Catalysts for Gender Inclusion in Innovation: The Role of Universities and their Top Inventors

Feb 9/2020

Authors: Mercedes Delgado[1]*, Fiona Murray[2]†

Affiliations:

[1]Copenhagen Business School and MIT Innovation Initiative

[2] MIT Sloan and NBER

*Correspondence to: md.si@cbs.dk

†fmurray@mit.edu

Abstract: Women's contribution to patenting is persistently low. We explore the catalysts for gender inclusion in innovation, focusing on factors that shape the presence of female inventors (including graduate students) at leading research universities. When universities promote new (first patent) female inventors, they facilitate women becoming productive inventors throughout their careers. We quantify gender inclusion in patents and show that universities have greater inventor inclusivity than the U.S. economy. However, the share of female new inventors in university patents is lower than the share of female STEM PhDs. Within universities, the presence of female new inventors is significantly higher in those patents with (versus without) a top (male or female) inventor. Further, female top inventors are key catalysts for increasing women's early participation in patenting.

One Sentence Summary: We analyze the role of leading research universities and their top academic inventors as catalysts for engaging women in patenting.

Main Text: 1995 words.

Main Text:

The innovation economy is driven by the transformation of ideas into solutions to pressing problems. To be effective, the most creative minds must be involved in innovation (*1*). Yet, while the rate at which women enter STEM education has increased, this has not translated into full participation in commercial science activities (*2*–*5*).

We examine the gender gap in STEM PhD participation and in patenting at universities. Although universities generate only about 4% of the patents in the United States, they train the majority of STEM students and generate many new (first patent) inventors each year. Universities play a critical role in shaping the skills and expectations of students. This is particularly important for innovation, given that attitudes toward research and patenting are shaped early in graduate experiences (*6*, *7*). If universities promote would-be female inventors, then these women will have the relevant skills and experience to become productive inventors during their careers. We study the role of leading research universities and their top inventors (patentees) as catalysts for increasing the participation of women in patenting.

U.S. Patenting and Female Inventor Inclusivity – A New Approach

To analyze women's inclusion in U.S. patents, we first assign the probable gender of each of the inventors on a patent using a name-gender match algorithm (*8*). This inventor-gender identification allows us to compute Female Inventor Inclusivity (FII) scores at multiple levels of analysis: for the entire pool of U.S. patents; at universities; and at the inventor-level for the relevant time period (2000-2015). We compute two scores: the *% Female Inventors* of total inventors; and the *% Female New Inventors* of total new (first patent granted) inventors (*8*).

To examine the role of universities as catalysts for inclusion, we focus on the top 25 universities filing patents in the United States. They generated 32,032 patents – over 50% of all university patents during our period (Table S1). We also produce a list of the top inventors (TIs) at each of these universities. During our period in the U.S. economy, the 90th percentile of the number of patents granted to inventor-organization pairs is seven (Table S2). Thus, we define TIs as having at least seven patents within a university during 2000-2015 (and explore other cut-offs, *8*). The 25 universities yield a total of 2,238 gender-matched TIs. The vast majority are faculty; for example, at MIT, there are 203 TIs of whom 82% are faculty and the rest are research scientists (MIT TLO).

The Persistent Inventor Gender Gap: Universities as Catalysts for Inclusion

There is a large and persistent inventor gender gap in the United States and elsewhere (*1*, *9*, *10*). The *% Female Inventors* was only 10% in the U.S. economy in 2015 (Table S1 and Fig. 1). This gap has been narrowing slowly, changing by 2 percentage points over 16 years (i.e., the share was 8% in 2000). The gap is also large for the *% Female New Inventors*: 14.3% of all U.S. new inventors were women in 2015.

At the current rate of change, parity in female inventors will be achieved by 2281 in the United States (Table S3)! This raises the central question of our paper: *How to accelerate gender inclusion in patenting?* We explore the role of leading research universities and their top inventors as catalysts for inclusivity in the innovation economy.

Based on our scores, universities are more inclusive than the U.S. economy, and their rate of change is faster (Table S1 and Fig.1). In 2015, the *% Female Inventors* in the top 25 universities was 16.9% (vs. 10% in the United States) and the *% Female New Inventors* was 22.3% (vs. 14.3%).

However, the patent composition in the U.S. economy and universities differs, and this explains part of the differences in inclusivity. In the top universities, *Drugs & Medical* is the largest technology class, with 45% of patents, versus only 15% of U.S. patents (Table S4). To account for this, we compute the inclusivity scores by technology class: the 25-universities are more inclusive for each class. We also build an "inclusivity index" (a weighted average of the technology-class scores) for the set of universities, each university, and each top inventor (*8*). We find that the 25-universities (and all-universities) inclusivity indices and their rate of change are positive. This confirms that universities are more inclusive and that their gap with the U.S. economy has widened.

Female Inventors versus Female STEM PhDs

Higher inclusion in universities is encouraging, but at the current rate it would take 169 years to reach parity in female inventors (Table S3). This raises the question of whether the limiting step lies in the supply of STEM PhDs – a pool of potential inventors who are critical to the research work of academic laboratories (*2*, *6*, *11*).

We know little about the participation of PhD students in patents. During our study period, the top universities generated 20,298 new inventors (Table S1) and our analysis suggests that many of them were graduate students when their first patent was filed. In a representative sample of about 800 new inventors, we found that 42% were students and 9% post-docs (Table S5). For the full sample of MIT patents in 2015, students represent 48% of new inventors and 26% of all co-inventors in a patent (*8*).

The large number of junior new inventors generated by universities suggests the long-lasting effect that universities can have on inventors' careers. But are these opportunities available equally to female and male PhDs? Existing research on the inventor gender gap has emphasized the lower selection of women into STEM fields (*1*). However, we find that the inventor gender gap in universities is higher than expected given the share of female STEM PhD holders (Tables S6-S7).

To study this question, we compute the STEM PhDs by gender granted by the top universities (*8*) and then compare patents granted in a given year (2000) to PhDs granted in that year and in the previous five years (1995-2000). Fig. 1 shows that while universities experienced increasing inclusion in patenting compared to the U.S. economy, the gap between the *% Female STEM PhDs* and *% Female New Inventors* remained large and increased: the *% Female STEM PhDs* was 1.5 times higher by 2015 (34% vs. 22%). Thus female PhDs are being included in university patenting at lower rates than they are completing PhD programs. We estimate that gender parity in the flow of STEM PhDs would take 22 years versus 103 years for new inventors.

Aggregate PhD statistics may hide field-specific differences. The *% Female STEM PhDs* granted by the 25-universities ranges from 19% in *Computer & Communications* to 54% in *Biological and Biomedical* sciences (Table S7). Focusing on the most inclusive field, we compare the *Biological and Biomedical* PhDs to the *Drugs & Medical* inventor data. There is still a significant gap: these patents have the highest % *Female New Inventors*, 29.5% in 2015, but this number is 24 percentage points lower than women's participation in PhDs.

During the full study period, each top university has inclusivity scores higher than the overall economy but much lower than the presence of women in their STEM PhD programs (Fig. S1). Overall, the rate at which young women PhDs engage in university patenting is about 62% that of their male counterparts.

The Role of Top Academic Inventors in Inclusion

Over 60% of female STEM PhDs work outside universities (Table S6), yet their patenting practices may be strongly shaped by their PhD training. Therefore, it is essential to understand and support the catalysts for inclusion within universities.

As principal investigators in grants, faculty members have a central role in training and mentoring that can spill over into later career behaviors (*3*, *11*). University patents are produced by a small number of top academic inventors (with 7+ patents): the 2,238 TIs represent only 6% of inventors, but contribute to 59% of our 25-university patents (Table S8). These individuals' autonomy, patenting intensity, and reputation give them a disproportionate role in shaping inventor inclusion.

Inclusivity of Patents with TI versus Patents without TI: We split our 25-university patents into those with TIs (a Female TI, a Male TI or Mixed-Gender TIs) and without TIs (Table S8). We compute inclusivity scores for each patent (excluding all the TIs). Our key finding is that the presence of new female inventors is statistically higher in those patents with TIs (female or male) than in patents without TIs (Fig. 2a). Patents with top inventors have a 1.8% higher percent of female new inventors, controlling for the team size, the technology class, university, and issued year. These findings are robust (*8*) and suggest that TIs are catalysts for increasing the presence of young female PhDs in patenting.

The Role of Female versus Male Top Academic Inventors in Inclusion

Do top inventors differ by gender in their levels of female inventor inclusivity? Prior work finds that homophily or cultural similarity may play a role in the formation of teams (*12*), especially among under-represented groups (*13*). In recent work, mentorship by women has been associated with selection and retention of women in STEM fields (*14*) and with female PhDs' publication productivity (*6*, *11*).

Female top inventors could attract and mentor more female PhDs and this is also true for patenting. In fact, we find that the *% Female New Inventors* is significantly higher for patents with only female TIs than with only male TIs: the gap is 5.9% (Fig. 2a).

To properly assess how top inventor gender affects inclusion, we next compare female versus male top inventors (FTIs vs. MTIs) for the presence of women co-inventors in their patent portfolio. Specifically, we compute our inclusivity scores for each TI across his/her patents in 2000-2015. Female TIs are statistically more inclusive on average for each of our scores (Table S12). The mean *% Female Inventors* (excluding the focal TI) is 21% among the 202 FTIs versus 15% among the 2,013 MTIs, a 6 percentage-point gap.

These results are also significant with regards to a TI's collaboration with female new inventors for whom mentorship is particularly salient. Here, we compute scores for each TI, excluding *all* TIs listed on the patents. The gap in the *% Female New Inventors* between FTI and MTI is larger (8 percentage points), suggesting that FTI mentorship is especially important for inclusion of new inventors (see Fig.3 by university). The mean *% Female New Inventors* is 30% for the FTIs and 22% for the MTIs, the latter score lower than the % *Female STEM PhDs* of 30% in the universities. On average, female TIs use the pool of young female PhDs as co-inventors more extensively.

These findings hold when we estimate the relationship between TI gender and inclusivity scores controlling for TI's university and main technology. In our specifications, FTIs have significantly higher inclusivity than MTIs, and that gap is larger for the presence of new women inventors. The estimated gap is 4.5 percentage points for *% Female Inventors* and 6.5 percentage points for *% Female New Inventors* (Fig. 2b). The gap is the same when using the inclusivity indices, and persists in fields with large presence of female PhDs (Fig. S3). These findings are very robust (*8*) and show that female top inventors are key mentors for women's early engagement in patenting. FTIs represent only about 9% of all top inventors but they serve as critical catalysts for inclusion.

Conclusion

Universities generate many new inventors (including graduate students) who go on to work in other organizations. Top inventors, and female top inventors in particular, are critical to incorporating the increasing but underutilized pool of young female STEM PhDs into patenting. To accelerate change in women's inclusion in patenting, we must recognize the critical role of universities. First, developing university inclusivity metrics could induce change. Second, initiatives that document and celebrate inclusive TIs and their best practices likely will shape inclusion. Finally, prior work shows that access to technology transfer offices and industry collaboration can increase female faculty patenting (*3*), suggesting opportunities for targeted interventions to increase the number of female top inventors.

References and Notes:

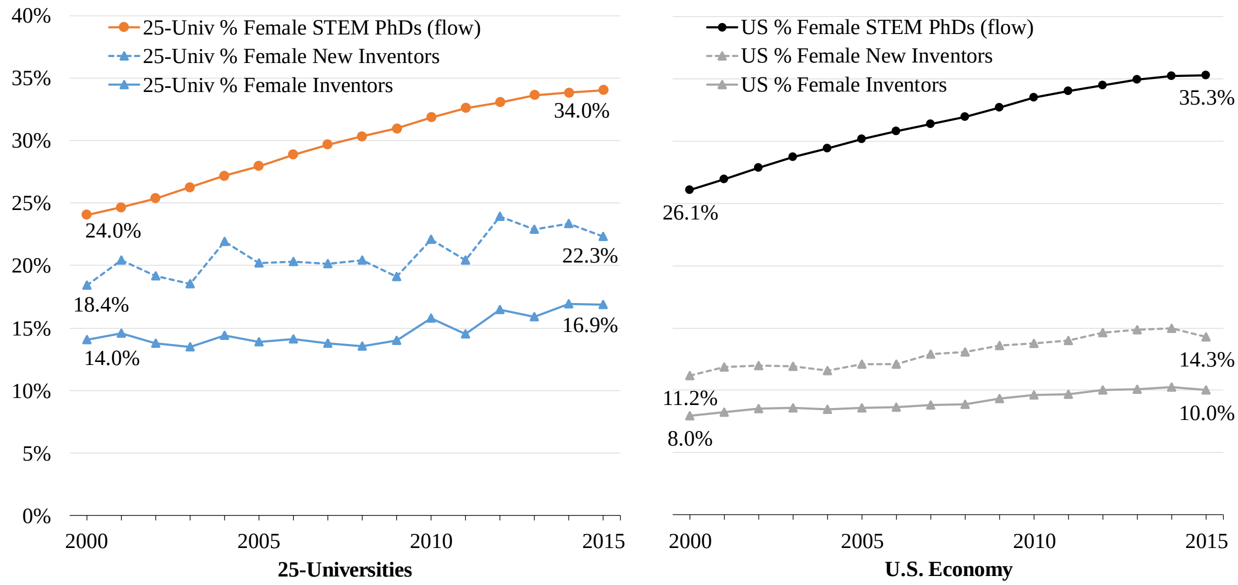
1. A. Bell, R. Chetty, X. Jaravel, N. Petkova, J. Van Reenen, Who becomes an inventor in America? The importance of exposure to innovation. *Q J Econ*. **134**, 647–713 (2019).
2. J. G. Thursby, M. C. Thursby, Gender patterns of research and licensing activity of science and engineering faculty. *J Technol Transfer*. **30**, 343–353 (2005).
3. W. W. Ding, F. Murray, T. E. Stuart, Gender differences in patenting in the academic life sciences. *Science*. **313**, 665–667 (2006).
4. J. Hunt, J.-P. Garant, H. Herman, D. J. Munroe, Why are women underrepresented amongst patentees? *Res Policy*. **42**, 831–843 (2013).
5. K. Jensen, B. Kovács, O. Sorenson, Gender differences in obtaining and maintaining patent rights. *Nat Biotechnol*. **36**, 307–309 (2018).
6. M. Pezzoni, J. Mairesse, P. Stephan, J. Lane, Gender and the publication output of graduate students: A case study. *PLOS One*. **11** (2016), doi:10.1371/journal.pone.0145146.
7. P. Azoulay, C. C. Liu, T. E. Stuart, Social influence given (partially) deliberate matching: Career imprints in the creation of academic entrepreneurs. *Am J Sociol*. **122**, 1223–1271 (2017).
8. Materials and methods are available as supplementary materials.
9. G. Lax Martínez, J. Raffo, K. Saito, "Identifying the gender of PCT inventors," *WIPO Economic Research Working Papers* (33, World Intellectual Property Organization - Economics and Statistics Division, 2016), (available at https://ideas.repec.org/p/wip/wpaper/33.html).
10. K. Hoisl, M. Mariani, It's a man's job: Income and the gender gap in industrial research. *Manage Sci*. **63**, 766–790 (2016).
11. P. Gaule, M. Piacentini, An advisor like me? Advisor gender and post-graduate careers in science. *Res Policy*. **47**, 805–813 (2018).
12. M. Ruef, H. E. Aldrich, N. M. Carter, The structure of founding teams: Homophily, strong ties, and isolation among U.S. entrepreneurs. *Am Sociol Rev*. **69**, 297–297 (2004).
13. M. McPherson, L. Smith-Lovin, J. M Cook, Birds of a feather: Homophily in social networks. *Annu Rev Sociol*. **27**, 415–444 (2001).
14. T. C. Dennehy, N. Dasgupta, Female peer mentors early in college increase women's positive academic experiences and retention in engineering. *P Natl Acad Sci USA*. **114**, 5964–5969 (2017).
15. B. H. Hall, A. B. Jaffe, M. Trajtenberg, "The NBER patent citation data file: Lessons, insights and methodological tools" (NBER Working Paper 8498, 2001), doi:10.3386/w8498.

Acknowledgements. We thank (especially) Peter Favaloro, Rich Bryden, Western Bonime, and Cory Ventres-Pake for their great research assistance with the data collection and visualization. We thank Lydia Snover for access to an anonymous list of first names and gender of students and applicants at MIT. We are grateful for the comments by Maryann Feldman, Michael Cima, Adam Jaffe, Shulamit Kahn, Myriam Mariani, Paula Stephan, Cathy Fazio, Patrick Gaule, Karin Hoisl, Valerie Karplus, Erin Scott, Stefan Sorg, Don Sull, and participants at the Geography of Innovation Conference 2018, Industry Studies Association Conference 2018, Innovation, Economic Complexity, and Economic Geography Workshop 2018, DRUID PDW on Female Scientists 2019, and seminars at the OECD and the Copenhagen Business School 2019. **Funding:** This project has been funded by a National Science Foundation, Science of Science and Innovation Policy Grant (Award #1757344). Authors are the PIs. **Competing interests:** Authors declare no competing interests.

List of Supplementary Materials:

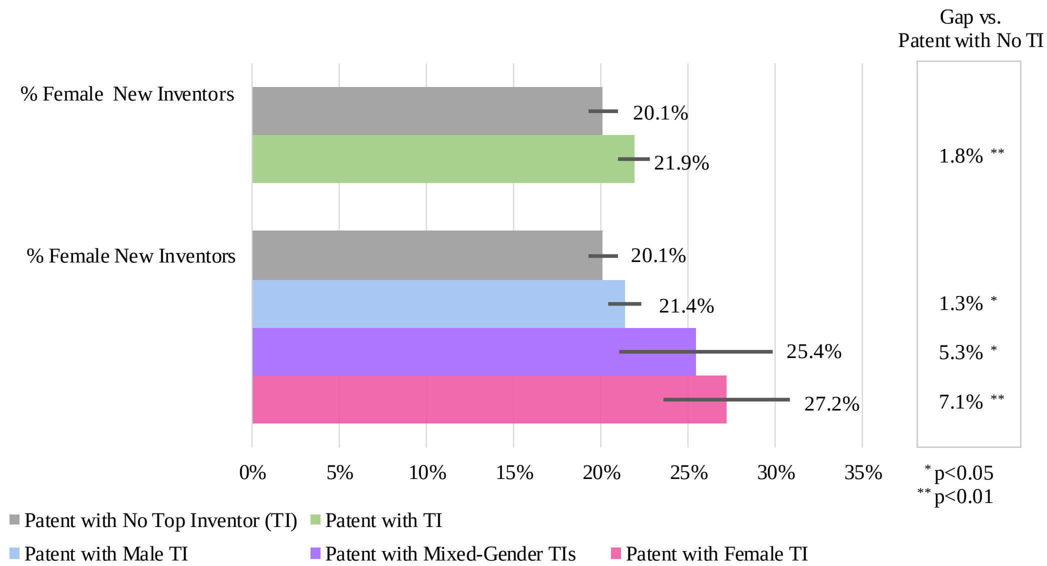
Materials and Methods
Fig. S1-S3
Tables S1 to S17
Reference (*15*)

Fig. 1.



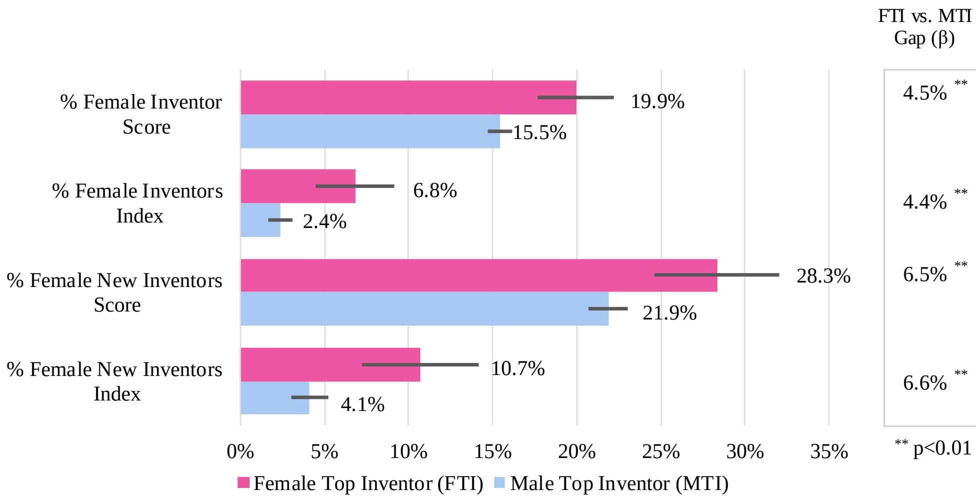
Trends in % Female STEM PhDs and % Female Inventors: 25-Universities vs. U.S. Economy. The PhDs data come from IPEDS and the NSF (*8*). Inventors in patents granted in yeart *t* are compared to PhDs granted in year *t* and the previous five years. This assumes that the PhD students at risk of filing a patent with their university are those 3-years prior to graduation to 2-years after graduation (i.e., for a patent granted in 2000, assuming it is filed 3 years earlier (1997), the PhD students are those graduating in 1995-2000).

Fig. 2.a



Patents with vs. without a Top Inventor: % Female New Inventors (25-Universities). Analysis of patents *p* granted to top universities during 2000-2015. Expected value (and 95% confidence intervals) of % *Female New Inventors* for each patent type (models control for Team Size, University, Tech. Class, and Granted Year; Tables S9-S10).

Fig. 2.b



Female vs. Male Top Inventors: Inclusivity Scores and Indices (25-Universities). Sample of Top Inventors (TIs) in 25-universities. Expected value (and 95% confidence intervals) of the inclusivity scores and indices of TIs by gender. The estimated scores control for a TI's university and main tech class (Table S14). The estimated indices (weighted average of inventor's tech-class sub-scores) control for a TI's university (Table S15).

Supplementary Materials for

Catalysts for Gender Inclusion in Innovation: The Role of Universities and their Top Inventors

Mercedes Delgado, Fiona Murray

This file includes:
Materials and Methods
Fig. S1-S3
Tables S1 to S17

Materials and Methods

University Sample: University-Patent Assignee Bridge

To create our university sample, we build a bridge to map USPTO patent assignee codes into individual universities. We identify the set of 201 universities with at least five patents in the 2011-2015 period (All Universities sample, and separate out the top-25 universities by patent count in the same period (25-Universities sample). The definition of the patents of a particular university is based on the first assignee listed in the patent. In our sample of 25-universities, about 90% of the patents granted (2000-2015) have only one assignee.

In the USPTO data, some universities assign their patents to a single entity, and thus are allocated a single assignee code, but others have multiple codes. For instance, it would be misleading to focus on the patent assignee "Cornell University," because the majority of Cornell University's patents were granted to the "Cornell Research Foundation" until 2013. Thus, we created a bridge to map patent assignee codes to universities. The process identifies 201 unique universities (i.e., institutions of higher education as defined by the National Science Foundation (NSF)) with at least five patents in the last five years, associated with 334 patent assignee codes. Sixty percent of the universities have only one assignee code, the remaining forty percent have 2-6 assignee codes.

To build the university-assignee bridge, we use the NSF HERD Survey 2015 to identify the top 200 institutions of higher education by STEM R&D expenses (i.e., STEM R&D is defined as every field in Table 18 of the HERD Survey except *Psychology*, *Social Science*, and *Non-S&E*). We supplement that list of universities by searching among all patent assignees with names that include keywords like "university" and "institute of technology" to identify any such assignees with significant patent activity. This supplemental search finds 76 additional assignees with at least five patents in the 2011-2015 period, 62 of which appear in the NSF lists of institutions of higher education. This gives us 262 NSF-listed institutions with either significant research expenditure or significant patent activity.

For each of the 262 universities, we generate candidate matches by searching for patent assignee names that share words with the focal university's name or its common abbreviations. A candidate's name either must share an unusual word (e.g. "Cornell") with the focal university, or else share a common word (e.g. "Ohio") and include a word that suggests university affiliation (e.g. "foundation"). These potential matches are then evaluated manually to determine whether they represent patent activity of the focal university. For example, the Cornell Research Foundation holds Cornell University's patents. Affiliated entities that share a location with the focal institution are included if they also share faculty or are likely to share patent assignee codes. Patents originating at university-affiliated medical centers and teaching hospitals are included in that university's patents, unless the medical center is affiliated with more than one university during our period. In some cases, such as the University of Massachusetts system or the Regents of University of California, multiple campuses use the same patent assignee code for a significant portion of their patents. In these cases, we aggregate the affected campuses into one entity. Thus the top 200 institutions by STEM R&D expenses, plus the 62 other institutions found in the assignee data, together yield 201 unique universities for our purposes. A table showing the assignee codes for each of the universities in our sample is available in the University-Patent Bridge Appendix.

STEM PhDs by Gender in the US Economy and by University

At the US-level, the NSF Survey of Doctoral Recipients offers data on the stock of STEM PhDs (Table S6), and the Survey of Earned Doctorates on the flow of STEM PhDs (Table S7). We define STEM to include the fields in Biological & Agr. Sciences; Physical Sciences; Mathematical Sciences; Computer, Communications & Info. Sciences; and Engineering. We excluded Social, Psychology, or Health Sciences since these fields are less likely to patent. The STEM stock variables are bi-annual.

At the university-level, the Integrated Postsecondary Educational Data System (IPEDS) offers annual data on PhD completions by field and gender for US institutions of higher education. We use institution names to associate IPEDS institutions (identified by "unitIDs") with our list of NSF-USPTO defined universities. We use the same keyword-matching process as in the main university bridge and manually remove false positives.

Our definition of STEM PhDs includes the following National Center for Education Statistics (NCES) fields that are more likely to patent: 01, 02, 03, 10, 11, 14, 15, 26, 27, 40, and 41. In the year-2000 classification these are named, respectively: *Agriculture, agriculture operations, and related sciences* (01); *Agricultural sciences* (02); *Natural resources and conservation* (03); *Communications Technologies/Technicians and Support Services* (10) *Computer and information sciences and support services* (11); *Engineering* (14); *Engineering technologies/technicians* (15); *Biological and biomedical sciences* (26); *Mathematics and statistics* (27); *Physical sciences* (40); and *Science technologies* (41). (See Table S7 and Fig. 1).

Name-Gender Match Algorithm

We used data on the U.S. Social Security Administration enrollees (SSA) and MIT students and applicants (MIT) to associate first names with genders. The SSA data lists the frequency of forename by gender and year (*5*), 1916-2016 (https://www.ssa.gov/oact/babynames/limits.html). We have access to similar data on MIT students and applicants from 1996-2016 (about 628,000 people with first names and self-reported gender). Combining both samples we have created a list of over 111,000 unique first names and their frequencies of being female/male names. The gender-name algorithm (underlying programs) can be accessed in the Online Appendix.

The SSA data includes 95,900 unique first names. The MIT data allows us to validate the gender distribution by name in a different population, and also adds a further 15,700 unique first names. (Another 49,000 unique names in the MIT data occur only once, and are excluded from our analysis for reasons described below.) The two data sources agree substantially on gender distribution: 86% of overlapping names are categorized into the same gender by both sources. When the two data sources disagree, we use the data where the forename occurs more frequently.

Most first names are strongly associated with either the male or female gender in the SSA and MIT data. To gauge our precision in estimating gender for rare names, we perform a cross-validation analysis within the MIT data. For instance, among first names that only occur twice in the MIT data, 85% are associated with the same gender in both instances. Thus, we focus on MIT unique names that occur with at least two instances.

The first names in the Social Security data have standardized capitalization, no punctuation, and no multi-word names (i.e., no spaces). We create a program to "clean" the MIT and USPTO first names to match this standard. The SSN database only includes first names with at least five instances for a given gender and year, so we cannot distinguish a count of four from a count of zero. Our analysis treats all zeros as true zeros.

We estimate an inventor gender only when 80% or more of the individuals with that first name are of a single gender. Using this approach about 91% of the U.S. inventors and 86% of university inventors were matched to one gender. For top academic inventors with first names with less than 90% frequency of a given gender, we searched online to identify their gender. This resulted in 2,238 gender-matched TIs and only 10 TIs with unmatched gender (Table S2).

Female Inventor Inclusivity Scores: All Inventors and New Inventors by Gender

The definition of inventor in the paper is organization specific (i.e., an individual with patents in two universities counts as two inventors). The inventor id is sourced from the USPTO PatentsView' rawinventor.tsv data accessed in June 2017 (https://www.uspto.gov/ip-policy/economic-research/patentsview).

We compute two female inventor inclusivity (FII) scores based on the gender-matched inventors. First, the *% Female Inventors* (FIs) of total inventors (i.e., the number of FIs divided by all gender-matched inventors). This captures the presence of women in the pool of inventors. Second, the *% Female New Inventors* (FNIs) of total new inventors (i.e., the number of new (first utility patent granted) FIs divided by all gender-matched new inventors). This captures the presence of women in the flow of new inventors. We compute FII scores for the entire pool of U.S. patents, the set of all patenting universities, our set of 25-universities, each top university, and each top academic inventor.

Who are the new inventors? To illuminate the important point that universities can influence inventorship early on scientists' careers, we implemented a detailed analysis for Massachusetts Institute of Technology (MIT), a top university with 197 new inventors in 263 patents granted in 2015. This analysis used the roster of all MIT graduate students and faculty and their patents (data provided by MIT Institutional Research) to build the measures reported in the paper: graduate students make up about 48% of new inventors in 2015. The majority of these students filed their first patent during the 3rd or 4th year in the program, and their average age when they started their graduate program was about 24.

The large presence of PhD students on university patents is not exclusive to MIT. We randomly selected 5% of the 20,298 new inventors generated by the top universities (a representative sample of 1,050 inventors). For each university, we sampled 42 new inventors: 21 during 2000-2009 and 21 after 2009, because in the full sample 50% of the new inventors were generated in the 2010-2015 period. For the random sample, the distribution across the six tech classes is very similar to that of the entire sample of new inventors (Table S5). We then implemented an online search to assess their status when their first patent was filed. We identified the status of 805 of the inventors: of those, 42% were students (including a few that graduated from the (assignee) university up to two years before filing the patent) and another 9% postdocs.

Female Inventor Inclusivity Indices

The patent composition across technology classes of the U.S. economy versus universities is different (Table S4), and the supply of women in STEM varies by field too (Table S7). Furthermore, individual universities and their top inventors also may differ in their patent composition. Thus, we compute the inclusivity scores by technology class and then build an "Inclusivity Index", which is a weighted average of the six technology-class inclusivity scores. We compute the indices at three levels of analysis, the pool of the top universities *u25* and all universities *uAll* (Equations 1a and 1b), each of the top universities *u* (Equation 2), and each of the Top Inventors *iu* in these universities (Equation 3):

$$Inclusivity\ Index_{u25} = \sum_{tech} Share\ Patents_{u25}^{tech} * (Inclusivity\ Score_{u25}^{tech} - Inclusivity\ Score_{US}^{tech}) \tag{1a}$$

$$Inclusivity\ Index_{uAll} = \sum_{tech} Share\ Patents_{uAll}^{tech} * (Inclusivity\ Score_{uAll}^{tech} - Inclusivity\ Score_{US}^{tech}) \tag{1b}$$

$$Inclusivity\ Index_{u} = \sum_{tech} Share\ Patents_{u}^{tech} * (Inclusivity\ Score_{u}^{tech} - Inclusivity\ Score_{US}^{tech}) \tag{2}$$

$$Inclusivity\ Index_{iu} = \sum_{tech} Share\ Patents_{iu}^{tech} * (Inclusivity\ Score_{iu}^{tech} - Inclusivity\ Score_{US}^{tech}) \tag{3}$$

The index first normalizes each score relative to the U.S. by technology class (e.g, difference between the 25-universities and the US score); and then weighs each normalized score based on the share of patents in the technology class ($Share\ Patents^{tech}$).

Equation 1 allows us to compare the inclusivity of top-25 universities (and all-universities) to that of the U.S. economy (Table S4). For example, in 2015, the % *Female Inventor index* shows a 4 percentage point greater presence of women inventors in top-university patents, with a 1.5 percentage point change in the index during our period (indicating a faster change in the 25-universities vs. the U.S. Economy).

Equation 2 allows us to compare across individual universities that can vary in the technology class composition of their patents (e.g., MIT has many patents in Mechanical which is a field with fewer women). We find that the vast majority of universities have positive inclusivity indices, which confirms the greater inclusion of universities than the U.S. economy. (See Fig. S1 for the *% Female New Inventors* by university).

Finally, Equation 3 allows us to compare across top inventors who specialize in different technology classes and/or patent in multiple classes (Tables S12 and S15 and Fig. 2b).

Patent-Level Models: Inclusivity of Patents Without vs. With Top Inventors

Top inventors (TI) are those with at least seven patents granted within a particular university during 2000-2015 (i.e., 90^{th} percentile in patenting proficiency in the U.S. economy; Table S2). We chose this cut-off (7+ patents) for two main reasons: we want a baseline definition of TI that applies to firms as well as to universities, and we look for potential catalyzers – those who have many patents and so could work with many co-inventors. In the sensitivity analysis, we use a lower cut-off value of 5+ patents (i.e., the 90^{th} percentile inventor-university patenting in the 25-universities).

Our definition of TI is not technology-class specific because many TIs patent in several classes, even if they concentrate the majority of patents in one (Table S11). Furthermore, the patenting proficiency of inventors by tech class is similar (slightly lower) to that of all patents. The 90^{th} percentile value of inventor patenting ranges from 4 to 6 patents across tech classes (4 patents for *Mechanical* and *Others* and 6 patents for *Drugs & Medical* and *Computers & Comm*.) Thus, the baseline definition of TI (7+ patents) and the additional lower cut-off (5+ patents) will capture well TIs across technology classes during the relevant period (2000-2015).

We compare patents with TIs to patents without TIs. The estimated patent-level model is as follows:

$$\begin{aligned}\text{Inclusivity Score}_p^{\textit{Exclude all TIs}} = \ & \beta \text{Patent with TI}_\text{p} + \lambda \text{Patent-Team Size}_p^{\text{Exclude all TIs}} + \\ & \delta + \delta_{iyear} + \delta_{tech} + \delta_u + \varepsilon_p .\end{aligned} \tag{4}$$

The dependent variables are the presence in each patent *p* of new female inventors (*Female New Inventors Share*, excluding all TIs). The key explanatory variable is a dummy equal to one for patents with at least a TI (*Patent with* TI_p) and zero otherwise. The model includes the size of the team of (non-top) inventors in the patent (*Patent-Team* $Size_p$), and dummies for the issued year (δ_{iyear}), technology class (δ_{tech}) and university (δ_u) of the patent.

We also want to compare the inclusivity of patents without top inventors versus patents with top inventors of different gender (female, male, or mixed-gender). Similarly to Equation (4), the estimated patent-level model is as follows:

$$
\begin{aligned}
\text{Inclusivity Score}_p^{\textit{Exclude all TIs}} = \quad & \beta^F \text{ Patent with Female TI}_p + \\
& \beta^{Mix} \text{Patent with Mixed-Gender TIs}_p + \\
& \beta^M \text{ Patent with Male TI}_p + \\
& \lambda \text{Patent-Team Size}_p^{\text{Exclude all TIs}} + \\
& \delta + \delta_{tech} + \delta_u + \delta_{iyear} + \varepsilon_p .
\end{aligned} \tag{5}
$$

The key explanatory variables are four dummies for mutually exclusive types of patent based on whether the patent has at least one female TI (*Patent with Female TI$_p$*); top inventors of different gender (*Patent with Mixed Gender TIs$_p$*); at least one male TI (*Patent with Male TI$_p$);* and the omitted category of patents with zero TI (*Patent with No TI*). Table S8 shows the count of patents in each of the four types. This specification allows us to compare the inclusivity of patents with female or male top-inventors versus patents without top inventors (β).

We estimate the equations (4-5) using OLS. The results are reported in Tables S9-S10 and Fig. 2a. These findings of the statistically higher inclusivity of patents by TIs (and especially by FTIs) are robust to including patent filed year (versus issued year) dummies; university-technology-class pair dummies ($\delta_{u,tech}$); and to clustering the standard errors by Top Inventor and by patent to allow for correlation of the inclusivity scores of the set of patents of each TI (i.e., patents with TIs are repeated for each focal TI for double clustering). Furthermore, the findings are robust to using a larger sample of patents by any university with at least one FTI and one MTI (85 universities versus our baseline 25 universities); and to changing the definition of Top Inventors as those with at least five patents (versus seven patents) granted within a particular university.

<u>Top Inventor-Level Models: Inclusivity of Female versus Male Top Inventors</u>

Table S11 shows the key attributes of TIs by gender, such as patent-team size, main technology class, patenting proficiency during our study period, and "age" or tenure in patenting. This allows us to assess if there are attributes of TIs that vary by gender that could relate to female inventor inclusivity.

We estimate the relationship between the gender of top inventors and their female inclusivity scores after controlling for other observable attributes of the TI. The baseline model is as follows:

$$\text{Inclusivity Score}_{iu} = \beta Female_i + \delta + \delta_{tech} + \delta_u + \varepsilon_{iu} \tag{6}$$

The unit of observation is a focal top inventor *i* within the university *u*. The dependent variables are the inclusivity scores of the focal inventor computed across all his/her patents in 2000-2015 (excluding the focal TI or excluding all the TIs in own patents). The key explanatory variable is the gender of the top inventor (*Female* dummy). The baseline model includes dummies for the main technology class of the inventor's patents (δ_{tech}) and for universities (δ_u); or alternatively university-tech-class dummies ($\delta_{u,tech}$). The findings are reported in Table S14 and Fig. 2b.

We implement several sensitivity analyses. First, our findings are not driven because FTI and MTIs focus on different types of patents (proxy for research interest) that would induce women to self-select to work only with FTIs. To test this, we compare the distribution of patents across all 3-digit United States Patent Classification (USPC) subclasses *s* within each of the six technology classes (Table S13). The similarity measure is the correlation coefficient of the two vectors of count of patents across subclasses (Correlation(Patent with $FTI_{s,tech}$, Patent with $MTI_{s,tech}$). The distribution of Patents with FTI and Patents with MTI across subclasses are very similar (with a correlation coefficient of about 0.90 within each technology class).

Second, we compute the *Inclusivity Indices*$_{iu}$ (see Equation 3) for each TI and use them as alternative dependent variables. This allows us to control for the fact that many inventors patent in multiple technology classes (Table S11) and to better compare across TIs with different patent composition (Table S15 and Fig. 2.b). Our findings are robust.

Third, we have also implemented analysis by technology class and find that the FTI vs. MTI positive gap in the *% Female New Inventors* is statistically significant in fields with many STEM women where we expect less homophily: *Drugs & Medical* patents and *Chemical* patents (Table S16 and Fig. S3). FTIs in Drugs & Medical have the highest estimated % *Female New Inventors* at 36% (8% higher than MTIs), but still lower than the *% Female Biological & Biomedical PhDs* at 49%. This further suggests that the higher inclusivity of FTIs is not due to low supply of female STEM PhDs or positive discrimination induced by homophily.

Finally, we control for other attributes of TIs (Table S11 and S17): the count of patents during 2000-2015 (on average slightly smaller for FTIs: 12 vs. 13.6 patents); the “age” in patenting (i.e., number of years since the inventor’s first patent) because FTIs are more junior in patenting (on average their first patent was granted in 2000 vs. 1997 for MTIs); and the average patent-team size which is somewhat larger for FTIs (3.8 vs. 3.5 inventors). Our results are robust to controlling for the TI’s attributes (see Table S17).

Table S1.

	Granted Year	Patents	Female Inventors (FIs)	Male Inventors (MIs)	Female New Inventors (FNIs)	Male New Inventors (MNIs)	% FIs	% FNIs
US	2000-15	1,394,632	104,317	921,941	72,435	476,723	10.2%	13.2%
	2000	68,956	8,162	94,344	3,711	29,442	8.0%	11.2%
	2015	127,300	18,386	165,343	6,185	37,092	10.0%	14.3%
	Change $_{2000-15}$	58,344	10,224	70,999	2,474	7,650	2.0%	3.1%
25-Univ	2000-15	32,032	5,531	26,454	4,309	15,989	17.3%	21.2%
	2000	1,539	399	2,445	190	843	14.0%	18.4%
	2015	3,167	1,045	5,154	459	1,597	16.9%	22.3%
	Change $_{2000-15}$	1,628	645	2,709	269	754	2.8%	3.9%
All Univ	2000-15	59,105	10,579	50,436	8,182	30,248	17.3%	21.3%
	2000	3,023	756	4,707	364	1,616	13.8%	18.4%
	2015	5,683	1,903	9,150	851	2,839	17.2%	23.1%
	Change $_{2000-15}$	2,660	1,146	4,443	487	1,223	3.4%	4.7%

Female Inventor Inclusivity in the U.S. Economy and Universities, 2000-2015. This analysis uses utility patents of U.S. origin granted to organizations and their inventors located in the United States (sourced from the USPTO). We identify 201 universities (All Univ) and separate out the top 25 by patenting (25-Univ). The definition of inventor is organization specific (i.e., an individual with patents in two organizations counts as two inventors). An inventor is "new" if his/her first patent has been granted in the particular year or period. In our sample, 91% of U.S. inventors and 86% of university inventors have a matched gender.

Table S2.

	Patents	Inventors*	Patents by Inventor, 2000-2015					Top Inventors (7+ Patents)
			Median	75 Pctile	90 Pctile	95 Pctile	99 Pctile	
U.S. Economy	1,394,632	1,130,834	1	3	7	11	29	114,071
25-Universities	32,032	37,314	1	2	5	7	18	2,248
All Universities	59,105	71,749	1	2	4	7	16	3,707

Patenting Proficiency of Inventor-Organization, 2000-2015. The analysis uses patents of U.S. origin granted to organizations and their inventors located in the US (USPTO). * The definition of inventor is organization specific (i.e., an individual with patents in two organizations counts as two inventors). In the U.S. sample the organization refers to the main assignee code in the patent. In the university sample, we identify 201 individual universities (All Univ) and separate out the top 25 by patenting (25-Univ). The 90th percentile value of the number of patents granted to an inventor-organization is seven in the U.S. Economy and 5 in the 25-Universities. Thus, we define top academic inventors as those with 7+ patents (baseline definition) or 5+ patents (alternative definition) within a particular university during 2000-2015.

Table S3.

	% Female Inventors				% Female New Inventors				% Female STEM PhDs $_{t-5,\,t}$			
	Annual Change	Years to Parity	15-year Change	Years to Parity	Annual Change	Years to Parity	15-year Change	Years to Parity	Annual Change	Years to Parity	15-year Change	Years to Parity
U.S.	0.15%	266	2.0%	293	0.25%	140	3.1%	173	0.63%	23	9.2%	24
25-Univ	0.20%	169	2.8%	176	0.27%	103	3.9%	106	0.71%	22	10.0%	24
All-Univ	0.22%	151	3.4%	146	0.30%	88	4.7%	86	0.72%	21	10.2%	22

Number of Years to Reach Parity in Female Inventors and Female STEM PhDs Shares. Parity means that the score is 50%. The estimated annual change is the slope in the 2000-2015 annual trends reported in Fig. 1. The 15-year change in the inclusivity scores are reported in Table S1. The *% Female STEM PhDs* is the recent pool of PhDs granted (year t and prior 5 years; Table S7).

Table S4.

	2000			2015			2000-2015 Change	
	% Patents	% FIs	% FNIs	% Patents	% FIs	% FNIs	% FIs	% FNIs
U.S. Score	**100%**	**8.0%**	**11.2%**	**100%**	**10.0%**	**14.3%**	**2.0%**	**3.1%**
Chemical	15%	10.6%	15.1%	8%	12.6%	18.0%	1.9%	2.9%
Computers & Comm	24%	6.4%	9.6%	41%	8.6%	12.5%	2.2%	2.9%
Drugs & Medical	14%	15.1%	22.8%	15%	17.1%	27.5%	2.0%	4.7%
Electrical/Electronic	19%	4.8%	7.2%	17%	6.7%	9.8%	1.8%	2.6%
Mechanical	14%	3.7%	5.6%	9%	5.2%	7.7%	1.5%	2.1%
Other	14%	6.3%	9.8%	10%	8.4%	13.0%	2.1%	3.2%
25-Univ Score	100%	14.0%	18.4%	100%	16.9%	22.3%	2.8%	3.9%
Chemical	21%	13.9%	20.4%	16%	17.6%	20.8%	3.7%	0.4%
Computers & Comm	9%	7.8%	9.0%	14%	12.9%	15.5%	5.1%	6.5%
Drugs & Medical	45%	17.7%	23.6%	45%	20.5%	29.5%	2.8%	5.9%
Electrical/Electronic	15%	6.6%	8.4%	18%	10.6%	15.6%	3.9%	7.2%
Mechanical	6%	7.4%	13.0%	4%	9.4%	12.2%	2.0%	-0.9%
Other	3%	7.3%	8.1%	3%	14.6%	14.9%	7.4%	6.8%
25-Univ Index	**100%**	**2.5%**	**2.0%**	**100%**	**4.0%**	**3.1%**	**1.5%**	**1.1%**
All-Univ Score	100%	13.8%	18.4%	100%	17.2%	23.1%	3.4%	4.7%
Chemical	23%	13.6%	20.3%	18%	17.5%	23.0%	4.0%	2.7%
Computers & Comm	7%	6.8%	8.3%	12%	11.8%	15.0%	5.0%	6.7%
Drugs & Medical	48%	17.2%	23.1%	47%	20.8%	29.2%	3.7%	6.1%
Electrical/Electronic	13%	6.9%	10.2%	16%	11.0%	15.8%	4.1%	5.7%
Mechanical	5%	6.0%	9.3%	4%	9.4%	13.4%	3.4%	4.0%
Other	4%	6.9%	8.2%	4%	15.2%	17.5%	8.3%	9.4%
All-Univ Index	**100%**	**2.1%**	**1.8%**	**100%**	**4.1%**	**3.4%**	**2.0%**	**1.6%**

Inventor Inclusivity by Technology Class in the U.S. Economy and 25-Universities: 2000 vs. 2015. The six technology classes are from (*15*). The definition of inventor is organization-tech-class specific. The 25-University index is a weighted average of technology-classes scores (see Equation 1).

Table S5.

	All New Inventors		Sampled NIs		NIs Status in Patent's Filed Year					
					Identified		PhD Student		Postdocs	
	#	%	#	%	#	%	#	%	#	%
25-Univ, 2000-2015	20,295	100%	1,050	100%	805	100%	341	42%	70	9%
Chemical	3635	18%	174	17%	134	17%	60	45%	13	10%
Computers & Comm.	3361	16%	184	18%	149	19%	70	47%	5	3%
Drugs & Medical	8137	39%	419	40%	310	39%	101	33%	36	12%
Electrical & Electronic	3753	18%	176	17%	146	18%	74	51%	12	8%
Mechanical	1042	5%	50	5%	37	5%	21	57%	4	11%
Other	811	4%	47	4%	29	4%	15	52%	0	0%

Presence of Graduate Students and Postdocs among New Inventors: 25-Universities, 1990-2015. Analysis based on a representative random sample of 1,050 NIs (42 by university). For a subset of 805 inventors we found their work status the year they filed their first patent.

Table S6

	2000 Female #	Male #	Female %	2015 Female #	Male #	Female %	2000-2015 Female % Change
U.S. STEM PhDs Stock	78,870	367,760	17.7%	162,100	464,650	25.9%	8.2%
U.S. STEM PhDs in University Jobs	32,730	123,110	21.0%	60,400	145,600	29.3%	8.3%

STEM PhDs Stock by Gender in the U.S. Economy: 2000-2015. The U.S. STEM PhDs data come from the National Science Foundation (NSF): the Survey of Doctoral Recipients. PhDs "stock" refers to PhD holders residing in the United States who are less than 76 years of age. By 2015, the % of women in the U.S. STEM PhD stock (29.3%) is much higher than the *% Female Inventor*s in top universities (16.9%) and the U.S. Economy (10%).

Table S7

	1995-2015			1995-2000			2010-2015			
	Female (F) 000s	Male (M) 000s	F %	F 000s	M 000s	F %	F 000s	M 000s	F %	F % Change
U.S. STEM PhDs Flow	148.8	320.9	31.7%	29.9	84.7	26.1%	58.2	106.8	35.3%	9.2%
25-Univ STEM PhDs Flow:	55.6	129.2	30.1%	10.3	32.7	24.0%	22.7	44.0	34.0%	10.0%
Agriculture [a]	1.8	2.6	41.4%	0.4	0.8	32.4%	0.7	0.7	49.1%	16.7%
Biological & Biomedical [b]	23.5	24.2	49.2%	4.3	5.8	42.4%	9.9	8.6	53.6%	11.2%
Computer & Comm. [c]	1.9	8.8	17.8%	0.3	1.7	14.2%	0.8	3.4	19.2%	5.0%
Engineering Tech. [d]	13.4	56.2	19.3%	2.2	14.0	13.7%	5.7	19.2	22.7%	9.1%
Math & Statistics [e]	2.7	8.3	24.3%	0.6	2.1	22.0%	1.0	3.0	24.8%	2.7%
Natural Resources [f]	1.0	1.4	41.7%	0.1	0.3	29.9%	0.4	0.5	46.9%	17.0%
Physical [g]	11.3	27.7	29.1%	2.5	8.0	23.6%	4.2	8.6	33.0%	9.4%

STEM PhDs Flow by Gender in the U.S. Economy and 25-Universities: 1995-2015. The U.S. STEM PhDs data come from the National Science Foundation (NSF): Survey of Earned Doctorates. University-level STEM PhDs data come from the Integrated Postsecondary Educational Data System (IPEDS). PhDs "flow" refers to PhDs granted in a given period (year t and prior 5 years; 1995-2000 vs. 2010-2015). Our definition of STEM includes the following National Center for Education Statistics (NCES) fields: [a] 01-02; [b] 26; [c] 10-11; [d] 14-15 and 41; [e] 27; [f] 03; and [g] 40. See Supplementary Materials.

Table S8.

25-Universities 2000-2015	Patents	%	New Inventor Patents*	%
All Patents	32,032	100%	13,234	100%
Patents with TI	**18,956**	**59%**	**6,095**	**46%**
Patents with Female TI (No MTI)	952	3%	376	3%
Patents with Male TI (No FTI)	16,644	52%	5,456	41%
Patents with Mixed-Gender TIs	1,280	4%	250	2%
Patents with No TI	**13,076**	**41%**	**7,139**	**54%**

Patent Type: With or Without Top Inventors. The 25-universities have 18,956 patents granted with at least one TI (7+ patents). * *New Inventor Patents* are those with a gender-matched new inventor (after excluding all TIs).

Table S9.

	Y_p= Female New Inventors Share in a patent (exc. all TIs)				
	25-Univ TI (7+ Patents)			25-Univ TI (5+ Patents)	85-Univ TI (7+ Patents)
Patents by Universities, 2000-2015	1	2	3	4	5
Patent with TI	.018**	.018**	.020**	.024**	.026**
Patent-Team Size (exc. all TIs)	.007**	.007**	.006**	.006**	.008**
Issued Year FEs	Yes	Yes	Yes	Yes	Yes
Tech Class FEs	Yes	No	Yes	Yes	Yes
University FEs	Yes	No	Yes	Yes	Yes
Univ-Tech FEs	No	Yes	No	No	No
R-squared	.044	.054	.043	.045	.045
Obs. (Patents)	13,221	13,221	15,396	12,606	20,602
# Patent clusters			13,221		
# TI clusters			1,947		
Mean Dept. Variable Y_p	.209	.209	.209	.211	.213

Presence of Female New Inventors in Patents with TI vs. without TI (Universities, 2000-2015). Patent-level analysis (Equation 4): patent *p* granted to the 25-universities (Models 1-4) or 85-universities with a (female and male) TI (Model 5). Y_p=*Female New Inventors Share* (exclude all the TIs in the patent). Patents are grouped into two mutually exclusive categories: with TI and No TI (the omitted category). TI has 7+ patents (baseline) or 5+ patents (Model 4) granted within the university during 2000-2015. Model 3 clusters std.errors by TI and patent. **, * Coefficient is significant at 1% level.

Table S10.

	Y_p= Female New Inventors Share in a patent (exc. all TIs)				
	25-Univ TI (7+ Patents)			25-Univ TI (5+ Patents)	85-Univ TI (7+ Patents)
Patents by Universities, 2000-2015	1	2	3	4	5
Patent with Female TI	.071**	.072**	.073**	.085**	.088**
Patent with Mixed-Gender TIs	.053*	.051*	.049	.077**	.078**
Patent with Male TI	.013*	.013	.015*	.016*	.019**
Patent-Team Size (exc. all TIs)	.007**	.007**	.006**	.007**	.008**
Issued Year FEs	Yes	Yes	Yes	Yes	Yes
Tech Class FEs	Yes	No	Yes	Yes	Yes
University FEs	Yes	No	Yes	Yes	Yes
Univ-Tech FEs	No	Yes	No	No	No
R-squared	.045	.055	.045	.046	.046
Obs. (Patents)	13,221	13,221	15,396	12,606	20,602
# Patent clusters			13,221		
# TI clusters			1,947		
Dif. Patent with FTI vs. MTI	.059**	.060**	.058*	.069**	.069**
Mean Dept. Variable Y_p	.209	.209	.209	.211	.213

Presence of Female New Inventors in Patents with TI (FTI, MTI, Mixed TIs) vs. without TI (Universities, 2000-2015). Patent-level analysis (Equation 5): patent *p* granted to the 25-universities (Models 1-4) or 85-universities with a (female and male) TI (Model 5). Y_p=*Female New Inventors Share* (exclude all the TIs in the patent). Patents are grouped into four mutually exclusive categories: with FTI, Mixed-Gender TIs, MTI, and No TI (the omitted category). TI has 7+ patents (baseline) or 5+ patents (Model 4) granted within the university during 2000-2015. Model 3 clusters std. errors by TI and patent. **, * Coefficient is significant at 1% and 5%, respectively.

Table S11.

Attributes of TIs in 25-Universities	All TIs	Female TIs	Male TIs	FTI vs. MTI
	N=2,238	N=206	N=2,032	
	Mean	Mean	Mean	Diff. in Means
Patents 2000-2015$_{iu}$	13.4	12.0	13.6	-1.6*
Average Team Size $_{iu}$	3.5	3.8	3.5	0.3**
Average Team Size (exc. all TIs) $_{iu}$	1.8	1.9	1.8	0.1
Main Tech Class Chemical$_{iu}$	17%	21%	16%	5%
Main Tech Class Computers & Communications$_{iu}$	12%	7%	12%	-5%*
Main Tech Class Drugs & Medical$_{iu}$	43%	52%	43%	9%**
Main Tech Class Electrical & Electronic$_{iu}$	22%	16%	23%	-7%**
Main Tech Class Mechanical$_{iu}$	3%	2%	3%	-1%
Main Tech Class Others$_{iu}$	2%	0%	2%	-2%
Number of Tech Classes$_{iu}$	2.2	2.1	2.2	-0.1
Multiple Tech Classes$_{iu}$	73%	75%	72%	3%
% Patents in Main Tech Class$_{iu}$	76%	75%	76%	-1%
Year of First Patent Granted$_{i}$	1997	2000	1997	3**
Age in Patenting (count of years since 1st patent) $_{i}$	19	16	19	-3**
Inventor Before 2000 $_{i}$	57%	43%	58%	-15%**
Patents Granted Before 2000 $_{i}$	4.5	1.7	4.8	-3.1**

Attributes of Top Inventors by Gender. The definition of top inventor (TI) is those with at least seven patents granted within a particular university during 2000-2015. In the 25-universities, the percent of gender-matched top inventors is 99%. TI's age is based on the year of first granted utility patent (organization or individual). It is computed as of 2015 and truncated in 1976 (i.e., max age is 40). *Patents Granted Before 2000* counts utility patents granted to the individual or his/her organization. **, * Difference in mean of variables (FTIs vs. MTIs) is significant at 1% and 5%, respectively.

Table S12.

25-Universites, 2000-2015		Top Inventor (TI) Inclusivity Score/Index All Co-Inventors (exc. TI) TIs #	 **% Female Inventors (FIs)** Mean	 New Co-Inventors (exc. all TIs) TIs #	 **% Female New Inventors (FNIs)** Mean
All Patents	Female TIs	202	21.1%	177	29.8%
	Male TIs	2,013	15.3%	1,772	21.7%
	Dif. FTIs vs. MTIs		*5.7%***		*8.0%***
Chemical	FTIs	130	22%	75	35%
	MTIs	1,077	16%	661	23%
	Dif. FTIs vs. MTIs		*6%***		*12%***
Computers & Comm.	FTIs	39	12%	31	23%
	MTIs	585	9%	426	14%
	Dif. FTIs vs. MTIs		*2%*		*8%*
Drugs & Medical	FTIs	143	22%	111	35%
	MTIs	1,235	19%	942	28%
	Dif. FTIs vs. MTIs		*4%**		*7%**
Electrical & Electronic	FTIs	68	15%	49	13%
	MTIs	873	9%	689	13%
	Dif. FTIs vs. MTIs		*6%***		*0%*
Mechanical	FTIs	34	13%	21	17%
	MTIs	400	8%	236	12%
	Dif. FTIs vs. MTIs		*5%*		*5%*
Other	FTIs	21	16%	12	17%
	MTIs	272	11%	155	19%
	Dif. FTIs vs. MTIs		*6%*		*-1%*
Index	FTIs	202	6.8%	177	10.6%
(Tech-class weighted avg.)	MTIs	2,013	2.4%	1,772	4.1%
	Dif. FTIs vs. MTIs		*4.4%***		*6.5%***

Female vs. Male Top Inventors: Mean Inclusivity Scores and Indices (25-Universities). We compute the TI scores using the pool of own patents granted in a given university *u* during 2000-2015. We also compute TI scores by tech class and indices (weighted average across tech classes; Equation 3). TIs can patent in multiple technology classes. % FIs exclude the focal TI (missing if all patents are solo-inventor); and % FNIs exclude all TIs in the university to measure inclusivity among non-top collaborators of the TI (missing if all patents are co-invented only with TIs). **, * Difference in mean scores/indices of FTIs vs. MTIs is significant at 1% and 5% levels, respectively. Benchmark: The *% Female STEM PhDs* is 30%; *% Female Computer & Comm. PhDs* is 18%; and the *% Female Biological & Biomedical PhDs* is 49% (Table S7).

Table S13.

25-Universities **Tech Class$_c$**	Tech Subclasses$_s$ (3-digit USPC) Num	Patents with a Female TI (PFTI) Num	Patents with a Male TI (PMTI) Num	Similarity in Patenting across Subclasses Corr(PFTI$_{s,tech}$, PMTI$_s$)
Chemical	71	496	3579	0.87**
Computers & Comm.	41	155	2132	0.95**
Drugs & Medical	14	1143	6899	0.96**
Electrical & Electronic	46	315	3945	0.88**
Mechanical	51	73	932	0.96**
Others	38	43	514	0.94**
Total	**261**	**2225**	**18001**	

FTI vs. MTI Patents: Distribution across Technology Subclasess. TI (7+ patents). We separate *Patents with FTIs* versus *Patents with MTIs*, and compare their distribution across all 3-digit United States Patent Classification (USPC) subclasses *s* within each technology class. The similarity measure is the correlation coefficient of the two vectors of count of patents across subclasses: Corr(PFTI$_{s,tech}$, PMTI$_{s,tech}$). ** The correlation coefficient is significant at the 1% level.

Table S14.

	Y_{iu}=Top Inventor Inclusivity Scores							
	Female Inventors Share (Exc. the TI)				**Female New Inventors Share** (Exc. *all* TIs)			
	TI (7+ pat) 25-Univ		TI (7+ pat) 85-Univ	TI (5+ pat) 25-Univ	TI (7+ pat) 25-Univ		TI (7+ pat) 85-Univ	TI (5+ pat) 25-Univ
	1	2	3	4	5	6	7	8
Female TI	.045**	.037**	.034**	.037**	.065**	.060**	.069**	.076**
Tech Class FEs	Yes	No	Yes	Yes	Yes	No	Yes	Yes
Univ FEs (25)	Yes	No	Yes	Yes	Yes	No	Yes	Yes
Univ-Tech FEs	No	Yes	No	No	No	Yes	No	No
R-sq.	.087	.144	.102	.071	.095	.151	.127	.078
Obs. (TIs Count)	2,215	2,215	3,291	3,670	1,949	1,949	2,842	2,979
Mean Y_{iu}	.159	.159	.164	.154	.225	.225	.233	.228
E(Y/FTI)	.199	.192	.195	0.187	.283	.279	.295	.296
E(Y/MTI)	.155	.155	.160	0.151	.219	.219	.226	.221

Female vs. Male Top Inventors: Inclusivity Scores. TI-level analysis in 25-universities (Equation 6). The baseline definition of TI is 7+ patents granted, but we also explore 5+ patents (models 4 and 8). Y_{iu}=inclusivity scores of a TI's patents granted in the 2000-15 period, excluding the focal TI *iu* from own patents (FI models); and excluding all the TIs in the university from own patents (FNI models). *Tech Class* are dummies for the largest technology class of the TI's patents and *Univ-Tech* are dummies for university-tech class pairs. Findings are robust to using the larger sample of TIs in the 85 universities with 1+ female and male TI (Models 3 and 7). **, * Coefficient is significant at 1% and 5% levels, respectively.

Table S15.

	Y_{iu}=Top Inventor Inclusivity Indices					
	Female Inventors Share (exc. the TI)			**Female New Inventors Share** (exc. all TIs)		
	TI (7+ pat) 25-Univ 1	TI (7+ pat) 85-Univ 2	TI (5+ pat) 25-Univ 3	TI (7+ pat) 25-Univ 4	TI (7+ pat) 85-Univ 5	TI (5+ pat) 25-Univ 6
Female TI	.044**	.030**	.036**	.066**	.067**	.072**
University FEs	Yes	Yes	Yes	Yes	Yes	Yes
R-sq.	.022	.050	.014	.038	.075	.028
Obs. (TIs Count)	2,215	3,291	3,670	1,949	2,842	2,979
Mean Y_{iu}	.028	.030	.023	.047	.049	.051
E(Y/FTI)	.068	.057	.055	.107	.109	.115
E(Y/MTI)	.024	.027	.019	.041	.042	.044

Female vs. Male Top Inventors: Inclusivity Indices. Top Inventor (TI) level analysis in 25-universities. The baseline definition of TI is 7+ patents granted, but we also explore 5+ patents (models 3 and 6). Y_{iu}=Inclusivity index of a TI's patents granted during 2000-15, excluding the focal TI (models 1-3) or all TIs (models 4-6) from own patents. Index is defined in Equation (3). **, * Coefficient is significant at 1% level. Findings are robust to using the larger sample of TIs in the 85 universities with 1+ female and male TI (models 2 and 5).

Table S16.

	Chemical		Computer & Comm.		Drugs & Med		Electrical		Mechanical		Other	
	% FI 1	% FNI 2	% FI 3	% FNI 4	% FI 5	% FNI 6	% FI 7	% FNI 8	% FI 9	% FNI 10	%FI 11	% FNI 12
Female TI	.056**	.117**	.011	.065	.035	.079**	.056**	-.002	.052	.061	.051	.015
University FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	.029	.045	.047	.053	.033	.052	.035	.032	.088	.189	.095	.164
Obs.	1,207	736	624	457	1,378	1,053	941	738	434	257	293	167
Mean Y_{iu}	.165	.240	.096	.150	.191	.286	.095	.129	.086	.124	.110	.186
E(Y/FTI)	.216	.345	.107	.210	.222	.357	.148	.126	.134	.180	.157	.200
E(Y/MTI)	.159	.228	.095	.145	.187	.278	.091	.129	.082	.119	.106	.185

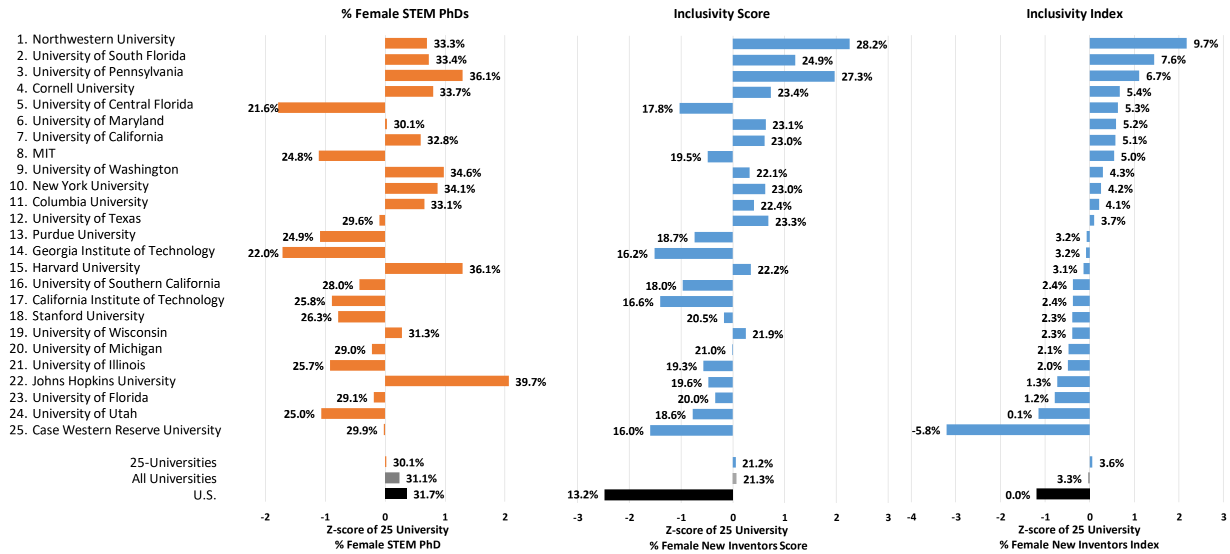
Female vs. Male Top Inventors: Inventor Inclusivity by Tech Class (25-Universities). TIs (7+ patents) can patent in multiple technology classes. We compute a TI's sub-scores by tech class: Female Inventors Share (% FI) and Female New Inventors Share (% FNI). Findings are robust to defining TI as those with 5+ patents during 2000-2015.

Table S17.

	Y_{iu}=TI Inclusivity Scores							
	Female Inventors Share (Exc. the TI)				**Female New Inventors Share** (Exc. *all* TIs)			
	1	2	3	4	5	6	7	8
Female TI	.045**	.046**	.045**	.046**	.064**	.062**	.064**	.061**
Ln $Patents_{00\text{-}15}$	-.002			-.004	-.009			-.006
Avg. Team Size		.000		.000		-.001		-.001
Age in Patenting			-.001	-.001			.004	.003
Tech Class FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
University FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-sq.	.087	.087	.087	.087	.095	.096	.095	.096
Obs.	2,215	2,215	2,215	2,215	1,949	1,949	1,949	1,949

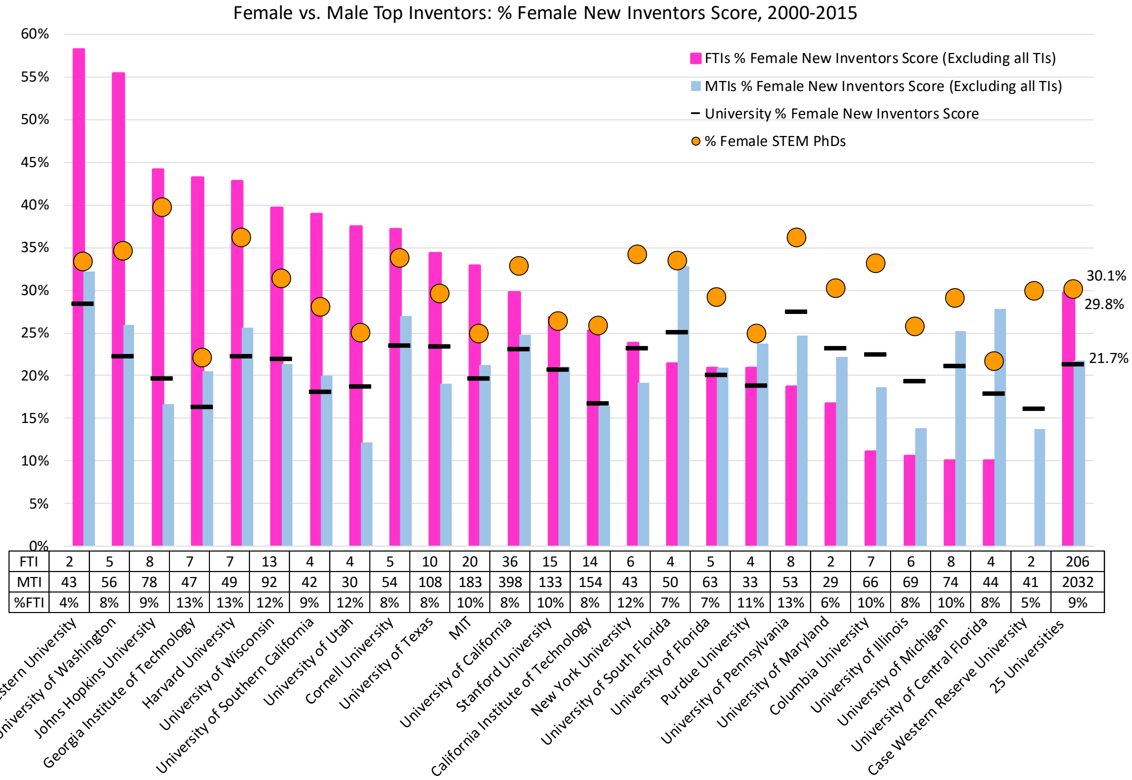
Female vs. Male Top Inventors in 25-Universities: Inclusivity Scores Controlling for TI's Attributes. Models control for a TI's attributes (Table S11): (log of) count of patents during 2000-2015(*Ln Patents*); Age in Patenting (i.e., number of years since the inventor's first patent) and average patent-team size (Avg. Team Size). ** Coefficient is significant at 1% level. The interaction effects of each of the control variables with the *Female TI* variable are statistically insignificant.

Fig. S1.



% Female New Inventors Score and Index by University, 2000-2015. The 25-University mean *% Female STEM PhDs* = 30%, *% Female New Inventors Score*=21.1%, and *% Female New Inventors Index*=3.5%. The inclusivity index is a weighted average of the university's tech-class scores (see Equation 2). The % Female STEM PhDs is based on PhDs granted during 1995-2015. We also report the indicators for the pooled PhDs and patents of the 25-Universities, All Universities, and the U.S. Economy. The correlation between % Female STEM PhDs and the inclusivity score is 0.65.

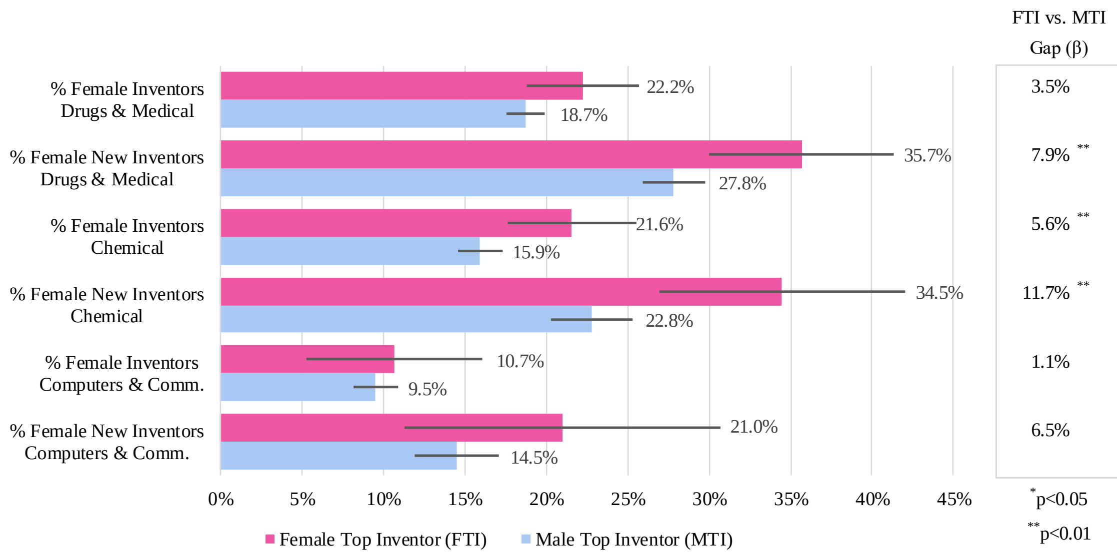
Fig. S2.



	Northwestern University	University of Washington	Johns Hopkins University	Georgia Institute of Technology	Harvard University	University of Wisconsin	University of Southern California	University of Utah	Cornell University	University of Texas	MIT	University of California	Stanford University	California Institute of Technology	New York University	University of South Florida	University of Florida	Purdue University	University of Pennsylvania	University of Maryland	Columbia University	University of Illinois	University of Michigan	University of Central Florida	Case Western Reserve University	25 Universities
FTI	2	5	8	7	7	13	4	4	5	10	20	36	15	14	6	4	5	4	8	2	7	6	8	4	2	206
MTI	43	56	78	47	49	92	42	30	54	108	183	398	133	154	43	50	63	33	53	29	66	69	74	44	41	2032
%FTI	4%	8%	9%	13%	13%	12%	9%	12%	8%	8%	10%	8%	10%	8%	12%	7%	7%	11%	13%	6%	10%	8%	10%	8%	5%	9%

Female vs. Male Top Inventors: Mean % Female New Inventors by University, 2000-2015. Mean % Female New Inventors score (exc. all TIs) across Female vs. Male Top Inventors (FTIs vs. MTIs) for each university and the set of 25-universities. The circles represent the *% Female STEM PhDs* by university (pool of 1995-2015 PhDs granted). The numbers below the bars are counts of FTIs and MTIs and the % FTIs by university. The median count of TIs is 61.

Fig. S3



Female vs. Male Top Inventors: Inclusivity Scores for Selected Tech Classes (25-Universities). Sample of TIs in 25-universities. Expected value (and 95% confidence intervals) of the inclusivity scores of TIs by gender and tech class. The estimated scores control for a TI's university (see Table S16).