ANNALS OF ECONOMICS AND STATISTICS
NUMBER 115/116, DECEMBER 2014

Exploring Creative Research in Emerging Research Domains. Worldwide Longitudinal Evidence for Nanoscale Science and Technology

Thomas HEINZE
Bergische Universität Wuppertal

Gerrit BAUER
Ludwig Maximilians Universität München

This paper explores research capabilities in nanoscale science and technology. We introduce a method for sampling a population of creative scientists based on two measures: network brokerage and multivocality. Subsequent data exploration shows how these scientists are distributed internationally, and how their creativity is connected to publishing and patenting behavior, institutional affiliation, availability of funding, and job mobility.*

I. Introduction

Science and technology based on unified concepts of matter at the nanoscale provide a new foundation for knowledge creation, technology development, and innovation. Most important, nanoscale science and technology ("nanoscale S&T") is widely anticipated as one of the key drivers of technology-based business and economic growth. As in other high-technology domains, such as biotechnology, economic opportunities abound with progress in scientific research conducted in both public and private laboratories. Yet, little is known about the distribution of these research capabilities. In which countries, regions, and organizations are innovations in basic nanoscale research being accomplished? Do creative scientists differ from their peers with respect to productivity, patenting behavior, or access to research funding? How does job mobility affect scientists' creativity? These are the questions that this paper sets out to answer. Several other studies have examined the mechanisms by which commercial activities spring from innovative basic research, including labor mobility of star scientists (ZUCKER, DARBY, and TORERO [2002]), the mechanisms by which interorganizational networks grow, including cumulative advantage and homophily (POWELL *et al.* [2005]), or technology transfer mechanisms in the university setting (COLYVAS [2007]). This paper, however, investigates where creative scientists are located, how the global distribution of creative potential has changed over time, and the mechanisms that explain the differences between creative scientists and their peers.

To begin with, nano S&T domain developed from a number of fundamental breakthroughs in diverse research fields. One such breakthrough in applied physics was a new type of spectroscopy

*JEL: Z18, I23, I28, L31, L38, O31, O38 / KEY WORDS: Creativity, Innovation, Nanotechnology, Job mobility, Network Analysis, Patents, Citations

based on quantum mechanics. Both the scanning tunnel microscope (STM) and the atomic force microscope (AFM) (BINNIG and ROHRER [1982]BINNIG and ROHRER [1986]) attain extremely high-quality resolution at the atomic level either in conductive materials (STM) or non-conductive matter (AFM). Other fundamental breakthroughs were accomplished in the field of inorganic chemistry with the synthesis of two new carbon materials: carbon nanoballs (HEATH *et al.* [1985]KROTO *et al.* [1985]) and carbon nanotubes (IIJIMA [1991]IIJIMA, AJAYAN, and ICHIHASHI [1992]). These carbon structures have interesting chemical and physical properties related to conductivity and stiffness. Recent developments building on these new materials are nanotube transistors at room temperature (TANS, VERSCHUEREN, and DEKKER [1998]) and nanotube-based circuits (COLLIER *et al.* [1999]).

Several of these breakthroughs won Nobel Prizes, and all of them are associated with outstanding scientists and their research groups (HEINZE *et al.* [2013]). Most important, these accomplishments opened up new research venues and even whole research fields that both build upon and extend the knowledge paths generated by the initial discoveries. In other words, the knowledge production emerging around these breakthroughs is cumulative (ZUCKER *et al.* [2007]), a fact that is also well known from studies on technological discontinuities (ANDERSON and TUSHMAN [1990]). The fact that production of scientific knowledge following initial breakthroughs is cumulative has two implications. First, there is a multitude of further improvements in theory, method, and instrumentation (HEINZE *et al.* [2007]). Second, there are significant capabilities in the research system from which these subsequent improvements arise. In this paper, we are concerned with the second implication. We identify and explore creativity using the contents of NanoBank, an open-access digital library in the field of nano S&T (ZUCKER *et al.* [2007]).

We approach creativity at the level of several thousand scientists over a period of fiveteen years (1990-2004). By means of NanoBank content, we delineated two populations of scientists using information on their network structure and publication patterns: a creative population, and a control population. The two populations are defined based on SIMONTON's [2004] probabilistic-evolutionary perspective on research outcomes and BURT's [2004] perspective of network positions as opportunity structures for novel ideas. Subsequently, these populations are examined with respect to several independent variables, such as productivity, national and institutional affiliation, availability of funding, and job mobility.

The next section introduces definitions of creativity and explains how the population of creative scientists and the control population are delineated (SECTION II). We then present data showing significant differences between these two populations. Variables include productivity measures, such as publication and patent counts, national and organizational affiliations, funding patterns, and job mobility. These findings are synthesized using multivariate panel regression models (SECTION 3). Furthermore, we show that the two populations differ markedly with respect to defining research frontiers. Using text-mining techniques, we examine the semantic spaces of these two populations (SECTION IV). The concluding section (SECTION V) briefly summarizes our findings.

II. Definition and Delineation of Creativity in Nano S&T

Scientific creativity is typically defined in terms of knowledge and capabilities that are new, original, surprising, and useful (HEINZE [2013]; HOLLINGSWORTH [2004]; LIGHTMAN [2005]; SIMONTON [2004]). In science, as in other fields, standards and norms are established against which claims for innovative contributions are assessed. However, science, more than other fields, has evolved procedures, disciplines, and institutions to accredit new knowledge (WHITLEY [2000]). There are varied approaches to examining and empirically measuring creativity, such as identifying creative individuals, the products or outcomes of creative work, creative processes, and creative knowledge environments (HEMLIN, ALLWOOD, and MARTIN [2004]; STUMPF [1995]).

Creative research has most frequently been identified either by research breakthroughs as manifested with the awarding of prestigious scientific prizes (HOLLINGSWORTH [2002]; ZUCKERMAN [1977]), or as productivity and citation measures (FLEMING and SZIGETY [2006]; SIMONTON [2004]). Also, creative research accomplishments were established by combining survey nominations made by several hundred experts with data on prize winners (HEINZE *et al.* [2007]). There have also been various attempts to explain creative outcomes, either within a probabilistic-evolutionary framework (SIMONTON [1999; 2004]) or by reference to network positions (BURT [2004]; FLEMING, SANTIAGO, and CHEN [2007]; UZZI and SPIRO [2005]) and organizational variables (ALLISON and LONG [1990]; HEINZE *et al.* [2009]; HOLLINGSWORTH [2004]).

A well-known approach to assessing scientific creativity through outcomes is publication and citation analysis within an evolutionary-probability theoretical framework (SIMONTON [1999; 2004]). Simonton argues that scientists who are highly productive in publishing papers encounter a greater likelihood that one or more of their papers will come to the attention of other scientists, be cited, and be recognized as creative. In other words, the more contributions to knowledge that a scientist produces, the higher the chances are that one of these contributions resonates well in the scientific community.

Another well-known approach highlights the opportunity structures in collaboration networks that facilitate the generation of novel ideas. BURT [2004] argues that people who are placed at the intersection of heterogeneous social groups have an increased likelihood of drawing upon multiple knowledge sources, leading to the generation of new ideas. This argument of multivocality has been successfully applied to managerial performance and innovativeness of companies (RODAN and GALUNIC [2004]; ZAHEER and BELL [2005]).

Both approaches have stimulated considerable debate. Regarding Simonton's argument, some contend that creative scientists sometimes publish only a few papers, and the relationship between citations and quality is not straightforward (AKSNES [2006]). Furthermore, while productive people are more likely to generate more ideas, they tend to favor the exploitation of familiar knowledge at the expense of exploration of new domains (AUDIA and GONCALO [2007]). Regarding Burt's argument, some contend that although brokering inventors are more likely to generate new ideas, the brokered network structure itself is less suited to diffusing these ideas (FLEMING, SANTIAGO, and CHEN [2007]).

In reference to this ongoing debate, we connect Burt's structural position argument with Simonton's evolutionary-probabilistic perspective. In a first step, we delineate a population of nano scientists (and a control population) in the NanoBank dataset by using two variables that capture Burt's argument: network brokerage and multivocality. Clearly, these two variables do not measure particular creative accomplishments. Rather, they describe characteristics in the work styles of individual scientists that – according to Burt's argument – are associated with creative accomplishment. Therefore, this population consists of *potentially creative* scientists. We theorize that if scientists have certain values in these two variables, they are more or less likely to accomplish work that is judged by their peer scientists as creative. In a second step, we test the validity of Simonton's argument by comparing the productivity of the *creative population* with that of the *control population.*

The delineating procedure is as follows: Starting with the entire NanoBank population of scientists in the years 1990-2004, we selected those authors publishing more than one paper in at least three of the following time periods: 1990-1992, 1993-1995, 1996-1998 1999-2001, and 2002-2004. For each scientist and for all time periods, network brokerage and multivocality were computed. The *creative population* (P1) was derived from combining the highest 25% of both variables (14,851 authors), while the *control population* (P2) was delineated by combining the lowest 25% for both variables (18,408 authors).

Network brokerage measures the extent to which a scientist connects his or her peers. The brokerage index is computed as follows: the number of possible ties between peers without the focal scientist (=$(n^2$-n)/2) minus the number of factual ties between peers without the focal scientist (=h_{ij}), divided by the number of possible ties (=$(n^2$-n)/2). If a scientist connects many of his peers, then h_{ij} is small, and the brokerage index high. If a scientist connects a few peers, then h_{ij} is large, and the brokerage index low.

$$\text{Equation 1: NETWORK BROKERAGE} = \frac{\frac{n^2-n}{2} - \sum_{i=1}^{n}\sum_{j=1}^{m} h_{ij}}{\frac{n^2-n}{2}}$$

Multivocality measures the publishing behavior of P1 and P2 scientists across journals listed in NanoBank (based upon Web of Science). The index combines the number of journals in which scientists publish and the concentration of their publications across these journals. The multivocality index increases when scientists publish in different journals, but it decreases when they publish most of their work in a few journals. Consequently, if scientists publish across various journals but show a high concentration in a few, they receive lower values than those scientists with a more equal distribution in their publishing activity. This relationship is stated in the formula below, where G denotes the Gini coefficient and x_i represents the frequency with which scientists publish in *n* different journals:

$$\text{Equation 2: MULTIVOCALITY} = (1-G) \times \sum_{i=1}^{n} x_i$$

The combination of the upper and lower quartiles of the two distributions generates P1 representing 11% and P2 representing 14% of the total NanoBank population of the years 1990-2004. These percentages are relatively constant over time. Combining the tails of the two distributions means that both indizes equally contribute to the delineation of P1 and P2. One might question whether the equal contribution is justified because the two variables measure different things. We address this argument by running multiple panel regression models

(SECTION III) to test whether the highest 25% of network brokerage (33,249 scientists) or the highest 25% of multivocality (33,205 scientists) is associated differently with outcome measures compared to the P1 population (14,851 scientists).

In the next section, we explore differences between these two populations, P1 and P2, using variables such as publication and patent counts, national and organizational affiliations, availability of funding, and job mobility. These findings are then synthesized in multivariate panel regression models.

III. Explorative Analysis of the Creative Population (P1)

According to SIMONTON [2004], creativity in science is a probabilistic consequence of productivity. Therefore, P1 scientists should have a significantly higher publication output than their P2 peers. In fact, the NanoBank data show that P1 scientists increased their average publication output significantly between 1990 and 2004 from 10 to 28 papers on average, while P2 scientists published not only much less than P1, but also their publication output remained fairly constant over time: fewer than 3 papers on average per period (TABLE I). It seems that the considerable publication growth in the nano S&T domain, as documented by HULLMANN and MEYER [2003], p. 510 and ZUCKER and DARBY [2005], p. 26 can be attributed significantly to P1.

One may argue that our delineation of P1 using the upper quartiles of brokerage and multivocality automatically selects people with a high publication record. Although we agree that people with many publications have more opportunities to broker their colleagues and to publish across a wide range of journals than people with fewer publications, publication is linked to brokerage and multivocality in a probabilistic fashion. First, because there are prolific scientists who connect few colleagues and who publish within a narrow range of specialized journals (in P2). Second, because there are scientists with relatively low publication output but high values in brokerage and multivocality (in P1). In addition, as we show below: although the publication variable is an important predictor of P1 versus P2, other institutional variables are influential as well (TABLE IX).[1]

TABLE I. – Mean Number of SCI Publications of P1 and P2

	1990–1992	1993–1995	1996–1998	1999–2001	2002–2004
P1	9.57 (0.22)	12.76 (0.21)	15.12 (0.21)	20.41 (0.31)	27.73 (0.48)
P2	2.86 (0.03)	2.94 (0.03)	2.42 (0.01)	2.54 (0.02)	2.69 (0.02)
N (P1)	1279	2407	3763	3726	3674
N (P2)	1516	2258	6294	4366	3974

Notes: Missing data not displayed, standard error in brackets.
P1 = creative population. P2 = control population. SCI = Science Citation Index.

1. In addition, one may ask if senior researchers are over-represented among P1 scientists, since they have a longer career and had time to publish more than young (not yet creative) researchers. An answer to this question would be possible with demographic data. However, the current version of NanoBank data does not contain such data.

TABLE II. – Country Shares in P1 and P2, in Percent

	1990–1992			2002–2004		
	Total	P1	P2	Total	P1	P2
JP	16.6	32.3	12.6	15.1	24.8	8.9
US	32.6	29.7	31.4	23.2	22.0	25.7
GER	10.2	9.4	8.2	9.4	7.7	9.6
CN	2.9	2.7	3.2	7.9	16.6	2.3
GB	6.2	4.8	5.7	5.1	3.3	5.7
RU	4.9	3.4	9.2	3.1	0.6	6.7
FR	5.5	2.5	5.3	5.5	2.8	7.3
CA	2.7	2.9	2.2	2.2	1.9	3.5
IT	2.8	2.2	4.2	3.8	1.0	5.7
CH	2.0	1.2	1.7	1.5	1.1	1.2
TW	0.5	0.6	0.3	1.9	4.1	0.5
Other	13.5	8.8	16.2	23.2	18.1	23.2
Total	100.0	100.0	100.0	100.0	100.0	100.0
N	13041	1279	1516	31824	3674	3974

Notes: Missing data not displayed. JP = Japan, US = United States, GER = Germany, CN = China, GB = Great Britain, RU = Russia, FR = France, CA = Canada, IT = Italy, CH = Switzerland, TW = Taiwan. P1 = creative population, P2 = control population.

Interesting patterns emerge when publication output is analyzed at the country level (TABLE II). In 1990-1992, more than 70% of P1 scientists were nationals from three countries with strong national R&D capacities: Japan, the United States, and Germany. In contrast, the United Kingdom, Russia, and France were underrepresented in this league of highly productive scientists from the beginning and remained in that position over time. More than a decade later, however, the top three countries were Japan, the United States, and China. Although Japan and the United States remained strongly represented in the P1 population, both lost roughly 25% of their share each, while at the same time China emerged as a new hub of scientific productivity. Its share in P1 sixtupled during the 15-year observation period. Another newcomer was Taiwan, which had higher shares than the United Kingdom and France in 2002-2004. Consequently, the global loci of both research productivity have shifted somewhat from North America and Europe to Asia, and in particular to China.

As in other high-technology domains, such as biotechnology, academic publications are the typical output of scientific research. However, following worldwide deregulation in the public research sector with regard to intellectual property rights, patents are now also recognized in bibliometrics as a major output of scientific research (MOED, GLÄNZEL, and SCHMOCH [2004]; NOYONS *et al.* [2003]). Consequently, we should also expect a significantly higher patent output of P1 scientists compared with their P2 peers. Indeed, our observations in the NanoBank dataset confirm this expectation. P1 scientists had significantly more granted patents and patents per publication compared to their P2 peers, indicating the intertwined relationship between science and technology dynamics within the P1 population (TABLE III). Furthermore, P1 scientists received consistently higher scores for patent claims and patent citations (both relative to patent output) over time.[2] Higher patent citation scores indicate that subsequent inventions drew much more frequently on patents from P1 scientists than on patents originating

2. The patent count variable measures "forward citations," i.e., references that other patents give to the focal patent in the years following the grant year.

TABLE III. – Mean Number of Patents, Patent Claims, and Patent Citations of P1 and P2

	1990–92		1993–95	
	P1	P2	P1	P2
npat	3.24 (0.18)	0.25 (0.03)	3.14 (0.14)	0.19 (0.02)
(npat/npub)*100	34.55 (1.82)	9.67 (1.21)	24.02 (0.94)	7.77 (0.82)
nclaims/npat	7.32 (0.30)	1.74 (0.16)	7.54 (0.21)	1.41 (0.12)
ncit/npat	0.11 (0.01)	0.01 (0.00)	1.99 (0.09)	0.11 (0.01)
N	1279	1516	2407	2258

Notes: Missing data not displayed, standard error in brackets. Patent counts are zero for periods after 1993-1995. P1 = creative population, P2 = control population. Mean differences between P1 and P2 are significant for all variables at the $p < 0.001$ level.

from P2 scientists. Consequently, patents granted to P1 scientists were more useful and fruitful for subsequent inventions. Therefore, P1 patents tended to reflect higher technical and, perhaps, economic value than patents by the P2 population.[3]

Patent analyses at the country level showed that the emergence of China and Taiwan as new publishing hubs was not mirrored with respect to patents. This might be attributed to the fact that NanoBank captures USPTO patents only; however, the relationship between the worlds of science and technology is closest in Japan, followed by the United States and Germany. P1 scientists based in Japan (not necessarily Japanese citizens) had on average 85.5 (1990-1992) and 59.1 (1993-1995) patents per 100 publications (afterwards declining rapidly), while scientists based in the United States had, on average, between 12.7 (1990-1992) and 12.8 (1993-1995) patents per 100 publications; and scientists based in Germany had values of 11.1 (1990-1992) and 8.0 (1993-1995), respectively. Therefore, two major competitors of the United States are filing patents at the United States Patent and Trademark Office (USPTO), indicating the high commercial expectations from their inventions. Furthermore, Japanese and German scientists are motivated to seek patents in the U.S. because these countries are important trade and foreign direct investment partners of the United States. In contrast, commercial expectations linking science and technology cannot (yet) be observed for China and Taiwan. Another interesting finding is that patents originating in Japan claimed significantly more intellectual property protection than average, and they received significantly higher citation scores (6 times more than US, 13 times more than Germany). Thus, the data reflect Japan's key technological position in the early stage of nano S&T field formation. Also, given that Japan has a higher share than other countries of nano scientists associated with companies, these scientists might seek broader patent scope, and their inventions tend to be more applied in nature.

The various institutions engaged in scientific research contribute differently to the development of the nano S&T field. Assuming that the academic heartland is institutionalized

3. It is also possible that higher patent citation scores mean that the patents generated by P1 scientists tend to relate to a broader range of technological areas upon which other inventors can build. However, we cannot test this claim with the current version of NanoBank data.

TABLE IV. – Institutional Affiliations of P1 and P2 Scientists, in Percent

	1990–1992			2002–2004		
	Total	P1	P2	Total	P1	P2
Company	11.8	15.4	7.7	4.3	4.8	5.9
University	64.4	67.8	68.3	70.9	73.0	66.3
Research institute	14.4	11.4	13.0	14.0	13.6	13.4
National laboratory	1.9	0.8	2.0	1.7	1.6	2.5
US government	1.6	0.6	1.5	1.2	0.8	1.3
Other	5.9	3.9	7.4	7.9	6.3	10.7
N	7364	970	783	21460	3401	1834

Notes: Missing data not displayed. P1 = creative population, P2 = control population.

in universities and other government-funded research organizations, one would expect these organizations to be strongly represented in the P1 population, while the share of firms would be relatively low in P1. In general, these expectations are confirmed. We found that firms had below-average representation in the P1 population, whereas scientists from public institutions had above-average representation in the P1 population (TABLE IV).

However, the comparison of organizational categories between the P1 and P2 populations revealed conspicuous institutional transformations over time. For example, in the early 1990s, firms were (despite their low values) represented twice as much in P1 (15.4%) as in P2 (7.7%), while universities and other research institutes had similar shares in both populations (TABLE IV). Yet, the company share in P1 had dropped significantly by 2002-2004 (4.8%), while their representation in P2 remained relatively stable (5.9%). Inspection of all five observation periods shows that firms had already lost their grip in P1 in 1993-1995 but that it took more than a decade for firm representation in P2 to exceed that of P1. We did not observe similar patterns for research institutes or other categories.

On the country level, the decreasing share of firms in the P1 population between 1990-1992 and 2002-2004 was most pronounced in Japan (from 29.5% to 10.0%), followed by the United States (from 14.6% to 7.7%), and Germany (from 5.5% to 2.3%) (TABLE V).

These results indicate that institutional change in national research systems has been profound. Whereas explorative research was commonplace in corporate R&D laboratories until the mid-1990s, particularly in Japan and the United States, today such research plays a much smaller role compared to application and development activities. The fact that the P2 shares of both Japan and the United States have remained relatively stable over time indicates that although corporate leaders have considerably reduced basic research capacities, they have apparently not reduced investments in more applied R&D activities (TABLE V). Our findings corroborate evidence that fundamental research labs of large, leading industrial companies were a magnet for creative scientists until the mid-1990s, while due to changes in market conditions and industry strategies, many of these labs no longer exist today, and some that do remain no longer allow the sort of work that made them an attractive place for fundamental research (HEINZE *et al.* [2009]; MUNARI, ROBERTS, and SOBRERO [2002]).

TABLE V. – Institutional Affiliations of P1 and P2 Scientists, in Percent

P1	1990–1992			2002–2004		
	JP	US	GER	JP	US	GER
Company	29.5	14.6	5.5	10.0	7.7	2.3
University	59.9	78.4	71.4	66.9	79.5	68.5
Research institute	8.3	2.1	22.0	20.8	—	25.4
National laboratory	—	2.8	—	—	7.1	—
US government	—	2.1	—	—	3.6	—
Other	2.2	—	1.1	2.3	2.1	3.8
N	312	287	91	840	745	260
P2						
Company	19.6	11.6	7.9	16.1	11.4	5.8
University	74.2	75.2	60.3	62.1	69.7	64.2
Research institute	5.2	1.2	28.6	19.3	1.1	26.0
National laboratory	—	6.6	—	—	9.7	—
US government	—	5.0	—	—	4.9	—
Other	1.0	0.4	3.2	2.6	3.2	4.1
N	97	242	63	161	465	173

Notes: Missing data not displayed. P1 = creative population, P2 = control population.

The emergence of the nano S&T domain builds on the availability of substantial R&D resources, both public and private. STM and AFM facilities, but also other laboratory systems are very expensive. Therefore, several countries launched dedicated nano S&T funding programmes in the late 1990s and early 2000s to meet these investment needs and to ensure international competitiveness of their research institutions. Total annual funding for 2004 has been estimated for the following countries. United States: 1.243 Mio €, Japan: 750 Mio €, European Commission: 370 Mio €, Germany: 293 Mio €, France: 224 Mio €, United Kingdom: 133 Mio €, and China: 83 Mio € (EUROPEAN COMMISSION [2005]).

New resources are typically not distributed equally across the landscape of research institutions and research groups. According to MERTON [1973] and ZUCKERMAN [1977], scientific recognition and the accumulation of resources tend to reinforce each other with the effect "(...) the rich get richer at a rate that makes the poor become relatively poorer. Thus, centers of demonstrated scientific excellence are allocated far larger resources for investigation than centers which have yet to make their mark" (MERTON [1973], p. 457). Using NanoBank grants data from the US National Science Foundation (NSF) and the US National Institutes of Health (NIH), this cumulative advantage hypothesis is supported when P1 scientists are compared with P2 peers (TABLE VI). We found that P1 scientists received not only significantly more grants from these two funding agencies in all observation periods but also that P1 scientists improved their share relative to P2 researchers. While the ratio of P1 to P2 was 2 to 1 for NIH grants and 3 to 1 for NSF grants in the early 1990s, in 2002-2004, these ratios increased considerably: 5 to 1 for both NIH and NSF grants (TABLE VI). The gap between P1 and P2 grew over time,

TABLE VI. – NSF and NIH Funding in the United States

	1990–1992		1993–1995		1996–1998		1999–2001		2002–2004	
	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2
NIH grants	0.08 (0.04)	0.04 (0.02)	0.12 (0.04)	0.03 (0.03)	0.29 (0.05)	0.10 (0.03)	0.25 (0.03)	0.03 (0.01)	0.43 (0.05)	0.09 (0.03)
NSF grants	0.30 (0.04)	0.11 (0.02)	0.44 (0.03)	0.08 (0.02)	0.45 (0.03)	0.10 (0.01)	0.62 (0.04)	0.13 (0.02)	0.83 (0.05)	0.17 (0.03)
N	287	242	501	319	721	657	673	491	745	465

Notes: P1 = creative population, P2 = control population. Mean differences between P1 and P2 are all significant at the $p < 0.001$ level except for NIH grants in 1990-1992.

TABLE VII. – International Mobility of P1 and P2 Scientists, in Percent

	1990–1992 to 1993–1995	1999–2001 to 2002–2004
P1	13	26
P2	3	18
Total	6	18
N (P1)	1177	3379
N (P2)	1224	3702
N (Total)	10954	29668

Notes: Missing data not displayed. P1 = creative population, P2 = control population.

indicating that P1 scientists were increasingly successful in attracting grant monies from both sponsors. There was a strong positive correlation between productivity and creativity levels on the one hand, and funding patterns on the other hand.

Finally, we explored NanoBank to assess the international mobility of scientists. By definition, scientific research transcends national borders. Scientists are not only engaged in cross-border research collaboration but they also work at foreign research institutions, either for a limited time period (research stays, sabbaticals) or permanently (when accepting job offers from foreign institutions). Increasing global collaboration is not specific to the domain of nano S&T but reflects a more general trend in the global research system (GEORGHIOU [1998]; GLÄNZEL [2001; 2002]; JAPPE [2007]; PERSSON, GLÄNZEL, and DANELL [2004]).

We investigated frequencies of country changes in the two populations, as visible in publication addresses (TABLE VII). In the total population, country changes tripled from 6% (1990-1992 to 1993-1995) to 18% (1999-2001 to 2002-2004). Hence, scientists in nano S&T were becoming increasingly connected internationally in the observation period 1990-2004. With regard to P1 and P2, we found that country changes were much more common within P1 than in P2. Over the entire observation period, P1 scientists moved significantly more often to other countries than did their P2 peers. Between 1990-1992 and 1993-1995, 13% of P1 scientists changed country compared to 3% in the P2 group. A decade later, almost 26% of P1 scientists changed country compared to 18% in the P2 group. The formidable gap in the relative frequencies of country changes between P1 and P2 scientists indicates differential access to knowledge and expertise at the international level.

Within the P1 population, Japan had the most favorable inflow–outflow relationship in the early 1990s, recruiting most of its foreign scientists from the United States and Germany. Among

TABLE VIII. – Country Changes in P1 Population (Cum. Counts for 1990-2004)

	US	GER	JP	RU	CN	FR	CH	NL	TW	GB	OTH	MIS	Total
US	1619	47	91	2	241	11	9	5	53	41	196	88	2403
GER	41	750	14	4	15	20	16	3	2	8	64	43	980
JP	90	10	2623	3	39	2	2	2	10	11	44	60	2896
RU	4	7	1	93	0	4	1	0	0	0	9	8	127
CN	160	17	38	1	580	2	0	0	26	20	67	16	927
FR	15	18	1	6	2	208	4	4	0	1	13	21	293
CH	9	19	2	2	0	5	91	0	0	1	6	8	143
NL	5	5	2	0	2	2	1	51	0	1	2	6	77
TW	31	3	3	0	26	0	0	0	137	0	11	1	212
GB	52	7	14	0	20	1	1	2	4	262	31	17	411
OTH	158	52	55	5	93	21	5	0	16	20	1315	51	1791
MIS	91	48	63	9	25	22	8	10	5	18	67	324	690
Total	2275	983	2907	125	1043	298	138	77	253	383	1825	643	10950

Notes: TABLE shows cumulated matrix of outflows (rows) and inflows (columns) from 1990-1992 to 2002-2004. OTH = other countries, MIS = missing value.

the countries with a strong inflow–outflow deficit was Russia, losing several top researchers to Germany, the United States, and France until the mid-1990s. A decade later, the picture had changed markedly. From 1999-2001 to 2002-2004, Japan had a clear inflow–outflow deficit, with several P1 scientists migrating to the United States and China. As a newcomer, China had the most favorable inflow–outflow relation in 2002-2004, attracting foreign scientists primarily from the United States and Japan (cum. counts in TABLE VIII).

Together with the results presented above, these figures demonstrate that the international mobility of scientists supports Merton's cumulative advantage hypothesis. However, broader institutional developments are also discernable; for example, the retreat of large corporations from fundamental research (most profoundly in Japan and the United States), deteriorating research conditions after political turmoil (evident in Russia), and the rise of new economic regions (China and Taiwan).

We synthesize our findings using fixed-effects logistic panel regression models that determined to what extent our independent variables influence the likelihood of scientists to belong to P1 in the five observation periods. In contrast to pooled regression models, fixed-effects models explain how the change in the values of independent variables affects the change in the dependent variable over time. Exponentiated coefficients (odds ratios) can be interpreted as the causal effect of X on Y (ALLISON and WATERMAN [2002]).

Apart from minor variations, the regression analyses supported our descriptive findings. Strong positive effects could be observed for the number of publications, patent citations, NSF grants, and international job mobility (both temporary and permanent). In contrast, there were no significant effects for the patent count variable (although there was a positive association), probably because this variable is almost zero for the years following 1995 (because of the time delay inherent in the granting procedure). As would be expected, there was no significant effect for the NIH grant variable (although there was a positive association). As was reported above, the mean difference for NIH grants between the P1 and P2 populations was not significant in the first two observation periods. Therefore, this variable's influence was not strong enough to have had an effect on the dependent variable. Surprisingly, however, we found positive significant

TABLE IX. – Fixed-Effect Logistic Panel Regression
Dependent Variable: P1 population (1), not–P1 population (0)

	(1)	(2)	(3)	(4)	(5)
Number of publications	1.102***	1.102***	1.100***	1.100***	1.099***
	(0.00275)	(0.00275)	(0.00276)	(0.00276)	(0.00276)
Number of patents		0.985	0.982	0.982	0.981
		(0.0107)	(0.0107)	(0.0107)	(0.0107)
Number of patent citations		1.004*	1.004*	1.004*	1.004*
		(0.00149)	(0.00149)	(0.00148)	(0.00148)
Number of NIH grants			1.008	1.008	1.008
			(0.0220)	(0.0220)	(0.0220)
Number of NSF grants			1.155***	1.154***	1.155***
			(0.0329)	(0.0329)	(0.0330)
Move from one country to another			1.190***	1.203***	1.219***
			(0.0509)	(0.0515)	(0.0523)
Move to another organization				0.882***	0.821***
				(0.0254)	(0.0255)
Move to a public research organization					1.408***
					(0.0983)
Move to a firm					1.524***
					(0.139)
Wave 1 (1990–1992)	Reference	Reference	Reference	Reference	Reference
Wave 2 (1993–1995)	1.297***	1.270***	1.243***	1.241***	1.245***
	(0.0623)	(0.0619)	(0.0608)	(0.0607)	(0.0609)
Wave 3 (1996–1998)	1.766***	1.718***	1.664***	1.660***	1.668***
	(0.0836)	(0.0839)	(0.0822)	(0.0820)	(0.0825)
Wave 4 (1999–2001)	1.323***	1.285***	1.220***	1.222***	1.226***
	(0.0659)	(0.0660)	(0.0638)	(0.0639)	(0.0642)
Wave 5 (2002–2004)	0.943	0.917	0.858**	0.857**	0.863**
	(0.0508)	(0.0506)	(0.0484)	(0.0484)	(0.0488)
N (observations)	30794	30794	30794	30794	30794
N (persons)	7330	7330	7330	7330	7330
pseudo R^2	0.128	0.128	0.130	0.131	0.133

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

effects for the two variables that measure job moves to either public research organizations or companies. Both effects were similarly strong, suggesting that changes from public to private research or vice versa were associated with upward mobility to P1.

In light of the findings reported above, moves to firms were more likely in the early phase of field formation (early 1990s), whereas academic institutions absorbed many researchers when the attractive conditions in industry labs deteriorated (mid- to late 1990s). This finding suggests that when scientists move, they tend to move to research units that offer an opportunity to change field or to address challenging research problems.

To test whether network position is more important than multivocality in explaining P1, we ran the same multiple panel regressions with these two alternative dependent variables separately. Instead of the smaller P1 population (14,851 scientists), the dependent variable was first the highest 25% of network brokerage (33,249 scientists), and second the highest 25% of multivocality (33,205 scientists). By and large, these regression analyses gave strong support to our delineation procedure (SECTION II). With minor exceptions, all independent variables had about the same values, strengths, and levels of significance as in the P1 models.[4]

Interestingly, however, while the patent count variable was positive but not significant in the P1 population, it had a positive and significant effect for the upper quartiles of network brokerage and multivocality. This finding suggests that although invention and scientific research were closely related in P1, both variables were less directly connected than in the more extensive alternative populations. In contrast to the results shown in TABLE IX, international job mobility was not a significant predictor for the multivocality population. Apparently, the capability of publishing across a wide range of journals does not require international job moves or affiliations with foreign research institutions to the same extent as brokering otherwise disconnected peer scientists.

IV. Summary and Discussion

What are the new findings reported in this paper? We argue that there are significant creative capabilities in the research system from which subsequent improvements of major research breakthroughs arise. We identify and explore these capabilities drawing on NanoBank, an open-access digital library in the field of nano S&T (ZUCKER *et al.* [2007]). The paper delineates a creative population of scientists, P1 (and a control population: P2), using network brokerage and multivocality as discriminators between creative and normal scientists.

In sum, our findings were as follows: (1) In comparison with the control population, the latent creative population had a significantly higher publication output that steadily rose in the 15-year observation period. (2) Similarly, P1 scientists received significantly higher patent counts (both absolute and relative), and their patents were cited much more frequently in subsequent inventions. (3) In the early 1990s, P1 scientists were nationals predominantly from Japan, the United States, and Germany, while a decade later, China and Taiwan had emerged as new productivity hubs in nano S&T. (4) With regard to institutional affiliations, corporate scientists had below-average representation (and scientists from public research institutions above-average) in the P1 population. However, while the company share in P1 was comparatively high in 1990-1992, it had dropped significantly by 2002-2004, indicating a profound decline in exploratory research in corporate R&D laboratories, particularly in Japan and the United States. (5) As to research sponsorship, P1 scientists (in the United States) received not only significantly more grants from NSF and NIH, but their share grew markedly at the expense of scientists in the P2 population, indicating a cumulative advantage effect. (6) P1 scientists moved significantly more often to (or had affiliations with research institutions in) other countries than their P2 peers. The gap in the relative frequencies demonstrates differential access to the knowledge base of international scientific communities. Today, China has the most favorable inflow–outflow relationship, attracting foreign scientists primarily from the United States and Japan. (7)

What are the lessons that can be learned from this paper? First, science analysts and policy makers are becoming increasingly aware of the inherent limitations in measuring scientific

4. The results of these regression analyses are available on request from the second author.

performance by publication and citation indicators only. There is a clear need for alternative indicators and methods that adequately capture the upper tail of the quality distribution in science (BONACCORSI [2007]). However, measurement of capabilities that are associated with breakthrough research is a challenging task. In this paper, we introduce a method for delineating such capabilities on the population level. The explorative data analysis suggests that this method extends the more conventional techniques of counting publications and citations for individual scientists. While these techniques have been used for selecting high-impact work in established and mainstream research areas, our method captures creative capabilities in multidisciplinary and rapid-growth research domains. In these domains, such as nano S&T, creative ideas emerge when communication networks provide sufficient opportunities for connecting heterogeneous knowledge sources, and when scientists and their groups have a work spectrum that goes beyond narrow thematic specialization. We are confident that using network brokerage and multivocality as population generators represents a considerable improvement in identifying the upper tail of scientific creativity on an aggregate scale.

The second lesson of this paper concerns the links between developments in creativity levels and institutional changes both within the research system and in its environment, such as the decline of fundamental research in industry or new opportunities for scientific research in rising economies. Our results support the view that the capabilities of research groups, departments, research organizations, regions, and entire nations are strongly influenced by various institutional factors and their combinations. While we do not challenge the research tradition that locates creativity mainly in highly gifted individuals (see, for example, HELLER *et al.* [2000]), we do claim and provide empirical evidence that studying the relationship between institutional mechanisms and factors (as independent variables) and different creativity capabilities in scientific research (as the dependent variable) yields valuable insights as to why some organizations or regions (or other units of analysis) are more successful in accomplishing ground-breaking work than others. Such insights are of particular interest for research policy makers and research managers in their efforts to improve the governance of research.

ACKNOWLEDGEMENTS

Data discussed in this paper are derived from Nanobank <Lynne G. Zucker and Michael R. Darby, Nanobank Data Description, release 1.0 (beta-test), Los Angeles, CA: UCLA Center for International Science, Technology, and Cultural Policy and Nanobank, January 17 [2007].> Nanobank data are derived from the Science Citation Index Expanded, Social Sciences Citation Index, and Arts & Humanities Citation Index of the Institute for Scientific Information®, Inc. (ISI®), Philadelphia, Pennsylvania, USA: © Copyright Institute for Scientific Information®, Inc. 2006. All rights reserved. The authors also gratefully acknowledge the support of Nicolai Mallig, computer scientist at the Karlsruhe Institute of Technology, Germany.

Correspondence:
Thomas Heinze,
Bergische Universität Wuppertal, Gaußstraße 20, DE-42119 Wuppertal,
theinze@uni-wuppertal.de (corresponding author)

Gerrit Bauer,
Ludwig Maximilians Universität München, Konradstraße 6 / 010, DE-80801 München,
gerrit.bauer@lmu.de

References

AKSNES, D. W. (2006): "Citation Rates and Perceptions of Scientific Contribution," *Journal of the American Society for Information Science and Technology*, **57**, 169-185. [109]

ALLISON, D. P., and J. S. LONG (1990): "Departmental Effects on Scientific Productivity," *American Sociological Review*, **55**, 469-478. [109]

ALLISON, D. P., and R. P. WATERMAN (2002): "Fixed-Effects Negative Binomial Regression Models," *Sociological Methodology*, **32**, 247-265. [117]

ANDERSON, P., and M. L. TUSHMAN (1990): "Technological Discontinuities and Dominant Designs: A Cyclical Model of Technological Change," *Administrative Science Quarterly*, **35**, 604-633. [108]

AUDIA, P. G., and J. A. GONCALO (2007): "Past Success and Creativity over Time: A Study of Inventors in the Hard Disk Drive Industry," *Management Science*, **53**, 1-15. [109]

BINNIG, G., and H. ROHRER (1982): "Scanning Tunneling Microscopy," *Surface Science*, **126**, 336-344. [108]

BINNIG, G., and H. ROHRER (1986): "Scanning Tunneling Microscopy," *IBM Journal Research Division*, **30**, 355-369. [108]

BONACCORSI, A. (2007): "Explaining Poor Performance of European Science. Institutions versus Policies," *Science and Public Policy*, **34**, 303-316. [120]

BURT, R. S. (2004): "Structural Holes and Good Ideas," *The American Journal of Sociology*, **110**, 349-399. [108, 109]

COLLIER, C. P., E. W. WONG, M. BELOHRADSKY, and F. J. RAYMO (1999): "Electronically Configurable Molecular-Based Logic Gates," *Science*, **285**, 391-394. [108]

COLYVAS, J. A. (2007): "From Divergent Meanings to Common Practices: The Early Institutionalization of Technology Transfer in the Life Sciences at Stanford University," *Research Policy*, **36**, 456-476. [107]

EUROPEAN COMMISSION (2005): *Some Figures about Nanotechnology R&D in Europe and Beyond*. www.cordis.europa.eu/nanotechnology. [115]

FLEMING, L., M. SANTIAGO, and D. CHEN (2007): "Collaborative Brokerage, Generative Creativity, and Creative Success," *Administrative Science Quarterly*, **52**, 443-475. [109]

FLEMING, L., and M. SZIGETY (2006): *Exploring the Tail of Creativity: An Evolutionary Model of Breakthrough Invention*. In: *Ecology and Strategy: Advances in Strategic Management*. Ed. by J. Baum, S. D. Dobrev, and A. van Witteloostuijn. **23**. Oxford: Elsevier, 335-359. [109]

GEORGHIOU, L. (1998): "Global Cooperation in Research," *Research Policy*, **27**, 611-626. [116]

GLÄNZEL, W. (2001): "National Characteristics in International Scientific Co-Authorship Relations," *Scientometrics*, **51**, 69-115. [116]

GLÄNZEL, W. (2002): "Coauthorship Patterns and Trends in the Sciences (1980-1998): A Bibliometric Study with Implications for Database Indexing and Search Strategies," *Library Trends*, **50**, 461-473. [116]

HEATH, J. R., S. C. O'BRIEN, Q. ZHANG, and Y. LIU (1985): "Lanthanum Complexes of Spheroidal Carbon Shells," *Journal of the American Chemical Society*, **107**, 7779-7780. [108]

HEINZE, T. (2013): "Creative Accomplishments in Science: Definition, Theoretical Considerations, Examples from Science History, and Bibliometric Findings," *Scientometrics*, **95**, 927-940. [109]

HEINZE, T., R. HEIDLER, H. HEIBERGER, and J. RIEBLING (2013): "New Patterns of Scientific Growth. How Research Expanded after the Invention of Scanning Tunneling Microscopy and the Discovery of Buckminsterfullerenes," *Journal of the American Society for Information Science and Technology*, **64**, 829. [108]

HEINZE, T., P. SHAPIRA, J. ROGERS, and J. SENKER (2009): "Organizational and Institutional Influences on Creativity in Scientific Research," *Research Policy*, **38**, 610-623. [109, 114]

HEINZE, T., P. SHAPIRA, J. SENKER, and S. KUHLMANN (2007): "Identifying Creative Research Accomplishments: Methodology and Results for Nanotechnology and Human Genetics," *Scientometrics*, **70**, 125-152. [108, 109]

HELLER, K. A., F. J. MÖNKS, R. F. SUBOTNIK, and R. J. STERNBERG (2000): *International Handbook of Giftedness and Talent*. Oxford: Elsevier Science. [120]

HEMLIN, S., C. M. ALLWOOD, and B. R. MARTIN (2004): *Creative Knowledge Environments. The Influences on Creativity in Research and Innovation*. Cheltenham: Edward Elgar. [109]

HOLLINGSWORTH, R. J. (2002): *Research Organizations and Major Discoveries in Twentieth-Century Science: A Case of Excellence in Biomedical Research*. WZB Discussion Paper P02-003. Berlin: WZB. [109]

HOLLINGSWORTH, R. J. (2004): *Institutionalizing Excellence in Biomedical Research: The Case of Rockefeller University*. In: *Creating a Tradition of Biomedical Research. Contributions to the History of The Rockefeller University*. Ed. by D. H. Stapleton. New York: Rockefeller University Press, 17-63. [109]

HULLMANN, A., and M. MEYER (2003): "Publications and Patents in Nanotechnology. An Overview of Previous Studies and the State of the Art," *Scientometrics*, **58**, 507-527. [111]

IIJIMA, S. (1991): "Helical Microtubules of Graphitic Carbon," *Nature*, **354**, 56-58. [108]

IIJIMA, S., P. M. AJAYAN, and T. ICHIHASHI (1992): "Growth-Model for Carbon Nanotubes," *Physical Review Letters*, **69**, 3100-3103. [108]

JAPPE, A. (2007): "Explaining International Collaboration in Global Environmental Change Research," *Scientometrics*, **71**, 367-390. [116]

KROTO, H. W., J. R. HEATH, S. C. O'BRIEN, R. F. CURL, and R. E. SMALLEY (1985): "C-60: Buckminsterfullerene," *Nature*, **318**, 162-163. [108]

LIGHTMAN, A. (2005): *The Discoveries. Great Breakthroughs in the 20th Century Science*. New York: Pantheon Books. [109]

MERTON, R. K. (1973): *The Matthew Effect in Science*. In: *The Sociology of Science: Theoretical and Empirical Investigations*. Ed. by N. W. Storer. Chicago: Chicago University Press, 439-459. [115]

MOED, H. F., W. GLÄNZEL, and U. SCHMOCH (2004): *Handbook of Quantitative Science and Technology Research*. The Use of Publication and Patent Statistics in Studies of S&T Systems. Dordrecht: Kluwer. [112]

MUNARI, F., E. B. ROBERTS, and M. SOBRERO (2002): "Privatization Processes and the Redefinition of Corporate R&D Boundaries," *Research Policy*, **31**, 31-53. [114]

NOYONS, E. C. M., R. K. BUTER, A. F. J. van RAAN, U. SCHMOCH, S. HINZE, T. HEINZE, and R. RAGNOW (2003): *Mapping Excellence in Science and Technology across Europe. Nanoscience and Nanotechnology*. Brussels: European Commission. [112]

PERSSON, O., W. GLÄNZEL, and R. DANELL (2004): "Inflationary Bibliometric Values: The Role of Scientific Collaboration and the Need for Relative Indicators in Evaluative Studies," *Scientometrics*, **60**, 421-432. [116]

POWELL, W. W., D. R. WHITE, K. W. KOPUT, and J. OWEN-SMITH (2005): "Network Dynamics and Field Evolution: The Growth of Interorganizational Collaboration in the Life Sciences," *The American Journal of Sociology*, **110**, 1132-1205. [107]

RODAN, S., and C. GALUNIC (2004): "More Than Network Structure: How Knowledge Heterogeneity Influences Managerial Performance and Innovativeness," *Strategic Management Journal*, **25**, 541-562. [109]

SIMONTON, D. K. (1999): *Origins of Genius: Darwinian Perspectives on Creativity*. New York: Oxford University Press. [109]

SIMONTON, D. K. (2004): *Creativity in Science. Change, Logic, Genius, and Zeitgeist*. Cambridge: Cambridge University Press. [108, 109, 111]

STUMPF, H. (1995): "Scientific Creativity: A Short Overview," *Educational Psychology Review*, **7**, 225-241. [109]

TANS, S. J., A. R. M. VERSCHUEREN, and C. DEKKER (1998): "Room-Temperature Transistor Based on a Single Carbon Nanotube," *Nature*, **393**, 49-52. [108]

UZZI, B., and J. SPIRO (2005): "Collaboration and Creativity. The Small World Problem," *The American Journal of Sociology*, **111**, 447-504. [109]

WHITLEY, R. (2000): *Introduction: Science Transformed? The Changing Nature of Knowledge Production at the End of the Twentieth Century*. In: *The Intellectual and Social Organization of the Sciences, 2nd Edition*. Ed. by R. Whitley. Oxford: Oxford University Press, IX-XLIV. [109]

ZAHEER, A., and G. G. BELL (2005): "Benefiting from Network Position: Firm Capabilities, Structural Holes, and Performance," *Strategic Management Journal*, **26**, 809-825. [109]

ZUCKER, L. G., and M. R. DARBY (2005): *Socio-Economic Impact of Nanoscale Science: Initial Results and NanoBank*. NBER Working Paper No. 11181. [111]

ZUCKER, L. G., M. R. DARBY, J. FURNER, R. C. LIU, and H. MA (2007): "Minerva Unbound: Knowledge Stocks, Knowledge Flows and New Knowledge Production," *Research Policy*, **36**, 850-863. [108, 119]

ZUCKER, L. G., M. R. DARBY, and M. TORERO (2002): "Labor Mobility from Academe to Commerce," *Journal of Labor Economics*, **20**, 629-660. [107]

ZUCKERMAN, H. (1977): *Scientific Elite: Nobel Laureates in the United States*. New York: Free Press. [109, 115]