Jointly published by Akadémiai Kiadó, Budapest and Springer, Dordrecht

Scientometrics
DOI: 10.1007/s11192-008-2101-0

Gender differences in patenting activity: An examination of the US biotechnology industry

G. STEVEN MCMILLAN

Penn State Abington, 1600 Woodland Road, Abington, PA 19001, USA

The gender gap in science and technology has received considerable attention by both researchers and policy makers. In an effort to better understand the quantity, quality, and underlying characteristics of female research efforts, I integrate three existing databases to uncover how female patenting activities differ from men's in the US biotechnology industry. Data on how much science the patents build upon, the author institutions of that science, and who funded the papers in which the science appears are all examined. In addition, using the NBER Patent Citation Data Files, I propose a possible gender-based life cycle model for patenting activity. The policy implications of my findings are also discussed.

Introduction

The gender gap in both academic and commercial 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 [DING & AL., 2006]. In addition, MURRAY & GRAHAM [2007] found that female academic scientists continue to lag their male counterparts in engaging in commercial science, i.e., patenting. However, some studies have found that, at least in the academic sciences, the gender gap has narrowed [NATIONAL SCIENCE FOUNDATION, 2004]. Other recent research has found that in both the academic and commercial settings, female scientists engage in and produce less commercial work than their male counterparts, and that the degree of disparity remains constant over time [BUNKER WHITTINGTON & SMITH-DOERR, 2005]. However, the quality and impact of the women's work remains the same or better than that of men scientists. Yet, to date, no study has examined the biotechnology industry in depth to uncover not only the patenting activity by gender, but also many of the underlying characteristics of those patents including the basic science upon which these patents build, and it is in this area that this study hopes to make a contribution.

Previous research has found that the biotechnology sector is a particularly attractive employment option for female scientists [EATON, 1999]. One reason is that women as a group have a very difficult time achieving employment equity in academic settings

Received June 16, 2008

Address for correspondence:
G. STEVEN MCMILLAN
E-mail: gsm5@psu.edu

0138–9130/US $ 20.00

though some strides have been made. EATON [1999] identified a number of aspects of the work structure in biotechnology that help create surprising opportunities for women. These include: academic science is a particularly difficult career option for women, making biotechnology look good in comparison; the way work itself is structured and scheduled in this industry seems to create potential flexibility in scheduling and work organization, which is especially important to working women in two-career families; and the structure of project management and other supervisory work creates opportunities for women to gain management skills and to earn promotions to managers' jobs.

Yet, given that the biotechnology industry appears to such an attractive employment option, little work has been done to examine how well women are succeeding in the research and development activities that are so important in that sector. In an effort to do so, I will integrate three existing databases to examine this subject.

Previous research

There have been substantial research efforts made in trying to better understand how female scientists are faring in commercial activities (patenting), however, much of this work has focused on female scientists in academia, where with the passage in 1980 of the Bayh-Dole Act in the US, commercial involvement has become much more pronounced and important in US academic institutions [DING & AL., 2006; BUNKER WHITTINGTON & SMITH-DOERR, 2005]. Also, one of the most comprehensive projects was conducted by NALDI & VANNINI PARENTI [2002] for the European Commission – DG Research aimed at assessing the feasibility of producing patent and bibliometric indicators by the gender of the inventor/author, though they did not segregate inventors/authors by academia or industry. Nevertheless, there has been research done focusing on female scientists employed in industry.

BUNKER WHITTINGTON & SMITH-DOER [2005] examined the patenting results for over 1,000 life scientists. They found that male scientists in industry patented more than females, but that, generally, the disparity was substantially less than that in academia. In addition, in industry, the difference in the quality of the patents, using two measures developed for the National Bureau of Economic Research (NBER) database [HALL & AL., 2001], were quite minimal.

Research methodology

By making use of a database developed by CHI Research in Haddon Heights, NJ, this research effort seeks to uncover some of the characteristics that underlie the patenting efforts, by gender, of scientists employed in the US biotechnology industry.

Some of these characteristics include: their reliance on public science (science that originates in universities, research institutes and other public institutions) versus private science, their reliance on basic versus applied science, and the funding sources for that public science.

Database #1

The following describes a database previously utilized for papers published in the journals *Research Policy* [MCMILLAN & AL., 2000] and *Scientometrics* [MCMILLAN & HAMILTON, 2007] . To develop the database, we requested the initial public offering (IPO) prospectuses of the 220 US biotechnology companies that were publicly traded as of 1993; 119 responded and formed our sample. Our examination in that study focused on their reliance on public science. Most of the biotechs had few, if any, pre-IPO patents so we focused on patents granted post-IPO. The 119 companies in our sample were granted 2,334 patents from their respective IPO dates though 1997; some of the larger ones were Chiron (408 since 1982), Mycogen (138 since 1986 and Genetic Institute (123 since 1986).

In our sample, the total numbers of patent-to-patent references were 10,335 and the number of non-patent references was 23,286. These non patent references (NPRs) yielded 20,752 science citations representing 12,477 scientific papers (papers could be cited more than once). Of these citations, we were not able to classify 2,745, usually because a reference might be incomplete; an author's name misspelled, or a wrong volume or page number given. The remaining 18,007 citations were matched with the Science Citation Index-based Science Literature Indicators Database (SLID) maintained at CHI Research for the National Science Foundation. The SLID is a comprehensive database containing information on article authors, institutions, level of journal, and, perhaps most importantly, funding sources.

This biotechnology database has been recently extended [MCMILLAN & HAMILTON, 2007] using the NBER Patent Citations Data File. This database contains all the patent citations (some 16 million) to over three million US utility patents. Using the NBER Patent Citations Data File, we weighted the patents based on their citation rates to arrive at a measure of their quality [HALL & AL., 2005].

Database #2

Patent and bibliographic databases do not contain coding on the gender of inventors and authors. Previous research has sought to overcome this problem by utilizing the first names of authors and inventors to sort the data by gender [NALDI & AL., 2004].

Geoff Peters of Simon Fraser University has developed a database that determines gender based on first names. It is called "First Name Gender Disambiguation" (FNGD)

and is available for free on the Web (http://www.gpeters.com/name-gender/). For example, the name John had a 99% likelihood of being male, while Emily had the same likelihood of being female. Names that are gender-neutral are assigned a "gender ratio" based on the algorithm that Peters has developed, and this ratio indicates the likelihood of a particular gender being represented. This database was used to determine the gender for the inventors noted in Database #1.

Database #3

The final database employed was the NBER Patent Citation Data Files. Two of its measures that I utilize in my analysis are generality and originality. Generality is measured by examining the technology of "forward" citations with the intuitive assumption being that future patents from a broad variety of technological classes have higher generality. Originality is a similar construct that explores the technological classes of "backward", cited patents with the intuition being that a patent that cites across a wide range of fields will have higher originality [BUNKER WHITTINGTON & SMITH-DOERR, 2005].

Results

Of the original 2,334 patents, I was able to classify 1, 903 by gender. The difference was generally due to foreign names that the FNGD database did not include. Table 1 provides a breakdown by gender of those classified patents.

Obviously, it is quite clear that men dominate the patenting activity with only 4% of the patents being discovered by women alone. Table 2 illustrates even more clearly the dominance of men in this activity.

Out of over 5,000 total inventors, less than 900 were women. However, as noted in Table 3, on patents that include both women and men, women fare rather well with their representation being over one-third.

As I move to Table 4, I believe that I uncover one of the most interesting initial findings.

On patents that include only female inventors, the average number of inventors is less than two-thirds of the average for men only, and less than half of that for patents that include both genders. This suggests that when women chose to research only with other women, that it is a small, perhaps close-knit group. Also, given that patents with both genders have such a dramatically higher average inventor indicator, it seems clear when working across genders, many more inventors are brought into the inventive process. Possible reasons for this finding will be discussed later.

Table 1. Total patents segregated by gender

Female only	Both genders	Male only	Total
74	612	1,217	1,903
4%	32%	64%	100%

Table 2. Total number of inventors identified in the 1,903 patents by gender

Female	Male	Total
897	4,259	5,156
17%	83%	100%

Table 3. Total number of inventors for 612 "Both" Patents by Gender

Female	Male	Total
794	1,419	2,213
36%	64%	100%

Table 4. Average number of inventors per patent

Female	Both	Male
1.53	3.68	2.4

My next level of analysis investigates the scientific publications that are cited in the patents under review. Table 5 shows that the reliance on public science in those publications is relatively constant across all three groups.

However, when I move to the field and level of journal in which the articles appeared, we begin to find some interesting differences.

Women cite articles that appear in biomedical journals substantially more than men, while they cite chemistry journals substantially less. Linked to this finding is that women cite journals that are much more basic than either men only or patents that include both genders by a substantial amount.[1] This suggests that women rely on research that is much more cutting edge than the other two groups.

The next level of analysis focuses on the funding sources that were acknowledged in the papers that were cited.

Not surprisingly, the various agencies associated with the National Institutes of Health are the primary funding sources across genders. However, the most interesting finding centers on the two funding sources for papers that the female inventors use much more than either both genders or just males. The Department of Defense and the National Institute of Mental Health (NIMH) make the top ten only for women. For the NIMH, it ranks 23rd for men only and 33rd for both genders. This suggests that research focusing on mental health issues figures much more prominently in the inventive activities for women versus men.

[1] The National Science Foundation and CHI Research jointly developed a classification scheme for the journals covered by ISI's Science Citation Index [PINKSI & NARIN, 1976; NARIN & AL., 1976]. Their classification scale ranges from 1 to 4 with 1 being the most applied journal and 4 being the most basic.

Table 5. Reliance on public or private science

Author institution type	Female	Both	Male
Public	69%	70%	69%
Public/Private	10%	11	11%
Private	21%	19%	20%
Totals	100%	100%	100%

Table 6. Field of journals

Field	Female	Both	Male
Biomedical Research	79%	67%	60%
Clinical Medicine	17%	30%	30%
Biology	3%	2%	2%
Chemistry	1%	1%	8%
Totals	100%	100%	100%

Table 7. Level of journal

Level	Female	Both	Male
4	81%	69%	68%
3	17%	22%	23%
2	2%	6%	6%
1	0%	3%	3%
Totals	100%	100%	100%

Table 8. Top ten identified funding sources (number of papers)*

Funding Source	Female	Both	Male
National Cancer Institute	71	3,400	4,056
Nat'l Inst Allergy & Infectious Dis	54	2,267	2,051
Nat'l Inst of Gen'l Med Sciences	46	2,098	2,954
National Science Foundation	20	566	900
Nat'l Inst of Health (gen)	20	593	737
American Cancer Society	16	940	1,135
Nat'l Heart, Lung & Blood Inst	16	1,096	1,466
Department of Defense	8	**	**
National Institute of Mental Health	6	**	**
Nat'l Inst Arthr & Muscloski	6	1,107	798

* Female column is sorted in descending order
** Not in the Top Ten

Next I move to an analysis of the author institutions for the papers that were cited in the patents.

Similar to the funding sources, there are substantial similarities between the three groups regarding the prominence of the University of Washington Medical School, Harvard Medical School, and the National Cancer Institute. However, I believe the finding for the Hoffman La Rouche Institute is particularly interesting. It ranks fourth for females only and tenth for both genders, but does not appear on the list for men only. This seems to indicate that much of the research done at the Hoffman La Rouche Institute is of particular interest to women.

The final level of analysis utilizes the NBER Patent Citations Data File, the third database.[2]

Table 9. Top ten identified author institutions (number of papers)

Female Only

UNIV WASHINGTON, MED	129
HARVARD UNIV, MED	74
NCI	51
HOFFMANN LA ROCHE IN	47
USDA, ARS	35
UNIV CALIF RIVERSIDE	30
IMMUNEX CORP	30
MASS GEN HOSP	27
OAK RIDGE NATL LAB	26
NEW ENGLAND BIOLABS	26

Both

HARVARD UNIV, MED	1242
UNIV WASHINGTON, MED	884
U CAL SAN FRANC, MED	872
MIT	791
GENETICS INST	456
MASS GEN HOSP	428
U CAL BERKELEY	425
STANFORD UNIV, MED	387
NCI	350
HOFFMANN LA ROCHE IN	285

Men Only

HARVARD UNIV, MED	1570
UNIV WASHINGTON, MED	1202
NCI	1123
MASS GEN HOSP	770
STANFORD UNIV, MED	758
U CAL SAN FRANC, MED	588
VIRGINIA MASON CLIN	569
GENENTECH INC	536
UNIV WASHINGTON	528
COLUMBIA UNIV COLL	480

Table 10. Patent characteristics from the NBER database

Characteristics	Female	Both	Male
Granting Year	1992.19	1993.08	1992.77
Citations Made	5.07	6.86	7.39
Citations Received	9.89	8.88	8.37
Generality	0.363	0.365	0.364
Originality	0.361	0.395	0.384

[2] I recognize the potential "right side" bias in my analysis, i.e., that patents that were granted later in the observation period would receive fewer citations. The NBER database ended with the year 1999, and an examination of my patents indicates that all of them were granted prior to 1998. In fact, 70% were granted prior to 1996. I would suggest that this ameliorates the possible bias.

The first finding of interest is the average granting year of the patents. While the averages between men and women are almost equal, the average granting year for the both genders group is almost a year later. I would argue that this indicates that early in my observation period women worked with women and men with men, but that later the two genders worked more collaboratively.

In the citations received and citations made data, the findings across genders are quite different. Men cite previous patents much more than women, yet they are cited less. Since citations received are one important measure of quality (for an extensive discussion on how citations are related to quality, see [HALL & AL., 2005]), it appears that while women may patent much less than men, the quality of their patents is higher.

In a final attempt to measure quality, I examine how the female inventors fare on generality and originality. My findings are similar to Bunker Whittington and Smith-Doerr in that there are no substantial differences across the three groups on either measure. Bunker Whittington and Smith-Doerr did find a significant relationship in generality but it was driven by a difference in academic patenting, an area I did not investigate. That citations received does not appear to correlate with generality and originality is deserving of future research.

Conclusions and implications

Having conducted an in-depth analysis of the role of gender in the US bio-technology industry, I believe there are some conclusions that can be drawn. First, it is apparent that, at least during the time under examination, there was a substantial gender gap even in an industry as attractive to women as biotechnology. Female inventors were patenting less by any measure that I used. However, the nature of the underlying science (type, funding sources, and author institutions) they cited in their patents was quite different from men and in many respects superior. In addition, using a citation methodology, I found that women's patents were valued higher than those of men or joint work between men and women. But, while the findings have some importance, the possible implications are much more.

Coupling the findings for the category "Both" on age of the patents (most recent) and number of inventors (most), I would argue that there is a life cycle that these results might elucidate. It seems possible that the biotechnology research efforts were moving from men working with men and women with women to more joint work. Future research might be conducted to see if such a trend did indeed continue.

Here I should note an obvious limitation to this study that centers on the age of the patents. It is quite likely that the biotechnology industry of 2008 is quite different from the early 1990's. However, I would argue that a better understanding of where biotechnology has been on this issue could perhaps inform today's R&D environment.

It is clear that women are certainly able, and willing, to do high-quality research that leads to important commercial outcomes. Yet their overall role in these endeavors remains, in many respects, quite modest. As high-technology companies continue to face stiff competition from both home and abroad, they must find mechanisms to deploy all of their scientific assets in the most efficient and effective ways as possible. It seems clear that assuring that female inventors figure prominently in their corporate research activities would be quite important. And while biotechnology companies have shown great desire to hire female scientists and provide them with the requisite resources to do high-caliber research, companies must make sure that women participate fully in one of the final aspects of the R&D process, acquiring a patent. In doing so, firms will quite likely reap many more benefits from their research activities.

References

BUNKER WHITTINGTON, K., SMITH-DOERR, L. (2005), Gender and commercial science: women's patenting in the life sciences, *Journal of Technology Transfer*, 30 : 355–370.

DING, W. W., MURRAY, F., STUART, T. (2006), Gender differences in patenting in the academic life sciences, *Science*, 313: 665–667.

EATON, S. C. (1999), Surprising opportunities: gender and the structure of work in biotechnology firms, *Annals New York Academy of Sciences,* 175–188.

HALL, B. JAFFE, A., TRAJTENBERG, M. (2001), The NBER patent citations data file: lessons, insights and methodological tools, *NBER Working Paper, 8498*.

HALL, B. JAFFE, A., TRAJTENBERG, M. (2005), Market value and patent citations, *RAND Journal of Economics*, 36 : 16–38.

MCMILLAN, G. S., NARIN, F., DEEDS, D. (2000), An analysis of the critical role of public science in innovation: The case of biotechnology, *Research Policy,* 29 : 1–8.

MCMILLAN, G. S., HAMILTON, R. (2007), The public science base of us biotechnology: A citation-weighted approach, *Scientometrics*, 72 : 3–10.

MURRAY, F., GRAHAM, L. (2007), Buying science and selling science: gender differences in the market for commercial science, *Industrial and Corporate Change*, 16 : 657–689.

NALDI, F, VANNINI PARENTI, I., *Scientific and Technological Performance by Gender- A Feasibility Study on Patents and Bibliometric Indicators*. Luxembourg: Office for Official Publications of the European Communities, 2002.

NALDI, F., LUZI, D., VALENTE, A., PARENTI, H. V. (2004), Scientific and technological performance by gender, In: H. F. MOED et al. (Eds), *Handbook of Quantitative Science and Technology Research,* Kluwer Academic Publishers, pp. 299–314.

NARIN, F., PINSKI, G., GEE, H. (1976), Structure of the biomedical literature, *Journal of the American Society for Information Science*, 27 : 25–45.

NATIONAL SCIENCE FOUNDATION (2004), *Gender Differences in the Careers of Academic Scientists and Engineers*, NSF 04-323. (www.nsf.gov/sbe/srs/ns04323/pdf/front.pdf)

PINSKI, G., NARIN, F. (1976), Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics, *Information Processing and Management*, 12 : 297–312.