Gender differences in obtaining and maintaining patent rights

Kyle Jensen, Balázs Kovács & Olav Sorenson

An examination of the prosecution and maintenance histories of approximately 2.7 million US patent applications indicates that women have less favorable outcomes than men.

Although women make up half of the population, they represent just 10% of US patent inventors and only 15% of inventors in the life sciences[1–4]. By tracking patent applications through the prosecution process, we found disparities between men and women inventors in the processes of obtaining and maintaining patent rights. Patent applications by women inventors were found to be more likely to be rejected than those of men, and those rejections were less likely to be appealed by the applicant team (inventor, assignee, and prosecuting attorney). Conditional on being granted, patent applications by women inventors had a smaller fraction of their claims allowed, on average, than did applications by men. Further, those claims allowed had more words added during prosecution, thus reducing their scope and value. The granted patents of women inventors also received fewer citations than those of men and were less likely to be maintained by their assignees. Surprisingly, many of these effects were larger in the life sciences than in other technology areas.

Methodology

Our study examined the individual prosecution histories of approximately 2.7 million US utility patent applications from the years spanning 2001 to 2014 (ref. 5). The US Patent and Trademark Office (USPTO) recently released these data in aggregate. In the past, researchers could access these data only one application at a time, through the USPTO's Patent Application Information Retrieval system (https://portal.uspto.gov/pair/PublicPair/), thus hindering the large-scale empirical study of patent-prosecution outcomes. We joined these prosecution histories with the maintenance-fee and full-text patent databases available from the USPTO. The joined data allowed us to inspect the communication between applicants and examiners, the manner in which application claims changed during prosecution, the dates of various communications, the payment of maintenance fees, the influx of forward citations, and other phenomena.

We determined the probable gender of each inventor by using forename gender distributions available from the US Social Security Administration (https://www.ssa.gov/oact/babynames/limits.html) and from two commercial databases (see **Supplementary Data**). In the covered population, 94.1% of forenames were associated at least 95% of the time with only one gender. If an inventor had a highly gendered forename, we accordingly classified that inventor as either a man or a woman. This approach allowed us to classify the probable genders of 89% of the inventors listed on the applications (detailed discussion and analysis of the classification process, including possible selection issues, in **Supplementary Data**). Because most applications listed multiple inventors, we calculated a 'proportion women' variable: the number of women inventors divided by the total number of inventors on each application. When we refer to effect sizes, the disparities between men and women represented a shift in this variable from 0% to 100%, from all men to all women inventors.

We used a series of linear regressions and Poisson count models to estimate the associations between gender and various patent-prosecution outcomes (**Fig. 1**). These models included controls for a variety of patent attributes, such as the number of claims, the number of inventors listed on the patent, and the size of the organizational assignee. Detailed descriptions of the models, as well as their robustness, are provided in the **Supplementary Data**. For example, we demonstrated the robustness of the results to using subsamples of patent applications—such as applications with only US inventors and only applications from large organizations—and to using alternative, nonlinear specifications of the inventor-team composition.

Results and discussion

In **Figure 1**, the dark-blue bars depict 'raw' gender differences, and the light-blue bars depict gender differences after introduction of a fixed effect for each application's primary technology class in the United States Patent Classification (USPC) system (https://www.uspto.gov/web/patents/classification/). (In the USPC, each application submitted to the USPTO is assigned to one or more of >400 different USPC classes, which reflect the subject matter of the application, in categories as diverse as 'apparel', 'music', 'surgery', and 'molecular biology and microbiology'. These classifications are used to assign the patent application to particular groups of patent examiners.) As the figure illustrates, men and women differed less in their outcomes after adjustment for the technology class. For example, the two topmost blue bars indicate that women inventors were 21% less likely than men inventors to have their application accepted, but that difference declined to 7% after technology-class fixed effects were included. This effect could be viewed as an example of 'Simpson's paradox'[6]; that is, two-thirds of the diminished probability of women's applications being accepted stemmed from women applying at higher rates than men to technology classes with lower acceptance rates. In those classes, it is

Kyle Jensen, Balázs Kovács and Olav Sorenson are at the Yale School of Management, Yale University, New Haven, Connecticut, USA. e-mail: kyle.jensen@yale.edu, balazs.kovacs@yale.edu or olav.sorenson@yale.edu

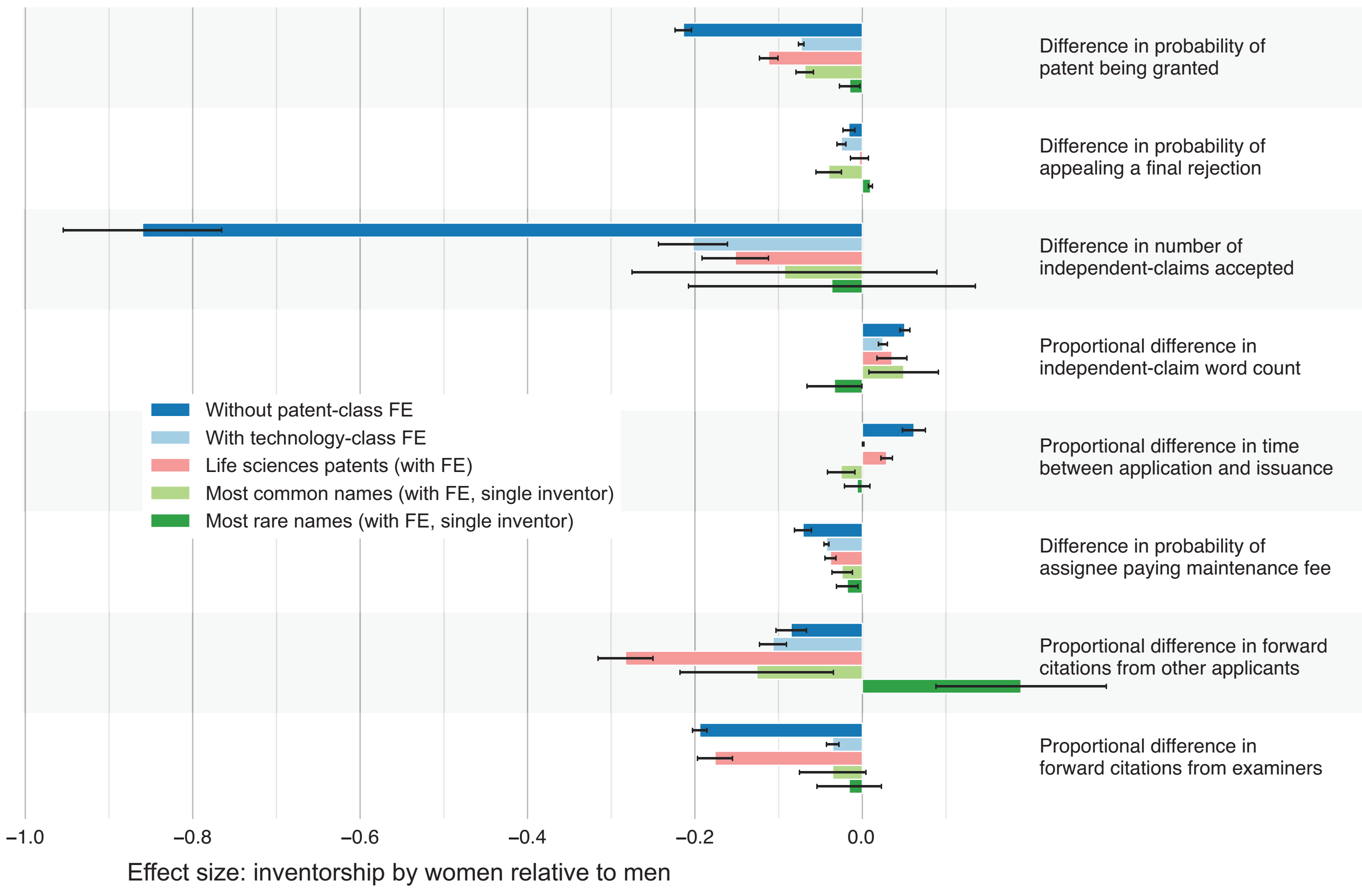


Figure 1 Estimated differences for teams of all women inventors relative to teams of all men, in the processes of obtaining and maintaining patent rights. Wide bars, point estimates; narrow bars, 95% confidence intervals (full model specifications in **Supplementary Data**). Teams with higher proportions of women had more negative outcomes during patent prosecution. For example, the topmost dark-blue bar indicates that patent applications by teams of all women inventors were 21% less likely to be granted than similar applications by teams of all men. The light-blue bar accounts for technology-class fixed effects (women are overrepresented in technology areas with lower acceptance rates); the topmost bar for example, indicates that even after accounting for technology-class fixed effects, all-women teams had a 7% lower probability of acceptance. The pink bars indicate the differences for patents in technology classes related to the life sciences. The final two bars (light and dark green) depict the estimated effects within two subsets of single-inventor patent applications. By examining the effects for inventors with common versus rare names, they provide an indication of the degree to which the gender differences stem from the applicant side—inventor, assignee, and attorney—versus the examiner side. The first two green bars, for example, suggest that approximately two-thirds of the lower probability of acceptance for applications with women inventors comes from the examiner side.

more challenging for anybody to get a patent approved, regardless of gender.

Even after adjustment for the differences across patent technology classes, however, women inventors still had less favorable experiences in nearly all outcomes. All else equal, relative to a team of all-men inventors, patent applications by teams of all women were 2.5% less likely to be appealed if rejected. Conditional on being granted, these applications, on average, had the number of independent claims reduced by one-fifth of a claim; had the number of words in their claims increased by 2.5%, thus narrowing the scope of these claims[7]; were 4.3% less likely to be maintained by their assignee; received 11% fewer citations from other patent applicants; and received 3.5% fewer citations from patent examiners. Forward citations trace the acknowledged contributions of prior art and are often used as measures of a patent's importance, scope, and value[8,9]. (These statistics all refer to the light-blue series in **Fig. 1**, which includes technology-class fixed effects, and appear in tabular form in the **Supplementary Data**.)

Although women might be expected to fare better in the life sciences, given their relatively higher representation in those fields, the data show no such pattern. The pink bars in **Figure 1** depict the gender differences within the subset of patents bearing life science classifications (description of how these are identified in **Supplementary Data**). For all outcomes that differed for the life sciences subset compared with the population of patents as a whole, the disparities in the life sciences appeared more disadvantageous to women. For example, in the life sciences, a team of all-women inventors was found to be 11% less likely than a team of all men to have its patent application accepted. Patents by women inventors in the life sciences also received 28% fewer forward citations from other inventors.

The data available did not allow us to isolate the mechanisms responsible for the gender differences in **Figure 1**—we were able to assess only the direction and magnitude of these differences. However, the natural variation in forename frequencies allowed us to gain some insight into the degree to which these differences arose from the applicant side—the inventor, assignee, and attorney—as compared with other parties. The inventors themselves are obviously aware of their own gender. Similarly, their employers and the attorneys representing them probably have firsthand knowledge of the inventors' genders. In contrast, the patent examiners and others must generally infer, either consciously or subconsciously, the gender of the inventors according to the forenames listed on patents and patent applications. (Examiners do sometimes meet with inventors in person and by telephone; robustness checks related to these scenarios are shown in the **Supplementary Data**.) For common names, such as 'Mary' and 'Robert,' outsiders can infer gender with a high level of confidence, but for thousands of rare names each held by only a few individuals, they cannot make such inferences. 'Jameire' and 'Kunnath', for example, are also strongly associated with gender, with the first being male and the second being female, but because they are rare names, few people would be aware of these associations. The gender differences associated with common names therefore should capture both differences in behavior on the applicant side as well as differences in treatment of those inventors by others. Any gender differences associated with rare names, in contrast, should stem only from the behavior of the applicant side.

The two series of green bars in **Figure 1** show how the frequency of an inventor's forename moderates the effects of gender on various outcomes. Because these models also include a control for forename frequency, they account for any association between the rarity of a name and the underlying quality of the patent, for example, because those patents might disproportionately represent foreign applicants. To avoid complications in aggregating across names of varying frequency, these models include only single-inventor patents. Among those, two outcomes had large and statistically significant differences between inventors with common forenames and those with rare forenames. First, among inventors with common names, women had an 8.2% lower probability of having their application accepted than did men. In contrast, among inventors with rare names, women had only a 2.8% lower probability of acceptance than did men. This combination suggests that approximately two-thirds of the lower probability of acceptance for applications with women inventors stemmed from the examiner side. Second, future patent applicants cited the patents of women with common names 30% less frequently than those of men with common names. The patents authored by women with rare forenames, and who were therefore not easily identified as women, were cited approximately 20% more often than the average patent by a male inventor with a rare forename, all else equal. To the extent that citations reflect patent quality, this result suggests that women inventors must clear a higher hurdle than men and therefore that the average patent granted to a woman inventor is of higher quality than the average patent granted to a man.

Conclusions

These results should interest inventors, patent holders, and policymakers. In advanced economies, technical progress appears to be the primary driver of economic growth[10]. The patent system, moreover, is one of the principal public-policy mechanisms for promoting this progress: governments grant patent holders a limited monopoly in exchange for a thorough disclosure of their inventions, so that others may build upon those inventions[11]. That women inventors are underrepresented in this system and appear disadvantaged in the process of obtaining and maintaining patents suggests that changes to the patent system and its prosecution process would increase fairness and might even stimulate economic growth.

A thorough discussion of possible adjustments to the patent system is beyond the scope of this paper, but we can imagine many possibilities worth consideration. It may help, for example, to make the patent-prosecution process more 'blind' to the identity of participants. Patents and patent applications could list only the initials of the forenames of the inventors on patent applications and could require communication between examiners and applicants to occur on a platform that would maintain the anonymity of the applicants. Such blind processes have eliminated gender inequality in other settings: For instance, when orchestras introduced opaque screens to conceal the identities of those auditioning, they hired more women and placed more women in leadership positions[12]. The introduction of such practices at the patent office could help to ameliorate the gender differences in patenting both in the life sciences and in other technological areas.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS
The authors are grateful to IPwe for access to patent data via Zuse Analytics and to T. Botelho and I. Fernandez-Mateo for comments on earlier versions of this paper.

AUTHOR CONTRIBUTIONS
The authors each contributed substantially to the work in this manuscript.

COMPETING INTERESTS
The authors declare no competing interests.

1. Ding, W.W., Murray, F. & Stuart, T.E. *Science* **313**, 665–667 (2006).
2. Wittington, K.B. & Smith-Doerr, L. *Gend. Soc.* **22**, 194–218 (2008).
3. Hunt, J. *et al. Res. Policy* **42**, 831–843 (2013).
4. Sugimoto, C.R. *et al. PLoS One* **10**, e0128000 (2015).
5. Marco, A.C. *et al. The USPTO Historical Patent Data Files: Two Centuries of Invention.* USPTO Economic Working Paper no. 2015-1 www.uspto.gov/sites/default/files/documents/USPTO_economic_WP_2015-01_v2.pdf (2015).
6. Albers, C.J. *Proc. Natl. Acad. Sci. USA* **112**, E6828–E6829 (2015).
7. Marco, A.C. *et al. Patent Claims and Patent Scope.* USPTO Economic Working Paper no. 2016–04 http://doi.org/10.2139/ssrn.2844964 (2016).
8. Trajtenberg, M. *Rand J. Econ.* **21**, 172–187 (1990).
9. Reitzig, M. *Res. Policy* **33**, 939–957 (2004).
10. Fagerberg, J. *J. Econ. Lit.* **32**, 1147–1175 (1994).
11. Mazzoleni, R. & Nelson, R.R. *Res. Policy* **27**, 273–284 (1998).
12. Goldin, C. & Rouse, C. *Am. Econ. Rev.* **90**, 715–741 (2000).

Supplementary Material for

Gender differences in obtaining and maintaining patent rights

Kyle L. Jensen, Balázs Kovács, and Olav Sorenson

This file includes:

Materials and Methods

Public Pair Patent application data

We used a dataset publicly available from the United States Patent and Trademark Office (Marco et al 2015). The dataset includes patent applications from as early as 1975 but it only has complete coverage of published applications beginning in 2001. We therefore restricted our analysis to the period from January 2001 to December 2014. This period had more than ten million patent-application events associated with more than three million patent applications. Because each stage in the process, such as a request for revision or an appeal, generates a separate event, the average application generates multiple events.

In line with the patenting literature [S1], we focused on applications for utility patents and disregarded those for design patents (utility applications account for 93.78 % of the applications in the sample). Also, because re-exam, re-issue, and provisional patent applications constitute special cases, we focused on "regular" patent applications, which account for 94.48% of the utility patents. The resulting sample includes 3,864,985 applications, 2,299,255 (59.49%) of which eventually received patents, 1,071,545 (27.72%) of which have been abandoned, and 494,185 (12.79%) of which remain in process.

For each application, the dataset identifies the inventor(s), the date of application, the set of decision rounds, the final outcome, the content of the application, the nationality/location of the inventor(s) and assignee(s), and the primary and secondary technology classifications to which the application has been assigned.

Gender coding of the inventors

We infer the gender of the inventors using three forename gender disambiguation datasets: (1) The U.S. Social Security application data report gender distributions for 87,703 unique forenames (e.g., Peter is 99.7% male). It includes the number of women with a particular forename and men with that same forename if the forename occurred more than five times among Social Security Number applicants. We also used two online services, (2) GenderAPI, and (3) genderize.io. Both services collect data from social media sites and other online sources that record the gender of their users and use this information to characterize the frequency and gender of forenames. These two databases allowed us to assign gender probabilities to an additional 137,779 and 19,063 forenames, respectively. We then used the combined set of 244,544 forenames to assign gender probabilities to the inventors listed on our sample of patents.

To avoid possible coding errors for gender ambiguous forenames, such as "Sasha," we adopted a conservative approach: We only considered an inventor a man (woman) if more than 95% of individuals with the same forename are men (or women). (See below for robustness checks and a discussion of possible sample selection issues). With this threshold, we assigned a gender to 88.6% of the inventors in the dataset. To ensure that our measures remained accurate at the patent application level – in other words, aggregating across all inventors listed on an application – we restricted our primary analysis to cases in which we could assign a gender to *all* of the listed inventors with at least 95% confidence. This restriction limited our analysis to roughly 70% of the applications. Within this subset, 8.8% of the inventors in our dataset have female forenames. See below for robustness checks using alternative inclusion criteria.

Of course, the gender distribution for each forename in our database primarily reflects the usage of those names in the U.S. because of our use of the Social Security Administration data. This geographic bias, however, fits well with our usage of the data, given that nearly all of the USPTO examiners themselves reside and have lived most of their lives in the U.S. One would therefore expect that their exposure to and beliefs about the gender associations of forenames would come primarily from the usage of those names in the U.S. (Finally, we wish to say that our binary coding of gender is to aid in the statistics of our analysis; the classifications are not intended as a commentary on the latent, possibly non-binary, gender identity of inventors.)

Table S1 reports the proportion of patents for which we could assign gender to all of the inventors as a function of the number of inventors listed on the patent, as well as for U.S.-based versus non-U.S.-based teams. One can readily see that our ability to characterize the gender of all members of the team declines with team size. Our gender assignment algorithms also have greater leverage among U.S.-based teams of inventors.

	Proportion of applications for which all inventors could be gender assigned			
Inventor count	Non-US-inventors	All US inventors	Count of cases	Proportion of cases
1	0.772	0.896	775,724	0.201
2	0.739	0.843	1,088,537	0.282
3	0.650	0.742	660,435	0.171
4	0.592	0.686	559,239	0.145
5	0.502	0.591	256,397	0.066
6	0.495	0.593	227,080	0.059
7	0.407	0.478	76,476	0.020
8	0.450	0.536	96,266	0.025
9	0.349	0.387	23,071	0.006
10	0.398	0.489	42,597	0.011
more than 10	0.319	0.340	59,160	0.015
Overall	0.646	0.763		

Table S1: Proportion of applications for which we could assign a gender with 95% confidence to all inventors, by inventor count and US/Non-US inventors.

The proportion of women inventors varies over time and across technological areas. Women also work with other women at rates well above chance. Fig. S1 depicts the number of men and women listed on the average patent over our sample period. Although both numbers rise over time, the small base for women means that they have been gaining slowly as a proportion of all inventors. Fig. S2 reports the proportions of women inventors in six broad technological areas. Women account for more than three times as large of a proportion of inventors in drugs and medicine as they do in mechanical technologies. These differences, moreover, appear relatively stable over the sample period.

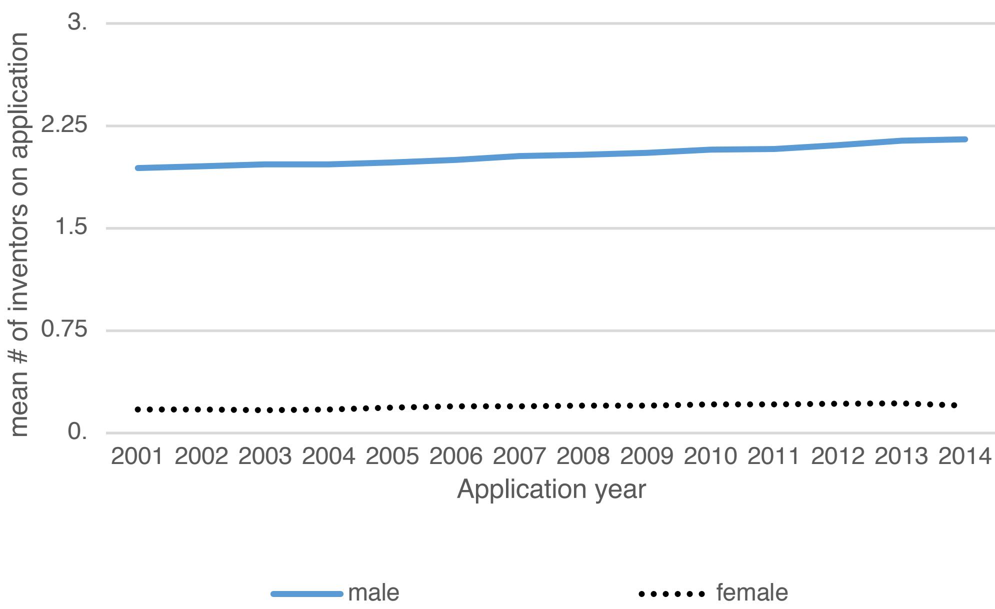


Fig. S1: Mean number of female and male inventors per patent application over time.

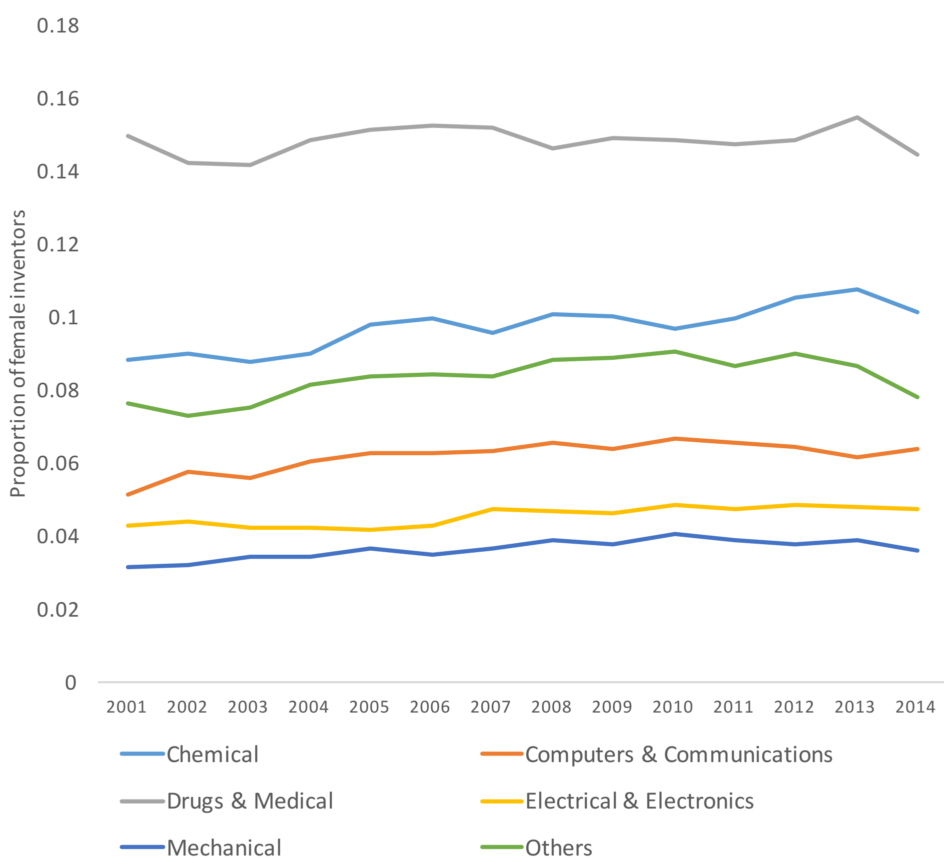


Fig. S2: Proportion of women inventors by technological area over time (as defined by the National Bureau of Economic Research, [S3]).

Table S2 meanwhile provides some sense of the extent to which men and women tend to invent with others of the same gender. If women sorted randomly into teams of inventors, one should see almost no teams of five or more, for example, composed only of women. For a team of five, the probability in any given year would amount to roughly 0.0005% (=$.088^5$); for a team of seven or more, it becomes an almost-negligible 4 (or fewer) out of 100 million. The table nevertheless reveals that teams with multiple women occur quite frequently.

Count of inventors on team	Count of female inventors: 0	1	2	3	4	5	6	7+	Total
1	607,704	38,299	0	0	0	0	0	0	**646,003**
2	773,826	51,051	32,277	0	0	0	0	0	**857,154**
3	377,125	57,152	19,557	3,637	0	0	0	0	**457,471**
4	276,918	43,506	27,096	3,808	1,769	0	0	0	**353,097**
5	95,510	27,256	10,802	3,525	944	146	0	0	**138,183**
6	85,043	14,738	15,955	2,380	2,474	233	262	0	**121,085**
7+	73,450	16,607	22,049	4,032	6,162	857	1,753	908	**125,818**
Total	**2,289,576**	**248,609**	**127,736**	**17,382**	**11,349**	**1,236**	**2,015**	**908**	**2,698,811**

Table S2. The table reports the distribution of the count of female inventors on inventor teams for applications for which we could identify the gender of all inventors.

Variables used in the analyses

Accepted (i.e., patent granted): An indicator variable with a value of one if the USPTO grants the patent requested in the patent application.

Appealed (i.e., request for re-examination): Applicants have the right to appeal a final rejection. A large proportion of applicants (55.3%) do so. If the applicant appeals, this indicator variable takes a value of one. To some extent, this outcome captures the perseverance of the applicant.

Change in the number of claims: As patent applications go through the examination process, the USPTO examiner may request changes to the original claims. These changes typically involve reducing the number of claims to narrow the scope of the patent.

Change in the claim (word) length: USPTO examiners may also request changes to the claims themselves. Those changes typically involve adding clauses that restrict the application of the claim to a particular area of application [11]. Because conditional clauses almost always increase the length of sentences, most claims become longer in length through the prosecution process.

Days between application and issuance: To measure the length of the process, we calculate the number of days between when the submission of the patent application and the issuance of the patent stemming from it.

Maintenance fees paid (maintained): To maintain a property right, the owner(s) of the patent must pay a maintenance fee at three, seven, and eleven years after issuance. Overall, about 88% of the patents get renewed at year three, 72% at year seven, and 47% at year 11. Our indicator variable has a value of one if the owner paid the maintenance fees at the first renewal (at three years after issuance).

Forward citations from examiners: USPTO examiners often request that patents add citations to prior art which they find in the course of their searches for whether the application represents a novel technology. These citations generally refer to ways an application builds on prior inventions and therefore effectively defer to those patents in terms of their coverage of particular intellectual property rights. Because of this fact, citation counts have been considered a good proxy for the value of a patent [12,13]. This measure counts the number of citations received by a granted focal patent from future patents (added to those patents by examiners).

Forward citations from applicants: Applicants also typically include their own citations to what they consider relevant prior art. These also effectively defer property rights to earlier patents and therefore increase the value of the patents receiving these citations. This measure counts the number of citations received by a granted focal patent from future patents (added to those patents by the applicants, either the inventors themselves or the lawyers representing them).

Various sets of models also control for a variety of other characteristics of the patents to adjust for any differences that might exist between men and women in the kinds of technologies for which they apply for patent protection:

Year of submission: Because the technological landscape changes over time, we adjust for the year of patent application submission, by including a vector of indicator variables for these years.

USPC Primary Class: The USPTO assigns patent applications to one or more of the roughly 400 classes in the US Patent Classification System. Because different technologies have substantially different acceptance, revision, and citation rates, in the more saturated models (reported as the dark bars in Fig. 1), we account for these differences by adjusting for the mean levels of the various dependent variables across all patent applications assigned to a primary class (equivalent to "fixed effects" for primary patent classes). Because patent classes evolve over time, we have implemented these intercepts as specific to the year of application (i.e., one can think of them as class-application-year fixed effects, more than 6000 intercepts). We should note that, even though the USPTO has changed the UPC classification system (in 2015), we use here the USPC classification system which had been in effect during our observation period.

Foreign/US domestic: As the country of origin of the applicants may influence patenting outcomes [S2], some of our models control for the correspondence address of the applicant (as an indicator variable).

Small entity: Prior research has found that small entities have lower than average acceptance rates. The USPTO defines a small entity as (i) independent inventors, (ii) companies with fewer than 500 employees, and (iii) organizations with non-profit status (such as universities). In our sample, 28% of the applications have been submitted by small entities. As size may influence patenting experience and resources, some of our models included an indicator variable to control for small entity status. We should note that women appear to sort into small entities: 8.4% of inventors in small entities are women (versus 6.9% of inventors in other categories).

Continuation patents and divisional patent applications: Some patent applications are filed as continuations of previous patents. Because these applications are likely to be handled differently, we controlled for whether the patent application is a continuation, or divisional, application.

Examiner experience. As the seniority and prior experience of patent examiners may influence the speed and eventual outcome of the application process [S4], we control for examiner's experience. We proxy examiners' experience with the (logged) cumulated count of patent applications that they handled prior to the focal application.

Foreign priority. Some patent applications to the USPTO are based on inventions that have already been patented in another jurisdiction. As these patent applications may be handled differently, some of our models include a binary variable to indicate whether the focal patent application has a foreign priority (i.e., is based on a patent or patent application previously submitted to a non-US patent office).

	ALL PATENT APPLICATIONS (N=3,864,982)		PATENT APPLICATIONS FOR WHICH WE COULD ASSIGN THE GENDER OF ALL APPLICANTS (N=2,698,811)	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Count of male inventors (95pct certainty)	2.443	2.050	2.617	1.867
Count of female inventors (95pct certainty)	0.247	0.684	0.233	0.646
Application accepted	0.595	0.491	0.598	0.490
Request for re-examination	0.248	0.432	0.248	0.432
Difference in claim vs issued application	-1.506	10.382	-1.394	8.883
Difference in avg. word length of independent claims in claimed vs issued application	1.537	1.739	1.534	1.612
Patent maintained	0.370	0.483	0.377	0.485
Total examiner-added citations	2.507	4.489	2.504	4.431
Total applicant-added citations	3.802	15.221	3.870	15.455
Application year	2006.985	3.575	2006.871	3.591
Days until acceptance	1112.485	550.125	1113.850	554.776
Small entity	0.250	0.433	0.260	0.439
Continuation patent	0.538	0.499	0.556	0.497
Has Foreign Priority	0.411	0.492	0.393	0.488

Table S3: Descriptive statistics of the main variables.

Table S3 reports descriptive statistics for these variables. It reports two sets of statistics, one for the population, the other for the subset for which we could assign gender to all inventors. One can see that the subset used for the analysis differs little from the population.

Methods of estimation

In most of our models (columns 2-6 in tables S4-S15), we estimated linear regression models with robust standard errors. Even for dichotomous outcomes, such as acceptance of the patent, these models (referred to as linear probability models for these cases) provide unbiased estimates of the conditional mean probability of an outcome and have an advantage relative to estimation using the logit or probit of avoiding incidental parameter bias due to the inclusion of a large number of intercepts [S5,S6]. In the models estimating citation counts and days in process (columns 7-9 in tables S4-S15), we estimated quasi-maximum likelihood Poisson models with robust standard errors [S7].

Specifically, we estimated two primary models:

(1) $Y = f(\beta \ \ Proportion\, Women + \tau + \varepsilon)$,

(2) $Y = f(\beta \ \ Proportion\, Women + \varphi + \varepsilon)$,

where β denotes the coefficient of interest, the partial correlation of the proportion of women on the team of inventors with the various outcomes, Y, τ represents a vector of application-year indicator variables, and φ denotes a vector of USPC-primary-class-year indicator variables (in others words, with a separate variable representing each class in each application year). We used STATA 13's xtreg and xtpoisson commands with the "fe" option to estimate the models.

Results

	Patent issued	Request for reexam[a]	Independent claim count change[b]	Independent claim word length change[b]	Maintained[c]	Examiner cites[d]	Applicant cites[d]	Days between application and issuance[d]
Prop. female	-0.130***	-0.009**	-1.001***	0.076***	-0.048***	-0.216***	-0.089***	0.060***
	(0.006)	(0.004)	(0.110)	(0.009)	(0.007)	(0.011)	(0.020)	(0.013)
Constant	0.608***	0.557***	-1.164***	1.499***	0.677***			
	(0.000)	(0.000)	(0.007)	(0.001)	(0.000)			
Fixed effects	Year	Year	Year	Year	Year	Year	Year	Year
Observations	2,698,811	1,117,714	1,425,515	1,425,439	1,508,816	1,613,836	1,613,836	1,613,832
Log-likelihood	-1.811e+06	-797977	-5.123e+06	-2.649e+06	-723011	-5.224e+06	-1.340e+07	-1.870e+08
Prop. female	-0.044***	-0.014***	-0.245***	0.037***	-0.029***	-0.036***	-0.113***	0.001
	(0.002)	(0.003)	(0.050)	(0.008)	(0.002)	(0.008)	(0.018)	(0.003)
Constant	0.601***	0.557***	-1.211***	1.501***	0.676***			
	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)			
Fixed effects	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear
Observations	2,698,579	1,117,706	1,425,506	1,425,430	1,508,806	1,613,346	1,612,826	1,613,664
Log-likelihood	-1.668e+06	-776781	-5.110e+06	-2.636e+06	-682568	-4.672e+06	-1.100e+07	-1.290e+08

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a] estimated on the subset of patent applications with at least one final rejection
[b] estimated on the subset of issued patents for which we could obtain claim comparison data
[c] estimated on the subset of patents issued before 2012
[d] estimated on the subset of issued patents

Table S4: The estimated effects of the proportion of women inventors listed on a patent application on the processes of obtaining, maintaining, and asserting patent rights. The first five columns report linear regression models while the last three report Poisson count models. For each outcome, the first specification adjusts for year, the second for primary class by year. These results have been used to construct the blue bars in Fig. 1, with the top panel being the light blue bars and the bottom panel the dark blue bars. As the linear regression models predict absolute differences, we transform them for the figure to percentage differences by rescaling the co-efficient estimate by the constant (the average outcome for an investor list with no women). For example, for the first column, we rescale the estimate -0.130 to -0.130/0.608=-0.214. To transform the Poisson models to percentage differences, we raise *e* to the power of the coefficients.

	Patent issued	Request for reexam[a]	Independent claim count change[b]	Independent claim word length change[b]	Maintained[c]	Examiner cites[d]	Applicant cites[d]	Days between application and issuance[d]
Prop. female	-0.103***	-0.024***	-1.441***	0.007	-0.031**	-0.868***	-1.112***	0.001
	(0.012)	(0.005)	(0.186)	(0.033)	(0.011)	(0.019)	(0.068)	(0.019)
Constant	0.463***	0.569***	-2.880***	1.557***	0.629***			
	(0.002)	(0.001)	(0.023)	(0.004)	(0.001)			
Fixed effects	Year	Year	Year	Year	Year	Year	Year	Year
Observations	349,857	170,791	140,550	140,540	145,115	156,781	156,781	156,780
Log-likelihood	-243435	-121198	-517206	-326092	-74837	-379784	-2.241e+06	-1.800e+07
Prop. female	-0.051***	-0.002	-0.456***	0.055**	-0.024***	-0.176***	-0.283***	0.029***
	(0.005)	(0.006)	(0.119)	(0.028)	(0.004)	(0.021)	(0.033)	(0.007)
Constant	0.455***	0.566***	-3.003***	1.551***	0.629***			
	(0.001)	(0.001)	(0.015)	(0.003)	(0.001)			
Fixed effects	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear
Observations	349,851	170,791	140,548	140,538	145,113	156,261	156,341	156,401
Log-likelihood	-234173	-120132	-515485	-325699	-72199	-320857	-1.768e+06	-1.640e+07

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a] estimated on the subset of patent applications with at least one final rejection
[b] estimated on the subset of issued patents for which we could obtain claim comparison data
[c] estimated on the subset of patents issued before 2012
[d] estimated on the subset of issued patents

Table S5: The table reports the estimated effect of the proportion of women inventors on the processes of obtaining, maintaining, and asserting patent rights. The table is identical to Table S4, except that these models have been estimated on the subset of patent applications that are classified as “Drugs&Medical” by the National Bureau of Economic Research (NBER). The results of the bottom panel (with class-year fixed effects) have been used to construct the pink bars in Fig. 1. The first five columns report linear regression models while the last three report Poisson count models. As the linear regression models predict absolute differences, we transform them for the figure to percentage differences by rescaling the coefficient estimate by the constant (the average outcome for an investor list with no women). For example, for the first column, bottom panel, we rescale the estimate -0.051 to -0.051/0.455=-0.112. To transform the Poisson models to percentage differences, we raise *e* to the power of the coefficients.

	Patent issued	Request for reexam[a]	Independent claim count change[b]	Independent claim word length change[b]	Maintained[c]	Examiner cites[d]	Applicant cites[d]	Days between application and issuance[d]
Prop. female	-0.092***	-0.021***	-0.178***	0.030**	-0.031***	-0.076***	-0.064***	0.001
	(0.003)	(0.004)	(0.033)	(0.011)	(0.004)	(0.011)	(0.020)	(0.007)
Constant	0.642***	0.528***	-0.714***	1.493***	0.811***			
	(0.000)	(0.000)	(0.002)	(0.001)	(0.000)			
Fixed effects	Year	Year	Year	Year	Year	Year	Year	Year
Observations	646,003	262,431	344,090	344,054	410,725	411,413	411,413	411,418
Log-likelihood	-436253	-188962	-1.053e+06	-579826	-190410	-1.463e+06	-3.360e+06	-5.780e+07
Prop. female	-0.044***	-0.016***	-0.044	0.009	-0.022***	-0.034***	-0.061**	-0.018***
	(0.003)	(0.004)	(0.047)	(0.011)	(0.003)	(0.011)	(0.029)	(0.005)
Constant	0.640***	0.528***	-0.721***	1.495***	0.810***			
	(0.000)	(0.000)	(0.002)	(0.001)	(0.000)			
Fixed effects	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear
Observations	645,934	262,428	344,087	344,051	410,720	410,645	410,112	410,960
Log-likelihood	-398543	-181051	-1.048e+06	-573165	-180992	-1.321e+06	-2.733e+06	-3.630e+07

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a] estimated on the subset of patent applications with at least one final rejection
[b] estimated on the subset of issued patents for which we could obtain claim comparison data
[c] estimated on the subset of patents issued before 2012
[d] estimated on the subset of issued patents

Table S6: The table reports the estimated effect of a woman inventor on the processes of obtaining, maintaining, and asserting patent rights. The table is identical to Table S4, except that these models have been estimated on the subset of patent applications that have a single inventor.

The light and dark green bars in Fig.1 depict how the effect sizes differ for solo inventors with rare versus common forenames. To estimate how the frequency of the forename influences patent application outcomes, we re-estimated the models of Table S6 but we also included as an additional variable the frequency of the forename of the inventor (measured as the logged number of counts of persons born in the US in a given year with that name) and an interaction between the popularity of the name and the gender of the name (male or female). Note that the frequency of forename variable should capture any differences in the quality of patents associated with the rarity of inventor names. For example, to the extent that foreign or immigrant inventors might both have less common names and applications less likely to succeed, these effects would appear in the "main" effect of this variable. The interaction effect between name frequency and the gender of the name captures the gender effect for those with rare names.

The values depicted in Fig. 1 reflect the values for the marginal effect of "female inventor" at the lowest and highest observed values of "logged forename popularity": -3 and 10.2. Note that the negative value means that these forenames are rarely, if ever, given to children born in the U.S. We gender-identified these names using the two commercial datasets described above and rescaled the forename popularity across datasets such that their frequencies for common names would be comparable to those in the U.S. Social Security database.

	Patent issued	**Request for reexam[a]**	**Independent claim count change[b]**	**Independent claim word length change[b]**	**Maintained[c]**	**Examiner cites[d]**	**Applicant cites[d]**	**Days between application and issuance**
Female inventor	-0.028***	-0.002	-0.050	-0.013	-0.020***	-0.020	0.113***	-0.011*
	(0.004)	(0.006)	(0.058)	(0.011)	(0.004)	(0.014)	(0.037)	(0.006)
Foretname popularity (logged)	-0.001**	-0.004***	-0.035***	0.001*	0.001***	0.003***	0.046***	0.002***
	(0.000)	(0.000)	(0.003)	(0.001)	(0.000)	(0.001)	(0.002)	(0.000)
Firstname popularity (logged) X Female inventor	-0.004***	-0.004***	-0.004	0.006**	-0.000	-0.001	-0.023***	-0.001
	(0.001)	(0.001)	(0.012)	(0.002)	(0.001)	(0.002)	(0.006)	(0.001)
Constant	0.642***	0.550***	-0.558***	1.489***	0.805***			
	(0.001)	(0.002)	(0.015)	(0.003)	(0.001)			
Fixed effects	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear
Observations	664,311	269,620	353,734	353,697	422,062	422,002	421,480	422,327
Log-likelihood	-410162	-185890	-1.076e+06	-587198	-186293	-1.357e+06	-2.773e+06	-3.720e+07

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a] estimated on the subset of patent applications with at least one final rejection
[b] estimated on the subset of issued patents for which we could obtain claim comparison data
[c] estimated on the subset of patents issued before 2012
[d] estimated on the subset of issued patents

Table S7: The table reports the estimated effect of having a woman inventor on the processes of obtaining, maintaining, and asserting patent rights. The table is identical to the bottom panel of Table S6, except that it includes as additional variables the popularity of the forename of the inventor and an interaction effect for name popularity and gender.

Additional analyses and robustness checks

We also estimated a variety of additional models to explore the sensitivity of the results to our coding and estimation choices. As noted in the manuscript, one of the central issues concerns whether men and women differ in terms of the kinds of technologies that they invent or whether the observed differences reflect a differential treatment of men and women for similar applications. The technology class fixed effects in the regressions above classify applications into more than 400 groups. Their inclusion dramatically reduces the observed differences between men and women.

These classes nevertheless remain relatively broad. The typical class has more than 5000 applications in our sample. Even within these classes, applications may therefore differ in meaningful ways. We assessed this possibility in two ways. The first involved using the far more detailed subclass system to adjust for technological differences. The second incorporated other attributes of the patent as covariates. These alternatives produce very similar results to the main analysis above.

Adding controls for a more refined classification system. We explored whether using more refined controls for the technological content of the patent applications would alter the main set of results. To do so, we relied on subclass information provided for each patent application. Each primary class has dozens or even hundreds of subclasses. Overall, the United States Patent Classification system has more 100,000 subclasses.

Table S8 reports the results using subclass-by-year fixed effects. The use of these more fine-grained controls does reduce the magnitude of the coefficients associated with the proportion of women on the team, by as much as 25%. But the pattern remains. This approach also uses more than one million additional degrees of freedom. Chi-squared tests for whether one should prefer these models to those with the broader class-by-year fixed effects all yield negative results ($p > .99$). Subclass-level adjustments therefore appear to overfit the data.

	Patent issued	Request for reexam[a]	Independent claim count change[b]	Independent claim word length change[b]	Maintained[c]	Examiner cites[d]	Applicant cites[d]	Days between application and issuance[d]
Prop. female	-0.035***	-0.010***	-0.211***	0.029***	-0.033***	-0.052***	-0.114***	0.003
	(0.002)	(0.003)	(0.046)	(0.008)	(0.002)	(0.008)	(0.018)	(0.002)
Constant	0.601***	0.557***	-1.213***	1.502***	0.676***			
	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)			
Fixed effects	SubclassXyear	SubclassXyear	SubclassXyear	SubclassXyear	SubclassXyear	SubclassXyear	SubclassXyear	SubclassXyear
Observations	2,696,841	1,117,706	1,425,506	1,425,430	1,508,806	1,573,850	1,558,026	1,584,145
Log-likelihood	-1.644e+06	-730840	-5.074e+06	-2.590e+06	-859186	-4.608e+06	-9.661e+06	-1.500e+08

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a] estimated on the subset of patent applications with at least one final rejection
[b] estimated on the subset of issued patents for which we could obtain claim comparison data
[c] estimated on the subset of patents issued before 2012
[d] estimated on the subset of issued patents

Table S8. The table reports the estimated effect of the proportion of women inventors listed on a patent application on the processes of obtaining, maintaining, and asserting patent rights, adjusting for the average outcomes in each subclass in each year. The first five columns report linear regression models while the last three report Poisson count models.

Adding control variables. Given that the subclass fixed effects did not improve the models, our next set of analyses reverted to using technology fixed effects. Here, however, we included as controls several attributes, such as being based in the U.S., which might influence either the quality of the patent or the resources available for pursuing a patent.

Table S9 reports the results of these analyses. The inclusion of these variables has little effect on the results, generally shifting the point estimates for the proportion women by less than 10%. The largest effect appears on changes in the number of claims. The disadvantage for a team of all women relative to one of all men declines by roughly 19% with the inclusion of these controls.

	Patent issued	Request for reexam	Independent claim count change	Mean independent claim word change	Maintain	Examiner cites	Applicant cites	Days between application and issuance[d]
Prop_female	-0.044***	-0.014***	-0.185***	0.036***	-0.029***	-0.060***	-0.170***	-0.001
	(0.002)	(0.002)	(0.048)	(0.008)	(0.002)	(0.008)	(0.020)	(0.003)
USaddress	0.038***	0.015***	0.006	0.018*	-0.002	0.007	0.066***	0.023***
	(0.004)	(0.005)	(0.043)	(0.010)	(0.004)	(0.011)	(0.026)	(0.004)
Inventor count	0.005***	0.006***	-0.095***	0.005***	0.001**	0.052***	0.071***	0.006***
	(0.000)	(0.000)	(0.008)	(0.001)	(0.000)	(0.001)	(0.002)	(0.000)
Small entity	-0.131***	-0.068***	-0.318***	0.045***	-0.063***	-0.020***	-0.169***	-0.017***
	(0.003)	(0.002)	(0.027)	(0.004)	(0.002)	(0.006)	(0.017)	(0.002)
Examiner experience	0.072***	-0.029***	0.316***	-0.100***	0.038***	-0.004**	0.016***	-0.080***
	(0.001)	(0.001)	(0.011)	(0.002)	(0.001)	(0.002)	(0.004)	(0.001)
Continuation or divisional patent	0.006***	0.010***	-0.167***	0.016***	0.039***	0.057***	0.304***	-0.072***
	(0.001)	(0.001)	(0.026)	(0.003)	(0.002)	(0.006)	(0.010)	(0.002)
Has foreign priority	-0.033***	0.024***	0.630***	0.022***	-0.023***	-0.150***	-0.636***	-0.019***
	(0.002)	(0.002)	(0.021)	(0.004)	(0.001)	(0.006)	(0.014)	(0.002)
Constant	0.240***	0.665***	-2.675***	1.949***	0.488***			
	(0.006)	(0.006)	(0.071)	(0.015)	(0.007)			
Fixed effects	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear
Observations	2,698,579	1,117,706	1,425,506	1,425,430	1,508,806	1,613,346	1,612,826	1,613,664
Log-likelihood	-1.610e+06	-771693	-5.108e+06	-2.633e+06	-671284	-4.633e+06	-1.050e+07	-1.220e+08

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a] estimated on the subset of patent applications with at least one final rejection
[b] estimated on the subset of issued patents for which we could obtain claim comparison data
[c] estimated on the subset of patents issued before 2012
[d] estimated on the subset of issued patents

Table S9: The table reports the estimated effect of the proportion of women inventors listed on a patent application on the processes of obtaining, maintaining, and asserting patent rights, adjusting for inventor location, total inventor count, small entity status, examiner experience, whether the application has a parent patent, and whether it has foreign priority. The first five columns report linear regression models while the last three report Poisson count models.

Robustness of gender identification

We explored the sensitivity of the results to a variety of approaches, such as relaxing the degree of confidence required, that allow us to include a larger proportion of the applications.

Using a lower gender-assignment probability threshold. We explored whether changing the 95% certainty threshold in the gender identification algorithm to 90% would affect our results. We have rerun our analyses with this gender-assignment algorithm and found that our results are robust to this specification. Table S10 reports these results.

	Patent issued	**Request for reexam**[a]	**Independent claim count change**[b]	**Independent claim word length change**[b]	**Maintained**[c]	**Examiner cites**[d]	**Applicant cites**[d]	**Days between application and issuance**[d]
Prop. female	-0.128***	-0.009*	-0.985***	0.066***	-0.047***	-0.214***	-0.078***	0.058***
	(0.006)	(0.004)	(0.108)	(0.009)	(0.006)	(0.010)	(0.020)	(0.013)
Constant	0.607***	0.558***	-1.179***	1.501***	0.676***			
	(0.000)	(0.000)	(0.007)	(0.001)	(0.000)			
Fixed effects	Year	Year	Year	Year	Year	Year	Year	Year
Observations	2,846,282	1,180,515	1,503,369	1,503,289	1,589,401	1,700,575	1,700,575	1,700,570
Log-likelihood	-1.910e+06	-842486	-5.386e+06	-2.897e+06	-761484	-5.515e+06	-1.410e+07	-1.970e+08
Prop. female	-0.043***	-0.013***	-0.237***	0.028***	-0.028***	-0.034***	-0.103***	0.001
	(0.002)	(0.003)	(0.047)	(0.008)	(0.002)	(0.008)	(0.018)	(0.003)
Constant	0.601***	0.558***	-1.227***	1.504***	0.675***			
	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)			
Fixed effects	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear
Observations	2,846,033	1,180,507	1,503,360	1,503,280	1,589,391	1,700,109	1,699,575	1,700,406
Log-likelihood	-1.759e+06	-820190	-5.373e+06	-2.885e+06	-718870	-4.929e+06	-1.160e+07	-1.360e+08

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a] estimated on the subset of patent applications with at least one final rejection
[b] estimated on the subset of issued patents for which we could obtain claim comparison data
[c] estimated on the subset of patents issued before 2012
[d] estimated on the subset of issued patents

Table S10: The table reports the estimated effect of the proportion of women inventors listed on a patent application on the processes of obtaining, maintaining, and asserting patent rights. The table is identical to Table S4, except that these models use a 90% cut-off in identifying the gender of inventors (Table S4 used 95%).

Using a continuous coding of gender. We also explored using the average probability of inventors being women based on their forename. In other words, instead of coding "Peter" as a one, it would receive a score of .003 to reflect the proportion of women with that name. This approach allows us to include every name that matched any of the three databases. Table S11 reports the results of this analysis, which yields nearly equivalent effect sizes.

Sample	**Patent issued**	**Request for reexam**[a]	**Independent claim count change**[b]	**Independent claim word length change**[b]	**Maintained**[c]	**Examiner cites**[d]	**Applicant cites**[d]	**Days between application and issuance**[d]
Prop. female	-0.110***	-0.025***	-0.779***	0.033***	-0.036***	-0.152***	-0.228***	0.023***
	(0.007)	(0.004)	(0.120)	(0.004)	(0.004)	(0.005)	(0.016)	(0.009)
Constant	0.609***	0.561***	-1.231***	1.502***	0.671***			
	(0.001)	(0.000)	(0.014)	(0.000)	(0.000)			
Fixed effects	Year	Year	Year	Year	Year	Year	Year	Year
Observations	3,833,730	1,583,690	2,029,912	2,029,822	2,127,275	2,281,242	2,281,242	2,281,234
Log-likelihood	-2.578e+06	-1.130e+06	-7.647e+06	-3.955e+06	-1.017e+06	-7.471e+06	-1.860e+07	-2.600e+08
Prop. female	-0.058***	-0.032***	-0.210***	0.024***	-0.028***	-0.027***	-0.216***	-0.010***
	(0.002)	(0.002)	(0.038)	(0.007)	(0.002)	(0.007)	(0.016)	(0.002)
Constant	0.602***	0.561***	-1.297***	1.503***	0.670***			
	(0.000)	(0.000)	(0.004)	(0.001)	(0.000)			
Fixed effects	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear
Observations	3,833,403	1,583,680	2,029,901	2,029,811	2,127,263	2,280,772	2,280,241	2,281,057
Log-likelihood	-2.378e+06	-1.102e+06	-7.632e+06	-3.939e+06	-959914	-6.660e+06	-1.560e+07	-1.800e+08

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a] estimated on the subset of patent applications with at least one final rejection
[b] estimated on the subset of issued patents for which we could obtain claim comparison data
[c] estimated on the subset of patents issued before 2012
[d] estimated on the subset of issued patents

Table S11: The table reports the estimated effect of the expected percentage of women inventors listed on a patent application on the processes of obtaining, maintaining, and asserting patent rights. The table is identical to Table S4, except that these models use continuous values for the gender assignment of inventors, ranging from 0 (inventor with a forename that has only been given to men) to 1 (inventor with a forename that has only been given to women), rather than a dichotomous coding.

Including teams for which we could not assign genders to all inventors. Requiring a coding on gender for all inventors drops a large number of cases for which we could code gender for at least one inventor. We therefore ran a set of analyses that excluded inventors with non-matched forenames from both the numerator and denominator of our gender ratio measure. In other words, we calculated the proportion women just among those with matched forenames. All of our results hold in this alternate sample and the effect sizes vary only slightly from those reported in the primary analyses (see Table S12).

	Patent issued	**Request for reexam**[a]	**Independent claim count change**[b]	**Independent claim word length change**[b]	**Maintained**[c]	**Examiner cites**[d]	**Applicant cites**[d]	**Days between application and issuance**[d]
Prop. female	-0.115***	-0.011***	-1.000***	0.060***	-0.039***	-0.206***	-0.077***	0.056***
	(0.007)	(0.003)	(0.112)	(0.007)	(0.005)	(0.006)	(0.020)	(0.013)
Constant	0.607***	0.562***	-1.268***	1.504***	0.671***			
	(0.001)	(0.000)	(0.008)	(0.000)	(0.000)			
Fixed effects	Year	Year	Year	Year	Year	Year	Year	Year
Observations	3,593,673	1,499,121	1,909,815	1,909,728	2,001,447	2,146,695	2,146,695	2,146,687
Log-likelihood	-2.410e+06	-1.068e+06	-7.243e+06	-3.732e+06	-952359	-7.057e+06	-1.790e+07	-2.460e+08
Prop. female	-0.039***	-0.016***	-0.251***	0.035***	-0.022***	-0.023***	-0.096***	0.004
	(0.002)	(0.002)	(0.033)	(0.008)	(0.002)	(0.007)	(0.016)	(0.002)
Constant	0.601***	0.562***	-1.323***	1.506***	0.670***			
	(0.000)	(0.000)	(0.002)	(0.001)	(0.000)			
Fixed effects	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear
Observations	3,593,368	1,499,111	1,909,804	1,909,717	2,001,435	2,146,227	2,145,713	2,146,513
Log-likelihood	-2.221e+06	-1.042e+06	-7.229e+06	-3.717e+06	-898211	-6.274e+06	-1.490e+07	-1.710e+08

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a] estimated on the subset of patent applications with at least one final rejection
[b] estimated on the subset of issued patents for which we could obtain claim comparison data
[c] estimated on the subset of patents issued before 2012
[d] estimated on the subset of issued patents

Table S12: The table reports the estimated effect of the proportion of women inventors listed on a patent application on the processes of obtaining, maintaining, and asserting patent rights. The table is identical to Table S4, except that these models use all patent applications for which we could identify the gender of at least one of the inventors (excluding non-matched individuals from both the numerator and denominator of our gender ratio measure).

Restricting the data to U.S. inventors. Table S13 reports the results restricting the analysis to patent applications with U.S.-based inventors (36.03% of all applications have only U.S.-based inventors). We explored the importance of this restriction for two reasons. First, although all of the patent applications have been handled by the USPTO, biases on the side of the applicant could vary from country to country. Second, our gender assignment algorithm, being based in large part on Social Security data, has greater fidelity for U.S. applicants. We found that our approach identifies the gender of inventors from English-speaking and Western countries better than inventors from non-English speaking and non-Western countries. We could successfully determine the gender of 89% of U.S.-based inventors, 87% of Canadian inventors, and 94% of the inventors form the UK and Australia. Our approach to inferring gender becomes noticeably worse for countries that do not use the Latin alphabet, such as China (39%) and Japan (71%). When converting their names to the Latin alphabet, the applicants from these countries may have used non-conventional conversions. Chinese inventors may also list their surnames first (instead of listing their forenames first).

	Patent issued	Request for reexam[a]	Independent claim count change[b]	Independent claim word length change[b]	Maintained[c]	Examiner cites[d]	Applicant cites[d]	Days between application and issuance[d]
Prop. female	-0.146***	-0.016***	-1.059***	0.106***	-0.051***	-0.191***	-0.224***	0.067***
	(0.005)	(0.005)	(0.151)	(0.009)	(0.005)	(0.009)	(0.022)	(0.012)
Constant	0.618***	0.543***	-1.493***	1.513***	0.693***			
	(0.000)	(0.000)	(0.010)	(0.001)	(0.000)			
Fixed effects	Year	Year	Year	Year	Year	Year	Year	Year
Observations	1,324,002	579,807	660,312	660,279	754,762	802,851	802,851	802,862
Log-likelihood	-895292	-415803	-2.554e+06	-1.229e+06	-360083	-2.792e+06	-8.249e+06	-1.020e+08
Prop. female	-0.055***	-0.018***	-0.274***	0.067***	-0.031***	-0.041***	-0.213***	0.011***
	(0.003)	(0.004)	(0.086)	(0.012)	(0.003)	(0.010)	(0.020)	(0.003)
Constant	0.611***	0.544***	-1.544***	1.516***	0.692***			
	(0.000)	(0.000)	(0.006)	(0.001)	(0.000)			
Fixed effects	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear
Observations	1,323,914	579,801	660,304	660,271	754,754	802,290	801,959	802,676
Log-likelihood	-823052	-399853	-2.549e+06	-1.220e+06	-338765	-2.470e+06	-6.532e+06	-7.010e+07

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a] estimated on the subset of patent applications with at least one final rejection
[b] estimated on the subset of issued patents for which we could obtain claim comparison data
[c] estimated on the subset of patents issued before 2012
[d] estimated on the subset of issued patents

Table S13: The table reports the estimated effect of the proportion of women inventors listed on a patent application on the processes of obtaining, maintaining, and asserting patent rights. The table is identical to Table S4, except that these models have been estimated on the subset of patent applications with inventors living in the U.S.

Restricting the data to patent applications that are not in process. Although the year of application fixed effects should account for differences in the processing time, to determine whether applications that have been in the process for a long time might differ from other applications, Table S14 reports the results restricting the analysis to patent application no longer in-process as of December 2014.

	Patent issued	**Request for reexam[a]**	**Independent claim count change[b]**	**Independent claim word length change[b]**	**Maintained[c]**	**Examiner cites[d]**	**Applicant cites[d]**	**Days between application and issuance[d]**
Prop. female	-0.140***	-0.022**	-1.001***	0.076***	-0.048***	-0.216***	-0.089***	0.060***
	(0.003)	(0.007)	(0.110)	(0.009)	(0.007)	(0.011)	(0.020)	(0.013)
Constant	0.693***	0.529***	-1.164***	1.499***	0.677***			
	(0.000)	(0.001)	(0.007)	(0.001)	(0.000)			
Fixed effects	Year	Year	Year	Year	Year	Year	Year	Year
Observations	2,363,247	944,351	1,425,515	1,425,439	1,508,816	1,613,836	1,613,836	1,613,832
Log-likelihood	-1.530e+06	-676375	-5.123e+06	-2.649e+06	-723011	-5.224e+06	-1.340e+07	-1.870e+08
Prop. female	-0.048***	-0.015***	-0.245***	0.037***	-0.029***	-0.036***	-0.113***	0.001
	(0.002)	(0.003)	(0.050)	(0.008)	(0.002)	(0.008)	(0.018)	(0.003)
Constant	0.686***	0.529***	-1.211***	1.501***	0.676***			
	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)			
Fixed effects	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear
Observations	2,363,067	944,344	1,425,506	1,425,430	1,508,806	1,613,346	1,612,826	1,613,664
Log-likelihood	-1.418e+06	-658178	-5.110e+06	-2.636e+06	-682568	-4.672e+06	-1.100e+07	-1.290e+08

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a] estimated on the subset of patent applications with at least one final rejection
[b] estimated on the subset of issued patents for which we could obtain claim comparison data
[c] estimated on the subset of patents issued before 2012
[d] estimated on the subset of issued patents

Table S14: The table reports the estimated effect of the proportion of women inventors listed on a patent application on the processes of obtaining, maintaining, and asserting patent rights. The table is identical to Table S4, except that these models have been estimated on the subset of patent applications that no longer in-process as of December 2014.

Robustness within the subset of large assignees

One of the primary concerns in comparing men and women on these outcomes is that their inventions differ in some way difficult for the researcher to observe, most notably in their quality. Our approach of using rare names gets around this issue by assuming that the rarity of a forename affects the odds that a patent examiner or other party external to the assignee could infer the gender of the inventor but that the rarity of a forename does not predict that quality of the patent – or more precisely that any association between forename rarity and patent quality would not be different for men versus for women.

An alternative approach to controlling for patent quality relies on the fact that large organizations typically have internal processes for assessing the value of patenting an invention. One would therefore expect less heterogeneity in the quality of the patent applications filed by these entities. This approach, however, also has a disadvantage. Those within the inventor's organization are likely aware of the gender of the inventors and they may bring their own implicit or explicit biases into play when determining which inventions should get invented. In other words, the applications from these large entities have probably already been subject to some selection based on the gender of the inventors. Estimates of gender differences within this subsample, therefore, essentially provide some indication of whether the patent examiners and other external parties, such as future patent applicants exhibit *more* bias than these organizations do in their internal processes.

Table S15 reports estimates using only applications filed by large assignees. As one would expect, the gender differences appear smaller within this subset. The data do not allow us to determine whether this decline in the magnitude of the effects stems from the fact that these applications vary less in their quality relative to the population versus from the possibility that women inventors within these organizations suffer from disparate support for their inventions from their employers. But the fact that gender differences persist even within this subset lends further confidence to the conclusion that unobserved quality differences across patent applications cannot account for these disparities.

	Patent issued	Request for reexam[a]	Independent claim count change[b]	Independent claim word length change[b]	Maintained[c]	Examiner cites[d]	Applicant cites[d]	Days between application and issuance
Prop. Female	-0.123***	0.016***	-0.987***	0.066***	-0.044***	-0.216***	-0.000	0.075***
	(0.008)	(0.004)	(0.116)	(0.012)	(0.008)	(0.011)	(0.023)	(0.017)
Constant	0.650***	0.579***	-1.040***	1.490***	0.690***			
	(0.001)	(0.000)	(0.007)	(0.001)	(0.001)			
Fixed effects	Year	Year	Year	Year	Year	Year	Year	Year
Observations	1,997,744	1,997,744	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Log-likelihood	-1.285e+06	-1.285e+06	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Prop. Female	-0.035***	0.001	-0.305***	0.023***	-0.023***	-0.024**	-0.055**	0.006
	(0.002)	(0.003)	(0.061)	(0.008)	(0.003)	(0.010)	(0.022)	(0.004)
Constant	0.644***	0.580***	-1.081***	1.492***	0.689***			
	(0.000)	(0.000)	(0.004)	(0.000)	(0.000)			
Fixed effects	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear
Observations	1,997,618	1,997,618	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Log-likelihood	-1.182e+06	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a] estimated on the subset of patent applications with at least one final rejection
[b] estimated on the subset of issued patents for which we could obtain claim comparison data
[c] estimated on the subset of patents issued before 2012
[d] estimated on the subset of issued patents

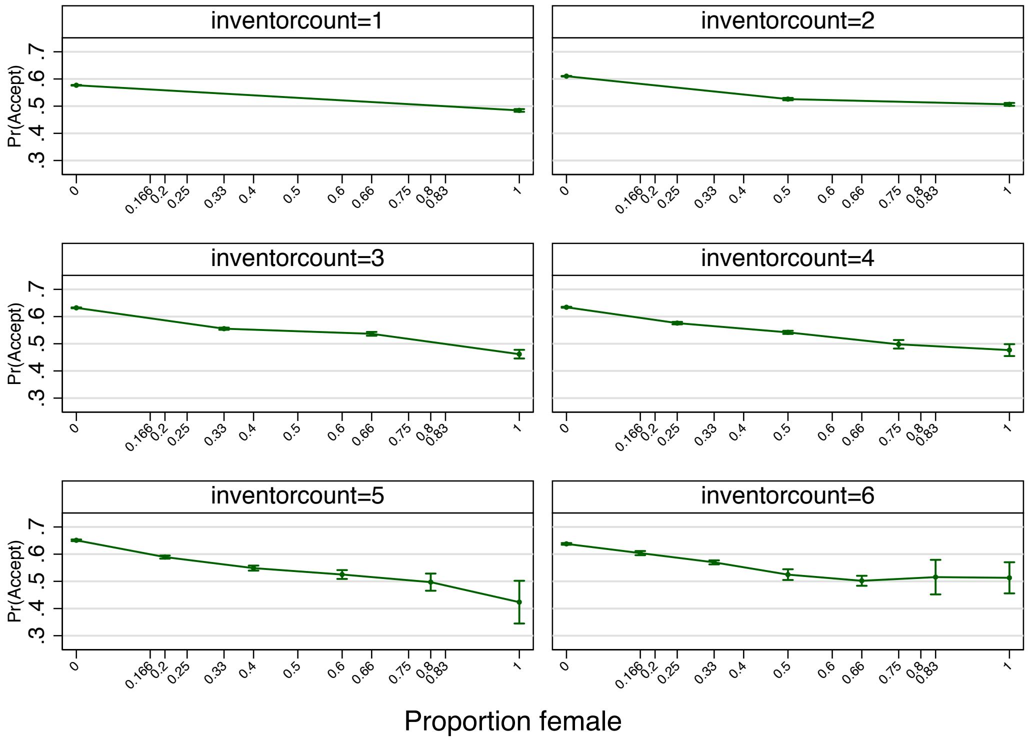
Table S15: The table reports the estimated effect of the proportion of women inventors listed on a patent application on the processes of obtaining, maintaining, and asserting patent rights. The table is identical to Table S4, except that these models have been estimated on the subset of patent applications from large assignees (i.e., assignees that do not qualify for the “small entity” classification of the USPTO).

Functional form of gender effects

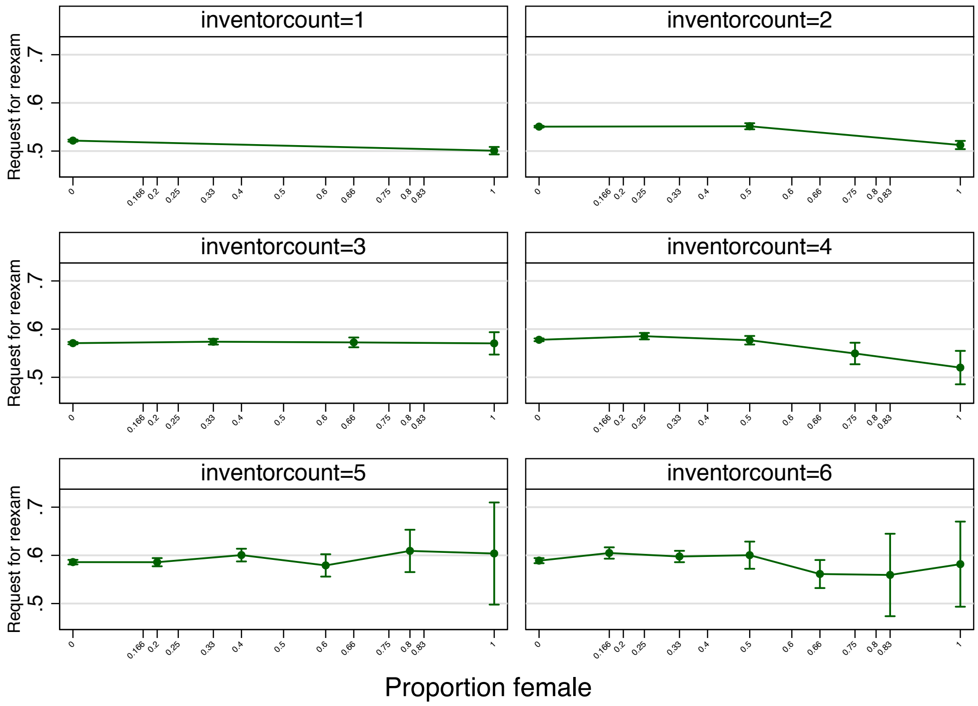
In the main set of models, we measured team composition with a "proportion of female inventors on the inventor team" variable, effectively assuming that this proportion has a linear relationship to the outcomes of interest. But perhaps a single woman on the team has a disproportionate effect. We therefore assess potential deviations from this assumption.

Non-parametric estimates of the proportion women. Fig. S3(a)-S3(f) explores the sensitivity of this assumption of linearity. We split the sample according to the number of inventors (to avoid confounding team size and gender composition effects). For each subsample (i.e., for each possible team size), we re-estimated the models in Table S4. But we estimated the effects non-parametrically, including a dummy variable for the number of women on the list of inventors and estimating a marginal effect for each level. This approach allows us to answer questions such as: "Is the presence of one man enough to save an application team?" or "Is the presence of one women enough to hurt a whole team?" and "Do the effects become stronger above some threshold?" As the figures below illustrate, however, the answers to these questions would generally be, no – in no instance can one reject the null that gender composition has a monotonic effect on every outcome. It generally appears linear.

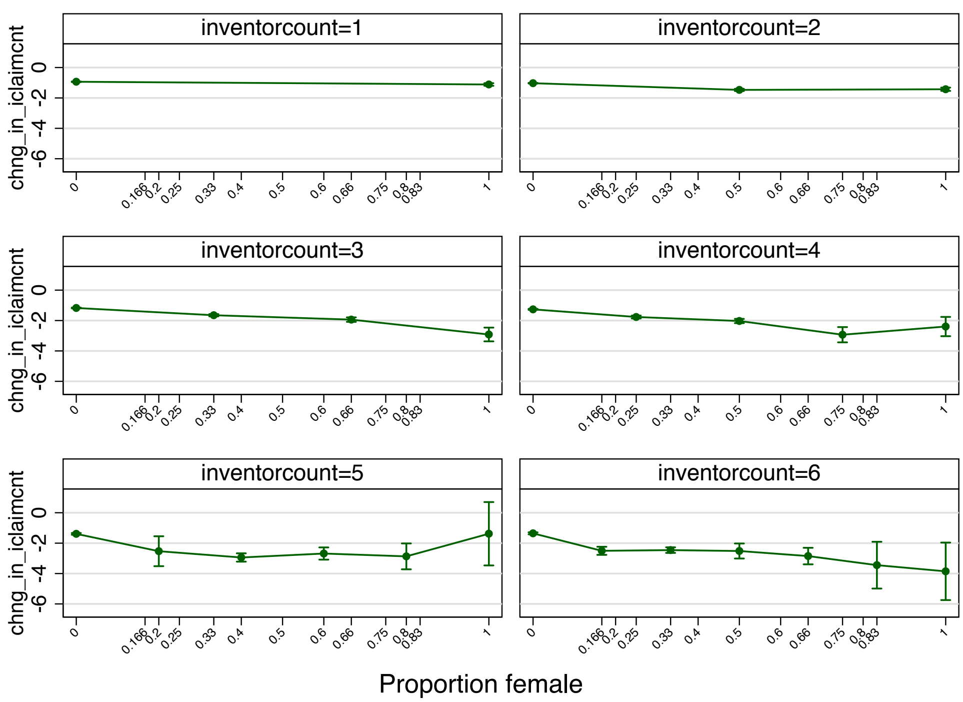
(a) Testing for non-linear effect of inventor team's gender composition on acceptance rate.



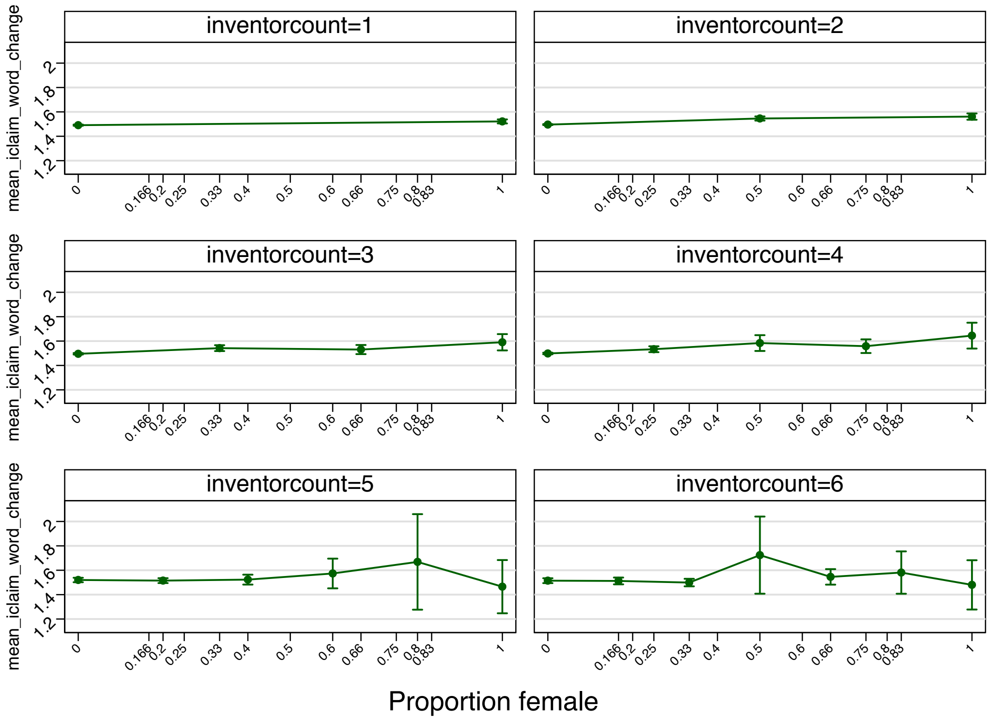
(b) Testing for non-linear effect of inventor team's gender composition on the probability that the inventors request examination after rejection.



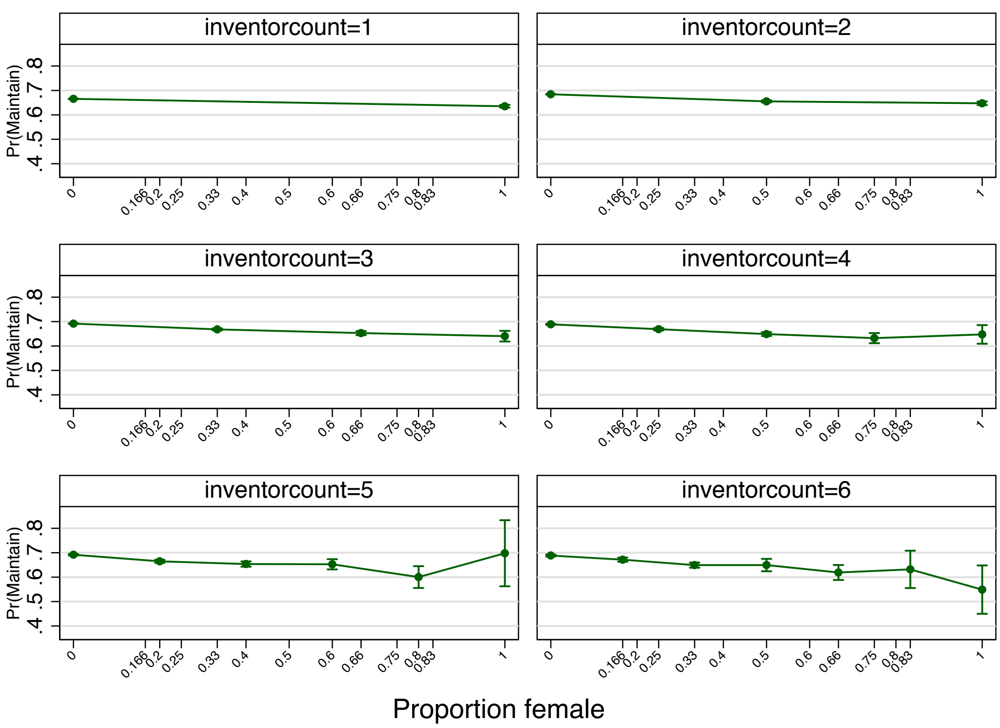
(c) Testing for non-linear effect of inventor team's gender composition on the change in independent claim counts from applications to the granted patent.



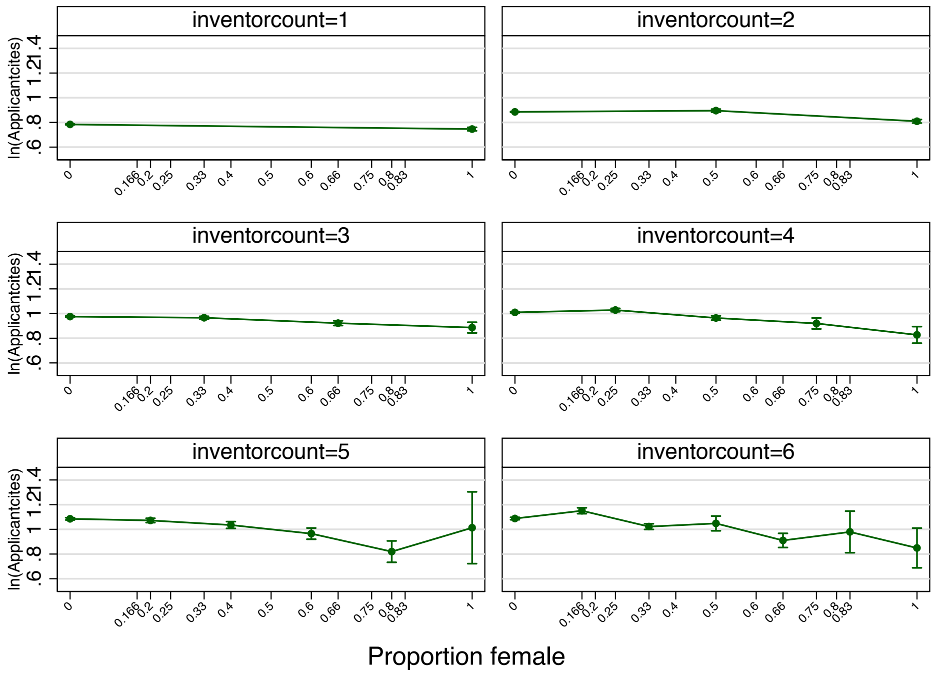
(d) Testing for non-linear effect of inventor team's gender composition on the proportional change in the average word counts of independent claim from applications to the granted patent (where claims with fewer words tend to reflect more general claims).



(e) Testing for non-linear effect of inventor team's gender composition on the probability that the inventors pay the maintenance fee.



(f) Testing for non-linear effect of inventor team's gender composition on the (logged) count of citations the accepted patent receives in the future (inventor added citations).



(f) Testing for non-linear effect of inventor team's gender composition on the (logged) count of citations the accepted patent receives in the future (examiner added citations).

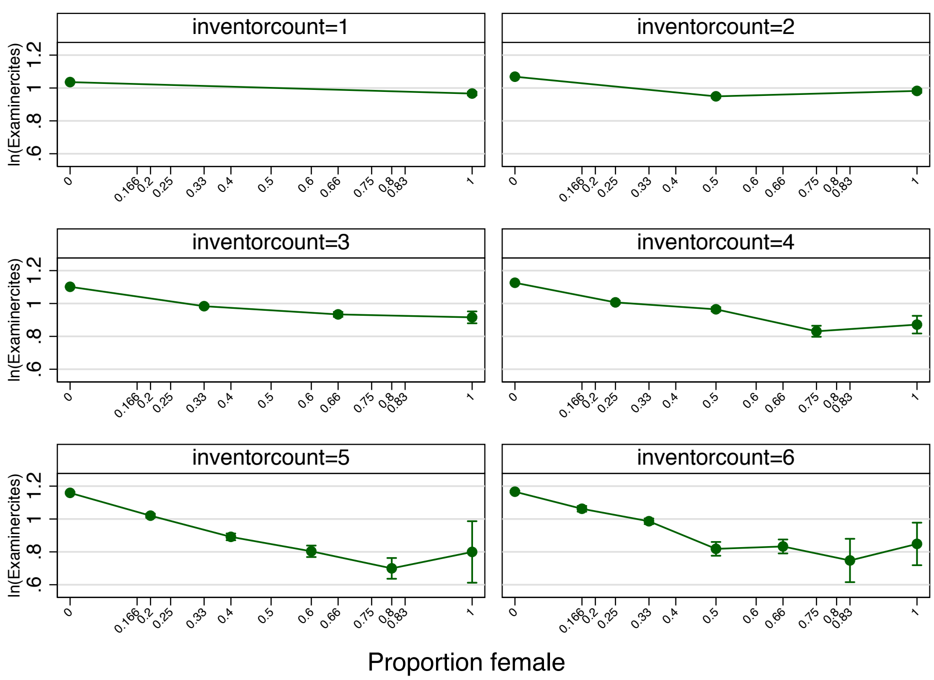


Fig. S3. The panels above display the estimated marginal effects of female inventor counts on various outcome variables, separately for various inventor counts.

Investigating disproportionate effects of the first listed inventor. Even if the proportion women, on average, has linear effects, the gender of the first listed inventor might prove unusually influential to examiners and others encountering the patent. To test this possibility, we included fixed effects for every possible gender combination (e.g., two men and one woman), and estimated net of these composition effects whether the gender of the first listed author matters.

Although all inventors listed on a patent share the same property rights, teams of inventors may nonetheless order individuals in terms of their contributions to the patent. To address that potential endogeneity in who gets listed first, we also estimated these effects on the subset of inventions in which the inventors had been listed in alphabetical order of their surnames. Within that set, whether or not the patent lists a woman as the first author ends up being random. Table S16 reveals that the gender of the first author has almost no measurable effect beyond the extent to which that author contributes to the gender composition of the team of inventors.

	Patent issued	Request for reexam[a]	Independent claim count change[b]	Independent claim word length change[b]	Maintained[c]	Examiner cites[d]	Applicant cites[d]	Days between application and issuance
Female inventor	0.003	0.005*	0.009	0.013	-0.002	0.039	0.123	1.494
	(0.002)	(0.003)	(0.052)	(0.009)	(0.002)	(0.026)	(0.098)	(2.434)
Constant	0.515***	0.510***	-1.090***	1.497***	0.641***	2.694***	3.287***	1,066.031***
	(0.003)	(0.005)	(0.086)	(0.015)	(0.003)	(0.041)	(0.155)	(3.860)
Fixed effects	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear
Gender composition fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,698,579	1,117,706	1,425,506	1,425,430	1,508,806	1,613,824	1,613,824	1,613,820
Log-likelihood	-1.664e+06	-776019	-5.110e+06	-2.635e+06	-681989	-4.739e+06	-6.884e+06	-1.210e+07
Female inventor	-0.000	-0.003	-0.132**	0.019	-0.006	0.030	-0.057	9.437**
	(0.003)	(0.005)	(0.063)	(0.014)	(0.004)	(0.045)	(0.153)	(4.308)
Constant	0.528***	0.519***	-0.909***	1.490***	0.654***	2.874***	3.597***	1,058.972***
	(0.004)	(0.007)	(0.077)	(0.018)	(0.005)	(0.054)	(0.185)	(5.212)
Fixed effects	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear	ClassXyear
Gender composition fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,632,406	659,084	844,972	844,923	908,261	968,818	968,818	968,817
Log-likelihood	-1.013e+06	-458616	-2.680e+06	-1.433e+06	-421609	-2.803e+06	-3.997e+06	-7.231e+06

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[a] estimated on the subset of patent applications with at least one final rejection
[b] estimated on the subset of issued patents for which we could obtain claim comparison data
[c] estimated on the subset of patents issued before 2012
[d] estimated on the subset of issued patents

Table S16: The table reports the estimated effect of the gender of the first listed inventor on the processes of obtaining, maintaining, and asserting patent rights, *after* controlling for the proportion of women inventors listed on a patent application. The upper panel shows the results based on the sample of all patent applications for which we could identify with 95% accuracy the gender of all inventors; the lower panel shows the results for the subset of these applications that list the inventors in alphabetical order of their names.

Estimating the effects by decision round

We also explored whether the effects of gender in terms of being accepted on that round varied by round. In essence, this analysis disaggregated the eventual acceptance variable into ten discrete variables: The first has a value of one if the application received approval in its first round of consideration. Conditional on not being accepted, the second has a value of one if the application received approval in its second round of consideration. And so forth. Because each round conditions on not being accepted in the previous round, the size of the sample declines with each round and our standard errors increase due to the loss of statistical power. We do not consider rounds after the 10th round because the sample size becomes so small (subsequent rounds represent less than 0.01% of cases).

Overall, the analysis suggests that the differences between men and women in their acceptance rates grow from the first round to the subsequent rounds but that they remain relatively stable from the second round onward. These results seem consistent with women inventors possibly having somewhat less persistence in the process.

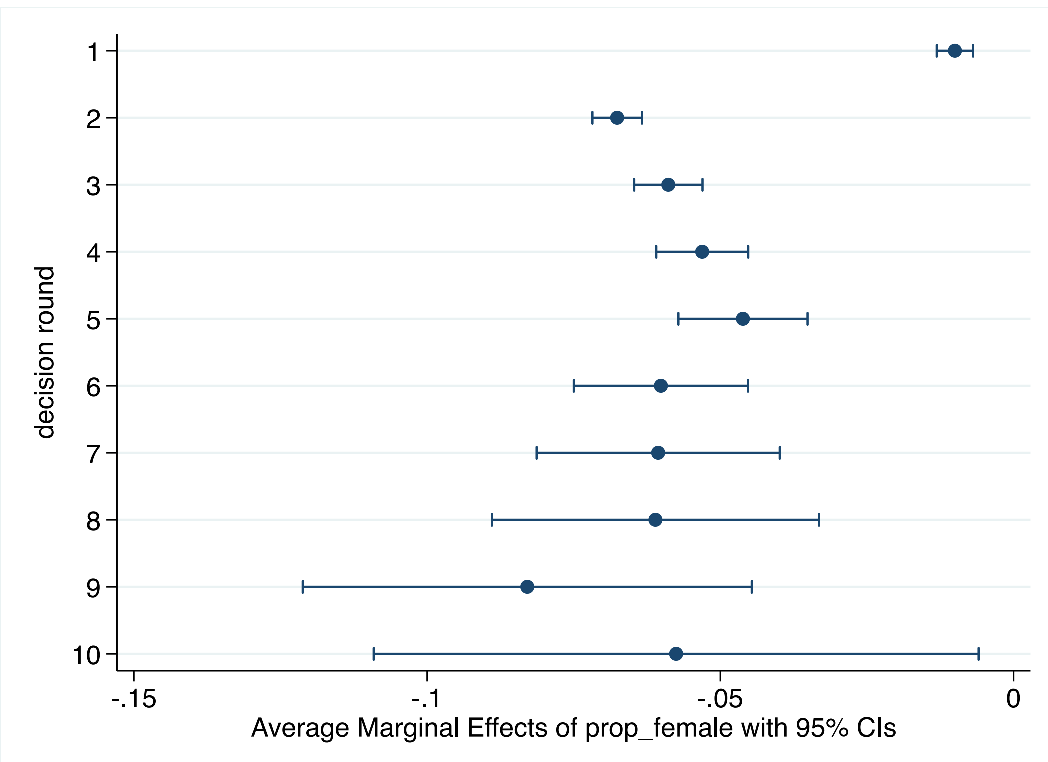


Fig. S4: Average marginal effect of the proportion of female inventors on acceptance rate at each decision round.

References in the Supplemental Materials

[S1] Bessen, J., Hunt, R.M., An empirical look at software patents. *J. of Econ. & Mgmt Strat.,* **16**, 157 (2007).

[S2] Thompson, P., Patent citations and the geography of knowledge spillovers: evidence from inventor-and examiner-added citations. *The Review of Economics and Statistics*, ***88***, 383-388. (2006)

[S3] Hall, B. H., Jaffe, A. B., & Trajtenberg, M., The NBER patent citation data file: Lessons, insights and methodological tools (No. w8498). *National Bureau of Economic Research* (2001).

[S4] Frakes MD, Wasserman MF. Is the time allocated to review patent applications inducing examiners to grant invalid patents?: Evidence from micro-level application data. *Review of Economics and Statistics*. (16 2014).

[S5] Angrist, J. D., & Pischke, J.-S., *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press: Princeton, NJ (2009).

[S6] Katz, E. Bias in conditional and unconditional fixed effects logit estimation. *Political Analysis*, ***9**,* 379-384 (2001).

[S7] Gourieroux, C., Monfort, A., & Trognon, A., Pseudo maximum likelihood methods: Applications to Poisson models. *Econometrica,* ***52***, 701-720 (1984).