10-2014

Gender Diversity in the Patent Bar

Saurabh Vishnubhakat
Texas A&M University School of Law, sv10@law.tamu.edu

THE JOHN MARSHALL
REVIEW OF INTELLECTUAL PROPERTY LAW

GENDER DIVERSITY IN THE PATENT BAR

SAURABH VISHNUBHAKAT

ABSTRACT

This article describes the state of gender diversity across technology and geography within the U.S. patent bar. The findings rely on a new gender-matched dataset, the first public dataset of its kind, not only of all attorneys and agents registered to practice before the United States Patent and Trademark Office, but also of attorneys and agents on patents granted by the USPTO. To enable follow-on research, the article describes all data and methodology and offers suggestions for refinement. This study is timely in view of renewed interest about the participation of women in the U.S. innovation ecosystem, notably the provision of the Leahy-Smith America Invents Act directing the USPTO to study diversity, including gender diversity, among patent applicants, and of related research by the National Women's Business Council on usage of the U.S. patent and trademark systems by U.S.-based female entrepreneurs. Analysis of gender data on the patent bar complements these studies and begins to provide a more complete picture of diversity in the U.S. patent system.

Cite as Saurabh Vishnubhakat, *Gender Diversity in the Patent* Bar, 14 J. MARSHALL REV. INTELL. PROP. L. 67 (2014).

Gender Diversity in the Patent Bar

Saurabh Vishnubhakat

GENDER DIVERSITY IN THE PATENT BAR

SAURABH VISHNUBHAKAT*

I. INTRODUCTION

This article describes the first public dataset gender-matching the attorneys and agents who are registered to practice before the United States Patent and Trademark Office. Significantly, all the underlying sources and methodologies used in developing this dataset are also from public sources. As a result, the contribution of this study is not merely a new and useful dataset to support empirical research into diversity within the U.S. patent system, but also a mechanism for ensuring that the dataset may be readily updated and more easily tailored to fit particular research needs.

Much current research on the participation of women in the intellectual property system has been doctrinal, focusing on the intersections of feminist theory with patent and copyright law and with intellectual property law more generally.[1] While this literature advances a conceptual framework and normative proposals,[2] empirical work in this area has been sparse.

This article invites quantitative research to fill this gap by gender-identifying USPTO practitioners generally and practitioners of record on granted patents specifically, publishing the dataset as well as the methodology, and drawing more detailed comparisons across geography and technology. By examination of these dimensions in particular, discussions of gender diversity in the patent bar will be more fully connected with the broader economic discourse on the geographical and technological specificity of innovation and entrepreneurship.

II. RELEVANT DIMENSIONS OF GENDER DIVERSITY

In describing gender diversity in the intellectual property system, the relevance of geography and technology is apparent in the broader empirical literature on diversity. Moreover, the importance of these dimensions to the patent bar proceeds from several upstream arenas of knowledge and skills development including

---

* © Saurabh Vishnubhakat 2014. Postdoctoral Associate, Duke University School of Law; Expert Advisor, United States Patent and Trademark Office. The arguments in this writing are the author's and should not be imputed to the USPTO or to any other organization. Sincere thanks to Michael Carley, Bryan Choi, Christine Haight Farley, John Golden, Stuart Graham, Jay Kesan, Alan Marco, Malwina Mejer, and Laura Pedraza-Fariña for thoughtful comments and suggestions. This research was presented at the 2014 Works in Progress IP Colloquium.

1 The American University Symposia, *IP/Gender: The Unmapped Connections*, represent a rich body of literature on this intersection. *See, e.g.*, Carys J. Craig, *Reconstructing the Author-Self: Some Feminist Lessons for Copyright Law*, 15 AM. U. J. GENDER SOC. POL'Y & L. 207 (2007); Debora Halbert, *Feminist Interpretations of Intellectual Property*, 14 AM. U. J. GENDER SOC. POL'Y & L. 431 (2006).

2 *See, e.g.*, Dan L. Burk, *Do Patents Have Gender?*, 19 AM. U. J. GENDER, SOC. POL'Y & L. 881 (2011); Malla Pollack, *Towards a Feminist Theory of the Public Domain, or Rejecting the Gendered Scope of United States' Copyrightable and Patentable Subject Matter*, 12 WM. & MARY J. WOMEN & L. 603 (2006).

undergraduate and graduate education in science and engineering disciplines as well as legal education.

In the case of geography, scholarly discussion of diversity has often treated geography as a subject of diversification on par with characteristics of race, ethnicity, socioeconomic class, age, and gender—and done so across a range of legally salient contexts such as democratic participation,[3] administrative decision-making,[4] and corporate governance.[5] A geography of gender itself is of somewhat recent vintage, particularly with respect to regional geography.[6] Professor Townsend's seminal paper arguing for such a literature placed it within the school of "new regional geography"[7] to which studies of gender had already contributed.[8]

Quantitative detailed scholarship on legal institutions and the legal profession, however, has been sparse in response to this call for geographically segmented examinations of gender diversity. More common have been qualitative discussions of gender equity in a variety of settings such as within law firms[9] and in-house legal departments,[10] on corporate boards,[11] and at law schools both among faculty and students.[12]

A principal concern of this literature is the so-called "pipeline problem" that limits the diversity of populations of interest according to the diversity of available candidates.[13] Thus, for example, large corporations recruit laterally from large law firms and consequently tend to inherit the relative gender homogeneity of those law firms.[14] Moreover, similar to concerns of diversity in the contents of a pipeline are the

---

[3] *See, e.g.*, Justin Levitt, *Democracy on the High Wire: Citizen Commission Implementation of the Voting Rights Act*, 46 U.C. DAVIS L. REV. 1041, 1076 (2013); Bruce E. Cain, *Redistricting Commissions: A Better Political Buffer?*, 121 Yale L.J. 1808, 1825 (2012).

[4] *See, e.g.*, Reeve T. Bull, *Making the Administrative State "Safe for Democracy": A Theoretical and Practical Analysis of Citizen Participation in Agency Decisionmaking*, 65 ADMIN. L. REV. 611, 641 (2013).

[5] *See, e.g.*, Sonja S. Carlson, *"Women Directors": A Term of Art Showcasing the Need for Meaningful Gender Diversity on Corporate Boards*, 11 SEATTLE J. FOR SOC. JUST. 337, 338–39 n.11 n.11 (2012).

[6] *See generally* Janet G. Townsend, *Toward a Regional Geography of Gender*, 157 GEOGR. J. 25 (1991).

[7] *See generally* Hans Holmén, *What's New and What's Regional in the 'New Regional Geography'?*, 77 GEOGR. ANN. 47 (1995) (characterizing the school by its focus on social-theoretic structure to explain observed spatial variations, as opposed to traditionally atheoretical approaches that were thought inadequate at organizing empirical observations into a conceptual framework).

[8] *See, e.g.*, Mary Beth Pudup, *Arguments Within Regional Geography*, 12 PROG. HUM. GEOGR. 369 (1988).

[9] *See, e.g.*, Deborah L. Rhode, *From Platitudes to Priorities: Diversity and Gender Equity in Law Firms*, 24 GEO. J. LEGAL ETHICS 1041 (2011).

[10] *See, e.g.*, Eli Wald, *In-House Myths*, 2012 WIS. L. REV. 407 (2012).

[11] *See, e.g.*, Thomas Lee Hazen & Lissa Lamkin Broome, *Board Diversity and Proxy Disclosure*, 37 U. DAYTON L. REV. 39 (2011).

[12] *See, e.g.*, Meera E. Deo, *The Promise of Grutter: Diverse Interactions at the University of Michigan Law School*, 17 MICH. J. RACE & L. 63 (2011); Patrick S. Shin & Mitu Gulati, *Showcasing Diversity*, 89 N.C. L. REV. 1017 (2011).

[13] *See generally* THE EDUCATION PIPELINE TO THE PROFESSIONS: PROGRAMS THAT WORK TO INCREASE DIVERSITY (Sarah E. Redfield, ed., 2012). Notably, the major pipeline problem appears to rest, not between stepping stones within a profession, but between education and entry into the profession itself.

[14] Wald, *supra* note 10, at 456.

mechanisms by which such future constituencies may prepare themselves for entry into the profession, positions of leadership, and other indicia of meaningful diversity.[15]

A related strain in this literature has also argued for a reframing of gender diversity from an end in itself to an instrumental mechanism for otherwise desired ends such as profit-maximization and cross-cultural competence.[16] It is perhaps not surprising then, that the shift away from gender diversity as a metric of primary interest inviting further geographic cross-tabulation has led instead to separate examination of gender and geography (among other traits of interest) in service of metrics that are themselves more complex and difficult to estimate, such as institutional influence[17] and professional satisfaction.[18]

The problem thus reframed, however, has received some empirical attention with regard to gender diversity in the legal profession generally. Studies in this literature include the emergence of gendered outcomes such as two-tiered partnership tracks,[19] the gendered role of personal priorities in advancement,[20] and the antecedent prevalence of gender bias that such other findings suggest.[21]

The dimension of technology, for its part, has received even less empirical attention with regard to gender diversity than has geography. A notable exception is Annette Kahler's 2011 article on trends in education and downstream invention and patenting by women.[22] Professor Kahler's empirical analysis focused primarily on

---

[15] *See, e.g.*, Claire McCarty Kilian, *Corporate Leadership: Building Diversity into the Pipeline*, ETHNIC MINORITY LEADERSHIP COMMUNIQUE (Am. Psych. Ass'n) (Aug. 2009) (identifying a "lack of mentors and role models" as a key barrier to the advancement of women in corporate settings).

[16] *See, e.g.*, David B. Wilkins, *From "Separate Is Inherently Unequal" to "Diversity Is Good for Business": The Rise of Market-Based Diversity Arguments and the Fate of the Black Corporate Bar*, 117 HARV. L. REV. 1548. Professor Wilkins extrapolates from the relationship between market efficiency and non-discrimination on the basis of race to non-discrimination on the basis of national origin and gender as well, citing arguments before the Supreme Court in *Grutter v. Bollinger*, 539 U.S. 306 (2003):

> It is therefore not surprising that blacks have made substantial inroads in heavily regulated industries such as communications and insurance, that Asian Americans are well represented in technology companies that do substantial business with Asia, and that companies that market consumer products for use in the home often have good records for hiring and promoting women.

Id. (citing Brief for Amici Curiae 65 Leading American Businesses in Support of Respondents, Grutter (No. 02-241), Gratz (No. 02-516)).

[17] *See, e.g.*, David B. Wilkins, *Partners Without Power? A Preliminary Look at Black Partners in Corporate Law Firms*, 2 J. INST. STUDY LEGAL ETHICS 15 (1999); Fiona M. Kay & John Hagan, *Cultivating Clients in the Competition for Partnership: Gender and the Organizational of Law Firms in the 1990s*, 33 LAW & SOC'Y REV. 517 (1999).

[18] *See, e.g.*, Joan C. Williams & Veta T. Richardson, NEW MILLENNIUM, SAME GLASS CEILING?: THE IMPACT OF LAW FIRM COMPENSATION SYSTEMS ON WOMEN (2010); Nancy J. Reichman & Joyce S. Sterling, *Sticky Floors, Broken Steps, and Concrete Ceilings in Legal Careers*, 14 TEX. J. WOMEN & L. 27 (2004).

[19] *See, e.g.*, William D. Henderson, *An Empirical Study of Single-Tier Versus Two-Tier Partnerships in the Am Law 200*, 84 N.C. L. REV. 1691 (2006).

[20] *See, e.g.*, Kathleen E. Hull & Robert L. Nelson, *Assimilation, Choice, or Constraint? Testing Theories of Gender Differences in the Careers of Lawyers*, 79 SOC. FORCES 229 (2000).

[21] Justin D. Levinson & Danielle Young, *Implicit Gender Bias in the Legal Profession: An Empirical Study*, 18 DUKE J. GENDER L. & POL'Y 1 (2010).

[22] Annette I. Kahler, *Examining Exclusion in Woman-Inventor Patenting: Historical, Economic and Social Perspectives*, 19 AM. U. J. GENDER, SOC. POL'Y & L. 773 (2011).

female inventors,[23] as have two subsequent reports published in 2012 by the National Women's Business Council.[24]

Congress, too, has expressed interest in studying with greater empirical detail the participation of minorities, including women, in the patent system, requiring in the Leahy Smith America Invents Act that the USPTO "establish methods for studying the diversity of patent applicants, including those applicants who are minorities, women, or veterans."[25] The USPTO duly published its methodology in March, 2012,[26] and is currently implementing it.

This initial focus on inventors is appropriate, and points to existing investigations of gender diversity in education geared toward science, technology, engineering, and math disciplines as a whole[27] as well as toward particular disciplines such as clinical medicine,[28] mechanical engineering and materials science,[29] computer science[30] and information technology,[31] and biomedical engineering.[32]

Moreover, the empirical interest in inventors as a link between science and engineering education on the one hand and the patent system on the other also invites complementary research on issues including the diversity of the patent bar. Indeed, Professor Kahler's article briefly addressed this question with a gender-identified dataset of registered USPTO practitioners as of 2008 and, for a subsample of practitioners, drawing comparisons across region, law firm type, and educational background.[33]

To build on this work, the dataset presented here provides an updated gender-identified USPTO practitioner roster and provides the first gender-identified data on that subset of patent attorneys and agents who have successfully prosecuted patents to issuance during the last five years. The result is a comprehensive empirical platform for exploring gender diversity across geography and technology and enabling a wide range of policy and market analyses.

---

[23] *Id.* at 782–784 (analyzing recent statistics from the USPTO Patent Technology Monitoring Team); *id.* at 784–791 (comparing educational and patenting trends).

[24] NAT'L WOMEN'S BUS. COUNCIL, INTELLECTUAL PROPERTY AND WOMEN ENTREPRENEURS (2012), *available at* http://www.nwbc.gov/issues-and-research.

[25] Pub. L. No. 112-29 § 29 (Sept. 16, 2011).

[26] USPTO, DIVERSITY OF APPLICANT METHODOLOGY (Mar. 16, 2012), (*available at* http://www.uspto.gov/aia_implementation/20120316-diversity_of_applicant_methodology.pdf).

[27] *See, e.g.*, WOMEN AND MINORITIES IN SCIENCE, TECHNOLOGY, ENGINEERING AND MATHEMATICS: UPPING THE NUMBERS (Ronald J. Burke & Mary C. Mattis, eds.) (2007).

[28] *See, e.g.*, Janet Bickel, *Gender Equity in Undergraduate Medical Education: A Status Report*, 10 J. WOMEN'S HEALTH & GENDER-BASED MED. 261 (2004).

[29] *See, e.g.*, Anne-Françoise Gilbert, *Disciplinary Cultures in Mechanical Engineering and Materials Science: Gendered/Gendering Practices?*, 28 EQUAL OPPORTUNITIES INT'L 24 (2009).

[30] *See, e.g.*, Elizabeth A. Larsen & Margaret L. Stubbs, *Increasing Diversity in Computer Science: Acknowledging, Yet Moving Beyond, Gender*, 11 J. WOMEN & MINORITIES SCI. & ENGINEERING 139 (2005).

[31] *See, e.g.*, LeAnne Coder et al, *Economic and Business Dimensions: Increasing Gender Diversity in the IT Work Force*, 52 COMM. ASS'N COMPUTING MACHINERY 25 (2009).

[32] *See, e.g.*, Naomi C. Chesler et al, *The Pipeline Still Leaks and More Than You Think: A Status Report on Gender Diversity in Biomedical Engineering*, 38 ANNALS BIOMED. ENG. 1928 (2010).

[33] Kahler, *supra* note 22, at 791–792.

III. THE PATENT BAR GENDER DATA FILE

A. Data and Methodology

Source data on practitioners registered to practice before the USPTO was matched with gender data on the frequencies with which given names occur among women and men. The result is a new dataset that estimates the genders of registered practitioners. The new dataset is matched to locational information, allowing for geographic analysis. Also used is the USPTO public full text data on patent grants, available until recently through Google[34] and going forward from Reed Technology and Information Services.[35]

As to gender data, the U.S. Census Bureau has published two dictionaries tabulating the frequencies at which names occur among women and men.[36] By matching the given names of a USPTO-registered practitioner to these name dictionaries, it is possible to estimate the likelihood that the practitioner is male or female.

For example, the name Ryan occurs with a frequency of 0.328% among men, making it the 49th-most common male name. It occurs with a frequency of 0.006% among women, making it the 1229th-most common female name. The population is roughly equal as to men and women,[37] and the name Ryan occurs with higher frequency among men than among women, so it is intuitive that a practitioner named Ryan is male. Precisely, there is a 98.1% likelihood that a practitioner named Ryan is male, and a 1.90% likelihood that such a practitioner is female:

$$\begin{aligned} P(Male|Ryan) &= \frac{P(Ryan|Male) \times P(Male)}{P(Ryan|Male) \times P(Male) + P(Ryan|Female) \times P(Female)} \\ &= \frac{(0.00328) \times (0.4854)}{(0.00328) \times (0.4854) + (0.00006) \times (0.5146)} \\ &= 0.9810 \\ P(Female|Ryan) &= \frac{P(Ryan|Female) \times P(Female)}{P(Ryan|Female) \times P(Female) + P(Ryan|Male) \times P(Male)} \end{aligned}$$

[34] GOOGLE, USPTO BULK DOWNLOADS: PATENTS, *available at* http://www.google.com/googlebooks/uspto-patents.html.

[35] REED TECH IP SERVICES, USPTO DATA SETS, *available at* http://patents.reedtech.com/.

[36] CENSUS BUREAU, *Genealogy Data: Frequently Occurring Surnames from Census 1990-Names Files* (Oct. 1995), *available at* http://www.census.gov/topics/population/genealogy/data/1990_census/1990_census_namefiles.html.

[37] CENSUS BUREAU, *Documentation and Methodology for Frequently Occurring Names in the U.S.-1990* (Oct. 1995), *available at* http://www2.census.gov/topics/genealogy/1990surnames/nam_meth.txt.

The total number of records are as follow:

	Records	Percent
Female	3,184,399	51.46%
Male	3,003,954	48.54%
Total	6,188,353	100.00%

$$= \frac{(0.00006) \times (0.5146)}{(0.00006) \times (0.5146) + (0.00328) \times (0.4854)}$$
$$= 0.019$$

Thus, where a match exists between the practitioner data and the Census dictionaries, the gender of each practitioner is estimated by which gender has the greater probability of being the correct one, conditional on the given name.

By this method of matching and estimation, one may attach gender information to data regarding all registered practitioners in the USPTO active roster, as well as to data regarding practitioners named as the attorneys or agents of record on granted patents.

B. Practitioners in the Active Roster

As to data on all active practitioners, the USPTO Office of Enrollment and Discipline maintains a listing of registered attorneys and agents. The listing is both searchable online[38] and downloadable in bulk.[39] The listing, updated daily, provides the following original variables about practitioners:

Original Variables

- Last Name
- First Name
- Middle Name
- Post Nominal Suffix
- Firm Name
- Street Address
- City
- State
- Postal Code
- Phone Number
- Registration Number
- Indication of Attorney or Agent Status
- Indication of Government Employee Status[40]

In addition, individual practitioner records available through the USPTO online roster search tool indicate each practitioner's date of admission to the USPTO as an agent or attorney. Where the practitioner's status has changed, both admission dates are available.

As to geographic information, gender trends on U.S.-based patent practitioners may be examined at varying levels of abstraction. The USPTO data itself provides information on each practitioner's state, city of record, and ZIP code. The state-by-

[38] USPTO, OED Patent Attorney/Agent Search, *available at* https://oedci.uspto.gov/OEDCI/.
[39] USPTO, OED Patent Attorney/Agent ZIP File, *available at* https://oedci.uspto.gov/OEDCI/.
[40] USPTO, *supra* note 38.

state comparison is useful particularly because it may enable comparison directly with more general data on attorneys admitted to the state bars.

The city of record, however, has limited value as a basis for comparison because of variations in how practitioners report their location. For example, an Atlanta-based attorney may report her city as Decatur, GA, a suburb adjoining the eastern edge of the Atlanta city limits. The separate reporting of each such suburb subtracts from the apparent population of women and men, attorneys and agents, in the city of Atlanta as it is broadly, commonly understood. This underreporting is a source of potential skew, as women and men may be differentially located in the suburbs of major cities, and these differentials may themselves vary across the United States.

The ZIP code has limited value for a different reason. While the boundaries of a ZIP code are specific and consistent in a way that the city of record may not be, the area of a ZIP code is often smaller even than that of a suburb. As a result, reporting by ZIP code results in the division of the practitioner population into even smaller groups,[41] and these reduced sample sizes do much to diminish the statistical precision of results that may be derived from them.

It is appropriate, then, to assign each practitioner to a stable geographic unit: the metropolitan or micropolitan statistical area (MSA or μSA) where the practitioner is located. The Missouri Census Data Center provides a crosswalk between 5-digit ZIP codes and MSA definitions set forth by the White House Office of Management and Budget.[42] This allows a mapping from 2010 ZIP code definitions[43] to 2010 MSA and μSA definitions[44] with weights to estimate that fraction of a ZIP code's population which resides in each corresponding MSA or μSA. Usefully, the practitioner populations defined by MSA/μSA are less skewed toward small populations than those by city of record or ZIP code.[45]

Thus, the data was geography-matched on practitioner ZIP code and gender-matched on practitioner name, so that the final dataset of registered then additionally comprised the following constructed variables about practitioners:

<u>Constructed Variables</u>

- MSA; and
- estimated gender.

[41] About 67% of states/territories had practitioner populations over 100. However, populations over 100 accounted for only about 1.4% of cities of record. Similarly, practitioner populations over 100 accounted for fewer than 1% of ZIP codes.

[42] Missouri Census Data Center, MABLE/Geocorr12: Geographic Correspondence Engine, *available at* http://mcdc.missouri.edu/websas/geocorr12.html.

[43] *See generally* Census Bureau, ZIP Code™ Tabulation Areas (ZCTAs™), *available at* http://www.census.gov/geo/reference/zctas.html.

[44] *See* generally Census Bureau, Metropolitan and Micropolitan Statistical Areas Main, *available at* http://www.census.gov/population/metro/.

[45] Practitioner populations over 100 accounted for 7% of MSAs/μSAs, nearly five times as large an upper tier as for ZIP codes and over seven times as large as for cities of record.

C. Practitioners Named on Granted Patents

As to practitioners named on patents granted by the USPTO, source information is available as part of full text patent grant data through Google and RTIS.[46] Parsing for bibliographic information about the patent provides the following original variables about granted patents:

Original Variables

- patent number
- patent grant date
- patent class
- patent subclass
- last name of attorney or agent
- first and middle names or initials of attorney or agent
- U.S. state or foreign country associated with the first-named inventor on the patent

Using the patent technology class, this dataset was matched to the familiar Hall-Jaffe-Trajtenberg aggregate technology category and subcategory system[47] by concordance to the U.S. Patent Classification system,[48] describing each granted patent as Chemical, Computers and Communications, Drugs and Medical, Electrical and Electronic, Mechanical, or Others. As with the roster of registered practitioners, the data was then gender-matched, so that the final dataset of patent grants then additionally comprised the following constructed variables regarding granted patents:

Constructed Variables

- number of attorneys or agents of record;
- gender of each attorney or agent of record;
- Hall-Jaffe-Trajtenberg technology category; and
- Hall-Jaffe-Trajtenberg technology subcategory.

The concordance uniquely assigns categories and subcategories to 418 U.S. patent classes, leaving unassigned 10 additional U.S. patent classes that were created after the last Hall-Jaffe-Trajtenberg concordance was developed. For these remaining U.S. patent classes, Hall-Jaffe-Trajtenberg categories and subcategories were manually assigned as shown in Table 1. The dataset of patent grants was further limited to those granted during the calendar-year period of 2008–2012.

---

[46] *See supra* notes 33–34 and accompanying text.

[47] *See generally* Bronwyn H. Hall et al., *The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools* (Nat'l Bureau of Econ. Research, Working Paper No. 8498, 2001). The category and subcategory definitions as well as the concordance are available at http://www.nber.org/patents.

[48] USPTO, PATENT CLASSIFICATION INFORMATION, *available at* http://www.uspto.gov/web/patents/classification/.

IV. DISCUSSION

As of December, 2012, the USPTO roster listed 41,833 actively registered practitioners. Of these, 40,640 had a U.S. address of record. Of these 40,640 U.S.-based practitioners, a gender match was possible for 35,562. The gender of another 5,078 remains indeterminate. The following preliminary findings are limited to U.S.-based practitioners, as the dictionaries used for gender-matching are themselves derived from U.S. census data.

The disparity between women and men is higher among registered attorneys than among registered agents (Table 2 and Figure 1). Conversely, while more registered practitioners in general tend to be attorneys than agents, among registered practitioners, men tend to be attorneys rather than agents by a higher margin than women do (Table 3 and Figure 2).

Among government-employed practitioners, the disparity between women and men is higher among attorneys than among agents (Table 4 and Figure 3). However, while more government-employed practitioners in general tend to be attorneys than agents, there is no statistically significant difference between women and men as to attorney status versus agent status (Table 5 and Figure 4).

From the different geographic levels of abstraction at which gender disparity may be estimated, findings confirm expectations regarding sample size and statistical precision.

The proportion of gender-matched practitioners who are female is consistently lower than the proportion who are male, whether aggregated by state (Figure 5), by city of record (Figure 6), by zip code (Figure 7), or by metropolitan and micropolitan statistical area (Figure 8). These gender gaps are also robust to the disaggregation by attorney or agent status, both at the state level (Figure 9) and the MSA/μSA level (Figure 10).

Moreover, at each geographic level, variation among the gender proportions of similarly sized practitioner populations is sensitive to practitioner population. As Figure 5 shows, e.g., variation is greatest among states with small gender-matched practitioner populations. As that population increases, variation diminishes. This variation appears to be distributed normally around the gender proportion of the most practitioner-populous state, California. The same is also true of gender proportions at the level of city of record, zip code, and metropolitan and micropolitan statistical area (Figures 6–8).

Similarly, comparing across technologies reveals marked differences among the levels of gender diversity in subsets of the patent bar. As shown in Figure 11, for patents granted during the five-year period of 2008–2012, the proportions of gender-matched attorneys and agents of record for granted patents were highest for Drugs and Medical inventions, at over 25% women and at times nearly 30%. Second-highest were for Chemical inventions, starting at just over 25% women, but generally declining toward 20% over the five-year observation window (Figure 12). The remaining categories were almost entirely below 15% in the representation of women among attorneys and agents of record, with Mechanical inventions the lowest at under 10% for most of the observation window (Figure 16).

Figures 12–17 show in further detail the gender gaps across subcategories in the Hall-Jaffe-Trajtenberg scheme. Among Drugs and Medical inventions, the higher representation of women was due primarily to Biotechnology inventions and Drug inventions (Figure 14). The trend among Chemical inventions was a wider spread in the degree of representation among subcategories as well as greater volatility in the gender gaps of the leading subcategories: Resins; Organic Compounds; and Agriculture, Food, and Textiles (Figure 12).

For the relatively gender-disparate Electrical and Electronic inventions, female attorneys and agents were better represented in Electrical Device inventions at 20–30% (Figure 15) while the other subcategories were largely below 15%. The remaining technology categories generally reflect low representation and appear to vary only with respect to the variation among subcategories. Such variation was broader for Other inventions, with proportions at 0–20% (Figure 17), than for Computers & Communications inventions (Figure 13) and Mechanical inventions (Figure 16), both below 15%.

V. CONCLUSION

The descriptive relationships between these rates of gender diversity and related trends over time, across geography, and among technologies invites considerable further study. Of particular value may be event studies comparing gender disparity before and after changes in policy, whether in legal or regulatory regimes pertaining to gender, or in private initiatives aimed at greater inclusion of women in historically or persistently underrepresented segments of industry and academia.

To these ends, the dataset described here is amenable both to methodological refinements and to matching with related data. The principal methodological refinement already available would be to abstain from estimating gender at the level of the individual practitioner, as is currently done.[49] In place of this simple estimated count, the count may instead be assigned the probability itself. For example, in the given sample calculation[50] where there is a 98.1% likelihood that a practitioner named Ryan is male, and a 1.90% likelihood that such a practitioner is female, the data would show 0.981 men named Ryan and 0.190 women named Ryan. Though counterintuitive with respect to individuals, in the aggregate this probabilistic estimation would yield a truer estimate of gender distributions and the gender gaps that result.

Beyond methodology, related data of interest to which this dataset may be matched includes the USPTO study in progress on the diversity of USPTO applicants,[51] not only with respect to gender diversity between applicants on one hand and practitioners on the other, but also with respect to gender as one indicium of diversity among others.

Other relevant data includes gender diversity information about state bars generally, of which registered patent attorneys may be considered a subset. Such analysis, taken together with examination and admission statistics to state bars around the country,[52] would shed valuable comparative light on the general pipeline

---

[49] *See supra* § III.A (after calculating the relative probabilities of each gender conditional on a given name, the individual record is assigned the gender with the greater probability of being the correct one).

[50] *Id.*

[51] *See supra* note 25–26 and accompanying text.

[52] *See, e.g.*, NAT'L CONF. BAR EXAMINERS, PUBLICATIONS: STATISTICS, *available at* http://www.ncbex.org/publications/statistics/.

from legal education into practice and the specific pipeline from scientific and intellectual property education into patent practice.

Not least, USPTO data on examiners of record for patent grants is already available through the same public full text data from which attorney and agent information was extracted for the present dataset.[53] Detailed analysis on interactions between practitioners and examiners, as well as between practitioners and inventors, holds rich potential for studying whether and how returns from innovation are appropriated in a gendered way.

It is hoped that descriptive studies, as well as the normative arguments and recommendations that they support for crafting diversity policy in government, industry, and academia, will proceed with quantitative precision on the basis of data such as this article presents.

[53] *See supra* notes 34–35, 45 and accompanying text.

APPENDIX: TABLES AND FIGURES

Table 1. Manual Concordance of Unassigned U.S. Patent Classes

U.S. Patent Class	**Hall-Jaffe-Trajtenberg Category and Subcategory**
Class 703 (Data Processing: Structural Design, Modeling, Simulation, and Emulation)	Category 2—Subcategory 22 (Computers & Communications—Computer Hardware & Software)
Class 715 (Data Processing: Presentation Processing of Document, Operator Interface Processing, and Screen Saver Display Processing)	Category 2—Subcategory 22 (Computers & Communications—Computer Hardware & Software)
Class 716 (Computer-Aided Design and Analysis of Circuits and Semiconductor Masks)	Category 2—Subcategory 22 (Computers & Communications—Computer Hardware & Software)
Class 717 (Data Processing: Software Development, Installation, and Management)	Category 2—Subcategory 22 (Computers & Communications—Computer Hardware & Software)
Class 718 (Electrical Computers and Digital Processing Systems: Virtual Machine Task or Process Management or Task Management/Control)	Category 2—Subcategory 22 (Computers & Communications—Computer Hardware & Software)
Class 719 (Electrical Computers and Digital Processing Systems: Interprogram Communication or Interprocess Communication (IPC))	Category 2—Subcategory 22 (Computers & Communications—Computer Hardware & Software)
Class 720 (Dynamic Optical Information Storage or Retrieval)	Category 2—Subcategory 24 (Computers & Communications—Information Storage)
Class 725 (Interactive Video Distribution Systems)	Category 4—Subcategory 49 (Electrical and Electronic—Miscellaneous-Elec)
Class 726 (Information Security)	Category 2—Subcategory 22 (Computers & Communications—Computer Hardware & Software)
Class 850 (Scanning-Probe Techniques or Apparatus; Applications of Scanning-Probe Techniques, E.G., Scanning Probe Microscopy [SPM])	Category 1—Subcategory 19 (Chemical—Miscellaneous-Chemical)

Table 2. Diversity as to Gender Among Registered Patent Attorneys and Agents

	Attorney	Agent	Total
Male	72.85%	58.72%	69.39%
Female	17.20%	20.93%	18.12%
Unknown	9.94%	20.35%	12.50%
Total	100.00% (n = 30679)	100.00% (n = 9961)	100.00% (n = 40640)

Table 3. Diversity as to Attorney or Agent Status Among Male and Female Practitioners

	Male	Female	Unknown	Total
Attorney	79.26%	71.68%	60.08%	75.49%
Agent	20.74%	28.32%	39.92%	24.51%
Total	100.00% (n = 28199)	100.00% (n = 7363)	100.00% (n = 5078)	100.00% (n = 40640)

Table 4. Diversity as to Gender Among Government Employees

	Gov't-Attorney	Gov't-Agent	Gov't-Total
Male	69.12%	58.18%	66.85%
Female	24.23%	23.64%	24.11%
Unknown	6.65%	18.18%	9.04%
Total	100.00% (n = 421)	100.00% (n = 110)	100.00% (n = 531)

Table 5. Diversity as to Government Employee Status Among Male and Female Practitioners

	Male	Female	Unknown	Total
Gov't-Attorney	81.97%	79.69%	58.33%	79.28%
Gov't-Agent	18.03%	20.31%	41.67%	20.72%
Gov't-Total	100.00% (n = 355)	100.00% (n = 128)	100.00% (n = 48)	100.00% (n = 531)

Figure 1. Diversity as to Gender Among Registered Patent Attorneys and Agents

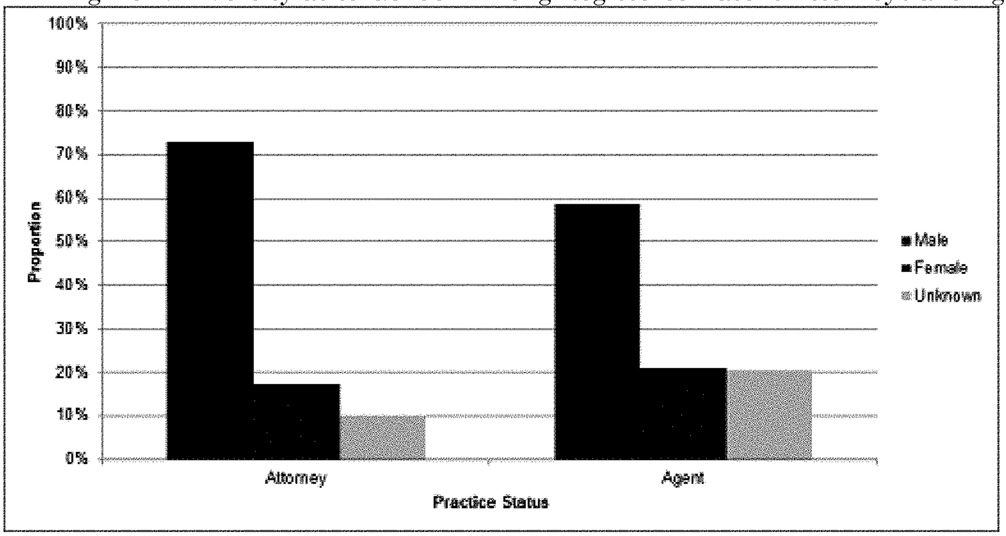

Figure 2. Diversity as to Attorney or Agent Status Among Male and Female Practitioners

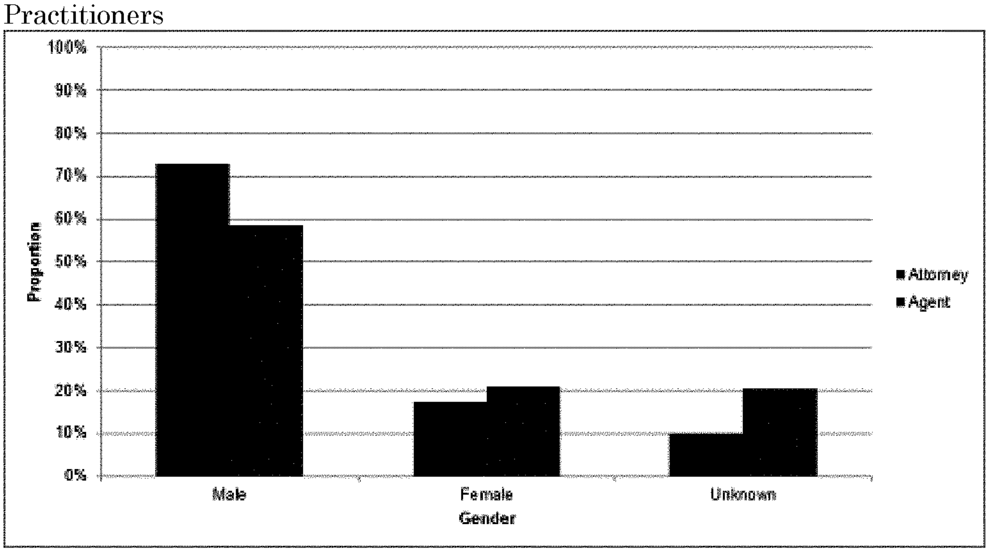

Figure 3. Diversity as to Gender Among Government Employees

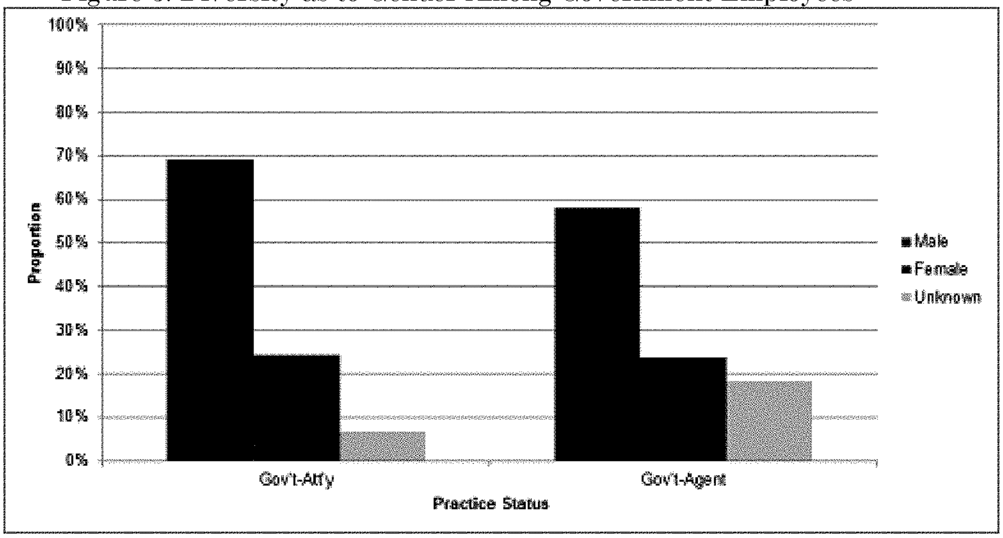

Figure 4. Diversity as to Government Employee Status Among Male and Female Practitioners

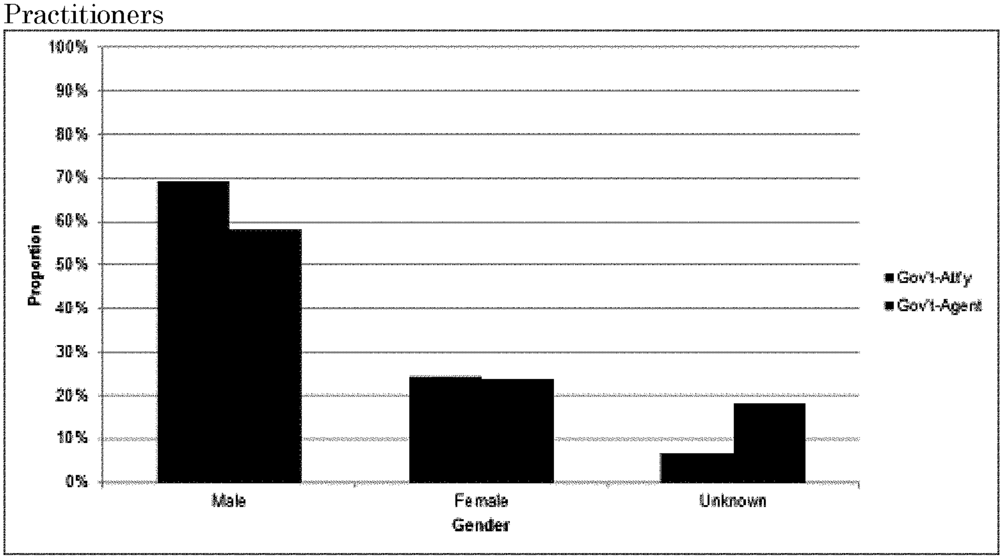

Figure 5. Gender-Matched Practitioners by State

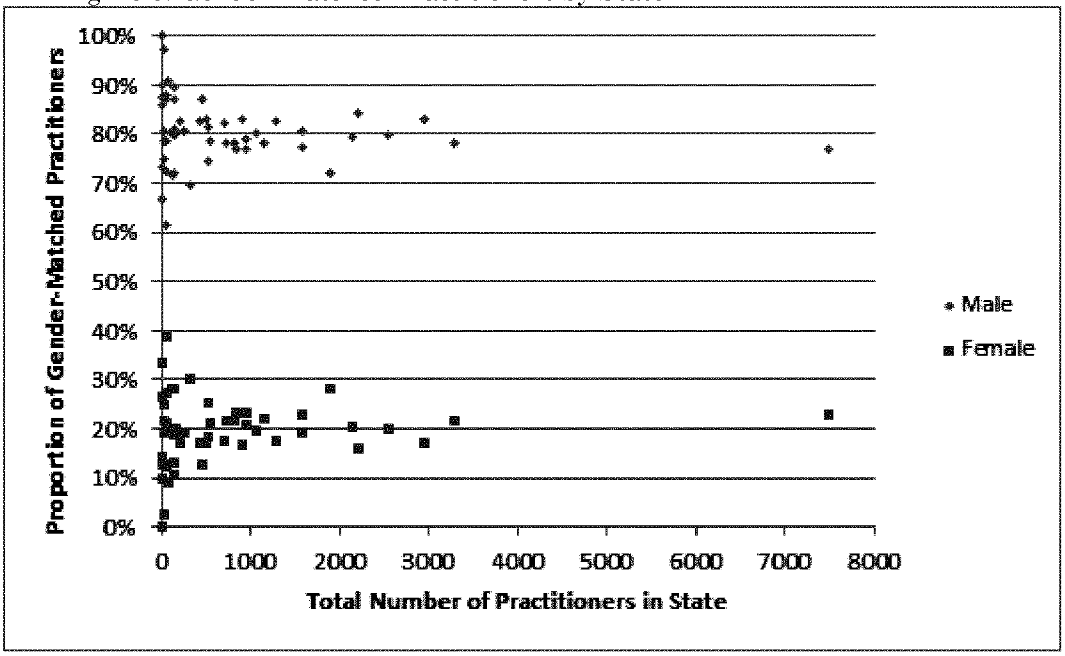

Figure 6. Gender-Matched Practitioners by City of Record

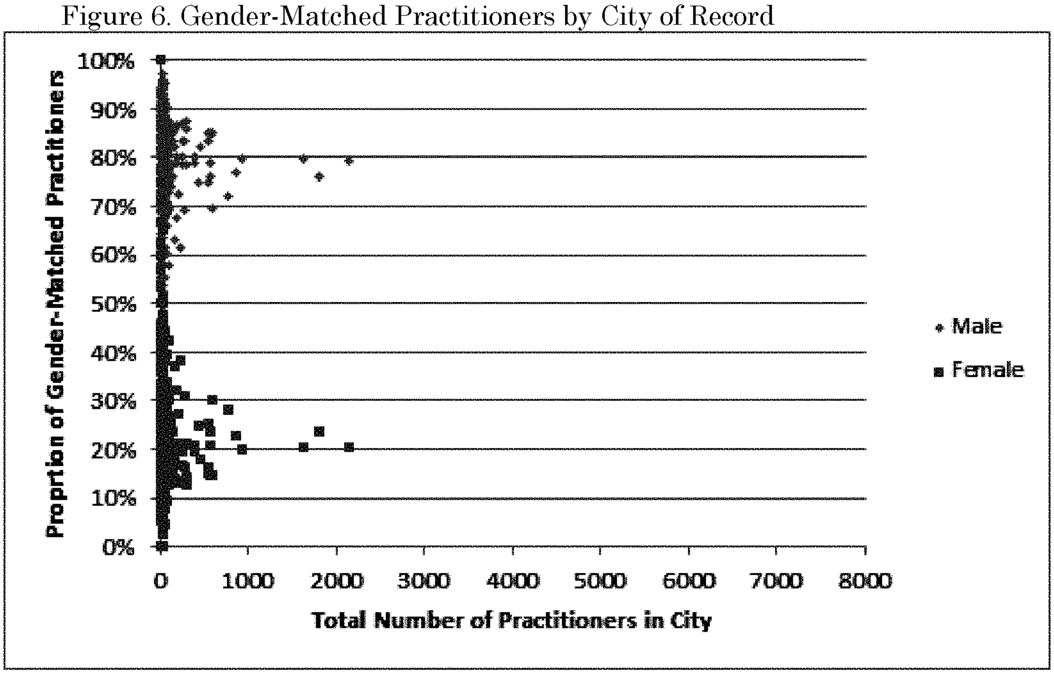

Figure 7. Gender-Matched Practitioners by ZIP Code

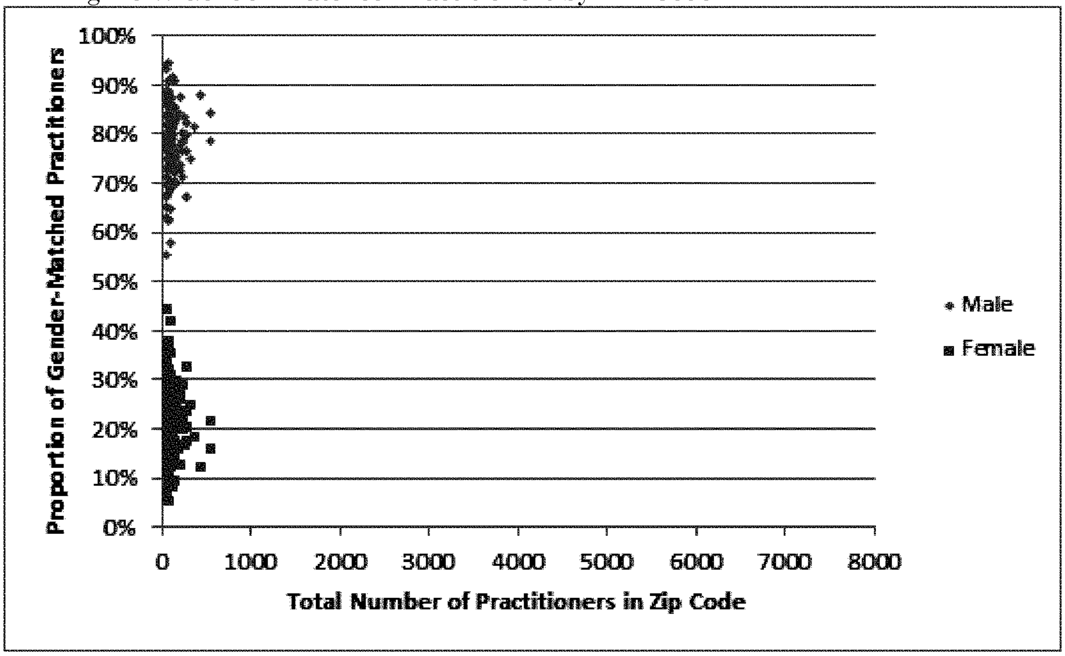

Figure 8. Gender-Matched Practitioners by MSA/μSA

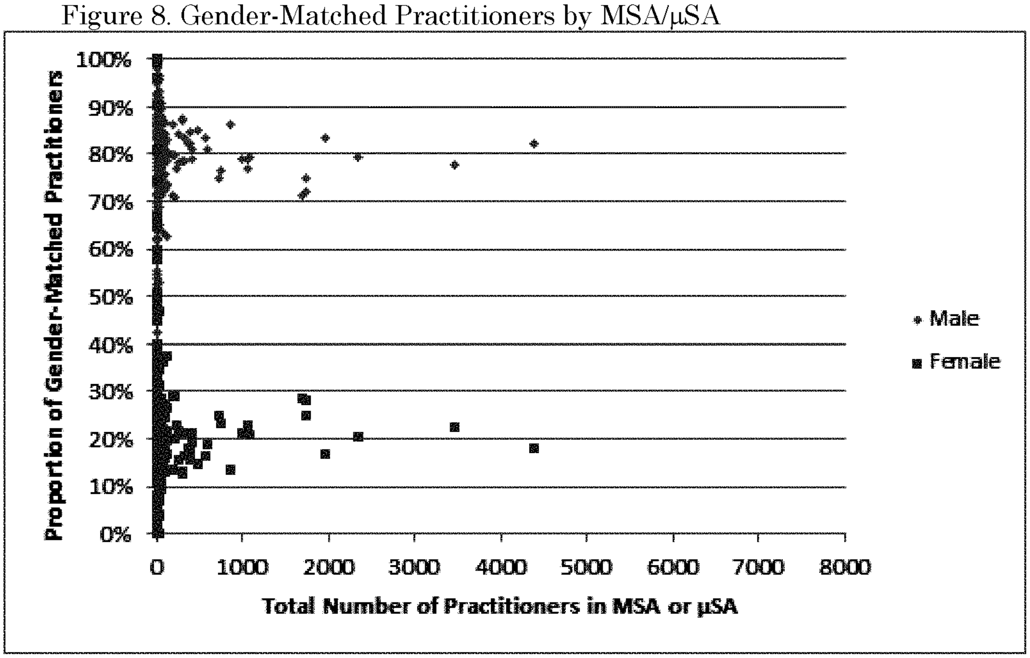

Figure 9. Gender-Matched Attorneys and Agents by State

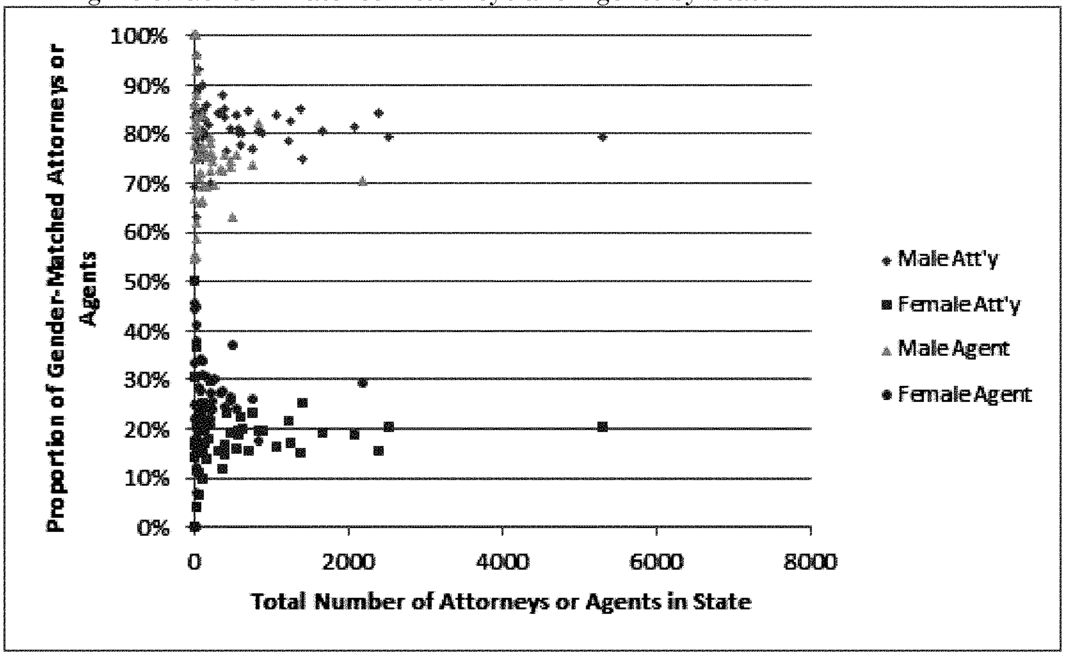

Figure 10. Gender-Matched Attorneys and Agents by MSA and μSA

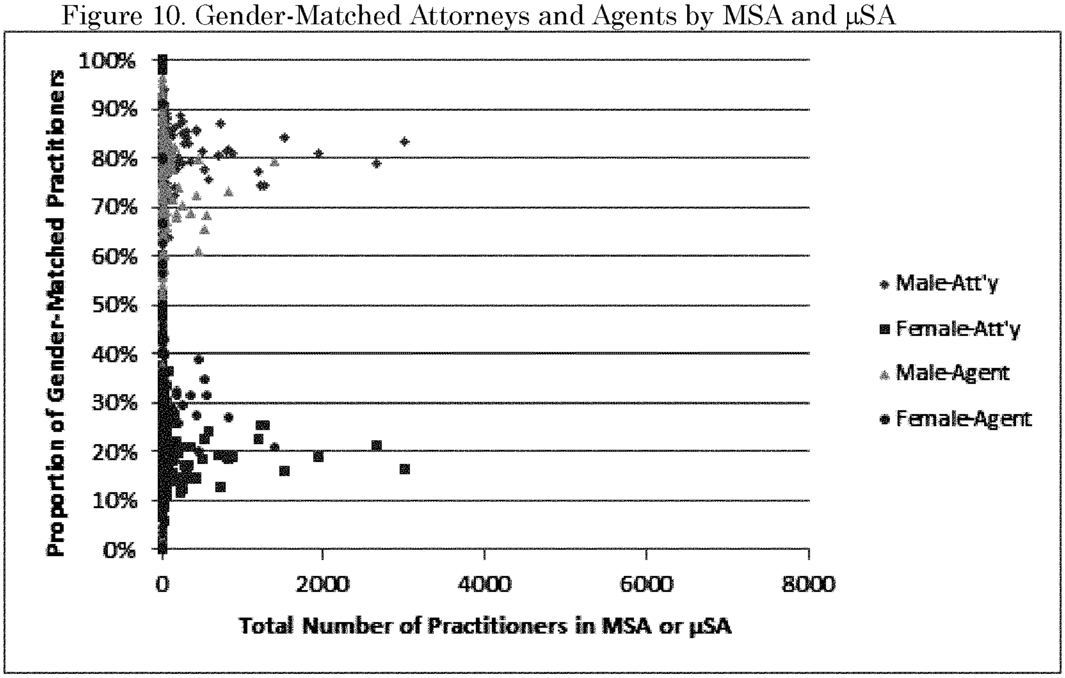

Figure 11. Female Attorneys/Agents Named on Granted Patents 2008–2012, All Technologies (backward-looking 6-month average)

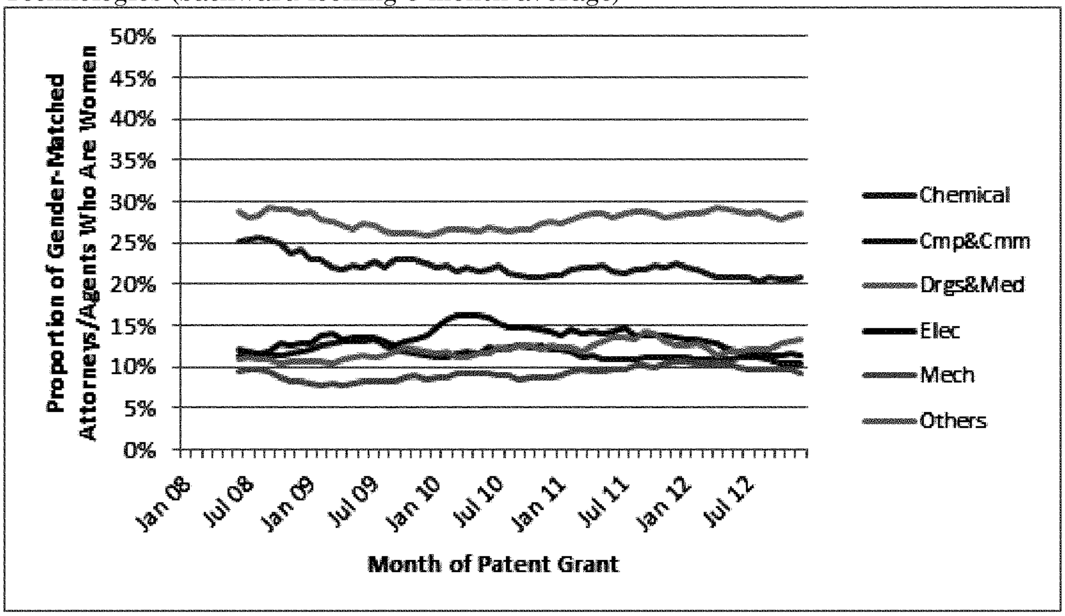

Figure 12. Female Attorneys/Agents Named on Granted Patents 2008–2012, Chemical (backward-looking 6-month average)

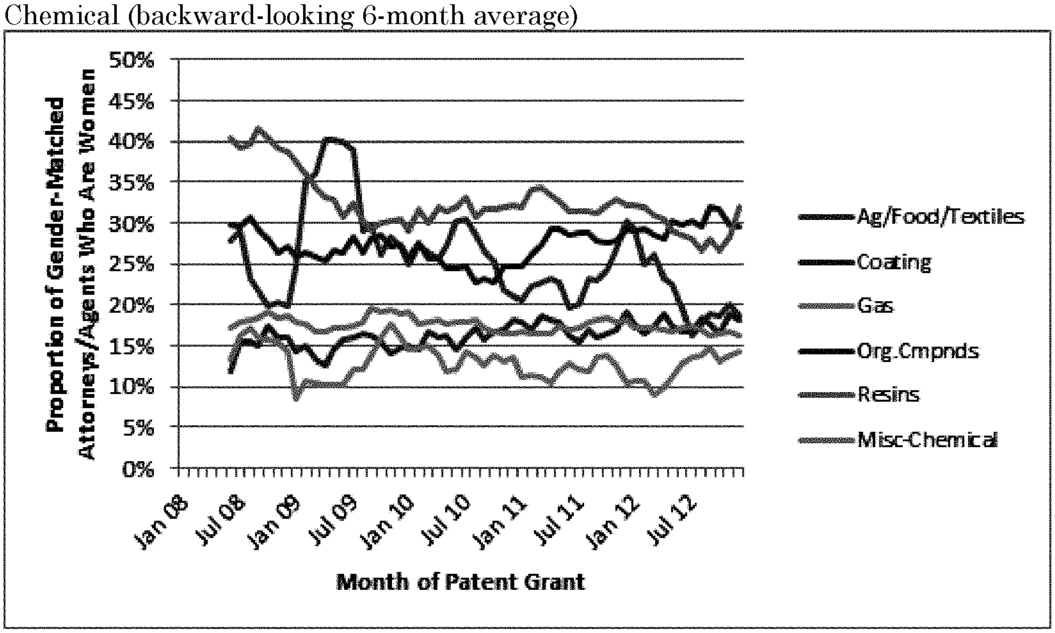

Figure 13. Female Attorneys/Agents Named on Granted Patents 2008–2012, Computers & Communications (backward-looking 6-month average)

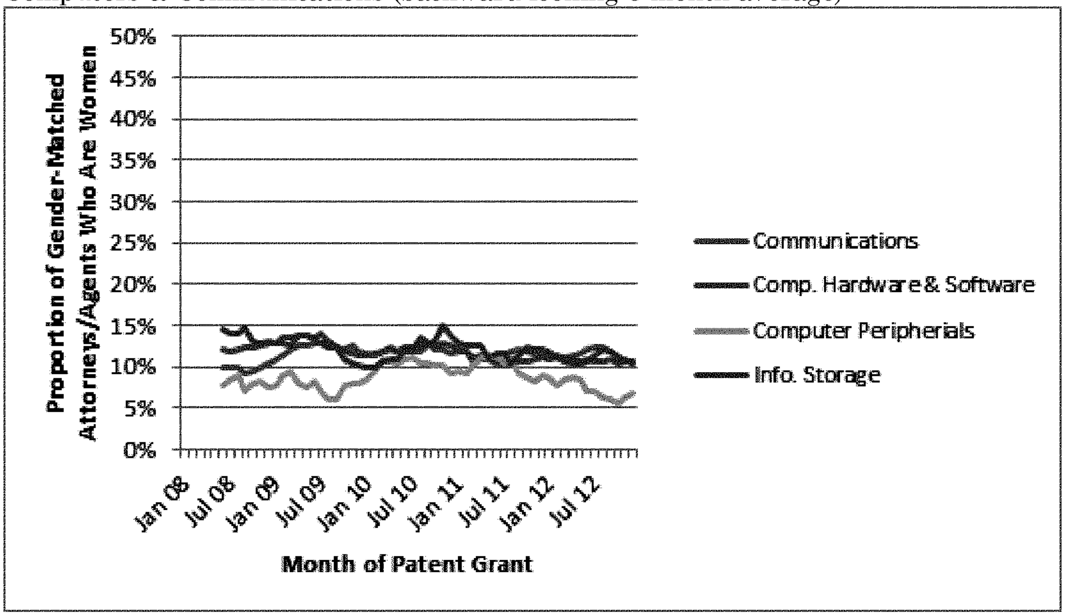

Figure 14. Female Attorneys/Agents Named on Granted Patents 2008–2012, Drugs & Medical (backward-looking 6-month average)

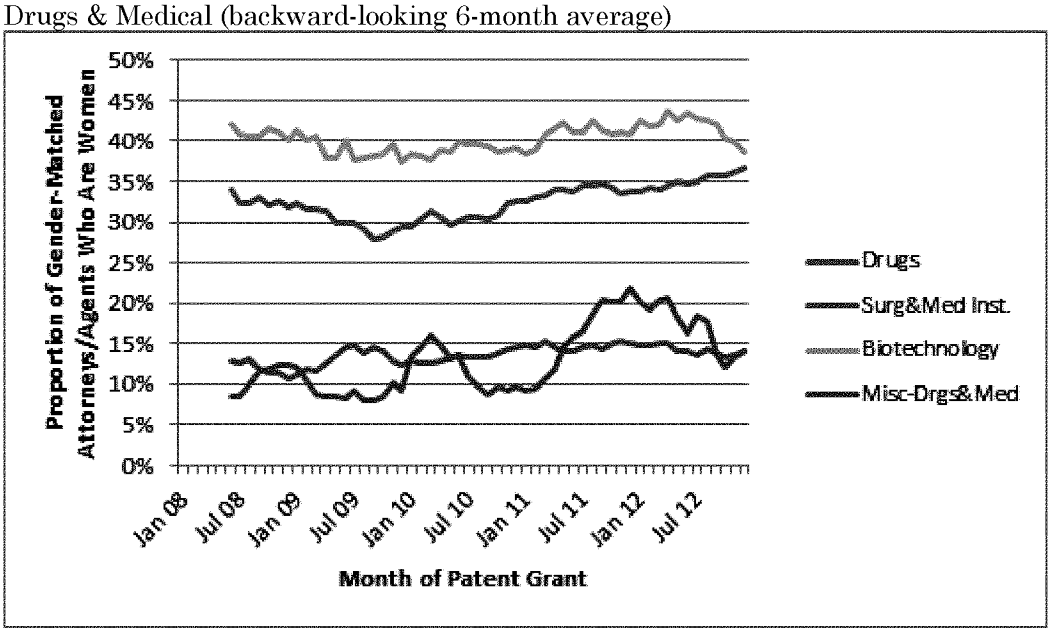

Figure 15. Female Attorneys/Agents Named on Granted Patents 2008–2012, Electrical & Electronic (backward-looking 6-month average)

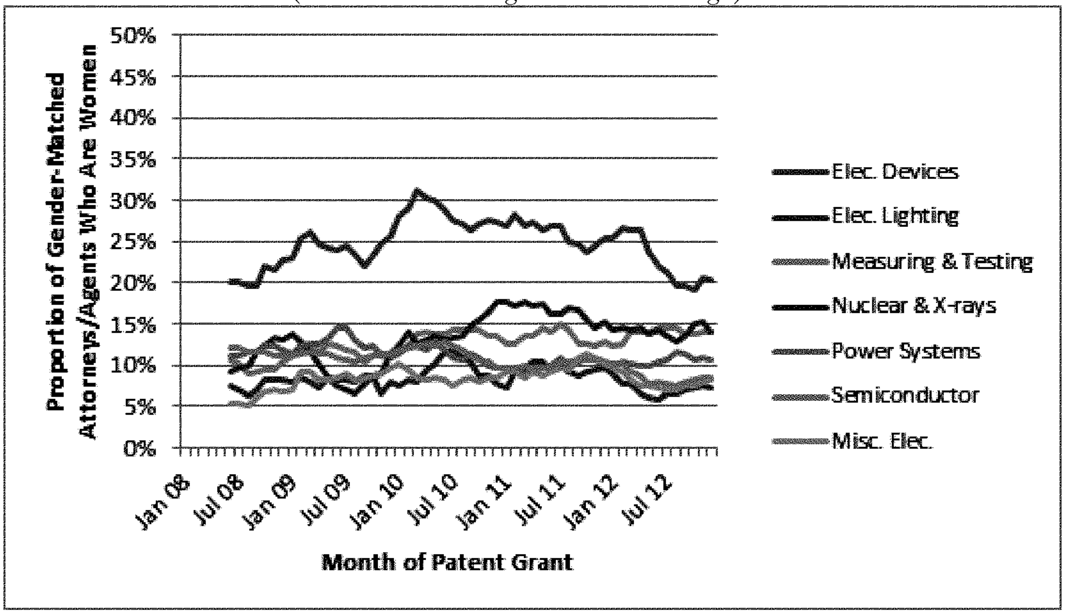

Figure 16. Female Attorneys/Agents Named on Granted Patents 2008–2012, Mechanical (backward-looking 6-month average)

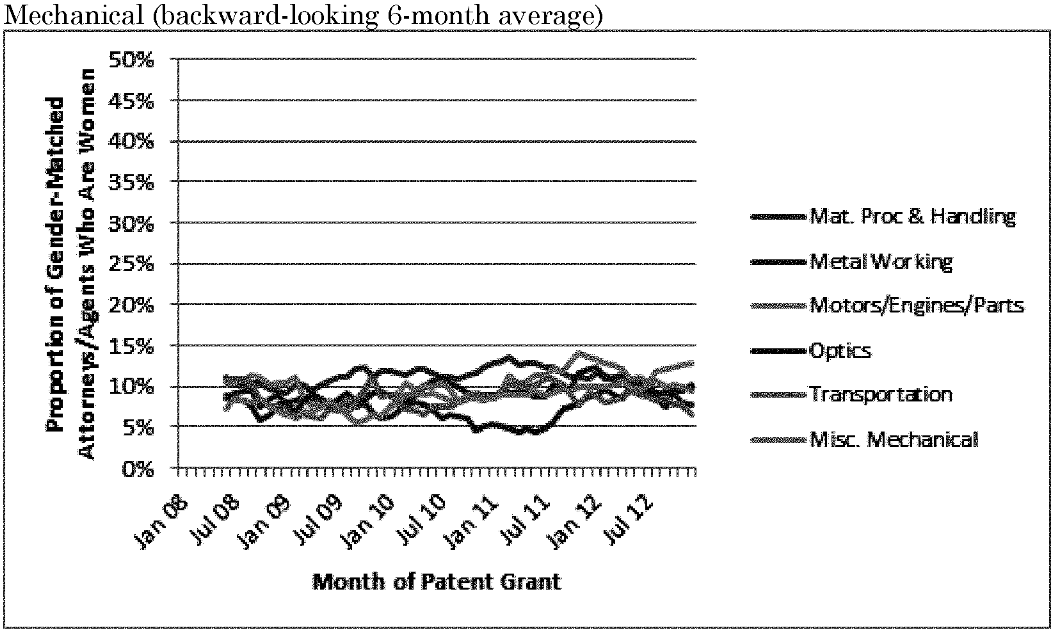

Figure 17. Female Attorneys/Agents Named on Granted Patents 2008–2012, Other Technologies (backward-looking 6-month average)