Gender Differences in Patenting in the Academic Life Sciences

Waverly W. Ding[1], Fiona Murray[2], and Toby E. Stuart[3,*]

[1]Haas School of Business
University of California, Berkeley, CA 94720

[2]MIT Sloan School of Management
Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02142

[3]Harvard Business School
Soldiers Field, Boston, MA 02163

[*]All authors contributed equally to this work. To whom correspondences should be addressed:
tstuart@hbs.edu

We analyze longitudinal data on academic careers and conduct interviews with faculty members to determine the scope and causes of the gender gap in patenting among life scientists. Our regressions on a random sample of 4,227 life scientists over a 30-year period show that women faculty members patent at about 40 percent of the rate of men. We find that the gender gap has been improving over time, although it remains large.

The gender gap in academic science is a topic of ongoing policy and scholarly debate. Studies in fields as diverse as engineering and biology have found that women scientists suffer from an attainment gap along at least three important dimensions – productivity, recognition and reward *(1-4)*. Fortunately, some recent evidence provides cause for optimism. Especially in fields within the academic life sciences, there has been a narrowing of the gender gap *(3, 5)*. Until recently, however, little research has explored an increasingly important source of non-salary remuneration for faculty: participation in commercial science (*6*). This omission is problematic given the growing opportunities for recognition and reward in the commercialization of scientific research. Although profiting from university research continues to generate controversy *(7)*, the reality is that commercial activities including patenting, consulting, and scientific advisory board (SAB) membership have become commonplace *(8)*. What limited evidence exists about "academic entrepreneurship" suggests a gender gap of considerable magnitude. For example, our research on SAB membership showed that of 771 SAB members in a large sample of young biomedical companies, only 6.5% were women *(9).*

In this article, we examine gender differences in one specific commercial activity—patenting. We conducted a longitudinal empirical analysis using a random sample of faculty in the life sciences employed in U.S. academic institutions. The analysis is complemented by interviews with life scientists at one prominent university. Although academic patents do not yield immediate financial returns to their inventors, frequently they are an avenue to a variety of rewards: royalty-bearing license agreements with established companies or startup formation with substantial equity participation. While not the only route to commercial engagement, our interviews and prior research suggest that patents are an important precursor to opportunities in industry.

We began the quantitative analysis by drawing a random sample of 12,000 life scientists from the *UMI Proquest Dissertations* database. We restricted the sample to those earning Ph.D.s between 1967 and 1995 in the scientific fields that have most fostered the commercial life sciences. We then used the *Science Citation Index* to collect the publications, coauthors, and employers of the individuals in the sample. As our interest is in academic careers, we retained only the 4,227 individuals with at least five years of post-Ph.D. publishing experience in academic institutions. Finally, we obtained the patents on which the scientists in the sample are listed as inventors. We then created a dataset of scientist-year observations with covariates for the individual's gender, annual publication activity, and annual patent count [supporting online text].

Of the scientists in our sample, 11.5% are listed as inventors on one or more patents. However, the full sample proportion masks a large gender difference: of the 903 women in the sample, 5.65% held patents as of the last year of the data. By contrast, 13% of the 3,324 male scientists in the data are listed on patents. Moreover, the 431 male patenters have amassed a total of 1,286 patents in our dataset. This compares to 92 patents produced by the 51 women patenters.

We have structured the data archive to enable survival analyses. Figure (1) displays gender-specific non-parametric survivor plots that show the likelihood that a scientist in the data has *not* patented up to a given year of professional tenure. The plots show that, at all career stages, the curve for male scientists is beneath that for women and the gender gap in survival probabilities increases over time.

As with other areas of scientific attainment, patenting is affected by a range of individual, field, and institutional factors, many of which may differ systematically between the sexes (*10-12*). As illustrated in the panels in Figure 2, after constructing four mutually exclusive sub-samples: male patent holders (MP), female patent holders (FP), men without patents (MNP), and women without patents (FNP) we observe considerable sub-sample differences in means (across levels of professional experience) in a) number of papers, b) amount of NIH grants, and c) number of papers coauthored with researchers in industry. Specifically, male patenters typically have the highest paper counts, most NIH grant money, and along with the women patenters, the most coauthorships with industry scientists (Table S1 for significance tests).

Given evidence of these disparities between male and female scientists, it is important to determine the conditional effect of gender on patenting—its effect net of other measured attributes of scientists. We therefore estimate scientist-level regressions of the rate of patenting. Formally, let $\lambda_i(t)$ designate the instantaneous transition rate, where t=1,…,35 years (we assume that 35 years is the maximum time any scientist is at risk). We estimate Cox proportional hazards regressions. Letting $\lambda_0(t)$ indicate the baseline hazard, $X_i(t\text{-}1)$ a vector of lagged, time-varying covariates, and V_i a vector of time-independent covariates, we estimate: $\lambda_i(t,X,V,\quad\gamma) = \lambda_0(t) \times \exp[\beta' X_i(t\text{-}1)+\gamma' V_i]$.

We include a number of variables in the $X_i(t\text{-}1)$ vector. One attribute that influences patenting is a scientist's research productivity. At the extreme, an unpublished scientist probably lacks novel findings to patent. We therefore include the number of articles each scientist has published in the previous five-year period and the square of this variable. Our interviewees suggested that scientists' employers also influence patenting by providing support for interactions with industry. To proxy for this workplace characteristic, we include the number of patents assigned to the scientist's employing university during the previous five years (excluding patents held by the focal scientist). Our logic is that universities with high patent counts will likely have an effective technology transfer office (TTO) and a culture that supports involvement in commercial endeavors.

Interviewees suggested that networks of colleagues influenced their patenting behavior, which accords with recent research on entrepreneurship *(8,13)*. Scientists (particularly male faculty) routinely mentioned consulting with co-authors, colleagues, and industry contacts for advice about the patent process. We capture faculty members' contacts with two variables. First, as a gauge of network reach, we include the average number of coauthors the scientist has had on previously published papers. Second, as a proxy for the richness of scientists' networks with industry, we include a dummy variable equal to one if a scientist has recently coauthored papers with one or more researchers in industry.

Figure 3 illustrates our regression findings (Table S2). The parameter estimates suggest that an increase in the five-year publication count of one standard deviation of the observed distribution (12.1 papers) multiplies the hazard rate of patenting by a factor of 1.81. Similarly, scientists that average a greater number of coauthors per paper, that work at universities that promote patenting, and those that have collaborative projects with scientists in industry patent at a greater rate. There also are statistically significant scientific field effects. Relative to the omitted category (biology, genetics) in the regression, molecular biologists, immunologists, and organic chemists (not shown) patent at a significantly higher rate.

After accounting for the substantial effects of productivity, networks, field, and employer attributes, what is the net effect of gender? There remains a large, statistically significant ($p<0.001$) effect of being female. The parameter estimate implies that, holding constant productivity, social network, scientific field, and employer characteristics, comparable women life scientists patent at only 0.40 times the rate of equivalent male scientists. This finding begs the question: what might cause such a large gender difference in patenting?

One possibility is that men and women do qualitatively different kinds of research. In particular, if women are risk averse in their research choices (*14*), there may be a gender difference in research "patentability". We believe such a difference would manifest in the extent of scholarly impact. To explore this possibility, we created a dataset of the 23,436 articles published by the women in our sample and matched each paper (by publication year) with a randomly drawn article from the pool of male scientists' papers. This yields a sample of articles with a 1:1 gender ratio. We then examined, by gender and year, the average number of citations and the *Journal Impact Factor* (JIF) of these papers. We find that the per-article mean citation count for male scientists is very similar to that of women (Table S3). Moreover, the gender gap in average JIF actually favors women (average JIF for male: 4.06; for female, 4.12). Overall, there is no evidence that women do less significant work based on standard measures of scientific impact.

If there is no measurable gender difference in the scholarly influence of research, what else might cause such a large gender difference in patenting? For clues, we turn to our faculty interviews, in which two factors loomed large. The first is lack of exposure to the commercial sector. Most (but not all) women had few contacts in industry. Lacking these connections, women found it time-consuming to gauge whether an idea was commercially relevant. In contrast, men often described an industry contact as a precursor to patenting. Hampered by their narrow networks and concerned about the time it would take to "shop" a patent around, several female faculty were deterred from completing a patent filing. Thus, differences in the composition of professional networks meant that the time cost of patenting was higher for many women faculty.

Several women suggested a second hurdle: concern that pursuing commercial opportunities might hinder their university careers. The women we interviewed were more likely to describe the challenges associated with balancing multiple career elements; teaching, research and commercialization. Unlike their male counterparts who described their patenting decisions as unproblematic and driven by translational interests, female faculty expressed concern about the potentially negative impact that patenting might have on education, collegiality, and research quality.

Our interviews also uncovered two factors that reduced the perceived costs of patenting for female faculty: collegial support and institutional assistance. Compared to men, female faculty were much more likely to be encouraged in patenting by their (typically male) co-authors, who often drove the patenting process. While men sought advice from their often broad-reaching networks, women frequently depended on close relationships with male collaborators to initiate the patenting process. Formal institutional sponsorship was also particularly important for women. Many women commented that their TTO provided industry contacts, advice, and encouragement to develop the commercial aspects of their research.

Our interviews also exposed differences between older and younger women scientists. Most senior female faculty we met perceived being excluded from industry relationships and therefore failed to develop an understanding of how commercial science works. Few made the transition to patenting. Some of the younger (but tenured) female life scientists had begun to incorporate patenting into their research strategy. Nonetheless, many still felt at a disadvantage to their male colleagues due to their limited experience at the academic-industry boundary. It is only among junior faculty that we found parity in attitudes, which were shaped by doctoral and post-doctoral experiences. Regardless of gender, those that experienced patenting during training were undaunted by the challenges of combining academic and commercial science.

Because our data spans 35 years, we can determine whether such generational distinctions are evident in the larger sample. For three Ph.D. cohorts (those earning degrees from 1967-1975; 1976-1985; and 1986-1995) we examined gender-specific non-parametric cumulative hazard plots. For each cohort, we also calculate the male-to-female ratio of the cumulative hazards. For illustration, at the 10th year after scientists earned their Ph.D., the cumulative hazard of patenting for male scientists was 4.4 times higher than women in the 1967-1975 cohort; 2.1 times higher in the 1976-1985 cohort; and 1.8 times larger in the 1986-1995 cohort (fig. S1). Thus, consistent with our interview findings, the archival data indicate that the gender gap in patenting rates has been declining.

Our analyses suggest that patenting has become common in the academic life sciences, particularly for highly productive and networked faculty. Among the most senior faculty, a large gender gap persists, reinforced by women's limited commercial networks and traditional views of academic careers. For younger cohorts patenting is widely embraced, although a gender gap remains. Increasingly, however, young female faculty are similar to their male colleagues: they view patents as accomplishments and as a legitimate means to disseminate research. If this trend continues, we may observe further declines in the magnitude of the gender gap in commercializing academic science.

References and Notes

1. J.R. Cole, S. Cole, *Social Stratification in Science* (Univ. Chicago Press, Chicago, IL, 1973).

2. J.S. Long, M.F. Fox, *Annu. Rev. Soc*. **21,** 45-71 (1995).

3. National Science Foundation, *Gender Differences in the Careers of Academic Scientists and Engineers* (NSF 04-323, 2004; http://www.nsf.gov/sbe/srs/ns04323/pdf/front.pdf).

4. G. Sonnert, G. Holton, *Who Succeeds in Science? The Gender Dimension* (Rutgers Univ. Press, New Jersey, 1995).

5. C. Holden, *Science*, **294**, 396-411 (2001). (http://sciencecareers.sciencemag.org/feature/salsurvey/v294i5541p396.htm)

6. J. Thursby, M. Thursby, *J. Technol Transfer* **30**, 343-353 (2005).

7. J. P. Walsh, A. Arora, W.M. Cohen, *Science* **299**, 1021 (2003).

8. T.E. Stuart, W. Ding, *Amer. Jnl. Soc.*, in press (2006).

9. W. Ding, F. Murray, T. Stuart, in preparation.

10. H. Etzkowitz, C. Kemelgor, M. Neuschatz, B. Uzzi, J. Alonzo, *Science* **266**, 51-54 (1994).

11. Y. Xie, K. A. Shauman *Women in Science: Career Processes and Outcomes* (Harvard Univ.Press, Cambridge, MA, 2003)

12. P. D. Allison, J. S. Long, *Amer. Soc. Rev*. **55**, 469-478 (1990).

13. S. Shane, T. Stuart, *Man. Sci*. **48**, 154-170 (January 2002).

14. R.L. Fisher, The Research Productivity of Scientists: How Gender, Organizational Culture and Problem Choice Processes Influence the Productivity of Scientists (University Press of America, 2005)

Supporting Online Material

www.sciencemag.org

Tables S1, S2, S3, S4

Fig. S1

Murray acknowledges financial support from the Cambridge-MIT Institute. Stuart acknowledges financial support from the Center for Entrepreneurial Leadership at the Ewing Marion Kauffman Foundation, Kansas City, MO. This material is partly based upon work supported by the National Science Foundation under Grant No. EEC-0345195 (Stuart).

Fig. 1: Gender-specific Kaplan-Meier estimates of the survivorship function of first patenting with confidence intervals. Shows the likelihood that scientists (blue line for male and red dashed line for female) have *not* patented up to a given year of professional experience. Both the stratified log-rank test and Wilcoxon test reject (p<0.01) the hypothesis that the survival functions are equal.

Fig. 2: Mean publication count (A), NIH grant totals (B), number of jointly authored papers with industry researchers (C) during first 20 years of scientists' careers. Groups are: male patenters (blue, squares), male without patents (light blue, triangles), female patenters (orange, circles) and female without patents (cranberry, squares). Although women patenters appear to have a higher mean grant total than male patenters beyond the 18th year of professional experience, this difference is not statistically significant. (Table S1 for equality of means tests across categories.)

Fig. 3: Hazard ratios and 95% confidence intervals from Cox regression of time until patenting. Hazard ratio *g* implies that the probability of patenting changes by a factor of *g* for a unit (dichotomous variables) or standard deviation (continuous variables) change in the covariate value. Predictors are sorted by effect magnitude and are statistically significant if 1.0 falls outside of the confidence interval. (Full regression results in Table S2.)

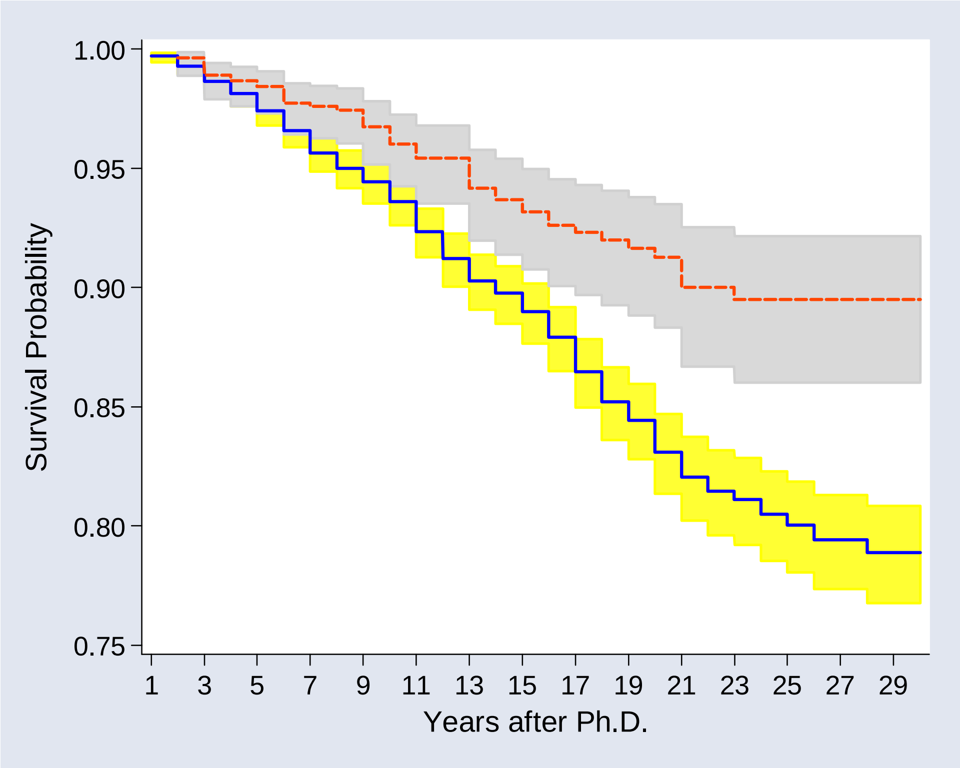

1.00
0.95
0.90
0.85
0.80
0.75
Survival Probability
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
Years after Ph.D.

Fig.1

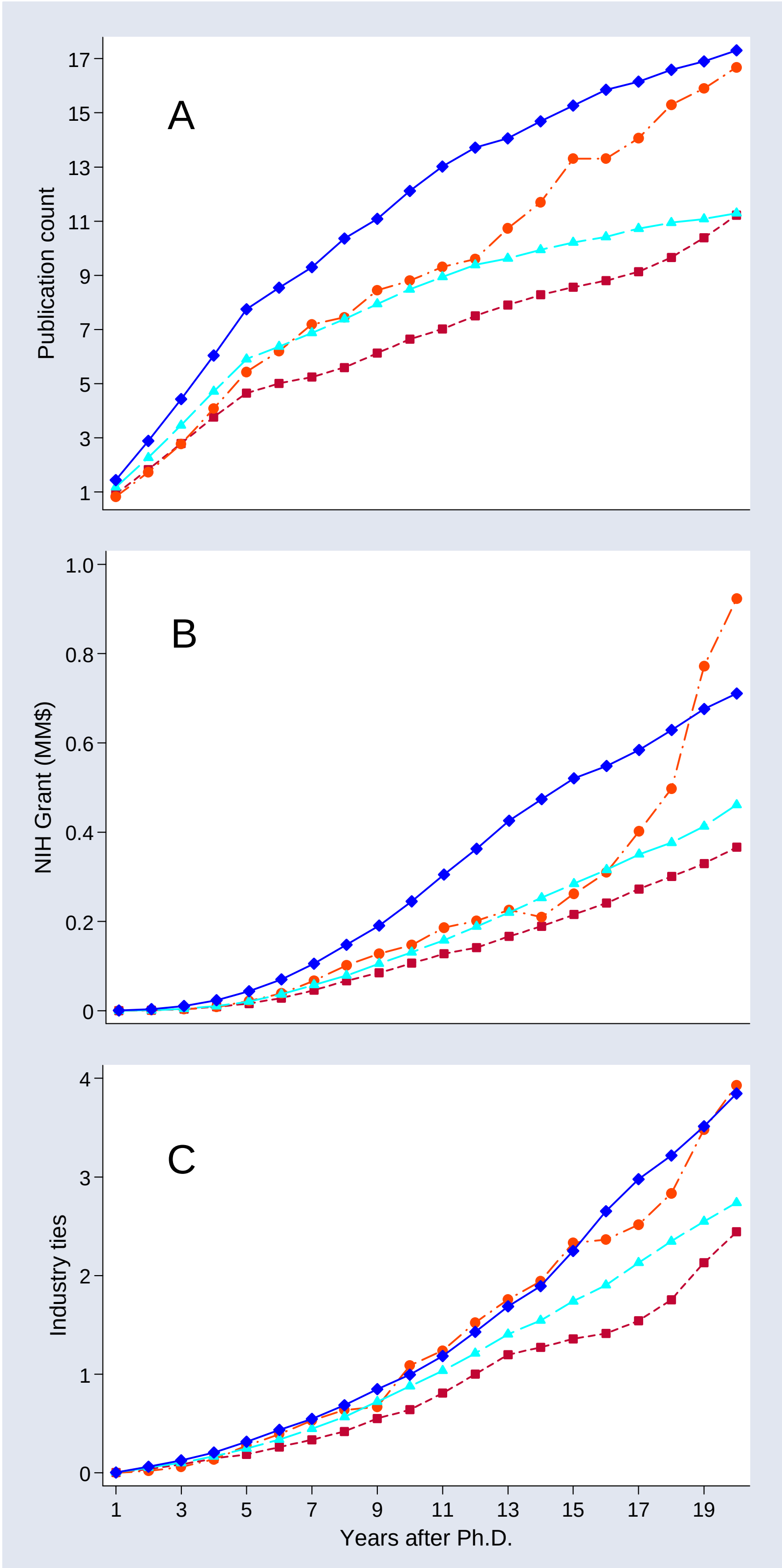

A
Publication count
17
15
13
11
9
7
5
3
1
B
NIH Grant (MM$)
1.0
0.8
0.6
0.4
0.2
0
C
Industry ties
4
3
2
1
0
1
3
5
7
9
11
13
15
17
19
Years after Ph.D.

Fig.2

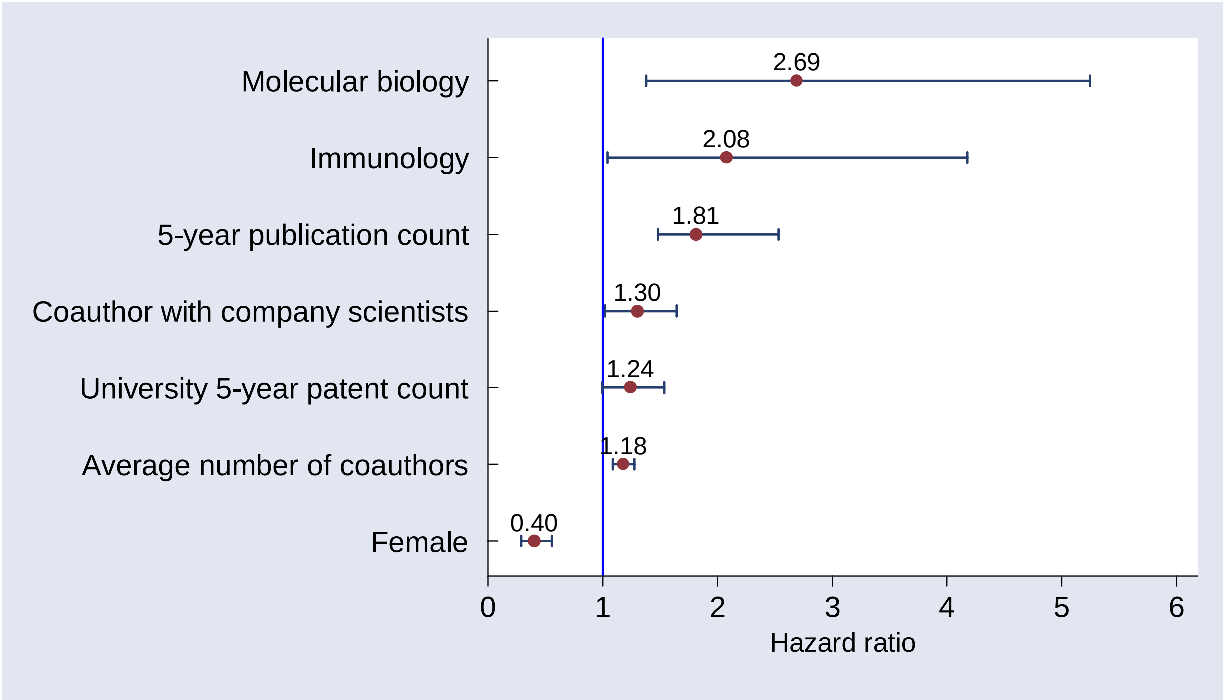

Molecular biology
Immunology
5-year publication count
Coauthor with company scientists
University 5-year patent count
Average number of coauthors
Female
2.69
2.08
1.81
1.30
1.24
1.18
0.40
0
1
2
3
4
5
6
Hazard ratio

Fig.3

Supporting Online Materials

As part of a broader project on university faculty members' engagement in the commercialization of academic science, we identified all members of the Scientific Advisory Boards (SABs) of biotechnology firms that have filed Securities and Exchange Commission papers to conduct an initial public offering (IPO) of stock. We included all life science firms that filed an IPO prospectus between the years 1970 and 2002. We then traced the career backgrounds of all scientific advisors that received Ph.D.s from U.S. universities and that held academic appointments at the time they joined the SAB. In particular, we identified the fields of their doctoral studies and the year in which they completed their degrees.

The information on SAB members' scientific disciplines and graduation years guided the construction of the random sample that we analyze in this paper. Specifically, we randomly drew scientists' names from UMI's *Proquest Dissertations* (described below). Names were selected in proportions that matched the distribution of Ph.D. fields and graduation years of the individuals in the SAB database. This insured that the scientists included in our dataset hail from scientific fields that are of demonstrable relevance to the commercial sector. Of course, the fact that we have chosen to focus on the subset of scientific disciplines that has been most important for the development of the biotechnology industry means that the base rate of patenting in our sample may be higher than if we had sampled scientists without regard to discipline.

The analyzed sample contains 4,227 individuals who received their doctorate degrees between 1967 and 1995. Table S4 reports, in order of their representation in the sample, the 15 scientific disciplines that appear most frequently in our data (i.e., that "contributed" the greatest number of individuals to the scientific advisory boards of recently public biotechnology firms).

Sample—Interviews

Our interviews were undertaken at one academic institution. We have chosen to keep its identity confidential given the sensitive nature of the interview material and the relatively small number of faculty involved. This institution was chosen from among the small set of universities that we identified as major contributors of SAB members to U.S. biotechnology firms. It has a mature and highly regarded

Technology Transfer Office and a high level of life science patenting. Thus, the faculty we interviewed had made many patenting decisions.

To choose individuals to interview, we identified all departments whose research broadly contributed to the life sciences. We did this by including departments in which the majority of the faculty have their Ph.D. training in the 15 scientific disciplines outlined in Table S4. Using faculty rosters, we identified all current faculty members in these departments. At the time of our interviews, there were a total of 88 male and 29 female faculty. We requested interviews with all the female faculty members via e-mail with a clear description of our research project and the goals of the interviews. Our response rate was 20/29, or 69%. We then conducted semi-structured interviews lasting between sixty and one hundred and thirty minutes. The interviews were recorded and transcribed.

We started our interviews by gathering biographical information including PhD supervisor, job history, and key academic collaborators. We then explored three broad areas. First, we asked faculty to describe their major research themes, including the genesis and the commercial relevance of their work. Second, we focused on whether the faculty member had participated in the commercialization of science, including patenting of research, consulting, SAB membership, sponsored research and start-up company formation. For each decision we discussed the source of the opportunity, any advice sought or given, and reasons for participating in or declining the commercial activity. Finally we examined overall attitudes toward commercial involvement, each individual's understanding of the role of industry in the development of her field, and the faculty member's perceptions of her colleagues' commercial opportunities.

At the end of the interview we asked each female faculty member to identify at least two male faculty who were in their cohort and whom they considered to be a peer. We then initiated a series of interviews with the "matched" sample of male faculty members. We followed a similar interview protocol except that our final discussion explored the possible interpretations of our preliminary findings that women are less engaged in commercial activities than their male colleagues.

Data Sources-Archival Analysis

We used three primary databases to assemble information about scientists and their professional achievements. First, we consulted the *UMI Proquest Dissertations* database to generate a roster of scientists' names. This database has catalogued more than 90 percent of all doctorate degrees awarded in the U.S. since 1861. Name, degree-granting institution, year of Ph.D., and discipline are available from this database. From UMI, we randomly drew names of scientists that earned degrees in the disciplines listed in Table S4.

Second, we made extensive use of the ISI's *Web of Science*, a detailed bibliometric database. The *Web of Science* includes information about most journal articles, including author(s), journal name, paper title, affiliations, abstracts, keywords, and citation information. We rely on the affiliation information gathered from this database to identify which of the scientists in the UMI sample had careers in academe, where they were employed, and when they changed employers. We also use the *Web of Science* database to create annual publication and citation counts for each scientist, as well as to identify each scientist's network of co-authors.

The third major input to the dataset we constructed for this project is the *National Bureau of Economic Research (NBER) patent database*, which contains all US patents issued between the early 1970s and 1999. The NBER data contain the patent application date (the date the patent application was submitted to the US Patent Office), the names of all of the inventors (scientists), and assignee names (the identity of the university that owns the patent). We use these data to identify if, when, and how often the academic scientists in our sample appear as inventors on patented technologies. For the period from 1999 to mid-2005, we retrieved scientists' patents directly from the U.S. Patent Office website.

One complexity that must be addressed is the matching of names in the patent database to the roster of scientists in the sample we created from *UMI* and *Web of Science*. We assumed that we had a valid name match under two scenarios. First, if there was a name match and the patent was assigned to the university with which the faculty member was affiliated, we assumed a match. Second, if there was a name match but the patent was assigned to a corporation, we included the patent in our database only when it was obviously related to a research area of the scientist (*i.e.*, the patent could be linked to a specific paper or group of papers). In all other cases, we considered name matches to be coincidental, and did not include the patent in the final dataset.

In structuring the datafile for analysis, we removed all individuals that (i) had no publications in any post-graduate year, (ii) published exclusively under corporate affiliations and (iii) exited academe within five years after receiving their Ph.D. In the statistical analysis, we assume that a scientist receives a patent at the time that the patent application arrives at the US patent office. On average, university-assigned life science patents are pending (under evaluation) for more than two years before the Patent Office determines whether to grant or deny an application. Following the convention in the empirical literature on patenting, we date patents to the year the application was submitted. We collected patents through early 2005, but it was necessary to censor the empirical analysis at the end of 2002 because of the lag between the dates of patent application and issuance.

Analysis

The article reports continuous time Cox proportional hazards regressions of the rate of patenting. Because we record patenting as taking place during the *year* in which a scientist's employing university originally filed for the relevant patent application, the outcome variable we model registers in discrete time intervals. The practical implication of this for the estimation routine is that we have a number of "tied" event times. The ties are handled with Breslow's approximation method. The Breslow method has been shown to be unproblematic as long as the ratio of events-to-observations-at-risk is small in all time intervals in the dataset. This is the case in our dataset. To verify that the findings are unaffected by our decision to employ a proportional hazards models and use a continuous time estimator, we have refit the regressions using discrete time panel logistic regression. In addition, for robustness, we have estimated a split population survival model. We used a logistic specification for the cure probability and a complementary log-log specification for the discrete-time proportional hazard of patent application. In neither case did we find a discrepancy with the results that are reported.

In unreported analyses, we have also incorporated many additional covariates in the regressions. These range from additional, time-changing attributes of scientists' professional achievements, such as cumulative NIH grant totals and cumulative citation counts, to characteristics of scientists' work contexts, such as whether they are employed in professional schools. None of these additional covariates produced a meaningful change in the estimated size of the gender gap in patenting.

In the regressions reported in the article, we treat patenting as a repeatable event. In unreported regressions, we have removed scientists from the risk set at the time they receive their first patent. In other words, we have run models of the hazard of the onset of first patenting. In these unreported regressions, the parameter estimate for the gender differences is =--0.728, implying a hazard ratio of 0.483=exp(-0.728). This implies a slightly smaller gender gap than that reported in the paper, a hazard ratio of 0.40. The difference is attributable to the fact that women are not just less likely to patent; among those who do patent, women are far less likely to patent multiple times throughout their careers.

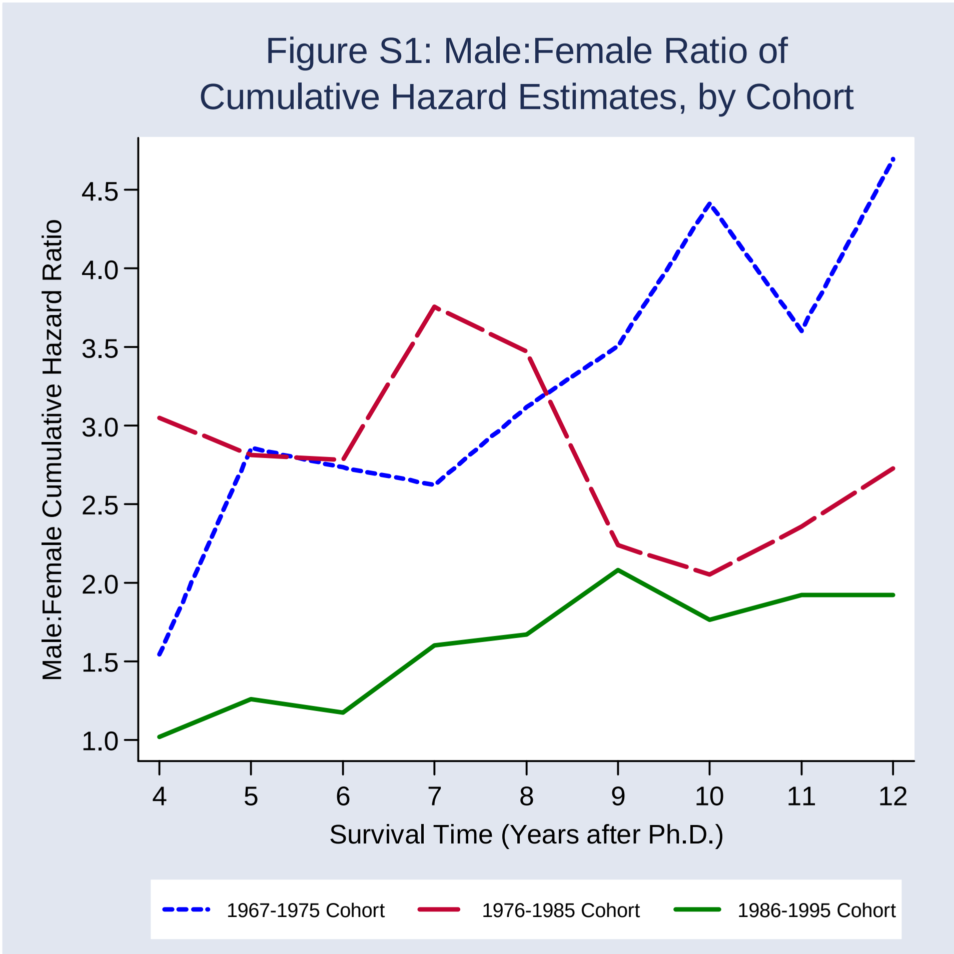
Results—Additional Tables and Interpretation

The supplement contains three tables and one figure with additional data analysis. Table S1 is paired with Figure 2 in the article. This table reports unpaired *t*-statistics for comparison of means tests across categories of scientists (male patenters, male non-patenters, female patenters, female non-patenters) at different levels of professional experience for three measures of scientific achievement. For example, the first row of the table indicates that Male Patenters (MP) have statistically higher five-year publication totals than do Female Patenters (FP) until the 13th year of professional experience. Beyond this point, the difference between male and female patenters becomes statistically insignificant, in large part due to the paucity of data points in the Female Patenter category at senior levels of professional experience. The top row in section two of the table shows the Male Patenters have statistically higher NIH grant totals until the 16th year of professional experience. Although Fig. 2B shows that FP>MP for the final few years of professional experience, Table S1 shows that this difference is not statistically significant. In fact, it is driven by a single outlier: one woman in our sample received a very large grant in the 19th year since earning her Ph.D. Because relatively few women patent, this single individual's high grant total significantly shifted the group mean.

Table S2 reports the results of three Cox regression models of the hazard of patenting. Column 1 reports estimates for background coefficients and col. 2 incorporates the *gender is female* dummy variable. The results that are graphically depicted in fig. 3 in the text correspond to the coefficients in col. 2 of the table. The final regression (col. 3) includes an interaction effect between scientist is female and the average number of coauthors per paper. This interaction effect is positive and statistically significant. Although average number of coauthors per paper is only a rough proxy for the reach of a scientist's network, this finding suggests that the existence of a broad collaboration network matters much more for women's transition to patenting than it does for men. This result is consistent with the data from our interviews: a number of the female faculty that patented reported being brought into the process by their male coauthors. While male faculty discussed the use of an advice network (including industry contacts) in decisions about patenting, they were much less dependent than women on the actions of their close academic coauthors.

One possible interpretation of the gender gap is that it simply reflects sex-based differences in the type or quality of scientists' research. In particular, if women faculty systematically tackle less challenging or important scientific questions, it is possible that their research is less likely to be eligible for patent protection and less likely to be of interest to potential patent licensees. To address this possibility, we examine gender differences in the scientific impact of research using citations in the scientific literature. As we discussed in the article, our logic for looking at citations is that if women are systematically undertaking less risky research projects, the scientific impact of women's research would be less than that of men. Table S3 breaks down the average citation counts and journal impact factors by publication year intervals and reports the *t*-tests of gender differences in each period. Overall, there is no evidence that women were doing less significant (or more incremental) work based on the scholarly impact data.

Finally, fig. S1 illustrates gender differences in patenting rates across successive cohorts of scientists. To construct the figure, we (arbitrarily) divided scientists into three groups based on the year they earned their doctorates (1967-1975, 1976-1985, and 1986-1995). We then computed gender- and cohort-specific Nelson-Aalen cumulative hazard curves and plotted the male-to-female ratio of the curves across the early years of professional experience. (Due to right censoring in the patent data, we have reliable information only for the first decade or so of the careers of the 1986-1995 cohort. As a result, we plot the data only through the 12th year of professional experience.) The figure illustrates that the gender difference in patenting propensities has declined across successive cohorts. The greatest gap is between the most senior (pre-1975) Ph.D.s and the subsequent cohort.



Caption: For each of three Ph.D. cohorts, at each professional age (survival time), we divide the Nelson-Aalen estimate of the cumulative hazard of first patenting for male scientists by that for female scientists. The ratios for all three cohorts are plotted until the 12th year of professional experience.

Table S1: *t*-Statistics for Comparisons of Research Characteristics Across Gender and Patenting Status, by Year of Professional Experience

	1	4	7	10	13	16	19
5-year publication count							
MP > FP	3.646^{**}	3.527^{**}	2.510^{**}	2.302^{*}	1.773^{*}	1.103	0.358
MNP > FNP	3.310^{**}	4.175^{**}	4.938^{**}	3.667^{**}	2.483^{**}	1.857^{*}	0.449
MP > MNP	2.403^{**}	4.361^{**}	5.434^{**}	6.377^{**}	6.217^{**}	6.036^{**}	5.256^{**}
FP > FNP	-0.863	0.459	2.309^{*}	$1.384^{\dagger}$	$1.374^{\dagger}$	1.836^{*}	1.796^{*}
FP > MNP	-2.499	-1.321	0.380	0.212	0.592	1.258	1.721^{*}
5-year NIH grant							
MP > FP	2.089^{*}	2.824^{**}	1.861^{*}	2.379^{**}	2.260^{*}	1.919^{*}	-0.333
MNP > FNP	-0.416	0.842	1.943^{*}	1.946^{*}	2.627^{**}	2.797^{*}	1.653^{*}
MP > MNP	$1.373^{\dagger}$	3.455^{**}	4.735^{**}	5.032^{**}	4.633^{**}	3.786^{**}	3.498^{**}
FP > FNP	-1.743	-0.172	1.131	1.121	0.750	0.609	$1.568^{\dagger}$
FP > MNP	-2.906	-0.534	0.518	0.457	0.071	-0.056	1.278
Count of collaborative ties with industry researchers							
MP > FP	1.000	0.816	0.087	-0.286	-0.141	0.391^{*}	0.022
MNP > FNP	1.261	0.617	$1.636^{\dagger}$	1.860^{*}	0.908	$1.606^{\dagger}$	0.769
MP > MNP	-1.579	0.854	1.129	0.783	1.263	2.234^{*}	2.128^{*}
FP > FNP	-1.579	0.854	1.129	0.783	1.263	2.234^{*}	2.128^{*}
FP > MNP	-3.932	-0.422	0.448	0.663	0.722	0.629	0.677

Caption: reports statistics of two sample unpaired *t* tests of equality in mean research characteristics across four groups of scientists, male patenters (MP), female patenters (FP), male non-paternters (MNP) and female non-patenters (FNP), at seven professional ages (i.e., years after Ph.D.). † indicates equality in mean characteristics across the two groups can be rejected at 90% confidence level, *at 95% confidence level and ** at 99% confidence level.

Table S2
Cox Proportional Hazards Models of Patenting

	(1)			(2)			(3)		
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)
	Coefficient	Standardized Hazard Ratio	Robust S.E.	Coefficient	Standardized Hazard Ratio	Robust S.E.	Coefficient	Standardized Hazard Ratio	Robust S.E.
Field – Molecular Biology	1.017	2.765	0.410*	0.988	2.687	0.406*	0.990	2.692	0.408*
Field – Health Sciences, Immunology	0.732	2.080	0.357*	0.734	2.083	0.355*	0.736	2.087	0.355*
Field – Organic Chemistry	0.753	2.123§	0.564**	0.687	1.989§	0.571**	0.653	1.921§	0.567**
5-year Publication Count	0.057	1.850‡	0.012**	0.055	1.806‡	0.011**	0.054	1.785‡	0.011**
5-year Publication Count^2	-0.0005	---	0.0002**	-0.0005	---	0.0002**	-0.0005	---	0.0002**
Average Number of Coauthors	0.133	1.160	0.040**	0.145	1.176	0.037**	0.097	1.115	0.046*
Coauthorships with Company Scientists	0.281	1.324	0.122*	0.259	1.295	0.122*	0.269	1.308	0.122*
Univ 5-year patent count (divided by 10)	0.037	1.223§	0.017*	0.039	1.237§	0.018*	0.039	1.237§	0.019*
Univ 5-year patent count (divided by 10) × ln (time)	-0.010	---	0.007	-0..011	---	0.007	-0..010	---	0.007
Female				-0.915	0.400	0.168**	-1.463	0.231	0.219**
Female × Avg Number of Coauthors							0.253	---	0.069**
Log pseudo-likelihood		-7537.9			-7498.7			-7491.7	
Wald χ-square		529.8			605.8			625.9	
Model d.f.		46			47			48	

Caption:
* All models include unreported calendar year dummies. Ph.D. field dummies are included in the regressions; only fields that have effects that are statistically different from the omitted category (biology, genetics) are reported.
† Time at risk = 71,664, number of scientists = 4,227, number of (repeated) patent application events = 1028 (Male = 946, Female =82).
‡ In columns (b) and (d), *Standardized hazard ratio* = exp (coefficient × standard deviation); for dummy variables, it equals exp (coefficient). The estimated hazard ratio of 5-year publication count is a standard deviation increase in publication count based on the joint effect of the linear and quadratic terms.
§ To correct for non-proportionality in four field dummies, we included interactions between ln(time) and Ph.D field dummies for "Pharmacology", "Organic chemistry", "Pathology", and the "Other" category. The coefficient and standardized hazard ratio of "Organic chemistry" are based on the joint effect of the main and the interaction term with log of the mean of clock (t=11.3). The standardized hazard ratio of "University 5-year patent count" are based on the joint effect of one standard deviation change in the covariate and its interaction term with log of the mean of the clock.
|| All time-varying covariates are lagged by one year.
¶ Standard errors (columns c and f) are clustered on individual scientists; †significant at 10%; *significant at 5%; **significant at 1%

Table S3

Comparison of Mean Research Article Impact by Gender

Publication Period	Citation Count			Average Journal Impact Factor		
	Male	Female	*t*	Male	Female	*t*
1967-1975	44.733	41.396	0.879	4.062	4.116	-0.275
1976-1980	45.446	49.483	-0.925	4.102	4.267	-1.351
1981-1985	48.471	49.918	-0.422	4.158	4.334	-1.850†
1986-1990	48.817	44.215	1.310	4.049	4.144	-1.185
1991-1995	36.765	34.222	1.829†	4.092	4.124	-0.422
1996-2002	12.784	13.261	-1.043	3.994	3.981	0.221
Total	33.767	32.949	0.834	4.061	4.123	-1.751†

Caption: mean citation counts and journal impact factors of 23,436 research papers published by female scientists in our data and a 1:1 random, matched (by year of publication) sample of papers published by men. Two-sample, unpaired *t*-test statistics (assuming unequal variance) are reported. † indicates that equality in mean values between male and female can be rejected at 90% confidence level.

Table S4

Top 15 Scientific Disciplines in the Sample

UMI Subject Code	UMI Subject Description	Frequency	
487; 303	Biochemistry	972	(22.8%)
306	Biology, General	611	(14.3%)
410	Biology, Microbiology	513	(12.0%)
419	Health Sciences, Pharmacology	248	(5.8%)
490	Chemistry, Organic	233	(5.5%)
786	Biophysics, General	231	(5.4%)
369	Biology, Genetics	221	(5.2%)
982	Health Sciences, Immunology	186	(4.4%)
433	Biology, Animal Physiology	185	(4.1%)
307	Biology, Molecular	114	(2.7%)
301	Bacteriology	67	(1.6%)
287	Biology, Anatomy	60	(1.4%)
571	Health Sciences, Pathology	53	(1.2%)
542	Engineering, Chemical	34	(0.8%)
572	Health Sciences, Pharmacy	34	(0.8%)

Caption: Reports the top 15 disciplines in our sample, including the number and proportion of scientists in the sample in each discipline.