Career-based influences on scientific recognition in the United States and Europe: Longitudinal evidence from curriculum vitae data

Jan Youtie, Enterprise Innovation Institute and School of Public Policy, Georgia Institute of Technology, 75 Fifth Street NW, Atlanta, GA 30332-0640, United States, jan.youtie@innovate.gatech.edu

Juan Rogers, School of Public Policy, Georgia Institute of Technology, 685 Cherry St. NW, Atlanta, GA 30332-0345, United States, jdrogers@gatech.edu

Thomas Heinze, Department of Educational and Social Sciences, University of Wuppertal, Gaussstrasse 20, D-42119 Wuppertal, Germany, theinze@uni-wuppertal.de

* Philip Shapira, Manchester Institute of Innovation Research, Manchester Business School, University of Manchester, M13 9PL, Manchester, United Kingdom, pshapira@mbs.ac.uk; and School of Public Policy, Georgia Institute of Technology, 685 Cherry St. NW, Atlanta, GA 30332-0345, United States

Li Tang, School of Public Economics and Administration, Shanghai University of Finance and Economics, Shanghai 200433, China, tang006@gmail.com; and Georgia Tech Program in Science, Technology and Innovation Policy, 685 Cherry St. NW, Atlanta, GA 30332-0345, United States.

*For correspondence
Corresponding author: Philip Shapira, Email: pshapira@mbs.ac.uk

April 23, 2013

Career-based influences on scientific recognition in the United States and Europe: Longitudinal evidence from curriculum vitae data.

Abstract

This paper examines how funding patterns, career pathways and collaboration networks influence scientific recognition. We analyze these institutional factors in the early and middle phases of academic careers through comparison of a group of researchers recognized as creative by their peers with a matched group of researchers. Measurement of scientific recognition is based on survey nominations and research prizes in two growing, laboratory-intensive research domains: nanotechnology and human genetics. Curriculum vitae data is used to compare researchers based in the United States and Europe. In the early career model for the United States, we find that scientific recognition is associated with broad academic education, fast completion of PhD, and a record of independent post-doctoral research, while in Europe these factors are much less prominent. The mid-career model suggests that both in the United States and Europe fast job promotion within academia is a strong predictor of future recognition. However, there is a clear divide across the Atlantic regarding other mid-career factors: work experience inside and outside academia, research leadership, external grant income, and prizes from professional associations are connected to scientific recognition in the United States, but are less influential in Europe.

Research Highlights

- Organizational and institutional context influences creative scientist recognition
- Institutional factors underlying scientific recognition differ in early and mid-career models
- Academic mobility, junior independence, and grants influence US scientific recognition
- Spatial mobility and disciplinary stability influence European scientific recognition
- Merit and research performance values are more institutionalized in the US than in Europe

1   Introduction

Scientific creativity is typically defined in terms of knowledge and capabilities that are both new and useful (Sternberg 2003; Hollingsworth, 2004; Simonton, 2004; Lightman, 2005). 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 (Stumpf, 1995; Amabile 1996, Hemlin, et al. 2004; Merton and Barber, 2004). Creative research has most frequently been identified either by research breakthroughs as manifested with the awarding of prestigious scientific prizes (Zuckerman, 1977; Hollingsworth, 2003), or as productivity and citation measures (Simonton, 2004; Fleming and Szigety, 2006). There have also been various attempts to explain creative outcomes, including within a probabilistic-evolutionary framework (Simonton, 1999; 2004), reference to network positions (Burt, 2004; Uzzi and Spiro, 2005; Fleming et al., 2007) or with regard to organizational and environmental variables (Hollingsworth, 2004; Heinze et al., 2009).

While there is a long tradition that views creativity as a personal or mental trait (Mayer 1999; Carter, 2004), another influential perspective stresses the social character of creativity, particularly in science (Csikszentmihalyi, 1997; 1999). Science, more than other fields, has evolved procedures, disciplines, and institutions to accredit new knowledge (Whitley, 2000). Whether a new theoretical idea, method, or research instrument is regarded as creative or not depends on the appraisal of peers. Scientific recognition is typically attributed by such peers according to individual creative contributions to the collective stock of knowledge.

There is an extensive body of work that examines the mechanisms of distribution and accumulation of scientific recognition. Following Merton's (1968) initial theoretical statement, the mechanism of "cumulative advantage" has been identified empirically in science and other areas of social life (DiPrete and Eirich, 2006; Zuckerman, 2010). Thus, minor initial differences in performance, for example in the number of publications, are transformed via continuous accumulation over time into substantial inequality in the distribution of scientific reputation. Initial differences in performance tend to be multiplied by institutional effects, such as funding, social capital, and organizational decisions, such as tenure or promotion. Therefore, the accumulation of recognition is connected – sometimes directly, sometimes indirectly – to various organizational and institutional factors operating at the "meso-level" of academic careers.

This paper examines how such career-based factors influence the recognition of scientists as contributors of creative pieces of work. Our analysis involves complex methodical issues. Since we are not interested in replicating previous findings on the cumulative advantage mechanism, we have to control for it in our measurements. We offer two solutions. First, building on a sample of scientists recognized as creative by a combined survey-prizewinner method, we identify a comparison group of scientists using early publication behavior variables. In this way, we build matched pairs that are almost identical in their early publication behavior. If there is a difference in performance and recognition between the "creative" scientists and the "matched" scientists later in their careers, these differences cannot be plausibly explained by early publication differentials accumulating over time, but should be due to other career-based factors. Second, we differentiate between early and mid- career factors, because we assume that the professional challenges scientists have to deal with in the years following the completion of their PhD are different from those that come with tenure. Thus, distinguishing between an early-career model and a mid-career model helps us in assessing the impact of institutional factors at different career stages.

We take on another complex methodical issue: the quantitative comparison between career-based factors in the United States and Europe. Institutional impacts on performance and recognition in science are typically examined either with data on US-based scientists (examples are Azoulay et al., 2009; Mumford et al., 2005; Dietz and Bozeman, 2005; Allison and Long, 1990) or with data on Europe-based scientists (examples are Sandström, 2009; Cañibano et al., 2008; Melin and Danell, 2006). In contrast, there are only few direct intercontinental comparisons (exceptions are Stephan, 2008; Laudel, 2006). A US-European comparison is valuable in at least two respects. First, there is considerable discussion in science policy circles on each side of the Atlantic regarding national and regional scientific competitiveness, in the context of evolving global research capabilities. The National Academy of Sciences and the National Academy of Engineering argue that the US can only maintain its position as a world leader in science and innovation if it takes action at improving science teaching, at increasing efforts in funding science and engineering research, and at putting in place effective incentives for scientific and technical innovations (NAS/NAE 2007). In Europe, there is significant discussion not only about individual country science capabilities but also about underlying factors and policy options in developing a cohesive and more powerful European Research Area (see, for example, ERAB 2009; Bonaccorsi, 2007). Our comparison of career-based factors between the US and Europe offers a contribution to this policy debate. Second, prior work has highlighted the interconnection between

institutional reforms and scientific eminence. For example, Ben-David (1971) in a qualitative study suggests that the rise of US leadership in science and technology in the early 20th century (overtaking Germany) was significantly aided by the establishment of graduate schools at leading American universities, which connected graduate education with scientific research, and because of the competitive and decentralized institutional context in which American universities were placed. Our analysis of career-based and institutional factors extends and updates this US-European comparative research theme, adding a fresh quantitative analysis. In this paper, we compare US-based and Europe-based researchers using a curriculum vitae (CV) based data source, in combination with publication data from Thomson Reuters Web of Science (WOS). This allows examination of differences in educational profile, employment, research grants, academic awards, and collaboration patterns between "creative" and "matched" researchers. The comparison between the US and Europe will be shown to be very important in underscoring the distinctive institutional and career features of both regions.

Our analysis covers two research domains: nanotechnology and human genetics. Nanotechnology took off in the early 1980s and today embraces fields, such as applied physics, materials science, physical chemistry, physics of condensed matter, molecular biology, biochemistry, and polymer science and engineering. Two major scientific innovations upon which the domain of nanotechnology is based are the invention of probe microscopy and the discovery of carbon fullerenes (Mody 2011, Hessenbruch, 2004, Aldersey-Williams, 1995). Human genetics has its intellectual roots in molecular biology developed in the 1940s (Kay, 1993; Kohler, 1991). It has undergone tremendous expansion following the start (1990) and the successful completion (2003) of the Human Genome Project (Sulston and Ferry, 2002). There are correspondences between the two fields. Both domains are laboratory-based, with relatively ubiquitous equipment demands, and each has grown considerably in recent years.[1] The two domains are also similar in that intellectual innovations often give rise to technological and medical applications (Porter and Youtie, 2009; Rafols and Meyer, 2010).

---

[1] Human genetics research grew from 1.9 percent of global research publications in 1989 to 10.1 percent by 2008 (Wellcome Trust, 2010). Nanotechnology publications increased from about 1 percent of global research publications in 1991 to 6.2 percent in 2010 (calculated from data in Arora et al., 2012). These results are based on Web of Science publications, although produced by different authors. While most nanotechnology papers are in such fields as materials science, chemistry, physics, and engineering (Porter and Youtie, 2009), some overlap in these reported publication estimates is likely where nanoscale research intersects with human genetics. Nonetheless, the fields of nanotechnology and human genetics have clearly emerged as major domains in the world of scientific research and together contribute a significant share of contemporary scientific output.

Our key results are these. In the early-career model, for the United States, we find that scientific recognition is associated with broad academic education, fast completion of PhD, and a record of independent post-doctoral research, while in Europe these factors are less prominent. The mid-career model suggests that both in the United States and Europe fast job promotion within academia is a strong predictor for future recognition. However, there is a clear divide across the Atlantic regarding other mid-career factors: work experience inside and outside academia, research leadership, external grant income, and prizes from professional associations are connected to high scientific recognition in the United States, but are less influential in Europe.

We begin the paper by reviewing contributions to the literature on institutional influences on scientific performance and reputation (Section 2). Then, we introduce our method and data including the early-career model and the mid-career model (Section 3). This is followed by a discussion of the regression analyses based on the two career models (Section 4). Finally, we put our findings in a broader theoretical perspective (Section 5).

2   Literature review

In this section, we concisely review selected works that deal with the question of how institutional and organizational factors influence scientific performance and reputation. Social scientists have long been concerned with explaining inequality in reputation and performance by reference to mechanisms of social stratification (see, for example, Merton, 1973, Münch 2008). We identify four major themes in the literature about how stratification in science is produced: institutional prestige, funding, career pathways and productivity, and collaboration structures. The literature on these topics is too extensive to be easily summarized here, so we introduce key examples from each of these literature themes to highlight key debates and insights.

The first theme in the literature about stratification of scientists relates to institutional prestige. We introduce two illustrative examples. The first example is a study by Allison and Long (1990) who find that the effect of department affiliation on productivity is more important than vice versa. Scientists moving to prestigious departments subsequently show substantial increases in their rate of publication and in the rate of citation to those publications, while those who move downward suffer decreases in performance. The second, more recent example is Hermanowicz's (2009) study which shows that the institutional setting affects how physicists appraise their career success, and how their self-evaluations are strengthened and reproduced by the stratified structure in which their work takes place. Using

4

longitudinal life-course interviews across three institutional types: elitist, pluralist, and communitarian colleges, this work tells a story of how physicists at elitist institutions focus on productivity and scientific accomplishment, whereas faculty at pluralist institutions have career ebbs and flows, and those at communitarian institutions withdraw at an early point in their careers.

Second, there is a literature on the effects of external funding on career success and scientific performance. This literature is somewhat younger than the literature on institutional prestige, because external funding – mostly distributed via projects grants or person grants within larger thematic programs – has become quite important in recent years. A first example from this literature describes the selection mechanism of external funding. The National Research Council (2006) evaluation of the Markey Scientists Program in the United States finds that awardees of successful proposals and applicants of highly-rated but unsuccessful proposals did not differ much on performance measures, but the successful awardees were more homogeneous than unsuccessful applicants in that they came from top universities, had already received tenure and been promoted, and received more research grants. The second example shows mid-term effects on research strategies of junior scientists: Melin and Danell's (2006) analysis of a funding scheme for highly talented junior scientists in Sweden shows that although the 40 candidates that were invited from a pool of 500 applicants were already quite homogenous in terms of their research productivity and quality, the performance profile of the 20 successful awardees became even more homogeneous due to the program's support, while at the same time the performance heterogeneity of the group of non-funded scientists grew substantially. Whereas successful candidates were able to enhance and strengthen their research profile, this process was more difficult and less certain for the non-successful candidates (Melin and Danell, 2006). The third example focuses on publication output and citation rates: Azoulay et al. (2009) find that investigators funded by the Howard Hughes Medical Institute (HHMI) Investigator Program generate more highly cited papers than do their peers who received grants from other sponsors. HHMI's focus on long-term institutional funding, which makes the program unique in the United States (Heinze, 2008), leads towards homogeneous performance among HHMI investigators.

Third, there is a literature on career pathways and scientific productivity. The first example is Mumford et al. (2005) who studied recognition as described in obituaries of several hundred scientists. The authors find that despite traditional emphasis in this literature on early career influences, such as education and mentoring, later career experiences, in particular work strategies, laboratory intellectual engagement, tangible research support and

5

active intellectual exchange, have strong effects as well. The second example is a comparison between France and the US regarding career pathways by Gaughan and Robin (2004). The authors find that in the US, postdoctoral positions neither help nor hinder entry into an academic career, while in France, postdoctoral positions and other temporary positions are perceived as barriers to academic careers. Similarly, a third study by Dietz and Bozeman (2005) shows that scientific productivity of postdoctoral researchers is significantly lower than that of colleagues with tenured faculty positions. Their finding is particularly relevant, because the number of postdoctoral positions has sharply risen since the 1970s not only in the US and Europe, but also in Australia (Giles et al., 2009), while the number of faculty positions has increased only moderately, results that are corroborated by Stephan (2008) who warns that employing an academic labor force with mostly temporary positions and limited career prospects might seriously diminish overall scientific productivity in western countries.

Fourth, there is a literature on the opportunity structures in collaboration networks that facilitate the generation of new ideas and their successful diffusion. Burt (2004) finds that managers who are placed at the intersection of heterogeneous social groups within their company have an increased likelihood of drawing upon multiple knowledge sources which, in turn, enables them to generate more and better evaluated ideas than their peers. In contrast to Burt, a study by Uzzi and Spiro (2005) argues that cohesive collaborative networks offer the benefits of trust, shared risk taking, and easy mobilization in facilitating information and knowledge transfer. Rather than brokering holes in collaboration networks, in their view innovative activities emerge from dense interaction among individuals who know each other well. A third example, a study by Fleming et al. (2007), integrates the two conflicting arguments of structural holes and cohesive collaboration. This study shows that although brokering inventors are more likely to generate new ideas, the brokered network itself is not well suited to diffusing these ideas. Rather, inventors depend on more cohesive and dense collaborations for effective diffusion and application of their ideas.

These four literatures on institutional prestige, funding programs, career pathways, and collaboration networks discuss factors in the institutional and organizational context of research that influence the performance and recognition of scientists. In this paper, we do not conceptualize institutional prestige in terms of institutional rankings or other similar organizational status measures. This is because the members of our sample are generally situated at prestigious universities or research laboratories. Instead we use measures of the wider institutional context in which respondents carry out their research, particularly regarding funding, career pathways, and collaboration networks. We examine the extent to

which these contextual factors distinguish recognition of highly creative research through supporting opportunities to pursue research funding, assembling work experience in non-academic jobs, providing for collaboration with colleagues, and fostering opportunities for the take-up of administrative leadership roles. In addition, we control for educational experience.

Most importantly, this paper studies these factors in combination. Specific contributions in the four literatures tend to give emphasis to network or organizational or career pathways or funding scheme variables. Relatively few have studied creativity and recognition in a broader field and country comparative framework. Therefore, this article contributes to combining the explanatory power of institutional variables using extensive curriculum vitae (CV) data. Following Mumford et al. (2005) in distinguishing between early and late career stages, we conceptualize an early-career model and a mid-career model, and we combine various organizational, external funding and collaboration measures. In so doing, we draw on the literature using CV data for empirical analyses (for example: Cañibano and Bozeman, 2009; Gaughan and Ponomariov, 2008; Cañibano et al., 2008; Dietz and Bozeman, 2005). The next section provides details of how we delineate the "creative" and "matched" population, and how we proceed with collecting and analyzing career-based variables.

3 Method and Data

3.1 Identification of creative scientists and matching technique

While previous studies usually rely on a single indicator to identify creative accomplishments or scientific recognition (or both), such as citation and publication data (Simonton, 2004) or prestigious science awards (Hollingsworth, 2003), we build our dependent variable on a combination of both survey nominations and prize winners. We apply this approach in the fields of nanotechnology and human genetics. Full details of the method and results have been previously published elsewhere (see Heinze et al., 2007), so we provide only a summary here. First, nomination data were collected through an international written survey in 2005 where several hundred experts in our two target scientific fields, among them highly cited scientists, active researchers from academia and industry and editors of major research journals, were asked to nominate creative research accomplishments in their respective fields. This yielded 301 nominations in nanotechnology (117 for Europe, 184 for US) and 101 nominations in human genetics (38 for Europe, 63 for US). Second, we compiled a dataset of scientific award winners in the period 1995-2004 by screening professional societies and major funding bodies in Europe and the United States. This yielded 247

individuals in nanotechnology (139 for Europe, 108 for US), and 173 individuals in human genetics (121 for Europe, 52 for US). Third, we combined nomination and prize winner data and thus identified 76 "creative researchers" or "CRs" who either received (a) multiple survey nominations (n=24), or (b) multiple prize awards (n=18), or (c) combinations of survey nominations and prize awards (n=34). It should be emphasized that our use of "creative" denotes a social attribute. Creativity in this study refers to social recognition by other researchers rather than research productivity as measured by publication.

Starting with these 76 CRs, we developed a technique for obtaining a comparison group after testing the performance of several different matching techniques (Heinze and Bauer, 2007; Youtie et al., 2009). As outlined above, we want to exclude as much of the cumulative advantage mechanism as possible. Therefore, we focus on minor initial differences in scientific achievement, such as publications and citations. The cumulative advantage mechanism is not directly about the accumulation of minor institutional advantages. So, variables, such as the doctoral advisor's reputation, or the prestige of the institution granting the terminal degree, should not be used for matching. We chose criteria of scientific outputs based on widely available and replicable bibliometric data. We matched creative researchers in nanotechnology and human genetics with a comparison group of researchers that have the same or very similar early career publishing characteristics: (a) same first year of publication (b) same subject category of the first publication, and (c) similar publication volume for the first six years in the specified emerging domain. This three-pronged approach yielded 463 scientists in nanotechnology cases and 249 scientists in human genetics. In addition, the pool from which the comparison group is eventually drawn has similar distributions in terms of full career article output and citation levels to those of the CRs. Therefore, both groups (CRs and matches) do neither differ much in terms of research productivity, as both groups operate at very high levels of scientific production, nor in the sheer number of citations, since the two groups have high numbers of citations as well, especially relative to that of a random sample (see histograms in Appendix Figure 1 and Appendix Figure 2).[2]

---

[2] We did test propensity score matching with the creative researchers to obtain a relevant comparison group of researchers (Youtie et al., 2009). In the propensity score matching, we focused on matching publication record characteristics; we did not try characteristics of the institution and career because that would preclude their use in the final model. In particular, we matched on the log of the number of citations divided by the number of years of publications in the given field (nanotechnology or human genetics). To properly estimate a propensity score model, one must stratify the sample into propensity score blocks and test that each covariate is balanced (no significant difference in means) within the blocks. Our propensity score matching effort did not satisfy this balancing requirement, however. For example, among creative researchers in nanotechnology, only 12 percent fell into the lowest interval while more than 70 percent fell into the top three intervals, whereas among the

In the next step, we developed a protocol to request CVs of the CRs and their matches. These procedures occurred simultaneously in the United States and in Europe and continued through the first two months of 2009. We initiated a process of coding the CV data we received from these requests. A coding scheme was developed, with more than 60 potential variables. A number of key variables including job start and end dates, research awards, prizes and editorial position were not universally available via CVs. The truncation of CVs is not a new problem (Cañibano and Bozeman, 2009). We addressed this problem by initiating a verification process through multiple rounds of email-based surveys to the scientists concerned in order to complete missing data. These surveys were administered from December 2009 through April 2010. The verification survey produced more complete data from 40 percent of the European sample and 25 percent of the United States sample. In addition, we supplemented missing information with Internet searches and excerpts from existing grant awards, dissertation abstracts, publication, and other relevant databases. Also, we added to the CV data Web of Science publications in order to have a common publication database rather than a heterogeneous list of CV publications. Based on these efforts, we were able to match all CRs with most optimal non-CR respondents. Final matches were selected for each CR based on the most complete CV or, if two or more CVs were equally complete, one was selected at random. The resulting database totaled 152 records: 76 CRs and 76 matches. Six of the CRs and nine of the matches are women (similarity in initial performance was used in matching without account of gender). Of the 152 records, 80 are US-based researchers and 72 are European-based with eight countries represented. Among the European-based researchers, 32 are from Germany, 14 from the UK, 12 from France, 4 each from the Netherlands and Switzerland, and 2 each from Austria, Spain and Sweden. For each country, there are equal numbers of CRs and matches. The country distribution of European researchers matches the national locations of the European-based researchers recognized as creative in the original survey.

---

comparison group in nanotechnology, 94 percent fell into the lowest interval and less than 1 percent into the highest interval. Thus we did not use propensity score matching to create a comparison group for this analysis. Instead we used the approach described in this paper, which requires that the comparison group have the same early career publishing characteristics as the creative researcher group in three areas: (1) same first year of publication, (2) same subject category of the first publication (where the subject category is the classification which the Web of Science uses to assign journals), and (3) similar numbers of publications for the first six years of articles in either nanotechnology or human genetics. The performance of our matching approach on a performance measure of the researchers, namely, citations, which was not used to select the matching researchers, can be observed from the histograms in the Appendix. This is a clear indication that the matching approach has worked as expected.

It should be emphasized that our matching technique offers a viable and robust approach to control for the cumulative advantage mechanism. Identification of a comparison group of scientists using early publication behavior variables implies that differences in performance and recognition between the "creative" scientists and their "matched" peers later in their careers – as measured by their publication and CV data – cannot be plausibly explained by recognition of early publication differentials accumulating over time, but should be due to other career-based and institutional factors. Thus, our models estimate organizational and institutional influences independently of early cumulative publication advantage. In addition, distinguishing between an early-career model and a mid-career model helps us in assessing the impact of institutional factors at different career stages, as is explained below.

3.2   Early Career Model

The dependent variable measures whether or not the respondent is a recognized as creative by peer scientists. The independent variables come in two models, an early-career and mid-career model. These models considered and described in the following two subsections. A summary list of variables and descriptors, data sources, and connections to the literature provided is in Table 1.

[INSERT TABLE 1 NEAR HERE]

Building two career models demanded that we construct variables representing consecutive time periods, measured via three year time windows. The first period captures the three-year period around the year when the researcher received her terminal degree. The second period is the three-year period six years after the researcher received his terminal degree. This is often the lead up time to achieving tenure or a permanent position in an academic or research career. The third period refers to three-year period twelve years after the researcher received her terminal degree representing the time to full professorship or other senior position in a research organization.

The general outcome variable Y is binary. Y=1 denotes that the observation represents a creative researcher (CR), Y=0 denotes a matched researcher. X is a k vector of observed explanatory variables; β is a k vector of unknown parameters. Our testing model for the binary outcome variable using probit regression can be written as: $Pr(Y_t = 1|X_t) = F(X_{t1}'\beta_{t1})$.

10

The early-career model posits that the probability of being recognized as a creative researcher is a function of several organizational context and collaboration variables. First of all, there are variables capturing the *career pathway*. This includes educational variables representing the breadth and diversity of university education:

- Degrees same area – indicates whether the researcher changes disciplines from bachelor to doctoral degrees.
- Count univ. – represents the number of different higher educational institutions at which researchers received their university degrees.
- Count countries univ. – indicates whether the educational institutions were in the same or different countries.

Another set of categorical career pathway variables captures the number of years required to receive a terminal degree, typically the PhD or the MD, after the bachelor's degree is earned. These variables measure whether the respondent worked in a supportive academic context in which a terminal degree could be earned after few years or if the work context demanded longer periods of PhD work. The distribution of the number of years required to receive a terminal degree was found to be non-normal (Shapiro-Wilk test W=.917, p<.01) because about an equal number of observations fell at the four-, five-, and six-year time periods, a few at the three- and seven-year periods, and then one or two at fewer (two years) or more (eight-to-fifteen years). To address this distribution, we developed four categorical variables (with all references to bachelor's degrees meaning bachelor's degrees or equivalent):[3]

- PhD 2-3 yrs. – terminal degrees received 2 to 3 years after the bachelor's degree, represents the reference group.
- PhD 4-5 yrs. – terminal degree received 4 to 5 years after the bachelor's degree
- PhD 6-8 yrs. – terminal degree received 6 to 8 years after the bachelor's degree.
- PhD 9+ yrs. – terminal degree received 9 to 15 years after the bachelor's degree, likely reflecting a period of interruption in the educational trajectory due to work or other factors.

In addition, we specified career pathway variables associated with working experience in research organizations. These variables all refer to the three-year period six years after the researcher received his terminal degree:

---

[3] We treated German and similar European 5-year diplomas as being the terminal equivalent to a master's degree. Diplomas were converted into a "bachelor's degree equivalent" (diploma year minus two years) and a "master degree equivalent" (diploma year).

- Early manager – a dummy variable representing whether the researcher held a management position within the second period, with a management position defined as a chair, vice president, dean, laboratory director or head, institute director or head.

- Postdoc – an indicator variable which registers whether or not the researcher ever held a post-doctoral position in academic, industry, government, or other sector in the second period.

- Early nonacademic – is a dummy variable representing whether or not the researcher worked outside of academia during the second period.

All variables reported so far refer to the career pathway. In addition, we constructed variables representing the wider institutional context in which respondents conducted their research. The institutional context comprises both opportunities for funding – this relates to the *funding program* literature outlined in the literature review (Section 2) – and collaboration with colleagues in other research organizations – this relates to the *collaboration network* literature:

- Early grant – a dummy variable representing whether a grant from outside the university is awarded in the second period.

- Early co-authors – represents the number of different authors or co-authors in articles published in the first period.

- Early specialist – is a measure of interdisciplinarity (cf. Porter et al. 2006, 2008), based on the journal subject category in the Web of Science.

3.3   Mid-career model

The mid-career model posits that the probability of being a creative researcher is a function of several mid-career variables in the third period. This is the three-year period twelve years after the doctoral degree year. This period is taken as a representation of the timeframe before which a researcher is promoted to full professorship or another senior position. Our testing model for the binary outcome variable using probit regression can be written as: $\Pr(Y_t = 1|X_t) = F(X_{t2}'\beta_{t2})$. The mid-career model posits that the probability of being recognized as a creative researcher is a function of several organizational context and collaboration variables. (See Table 1 for a summary listing of variables in this model).

First of all, we suggest that there are again *career pathway* variables. These include whether respondents were promoted to positions in which they could build their own research

agenda, typically a tenured position, whether researcher held a management position or worked outside academia, and how many different jobs a respondent held:

- Time-to-tenure – measures whether or not it took less than seven years from the terminal degree to the first tenured and/or permanent position in academia or senior position in government/industry/hospital sectors. [4]

- Mid manager – a binary categorical variable representing whether the researcher held a management position in the third period, with a management position defined as a chair, vice president, dean, laboratory director or head, institute director or head.

- Mid academic – registers whether or not the researcher worked in a non-academic institution in the third period.

- Job type count – a count of the number of different job positions that the researcher held in the third period.

In addition, there are variables representing the wider institutional context in which respondents conducted their research. The institutional context comprises opportunities to participate in *funding programs*:

- Mid grant – a binary categorical variable representing whether the researcher received a grant from a funding sponsor (not the home university) in the third period.

- Grant count – an ordinal categorical variable that indicates the number of organizational grant sponsors in the third period.

- Grant diversity – an ordinal categorical variable indicating the diversity of grant sponsors during the 2nd 6-years after terminal degree

Furthermore, the institutional context comprises collaboration with colleagues in other research organizations – these variables relate to the *collaboration network* literature:

- Count jobs – counts different organizations where the researcher held a position until the third period.

- Mid co-authors – represents the number of different authors or co-authors in articles published in the third period.

---

[4] We acknowledge that "tenure" might have different meanings in different national research systems. Additionally, in some countries, as in France or the United States, tenure can be attained earlier than in others, such as Germany or Switzerland. For the countries that appear in our CV data, we coded tenure as tenure or permanent academic position or senior position for researchers in government, industry or medical sectors.

- Mid specialist – is a measure of interdisciplinarity in the third period based on the journal subject category in the Web of Science.

Finally, there are two control variables:

- Professional prize – represents whether or not the researcher received a prize from a professional or scientific organization in the third period.

- Prize – represents whether or not the researcher received a prize from any organization (including the university, government laboratory, etc.) in the third period.

4   Results

The CR and matched group are very similar in terms of their academic age (as represented by the year of first receipt of the bachelor's degree) and type of sectoral experience. The year of receipt of the bachelor's or analogous first university degree is similar between the CR and comparison group. For both CRs and matches, 8 percent received bachelor's degrees before 1960. For the CRs, 22 percent and 31 percent received bachelor's degrees in the 1960s and 1970s respectively, compared with 18 percent and 31 percent for the matches. For the balance of CRs, 27 percent and 12 percent received their bachelor's degrees in the 1980s and 1990s respectively, compared with 34 percent and 8 percent for the matches. Additionally, 62 percent of CRs spent their entire career in academia while this figure is 63 percent for the comparison group. Thirty-seven percent of both the CRs and the comparison group members had worked in a governmental institution, while 32 percent of CRs and 34 percent of the matches had private sector work experience.

Descriptive statistics show additional early career characteristics of the researchers in our sample (see Table 2). CRs tend to finish their PhD significantly faster than matched researchers: their share of PhD 2-3 yrs. is more than twice as high, while their share of PhD 6-8 yrs. is lower. In addition, CRs tend to be less disciplinary in focus (Degrees same area) than the matched group, although this difference is not striking. More important, however, is that CRs specialize earlier in their scientific work when compared to their matched counterparts (Early specialist). This suggests early identification of their research concentrations by CRs and also early development of independent research, a finding that is also corroborated by CR's lower number of co-authors in their early publications (Early co-authors).

Interestingly, CRs have more often held postdoctoral positions (Postdoc), but at the same time, they have been promoted into tenured or permanent positions more quickly than

the matched counterparts (Time-to-tenure). This suggests that career speed seems to be important in terms of scientific recognition.

In terms of mid-career model variables, CRs tend to be more exposed to leadership roles (Mid manager) than their matched counterparts. We also find that CRs tend to be funded by fewer sponsors (typically two) indicating stability and continuity in their research sponsorship, while the matched counterparts show a much higher diversity of (about six) grant sources suggesting that additional effort is being placed by matches on pursuing sponsors in generating support for their research.

Breakdowns of these variables for CRs versus the comparison group and for the US versus Europe are shown in Appendix Table 1 and Appendix Table 2 respectively. Significant differences are not observed for most of these variables in simple binary significance tests between these two groups (using the Kolmogorov-Smirnov test). However, CRs versus comparison group researchers differ significantly in terms of early career specialization in publication outlets, grant type diversity, and time-to-tenure. US versus European researchers differ significantly in terms of number of educational institutions, early specialization in publication outlets, mid-career grant acquisition and diversity, and mid-career co-authorship size. A correlation matrix was developed to examine the extent of association between the covariates in the models (see Appendix Table 3 and Appendix Table 4); all correlation coefficients were found to be relatively small.

Although the sample size is the same for early and mid-career periods, the descriptive statistics indicate that missing observations are more prevalent in early career than in mid-career variables. This is not unexpected in that a greater period of time has elapsed for the early career data variables than for those in the mid-career model. Many of the CVs we received were missing information about the date and of the bachelor's (or equivalent) degree and the major area of concentration for this degree. While we did get additional information through follow-up surveys (as noted earlier), we could not complete all the missing information from the surveys and other sources. As a result, variables associated with the time between bachelor's and PhD and disciplinary similarity or diversity have more missing observations than the other variables under analysis.

[INSERT TABLE 2 NEAR HERE]

4.1 Early-career model results

In the following, we present results of probit regressions for the early and mid-to-later career models in order to estimate the common influence of all independent variables (Tables 3 and 4). In addition to ordinary probit regression, we also ran conditional logistic regressions for both the early career and the mid-career models. The results were consistent. We first fit a full model with all the variables of interest and used stepwise elimination to find the parsimonious model that had both the largest number of significant variables and was statistically equivalent to the full model in global model statistical tests. All parsimonious models were tested for equivalence with the full models so the null hypothesis that all the missing coefficients in the parsimonious model are zero was not rejected.[5] Since we are interested in career differences between researchers based either in the United States or Europe, we interpret the early career model separately for the US and European subsamples (Table 3).

The time taken to complete doctoral studies has no relevance to the European CR status but it remains strong and negative on the United States side. Having held a post-doctoral position has a similar effect on both, though the evidence on the European side is a little weaker (Postdoc). Continuing education in the same discipline (Degrees same area) has opposite effects on the two samples: it reduces the probability of being a CR in the United States but it increases in Europe. However, having a multinational study experience (Count countries univ.) emerges as a positive influence on the probability of being a CR for the European subsample, which coincides with the perception that more European scientists move away from their countries than Americans do. The early patterns of publication have an effect on the United States side: a small but negative incidence on the probability of being a CR is given by the larger number of co-authors during the early stage of their career. This means that those researchers who publish their research independently of their doctoral advisor have a higher probability to be recognized as creative. Independent research early in one's career is a signal in the United States that the respondent is a promising young scientist.

[INSERT TABLE 3 NEAR HERE]

---

[5] We also tested the models with other classification algorithms: linear discriminant analysis, impurity function classification trees, and running the probit models with the entire sample using the EU_US dummy as an interaction variable on all the independent variables. The results were very similar and only the classification trees slightly outperformed the models we report in the paper on classification accuracy. The significant variables were the same.

4.2   Mid-career model results

The same stepwise regression procedure as the early-career models was used here. The measures of model fit are much better for the late career models of the overall sample. Overall classification rates are close to 80 percent and the measure of improvement of the models over the null model (McFadden pseudo-R2) is twice to four times better than the previous cases, which also relates to the lowest probabilities of a type I error. Separate models for the United States and Europe are slightly better at detecting non CR cases and classifying them correctly, in other words, they have higher specificity, than correctly classifying CR cases (sensitivity), giving slightly greater confidence in the effects of the variables that reduce the probability of being a CR in the model (Table 4).[6]

[INSERT TABLE 4 NEAR HERE]

Receiving a tenured or other senior position within the first six years of the doctoral degree (Time-to-tenure) increases the probability of being an CR for both Americans and Europeans. However, other features of job trajectory of researchers that make it more likely to be a CR are only important on the US side. The US subsample model shows that the set of factors related to job history that increase the probability of being a CR at this stage of their career includes non-academic positions (Mid nonacademic), management positions (Mid manager), and the number of positions held (Count jobs). However, the probability of being a CR in the US diminishes at a remarkable rate if many different types of positions across diverse sectors were held in the third time period (Job type count). None of these variables seem to have an impact on the EU researchers' chances of being a CR.

Grants have an effect on both the US and the EU researchers' probability of being a CR but in different ways. On the US side, receiving research grants (Mid grant) increase their chances of being a CR but, and to a lesser extent, receiving many different types of grants (Grant type) reduce the probability of being recognized as a CR. On the European side, grants have a negative, but insignificant effect, while an increase in the variety of research grants (Grant type) has a negative and significant effect on the probability of being a CR. The co-occurrence of having many different types of jobs (Job type count), non-academic jobs (Mid nonacademic) and research grants during this period (Mid grant), represented by an interaction variable of the three categorical variables, also diminishes rather strongly the

---

[6] The same approach mentioned in the case of the early-career model with other classification algorithms was used for the mid-career model and the results were shown also to be consistent with those reported here.

probability of being a CR in Europe but has no effect on US researchers. Finally, having many co-authors (Mid co-authors) also reduces the probability of being recognized as a CR in the US but does not appear to have an incidence on Europeans.

5   Summary and Discussion

This paper compares two groups of scientists to understand three dimensions in the institutional context that differentiates recognition of creative research in nanotechnology and human genetics: career pathway, funding and research collaboration. The first group consists of researchers that were recognized as "creative" by their peers, either per individual nomination in an international survey, and/ or because these researchers were awarded prestigious research awards in their field. The second group consists of researchers that were "matched" according to early publication behavior variables. In this way, we build matched pairs that are almost identical in their early publication behavior. These two groups of scientists are similarly high achievers across scientific outcomes such as scientific publications and citation of these publications. Both groups are at the top of the distribution on these measures because the matching strategy used to select the control group was based on many attributes of achievement such as early productivity, the time in which their research activity began, and their focused field. We probe which career-based factors differentiate these two groups of scientists in terms of scientific recognition. Our particular focus is on a comparison between US-based scientists and Europe-based scientists.

Our paper presents several key results. We find that *funding context* in the early-career model does not significantly explain differences between CRs and matches, whether in the US or in Europe. In contrast, there are several *career pathway* and *network* variables that explain differences between CR and matches. In the US, early scientific recognition is associated with broad academic education, fast completion of PhD, and a record of independent post-doctoral research, whereas in Europe these factors are much less influential or even point in the opposite direction. For example, early scientific recognition in Europe is often associated with a disciplinary affiliation. While creative European scientists may study in various universities abroad, they tend to stay in one disciplinary context. In other words, for the Europeans, institutional diversity is coupled with disciplinary homogeneity.

Regarding the *funding context* in the mid-career model, we find that creative accomplishments are enabled via stable grant money in the US, whereas grants have an

overall negative effect of being recognized as creative in Europe. The more grant money and the more diverse grant resources, the less likely European scientists are to be recognized as creative. Regarding *career pathway* variables, both in the United States and Europe fast job promotion within academia is a strong predictor for future recognition. However, there is – again – a clear divide across the Atlantic regarding other mid-career factors: work experience inside and outside academia and research leadership are connected to scientific recognition in the United States, but negligible and even negative in Europe. This means that changing job to a company and then returning to academia helps US scientists to be recognized as creative, while this is not observed as strongly in Europe where scientists have higher chances of being recognized as creative if they stay their whole career inside the university or other academic research institutes. Regarding *collaboration network* variables, US scientists benefit from being mobile in the labor market, while this finding is not evident for European scientists.

How can these results be interpreted? Our study confirms earlier findings in the literature that educational experience and early independent research are important factors that regulate the distribution of scientific recognition. Our confirmatory evidence is noteworthy because we operationalize scientific recognition not as publication or citation scores, but via a combined survey-prizewinner identification method. Empirical results that hold across different approaches are indicative of substantive sociological findings. In addition, we believe that our results are particularly robust because we use a comparison group based, among other variables, on early publication behavior. Having a control group design is particularly important in order to hold constant the cumulative advantage effect that has been shown to be pervasive in studies of careers in academia and other areas of social life.

In addition to confirmatory results, our study also adds and qualifies earlier findings. First, we directly compare groups of researchers based in the US and Europe. Our comparison suggests that certain strengths and advantages in the institutional context of research in the US that emerged in the beginning of the 20th century are still intact. In that respect, our results show that classical analyses on the American University (Parsons and Platt, 1973) or on the historical emergence of the United States as the leading scientific nation (Ben-David, 1971), are still highly informative because they point to a close relationship between institutional structures, such as the graduate school, and scientific performance (Cole, 2010). For example, our results suggest that the academic labor market in the United States – compared to Europe – offers an open arena for developing scientific reputation because universalistic criteria of merit and individual performance receive strong institutional support. Both native- and

19

foreign-born scientists working in the United States are able to engage in this arena (Stephan and Levin, 2001).

A second qualification relates to the literature on *career pathways* and scientific performance. Since the 1970s the number of postdoctoral positions has expanded at a much higher rate than tenured faculty positions, particularly in growth fields like the biomedical sciences; however, these temporary positions are believed to be problematic because they tend to be associated with lower academic performance and delayed entry in the academic labor market (Dietz and Bozeman, 2005; Stephan, 2008). Our results show, however, that CRs not only hold more often postdoctoral positions than matched scientists, but that American CRs use these positions more effectively as opportunity for the development of independent research: they more often receive tenure within the first seven years after completion of their PhD than researchers in the matched group. While post-doctoral work is associated with independent research in the United States, in Europe younger scientists publish more often with their doctoral advisors. This impedes their chances to be recognized as independent producers of creative research.

Third, our findings qualify findings from the *collaboration network* literature. In the US, CRs collaborate less intensively with their academic mentors than in Europe. This suggests that doctoral advisors play a less central role in their collaboration networks in the US than in Europe. On the other hand, while European CRs are much more likely to have remained in the same academic discipline in their educational stage, the United States context offers more freedom for students and scientists to make their own way and mix and match disciplines. Our findings suggest that in Europe, if the researcher does not publish and collaborate in a clearly defined (sub-) discipline, it is to their disadvantage in gaining scientific recognition. These findings suggest that academic networks in Europe tend to cluster more around (sub-) disciplines and tend to reflect hierarchical work relationships, while academic networks in the US tend to be more multidisciplinary and show more structural holes between mentors and students. In sum, these results tend to confirm findings from the network literature about structural holes and collaborative brokerage primarily for the institutional context of the US. Therefore, findings on collaboration networks in academia may not be fully generalizable outside the institutional context of the US.

The critical lesson from our data is that organizational and institutional context has very important consequences for the career paths of researchers and makes an enormous difference in becoming recognized as creative scientist. We must keep in mind that the researchers in the matched set were carefully selected to have very similar potential during the

early stages of their career. So the differences reported by this study highlight the role of institutional influences. This suggests that if policy measures are considered, they need to be tailored to the broad institutional patterns that exist in the US and Europe. For example, one could argue that appropriate policy measures would be directed towards increasing collaboration across disciplinary and institutional boundaries in Europe. In countries with strong hierarchical work relations in research, like Germany or France, appropriate policy measures could include to better support the early independence of young scientists by providing them with their own research money and/or by encouraging research organizations to monitor and guide publication behavior of their scientific staff.

Another set of opportunities lies in the reform of advanced education and training, primarily at European universities. These include streamlining doctoral requirements, and supporting broader disciplinary curriculum in the early years of university education. Also relevant, both for Europe and the United States, is in the nature of research sponsorship. Our finding that creative research is associated with consistent research funding suggests that research policymakers should consider how they can support long-term research programs that not only reduce the need for grant hopping but which also provide the time to develop novel research approaches.

In interpreting the career and institute differences reported in this paper, it is important to keep in mind that both the US and European scientists included in our study are all successful. While CRs have received exceptional scientific acknowledgement by their peers through nominations and prizes, the control group also comprises a highly productive and well-cited set. Our findings suggest that there is more homogeneity in the ways in which productive scientists gain recognition as being creative researchers in the US whereas in Europe there is more heterogeneity with fewer consistently significant factors observed. We recognize that in treating Europe as an entity, we are unable to probe institutional differences and variations regarding career pathways within Europe. The inherently limited number of scientists recognized as being highly creative in our initial nomination process, even in the relatively large fields of nanotechnology and human genetics, restricts the ability to breakdown our continental-level samples into smaller geographical units (this is true for the US as well as Europe).

Extensions of our approach, including identifying more CRs from other scientific domains, and employing additional methodologies (including qualitative case studies) will be helpful in further pursuing the effects of national and regional institutional and career pathway differences. This includes extensions of the approach to other countries and regions in the

world, beyond the US and Europe. Replicating and applying both the CR identification method and the matching procedure to additional research domains would, with increased numbers of scientists, would also improve the explanatory power of the early and mid-career models. We further recognize that our paper is exploratory. We have operationalized concepts in the extant literature, including notions related to funding programs, career pathways, and collaboration networks, to investigate relationships between scientific recognition and institutional context. Our aim has been to offer a footing on which to develop more systematic hypotheses, further methodological advances, and additional evidence to improve our understanding of the complex relationships between organizational and institutional contexts and scientific creativity.

Acknowledgements

References

Aldersey-Williams, H., 1995. The Most Beautiful Molecule. An Adventure in Chemistry. London: Aurum Press.

Allison, P.D., Long, S., 1990. Departmental effects on scientific productivity. American Sociological Review 55, 469-478

Amabile, T.M., 1996. Creativity in Context. Westview Press, Boulder, Colorado.

Arora, S.K., Porter, A.L., Youtie, J., Shapira, P., 2012. Capturing new developments in an emerging technology: An updated search strategy for identifying nanotechnology research outputs. Atlanta, GA, Georgia Tech Program in Science, Technology and Innovation Policy.

Azoulay, P., Graff Zivin, J.S., Manso, G., 2009. Incentives and Creativity: Evidence from the Academic Life Sciences, NBER Working Paper No. 15466, National Bureau of Economic Research Cambridge, M.A.

Ben-David, J., 1971. The Scientist's Role in Society. Chicago, University of Chicago Press.

Bonaccorsi, A., 2007. Explaining poor performance of European science: Institutions versus policies. Science and Public Policy 34, 5, 303-316.

Burt, R. S., 2004. Structural holes and good ideas. American Journal of Sociology 110, 349-399.

Carter, R. 2004. Language and creativity: The art of common talk. Abingdon, UK, Routledge.

Cañibano, C., Bozeman, B., 2009. Curriculum vitae method in science policy and research evaluation: The state-of-the-art. Research Evaluation 18, 86-94.

Cañibano, C., Otamendi, J., Andújar, I., 2008. Measuring and assessing researcher mobility from CV analysis: The case of the Ramón y Cajal programme in Spain. Research Evaluation 17, 17-31.

Cole, J.R., 2010. The Great American University: Its Rise to Preeminence, Its Indispensable National Role, Why It Must Be Protected. New York: Basic Books

Csikszentmihalyi, M., 1997. Creativity. Flow and the Psychology of Discovery and Invention. New York, Harper Collins.

Csikszentmihalyi, M., 1999. Implications of a systems perspective for the study of creativity, in: Sternberg, R.J. (Ed), Handbook of Creativity. Cambridge: Cambridge University Press, pp. 313-335.

Dietz, J. S., Bozeman, B., 2005. Academic careers, patents, and productivity: industry experience as scientific and technical human capital. Research Policy 34, 349-367.

DiPrete, T. A., Eirich, G. M., 2006. Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. Annual Review of Sociology 32, 271–297.

ERAB, 2009. Preparing Europe for a New Renaissance. A Strategic View of the European Research Area. European Research Area Board. Brussels, European Commission, Directorate-General for Research.

Feist, G., Gorman, M., 1998. The Psychology of Science: Review and Integration of a Nascent Discipline. Review of General Psychology. 2, 3-47.

Fleming, L., Mingo, S., Chen, D., 2007. Collaborative brokerage, generative creativity, and creative success. Administrative Science Quarterly 52, 443-475

Fleming, L., Szigety, M., 2006. Exploring the tail of creativity. An evolutionary model of breakthrough innovation. Ecology and Strategy 23, 335-359

Gaughan, M., Ponomariov, B., 2008. Faculty publication productivity, collaboration, and grants velocity: Using curricula vitae to compare center-affiliated and unaffiliated scientists. Research Evaluation 17, 103-110

Gaughan, M., Robin, S., 2004. National science training policy and early scientific careers in France and the United States. Research Policy 33, 569-581.

Giles, M., Ski, C., Vrdoljak, D., 2009. Career pathways of science, engineering and technology research postgraduates. Australian Journal of Education 53, 69-86.

Heinze, T., Bauer, G., 2007. Characterizing creative scientists in nano S&T: Productivity, multidisciplinarity, and network brokerage in a longitudinal perspective. Scientometrics 70, 811-830.

Heinze, T., 2008. How to sponsor ground-breaking research: A comparison of funding schemes. Science and Public Policy 3, 5, 302-318.

Heinze, T., Shapira, P., Rogers, J.D., Senker, J.M., 2009. Organizational and institutional influences on creativity in scientific research. Research Policy 38, 610-623.

Heinze, T., Shapira, P., Senker, J., Kuhlmann, S., 2007. Identifying creative research accomplishments: Methodology and results for nanotechnology and human genetics. Scientometrics 70, 1, 125-152.

Hemlin, S., Allwood, C.M., Martin, B.R., 2004. Creative Knowledge Environments. The Influences on Creativity in Research and Innovation. Cheltenham, Edward Elgar.

Hermanowicz, J.C., 2009. Lives in Science: How Institutions Affect Academic Careers. Chicago, University of Chicago Press.

Hessenbruch, A., 2004. Nanotechnology and the Negotiation of Novelty. In D. Baird, A. Nordmann & J. Schummer, Discovering the Nanoscale (pp. 135-144). Amsterdam: IOS Press.

Hollingsworth, J.R., 2003. Research organizations and major discoveries in twentieth-century science: A case of excellence in biomedical research. WZB Discussion Paper P02-003, Social Science Research Center, Berlin.

Hollingsworth, J.R., 2004. Institutionalizing excellence in biomedical research: The case of the Rockefeller University, in: Stapleton, D.H. (Ed), Creating a Tradition of Biomedical Research. New York, Rockefeller University Press, 1–49.

Kay, Lily E. 1993. The Molecular Vision of Life: Caltech, the Rockefeller Foundation, and the Rise of the New Biology Oxford: Oxford University Press.

Kohler, Robert E. 1991. Partners in Science: Foundations and Natural Scientists 1900-1945. Chicago/London: University of Chicago Press.

Laudel, G., 2006. The art of getting funded: how scientists adapt to their funding conditions. Science and Public Policy, 33, 489–504.

Lightman, A., 2005. The Discoveries. Great Breakthroughs in 20th Century Science. New York, Pantheon Books

Melin, G., Danell, R., 2006. The top eight percent: Development of approved and rejected applicants for a prestigious grant in Sweden. Science and Public Policy, 33, 702-712.

Merton, R. K., 1968. On the history and systematics of sociological theory, in: Merton, R.K. (Ed), Social Theory and Social Structure. New York, The Free Press, 1–38.

Merton, R.K., 1973.The Matthew Effect in Science, in Merton, R.K. (Ed.) The Sociology of Science. Theoretical and Empirical Investigations. Chicago, Chicago University Press, 439-459.

Merton, R.K., Barber, E., 2004. The travels and adventures of serendipity: A study in socio-logical semantics and the sociology of science. Princeton, N.J., Princeton University Press.

Mayer, R.E. 1999. Fifty years of creativity research, in: Sternberg, R.J. (Ed), Handbook of Creativity. Cambridge: Cambridge University Press, 449-460.

Mody, C.C.M., 2011. Instrumental Community. Probe Microscopy and the Path to Nanotechnology. Cambridge: MIT Press

Mumford, M.D., Conelly, M.S., Scott, G., Espejo, J., Sohl, L.M., Hunter, S.T., Bedell, K.E., 2005. Career experiences and scientific performance: A study of social, physical, life and health sciences. Creativity Research Journal 17, 105-129.

Münch, R., 2008. Stratification by evaluation: Mechanisms of constructing status hierarchies in research. Zeitschrift für Soziologie 37, 60-80.

National Research Council, 2006. Evaluation of the Markey Scholars Program. Washington D.C., The National Academies Press.

Parsons, T., Platt, G. M., 1973. The American University. Cambridge, M.A., Harvard University Press.

Porter, A. L., Roessner, J. D., Heberger, A. E. 2008. How interdisciplinary is a given body of research? Research Evaluation 17, 4, 273-282.

Porter, A.L., Roessner, J. D., Cohen, A. S., Perreault, M., 2006. Interdisciplinary research: Meaning, metrics, and nurture. Research Evaluation 15, 3, 187-195.

Porter, A.L., Youtie, J., 2009, How interdisciplinary is nanotechnology? Journal of Nanoparticle Research 11, 5, 1023-1041.

Rafols, I., Meyer, M., 2010. Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience, Scientometrics 82, 12, 263-287.

Sandström, U., 2009. Combining curriculum vitae and bibliometric analysis: Mobility, gender and research performance. Research Evaluation 18, 135-142

Simonton, D. K., 1999. Origins of genius: Darwinian perspectives on creativity. Oxford, Oxford University Press.

Simonton, D. K., 2004. Creativity in Science: Chance, Logic, Genius, and Zeitgeist. Cambridge University Press, Cambridge, UK.

Stephan, P.E., Levin, S.G., 2001. Exceptional contributions to US science by the foreign-born and foreign-educated. Population Research and Policy Review 20, 59-79

Stephan, P.E., 2008. Job market effects on scientific productivity, in Albert, M., Schmidtchen, D., Voigt, S., (Eds.), Scientific competition. Tübingen, Mohr Siebeck, 11-29.

Sternberg, R. J., 2003. Wisdom, Intelligence, and Creativity Synthesized. Cambridge, Cambridge University Press.

Stumpf, H., 1995. Scientific creativity: A short overview. Educational Psychology Review 7, 3, 225-241.

Sulston, John, und Georgina Ferry. 2002. The Common Thread: A Story of Science, Politics, Ethics and the Human Genome. Washington DC: Joseph Henry Press.

Uzzi, B., J. Spiro, 2005. Collaboration and creativity. The small world problem. American Journal of Sociology, 111, 447–504.

Wellcome Trust, 2010. Human Genetics 1990-2009. Portfolio Review. London, The Wellcome Trust.

Whitley, R., 2000. The Intellectual and Social Organization of the Sciences, 2nd edition. Oxford, Oxford University Press.

Youtie, J., Shapira, P., Rogers, J., 2009. Blind matching versus matchmaking: Comparison group selection for creative researchers. IEEE Conference proceedings, Science and Innovation Policy, 2009 Atlanta Conference. (downloaded on 9 September 2011 from http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5367848).

Zuckerman, H., 1977. Scientific Elite: Nobel Laureates in the United States. New York, Free Press.

Zuckerman, H., 2010. The Matthew effect writ large and larger. A study in sociological semantics. Zeitschrift für Soziologie 20, 309-340.

Table 1 Variable Descriptions

Variables	Data sources	Literature connection	Variable description
CR-Comparison			Researcher type: creative researcher or matched researcher. 1 for CR and 0 for Match
Early-Career Model			
PhD 2-3 yrs.	CV	Career pathway literature	Dummy variable, 1 if the year lag between bachelors' degree and PhD degree ranges from 2-3, 0 otherwise
PhD 4-5 yrs.	CV		Dummy variable, 1 if the year lag between bachelors' degree and PhD degree ranges from 4-5, 0 otherwise
PhD 6-8 yrs.	CV		Dummy variable, 1 if the year lag between bachelors' degree and PhD degree ranges from 6-8, 0 otherwise
PhD 9+ yrs.	CV		Dummy variable, 1 if the year lag between bachelors' degree and PhD degree ranges from 9 and above, 0 otherwise
Count univ.	CV		Number of institutions attended from bachelor's degree to PhD degree
Postdoc	CV		Dummy variable, 1 if post-doctoral experience, 0 otherwise
Early nonacademic	CV		Dummy variable, 1 if worked in non-academic institution within the first 6 years of getting terminal degree, otherwise 0
Early manager	CV		Dummy variable, whether held a management position within the first 6 years of getting terminal degree, yes = 1, no = 0
Early grant	CV	Funding program literature	1 if awarded non-university grant in the first 6-years since getting final degree, otherwise 0
Degrees same area	CV	Network literature	Compares the major discipline of bachelors' degree and terminal degree (PhD/MD). 1 if same, 0 if different
Count countries univ.	CV		Dummy variable, 1 if studies in two or more countries for education, 0 otherwise
Early co-authors	WOS		Number of different coauthors in Phase 1, i.e. from one year prior to getting the first PhD/MD degree to one year after getting his first terminal degree
Early specialist	WOS		Specialization score of all Web of Science indexed publications in Phase 1

Table 1 Variable Descriptions (Continued)

Mid-Career Model			
Time-to-tenure	CV	Career pathway literature	Dummy variable, 1 if it took less than 7 years from getting PhD/MD degree to get the first tenured position or senior position in government labs/industry/hospital; otherwise 0
Mid manager	CV		Dummy variable, whether held a management position within the 2nd 6-years after terminal degree, yes 1 or no 0
Job type count	CV		Count of job types within the 2nd 6-years after terminal degree
Mid nonacademic	CV		Dummy variable, 1 if worked in non-academic institution within the 2nd 6-years of terminal degree, otherwise 0
Grant count	CV	Funding program literature	Ordinal variable, 1 if zero grant awarded; 2 if received grants from one or two different organizations; 3 if received grants from three or four different organizations; 4 if received grants from five or above different organizations
Mid grant	CV		Dummy variable: 1 if awarded non-university grant in the 2nd 6-years after final degree, otherwise 0
Grant diversity	CV		Ordinal categorical variable indicating the diversity of grant sponsors during the 2nd 6-years after terminal degree
Count jobs	CV	Network literature	Number of different organizations worked
Mid co-authors	WOS		Number of different coauthors in Phase 2, i.e. from the 5th year to the 7th year after getting the first PhD/MD degree
Mid specialist	WOS		Specialization score of all Web of Science indexed publications in Phase 2
Professional prize	CV	Control variables	Dummy variable, 1 if awarded professional prize in late career, otherwise 0
Prize	CV		Dummy variable, 1 if awarded any prize in late career, otherwise 0

Note: CV = Curriculum Vitae; WOS = Web of Science

Table 2 Descriptive Statistics

Variable	Observations	Mean	Std. Dev.	Min	Max
CR-Comparison	152	0.50	0.50	0	1
PhD 2-3 yrs.	119	0.14	0.35	0	1
PhD 4-5 yrs.	119	0.40	0.49	0	1
PhD 6-8 yrs.	119	0.37	0.48	0	1
PhD (+ yrs.	119	0.08	0.28	0	1
Degrees same area	120	0.46	0.50	0	1
Count countries univ.	150	0.17	0.37	0	1
Postdoc	152	0.56	0.50	0	1
Count jobs	149	5.49	3.06	1	19
Early nonacademic	142	0.59	0.49	0	1
Early manager	145	0.08	0.27	0	1
Early grant	147	0.33	0.47	0	1
Count univ.	148	1.76	0.67	1	4
Mid grant	147	0.54	0.50	0	1
Early co-authors	147	9.35	13.76	0	102
Early specialist	141	0.59	0.29	0	1
Time-to-tenure	152	0.41	0.49	0	1
Mid management	147	0.29	0.46	0	1
Job type count	146	1.36	0.55	1	4
Mid nonacademic	146	0.48	0.50	0	1
Professional Prize	150	0.20	0.40	0	1
Prize	150	0.47	0.50	0	1
Grant diversity	146	4.08	5.74	0	35
Grant count	145	2.59	1.21	1	4
Mid co-authors	149	26.59	49.45	0	546
Mid specialist	147	0.52	0.22	0	1

Table 3 Early Career Model: United States versus European Countries

	United States		European Countries	
	Full Model	Parsimonious Model	Full Model	Parsimonious Model
PhD 6-8 yrs.	-1.464**	-.930**	-0.189	
	(0.728)	(0.401)	(0.806)	
PhD 4-5 yrs.	-0.682		-0.027	
	(0.694)		(0.827)	
PhD 9+ yrs.	-1.141		-1.253	
	(0.970)		(1.061)	
Count univ.	-0.217		0.551	
	(0.334)		(0.57)	
Degrees same area	-0.868*	-0.883**	0.941*	.852*
	(0.461)	(0.399)	(0.553)	(0.462)
Count countries univ.	0.329		1.480**	1.368**
	(0.489)		(0.734)	(0.590)
Postdoc	0.749*	.870**	0.789	.787*
	(0.433)	(0.400)	(0.592)	(0.457)
Early nonacademic	0.209		0.145	
	(0.528)		(0.623)	
Early grant	0.553		-0.455	
	(0.540)		(0.687)	
Early co-authors	-0.087*	-.083**	0.001	
	(0.045)	(0.035)	(0.031)	
Early specialist	0.142		0.910	
	(0.955)		(0.995)	
Constant	1.310**	0.731	-2.380*	-1.250***
	(0.363)	(0.422)	(1.340)	(0.467)
Observations	61	61	39	39
Prob > chi2	0.0108**	.0003***	0.1916	.0131**
Pseudo-R2	0.2896	0.251	0.275	0.200
Sensitivity Pr(+ D)	80.00%	71.43%	72.22%	61.11%
Specificity Pr(-~D)	74.19%	68.75%	76.19%	76.19%
Correctly classified	77.05%	70.15%	74.36%	69.23%

Standard errors in parentheses; * p<0.10, ** p<0.05, *** p<0.01

Table 4 Mid-Career Model: United States versus European Countries

	United States		European Countries	
	Full Model	Parsimonious Model	Full Model	Parsimonious Model
Time-to-tenure	1.040**	1.080**	0.821*	0.850*
	(0.475)	(0.462)	(0.475)	(0.442)
Mid manager	1.918***	1.786***	0.143	
	(0.740)	(0.689)	(0.576)	
Job type count	-2.512***	-2.280***	0.673	0.721
	(0.895)	(0.835)	(0.589)	(0.520)
Mid nonacademic	1.630**	1.421*	-0.649	
	(0.786)	(0.750)	(0.597)	
Count jobs	0.268**	.260***	-0.0344	
	(0.109)	(0.099)	(0.0778)	
Mid grant	1.912**	2.065***	-0.842	
	(0.800)	(0.736)	(0.525)	
Grant diversity	-0.217***	-0.187***	-0.449	-.492***
	(0.063)	(0.051)	(0.509)	(0.140)
Professional prize	1.152*	1.349**	-1.295*	-1.518**
	(0.602)	(0.542)	(0.706)	(0.722)
Prize	0.347		0.969	1.192**
	(0.516)		(0.612)	(0.601)
Grant count	0.156		0.141	
	(0.292)		(0.968)	
Mid co-authors	-0.040**	-.0370**	-0.002	
	(0.018)	(0.016)	(0.004)	
Mid-specialist	-0.961		0.0558	
	(1.032)		(0.968)	
(Job type count) x (Mid non-academic) x (Mid grant)				-0.960**
				(0.415)
Constant	0.768	0.342	-0.044	-0.378
	(1.454)	(0.749)	(1.389)	(0.679)
Observations	73	73	58	58
Prob > chi2	0.0000***	0.0000***	0.0035***	0.0000***
Pseudo-R2	0.431	0.414	0.365	0.382
Sensitivity Pr(+ D)	78.38%	75.68%	75.86%	75.86%
Specificity Pr(-~D)	80.56%	77.78%	79.31%	82.75%
Correctly Classified	79.45%	76.71%	77.59%	79.31%

Standard errors in parentheses; * p<0.10, ** p<0.05, *** p<0.01

Appendices

Appendix Figure 1. Histogram of Logged Number of Full Career Publications:
Creative Researchers versus Comparison Group Pool and Random Sample
(logged distribution and unlogged statistics shown)

Comparison Group Pool: Number of cases=757, Mean=41.0 (s.d. 46.4), Median=27; Creative Researcher: Number of cases=76, Mean=85.5 (s.d. 80.0), Median=66. Random Sample: Number of cases=2000, Mean=4.7 (s.d. 8.8), Median=2.

**Appendix Figure 2. Histogram of Logged Number of Full Career Citations:
Creative Researchers versus Comparison Group Pool and Random Sample**
(logged distribution and unlogged statistics shown)

Comparison Group Pool: Number of cases=756, Mean=1087.5 (s.d. 1441.2), Median=560; Creative Researcher: Number of cases=76, Mean=4889.5 (s.d. 5616.9), Median=2565.5; Random Sample: Number of cases=2000, Mean=49.5 (173.4), Median=2.

Appendix Table 1. Descriptive Statistics: Creative Researchers versus Comparison Group

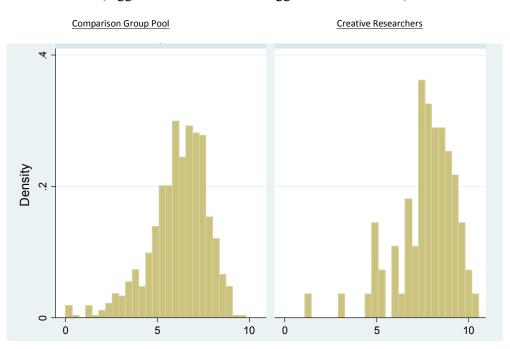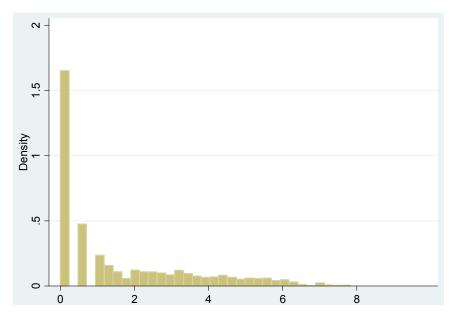| Variable	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max	
		Creative Researcher					Comparison Group				
PhD 2-3 yrs.	59	0.20	0.41	0	1	60	0.08	0.28	0	1	
PhD 4-5 yrs.	59	0.42	0.50	0	1	60	0.38	0.49	0	1	
PhD 6-8 yrs.	59	0.29	0.46	0	1	60	0.45	0.50	0	1	
PhD 9+ yrs.	59	0.08	0.28	0	1	60	0.08	0.28	0	1	
Degrees same area	60	0.42	0.48	0	1	60	0.50	0.50	0	1	
Count countries univ.	74	0.20	0.40	0	1	76	0.13	0.34	0	1	
Postdoc	76	0.61	0.49	0	1	76	0.51	0.50	0	1	
Count univ.	74	1.73	0.65	1	3	74	1.80	0.70	1	4	
Early nonacademic	68	0.62	0.49	0	1	74	0.57	0.50	0	1	
Early manager	71	0.08	0.28	0	1	74	0.07	0.25	0	1	
Early grant	73	0.34	0.48	0	1	74	0.31	0.47	0	1	
Early co-authors	76	7.46	8.28	0	43	71	11.37	17.71	0	102	
Early specialist	71	0.64	0.25	0	1	70	0.54	0.32	0	1	**
Time-to-tenure	76	0.51	0.50	0	1	76	0.30	0.46	0	1	**
Mid manager	71	0.35	0.48	0	1	76	0.24	0.43	0	1	
Job type count	71	1.37	0.59	1	4	75	1.35	0.51	1	3	
Mid nonacademic	71	0.51	0.50	0	1	75	0.45	0.50	0	1	
Count jobs	73	5.95	3.23	2	19	76	5.05	2.84	1	17	
Professional prize	74	0.24	0.43	0	1	76	0.16	0.37	0	1	
Prize	74	0.55	0.50	0	1	76	0.39	0.49	0	1	
Mid grant	73	0.47	0.50	0	1	74	0.62	0.49	0	1	
Grant diversity	75	2.29	3.74	0	29	71	5.96	6.81	0	35	***
Grant count	75	2.25	0.96	1	4	71	2.94	1.15	1	4	***
Mid co-authors	76	21.68	21.00	1	107	73	31.70	67.20	0	546	
Mid specialist	75	0.51	0.20	0.16	1	72	0.53	0.24	0.16	1	

* p<0.10, ** p<0.05, *** p<0.01

Appendix Table 2. Descriptive Statistics: Europe versus US

Variable	Obs	Mean	US Std. Dev.	Min	Max	Obs	Mean	Europe Std. Dev.	Min	Max	
PhD 2-3 yrs.	72	0.14	0.35	0	1	47	0.15	0.36	0	1	
PhD 4-5 yrs.	72	0.40	0.49	0	1	47	0.40	0.50	0	1	
PhD 6-8 yrs.	72	0.39	0.49	0	1	47	0.34	0.48	0	1	
PhD 9+ yrs.	72	0.07	0.26	0	1	47	0.11	0.31	0	1	
Degrees same area	73	0.44	0.50	0	1	47	0.49	0.51	0	1	
Count countries univ.	78	0.21	0.41	0	1	72	0.13	0.33	0	1	
Postdoc	80	0.61	0.49	0	1	72	0.50	0.50	0	1	
Count univ.	76	2.013	0.60	1	4	72	1.50	0.65	1	3	***
Early nonacademic	76	0.54	0.50	0	1	66	0.65	0.48	0	1	
Early manager	76	0.04	0.20	0	1	69	0.12	0.32	0	1	
Early grant	79	0.38	0.49	0	1	68	0.26	0.44	0	1	
Early co-authors	75	7.04	7.41	0	41	72	11.75	17.91	0	102	
Early specialist	70	0.65	0.28	0	1	71	0.53	0.29	0	1	*
Time-to-tenure	80	0.38	0.49	0	1	72	0.44	0.50	0	1	
Mid manager	77	0.26	0.44	0	1	70	0.33	0.47	0	1	
Job type count	77	1.38	0.51	1	3	69	1.33	0.59	1	4	
Mid nonacademic	77	0.51	0.50	0	1	69	0.45	0.50	0	1	
Count jobs	79	5.14	3.02	1	19	70	5.89	3.08	2	17	
Professional prize	80	0.23	0.42	0	1	70	0.17	0.38	0	1	
Prize	80	0.54	0.50	0	1	70	0.40	0.49	0	1	
Mid grant	79	0.66	0.48	0	1	68	0.41	0.50	0	1	**
Grant diversity	80	5.99	7.05	0	35	66	1.76	1.75	0	6	***
Grant count	80	2.94	1.14	1	4	66	2.17	0.90	1	4	***
Mid co-authors	77	17.74	14.58	1	81	72	36.06	68.52	0	546	*
Mid specialist	77	0.53	0.21	0.2	1	70	0.51	0.23	0.2	1	

* p<0.10, ** p<0.05, *** p<0.01

Appendix Table 3: Correlation Matrix for Early Career Model

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. CR-comparison	1.00													
2. PhD 2-3 yrs.	0.17	1.00												
3. PhD 4-5 yrs.	0.08	-0.32	1.00											
4. PhD 6-8 yrs.	-0.21	-0.30	-0.66	1.00										
5. PhD 9+ yrs.	0.01	-0.11	-0.25	-0.23	1.00									
6. Degrees same area	-0.07	0.05	-0.16	0.12	0.01	1.00								
7. Count countries univ.	0.17	-0.06	-0.06	-0.01	0.20	-0.02	1.00							
8. Postdoc	0.22	-0.04	-0.12	0.18	-0.06	-0.02	0.00	1.00						
9. Count univ.	0.00	-0.12	0.06	-0.06	0.14	-0.15	0.30	-0.21	1.00					
10. Early nonacademic	0.06	0.03	-0.05	0.05	-0.04	-0.18	-0.08	0.10	-0.10	1.00				
11. Early manager	0.05	-0.10	0.17	-0.13	0.06	0.05	-0.14	-0.17	-0.04	0.00	1.00			
12. Early grant	0.04	-0.13	0.02	0.04	0.06	-0.12	0.11	0.09	0.14	-0.29	0.02	1.00		
13. Early co-authors	-0.08	-0.03	-0.19	0.17	0.08	0.01	0.19	0.01	0.01	0.10	-0.06	0.30	1.00	
14. Early specialist	0.21	0.18	0.01	-0.09	-0.07	0.19	0.02	0.09	0.05	0.01	-0.11	-0.03	-0.29	1.00

Column number refers to variable with same row number.

Number of observations=101

Appendix Table 4: Correlation Matrix for Mid-Career Model

Varaibles	1	2	3	4	5	6	7	8	9	10	11	12	13
1. CR-comparison	1.00												
2. Time-to-tenure	0.24	1.00											
3. Mid manager	0.13	0.20	1.00										
4. Job type count	0.01	-0.02	0.25	1.00									
5. Mid nonacademic	0.02	-0.02	0.24	0.68	1.00								
6. Job type count	0.15	0.00	0.21	0.33	0.29	1.00							
7. Professional prize	0.11	0.11	0.01	-0.04	-0.08	-0.06	1.00						
8. Prize	0.11	0.17	0.12	0.00	-0.01	-0.07	0.53	1.00					
9. Mid grant	-0.22	-0.01	-0.05	-0.04	-0.19	-0.20	0.04	0.21	1.00				
10. Grant diversity	-0.30	0.04	0.05	0.01	-0.07	-0.16	0.03	0.03	0.43	1.00			
11. Grant count	-0.29	-0.06	0.04	-0.02	-0.11	-0.18	-0.01	0.09	0.58	0.68	1.00		
12. Mid co-authors	-0.09	0.19	0.14	0.10	0.11	0.05	0.03	0.12	0.09	-0.02	0.06	1.00	
13. Mid specialist	-0.01	-0.07	0.07	-0.05	0.03	-0.07	-0.17	-0.16	-0.12	-0.14	-0.07	-0.22	1.00

Column number refers to variable with same row number.

Number of observations=131