Matthew: Effect or Fable?

Pierre Azoulay
MIT & NBER
Sloan School of Management
100 Main Street, E62-482
Cambridge, MA 02142

Toby Stuart
Harvard Business School
211 Rock Center
Soldiers Field
Boston, MA 02163

Yanbo Wang
Boston University
School of Management
595 Commonwealth Avenue
Boston, MA 02215

October 14, 2010

Abstract

Status effects occur when actors are accorded differential recognition for their efforts depending on their location in a status ordering, *holding constant the quality of these efforts.* In practice, because it is difficult to measure quality, this ceteris paribus proviso often precludes convincing empirical assessments of the magnitude of status effects. We address this problem by examining the impact of a major status-conferring prize that shifts actors' positions in a prestige ordering. Specifically, using a precisely constructed matched sample, we estimate the effect of a scientist becoming a Howard Hughes Medical Investigator (HHMI) on citations to articles the scientist published *before* the prize was awarded. We find evidence of a post-appointment citation boost, but the effect is small and confined to a short window of time. Consistent with theories of status, however, the effect of the prize is significantly larger when there is uncertainty about producer and product quality.

Keywords: sociology of science, status, stratification, Matthew effect.

* We gratefully acknowledge the financial support of the National Science Foundation through its SciSIP Program (Award SBE-0738142). We thank Matthew Bothner, Benjamin Jones, Casey Ichniowski, and Ezra Zuckerman, as well as seminar participants at Berkeley and Dartmouth, for many helpful comments. The usual disclaimer applies.

1 Introduction

An accepted truth in sociology is that actors in social systems sort into status orderings. These ubiquitous hierarchies are of interest in the discipline because of their role in generating and reproducing inequality in social outcomes, which occurs because an actor's status often is a lens through which critical audiences form judgments about the actor's quality. In consequence, status orderings can become self-perpetuating; because status influences perceptions of quality, those of high status often garner favorable assessments, which then reifies their positions in the very status ordering that serves as an allocation mechanism for quality appraisals. Of course, the converse is true for those on the lower rungs of the status ladder.

This idea has animated a wide array of sociological research. Perhaps most famously, Merton (1968) developed this argument in the sociology of science. He posited that small differences in initial status amplify over time to generate cumulative advantages. In Merton's classic account, not only does status itself influence perceptions of quality, but high status scientists are more likely to attract tangible resources, such as research funding and outstanding graduate students, which are parlayed into better quality scientific outputs. Of course, although it has proved to be a fertile research site, work on status extends well beyond the study of science. For example, Podolny (1993; 2005) and colleagues (e.g., Podolny, Stuart, and Hannan 1996; Benjamin and Podolny, 1999) have examined these ideas in a variety of economic markets, finding a strong correlation between an organization's status position and its performance. In a study of the diffusion of internet standards, Simcoe and Waguespack (2010) show that concealing the names of eminent electrical engineers responsible for drafting the standards greatly diminishes their spread.

Yet, despite the general consensus about status dynamics, much of the empirical evidence on the consequences of social status is fragile. This is because of the intricate coupling between an actor's quality and status, which leads to the question: Does status itself affect outcomes, or is status simply a byproduct of quality? In much of the sociological work on the subject, there is an intricate feedback loop between these two constructs; status rankings may first emerge from quality distinctions among actors or differences in their social or ascriptive attributes, but these characteristics then endogenously interact (cf. Gould 2002; Podolny 2005; Lynn, Podolny, and Tao 2009). In many settings, causality seems as likely to flow from status to quality as it is to travel in the reverse direction. Our concern is that the (non-experimental) empirical literature contains many studies with plausible, but rarely definitive, proxies for actors' social status, and often less convincing controls for other actor attributes. These measurement challenges are endemic to the empirical literature on the consequences of social status and are an inevitable byproduct of the fact that theory points to a reciprocal relationship between quality and status, and it is extremely challenging to empirically disentangle their interaction. In consequence, few of the archival studies of social status present truly persuasive evidence of its effect.

Our contribution in this paper is a research design that offers a more definitive test of the effect of changes in status on actor performance (which, in our setting, is equivalent to perceptions of quality). The image we have in mind is that there exists a set of producers, each of whom generates a set of products. Producers might be investment banks and their products might be the securities they underwrite, as in Podolny (1993); producers could be academic scientists and their products journal articles, as in Merton (1968); producers could be academic departments whose products are graduate students (Burris 2004); they could be semiconductor firms that produce innovations (Podolny and Stuart 1995); or they could be law firms that employ attorneys from different calibers of law schools (Phillips and Zuckerman 1999). In the typical archival study of social status, there is a measure of producer-level status that is either derived from a social network among producers (e.g., Podolny 1993; Stuart, Hoang, and Hybels 1999) or that is an aggregation over measures of product-level status (e.g., Phillips and Zuckerman 1999; Podolny and Stuart 1995).

Here we examine the consequence of scientists' winning a highly coveted appointment, which results in an immediate shift in their status. We analyze the effect of this prize on the subsequent-to-award citation rates to journal articles that were published *before* the award was granted. This research design has two compelling advantages. First, because the producers (scientists) in the data have created many thousands of products (journal articles), we can generate a control group of papers that contains nearly exact matches to articles written by prize-winning scientists. This enables us to net out the effect of product-level (article-level) quality when we estimate the effect of a change in producer-level status. Second, because the prizes we study arrive after scientists already are well established in their careers, we can further restrict the analysis to the effect of the status shocks on articles that were written before the awards were received. In essence, the research setting offers unambiguous measures of status changes and product quality. This allows us to isolate what we believe to be a pure status effect. The payoff of the research design is that we are able to present the cleanest test we can devise of Merton's (1968; p. 58) often-quoted description:

> *"...the Matthew effect consists in the accruing of greater increments of recognition for particular scientific contributions to scientists of considerable repute and the withholding of such recognition from scientists who have not yet made their mark."*

In our research design, we literally measure increments in recognition (changes in citations to papers) as the dependent variable, and a change in status as the central explanatory variable.

We present three findings. First, we will show that results are very misleading when, as is typical in the literature, the effect of status is measured at the *producer* level without adjustments for product-level quality. When we analyze citation rates to articles while controlling only for scientist quality, we observe a very large effect of the status shock. However, this effect falls precipitously when we isolate the result to be net of both scientist- and article-level controls for quality. This gives rise to our second, notable result: the effect of changes of status on the attention given to a producer's products is smaller than we had anticipated, and it is apparent only for products that were produced near the time that the status shock occurred. Finally, as theory suggests, we find the magnitude of the status effect is very much contoured by the level of uncertainty around producer and product quality. For instance, the consequence of a shift in status is much larger for younger producers, for those whose positions in the status ordering are less well established at the time they receive the prize, for papers published in lower quality journals, and for articles written in novel areas of science. Therefore, our findings strongly support the idea that status is a social cue that conveys the greatest information to audiences when there is significant uncertainty about producer or product quality.

2 Theoretical Background

Sociologists and social psychologists have demonstrated that status orderings are an ever-present feature of social life (Gould 2002; Bothner et al. 2009). For instance, those who study small group dynamics find that even in a collection of strangers, a status hierarchy emerges almost immediately upon the designation of the individuals as members of a common group (e.g., Bales et al. 1951; Blau and Scott 1962; Burger et al. 1980; Anderson, John, Keltner, and Kring 2001; Hogg 2001; Anderson, Spataro, and Flynn 2008). The literature is rich with descriptions of status hierarchies across all types of social settings, from the schoolyard to the marketplace, from ephemeral groups to occupational communities, from hospitals to street gangs. The differentiation of actors into positions in status orderings truly does permeate social life.

A number of theories explain the emergence of these hierarchies. In Merton (1968), Podolny (1993), Gould (2002) and other formulations of status theory, the prestige hierarchy takes shape from audience members' *perceptions* of quality, but it can quickly diverge from the distribution of *actual* quality. In theories in which perceived quality determines position in the status ordering, there is one essential—albeit not particularly limiting—boundary condition: there must be some degree of uncertainty among the members of the social system regarding their ability to assess actual quality. Given this uncertainty, actors resort to social cues to resolve their uncertainty about a producer's underlying quality. Through a variety of mechanisms, the reliance on social cues to form perceptions of quality can drive a wedge between the distributions of status (as perception) and quality (as reality).[1]

[1]Of course, as many students of stratification processes have observed, one consequence of this fact is that rewards pegged to a status ordering need not reflect a meritocratic distribution based on quality, as opportunities (or constraints) do not depend on the means by which status positions are obtained; they simply are tied to the positions themselves.

One approach actors use to resolve uncertainty about another's quality is to infer it from the affiliations and exchange relationships in which a producer is embedded (Blau 1955). Therefore, the perception of a focal producer is an attribution made in part from the prominence of its affiliates. The implicit endorsement conveyed by the fact that an actor of high regard enters an association with a focal producer inevitably shapes how others perceive that actor.

This dynamic is central in sociological accounts of scientific careers. For instance, given their short and often non-diagnostic tenure in the profession, recently minted PhDs frequently are assessed according to the prestige of their mentors (Merton 1968), the department from which they earned their doctorates (Hagstrom 1971; Keith and Babchuk 1994; Long and Fox 1995; Burris 2004), and the status of the university in which they gain employment (Crane 1965). In studies of the marketplace, quality assessments have been shown to depend on the prestige of a focal organization's strategic partners (Stuart, Hoang, and Hybels 1999) or on recognition from certifying bodies (Baum and Oliver 1991). Indeed, in universities, the prestige of departments may depend more on the status of the university than on merit-based factors, such as the productivity of the department's faculty (Keith and Babchuk 1998). In each of these instances, status is based on the identities of a producer's affiliates, rather than on direct assessments of the producer's quality.

Another reason that a producer's status and quality may decouple—and one that is central to the research design in this paper—is that status often is amplified by designations, such as prestigious prizes, that create arbitrary break points in a smooth quality distribution. This is an important element of Merton's (1968) discussion of the Matthew effect. He writes that widely coveted accolades such as the Nobel Prize engender almost capricious status distinctions. To illustrate this point, Merton (1968; p. 56) identified what he labeled, "the phenomenon of the 41^{st} chair." He wrote:

> *"The French Academy, it will be remembered, decided early that only a cohort of 40 could qualify as members and so emerge as immortals. This limitation of numbers made inevitable, of course, the exclusion through the centuries of many talented individuals."*

In other words, the fixed number of places in the French Academy (and all other highly coveted recognitions) causes a major disjuncture between status and quality, in which the status of those who are awarded membership jumps significantly, and far above any actual quality difference that separates them from others who were on the cusp of gaining the recognition. Merton's discussion of the 41^{st} chair suggests a fascinating natural experiment: How does the post-award perception of status-enhancing prize winners compare to that of the producers who were in the consideration set, but failed to make the final cut? Or, conversely, if we contemplate the counterfactual, how would the career outcomes of the 41^{st} chairs have differed if, contrary to the fact, they had been elected to the French Academy?

In attempting to answer this question, it is important to recognize that there are two very different routes through which a change in a producer's status (such as winning a status-enhancing prize) can affect outcomes. First, for reasons already discussed, changes to a producer's status may directionally influence perceptions of the quality of the producer's products. But second, a change in status often affects actual (versus just perceptions of) quality. Most importantly, this occurs because one of the benefits of status is that those who possess it attract resources, and thereby experience an enhanced ability to create high quality goods.

Once again, Merton (1968) describes both pathways in the context of scientific careers. In addition to the fact that high status scientists garner greater acknowledgment for an article of a given level of quality than would a lower-status scientist, Merton also argues that the higher status scientist is more likely to gain access to the resources that are consumed in scientific production. For instance, a high status scientist may negotiate a lighter teaching load, which opens the time to produce more and better science. Likewise, prestigious scientists are more likely to find positions at elite universities, which attract the best students and possess state-of-the-art facilities. Through these and other resources that disproportionately flow to high status producers, those who occupy positions at the top of the prestige hierarchy often have the means to produce higher quality goods.

In the empirical work that follows, we present an analysis that very closely conforms to Merton's 41^{st} chair thought experiment. Formulating the inquiry in broader terms, we seek the answers to two primary questions:

1. Does a shock to a producer's status cause others to change their perceptions of the quality of a producer's products?
2. Does the extent to which audience members update their perceptions of quality as a function of a status shock depend on the uncertainty surrounding producer and/or product quality?

In seeking an answer to these questions, we are able to narrow our focus to the question of how a shock to status influences *perceptions* of quality. That is, in the two pathways through which a change in status can affect outcomes—by altering other-party perceptions of a given level of quality or by attracting the resources that can be invested to produce higher quality goods—we focus the empirical test on the former mechanism.

3 Empirical Approach

Our research design is quite different from most of the archival research on the effects of status. Therefore, before providing the full details of the methodology, we present a high-level roadmap of the empirical approach.

First, we have identified a set of producers/scientists who are recipients of a status-conferring prize that results in a meaningful boost to their status. Second, we create a control group of scientists who were close contenders for the prize, but were not selected. Third, we collect data on all products/journal articles written by prize winners and control group members. For prize winners, we limit the data analysis to articles that were written before the prize was awarded, which guarantees that the quality of the articles in the dataset could not have been influenced by resources that come with or that follow the prize. This is how we eliminate the resource-based pathway on the measured effect of status.

Fourth, we develop an algorithm to create a sample of matched pairs at the *product* level, in which we very closely match prize winners' articles to those of control group members on a set of criteria that, we will assert, virtually assures that the actual quality of the two papers in a pair are nearly identical. This results in a matched sample comprising pairs of very similar articles, with one belonging to a prize winner and the second authored by a prize contender who did not receive an award. The final step of the analysis is to assess whether the prize—a shock to a producer's status—affects third party perceptions of quality, relative to the control group of equal quality products. At this point, the analysis becomes quite simple and it is possible to present the findings in a set of graphs: we simply compare the post-award citation trajectories of papers that were written by prize winners to the citation paths of almost exactly matching papers that were written by a member of the control group. If there is a difference in post-prize citation trajectories, it is highly likely to have been caused by the status change attributable to the award itself.

3.1 Approaches to Identifying Status Effects

Returning to the existing literature, the most widespread approach to identifying the magnitude of status effects is to focus on a set of organizations, people, or products that vary in status. When the researcher can observe this variation, the analyst can then regress various measures of performance on a measure of status using standard statistical techniques. With detailed controls for other actor characteristics that might correlate with both status and organizational performance, it is possible in principle to quantify the benefits of status. However, because variation in status often occurs only in the cross-sectional dimension of the data, these estimates are susceptible to the critique that they may largely reflect the confounding influence of omitted variables, such as producers' true quality or resource endowments.[2]

[2]A number of studies estimate the effect of status on a performance outcome while including actor fixed effects (e.g., Podolny, Stuart and Hannan 1996). Under a strict set of assumptions, these studies address the problem of unobserved quality differences among producers. For the fixed effects estimator to yield accurate coefficient estimates, producer status positions must change meaningful during the time period of the study so that there is adequate within-producer variation to estimate the effect of status. In addition, quality differences must be non-time-varying, or the fixed effects will not resolve the problem of unobserved

We adopt a different approach: we focus on a setting in which we can unambiguously isolate the timing a significant status shock, and examine the benefits that accrue to the same producer and his/her products, before and after the shock. This longitudinal contrast purges our estimates of most sources of omitted variable bias that plague cross-sectional comparisons, but it remains possible that the timing of the shock is itself endogenous. Specifically, when status shocks incorporate forecasts about actors' expected performance improvements, then fixed effects estimates are unreliable (because the estimated change simply may reflect the selection of the producer into the treatment condition, rather than a causal effect of the treatment/status change per se). To remedy this problem, we pair each producer/product with a carefully chosen control which, in addition to providing a very close match based on time-invariant characteristics, also shares very similar performance trends prior to the shock. By analyzing the data at the matched-pair level of analysis, this simple difference-in-difference framework provides a flexible and non-parametric methodology to evaluate the effects of the shock. In fact, conditional on the assumption that the matching algorithm we employ successfully pairs products of comparable quality, we are able to present the findings in a straightforward, graphical form.

3.2 Status Shocks and their Associated Counterfactuals

The academic sciences provide an ideal "laboratory" for our study. First, the production of scientific knowledge has been the *locus classicus* of pioneering investigations of status and its effects, beginning with Merton's original statement of the Matthew effect, and continuing with many of his intellectual disciples (e.g., Cole and Cole 1968; Cole 1970; Allison et al. 1982). Second, this setting provides a clear distinction between individual producers (scientists) and their products (scientific articles). Third, scientists may receive shocks to their status at several distinct career stages, in the form or prizes (such as the Nobel Prize, the Fields Medal, or the Lasker Award) or election to prestigious societies (such as the National Academy of Science in the United States, the Royal Society in the United Kingdom, or the

differences between producers. In general, status theory suggests that status positions are far stickier than producer quality (Podolny 2005), which calls into question the ability of fixed effects estimators to solve the measurement problems in the status literature.

Académie Française in France). Finally, the flow of citations to scientific articles provides a metric to evaluate the effects of status, since by citing another piece of knowledge, producers inscribe into their own products esteem for their peers.

HHMI Appointment. We analyze a salient shock in the status of elite, mid-career academic life scientists in the United States—appointment to be investigators of the Howard Hughes Medical Institute (HHMI). HHMI, a non-profit medical research organization, is a major participant in biomedical research in the United States. The Institute's annual budget is larger than the amount the NSF typically commits to the biological sciences. During periodic, open competitions, the Institute solicits applications from researchers at universities, medical schools, and other research institutions across the United States. The selection committee for HHMI almost exclusively comprises members of the National Academy of Sciences, so the profession's most elite scientists choose the prize winners. Once selected, awardees continue to be based at their home institutions, but they are entitled to append "& HHMI" to their affiliation in the papers they publish, so that other scientists are constantly reminded of their enhanced status.

The subfields of the life sciences of interest to HHMI investigators are quite broad, but have tended to concentrate on cell and molecular biology, neurobiology, immunology, and biochemistry. Career-stage considerations have varied over time, although HHMI typically has not appointed scientists until they have had enough independent experience so that their work can be distinguished from that of their postdoctoral or graduate school adviser.

Appointment to HHMI is a major, status-conferring prize. Indeed, it is the most prominent accolade that a U.S. life scientist can receive relatively early in his/her career. Consistent with its stature, HHMI appointment is a harbinger of even greater accomplishments: the current group of HHMI investigators includes 14 Nobel laureates and 131 members of the National Academy of Sciences.

HHMI's award cycles last five years and typically are renewed at least once. Appointees also are recipients of large research budgets and may be the beneficiaries of intangible forms of assistance such as editorial goodwill and privileged access to sources of advice from an elite peer group. As such, HHMI appointment combines status with other forms of resources, and it is therefore likely to affect both the perceived and actual quality of a prize winner's work. To separate these two effects, we limit our analysis to the consequence of the prize to the citation trajectories to articles that were written before the award was granted. We do this because future appointment to HHMI cannot influence the actual quality of pre-existing work.

The Producer Control Group: Early Career Prize Winners. HHMI investigators are an elite group drawn from the large population of academic life scientists. The details of the selection process are important insofar as they elucidate the characteristics that a control group should exhibit to create the most similar counterfactuals for the articles published by HHMI investigators.

In addition to a subjective assessment of the originality and promise of nominees' future research plans, the selection committee looks for evidence that nominees have stepped out of the shadow cast by their mentors, as indicated by what we call a "big hit," *i.e.*, a high-impact publication in which the mentor's name does not appear on the coauthorship list. This is important because it suggests that the citation trajectories of papers that future HHMI awardees had published in the pre-appointment period might already incorporate peers' expectations of unusually positive impact. It is therefore essential to contrast the citation dynamics for HHMI's articles with those of a control group of articles exhibiting similar pre-appointment trends. This requires a method for identifying a control group of comparable scientists, and a method for culling from their research output a set of articles with citation profiles that are similar to that of HHMI investigators in the period preceding appointment.

Given the high degree of accomplishment exhibited by HHMI investigators at the time of their appointment, a random sample of scientists of the same age, working in the same fields, would not be appropriate. We therefore construct a control group comprising only scientists who received prestigious early career prizes that are awarded in the same subfields of the life sciences as HHMI. The Pew, Searle, Beckman, Packard, and Rita Allen Scholarships all are early career prizes that target scientists in the same life science subfields and similar research institutions as HHMI. We label members of this control group "Early Career Prize Winners," or ECPWs. Every year, these charitable trusts provide seed funding to around 60 life scientists in the first two years of their independent careers.

3.3 From Control Producers to Control Products: A Non-Parametric Matching Procedure

A key departure of our empirical approach is to unbundle producers' status from their products. Empirically, we begin with all products (articles) of "treated" producers (HHMI investigators) and then we search for nearly exactly matching products (articles) written by control group producers (EPCW scientists). The goal of the construction of this matched sample is to select a set of articles that pin down the citation trajectories associated with HHMI investigators' papers had they, contrary to the fact, not been awarded this prize.

In practice, identifying "close matches" is difficult. Because we are interested in the fate of individual articles, but the status shock we observe operates at the scientist-level of analysis, semi-parametric matching techniques such as the propensity score and its variants are of limited use in our context.[3] We propose instead a non-parametric matching approach, a so-called "Coarse Exact Matching" (CEM) procedure (Blackwell et al. 2009).

The selection of controls proceeds in a series of sequential steps. The first task is to select a relatively small set of covariates on which we would like to guarantee balance between the treatment and control group. The second step is to create a large number of strata to cover the entire support of the joint distribution of the covariates selected in the previous step. Next, each observation is allocated to a unique stratum; any stratum that either has no articles written by an HHMI, or that has less than five potential control articles, is then dropped from the data. In a fourth and final step, we select in each stratum a unique control

[3] A propensity score approach would entail estimating the probability that the scientists in the data win an HHMI, and then using the inverse of this estimated probability to weight the data in a second stage analysis of the effect of the HHMI on subsequent citation rates. However, because citations occur at the product/article level, achieving covariate balance by weighting the data by the producer-level likelihood of winning the prize, even if the determinants of winning were observable, would not resolve the problem of controlling for product-level quality.

article such that the sum of squared differences in citation flows between the treated and control article from the year of publication up to the year preceding the appointment is minimized. We break ties at random when there are several candidate articles that minimize this distance metric.

The procedure is coarse because we do not attempt to precisely match on covariate values; rather, we coarsen the support of the joint distribution of the covariates into a finite number of strata, and we match a treated observation if and only if a control observation can be found in the same stratum. An important advantage of CEM is that the analyst can guarantee the degree of covariate balance *ex ante*. However, the more fine-grained the partition of the support for the joint distribution (i.e., the higher the number of strata), the larger the number of unmatched, treated observations. In general, the analyst must trade off the quality of the matches with external validity: the longer the list of matching covariates, the more difficult it is to identify an "identical twin" for each article in the treatment group.

We distill the essence of the matching procedure in Figure I. Control articles are selected if they share the following characteristics with treated articles: (1) year of publication; (2) specific journal (e.g., *Cell* or the *New England Journal of Medicine*); (3) number of authors; (4) focal-scientist position on the authorship list; and (5) the cumulative number of citations the articles received between the time they were published and the year the treated scientist was appointed to HHMI. Implementation details can be found in Appendix I.

We start from a universe of 145,855 articles published by HHMI or ECPW scientists. Out of these 145,855 papers, 12,014 (8.24%) are pre-appointment publications written by HHMI investigators. It is this group of 12,014 treated articles for which we seek matches from the set of papers written by control group members. We successfully match 5,385 out of these 12,014 publications (44.82%). This relatively low match rate is to be expected because non-parametric matching procedures such as CEM are prone to a version of the "curse of dimensionality" whereby the proportion of matched units decreases rapidly with the number of strata. For instance, requiring a match on an additional indicator variable (e.g., matching on focal scientist gender in addition to the covariates mentioned above) would result in a match rate of about 30%. Conversely, if we did not require that control and treated articles are drawn from the same scientific journal, the match rate would jump to 70%. Relaxing this constraint, however, would come at the expense of internal validity.

3.4 Data Sources

We describe the data sources used to assemble our multilevel panel dataset, drilling down to increasingly finer-grained levels of analysis: from the individual scientist to the journal article; from the journal article to the articles that cite it; and even from a citing article to the characteristics of citing scientists and institutions.

Individual scientist data. We start from the set of all 446 HHMI investigators appointed between 1984 and 2003. We track these scientists' careers from the time they obtained their first position as independent investigators (typically after a postdoctoral fellowship) until 2006. We do so through a combination of curriculum vitæ, NIH biosketches, *Who's Who* profiles, accolades/obituaries in medical journals, National Academy of Sciences biographical memoirs, and Google searches. For each one of these individuals, we record employment history, degree held, date of degree, gender, and up to three departmental affiliations.

The 427 investigators who are the focus of this paper constitute a subset of this larger pool. We impose two screens to derive the final list. First, we eliminate from the sample 16 scientists who transition from academic positions to jobs in industry; second, we delete three investigators who work in a small field (computational biology), since our ability to find control scientists in the same field is limited.

To construct the control sample we proceed in parallel fashion to track the careers of ECPW scientists. The final sample contains 2,784 early career prize winners. In addition to the criteria listed above, we have deleted from the sample investigators earning their highest degree before 1956 or after 1998 (the lower and higher bounds for the HHMI sample), as well as those investigators whose main department affiliation has no counterpart among HHMIs (such as veterinary sciences, public health and preventive medicine, otolaryngology, or anesthesiology).

The timing of appointment for HHMI investigators is identified from the HHMI web site and scientists' CVs, rather than inferred from affiliation information in published articles. To be precise, we know the calendar year in which each investigator was first on HHMI's payroll, but not the exact date. We adopt the following convention: the baseline year for each treated scientist is the year that precedes the appointment year listed on HHMI's web site. It is possible that some publications that appear in the year of appointment in fact correspond to pre-appointment output. However, because appointment carries access to resources that may enhance the actual quality of work, we are more concerned about mislabeling post-appointment articles as having been published prior to appointment. In this respect, we err on the side of caution: our assignment of the treatment date insures that all articles in the matched pair sample were written before the treated scientist was appointed to HHMI.

Article data. The second step in the construction of our dataset is to link scientists to journal articles. We collect data on articles from PubMed, a comprehensive bibliographic database covering the fields of medicine and the life sciences. In practice, the challenge in collecting these data is the thorny issues of name uniqueness: common names make it difficult to distinguish between scientists, and scientists with relatively rare names sometimes are inconsistent in their use of publication names. We resolve this problem by designing a set of customized search queries for all 427 HHMIs in the treated group and all 2,784 EPCWs in the control groups, which enhances the accuracy of each scientist's bibliome. Further details on the data and linking process are provided in Appendix II.

We begin by downloading all publications of the HHMI and EPCW scientists using the customized queries. We then eliminate from the consideration set letters, comments, and editorials. Next, we eliminate all articles published 11 or more years prior to the date of the earliest appointment to HHMI in the sample (1984); similarly, we eliminate all articles published after 2003 (the latest HHMI competition we record) so that we always observe a minimum of three years of citation information for each article.

Citation data. PubMed does not contain citation data but we were able to retrieve this information from the *Web of Science* using a perl script. We further process these data to make them amenable to statistical analysis. First, we eliminate all self-citations, where self-citation is inferred by overlap between any of the cited authors with any of the citing authors (an author name is the tuple formed by the last name and the first initial for the purpose of this filter). Second, we match the citing article data with another database we have assembled containing all publications for members of the National Academy of Science and HHMI investigators. Whenever a citing article has on its authorship roster at least one scientist who was a member of the NAS or a previously-appointed HHMI investigator at the time of publication, we flag this citation as being "high-status." We then aggregate this information at the article-year level of analysis. In other words, we can decompose the total number of citations flowing to individual articles at a given point in time into an "ordinary" and a "high status" set.

3.5 Descriptive Statistics

The final sample contains 5,385 treated articles and 5,385 control articles, for a total of $2 \times 98,269 = 196,538$ article-year observations. (The average article is written a number of years before HHMI appointment and is observed for multiple years after appointment.) We report descriptive statistics in Table 1. In the table, all time-varying covariates are measured at baseline, which we define to be the year preceding appointment (for HHMI investigators) or counterfactual appointment (for ECPWs), making it easier to assess balance between treated and control articles.[4] Four facts merit attention. First, article-level, time-invariant characteristics are very closely matched between treated and control groups. For some covariates (e.g., number of authors, focal author position, article age), this is a mechanical reflection of the CEM procedure, but for others (such as the article's novelty, as assessed by the average vintage of the keywords that tag the article), the close match occurs by chance. Second, the distribution of citations received at baseline is also very similar between the two groups, as can be seen in Figure 2, Panel A. Third, as we would expect when we create a paper-level control group, balance does not extend to scientist characteristics, such

[4]A control article inherits the appointment year associated with its treated article match, resulting in a counterfactual appointment year for the ECPW scientist who is the focal author of this control article.

as gender and graduation year, though the two groups appear well-balanced on the number of "hit articles" they have previously published at baseline.[5] Fourth, one can at best discern a very small difference in the number of citations received cumulatively up until 10 years *after* appointment—the dependent variable of interest.

In short, the comparisons between control and treated observations bring to light that our matching procedure is product-level, rather than producer-level. Imposing a match on a full suite of producer characteristics in addition to article-level covariates would result in a very low match rate. Conversely, one could modify the procedure to achieve a closer match on focal scientist characteristics, but the articles matched in this way would differ in the pre-appointment flow of citations. By restricting the set of potential control producers to early career prize winners, and then insisting on a very close match at the article level, we have sought a balance between internal and external validity.

3.6 Statistical Approach

A natural starting point for an analysis of the effect of HHMI appointment on citation trajectories would be to run regressions using all article-year observations (treated and control) as the estimation sample, with article fixed effects. Because we have several cohorts of HHMI investigators in the sample (appointment years are staggered from 1984 to 2003), the control group that pins down the counterfactual vintage and calendar time effects for the articles that were written by currently appointed HHMI investigators contains three categories of articles: (i) articles written by early career prize winners; (ii) articles by scientists who will become HHMI investigators in the future; and (iii) articles written by HHMI investigators appointed in earlier periods. The articles that are part of the last two categories are problematic "controls," since they were treated in the past or will be treated in the future. If HHMI appointment events influence citation trends (rather than just levels), the fixed effects estimates will reflect in part this unwanted source of variation, occluding the interpretation of the results.

[5] We classify a paper as a hit if its cumulative number of citations up until 2008 places it above the 95^{th} percentile of the article-level citation distribution.

To produce an analysis in which the control group solely consists of articles written by ECPW scientists, we perform the statistical analysis at the article-pair level. Specifically, the outcome variable is the difference between the citations received in a given year by a treated article and its associated control identified in the matching procedure described above. Let i denote an article associated with an HHMI scientist and let $i^{'}$ index the corresponding control article. Then our estimating equation relates $\Delta CITES_{ii't} = CITES_{it} - CITES_{i't}$ with the timing of HHMI appointment in the following way:

$$E\left[\Delta CITES_{ii't}|X_{ijt}\right] \;=\; \beta_0 + \beta_1 AFTER\_HHMI_{jt} + f(AGE_{jt}) + \gamma_{ii'} \qquad (1)$$

where $AFTER\_HHMI$ denotes an indicator variable that switches to one in the year scientist j becomes an HHMI, $f(AGE)$ is a flexible function of the scientist's age, and the $\gamma_{ii'}$ are article-pair fixed effects, consistent with our approach to analyze *changes* in the citation rates to articles in each pair following the appointment of investigator j to HHMI.[6] We also run slight variations of this specification in which the dependent variable has been parsed so that we can break down citation flows by citer status (i.e., citations associated with members of the NAS vs. non-members).

There is another benefit to conducting the analysis at the product-pair level: since treated and control products always originate in the same year, experimental time and calendar time coincide, making it simple to display the results of the analysis graphically. The graphical approach is advantageous because it enables to us go beyond a temporal averaging of status effects (i.e., a single point estimate of the treatment effect that averages its impact over time) to illustrate their dynamics. The regression analysis, however, will prove useful when exploring interactions between the treatment effect and various scientist or article characteristics.

[6]It is conventional to include life cycle and period effects in studies of scientific productivity (Levin and Stephan 1991), but our empirical approach obviates the need for these because, by construction, both articles in each matched pair were written in the same year.

4 Results

4.1 Main Effect of HHMI Appointment on Citation Rates to Articles Published Post-Appointment

Before we present our analysis of the treatment effect of appointment to HHMI using the article-pair matched sample, we report the effect of becoming an HHMI that is comparable to what we would find if we followed the standard methodology to estimate its effect. This analysis differs from what will follow in two essential respects. First, in replicating the conventional approach, the match we impose is most stringent at the producer level, rather than at the product level. In this analysis, when estimating the effect of status, we control for producer-level but not product-level quality. Second, the products/articles in these data were written *after* the treated scientist won the prize, rather than before. In other words, the first set of results are akin to an estimate of the effect of the shock to producer status that, (i) does not include careful controls for product quality, and (ii) allows the effect to depend on the producer's subsequent-to-award output.

Here, we pair an HHMI winner with an ECPW scientist who is very similar on covariates that we know matter for selection into the HHMI program. These covariates are (i) year of highest degree (coarsened in three-year intervals); (ii) gender; and (iii) number of independent 'hits' scored up to the appointment year. By limiting the control group to early career prize winners and then further matching on the number of hits, we effectively include excellent controls for scientist-level quality. In addition, we match on a few basic article characteristics, including the length of the authorship roster, the focal author's position, and publication year.

The results of this analysis, which reveals the effect of appointment to HHMI on scientists' future performance while controlling for scientist characteristics, are presented in Panel A of Figure 3. In this and subsequent figures, we display the difference in average citation trends for the article pairs in the sample (the solid red line), along with a 95^{th} confidence interval (the dashed blue lines). Panel A shows that articles written by HHMIs are cited more frequently than articles written by early career prize winners. This premium exists in the first year of an article's life, increases in the article's second year, and declines over the next 10 years without ever vanishing. By 2006 (the end of our observation period), articles written by HHMIs have garnered 40 more citations than those of early career prize winners. Considering the controls implicit in the matching process, the treatment effect of appointment to HHMI appears to be very large.

As we have argued, however, interpreting appointment to HHMI as causing a change in the perceived quality of prize winners' work is problematic because of two alternative possibilities. The first is that the premium reflects the presence of correlated resource endowments. Relative to ECPW scientists, who must worry about funding the work of their laboratories, HHMI appointees are recipients of large research budgets. Furthermore, the actual quality of their post-appointment publications might increase relative to a control group of non-prize-winners because HHMI's may benefit from access to higher quality graduate students, better-equipped laboratories, advice from an elite peer group, and so forth. In other words, resource flows tied to the award may result in actual improvements to the quality of prize winners' articles, rather than simply changes in others' perception of their quality.

In fact, we present some evidence to suggest that this is indeed the case. Specifically, in Panel B we repeat the analysis presented in Panel A, but this time we add one additional criterion to the matching algorithm: instead of matching on just a *scientist*-level measure of quality (the number of hits), we also require that both of the papers in a pair were published in the same journal. In other words, we incorporate the equivalent of a product-level quality control to the matching algorithm, so that we better account for potential quality differences between the articles written by treated and control group members. When we do this, although the average premium for HHMI-authored papers is still evident, its magnitude is greatly reduced (to about one citation per year). Immediately, we can see the large potential bias in estimating the effect of status while controlling only for producer-level quality, without accounting for quality differences at the product level. The inclusion of a product-level control erases about three-fourths of the estimated status effect.

There is, however, a second complication that raises further questions about the interpretation of the residual treatment effect: it remains possible that the citation increase could be an artifact of a well-functioning selection process. Even if ECPW and HHMI scientists were perfectly matched on *past* achievement, the HHMI appointment event may also incorporate information about scientists' future *potential*. If this were the case, one would expect to observe a citation premium for articles by HHMI investigators, even in the absence of any status effect. By focusing on changes in citation rates following HHMI appointment for articles published before appointment while matching on the pre-appointment citation trajectory, we believe that our research design enables us to isolate the operation of status-based changes in perception from these competing effects.

4.2 Main Effect of HHMI Appointment on Citation Rates to Articles Published Pre-Appointment

The comparisons in Figure 3 contrast articles written after a scientist is appointed to HHMI with a paired article by an ECPW. By contrast, Figure 4 and 5—the core of our analysis—confine the analysis to articles written before the HHMI is appointed, which are each paired with a matching ECPW article. This and subsequent figures report the difference in citation trajectories for the 10-year period following the treated scientists' appointment to HHMI. In all graphs, the zero point is the year the HHMI is appointed; negative years indicate the pre-appointment period and positive years correspond to the post-appointment period. Recall that the fact we limit the analysis to articles that were written before appointment enables us to incorporate an additional, stringent control for article-level quality in the matching procedure: we now restrict the dataset to pairs of HHMI and ECPW articles that were: (i) published in the same year; (ii) in the same journal; (iii) with approximately the same number of authors; (iv) in which the HHMI and ECPW scientists occupy the same authorship position; and (v) that were following a nearly identical citation trajectory up to and including the year that precedes HHMI appointment. We then investigate whether there is a citation boost associated with HHMI appointment.

In Figure 4, the average citation premium hugs the horizontal axis of the graph until the year the HHMI is appointed. The lack of any detectable difference in the years between when a pair of articles was published and when one scientist in the pair is appointed to HHMI confirms that the matching algorithm indeed selects control articles with citation trends that were nearly identical to treated articles. The magnitude of the status effect for appointment to HHMI in the overall sample is captured as the difference in the curves in the post-appointment (after time zero) period. Inspection of the figure reveals that the effect is quite small: one can observe a very slight uptick in the citation rate in the first post-appointment year, with a gradual decrease in subsequent years. Based on this evidence, one might conclude that the overall effect of the status shock on the perceived quality of scientists' existing work is *de minimis*.

While this may be the overall conclusion, it is altered by cutting the data into different subsets to examine contingencies in the effect of the HHMI status shock. First, we find that the results very much depend on the vintage of the papers being examined. Figure 5, Panel A performs an identical analysis, but limits the sample to articles published at least three years before appointment. For this sample, there is absolutely no evidence of an HHMI citation boost. Figure 5, Panel B limits the sample to articles published exactly two years before the year of HHMI appointment. Once again, there is no hint in this subsample of an HHMI citation premium. Finally, Panel C focuses on articles published in the year immediately preceding the year of appointment.[7] In this sample, there is a post-appointment citation increase. HHMI articles receive 2.5 citations more on average than their ECPW counterparts in the year following appointment. This premium decreases steadily until the fifth year after appointment, after which it is statistically indistinguishable from zero. On average, HHMI articles appear to receive ten extra citations over the ten-year period that follows appointment. To contextualize this effect, the median number of cumulative citations in the control sample is equal to 36, and $36+10 = 46$ citations map into the 57^{th} percentile of the distribution. This seven-percentile point rightward shift strikes us as being meaningful.

[7] Note that the articles in each pair are matched on publication month. This is necessary because two articles with an equal number of citations in, say 1990, the first appearing in February, and the second appearing in October, might, in fact be on very different citation trends.

Based on these first results, one might conclude that major status shocks like HHMI do have the effect of directing the attention of other scientists, but its effect is limited in time to recent product introductions.

4.3 Variation in the HHMI Citation Premium

The foregoing results pertain to the *average* citation premium accruing to appointed individuals in the data. In Table 2, we now examine whether the magnitude of the premium correlates with specific attributes of articles or scientists. The regressions in this and the subsequent table include article-pair fixed effects, corresponding to the estimating equation above. Robust standard errors, clustered at the level of the focal HHMI scientist, appear below the coefficient estimates in brackets. In a first step, we explore whether status and tangible markers of quality are substitutes. That is, we examine whether the effect of status is amplified when there is greater uncertainty surrounding quality, either at the producer- or product-level.[8]

We begin at the product level, and implement two splits of the data to examine the impact of uncertainty in product quality on the magnitude of the HHMI treatment effect. First, we split the sample around the median of the distribution of Journal Impact Factor (JIF)—a measure of the frequency with which the "average article" in a journal has been cited in a particular year (Models 1a and 1b). We then estimate the treatment effect in each subsample. Our logic is that journal quality is a strong signal of article quality, so the effect of author status on perceptions of article quality should be greater in lower quality journals. We find evidence of an HHMI premium in both subsamples, but comparing the coefficients, the effect is indeed more pronounced for articles appearing in less prestigious journals.

Second, we explore whether the effect of status is larger when market participants face difficulty in evaluating the inherent worth of particular products because these products embody relatively novel concepts or categories (Podolny, Stuart, and Hannan 1996; Zucker-

[8]Although it is possible to present the results of all of these interaction effects in graphical form, to economize on space, we report the results in regression tables. Graphs are available from the authors upon request.

man 1999). To measure novelty at the article level, we first calculate the vintage of all the MeSH keywords tagging the articles in our sample.[9] Concretely, we define the birth year of a keyword as the year in which it first tags a paper indexed by PubMed. For each article, we then difference the average MeSH vintage from the publication year to produce our index of novelty. We assume that more novel articles are ones that are tagged by MeSH keywords that were first used in the recent past.

We again split the sample according to the median of this novelty measure to explore whether appointment to HHMI has a greater effect on subsequent citations to pre-existing articles when those articles were published in more novel areas of science, and this indeed is the case (Table 2, Models 2a and 2b): the post-HHMI citation premium is larger when the focal article in the pair is relatively more novel.

Journal Impact Factors and the MeSH keyword novelty are product-level measures of uncertainty. In addition, our data also contain significant variation in the level of uncertainty around producer quality. This is because there are major differences in the scientific track records of HHMI awardees at the time of their appointment. Some are appointed only a few years after starting their independent careers, while others are much more senior in their fields. In some cases, HHMI appointment consecrates a landmark scientific contribution;[10] in other instances, HHMI investigators are recruited more for their promise than past achievement.

Once again, the literature offers a clear expectation: the effect of a jump in status will be larger for producers when market participants face greater uncertainty regarding producer quality. In Models 3a and 3b, we carry out a median sample split by the age of the focal HHMI scientist. As theory suggests, we find that the citation premium is larger in magnitude

[9] MeSH is the National Library of Medicine's controlled vocabulary thesaurus. It consists of sets of terms naming descriptors in a hierarchical structure that permits searching at various levels of specificity. There are 24,767 descriptors in the 2008 MeSH (new terms are added to the dictionary as scientific advances are made). To fix ideas, the following MeSH terms appear in the publications of HHMI scientists in our sample: `Sphingolipidoses`, `RNA, Small Nucleolar`, and `Tachycardia, Ventricular`.

[10] For example, Craig Mello from the University of Massachusetts was appointed HHMI investigator in 2000, two years after the discovery, together with Andrew Fire from Stanford, of a remarkable biological effect called RNA interference (Fire et al., 1998). Six years later the pair was awarded the Nobel Prize in Medicine for this achievement.

for younger investigators. Similarly, in Models 4a and 4b, the sample is split according to the scientific eminence of the focal HHMI scientist at the time the award is granted, as measured by the cumulative stock of forward citations to all articles published up to the year before appointment. Again, we find that the articles published by (relatively) less eminent scientists benefit more from the status shock.

5 Robustness Checks

We present three robustness checks to further assess the integrity of the results.

Salience of HHMI status marker. First, we conduct a falsification test of sorts, by examining whether the citation premium accruing to HHMI awardees varies with authorship credit for the focal scientists. If HHMI appointment represents a genuine status shock and if future citations to past articles is a good measure of perceptions of quality, the strength of the results should depend on which position the HHMI holds on the authorship roster of an article. In particular, we exploit a strong social norm in the academic life sciences, which assigns last authorship to the principal investigator, first authorship to the junior author who was responsible for the actual conduct of the investigation, and apportions the remaining credit to authors in the middle of the authorship list, generally as a decreasing function of the distance from the list's extremities.

In our methodology, recall that articles in each pair are matched on authorship position, *i.e.*, first, middle, or last author. In Table 3, Models 2a and 2b, we split the cited-article sample in two by consolidating the first and last authorship categories, and contrasting it with those article-pairs in which the focal scientists appear in the middle of the authorship list. We find clear evidence of a more pronounced status effect for articles pairs in which the HHMI scientist is either first or last author. The evidence for middle-position authors is much smaller in magnitude, and also more fleeting when we display the results graphically (figures available from the authors). This is reassuring because the level of contribution of middle authors is often sufficiently small that a change in their status should not cause a significant reassessment of the quality of the article. We view this result as an important confirmation of the HHMI effect.

Appointment panel effects. We observed a post-citation boost only for articles published in the year immediately preceding appointment. If these recent articles are precisely those that convinced the selection panel to choose these particular nominees, then we run the risk that the results could be influenced by the citation patterns of the panelists themselves, as they are active publishers who may become aware of applicants' work in the selection process. If this were the case, it may be stretching the meaning of the Matthew Effect to interpret our results through its lens.

Although we cannot identity the panelists by name, we do know they are recruited from the ranks of the National Academy of Science and long-established HHMI investigators. In Table 3, Models 3a and 3b, we split the pool of *citing* articles into two separate buckets. The first includes articles in which the authorship roster does not list any member of the NAS or a past HHMI investigator. Conversely, the second is limited to articles in which at least one author is a member of the academic "super-elite." The results show that existing HHMI investigators and NAS members do not cite papers of soon-to-be-appointed HHMI investigators more than would be expected given their relatively tiny share of the overall pool of citing articles.

"Average" scientist control group. One potential concern about the analysis we have presented thus far is that the results may be heavily influenced by our choice of the control group. In particular, one might wonder whether selecting the control articles from among those published by early career prize winners makes it unlikely to detect an effect of HHMI appointment, because EPCW members themselves possess high status. (We return to this point in the concluding section.)

We maintain that a tight control for article-level quality is necessary to put Merton's celebrated hypothesis to a rigorous test. To probe the external validity of our results, however, we perform an identical battery of analyses except that we replace the EPCW control group with a second control group of articles, but this time all written by 6,272 "average" medical school faculty members (details on the construction of this second control group can be found in Appendix III).

Although we begin with a large sample of scientists, the net result is a much smaller sample of matched article pairs because the papers written by "average" scientists rarely follow the same citation trajectories or appear in the same journals as those written by HHMI awardees. To wit, the average number of hits (i.e., papers in the top 5% of the article-level citation distribution) at baseline is 28 on average for HHMI investigators, but only 3 for average scientists. Nonetheless, we successfully matched 736 article pairs.

Descriptive statistics for these articles are provided in Table 4. Once again, for these 736 pairs, we successfully balanced citation outcomes prior to appointment (Figure 2, Panel B).

The analysis we conduct with the control group of average scientists mirrors that performed in section 4.1 with the ECPW control group. First, we focus on article pairs in which the treated publication was published after appointment. Figure 6 displays the results graphically. Relative to average scientists, the HHMI citation premium is extremely large (peaking at close to 10 citations a year, declining thereafter) when we do not constrain the treated and control article to appear in the same scientific journal (Figure 6, Panel A). Comparing Panel A, Figure 6, to Panel A, Figure 3, we can see that matching on producer/scientist quality greatly diminishes the effect of the prize on the estimate of the citation premium to future work. When we impose a further constraint to insist that the papers in each pair are published in the same journal, we observe a more modest effect of the award (Panel B). Yet, it is significantly larger than the one displayed in Figure 3, Panel B. These results underscore that controlling for scientist quality does absorb a significant amount of the variance that would otherwise could be attributed to the effect of status.

Second, we examine citation dynamics for articles published prior to appointment. Figure 7 commingles all article vintages. It incorporates publications appearing from one year to ten years prior to the focal scientist's HHMI appointment. We can discern the existence of a post-appointment HHMI premium that is larger than the one depicted in Figure 4. When breaking down the analysis by article vintage (Figure 8), we find evidence of the same general pattern observed in Figure 5: the post-appointment treatment effect is limited in time, and affects only articles that were written recently.

To conclude, the patterns uncovered in Section 4 appear to be robust; these supplemental analyses do not alter the substantive interpretation of the core results. The precise magnitude of the effect of the status-enhancing prize is, however, sensitive to the choice of control group.

6 Conclusion

This paper presents a novel research design that enables a narrow test of the Mertonian hypothesis that a producer's status is a lens through which audience members assess the quality of its work. Specifically, the research design first zeroes in on the effect of a change in status caused by winning a major prize on other-party *perceptions* of the quality of a focal producer's goods. To identify the effect of the change in status that is independent of its potential influence on the actual quality of a producer's outputs, we limit the analysis to the effect of the change to evaluations of outputs that were produced prior to the time the prize was granted. To further insure that the results reflect truly causal changes in perceptions (versus forecasts that endogenously relate to the selection of specific producers as prize winners), we implement a product-based, matched sample design that pairs each treated product to a nearly identical twin that is almost perfectly matched on product quality.

We present a few, central findings. Our results suggest that, for two reasons, the standard approach for estimating the effect of status on performance outcomes is likely to overstate its true, causal influence. One reason is that controls for quality often are inadequate, particularly if quality is held constant at the producer level but performance is measured at the product level. The second is that changes in a focal actor's status follow acts of deference from high status actors, whether through the awarding of prizes or other forms of recognition, or through creation of some form of affiliation that conveys an implicit endorsement. While these actions on the part of already high status actors do produce changes in status, their intentionality is often rooted in forecasted changes in performance. If high status actors (such as the members of the HHMI selection committee) bestow recognitions or affiliations because they anticipate that the recipients of these acts of deference are on positive performance trajectories, status changes may reflect—rather than cause—changes in performance. For both reasons, much of the existing empirical literature on status may overestimate its true effect. The standard approach for estimating the effect of status certainly results in an untenably large estimate in our data.

Despite this core result, we still find that appointment to HHMI results in increases in citation rates to articles written before the award was granted. In a strict test of the Mertonian hypothesis that prestigious scientists garner greater recognition for outputs of a given level of quality, we find modest support for a main effect of a change in status. However, consistent with the literature, we show that the effect of status is much larger when there is significant uncertainty surrounding producer- or product-level quality.

The narrowness of the empirical test in the paper is both its core strength and weakness. On one hand, we believe that it is one of the cleanest tests yet of Merton's famous hypothesis. Moreover, we believe the research design is very much aligned with the spirit of Merton's (1968) thought experiment in which he compares the careers of the 40^{th} to the so-called 41^{st} chair, when the former is actually elected to the French Academy and the latter barely misses the cut.

By limiting the empirical analysis to previously published articles, we attempt to tightly estimate the effect of a shock to an actor's status on changes in perceptions of the actor's quality. The consequence of this tight comparison is that we neglect other pathways through which changes in status influence performance outcomes. It may be, for instance, that through the implicit anointment into the academic super-elite that co-occurs with appointment to HHMI investigatorship, prize-winners gain preferential access to the most prominent journals in their fields. Or, they may benefit from privileged access to very tangible forms of resources, such as state-of-the-art laboratory equipment. Insofar as these forms of resource access can be causally related to changes in status, our analysis may significantly understate the full consequence of gains in status, even if it correctly spotlights its effect through changes in other-party perceptions of a focal actor's outputs. Our goal in this paper was to present this narrow test focused on a specific mechanism. In future work, similar research designs can be developed to illuminate the other routes through which status affects attainment.

References

Allison, Paul D., J. Scott Long, and Tad K. Krauze. 1982. "Cumulative Advantage and Inequality in Science." *American Sociological Review* **47**(5): 615-625.

Anderson, Cameron, Oliver P. John, Dacher Keltner, and Ann M. Kring. 2001. "Who Attains Social Status? Effects of Personality and Physical Attractiveness in Social Groups." *Journal of Personality and Social Psychology* **81**(1): 116–132.

Anderson, Cameron, Sandra E. Spataro, and Francis J. Flynn. 2008. "Personality and Organizational Culture as Determinants of Influence." *Journal of Applied Psychology* **93**(3): 702–710.

Azoulay, Pierre, Andrew Stellman, and Joshua Graff Zivin. 2006. "PublicationHarvester: An Open-source Software Tool for Science Policy Research." *Research Policy* **35**(7): 970-974.

Bales, Robert F., Fred L. Strodtbeck, Theodore M. Mills, and Mary E. Roseborough. 1951. "Channels of Communication in Small Groups." *American Sociological Review* **16**(4): 461-488.

Baum, Joel A.C., and Christine Oliver. 1991. "Institutional Linkages and Organizational Mortality." *Administrative Science Quarterly* **36**(2): 187-218.

Benjamin, Beth A., and Joel M. Podolny. 1999. "Status, Quality, and Social Order in the California Wine Industry." *Administrative Science Quarterly* **44**(3): 563-589.

Berger, Joseph, Susan J. Rosenholtz, and Morris Zelditch, Jr. 1980. "Status Organizing Processes." *Annual Review of Sociology* **6**: 479–508.

Blackwell, Matthew, Stefano Iacus, Gary King, and Giuseppe Porro. 2009. "cem: Coarsened Exact Matching in Stata." *The Stata Journal* **9**(4): 524-546.

Blau, Peter M. 1955. *The Dynamics of Bureaucracy*. Chicago, IL: University of Chicago Press.

Blau, Peter M., and W. Richard Scott. 1962. *Formal Organizations: A Comparative Approach*. San Francisco, CA: Chandler.

Burris, Val. 2004. "The Academic Caste System: Prestige Hierarchies in PhD Exchange Networks." *American Sociological Review* **69**(2): 239-264.

Cole, Stephen. 1970. "Professional Standing and the Reception of Scientific Discoveries." *American Journal of Sociology* **76**(2): 286-306.

Cole, Stephen, and Jonathan R. Cole. 1968. "Visibility and the Structural Bases of Awareness of Scientific Research." *American Sociological Review* **33**(3): 397-413.

Crane, Diane. 1965. "Scientists at Major and Minor Universities: A Study of Productivity and Recognition." *American Sociological Review* **30**(5): 699-714.

Fire, Andrew, SiQun Xu, Mary K. Montgomery, Steven A. Kostas, Samuel E. Driver, and Craig C. Mello. 1998. "Potent and Specific Genetic Interference by Double-stranded RNA in Caenorhabditis Elegans." *Nature* **391**(6669): 806-811.

Furman, Jeffrey, and Scott Stern. Forthcoming. "Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Knowledge Production." *American Economic Review*.

Gould, Roger V. 2002. "The Origins of Status Hierarchies: A Formal Theory and Empirical Test." *American Journal of Sociology* **107**(5): 1143-1178.

Hagstrom, Warren. 1971. "Inputs, Outputs and the Prestige of University Science Departments." *Sociology of Education* **44**(4): 375-397.

Hogg, Michael A. 2001. "A Social Identity Theory of Leadership." *Personality and Social Psychology Review* **5**(3): 184–200.

Keith, Bruce, and Nicholas Babchuk. 1998. "The Quest for Institutional Recognition: A Longitudinal Analysis of Scholarly Productivity and Academic Prestige among Sociology Departments." *Social Forces* **76**(1): 495-1533.

Keith, Bruce, and Nicholas Babchuk. 1994. "A Longitudinal Assessment of Productivity in Prominent Sociology Journals and Departmental Prestige." *American Sociologist* **25**(4): 4-27.

Levin, Sharon G., and Paula E. Stephan. 1991. "Research Productivity over the Life Cycle: Evidence for Academic Scientists." *American Economic Review* **81**(1): 114-32.

Long, J. Scott, and Mary Frank Fox. 1995. "Scientific Careers: Universalism and Particularism." *Annual Review of Sociology* **21**: 45-71.

Lynn, Freda B., Joel M. Podolny, and Lin Tao. 2009. "A Sociological (De)Construction of the Relationship between Status and Quality." *American Journal of Sociology* **115**(3): 755-804.

Merton, Robert K. 1968. "The Matthew Effect in Science." *Science* **159**(3810): 56-63.

Phillips, Damon J., and Ezra W. Zuckerman. 2001. "Middle Status Conformity: Theoretical Restatement and Empirical Demonstration in Two Markets." *American Journal of Sociology* **107**(2): 379-429.

Podolny, Joel M. 1993. "A Status-based Model of Market Competition." *American Journal of Sociology* **98**(4): 829-872.

Podolny, Joel M. 2005. *Status Signals: A Sociological Study of Market Competition.* Princeton, NJ: Princeton University Press.

Podolny, Joel M., and Toby E. Stuart. 1995. "A Role-Based Ecology of Technological Change." *American Journal of Sociology* **100**(5): 1224-1260.

Podolny, Joel M., Toby E. Stuart, and Michael T. Hannan. 1996. "Networks, Knowledge, and Niches: Competition in the Worldwide Semiconductor Industry, 1984-1991." *American Journal of Sociology* **102**(3): 659-689.

Simcoe, Timothy S., and David Waguespack. 2010. "What's in a (Missing) Name? Status and Signaling in Open Standards Development." Working Paper, Boston University.

Stuart, Toby E., Ha Hoang, and Ralph C. Hybels. 1999. "Interorganizational Endorsements and the Performance of Entrepreneurial Ventures." *Administrative Science Quarterly* **44**(2): 315-349.

Zucker, Lynne G., Michael R. Darby, and Marilynn B. Brewer. 1998. "Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises." *American Economic Review* **88**(1): 290-306.

Zuckerman, Ezra W. 1999. "The Categorical Imperative: Securities Analysts and the Illegitimacy Discount." *American Journal of Sociology* **104**(5): 1398-1438

Table 1: Descriptive Statistics for Pre-Appointment Articles (n=2×5,385) ECPW Producer Control Group

		Mean	Median	Std. Dev.	Min.	Max.
ECPWs						
Matched by CEM Procedure	Number of Authors	4.320	4	2.250	1	22
	Focal Author is Last	0.567	1	0.496	0	1
	Journal Impact Factor	8.926	7	5.805	0	35
	Article Baseline Stock of Citations	22.737	11	30.535	0	316
	Article Baseline Stock of Citations from High Status Citers	1.374	0	3.043	0	39
	Article Publication Year	1988.751	1988	5.298	1974	2002
	Appointment Year	1992.885	1993	4.981	1984	2003
Unmatched by CEM Procedure	Focal Author Graduation Year	1971.479	1971	8.274	1956	1998
	Focal Author Gender	0.112	0	0.315	0	1
	Focal Author Number of 'Hits'	29.178	20	30.281	0	341
	Article Novelty	16.199	16	6.151	1	42
	Article Stock of Citations up to Year 10	66.953	41	86.733	0	1,491
HHMIs						
Matched by CEM Procedure	Number of Authors	4.340	4	2.381	1	32
	Focal Author is Last	0.567	1	0.496	0	1
	Journal Impact Factor	8.926	7	5.805	0	35
	Article Baseline Stock of Citations	23.551	12	31.084	0	322
	Article Baseline Stock of Citations from High Status Citers	1.933	0	3.775	0	42
	Article Publication Year	1988.751	1988	5.298	1974	2002
	Appointment Year	1992.885	1993	4.981	1984	2003
Unmatched by CEM Procedure	Focal Author Graduation Year	1977.294	1978	7.605	1956	1998
	Focal Author Gender	0.140	0	0.347	0	1
	Focal Author Number of 'Hits'	31.528	25	23.492	0	119
	Article Novelty	15.467	15	6.089	0	40
	Article Stock of Citations up to Year 10	71.346	43	95.713	0	2,213

Note: The match is "article-centric," i.e., the control article is always chosen from the same journal in the same publication year. The control article is coarsely matched on the number of authors (exact match for one, two, and three authors; four or five authors; between six and nine authors; and more than nine authors). We also match on focal scientist position in the authorship roster (first author; last author; middle author). For articles published one year before appointment, we also match on the month of publication. For articles published two years before appointment, we also match on the quarter of publication. In addition, control and treatment articles are matched on citation dynamics up to the year before the (possibly counterfactual) appointment year. The cost of a very close, non-parametric match on article characteristics is that author characteristics do not match closely. Imposing a close match on focal scientist age, gender, and overall eminence at baseline results in a match rate which is unacceptably low. A possible compromise is to not match on journal, but to match on author characteristics. This alternative does not change our overall message.

Table 2: Variation in the HHMI Post- Appointment Citation Boost

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
	Journal Prestige		**Novel vs. Not**		**Young vs. Not at Appnt.**		**Well-cited vs. Not at Appnt.**	
	High JIF	Low JIF	Novel	Not	Young	Old	Low	High
After Appointment	1.218*	3.458**	3.083**	1.361	2.611**	1.481	2.745**	1.929*
	(0.505)	(1.279)	(0.924)	(0.949)	(0.830)	(1.244)	(1.000)	(0.953)
Nb. of Observations	4,733	5,162	5,578	4,317	6,612	3,283	6,256	3,639
Nb. of Article Pairs	453	491	538	406	634	310	585	359
Nb. of Scientists	211	234	234	203	167	143	168	142
Adjusted R^2	0.642	0.717	0.724	0.692	0.725	0.684	0.720	0.673

Standard errors in parentheses, clustered by scientists. All specifications are estimated by OLS; the models include focal scientist career age indicator variables, as well as article-pair fixed effects.
† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 3: Effects of HHMI Appointment on Citation Rates

	(1)	(2a)	(2b)	(3a)	(3b)
	All	**PI vs. Non-PI Pubs**		**Citer Status**	
		First/Last	Middle	Non-Elite	Elite
After Appointment	2.321**	4.087**	1.553*	2.095**	0.226**
	(0.703)	(1.514)	(0.646)	(0.665)	(0.077)
Nb. of Observations	9,895	3,094	6,801	9,895	9,895
Nb. of Article Pairs	944	298	646	944	944
Nb. of Scientists	310	156	260	310	310
Adjusted R^2	0.710	0.730	0.687	0.713	0.428

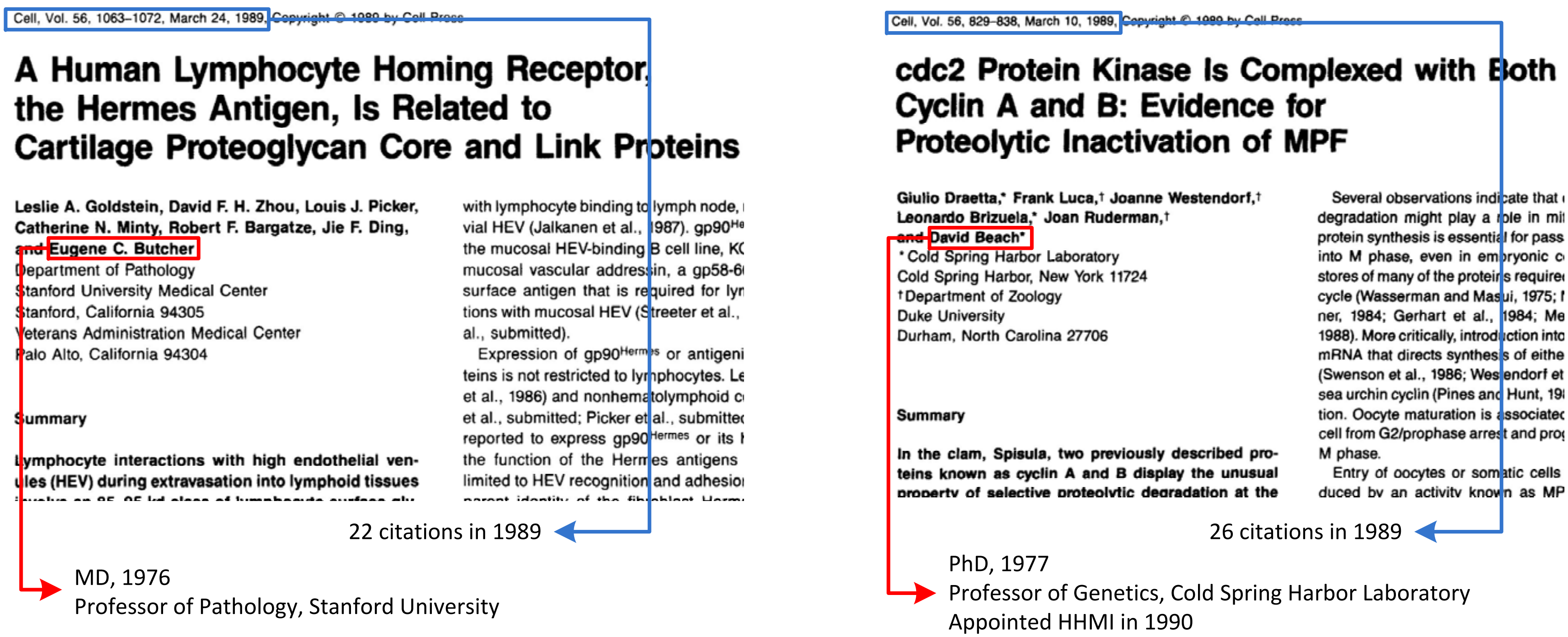
Standard errors in parentheses, clustered by scientists. All specifications are estimated by OLS; the models include focal scientist career age indicator variables, as well as article-pair fixed effects.
† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

Table 4: Descriptive Statistics for Pre-Appointment Articles (n=2×736) "Average" Producer Control Group

		Mean	Median	Std. Dev.	Min.	Max.
"Average" Scientists						
	Number of Authors	4.524	4	2.319	1	25
	Focal Author is Last	0.311	0	0.463	0	1
Matched by	Journal Impact Factor	6.460	6	3.410	1	30
CEM Procedure	Article Baseline Stock of Citations	7.245	1	15.288	0	140
	Article Publication Year	1990.514	1990	5.113	1977	2002
	Appointment Year	1993.579	1994	4.978	1984	2003
	Focal Author Graduation Year	1980.909	1981	6.978	1965	1998
Unmatched by	Focal Author Gender	0.224	0	0.417	0	1
CEM Procedure	Focal Author Number of 'Hits'	3.105	2	3.521	0	22
	Article Stock of Citations up to Year 10	40.141	19	95.759	0	1,654
HHMIs						
	Number of Authors	4.556	4	2.372	1	26
	Focal Author is Last	0.311	0	0.463	0	1
Matched by	Journal Impact Factor	6.460	6	3.410	1	30
CEM Procedure	Article Baseline Stock of Citations	7.777	1	15.942	0	157
	Article Publication Year	1990.514	1990	5.113	1977	2002
	Appointment Year	1993.579	1994	4.978	1984	2003
	Focal Author Graduation Year	1979.185	1980	7.605	1956	1998
Unmatched by	Focal Author Gender	0.171	0	0.377	0	1
CEM Procedure	Focal Author Number of 'Hits'	28.423	21	24.164	0	119
	Article Stock of Citations up to Year 10	49.689	25	81.267	0	878

Note: The match is "article-centric," i.e., the control article is always chosen from the same journal in the same publication year. The control article is coarsely matched on the number of authors (exact match for one, two, and three authors; four or five authors; between six and nine authors; and more than nine authors). We also match on focal scientist position in the authorship roster (first author; last author; middle author). For articles published one year before appointment, we also match on the month of publication. For articles published two years before appointment, we also match on the quarter of publication. In addition, control and treatment articles are matched on citation dynamics up to the year before the (possibly counterfactual) appointment year.

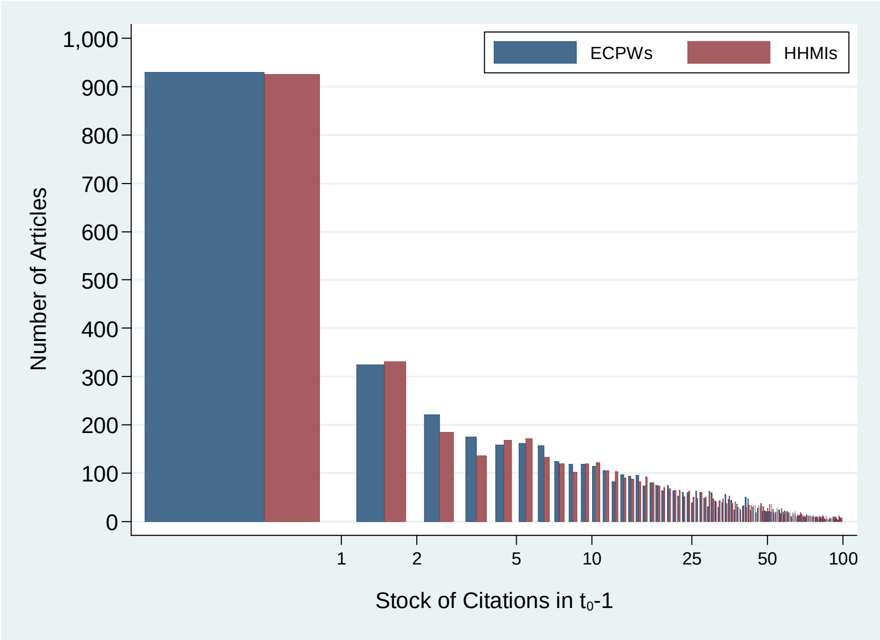
Figure 1: Matching Procedure



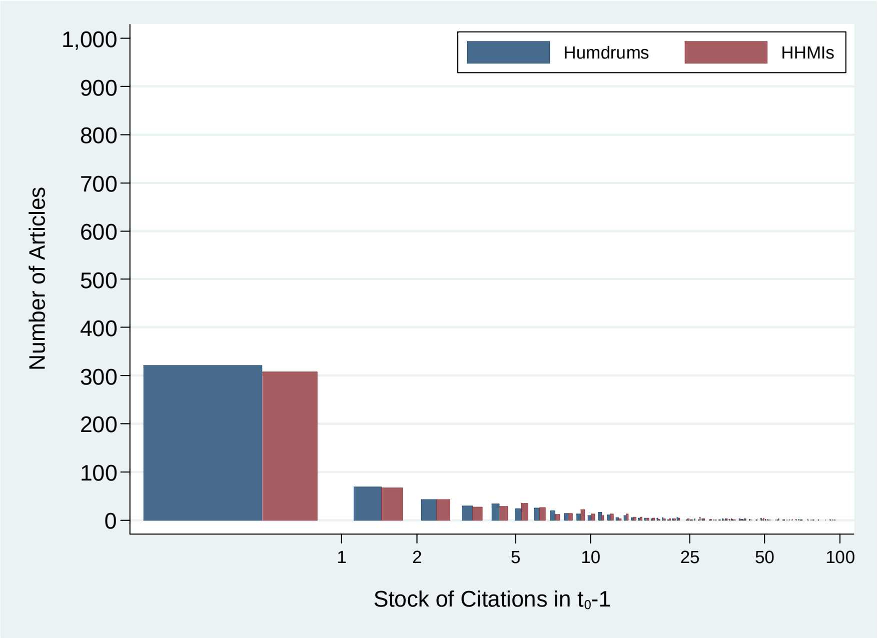
Note: The two articles above illustrate the essence of the Coarse Exact Matching procedure. These two articles appeared in the journal *Cell* in March 1989. They received a very similar number of citations in the birth year (1989): 22 citations for Goldstein et al.; 26 citations for Draetta et al. David Beach, the PI on the article on the right-hand side, was appointed a Howard Hughes Medical Investigator in 1990.

Figure 2
Covariate Balance, Cumulative Citations at Baseline
Pre-Appointment Articles

A. ECPW Producer Control Group



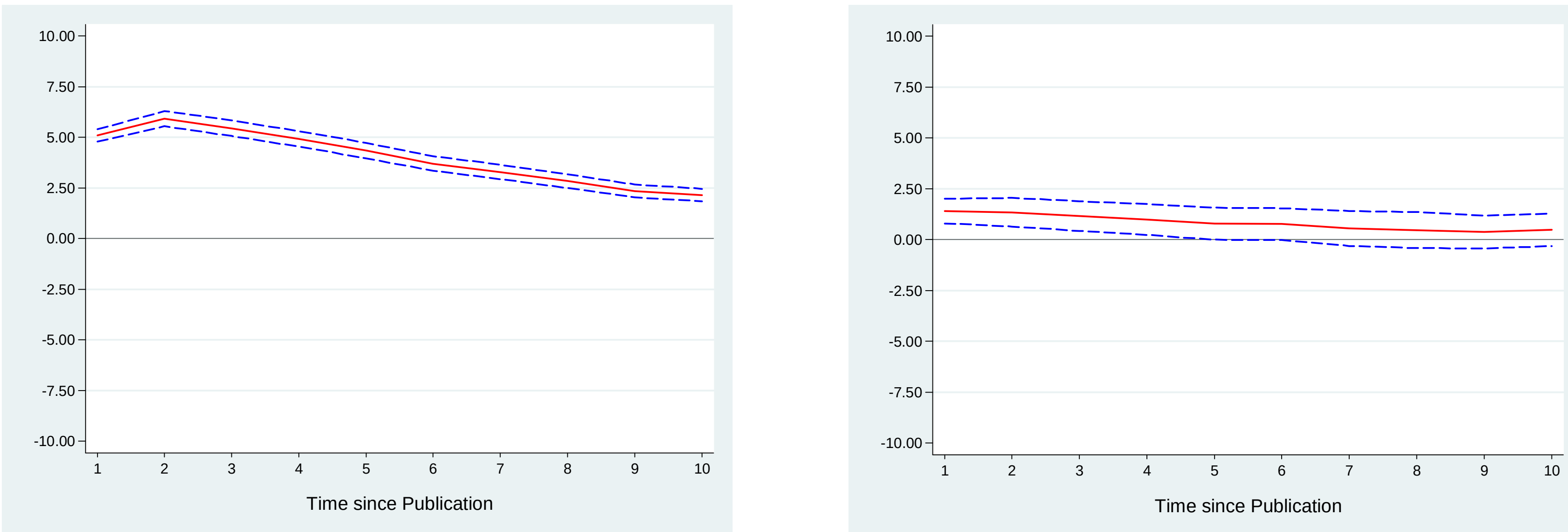
B. "Average" Producer Control Group



Note: We compute the cumulative number of citations for treatment and control articles, respectively, up to the year that immediately precedes the year of appointment. Panel A compares the distribution for 2×5,385 articles corresponding to articles associated with HHMI scientists and ECPW scientists. Panel B compares the distribution for 2×736 articles corresponding to articles associated with HHMI scientists and "Average" scientists.

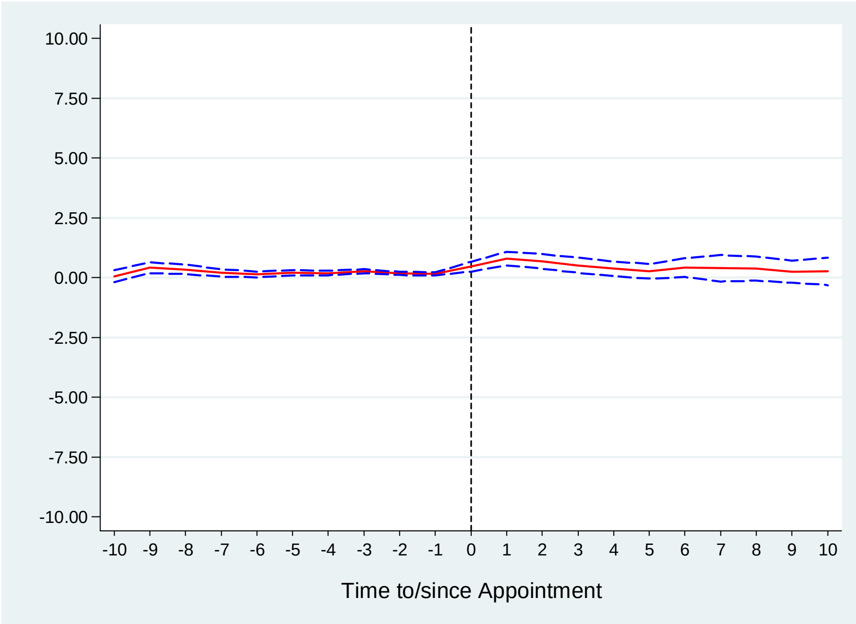
Figure 3
Effect of HHMI Appointment on Citation Rates
Post-Appointment Articles, ECPW Producer Control Group

A. Without Journal Match **B. With Journal Match**



Notes: Dynamics for the difference in yearly citations between HHMI-ECPW matched articles written in the post-appointment period. Articles in each pair are published in the same year, and the focal scientists are matched on degree year, gender, and eminence as indicated by the number of articles they published up to the year of appointment that fall in the top ventile of the vintage-specific article-level distribution of citations. In Panel B, the match is further constrained so that the two articles in each pair appeared in the same scientific journal.

Figure 4
Effect of HHMI Appointment on Citation Rates
Pre-Appointment Articles, ECPW Producer Control Group



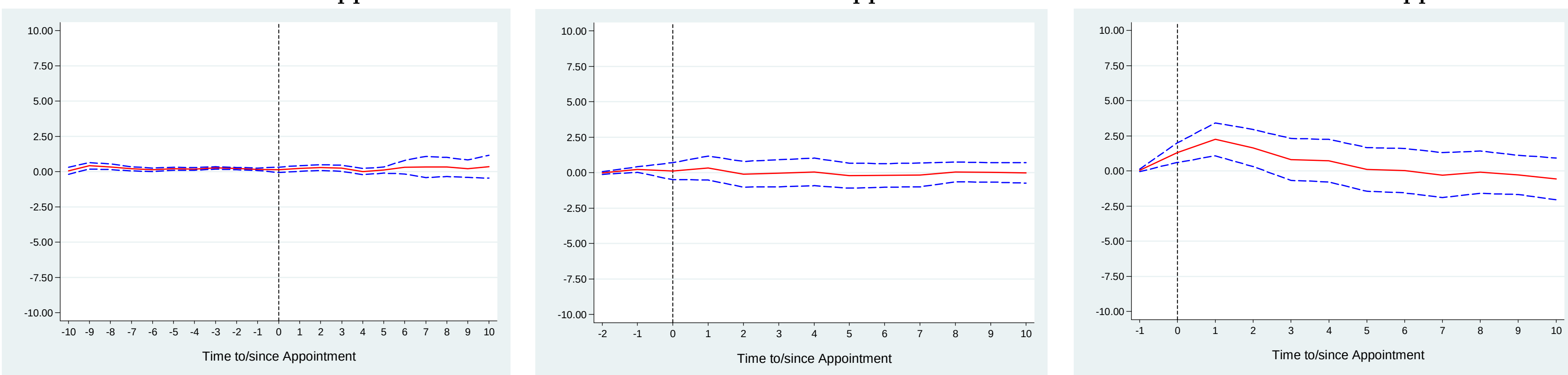
Note: Dynamics for the difference in yearly citations between HHMI-ECPW matched articles. The sample includes articles of vintage t_0-10 to t_0-1 — where t_0 is the year of (possibly counterfactual) appointment.

Figure 5
Effect of HHMI Appointment on Citation Rates
Pre-Appointment Articles, ECPW Producer Control Group

A. Articles Published Between 3 & 10 Years Before Appointment

B. Articles Published Two Years Before Appointment

C. Articles Published One Year Before Appointment

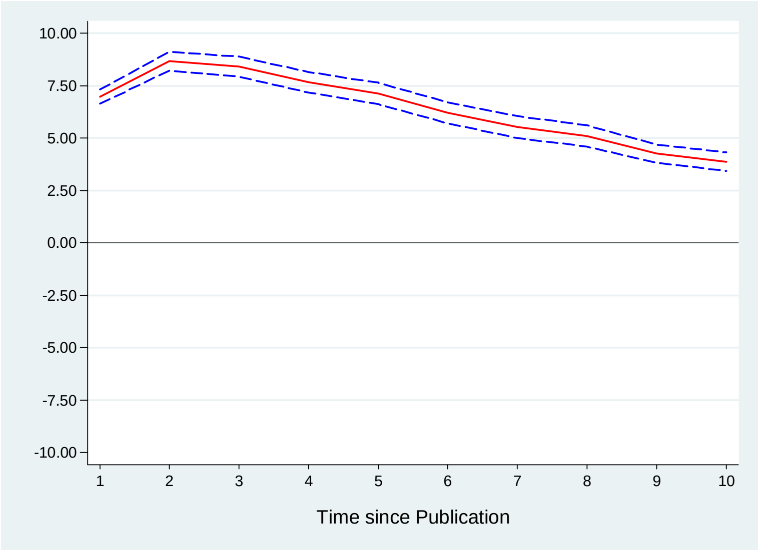


Note: Dynamics for the difference in yearly citations between HHMI-ECPW matched articles. Articles in each pair appeared in the same year and journal, and are also matched on focal scientist position on the authorship list, as well as overall number of authors. Further, control articles are selected such that the sum of squared differences in citations between control and treated article up to year t_0-1 is minimized — where t_0 is the year of (possibly counterfactual) appointment.

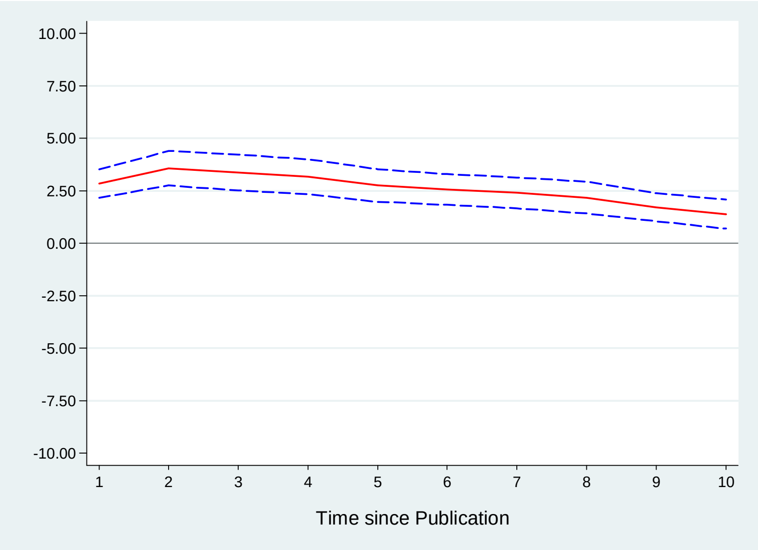
In Panel A, the sample is limited to articles published between year t_0-3 and year t_0-10. In Panel B, the sample is limited to articles published in year t_0-2. In addition to being matched on journal, focal scientist position on the authorship list, and overall number of authors, the articles in each pair appeared in the same quarter. In Panel C, the sample is limited to articles published in year t_0-1. In addition to being matched on journal, focal scientist position on the authorship list, and overall number of authors, the articles in each pair appeared in the same month.

Figure 6
Effect of HHMI Appointment on Citation Rates
Post-Appointment Articles, “Average” Producer Control Group

A. Without Journal Match

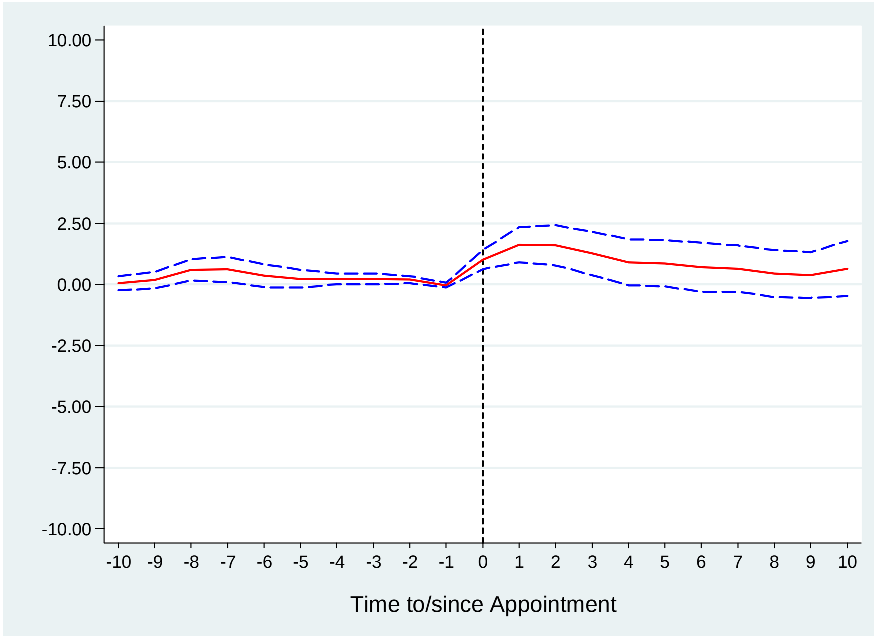


B. With Journal Match



Notes: Dynamics for the difference in yearly citations between HHMI-“Average” scientist matched articles written in the post-appointment period. Articles in each pair are published in the same year, and the focal scientists are matched on degree year, gender, and eminence as indicated by the number of articles they published up to the year of appointment that fall in the top ventile of the vintage-specific article-level distribution of citations. In Panel B, the match is further constrained so that the two articles in each pair appeared in the same scientific journal.

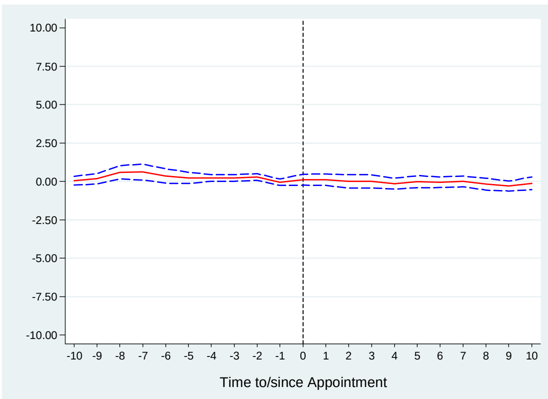
Figure 7
Effect of HHMI Appointment on Citation Rates
Pre-Appointment Articles, "Average" Producer Control Group



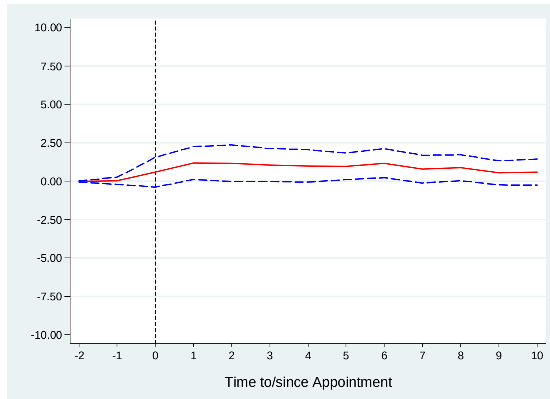
Note: Dynamics for the difference in yearly citations between HHMI-"Average" matched articles. The sample includes articles of vintage t_0-10 to t_0-1 — where t_0 is the year of (possibly counterfactual) appointment.

Figure 8
Effect of HHMI Appointment on Citation Rates
Pre-Appointment Articles, "Average" Producer Control Group

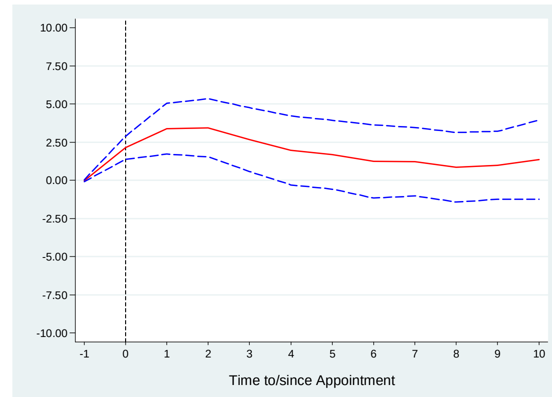
A. Articles Published Between 3 & 10 Years Before Appointment



B. Articles Published Two Years Before Appointment



C. Articles Published One Year Before Appointment



Note: Dynamics for the difference in yearly citations between HHMI-"Average" matched articles. Articles in each pair appeared in the same year and journal, and are also matched on focal scientist position on the authorship list, as well as overall number of authors. Further, control articles are selected such that the sum of squared differences in citations between control and treated article up to year t_0-1 is minimized — where t_0 is the year of (possibly counterfactual) appointment.

In Panel A, the sample is limited to articles published between year t_0-3 and year t_0-10. In Panel B, the sample is limited to articles published in year t_0-2. In addition to being matched on journal, focal scientist position on the authorship list, and overall number of authors, the articles in each pair appeared in the same quarter. In Panel C, the sample is limited to articles published in year t_0-1. In addition to being matched on journal, focal scientist position on the authorship list, and overall number of authors, the articles in each pair appeared in the same month.

Appendix I:
Construction of Article Control Group

We detail the procedure implemented to identify the set of control articles from among the set of articles published by early career prize winning (ECPW) scientists.

The sample of control articles is constructed such that the following two conditions are met:

1. treated articles exhibit no differential citation trends relative to control articles up to the time of appointment;
2. treated and control articles match on a number of time-invariant article characteristics;

We identify controls based on the following set of covariates: (1) year of publication; (2) specific journal (e.g. *Cell* or the *New England Journal of Medicine*); (3) number of authors (the distribution is coarsened into six bins: one, two, three, four or five, between six and nine, and ten or more authors); (4) focal-scientist position on the authorship list (first author, middle author, or last author). In the case of articles published in the year immediately preceding HHMI appointment, the list of matching covariates is expanded to also include the month of publication. In the case of articles published two years before appointment, the list of matching covariates is expanded to also include the quarter of publication. To ensure that pre-appointment citation trends are similar between articles written by HHMIs and their twins drawn from EPCW-authored papers, we also match on cumulative number of citations at the time of appointment, coarsened into 7 strata (0 to 10^{th}; 10^{th} to 25^{th}; 25^{th} to 50^{th}; 50^{th} to 75^{th}; 75^{th} to 95^{th}; 95^{th} to 99^{th}; and above the 99^{th} percentile).

We create a large number of strata to cover the entire support of the joint distribution of the covariates mentioned above. Each observation is allocated to a unique stratum. We then drop from the data all observations corresponding to strata in which there is no treated article and all observations corresponding to strata in which there are less than 5 potential controls. We have found that matching on cumulative citations at baseline/time of treatment is not enough to eliminate pre-move citation trends. To ensure that citation dynamics coincide for treated and control observations, we select among potential matches a single article that further minimizes the sum of squared differences in the number of citations between treated and control articles up until the year that precedes the appointment year.

The procedure is coarse because we do not attempt to precisely match on covariate values; rather, we coarsen the support of the joint distribution of the covariates into a finite number of strata, and we match a treated observation if and only if a control observation can be recruited from this stratum. An important advantage of CEM is that the analyst can guarantee the degree of covariate balance *ex ante*, but this comes at a cost: the more fine-grained the partition of the support for the joint distribution (i.e., the higher the number of strata), the larger the number of unmatched treated observations.

We implement the CEM procedure year by year, without replacement. Specifically, in year of appointment t, $1984 \leq t \leq 2003$, we:

1. eliminate from the set of potential controls all articles published by ECPW scientists who have collaborated with HHMI scientists prior to year t;
2. for each year of publication $t-k$, $1 \leq k \leq 10$;
   (a) create the strata;
   (b) identify within strata a control for each treated unit; break ties at random;
   (c) repeat these steps for year of publication $t-(k+1)$.
3. repeat these steps for year of appointment $t+1$.

Appendix II: Linking Scientists with their Journal Articles

The source of our publication data is PubMed, a publicly available bibliographic database maintained by the U.S. National Library of Medicine.[i] PubMed contains over 14 million articles from 4,800 journals published in the United States and more than 70 other countries from 1950 to the present. We have mined these data using PubHarvester, an open-source software tool that automates the process of gathering publication information for individual life scientists (Azoulay et al. 2006).

There are two major challenges that must be overcome to accurately link scientists to their publications. The first relates to what one might term "Type I Error," whereby we mistakenly attribute to a scientist a journal article actually authored by a namesake. The second relates to "Type II error," whereby we conservatively exclude from a scientist's publication roster legitimate articles:

Namesakes and popular names. PubMed does not assign unique identifiers to the authors of the publications they index. They identify authors simply by their last name, up to two initials, and an optional suffix. This makes it difficult to unambiguously assign publication output to individual scientists, especially when their last name is relatively common.

Inconsistent publication names. The opposite error—recording too few publications—occurs because scientists often are inconsistent in the choice of names they choose to publish under. By far the most common source of error is the haphazard use of a middle initial. Other errors stem from inconsistent use of suffixes (Jr., Sr., 2nd, etc.), or from multiple patronyms due to changes in spousal status.

To address with these measurement problems, we designed individual search queries that rely on relevant scientific keywords, the names of frequent collaborators, journal names, as well as institutional affiliations. Althogh the process of query design is very time consuming, it is feasible because we have scientists' CVs and biosketches. PubHarvester provides the option to use such custom queries in lieu of a completely generic query. For example, one can examine the publications of Scott A. Waldman, an eminent pharmacologist located in Philadelphia, PA at Thomas Jefferson University. Waldman is a relatively frequent name in the United States (with 208 researchers with an identical patronym in the American Association of Medical Colleges faculty roster); the combination `"waldman s"` is common to 3 researchers in the same database. A simple search query for `"waldman sa"[au] OR "waldman s"[au]` returns 302 publications at the time of this writing. However, a more refined query, based on Professor Waldman's biosketch returns only 210 publications.[ii]

The above example also makes clear how we deal with the issue of inconsistent publication names. PubHarvester gives the end-user the option to choose up to four PubMed-formatted names under which publications can be found for a given researcher. For example, Louis J. Tobian, Jr. publishes under `"tobian l"`, `"tobian l jr"`, and `"tobian lj"`, and all three names need to be provided as inputs to generate a complete publication listing.

We are very confident that such a degree of customization ensures the accuracy of treated and control scientists' bibliomes.

[i] `http://www.pubmed.gov/`

[ii] `(((("waldman sa"[au] NOT (ether OR anesthesia)) OR ("waldman s"[au] AND (murad OR philadelphia[ad] OR west point[ad] OR wong p[au] OR lasseter kc[au] OR colorectal))) AND 1980:2010[dp])`

Appendix III:
Identifying "Average" Scientists and their Publications

We identify 6,272 scientists who are, by construction, typical members of the profession. We label these scientists "average" to distinguish them from the small, elite sample of early career prize winners.

More precisely, we draw the names of these scientists from the Faculty Roster of the Association of American Medical Colleges (AAMC), to which we secured licensed access for the years 1975 through 2006. This roster is an annual census of all U.S. medical school faculty in which each faculty is linked across yearly cross-sections by a unique identifier. When all cross-sections are pooled, we obtain a matched employee/employer panel dataset. For each of the 230,000 faculty members that appear in the roster, we know the full name, the type of degrees received and the years they were awarded, gender, up to two departments, and medical school affiliation.

Because the roster only lists medical school faculty, however, it is not a complete census of the academic life sciences. For instance, it does not list information for faculty at institutions such as MIT, University of California at Berkeley, Rockefeller University, the Salk Institute, or the Bethesda campus of the NIH; and it also ignores faculty members in Arts and Sciences departments — such as biology and chemistry — if they do not hold joint appointments at a local medical school.

We begin by deleting from the Roster all past and present HHMI investigators; members of the National Academy of Sciences; and early career prize winners. We then delete all faculty members who earned their highest degree before 1956 or after 1998 (the lower and higher bounds for the HHMI sample), as well as those faculty members whose main department affiliation has no counterpart among HHMIs (such as veterinary sciences, public health and preventive medicine, otolaryngology, or anesthesiology).

In the absence of detailed biographical information on these individuals, our remaining concern is measurement error when linking them to publication output. To guard against incorrect bibliomes, we perform two additional screens. First, we eliminate from the data all faculty members whose last name appears more than once in the roster. Second, we drop faculty members for whom the roster does not provide a middle initial. In a final step, we impose the criterion that these scientists should have at least five career publications. This ensures that the 6,272 medical school faculty members who survive this succession of screens have more than a fleeting attachment to the academy.

We use the PubHarvester software tool described in Appendix II to collect these scientists' journal articles. The search queries supplied to PubHarvester are completely generic (e.g., `"seljeskog el"[au]` or `"abeyounis cj"[au]`), but this should be fairly innocuous given the rarity of the patronyms involved.

The collection of citation data proceeds in a fashion identical to that followed in the case of HHMI and ECPW scientists.