paper (Zucker and Darby 2006b) produced to meet a conference deadline, we reported results for the first 1,838 star scientists which we had downloaded from isihighlycited.com, parsed, cleaned, and matched. Since our downloading procedure was based on ISI's serial ID numbers, we downloaded all 888 of the individuals collected by ISI in the years 2000 through 2001 and the first 950 collected in the year 2000. ISI appears to have selected only the very most highly cited authors for inclusion in the first year of ISIHIGHLYCITED, with those appearing in later years clearing progressively lower hurdles. The combination of our systematic download procedure and first-draft deadline inadvertently produced an enriched sample of more stellar stars than the full set of 5,401 on which we focus for this paper. Table A1 reports sample statistics for this subset corresponding to text Table 1.¹⁵ Tables A2-A4 report the regression results obtained for this smaller set of stars and correspond exactly in format and order to the text Tables 2-4

We will not trespass on the patience of reader or editor by going through a detailed discussion of the results. We believe that they may be fairly summarized as at least as strong and sometimes stronger evidence than the full-sample results for the direct importance of top scientists and engineers for firm creation and patenting. Since these stars are princes and princesses of serendipity, we are doubly grateful for our fortunate confirmation of the theory.

An alert referee asked whether our results on firm entry might not be an artifact of our method of detecting firms newly entering an area by publication or patenting therein. Could the large productivity of the star scientists account for many of the articles and patents which we use to detect entry? Tables A5-A7 correspond to Tables 2-4 with the exception that we exclude any entry for a firm for which we detected its entry due to an article or patent which included a star scientist as either author or inventor. This is an acid test since we are excluding firms for which we might detect entry from a subsequent non-star article or patent. Since the results in Tables A5-A7 correspond closely to those in Tables 2-4, we conclude that the reported effect on entry is real and not an artifact of our measurement methods.

¹⁵ These Tables differ slightly from the tables originally reported in Zucker and Darby (2006b) due to continued cleaning of the data and correction of typographical errors.

Table A1. Summary Statistics for Variables Used in Empirical Analysis Based on Only First 1,838 Stars, 1981-2004

Variables	U.S. Regions					Top-25 Sci. & Tech Countries					Top-24 non-U.S S&T Countries				
	N	Mean	S.D.	Min	Max	N	Mean	S.D.	Min	Max	N	Mean	S.D.	Min	Max
<u>Entry of Firms into:</u>															
Biology/Chemistry/Medicine	4296	7.2	15.9	0	165	600	45.8	149.4	0	1067	576	21.8	24.9	0	125
Computing/Information Technology	4296	3.2	9.8	0	203	600	20.7	80.5	0	738	576	7.3	10.3	0	56
Nanotechnology	4296	1.9	6.2	0	87	600	12.5	53.5	0	483	576	5.4	9.7	0	71
Semiconductors	4296	2.7	7.3	0	90	600	24.5	80.7	0	657	576	9.3	12.2	0	73
Other Sciences	4296	3.6	8.3	0	104	600	24.5	89.0	0	609	576	9.5	13.5	0	66
Other Engineering	4296	12.8	24.7	0	248	600	72.8	287.0	0	2107	576	29.6	42.3	0	207
<u>Non-university patenting:</u>															
Biology/Chemistry/Medicine	4296	30.6	147.7	0	1986.4	600	729.8	2081.8	0	14783.0	576	344.7	663.8	0	4079.3
Computing/Information Technology	4296	20.4	165.2	0	4130.3	600	446.5	1709.5	0	14101.2	576	214.9	828.1	0	9188.5
Nanotechnology	4296	9.8	73.7	0	2089.8	600	247.8	966.5	0	8245.5	576	141.2	625.5	0	7967.5
Semiconductors	4296	5.4	46.7	0	1126.7	600	162.1	544.0	0	4453.6	576	96.1	347.4	0	3039.5
Other Sciences	4296	11.1	50.4	0	946.4	600	311.7	881.2	0	6453.9	576	164.3	446.2	0	4102.3
Other Engineering	4296	71.9	262.1	0	3422.7	600	2286.5	6261.7	0	41986.1	576	1241.1	3133.4	0	23161.3
<u>Star Scientists & Engineers Active</u>															
Biology/Chemistry/Medicine	4296	5.0	15.3	0	121	600	57.1	186.8	0	1140	576	19.6	26.0	0	130
Computing/Information Technology	4296	0.8	3.0	0	30	600	8.2	28.0	0	169	576	2.6	3.4	0	17
Nanotechnology	4296	0.2	0.9	0	12	600	2.0	7.2	0	58	576	0.8	1.7	0	10
Semiconductors	4296	1.1	3.8	0	43	600	11.2	40.7	0	270	576	3.1	3.3	0	19
Other Sciences	4296	0.6	2.4	0	26	600	6.3	21.5	0	132	576	2.0	3.3	0	17
Other Engineering	4296	0.3	1.2	0	11	600	3.3	11.6	0	76	576	1.0	1.8	0	11
<u>High Impact Articles Knowledge Stock</u>															
Biology/Chemistry/Medicine	4296	13.1	55.0	0	518.9	600	277.3	834.2	0	4,911.6	576	115.0	156.2	0	870.3
Computing/Information Technology	4296	2.5	11.5	0	123.5	600	58.4	162.4	0	1,194.3	576	27.0	34.6	0	248.2
Nanotechnology	4296	0.5	2.9	0	65.9	600	15.1	53.7	0	513.1	576	7.6	14.4	0	96.9
Semiconductors	4296	2.0	9.7	0	115.4	600	53.4	146.1	0	907.9	576	25.3	32.0	0	137.1
Other Sciences	4296	5.0	25.1	0	246.0	600	96.7	324.2	0	1,994.1	576	34.7	52.6	0	269.4
Other Engineering	4296	1.3	7.3	0	75.2	600	30.4	91.2	0	528.3	576	12.6	18.7	0	103.4
<u>University Articles Knowledge Stock</u>															
Biology/Chemistry/Medicine	4296	1485.0	5068.2	0	61,361.2	600	46272.6	108860.9	64.0	729,673.6	576	26210.2	31255.5	64.0	159,973.2
Computing/Information Technology	4296	90.9	317.1	0	3,803.9	600	3501.5	7268.2	6.0	49,950.6	576	2207.5	2476.7	6.0	12,449.2
Nanotechnology	4296	20.0	94.2	0	1,509.7	600	953.9	2557.6	0	24,685.3	576	654.0	1338.9	0	9,289.5
Semiconductors	4296	159.7	577.8	0	6,030.3	600	7438.5	13386.2	24.4	81,232.2	576	5198.3	6455.8	24.4	36,102.0
Other Sciences	4296	99.7	374.2	0	4,161.1	600	3245.2	7703.8	3	48,111.6	576	1816.8	2207.4	3	11,874.8
Other Engineering	4296	44.0	158.6	0	1,691.6	600	1534.7	3622.9	1	23,672.5	576	875.5	1166.2	1	7,241.4
<u>University Patents Knowledge Stock</u>															
Biology/Chemistry/Medicine	4296	8.7	40.4	0	564.3	600	134.7	672.2	0	5259.8	576	17.4	37.4	0	262.0
Computing/Information Technology	4296	0.7	3.9	0	66.6	600	11.3	61.0	0	576.3	576	1.3	3.6	0	35.5
Nanotechnology	4296	3.3	21.7	0	415.8	600	53.5	334.6	0	3599.1	576	6.3	18.0	0	237.4
Semiconductors	4296	0.3	2.4	0	44.5	600	5.8	28.5	0	294.8	576	0.8	4.4	0	92.8
Other Sciences	4296	2.2	12.3	0	224.4	600	34.8	163.4	0	1281.0	576	4.7	10.4	0	58.9
Other Engineering	4296	3.4	15.3	0	255.8	600	62.1	291.6	0	2330.0	576	9.5	20.0	0	129.1
Total Employment in Region/Country	4296	0.8	1.4	0.0	12.9	600	47.3	123.3	1.0382	752	576	44.2	124.9	1.0	752.0
Average Wage per Job in Region	4296	26.2	4.1	18.3	51.0	—	—	—	—	—	—	—	—	—	—

Notes: 1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
 2. U.S. regions are the 179 functional economic areas defined by the U.S. Bureau of Economic Analysis (Johnson and Kort 2004).

Table A2. Firm Entry into New Technologies – Only First 1,838 Stars Used for Analysis
Poisson Regressions – U.S. Functional Economic Regions, 1981-2004

Explanatory Variables	Science and Technology Areas of Firm Entry					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Region/ Country in Same S&T Area as Entry	0.0194*** (0.0016)	0.0384** (0.0131)	0.1281*** (0.0287)	0.0450*** (0.0083)	0.0493*** (0.0083)	0.1326*** (0.0135)
High Impact Articles Knowledge Stock in Same S&T Area as Entry	-0.0026*** (0.0006)	0.0265*** (0.0038)	0.0355^ (0.0199)	0.0256*** (0.0061)	0.0049*** (0.0011)	-0.0373*** (0.0061)
University Articles Knowledge Stock in Same S&T Area as Entry	0.0000*** (0.0000)	-0.0014*** (0.0002)	-0.0003 (0.0007)	-0.0007*** (0.0001)	-0.0007*** (0.0001)	0.0017*** (0.0004)
University Patents Knowledge Stock in Same S&T Area as Entry	-0.0047*** (0.0008)	-0.0022 (0.0081)	-0.0077^ (0.0199)	0.0332*** (0.0096)	0.0045* (0.0021)	-0.0068^ (0.0028)
Total Employment in Region/Country (millions of persons)	0.0860*** (0.0173)	0.1441*** (0.0218)	0.0664** (-0.0235)	0.1249*** (0.0208)	0.1832*** (0.0194)	0.1252*** (0.0186)
Average Wage per Job in Region (thousands of 2000 dollars)	0.1308*** (0.0074)	0.1792*** (0.0092)	0.2003*** (0.0097)	0.1439*** (0.0280)	0.1159*** (0.0083)	0.1259*** (0.0066)
Constant	-1.9523*** (0.1940)	-4.1337*** (0.2392)	-5.1485*** (0.2597)	-3.2997*** (0.2220)	-2.2145*** (0.2218)	-1.1115*** (0.1743)
Dummy = 1 in 2002, else 0	-0.8262*** (0.1109)	-0.9165*** (0.1049)	-0.7597*** (0.1153)	-0.6797*** (0.1054)	-0.9421*** (0.1379)	-1.0670*** (0.1008)
Dummy = 1 in 2003, else 0	-1.2677*** (0.1116)	-1.5988*** (0.1234)	-1.5156*** (0.1267)	-0.8646*** (0.1097)	-1.2416*** (0.1185)	-1.7126*** (0.1113)
Dummy = 1 in 2004, else 0	-2.0592*** (0.1431)	-2.3893*** (0.1293)	-2.7683*** (0.1986)	-1.5379*** (0.1395)	-2.2576*** (0.1370)	-3.1119*** (0.1432)
Pseudo R ²	0.6261	0.5827	0.5374	0.5348	0.5303	0.5938

Notes: Robust standard errors in parentheses below coefficient estimates. N = 4296. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

Table A3. Firm Entry into New Technologies – Only First 1,838 Stars Used for Analysis
Poisson Regressions – Top-25 Science & Technology Countries, 1981-2004

Explanatory Variables	Science and Technology Areas of Firm Entry					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Region/ Country in Same S&T Area as Entry	0.0068*** (0.0005)	0.0360*** (0.0040)	0.0924*** (0.0097)	0.0180* (0.0079)	0.0354*** (0.0057)	0.0584*** (0.0131)
High Impact Articles Knowledge Stock in Same S&T Area as Entry	-0.0020*** (0.0002)	-0.0072*** (0.0016)	-0.0004 (0.0080)	-0.0032 (0.0028)	-0.0030*** (0.0008)	-0.0077^ (0.0041)
University Articles Knowledge Stock in Same S&T Area as Entry	0.0000*** (0.0000)	0.0002*** (0.0000)	0.0003** (0.0001)	0.0000* (0.0000)	0.0002*** (0.0001)	0.0004*** (0.0001)
University Patents Knowledge Stock in Same S&T Area as Entry	-0.0008*** (0.0001)	-0.0031* (0.0014)	-0.0017* (0.0008)	-0.0007 (0.0033)	-0.0035*** (0.0006)	-0.0014** (0.0004)
Total Employment in Region/Country (millions of persons)	-0.0000*** (0.0000)	-0.0000** (0.0000)	0.0000 (0.0000)	.0000 (0.0000)	-0.0000* (0.0000)	-0.0000 (0.0000)
Constant	2.5835*** (0.0570)	1.5148*** (0.0699)	1.3556*** (0.1291)	2.1523*** (0.1253)	1.7004*** (0.1065)	2.8651*** (0.0957)
Dummy = 1 in 2002, else 0	-0.1754 (0.1161)	-0.4546** (0.2045)	-1.0613*** (0.2164)	0.1913 (0.1795)	-0.5989** (0.2042)	-0.8596*** (0.1729)
Dummy = 1 in 2003, else 0	-0.2512 (0.1555)	0.5873** (0.2083)	-2.0177*** (0.3995)	0.2526 (0.1894)	-1.1576*** (0.2859)	-1.2735*** (0.1375)
Dummy = 1 in 2004, else 0	-2.0691*** (0.5343)	-0.9691*** (0.2637)	-2.5840^ (1.3294)	0.3677 (0.2937)	-2.5481*** (0.5378)	-2.4954*** (0.4110)
Pseudo R ²	0.8803	0.8732	0.7604	0.7701	0.8690	0.8690

Notes: Robust standard errors in parentheses below coefficient estimates. N = 600. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

Table A4. Firm Entry into New Technologies – Only First 1,838 Stars Used for Analysis
Poisson Regressions – Top-24 Non-U.S. Science & Technology Countries, 1981-2004

Explanatory Variables	Science and Technology Areas of Firm Entry					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Region/ Country in Same S&T Area as Entry	0.0174*** (0.0033)	0.0467** (0.0151)	0.0445 (0.0364)	0.0450** (0.0189)	0.0165 (0.0249)	-0.0083 (0.0324)
High Impact Articles Knowledge Stock in Same S&T Area as Entry	-0.0015* (0.0007)	-0.0040 (0.0033)	0.0340*** (0.0090)	0.0132*** (0.0028)	0.0017 (0.0027)	0.0155** (0.0058)
University Articles Knowledge Stock in Same S&T Area as Entry	0.0000*** (0.0000)	0.0003*** (0.0000)	-0.0000 (0.0001)	0.0000*** (0.0000)	0.0003*** (0.0001)	0.0001 (0.0001)
University Patents Knowledge Stock in Same S&T Area as Entry	0.0011 (0.0008)	0.0021 (0.0101)	0.0136*** (0.0020)	-0.0057 (0.0067)	-0.0156* (0.0076)	0.0105*** (0.0023)
Total Employment in Region/Country (millions of persons)	-0.0000*** (0.0000)	-0.0000*** (0.0000)	-0.0000^ (0.0000)	-0.0000^ (0.0000)	-0.0000* (0.0000)	-0.0000^ (0.0000)
Constant	2.3409*** (0.0445)	1.0949*** (0.6406)	1.1505*** (0.0642)	1.2966*** (0.0586)	1.5624*** (0.0599)	2.8987*** (0.0630)
Dummy = 1 in 2002, else 0	-0.2373^ (0.1443)	-0.5619*** (0.1669)	-0.9431*** (0.2465)	0.2295* (0.1015)	-0.0608 (0.1563)	-0.9973*** (0.2861)
Dummy = 1 in 2003, else 0	-0.3317* (0.1499)	-0.6279** (0.2163)	-1.5847*** (0.2483)	0.0374 (0.1377)	-0.5147** (0.1767)	-1.400*** (0.2524)
Dummy = 1 in 2004, else 0	-0.8708*** (0.1706)	-1.4681*** (0.2600)	-1.8348*** (0.3138)	-0.1053 (0.1133)	-1.0134*** (0.2556)	-2.1298*** (0.2105)
Pseudo R ²	0.5879	0.4592	0.4377	0.5081	0.4444	0.4568

Notes: Robust standard errors in parentheses below coefficient estimates. N = 576. Significance levels: ^ 0.10, * 0.05, ** 0.01, *** 0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

Table A5. Firm Entry into New Technologies – Firm Entries Detected on Stars' Articles & Patents Excluded
Poisson Regressions – U.S. Functional Economic Regions, 1981-2004

Explanatory Variables	Science and Technology Areas of Firm Entry					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Region/ Country in Same S&T Area as Entry	0.0076*** (0.0008)	0.0030 (0.0047)	0.0455*** (0.0140)	0.0048 (0.0059)	0.1700*** (0.0031)	0.0353*** (0.0062)
High Impact Articles Knowledge Stock in Same S&T Area as Entry	-0.0012* (0.0006)	0.0196*** (0.0034)	0.0435* (0.0222)	0.0404*** (0.0055)	0.0037*** (0.0011)	-0.0292*** (0.0069)
University Articles Knowledge Stock in Same S&T Area as Entry	0.0000* (0.0000)	-0.0008*** (0.0001)	-0.0012 (0.0008)	-0.0010*** (0.0001)	-0.0006*** (0.0001)	0.0016*** (0.0004)
University Patents Knowledge Stock in Same S&T Area as Entry	-0.0048*** (0.0009)	-0.0213* (0.0069)	-0.0053 (0.0049)	0.0373*** (0.0095)	0.0049*** (0.0021)	-0.0090** (0.0032)
Total Employment in Region/Country (millions of persons)	0.0978*** (0.0186)	0.1936*** (0.0245)	0.0963*** (0.0232)	0.2399*** (0.0342)	0.1947*** (0.0190)	0.1437*** (0.0170)
Average Wage per Job in Region (thousands of 2000 dollars)	0.1280*** (0.0071)	0.1418*** (0.0068)	0.1995*** (0.0098)	0.1314*** (0.0088)	0.1150*** (0.0083)	0.1210*** (0.0064)
Constant	-1.8924*** (0.1856)	-1.2861*** (0.1734)	-5.1816*** (0.2623)	-3.0134*** (0.2273)	-2.2083*** (0.2223)	-0.9841*** (0.1707)
Dummy = 1 in 2002, else 0	-0.8256*** (0.1109)	-1.1006*** (0.0754)	-0.7353*** (0.1300)	-0.7262 (0.1168)	-0.9360*** (0.1397)	-1.1001*** (0.1124)
Dummy = 1 in 2003, else 0	-1.2731*** (0.1143)	-1.9574*** (0.0953)	-1.4325*** (0.1306)	-0.9287*** (0.1197)	-1.2565*** (0.1212)	-1.7302*** (0.1201)
Dummy = 1 in 2004, else 0	-2.0281*** (0.1362)	-3.3241*** (0.1051)	-2.6212*** (0.1728)	-1.5522*** (0.1378)	-2.2583*** (0.1403)	-3.1067*** (0.1427)
Pseudo R ²	0.6148	0.6330	0.5245	0.5248	0.5237	0.5805

Notes: Robust standard errors in parentheses below coefficient estimates. N = 4296. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

Table A6. Firm Entry into New Technologies – Firm Entries Detected on Stars' Articles & Patents Excluded
Poisson Regressions – Top-25 Science & Technology Countries, 1981-2004

Explanatory Variables	Science and Technology Areas of Firm Entry					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Region/ Country in Same S&T Area as Entry	0.0033*** (0.0001)	-0.0013 (0.0096)	0.0595*** (0.0048)	0.0137* (0.0054)	0.0178*** (0.0037)	0.0246** (0.0082)
High Impact Articles Knowledge Stock in Same S&T Area as Entry	-0.0019*** (0.0001)	-0.0034 (0.0063)	-0.0005 (0.0039)	-0.0033 (0.0027)	-0.0038*** (0.0006)	0.0003 (0.0020)
University Articles Knowledge Stock in Same S&T Area as Entry	0.0000*** (0.0000)	0.0002*** (0.0000)	0.0003*** (0.0000)	0.0000 (0.0000)	0.0002*** (0.0001)	0.0002^ (0.0001)
University Patents Knowledge Stock in Same S&T Area as Entry	-0.0010*** (0.0001)	-0.0097*** (0.0022)	-0.0023*** (0.0004)	-0.0039 (0.0027)	-0.0046*** (0.0005)	-0.0023*** (0.0005)
Total Employment in Region/Country (millions of persons)	-0.0015*** (0.0003)	-0.0013 (0.0002)	0.0000 (0.0004)	-0.0001 (0.0004)	-0.0005 (0.0003)	-0.0004 (0.0003)
Constant	2.7651*** (0.0477)	3.2904*** (0.0706)	1.5902*** (0.1008)	2.1965*** (0.1163)	1.9296*** (0.0998)	3.3105*** (0.0916)
Dummy = 1 in 2002, else 0	-0.2534^ (0.1078)	-0.8858*** (0.1560)	-0.6278*** (0.0989)	0.3351 (0.1803)	-0.2906^ (0.1719)	-0.4265* (0.2057)
Dummy = 1 in 2003, else 0	-0.4536** (0.1388)	-1.1131^ (0.4314)	-1.3859*** (0.1627)	0.4131^ (0.2097)	-0.9210*** (0.2367)	-0.8961*** (0.1905)
Dummy = 1 in 2004, else 0	-2.2415*** (0.5037)	-2.3102** (0.4303)	-1.8297*** (0.4601)	-0.1666 (0.2080)	-2.4620*** (0.4921)	-1.7683*** (0.1969)
Pseudo R ²	0.8801	0.6021	0.7612	0.7705	0.8518	0.8562

Notes: Robust standard errors in parentheses below coefficient estimates. N = 600. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

Table A7. Firm Entry into New Technologies – Firm Entries Detected on Stars' Articles & Patents Excluded
Poisson Regressions – Top-24 Non-U.S. Science & Technology Countries, 1981-2004

Explanatory Variables	Science and Technology Areas of Firm Entry					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Region/ Country in Same S&T Area as Entry	0.0093*** (0.0014)	0.0322*** (0.0076)	0.0419^ (0.0226)	0.0044 (0.0099)	-0.0096 (0.0140)	0.0717*** (0.0187)
High Impact Articles Knowledge Stock in Same S&T Area as Entry	-0.0014* (0.0006)	-0.0079** (0.0038)	0.0306*** (0.0089)	0.0176*** (0.0027)	0.0042 (0.0026)	0.0028 (0.0073)
University Articles Knowledge Stock in Same S&T Area as Entry	0.0000*** (0.0000)	0.0003*** (0.0000)	-0.0000 (0.0001)	0.0000* (0.0000)	0.0002*** (0.0001)	0.0001 (0.0001)
University Patents Knowledge Stock in Same S&T Area as Entry	-0.0003 (0.0009)	-0.0201 (0.0150)	0.0136*** (0.0021)	-0.0070 (0.0078)	-0.0134^ (0.0078)	0.0068** (0.0027)
Total Employment in Region/Country (millions of persons)	-0.0009*** (0.0002)	-0.0011*** (0.0002)	-0.0003 (0.0003)	-0.0004 (0.0003)	-0.0007** (0.0003)	-0.0003 (0.0003)
Constant	2.3185*** (0.0424)	2.8089*** (0.0637)	1.1037*** (0.0610)	1.3611*** (0.0571)	1.5685*** (0.0611)	2.8509*** (0.0565)
Dummy = 1 in 2002, else 0	-0.1762 (0.1568)	-1.1293*** (0.1578)	-0.8827*** (0.2448)	0.1954^ (0.1014)	-0.0859 (0.1509)	-0.7907** (0.2617)
Dummy = 1 in 2003, else 0	-0.2672 (0.1632)	-1.3410*** (0.2514)	-1.5007*** (0.2535)	-0.0524 (0.1448)	-0.5255** (0.1699)	-1.1793*** (0.2531)
Dummy = 1 in 2004, else 0	-0.8054*** (0.1872)	-2.7784*** (0.2577)	-1.7391*** (0.3063)	-0.1210 (0.1124)	-0.9535*** (0.2176)	-2.0366*** (0.2503)
Pseudo R ²	0.5875	0.5378	0.4379	0.5013	0.4424	0.4773

Notes: Robust standard errors in parentheses below coefficient estimates. N = 576 . Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

References

- [1] Azoulay, Pierre, Graff Zivin, Joshua, Wang, Jialan (2010) "Superstar Extinction," *Quarterly Journal of Economics*, May, 125, 549-589.
- [2] Baba, Yasunori, Shichijo, Naohiro, Sedita, Silvia Rita (2009) "How Do Collaborations with Universities Affect Firm' Innovative Performance? The Role of 'Pasteur Scientists' in the Advanced Materials Field," *Research Policy*, June, 38(5), 756-764.
- [3] Darby, Michael R., Zucker, Lynne G. (1999) *California's Science Base: Size, Quality and Productivity*, Sacramento, CA: California Council on Science and Technology.
- [4] Darby, Michael R., Zucker, Lynne G. (2001) "Change or Die: The Adoption of Biotechnology in the Japanese and U.S. Pharmaceutical Industries," *Comparative Studies of Technological Evolution*, 7, 85-125.
- [5] Darby, Michael R., Zucker, Lynne G. (2003) "Growing by Leaps and Inches: Creative Destruction, Real Cost Reduction, and Inching Up," *Economic Inquiry*, January, 41(1), 1-19.
- [6] Darby, Michael R., Zucker, Lynne G. (2005) "Grilichesian Breakthroughs: Inventions of Methods of Inventing in Nanotechnology and Biotechnology," *Annales d'Economie et Statistique*, July-December, 79/80, 143-163.
- [7] Elfenbein, Daniel W., Hamilton, Barton H., Zenger, Todd R. (2010) "The Small Firm Effect and the Entrepreneurial Spawning of Scientists and Engineers," *Management Science*, April, 56(4), 659-681.
- [8] Griliches, Zvi (1990) "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, December, 28(4), 1661-1707.
- [9] Harberger, Arnold C. (1998) "A Vision of the Growth Process," *American Economic Review*, March, 88(1), 1-32.
- [10] Institute for Scientific Information (1981-1997) *Science Citation Index*, machine-readable data bases, Philadelphia: Institute for Scientific Information.
- [11] Institute for Scientific Information (2005) *Science Citation Index Expanded, Social Sciences Citation Index, Arts & Humanities Citation Index, and High Impact Papers*, machine-readable data bases, Philadelphia: Institute for Scientific Information.
- [12] James, William (1911) "Stanford's Ideal Destiny," pages 356-367 in *Memories and Studies*, New York, NY: Longmans, Green & Co.
- [13] Jiang, Lin, Tan, Justin, Thursby, Marie (2011) "Incumbent Firm Invention in Emerging Fields: Evidence from the Semiconductor Industry," *Strategic Management Journal*, January, 32(1), 55-75.

- [14] Johnson, Kenneth P., Kort, John R. (2004) "2004 Redefinition of the BEA Economic Areas," *Survey of Current Business*, November, pp. 68-75.
- [15] Kennedy, Peter (1998) *A Guide to Econometrics*, fourth edition, Cambridge, MA: The MIT Press.
- [16] Kim, Sunwoong (2010) "From Brain Drain to Brain Competition: Changing Opportunities and the Career Patterns of US-Trained Korean Academics," in Charles T. Clotfelter, ed., *American Universities in a Global Market*, Chicago, IL: University of Chicago Press.
- [17] Ma, Hongyan, Furner, Jonathan, Zucker, Lynne G., Darby, Michael R. (2006) "Automatic Selection of Patents and Scientific Articles for NanoBank: An Integrated Probabilistic (IP) Approach," paper for presentation at the ACM & IEEE Computing Joint Conference on Digital Libraries.
- [18] National Research Council (1995) *Research-Doctorate Programs in the United States: Data Set*, machine-readable data base, Washington, DC: National Academy Press.
- [19] Saxenian, AnnaLee (2005) "From Brain Drain to Brain Circulation: Transnational Communities and Regional Upgrading in India and China," *Studies in Comparative International Development*, 40(2), 35-61.
- [20] Tushman, Michael L., Anderson, Philip (1986) "Technological Discontinuities and Organizational Environments," *Administrative Science Quarterly*, March, 31(1), 439-465.
- [21] Wooldridge, Jeffrey M. (1991) "On the Application of Robust, Regression-Based Diagnostics to Models of Conditional Means and Conditional Variances," *Journal of Econometrics*, January, 47(1), 5-46.
- [22] Zucker, Lynne G., Darby, Michael R. (1996) "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, 93(23), 12709-12716.
- [23] Zucker, Lynne G., Darby, Michael R. (1999a) *California's Inventive Activity: Patent Indicators of Quantity, Quality, and Organizational Origins*, California Council on Science and Technology.
- [24] Zucker, Lynne G., Darby, Michael R. (1999b) "Star Scientist Linkages to Firms in APEC and European Countries: Indicators of Regional Institutional Differences Affecting Competitive Advantage," *International Journal of Biotechnology*, 1(1), 119-131.
- [25] Zucker, Lynne G., Darby, Michael R. (2001) "Capturing Technological Opportunity Via Japan's Star Scientists: Evidence from Japanese Firms' Biotech Patents and Products," *Journal of Technology Transfer*, January, 26(1/2), 37-58.
- [26] Zucker, Lynne G., Darby, Michael R. (2006a) "Real Effects of Policy Change in Industry Development in Stem Cells," paper presented at the Stem Cells: Promise and Peril in

Regenerative Medicine Conference, UCLA Center for Society and Genetics, Los Angeles, CA, February 3-4.

- [27] Zucker, Lynne G., Darby, Michael R. (2006b) “Movement of Star Scientists and Engineers and High-Tech Firm Entry,” National Bureau of Economics Research Working Paper No. 12172, April.
- [28] Zucker, Lynne G., Darby, Michael R. (2007) “Socio-economic Impact of Nanoscale Science: Initial Results and NanoBank,” in Mihail C. Roco and William S. Bainbridge, eds., *Nanotechnology: Societal Implications II — Individual Perspectives*, Dordrecht, The Netherlands: Springer.
- [29] Zucker, Lynne G., Darby, Michael R. (2008) “Categorizing Science and Technology across Articles, Patents, Doctoral Programs, and Industries,” machine-readable concordance posted March 3, 2008, Social Science Research Network (SSRN) at <http://ssrn.com/abstract=1109109>. Also at <http://www.nanobank.org/downloads.php> under “Documentation for Zucker-Darby Science Area Categorizations.”
- [30] Zucker, Lynne G., Darby, Michael R., Armstrong, Jeff (1998) “Geographically Localized Knowledge: Spillovers or Markets?” *Economic Inquiry*, January, 36(1), 65-86.
- [31] Zucker, Lynne G., Darby, Michael R., Armstrong, Jeff (2002) “Commercializing Knowledge: University Science, Knowledge Capture, and Firm Performance in Biotechnology,” *Management Science*, January, 48(1), 138-153.
- [32] Zucker, Lynne G., Darby, Michael R., Brewer, Marilyn B. (1998) “Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises,” *American Economic Review*, March, 88(1), 290-306.
- [33] Zucker, Lynne G., Darby, Michael R., Fong, Jason (2014) “Communitywide Database Designs for Tracking Innovation Impact: COMETS, STARS and Nanobank,” *Annals of Economics and Statistics (Annales d’Economie et Statistique)*, this issue.
- [34] Zucker, Lynne G., Darby, Michael R., Furner, Jonathan, Liu, Robert C., Ma, Hongyan (2007) “Minerva Unbound: Knowledge Stocks, Knowledge Flows and New Knowledge Production,” *Research Policy*, July, 36(6), 850-863.
- [35] Zucker, Lynne G., Darby, Michael R., Torero, Máximo (2002) “Labor Mobility from Academe to Commerce,” *Journal of Labor Economics*, July, 20(3), 629-660.

**Movement of Star Scientists and Engineers and High-Tech Firm Entry
Appendix B: Supplementary Material for Electronic Publication Only**

Lynne G. Zucker and Michael R. Darby

Figures B1-B4 compare cumulative active star years (stars) and firm entry (circles) in U.S. and world for biology/chemistry/medicine and nanotechnology. The size of the stars and circles indicate the numbers of each, but note the change in scale in moving from Figure B1 to B2 or B3 to B4.

Tables B1-B6 report our attempts to apply the same models to explain non-university patenting by regions or nations: The first three of these tables report the estimates for the full set of stars as used in the text of the printed article. Tables 4-6 report the results substituting the first 1,838 stars collected as discussed in Appendix A around Tables A1-A4. We believe that these results show insufficient consistency to justify drawing conclusions, but we report them here for those who sat through – and rightly criticized – our efforts to do so in presentations of earlier versions of this paper.

Table B7 provides the interested reader correlation coefficients for variables used in the empirical analysis reported in the main text.

The Web of Science subject category codes, International Patent Classes, and National Research Council doctoral program names corresponding to each of these six science and technology areas are listed in Table B8.

Figure B1. Biology/Chemistry/Medicine Star Scientists & Firm Entry, U.S. Regions, 1981-2004

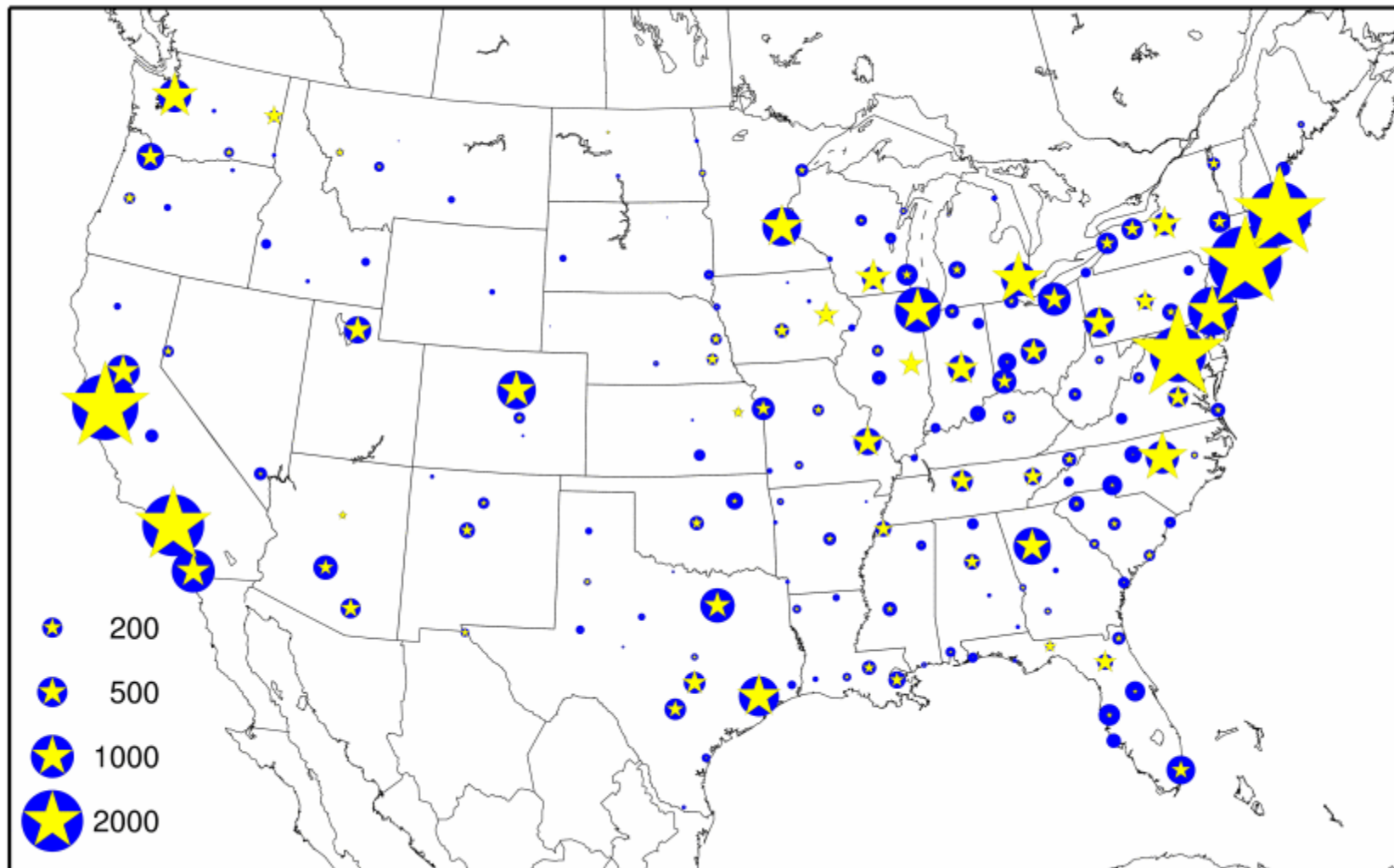


Figure B2. Nanoscale Science and Technology Star Scientists & Firm Entry, U.S. Regions, 1981-2004

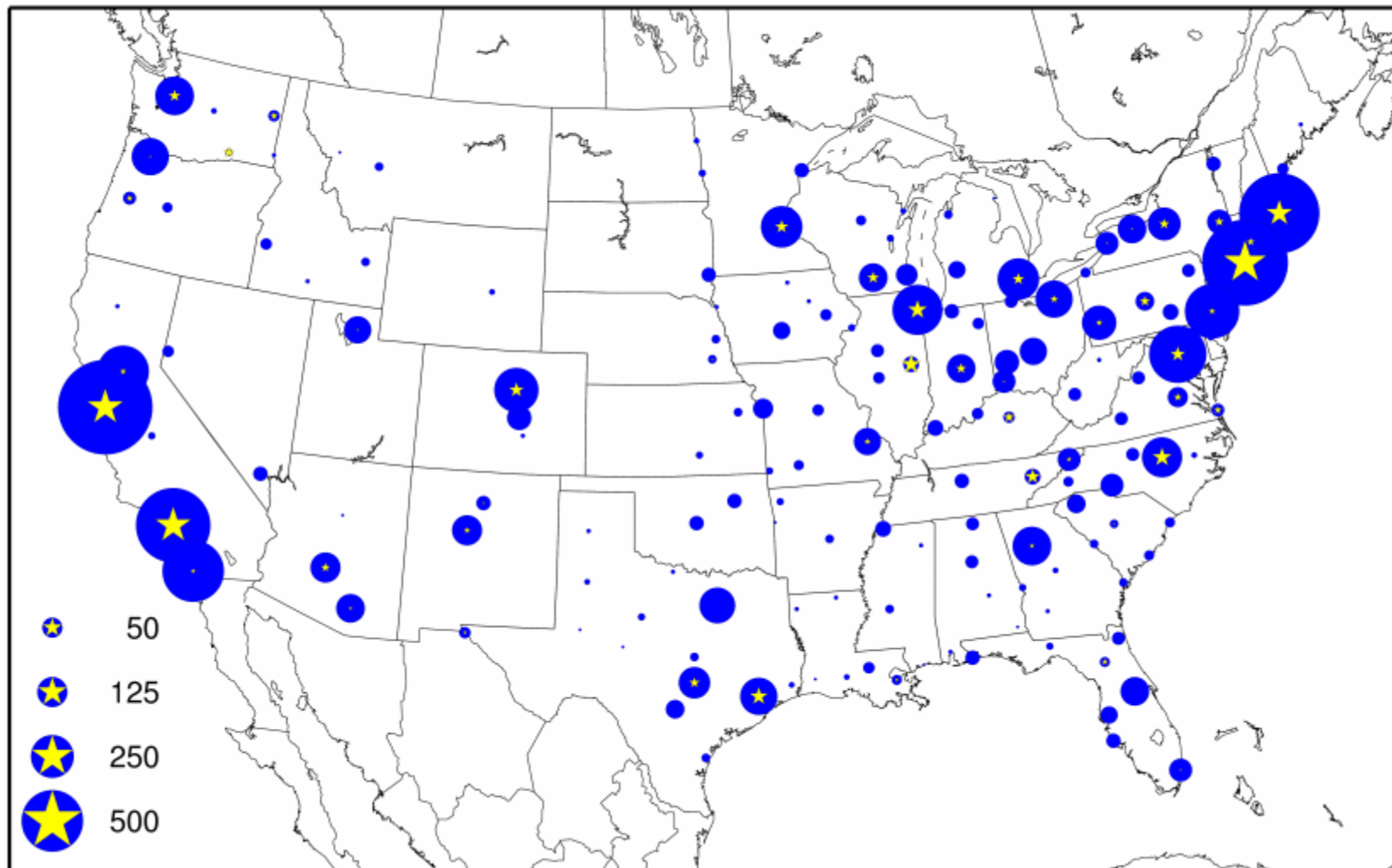


Figure B3. Biology/Chemistry/Medicine Star Scientists & Firm Entry, 25 Countries, 1981-2004

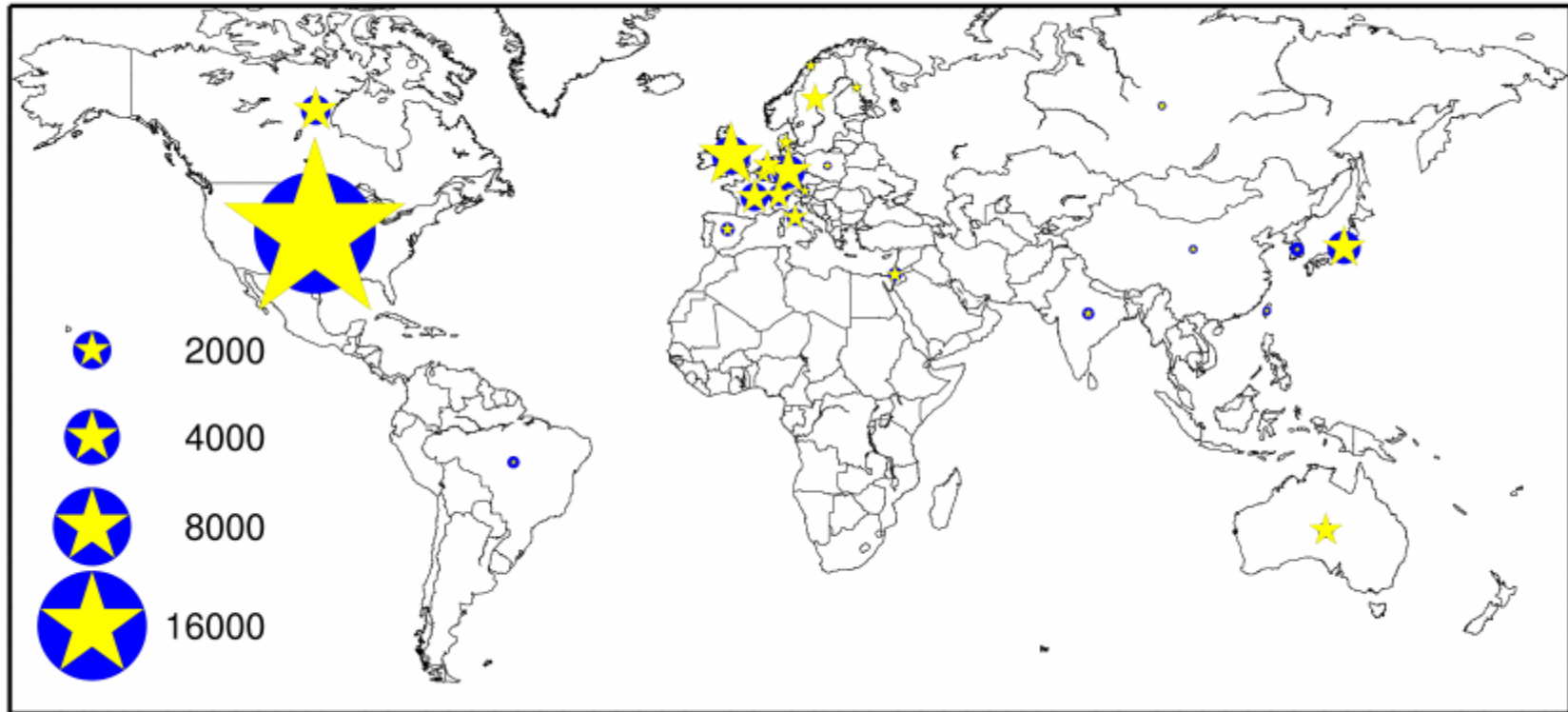


Figure B4. Nanoscale Science and Technology Star Scientists & Firm Entry, 25 Countries, 1981-2004

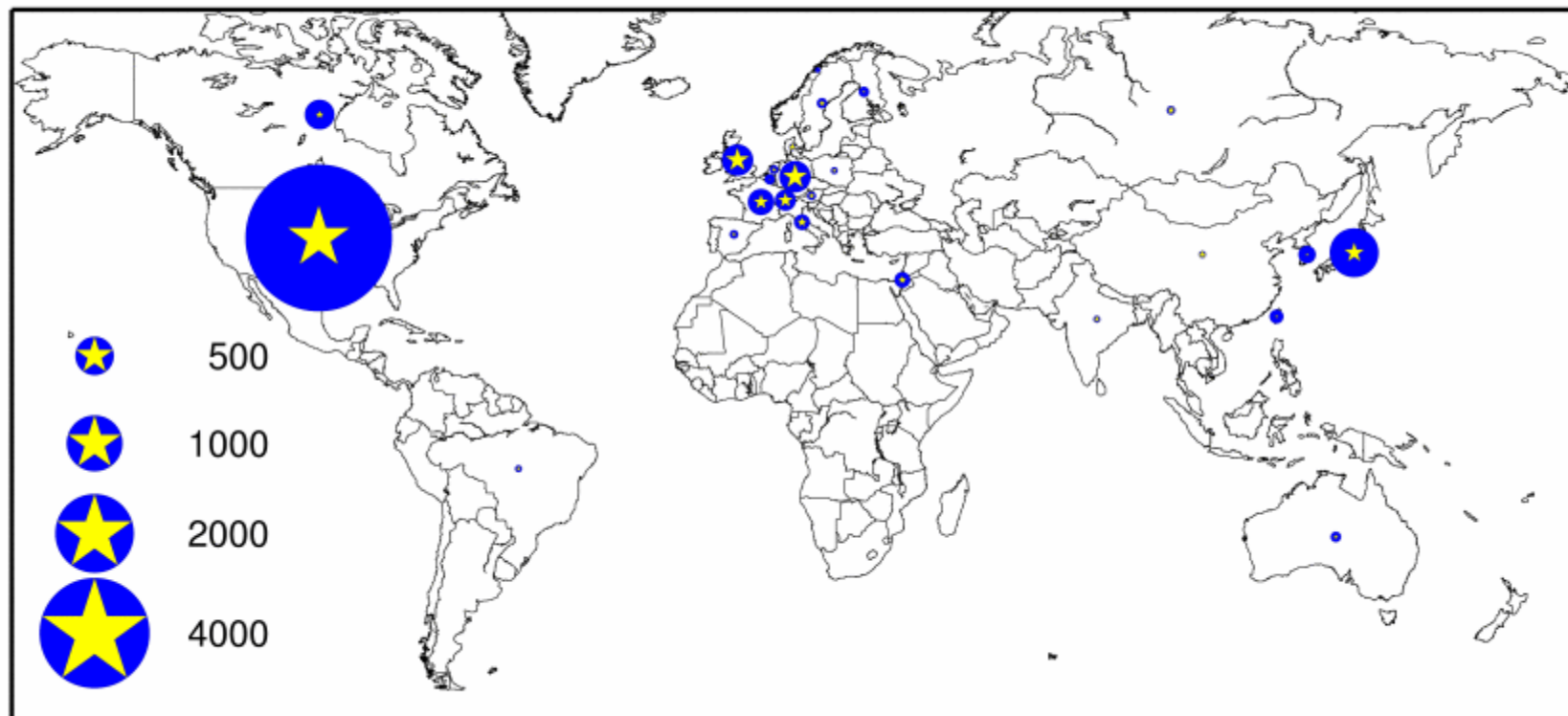


Table B1. Non-University Patenting by Science & Technology Areas – Poisson Regressions
U.S. Functional Economic Regions, 1981-2004

Explanatory Variables	Science and Technology Areas of Patents					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Region in Same S&T Area as Patent	0.0100*** (0.0017)	-0.0452*** (0.0087)	0.0190 (0.0374)	-0.0116* (0.0059)	-0.0952*** (0.0249)	-0.0534** (0.0172)
High Impact Articles Knowledge Stock in Same S&T Area as Patent	-0.0055*** (0.0015)	0.0507*** (0.0047)	0.0120 (0.0188)	0.0245*** (0.0061)	0.0105*** (0.0019)	-0.0376*** (0.0078)
University Articles Knowledge Stock in Same S&T Area as Patent	0.0002*** (0.0000)	0.0005* (0.0002)	0.0018** (0.0007)	0.0002 (0.0002)	0.0002 (0.0003)	0.0054*** (0.0004)
University Patents Knowledge Stock in Same S&T Area as Patent	-0.0113*** (0.0015)	-0.0186** (0.0071)	-0.0030 (0.0036)	0.0319** (0.0115)	0.0166*** (0.0044)	-0.0193*** (0.0016)
Total Employment in Region (millions of persons)	-0.2855** (0.1000)	-0.0882 (0.0698)	-0.0511 (0.0511)	-0.0158 (0.0801)	0.1741 (0.1124)	0.0333 (0.0223)
Average Wage per Job in Region (thousands of 2000 dollars)	0.1677*** (0.0145)	0.2088*** (0.0219)	0.2085*** (0.0161)	0.1282*** (0.0214)	0.0989*** (0.0243)	0.1397*** (0.0223)
Constant	-1.7754*** (0.3780)	-3.6542*** (0.5808)	-4.1676*** (0.4433)	-2.9353*** (0.5384)	-0.9259 (0.5992)	-0.0538 (0.3627)
Dummy = 1 in 2002, else 0	-0.3401 (0.2421)	0.1380 (0.1842)	0.0371 (0.2109)	0.2757 (0.2181)	-0.2128 (0.2473)	-0.2255 (0.1406)
Dummy = 1 in 2003, else 0	-0.6838* (0.3445)	-0.5419* (0.2345)	-0.2762 (0.2572)	0.2475 (0.2286)	0.0229 (0.1934)	-0.2883* (0.1450)
Dummy = 1 in 2004, else 0	1.2214*** (0.3787)	0.0912 (0.2115)	-0.2872 (0.3530)	0.3819 (0.2407)	0.1532 (0.1889)	-0.8200*** (0.2248)
Pseudo R ²	0.6624	0.7559	0.6755	0.7417	0.6088	0.5884

Notes: Robust standard errors in parentheses below coefficient estimates. N = 4296. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

Table B2. Non-University Patenting by Science & Technology Areas – Poisson Regressions
Top-25 Science & Technology Countries, 1981-2004

Explanatory Variables	Science and Technology Areas of Patents					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Country in Same S&T Area as Patent	0.0031*** (0.0003)	-0.0045 (0.0041)	0.0342*** (0.0077)	0.0010 (0.0032)	0.0032 (0.0069)	0.0023 (0.0090)
High Impact Articles Knowledge Stock in Same S&T Area as Patent	-0.0029*** (0.0002)	-0.0062* (0.0027)	-0.0010 (0.0028)	-0.0070*** (0.0020)	-0.0038*** (0.0011)	-0.0060** (0.0023)
University Articles Knowledge Stock in Same S&T Area as Patent	0.0000*** (0.0000)	0.0004*** (0.0000)	0.0006*** (0.0000)	0.0001*** (0.0000)	0.0003*** (0.0001)	0.0005*** (0.0001)
University Patents Knowledge Stock in Same S&T Area as Patent	-0.0014*** (0.0001)	-0.0079*** (0.0015)	-0.0034*** (0.0003)	-0.0083*** (0.0013)	-0.0037*** (0.0006)	-0.0022*** (0.0004)
Total Employment in Country (millions of persons)	-0.0013*** (0.0003)	-0.0011^ (0.0006)	-0.0025*** (0.0007)	-0.0035*** (0.0006)	-0.0002 (0.0003)	-0.0003 (0.0004)
Constant	5.1431*** (0.0629)	4.4491*** (0.1347)	4.0109*** (0.1162)	3.4846*** (0.1289)	4.5586*** (0.1096)	6.6587*** (0.1061)
Dummy = 1 in 2002, else 0	0.0967 (0.1551)	0.0246 (0.2940)	-0.3698^ (0.1900)	0.3867^ (0.2031)	0.2958 (0.2954)	0.2709 (0.2930)
Dummy = 1 in 2003, else 0	-0.0804 (0.1139)	0.8971^ (0.4911)	-1.0543** (0.3415)	0.5924* (0.2358)	0.1549 (0.3737)	0.1851 (0.3179)
Dummy = 1 in 2004, else 0	-2.5380*** (0.4526)	0.2418 (0.4151)	-0.9656** (0.3458)	0.3219 (0.2811)	-0.6482 (0.7875)	-0.2391 (0.3688)
Pseudo R ²	0.8622	0.6993	0.8014	0.7252	0.6370	0.6513

Notes: Robust standard errors in parentheses below coefficient estimates. N = 600. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

Table B3. Non-University Patenting by Science & Technology Areas – Poisson Regressions
Top-24 Non-U.S. Science & Technology Countries, 1981-2004

Explanatory Variables	Science and Technology Areas of Patents					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Country in Same S&T Area as Patent	0.0118*** (0.0019)	-0.0637*** (0.0151)	-0.0244 (0.0315)	0.0431** (0.0157)	-0.1912*** (0.0213)	-0.1309* (0.0520)
High Impact Articles Knowledge Stock in Same S&T Area as Patent	-0.0036*** (0.0007)	-0.0253*** (0.0057)	0.0063 (0.0075)	-0.0205* (0.0090)	0.0143*** (0.0039)	-0.0099 (0.0142)
University Articles Knowledge Stock in Same S&T Area as Patent	0.0000*** (0.0000)	0.0007*** (0.0001)	0.0006*** (0.0000)	0.0002*** (0.0000)	0.0006*** (0.0001)	0.0010*** (0.0001)
University Patents Knowledge Stock in Same S&T Area as Patent	-0.0090*** (0.0013)	0.0658*** (0.0101)	0.0116*** (0.0012)	0.0406*** (0.0063)	-0.0110 (0.0110)	0.0006 (0.0047)
Total Employment in Country (millions of persons)	-0.0013*** (0.0003)	-0.0029*** (0.0006)	-0.0032*** (0.0008)	-0.0042*** (0.0009)	-0.0017*** (0.0004)	-0.0016*** (0.0004)
Constant	4.6216*** (0.0636)	4.2897*** (0.1406)	3.8359*** (0.1071)	3.1383*** (0.1996)	4.0743*** (0.1078)	6.5280*** (0.0871)
Dummy = 1 in 2002, else 0	0.1375 (0.1261)	-0.6026** (0.2231)	-0.5678* (0.2308)	0.1328 (0.3038)	0.0284 (0.1931)	-0.4112 (0.3394)
Dummy = 1 in 2003, else 0	0.1252 (0.1320)	0.1305 (0.3270)	-0.9598*** (0.2692)	0.2010 (0.3270)	-0.0257 (0.2044)	-0.5316 (0.3505)
Dummy = 1 in 2004, else 0	-0.4003** (0.1372)	-0.7148* (0.2900)	-0.9361*** (0.2310)	-0.7544^ (0.3863)	0.2681 (0.2589)	-0.6360 (0.3931)
Pseudo R ²	0.7355	0.6380	0.7380	0.6687	0.5359	0.4184

Notes: Robust standard errors in parentheses below coefficient estimates. N = 576. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

Table B4. Non-University Patenting by Science & Technology Areas – Only First 1,838 Stars Used for Analysis
Poisson Regressions – U.S. Functional Economic Regions, 1981-2004

Explanatory Variables	Science and Technology Areas of Patents					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Region in Same S&T Area as Patent	0.0226*** (0.0042)	-0.1582*** (0.0305)	0.1585*** (0.0405)	0.0200 (0.0135)	-0.0263 (0.0270)	0.0796^ (0.0426)
High Impact Articles Knowledge Stock in Same S&T Area as Patent	-0.0064*** (0.0017)	0.0546*** (0.0053)	-0.0130 (0.0170)	0.0138* (0.0062)	0.0046*** (0.0012)	-0.0639*** (0.0085)
University Articles Knowledge Stock in Same S&T Area as Patent	0.0002*** (0.0000)	0.0005* (0.0002)	0.0026*** (0.0008)	0.0004* (0.0002)	0.0004^ (0.0002)	0.0058*** (0.0003)
University Patents Knowledge Stock in Same S&T Area as Patent	-0.0109*** (0.0015)	-0.0364*** (0.0074)	-0.0045 (0.0031)	0.0210* (0.0104)	0.0057^ (0.0034)	-0.0171*** (0.0017)
Total Employment in Region (millions of persons)	-0.2837** (0.1046)	-0.0542 (0.0675)	-0.1103** (0.0428)	-0.1401* (0.0663)	0.0080 (0.0938)	-0.0053 (0.0193)
Average Wage per Job in Region (thousands of 2000 dollars)	0.1722*** (0.0144)	0.2059*** (0.0221)	0.2129*** (0.0154)	0.1435*** (0.0210)	0.1031*** (0.0205)	0.1315*** (0.0115)
Constant	-1.8911*** (0.3744)	-3.5819*** (0.5818)	-4.2800*** (0.4182)	-3.3132*** (0.5369)	-0.9546^ (0.5063)	0.1305 (0.2944)
Dummy = 1 in 2002, else 0	-0.3273 (0.2408)	0.1145 (0.1771)	0.1363 (0.2007)	0.3506^ (0.2018)	-0.1984 (0.2020)	-0.2024 (0.1365)
Dummy = 1 in 2003, else 0	-0.6598^ (0.3402)	-0.5144* (0.2296)	-0.2590 (0.2690)	0.3504^ (0.2057)	-0.0527 (0.1812)	-0.2612^ (0.1411)
Dummy = 1 in 2004, else 0	-1.2569*** (0.3906)	0.1038 (0.1985)	-0.5283 (0.3984)	0.4015^ (0.2407)	-0.0652 (0.1769)	-0.9810*** (0.2522)
Pseudo R ²	0.6616	0.7552	0.6853	0.7416	0.5953	0.5852

Notes: Robust standard errors in parentheses below coefficient estimates. N = 4296. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

Table B5. Non-University Patenting by Science & Technology Areas – Only First 1,838 Stars Used for Analysis
Poisson Regressions – Top-25 Science & Technology Countries, 1981-2004

Explanatory Variables	Science and Technology Areas of Patents					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Country in Same S&T Area as Patent	0.0058*** (0.0010)	0.0024 (0.0169)	0.0391*** (0.0111)	0.0040 (0.0041)	0.0093 (0.0099)	0.0443** (0.0151)
High Impact Articles Knowledge Stock in Same S&T Area as Patent	-0.0030*** (0.0003)	-0.0088^ (0.0050)	0.0024 (0.0024)	-0.0080*** (0.0018)	-0.0037*** (0.0010)	-0.0150*** (0.0042)
University Articles Knowledge Stock in Same S&T Area as Patent	0.0000*** (0.0000)	0.0003*** (0.0001)	0.0006*** (0.0000)	0.0002*** (0.0000)	0.0003*** (0.0001)	0.0006*** (0.0001)
University Patents Knowledge Stock in Same S&T Area as Patent	-0.0012*** (0.0001)	-0.0075*** (0.0019)	-0.0034*** (0.0003)	-0.0080*** (0.0013)	-0.0036*** (0.0006)	-0.0018*** (0.0004)
Total Employment in Country (millions of persons)	-0.0000*** (0.0000)	-0.0000* (0.0000)	-0.0000** (0.0000)	-0.0000*** (0.0000)	-0.0000 (0.0000)	-0.0000 (0.0000)
Constant	5.1655*** (0.0635)	4.4882*** (0.1466)	4.0602*** (0.1155)	3.4831*** (0.1282)	4.5652*** (0.1077)	6.6368*** (0.0828)
Dummy = 1 in 2002, else 0	0.0925 (0.1550)	0.0673 (0.3025)	-0.5232** (0.1760)	0.3815^ (0.2025)	0.2881 (0.2960)	0.2481 (0.2773)
Dummy = 1 in 2003, else 0	-0.0562 (0.1290)	1.2849** (0.4147)	-1.3249*** (0.3781)	0.5975* (0.2373)	0.1433 (0.3828)	0.2181 (0.3007)
Dummy = 1 in 2004, else 0	-2.5142*** (0.4673)	0.4673 (0.3993)	-1.0230** (0.3568)	0.2524 (0.2918)	-0.6716 (0.7999)	-0.7497 (0.5178)
Pseudo R ²	0.8555	0.6969	0.7907	0.7264	0.6391	0.6697

Notes: Robust standard errors in parentheses below coefficient estimates. N = 600. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

Table B6. Non-University Patenting by Science & Technology Areas – Only First 1,838 Stars Used for Analysis
Poisson Regressions – Top-24 Non-U.S. Science & Technology Countries, 1981-2004

Explanatory Variables	Science and Technology Areas of Patents					
	Bio/Chem/Med	Computing/IT	Nanotechnology	Semiconductors	Other Sciences	Other Engineering
Star Scientists & Engineers Active in Country in Same S&T Area as Patent	0.0183*** (0.0048)	-0.1480*** (0.0334)	-0.1675*** (0.0433)	0.0288 (0.0322)	-0.4189*** (0.0447)	-0.4857*** (0.0577)
High Impact Articles Knowledge Stock in Same S&T Area as Patent	-0.0033*** (0.0010)	-0.0265*** (0.0058)	0.0195* (0.0083)	-0.0118 (0.0090)	0.0173*** (0.0041)	-0.0002 (0.0060)
University Articles Knowledge Stock in Same S&T Area as Patent	0.0000*** (0.0000)	0.0006*** (0.0001)	0.0005*** (0.0001)	0.0002*** (0.0000)	0.0004*** (0.0001)	0.0010*** (0.0001)
University Patents Knowledge Stock in Same S&T Area as Patent	-0.0059*** (0.0013)	0.0674*** (0.0105)	0.0119*** (0.0011)	0.0405*** (0.0064)	-0.0254* (0.0129)	0.0028 (0.0035)
Total Employment in Country (millions of persons)	-0.0000*** (0.0000)	-0.0000*** (0.0000)	-0.0000*** (0.0000)	-0.0000*** (0.0000)	-0.0000* (0.0000)	-0.0000*** (0.0000)
Constant	4.6568*** (0.0694)	4.2791*** (0.1364)	3.8805*** (0.1135)	3.1278*** (0.2031)	4.1023*** (0.1144)	6.3986*** (0.0945)
Dummy = 1 in 2002, else 0	-0.0088 (0.1281)	-0.5639** (0.2158)	-0.5674* (0.2406)	0.0269 (0.3043)	0.0433 (0.1946)	-0.5582** (0.2041)
Dummy = 1 in 2003, else 0	-0.0416 (0.1354)	0.2446 (0.3250)	-1.0276*** (0.2836)	-0.0374 (0.3210)	0.0836 (0.2237)	-0.7308*** (0.2227)
Dummy = 1 in 2004, else 0	-0.5367*** (0.1254)	-0.5884* (0.2782)	-0.8100*** (0.2168)	-0.8070* (0.3971)	0.4740^ (0.2791)	-0.7051** (0.2541)
Pseudo R ²	0.7234	0.6419	0.7492	0.6634	0.5465	0.5230

Notes: Robust standard errors in parentheses below coefficient estimates. N = 576. Significance levels: ^ 0.10, * 0.05, ** 0.01, ***0.001

1. The science and engineering areas are Biology/Chemistry/Medicine; Computing & Information Technology; Semiconductors, Integrated Circuits & Superconductors; Nanoscale Science & Technology; Other Sciences; and Other Engineering. Nanoscale Science & Technology articles & patents as defined for NanoBank.org are removed from the other five areas into which they would otherwise be classified.
2. Knowledge stocks are computed as a perpetual inventory of the indicated series with 20% depreciation applied to the prior year's stock.

Table B7. Correlation Coefficients for Variables Used in Main Empirical Analysis, 1981-2004

	U.S. Regions								Top - 25 Sci & Tech. Countries							Top - 24 Non-US S&T Countries						
	I	II	III	IV	V	VI	VII	VIII	I	II	III	IV	V	VI	VII	I	II	III	IV	V	VI	VII
BIOLOGY/CHEMISTRY/MEDICINE																						
I. Entry of Firms	1								1							1						
II. Non-university Patenting	0.82	1							0.86	1						0.72	1					
III. Stars Scientists and Engineers Active	0.88	0.78	1						0.95	0.94	1					0.86	0.57	1				
IV. High Impact Articles Knowledge Stock	0.85	0.76	0.98	1					0.92	0.94	0.99	1				0.84	0.54	0.96	1			
V. University Articles Knowledge Stock	0.84	0.85	0.89	0.91	1				0.86	0.97	0.95	0.97	1			0.81	0.81	0.78	0.82	1		
VI. University Patents Knowledge Stock	0.77	0.84	0.82	0.84	0.92	1			0.76	0.92	0.88	0.91	0.93	1		0.49	0.27	0.54	0.59	0.60	1	
VII. Total Employment in Region/County	0.86	0.86	0.83	0.82	0.89	0.79	1		0.10	0.11	0.10	0.10	0.11	0.11	1	-0.08	-0.03	-0.12	-0.11	-0.03	-0.01	1
VIII. Average Wage per Job in Region	0.58	0.61	0.55	0.55	0.66	0.62	0.67	1														
COMPUTING/INFORMATION TECHNOLOGY																						
I. Entry of Firms	1								1							1						
II. Non-university Patenting	0.66	1							0.69	1						0.49	1					
III. Stars Scientists and Engineers Active	0.76	0.56	1						0.93	0.70	1					0.58	0.04	1				
IV. High Impact Articles Knowledge Stock	0.78	0.58	0.95	1					0.90	0.73	0.98	1				0.62	0.19	0.78	1			
V. University Articles Knowledge Stock	0.71	0.53	0.87	0.91	1				0.87	0.81	0.93	0.96	1			0.71	0.38	0.60	0.87	1		
VI. University Patents Knowledge Stock	0.66	0.65	0.74	0.78	0.82	1			0.78	0.89	0.84	0.86	0.90	1		0.47	0.63	0.30	0.41	0.56	1	
VII. Total Employment in Region/County	0.67	0.44	0.84	0.80	0.87	0.67	1		0.11	0.10	0.11	0.12	0.16	0.11	1	-0.03	0.01	-0.11	-0.02	0.13	0.01	1
VIII. Average Wage per Job in Region	0.51	0.46	0.52	0.54	0.64	0.54	0.67	1														
NANOTECHNOLOGY																						
I. Entry of Firms	1								1							1						
II. Non-university Patenting	0.63	1							0.64	1						0.51	1					
III. Stars Scientists and Engineers Active	0.72	0.69	1						0.89	0.78	1					0.71	0.42	1				
IV. High Impact Articles Knowledge Stock	0.67	0.72	0.81	1					0.82	0.82	0.97	1				0.71	0.52	0.87	1			
V. University Articles Knowledge Stock	0.64	0.66	0.82	0.93	1				0.70	0.88	0.89	0.94	1			0.58	0.67	0.71	0.88	1		
VI. University Patents Knowledge Stock	0.72	0.81	0.82	0.91	0.88	1			0.76	0.78	0.93	0.96	0.87	1		0.39	0.46	0.38	0.50	0.55	1	
VII. Total Employment in Region/County	0.67	0.52	0.71	0.60	0.64	0.62	1		0.10	0.08	0.10	0.12	0.20	0.10	1	-0.02	0.01	0.00	0.07	0.23	0.00	1
VIII. Average Wage per Job in Region	0.54	0.50	0.48	0.51	0.56	0.55	0.67	1														

(continued)

Table B7. Correlation Coefficients for Variables Used in Main Empirical Analysis, 1981-2004
(concluded)

	U.S. Regions								Top - 25 Sci & Tech. Countries							Top - 24 Non -US S&T Countries						
	I	II	III	IV	V	VI	VII	VIII	I	II	III	IV	V	VI	VII	I	II	III	IV	V	VI	VII
SEMICONDUCTORS																						
I. Entry of Firms	1								1							1						
II. Non-university Patenting	0.48	1							0.57	1						0.49	1					
III. Stars Scientists and Engineers Active	0.80	0.41	1						0.94	0.60	1					0.79	0.44	1				
IV. High Impact Articles Knowledge Stock	0.81	0.44	0.93	1					0.90	0.65	0.98	1				0.85	0.48	0.89	1			
V. University Articles Knowledge Stock	0.76	0.43	0.88	0.93	1				0.81	0.79	0.87	0.92	1			0.79	0.65	0.73	0.82	1		
VI. University Patents Knowledge Stock	0.60	0.49	0.60	0.69	0.72	1			0.77	0.79	0.84	0.85	0.84	1		0.17	0.43	0.13	0.17	0.28	1	
VII. Total Employment in Region/County	0.78	0.41	0.85	0.80	0.83	0.53	1		0.11	0.08	0.12	0.11	0.20	0.12	1	-0.02	0.00	-0.02	-0.04	0.21	0.03	1
VIII. Average Wage per Job in Region	0.54	0.39	0.47	0.49	0.58	0.41	0.67	1														
OTHER SCIENCES																						
I. Entry of Firms	1								1							1						
II. Non-university Patenting	0.81	1							0.79	1						0.60	1					
III. Stars Scientists and Engineers Active	0.80	0.80	1						0.94	0.86	1					0.68	0.19	1				
IV. High Impact Articles Knowledge Stock	0.82	0.79	0.94	1					0.94	0.85	0.98	1				0.79	0.37	0.90	1			
V. University Articles Knowledge Stock	0.82	0.83	0.93	0.93	1				0.90	0.87	0.97	0.98	1			0.77	0.44	0.83	0.91	1		
VI. University Patents Knowledge Stock	0.76	0.90	0.83	0.82	0.83	1			0.84	0.87	0.96	0.95	0.95	1		0.53	0.26	0.64	0.63	0.75	1	
VII. Total Employment in Region/County	0.80	0.82	0.76	0.78	0.83	0.67	1		0.11	0.11	0.11	0.11	0.13	0.12	1	-0.05	0.00	-0.08	-0.07	0.06	-0.02	1
VIII. Average Wage per Job in Region	0.57	0.64	0.50	0.51	0.58	0.55	0.67	1														
OTHER ENGINEERING																						
I. Entry of Firms	1								1							1						
II. Non-university Patenting	0.79	1							0.72	1						0.56	1					
III. Stars Scientists and Engineers Active	0.66	0.58	1						0.90	0.85	1					0.72	0.25	1				
IV. High Impact Articles Knowledge Stock	0.73	0.66	0.93	1					0.89	0.84	0.98	1				0.63	0.29	0.82	1			
V. University Articles Knowledge Stock	0.76	0.78	0.81	0.89	1				0.81	0.89	0.95	0.96	1			0.65	0.50	0.79	0.79	1		
VI. University Patents Knowledge Stock	0.67	0.81	0.67	0.76	0.84	1			0.71	0.87	0.93	0.91	0.94	1		0.51	0.36	0.62	0.45	0.73	1	
VII. Total Employment in Region/County	0.84	0.87	0.70	0.74	0.81	0.71	1		0.10	0.11	0.11	0.10	0.14	0.12	1	-0.05	0.00	-0.10	-0.09	0.07	0.03	1
VIII. Average Wage per Job in Region	0.56	0.70	0.44	0.48	0.59	0.58	0.67	1														

Table B8. Science-Area Concordance: Doctoral Programs, Web of Science Subject Categories & International Patent Codes

Darby-Zucker (1999) Area Name	Corresponding NRC (1995) Doctoral Programs	Corresponding Web of Science Subject Category Codes	Corresponding International Patent Classes
Biology, Chemistry, & Medicine	Biochemistry & Molecular Biology Cell & Developmental Biology Molecular & General Genetics Ecology, Evolution & Behavioral Biomedical Engineering Pharmacology Chemistry Neurosciences Physiology Chemical Engineering	AD,AE,AF,AH,AK,AM,AQ,AY,AZ,BA,BD,CN,CO, CQ,CU,CX,DA,DB,DE,DM,DQ,DR,DS,DW,DX,DY, EA,EC,EE,EI,EY,FF,FI,FQ,FY,GA,GM,GU,HB, HE,HL,HQ,HT,HY,IA,IG,IH,II,IY,JA,JY,KA,KI,KM, LI,LJ,LQ,MA,MU,NE,NI,NN,OI,OO,OP,PT,PW, PY,QA,QB,QU,RQ,RT,RU,RX,RZ,SA,SD,SU,TA, TC,TD,TI,TM,TQ,TU,UH,UM,UY,VE,VY,WC,WE, WF,WH,WV,XE,XW,YA,YO,YP,YU,ZA,ZC,ZD, ZE,ZM,ZR	A 61 B,A 61 C,A 61 D,A 61 F,A 61 G,A 61 H,A 61 J,A 61 K,A 61 L,A 61 M, A 61 N,A 61 P,B 01 J,B 01 L,C 01 B,C 01 C,C 01 D,C 01 F,C 01 G,C 02 F, C 03 B,C 03 C,C 04 B,C 05 B,C 05 C,C 05 D,C 05 F,C 05 G,C 06 B,C 06 C, C 06 D,C 06 F,C 07 B,C 07 C,C 07 D,C 07 F,C 07 G,C 07 H,C 07 J,C 07 K, C 07 M,C 08 B,C 08 C,C 08 F,C 08 G,C 08 H,C 08 J,C 08 K,C 08 L,C 09 B, C 09 C,C 09 D,C 09 F,C 09 G,C 09 H,C 09 J,C 09 K,C 10 B,C 10 C,C 10 F, C 10 G,C 10 H,C 10 J,C 10 K,C 10 L,C 10 M,C 10 N,C 11 B,C 11 C,C 11 D, C 12 C,C 12 F,C 12 G,C 12 H,C 12 J,C 12 L,C 12 M,C 12 N,C 12 P,C 12 Q, C 12 R,C 12 S,C 13 C,C 13 D,C 13 F,C 13 G,C 13 H,C 13 J,C 13 K,C 14 B, C 14 C
Computing & Information Technology	Computer Sciences Mathematics	AC,EP,ER,ES,ET,EV,EW, EX,PE,PN,PQ,RB,XY,YE	G 06 C,G 06 D,G 06 E,G 06 F,G 06 G,G 06 J,G 06 K,G 06 N,G 06 T, G 09 C,G 11 B,G 11 C
Semiconductors Integrated Circuits, High-temperature Superconductors	Physics Electrical Engineering Materials Science Mechanical Engineering	AA,DT,IQ,IU,PJ,PK,PM, PU,PZ,QF,QG,QH,QJ, QM,SR,SY,UB,UE,UF, UI,UK,UN,UP,UR,XQ,ZI	H 01 L
Other Sciences	Oceanography Astrophysics/Astronomy Statistics/Biostatistics Geosciences	GC,ID,JU,KV,KY,LE, OU,PI,QE,QQ,RA,RE, RO,SI,TE	C 30 B,G 01 B,G 01 C,G 01 D,G 01 F,G 01 G,G 01 H,G 01 J,G 01 K, G 01 L,G 01 M,G 01 N,G 01 P,G 01 R,G 01 S,G 01 T,G 01 V,G 01 W, G 02 B,G 02 C,G 02 F,G 21 B,G 21 C,G 21 D,G 21 F,G 21 G,G 21 H, G 21 J,G 21 K,H 01 S
Other Engineering	Aerospace Engineering Civil Engineering Industrial Engineering	AI,BU,FA,IF,IJ,IK,IL,IM, IO,IP,IX,OA,RY,YR,ZQ	All others — see note for current list.

Note: International Patent Classes corresponding to Other Engineering: A 01 B,A 01 C,A 01 D,A 01 F,A 01 G,A 01 H,A 01 J,A 01 K,A 01 L,A 01 M,A 01 N,A 21 B,A 21 C,A 21 D,A 22 B,
A 22 C,A 23 B,A 23 C,A 23 D,A 23 F,A 23 G,A 23 J,A 23 K,A 23 L,A 23 N,A 23 P,A 24 B,A 24 C,A 24 D,A 24 F,A 41 B,A 41 C,A 41 D,A 41 F,A 41 G,A 41 H,A 42 B,A 42 C,
A 43 B,A 43 C,A 43 D,A 44 B,A 44 C,A 45 B,A 45 C,A 45 D,A 45 F,A 46 B,A 46 D,A 47 B,A 47 C,A 47 D,A 47 F,A 47 G,A 47 H,A 47 J,A 47 K,A 47 L,A 62 B,A 62 C,A 62 D,
A 63 B,A 63 C,A 63 D,A 63 F,A 63 G,A 63 H,A 63 J,A 63 K,B 01 B,B 01 D,B 01 F,B 02 B,B 02 C,B 03 B,B 03 C,B 03 D,B 04 B,B 04 C,B 05 B,B 05 C,B 05 D,B 06 B,B 07 B,
B 07 C,B 08 B,B 09 B,B 09 C,B 21 B,B 21 C,B 21 D,B 21 F,B 21 G,B 21 H,B 21 J,B 21 K,B 21 L,B 22 C,B 22 D,B 22 F,B 23 B,B 23 C,B 23 D,B 23 F,B 23 G,B 23 H,B 23 K,
B 23 P,B 23 Q,B 24 B,B 24 C,B 24 D,B 25 B,B 25 C,B 25 D,B 25 F,B 25 G,B 25 H,B 25 J,B 26 B,B 26 D,B 26 F,B 27 B,B 27 C,B 27 D,B 27 F,B 27 G,B 27 H,B 27 J,B 27 K,
B 27 L,B 27 M,B 27 N,B 28 B,B 28 C,B 28 D,B 29 B,B 29 C,B 29 D,B 29 K,B 29 L,B 30 B,B 31 B,B 31 C,B 31 D,B 31 F,B 32 B,B 41 B,B 41 C,B 41 D,B 41 F,B 41 G,
B 41 J,B 41 K,B 41 L,B 41 M,B 41 N,B 42 B,B 42 C,B 42 D,B 42 F,B 43 K,B 43 L,B 43 M,B 44 B,B 44 C,B 44 D,B 44 F,B 60 B,B 60 C,B 60 D,B 60 F,B 60 G,B 60 H,B 60 J,
B 60 K,B 60 L,B 60 M,B 60 N,B 60 P,B 60 Q,B 60 R,B 60 S,B 60 T,B 60 V,B 61 B,B 61 C,B 61 D,B 61 F,B 61 G,B 61 H,B 61 J,B 61 K,B 61 L,B 62 B,B 62 C,B 62 D,B 62 H,
B 62 J,B 62 K,B 62 L,B 62 M,B 63 B,B 63 C,B 63 G,B 63 H,B 63 J,B 64 B,B 64 C,B 64 D,B 64 F,B 64 G,B 65 B,B 65 C,B 65 D,B 65 F,B 65 G,B 65 H,B 66 B,B 66 C,B 66 D,
B 66 F,B 67 B,B 67 C,B 67 D,B 68 B,B 68 C,B 68 F,B 68 G,B 81 B,B 81 C,B 82 B,C 21 B,C 21 C,C 21 D,C 22 B,C 22 C,C 22 F,C 22 K,C 23 C,C 23 D,C 23 F,C 23 G,C 25 B,
C 25 C,C 25 D,C 25 F,D 01 B,D 01 C,D 01 D,D 01 F,D 01 H,D 02 G,D 02 H,D 02 J,D 03 C,D 03 D,D 03 J,D 04 C,D 04 D,D 04 G,D 04 H,D 05 B,D 05 C,D 06 B,
D 06 C,D 06 F,D 06 H,D 06 J,D 06 L,D 06 M,D 06 N,D 06 P,D 06 Q,D 07 B,D 21 B,D 21 C,F 23 D,F 23 G,F 23 H,F 23 J,F 23 K,F 23 L,F 23 M,F 23 N,F 23 Q,F 23 R,F 24 B,
E 02 B,E 02 C,E 02 D,E 02 F,E 03 B,E 03 C,E 03 D,E 03 F,E 04 B,E 04 C,E 04 D,E 04 F,E 04 G,E 04 H,E 05 B,E 05 C,E 05 D,E 05 F,E 05 G,E 06 B,E 06 C,E 21 B,E 21 C,
E 21 D,E 21 F,F 01 B,F 01 C,F 01 D,F 01 K,F 01 L,F 01 M,F 01 N,F 01 P,F 02 B,F 02 C,F 02 D,F 02 F,F 02 G,F 02 H,F 02 J,F 02 K,F 02 M,F 02 N,F 02 P,F 03 B,F 03 C,F 03 D,F 03 G,F 03 H,
F 04 B,F 04 C,F 04 D,F 04 F,F 15 B,F 15 C,F 15 D,F 16 B,F 16 C,F 16 D,F 16 F,F 16 G,F 16 H,F 16 J,F 16 K,F 16 L,F 16 M,F 16 N,F 16 P,F 16 S,F 16 T,F 17 B,F 17 C,F 17 D,
F 21 H,F 21 K,F 21 L,F 21 S,F 21 V,F 21 W,F 21 Y,F 22 B,F 22 D,F 22 G,F 23 B,F 23 C,F 23 D,F 23 G,F 23 H,F 23 J,F 23 K,F 23 L,F 23 M,F 23 N,F 23 Q,F 23 R,F 24 B,F 24 C,
F 24 D,F 24 F,F 24 H,F 24 J,F 25 B,F 25 C,F 25 D,F 25 J,F 26 B,F 27 B,F 27 D,F 28 B,F 28 C,F 28 D,F 28 F,F 28 G,F 41 A,F 41 B,F 41 C,F 41 F,F 41 G,F 41 H,F 41 J,F 42 B,
F 42 C,F 42 D,G 03 B,G 03 C,G 03 D,G 03 F,G 03 G,G 03 H,G 04 B,G 04 C,G 04 D,G 04 F,G 04 G,G 05 B,G 05 D,G 05 F,G 05 G,G 07 B,G 07 C,G 07 D,G 07 F,G 07 G,
G 08 B,G 08 C,G 08 G,G 09 B,G 09 D,G 09 F,G 09 G,G 10 B,G 10 C,G 10 D,G 10 F,G 10 G,G 10 H,G 10 K,G 10 L,H 01 B,H 01 C,H 01 F,H 01 G,H 01 H,H 01 J,H 01 K,
H 01 M,H 01 P,H 01 Q,H 01 R,H 01 T,H 02 B,H 02 G,H 02 H,H 02 J,H 02 K,H 02 M,H 02 P,H 03 B,H 03 C,H 03 D,H 03 F,H 03 G,H 03 H,H 03 J,H 03 K,H 03 L,H 03 M,
H 04 B,H 04 H,H 04 J,H 04 K,H 04 L,H 04 M,H 04 N,H 04 Q,H 04 R,H 04 S,H 05 B,H 05 C,H 05 F,H 05 G,H 05 H