Main Manuscript for

Faculty as Catalysts for Training New Inventors: Differential Outcomes for Male and Female PhD Students.

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

Affiliations:

[1]Copenhagen Business School and MIT

[2] MIT Sloan and NBER

*Corresponding author: Mercedes Delgado

Email: md.si@cbs.dk

Author Contributions: M.D. and F.M. designed the research and wrote the paper; M.D. implemented the empirical analysis and wrote the SI Appendix.

Competing Interest Statement: Authors declare no competing interests.

Keywords: Science of Innovation | STEM PhD students | Training in Commercial Science | New Inventors | Gender Diversity

This PDF file includes:

Main Text
Figures 1 to 5

Supporting Information Appendix (SI):

Materials and Methods: Sections S1-S5
Results: Sections S6-S7
Figures S1-S3
Tables S1 to S16

Abstract: STEM PhDs are a critical source of human capital in the economy, contributing to commercial as well as academic science. We examine whether STEM PhD students become new inventors (file their first patent) during their doctoral training at the top 25 U.S. universities (by patenting). We find that 4% of PhDs become new inventors. However, among PhDs of faculty who are themselves top (prolific) inventors, this figure rises to 23%. These faculty train 44% of all the new inventor PhDs by co-patenting with their advisees. We also explore whether new inventor PhDs are equally distributed by gender. In our university sample, the female share of new inventors is 9 percentage points (pp) lower than the female share of PhDs. Several channels contribute to this: First, female PhDs are less likely to be trained by top inventor advisors (TIs) than male PhDs. Second, they are less likely to be trained by (the larger number of) male top inventors: the estimated gap in the female % of PhDs between female and male top inventor advisors is 7-9 pp. Third, female PhDs (supervised by top inventors and especially by other faculty) have a lower probability of becoming new inventors relative to their male counterparts. Notably, we find that male and female top inventors have similar rates of transforming their female advisees into new inventors at 4-8 pp lower (17-26% lower rate) than for male advisees. The gap remains at 4 pp comparing students of the same advisor and controlling for thesis topic.

Statement of Significance: STEM PhDs increasingly contribute to commercial science, such as patenting. We analyze faculty's role in training STEM PhD students as new inventors on patents at leading research universities, emphasizing the drivers of gender differences. We show that faculty advisors who are themselves top inventors play a key role in training PhDs to be new inventors by co-patenting; the chances of patenting for students with a top inventor advisor are ten times higher than with other advisors. However, female PhDs have a lower likelihood of being matched with top inventor advisors. Furthermore, on average, male and female top inventor advisors transform their female advisees into new inventors at 4-8 percentage points lower rate than for male advisees.

Main Text:

The innovation economy is driven by the transformation of ideas into solutions to pressing problems. Highly trained human capital increasingly drives innovation (*1*, *2*). This underscores the national and global need for a strong science, technology, engineering, and mathematics (STEM) educational pipeline, and highlights training especially at the doctoral level (*3*).

Traditional approaches to doctoral education have emphasized academic science, with outputs such as publishing (*4*). However, about 60% of STEM PhDs in fields likely to lead to patenting (i.e., excluding psychology, social, and health sciences) work outside universities (*5*), and they are in increasing demand by employers from national labs to corporate research and startup ventures (*6*, *7*). With the shift in career paths, the policy emphasis on STEM as a source of competitiveness, and the fact that over 25% of inventors have PhDs (*2*), there is pressure for STEM PhDs to enter the economy with the skills needed to be productive contributors to inventive activity, not just traditional academic activities. Universities themselves are also emphasizing academic patenting as an important output (*8–10*).

The changing roles and expectations of STEM PhDs suggest that early experience in commercial science, such as patenting, during PhD training is vital (*11–13*). However, at present we have few insights into whether PhDs participate in patenting and become "new" (first time) inventors during their doctoral education. Moreover, it is important to understand whether opportunities for such participation are equally available, regardless of demographic background (*14–16*). There is reason for concern in this regard as the increasing rates of female entry into STEM education (including PhD training) have not translated into full participation in academic science (*4*, *17–21*) nor commercial science (*3*, *20*, *22–25*).

Our focus on patenting during PhD education provides a window into whether gender disparities in invention arise at the early, formative stage of training. This is particularly important because early success in innovation can build cumulative advantage to individuals over time (*11–13*, *10*). Furthermore, total career productivity of scientists is related to their career length (*19*), which begins with the first occurrence of research outcomes (e.g., first patent).

Universities are the critical organizational context for PhD training. Faculty advisors play an essential role in shaping the skills and attitudes toward publishing and patenting of their PhD advisees (*4*, *17*, *26*). We therefore examine the participation in patenting among STEM PhD students at the top 25 U.S. universities (by patenting) conditional on the characteristics of their faculty advisors. We determine the particular role of advisors who are themselves prolific inventors (referred to as "top inventors") as they may serve as catalysts for encouraging their PhD advisees to become new inventors through co-patenting. Our novel dataset and empirical approach allow us to examine potential differences in the rate of becoming new inventors among male and female PhD students supervised by different types of faculty: top inventors versus others; and female versus male top inventor advisors.

Our results illuminate the dynamics of patenting during PhD training and have implications for the design of individual and organizational policies to support the inclusive transformation of graduate STEM education for the twenty first century.

STEM PhDs and U.S. Patenting: Learning Commercial Science during Doctoral Training

In the U.S. economy, patenting is highly concentrated among a few organizations and a few prolific inventors within these organizations (*3*). Thus, to examine the extent to which PhD students are becoming new inventors during their doctoral training, we focus on PhDs trained at the top 25 universities by patent count, and inventors in their 2000-2015 granted utility patents (Section S1in the Supporting Information Appendix (SI) (*27*)). We consider PhDs who graduate during the 1995-2015 period in fields most likely to patent (*3*, *10*): Agricultural Sciences and Natural Resources, Biological and Biomedical Sciences, Computer and Information Sciences, Engineering, Mathematics and Statistics, and Physical Sciences (Section S2). We refer to this set of fields as "STEM" for the purpose of our analysis.

The 25-universities (including 40 individual campuses) account for half of all university patents (Tables S1 and S3), and train 41% of the STEM PhDs in the U.S. economy. They produce only 2% of the U.S. patents granted in 2000-2015 (Table S1), but 4% of *New Inventors* in the economy (i.e., individuals whose first ever granted patent is assigned to these universities). Such new inventors include graduate students, postdocs, faculty, and research staff. Our analysis focuses on one type of new inventor: PhD students. We define *New Inventor-PhDs* (NI-PhDs) as those who filed their first ever granted patent during their doctoral studies (or within two years after graduation), and the patent is assigned to their graduating university (See Equation 1 in Section S4).

We are especially interested in PhDs trained by top inventor faculty advisors (referred as TIs) and becoming new inventors by co-patenting with their advisors. We define TIs as those granted at least seven patents within a university during our full period (the 90th percentile value of the number of patents granted to inventor-organization pairs in the U.S. economy (Table S2)). In our sample these TIs account for 39% of all 25-universities patents (Table S3).

Our analysis determines the prevalence of new inventor PhD students. We examine whether the likelihood of PhDs becoming New Inventor-PhDs during their training is shaped by having a top inventor advisor. Importantly, we also explore how these likelihoods change with the gender of the PhD student and of the top inventor advisor.

Matching STEM PhDs and New Inventor-PhDs by Gender: New Approach

We construct a unique dataset to compare the population of STEM PhDs to the New Inventor-PhDs for the 25-universities and their top inventor advisors. First, we count the STEM PhD graduates by gender at the 25-universities (IPEDS data; Section S2 and Table S5). We focus on the almost 185,000 total PhD graduates in the 1995-2015 period (at risk of becoming inventors in the 2000-2015 granted patents). We then identify those PhD students among all the new inventors at our university patents by matching the ProQuest data (the largest repository of graduate theses) with the USPTO inventor data for the 32,032 university utility patents granted in the 2000-2015 period. Our set of New Inventor-PhDs includes 6,847 individuals in the university patents. (See Section S4 (*27*)).

To examine the role of top inventor advisors, we match the ProQuest and USPTO data to create a list of top inventor faculty who have trained at least one new inventor among their PhD students (Section S4). We identify 876 TIs who trained 13,286 gender-matched PhD graduates (1995-2015). We assume (and validate) that TIs and their advisees are in our STEM fields (Section S2). We recognize that students may be interested in different research topics (even within a given advisor lab), and that these topics may vary in their patentability. We therefore capture advisees' research topic using their thesis title (ProQuest) and transforming these into topics that are included as controls in some of our analysis (Section S5). The TIs have 3,189 New Inventor-PhDs among their advisees, and we focus on the 94% of them (3,008) who are co-patenting their first patent with their advisors.

Our exploration of patenting by PhD students at the 25-universities turns to the question of gender inclusivity. This analysis is grounded in our larger study of the presence of women in U.S. patents (*3*). We assign the probable gender of inventors on each patent using a name-gender match algorithm (Section S3). This inventor-gender identification allows us to create a novel inclusivity score to help us understand how universities and faculty advisors influence inventorship early on in female PhD students' careers: the *Female % of New Inventor-PhDs* (i.e., the percentage of women among (gender-matched) New Inventor-PhDs). We compute this score at multiple levels of analysis: the 25-university patents; each of these universities; the set of patents produced by the TIs; and for each of these advisors.

We also use our algorithm to determine the gender of the PhD advisees of the TIs (and use IPEDS data to count STEM PhDs by gender at our universities). This allows us to calculate the *Female % of PhDs* (i.e., the percentage of women among (gender-matched) PhDs) for the 25-universities, each university, the pool of TIs, and for each of these advisors. Finally, by matching PhDs and NI-PhDs by gender we can also compute the *Probability of New Inventor-PhDs* (i.e., the percentage of NI-PhDs among PhDs) for female and male PhDs at multiple levels of analysis.

Patenting during PhD Training

To date, we have limited understanding of the participation of PhD students in patenting. Based on our analysis we find that the 25-universities generate 6,847 New Inventor-PhDs – over 500 in 2015 alone (Fig. 1). PhD students represent 32% of all New Inventors (21,612) at universities: from 14% at Johns Hopkins University to 44% at MIT (Table S3). The PhD students' % of new inventors also varies across patent classes: from 22% in Drugs & Medical to 40% in Electrical and Electronic patents (Table S4). This analysis illustrates that for the 25-universities, PhD students are important new inventors, and (some) learn to patent during their training.

How prevalent has patenting by STEM PhD students become as part of PhD training? Overall, 3.7% of the STEM PhDs become new inventors (6,847 out of 184,865 PhDs) during our full period. In addition to the large number of New Inventor-PhDs identified in the set of 2000-2015 patents, we also examine annual trends (Fig. 1). We compute New Inventor-PhDs in patents granted in a given year (e.g., 2015) who graduated that year and in the previous five years (e.g., 2010-2015), and compare them to STEM PhD graduates in the same period (e.g., 2010-2015). The number of New Inventor-PhDs has increased rapidly, at 7% annual growth rate: from 212 in 2000 to 561 in 2015 (Fig. 1A). The supply of STEM PhD graduates has also increased, but more slowly, at 4% annual growth rate (Fig. 1B). Thus, over time universities are increasing the engagement of PhD students in patenting.

Female New Inventor-PhDs versus Female STEM PhD Graduates at Universities

Turning to the topic of gender inclusion, we examine the female share of New Inventor-PhDs at the 25-universities and compare this to the female share of all new inventors in the U.S. economy. Our analysis shows that universities are more inclusive than the U.S. economy (Table S4): women were 20.7% of New Inventor-PhDs in the university patents granted in 2000-2015 (1,419 out of 6,847) versus 13.1% (for new inventors overall) in the U.S. economy. The higher score for the universities holds across patent technology classes (using the classes defined by (*28*)). The *Female % of New Inventor-PhDs* also was increasing over our period, and by 2015 reached 25% in universities (Fig. 1C) versus 14.3% *Female % of New Inventors* in the U.S. economy (Table S1).

The higher female share of New Inventor-PhDs in universities is encouraging but remains far from parity. This raises the question of whether the limiting factor is the supply of female STEM PhD graduates – a pool of potential female inventors who may be critical to the research work of academic laboratories (*4, 17, 22*) and to the associated patenting.

To address this question, we compare the *Female % of New Inventor-PhDs* in patents granted in a given year (e.g., 2015) to the *Female % of PhDs* who graduated that year and the previous five years (e.g., 2010-2015) at the 25-universities during our study period. Fig. 1C shows a substantial gap: by 2015 the *Female % of New Inventor-PhDs* was 9 percentage points (pp) lower than the *Female % of PhDs* (25% versus 34%). Similarly, when we examine the full period, the gap is -9.4 pp (21% versus 30%) (Fig. 2). This gap holds across universities, ranging from 5 to 15 pp (Fig. S1). Aggregate PhD statistics may hide field-specific differences: the *Female % of PhDs* granted by universities in our sample ranges from 18% in *Computer & Communications* to 49% in *Biological and Biomedical* sciences (Table S5). The *Female % of New Inventor-PhDs* also varies across fields (Table S4). Comparing the *Biological and Biomedical* PhDs in 1995-2015 to the *Drugs & Medical* New Inventor-PhDs in 2000-2015 patents, there is still a significant gap. This field has the highest *Female % of New Inventor-PhDs*, at 31%, but this score is 18 pp lower than women's participation in PhDs.

Thus, in contrast to the view that STEM PhDs are the limiting factor to inclusion in patenting, even at the start of the PhD pipeline, we find that the new inventor gender gap in top universities is higher than one would expect given the share of female STEM PhD holders.

The Role of Top Inventor Faculty Advisors in PhD Student Patenting

To understand the gap between the participation of women in STEM PhDs versus in patenting during their training, we focus the rest of our analysis on the role of faculty advisors. As principal investigators in grants and labs, they have a central role in mentoring and training that spills over into later career behaviors (*4*, *17*, *26*).

We emphasize the role of faculty advisors who are themselves top inventors. The 876 TIs represent only 2% of inventors at the 25-universities, but they contribute to almost 40% of the university patents (Tables S6). Their autonomy, reputation, and patenting intensity give them a disproportionate role in bringing new inventors into the economy as part of their research teams; and they guide attitudes towards commercial science in their labs and universities (*3*, *23*). In our analysis, co-patenting with an advisor is the key channel for PhD advisees to become new inventors. Advisees patenting without their advisors is rare, reinforcing the notion that patenting is a key element of PhD training. These interactions not only represent the first stage where the pipeline of STEM PhDs are transformed into new inventors but also a focal point where disparities might arise, creating cumulative disadvantage and reducing the total productivity of PhD scientists across their careers (*10*, *11*, *19*).

Recall that our sample of TI faculty trained 13,286 PhD graduates during 1995-2015 of whom 3,008 became new inventor PhDs as co-inventors on their patents (Table S6). These TIs account for over 44% of all New Inventor-PhDs at the 25-universities. Thus, they have a significant impact on the likelihood of their PhD students becoming new inventors (*Probability of New Inventor-PhDs*): 22.6% for PhD students of TIs compared to 3.7% overall in the 25-university sample, and 2.2% for PhD students of other (non-TI) advisors (Fig. 2A and Table S8).

The Role of Top Inventor Faculty Advisors in Gender Inclusion

Is co-patenting with top inventor advisors equally available to female and male student advisees? To answer this, we measure the *Female % of New Inventor-PhDs* for the set of TIs and assess whether this score is aligned with their *Female % of PhDs*. We also compute the TIs' *Probability of New Inventor-PhDs* by advisee gender. This descriptive analysis is shown in Fig. 2A that compares STEM PhD graduates and New Inventor-PhDs by gender for the set of TIs versus at their universities in total.

Several findings are notable. First, the *Female % of PhDs* is lower for TIs than their universities overall (25.5% versus 30.1%). This means that female PhDs are less likely to be trained by TIs than male PhDs (6.1% versus 7.7%, Table S8). Second, just as for the 25-university sample, we find that the *Female % of New Inventor-PhDs* is 21% for all TIs. This score is lower than their supply of female PhDs: the gap between *Female % of New Inventor-PhDs* and *Female % of PhDs* is -4.6 pp for TIs. Nonetheless, this gap is narrower than for universities overall (-9.4 pp). The greater inclusion among top inventor advisors holds across most universities (Fig. S1).

Relatedly, if we compare the likelihood of first-patenting for TIs' female-versus-male PhDs, we find a 23% lower rate for female PhDs: 18.6% probability of becoming New Inventor-PhDs versus 24% for male PhDs. This differential probability is lower than for universities overall where female PhDs face a 39% lower rate of first-patenting (2.6% versus 4.2% probability). To put this starkly: while TIs have a seven times higher rate of training female PhDs to become new inventors than their universities (18.6% versus 2.6%), the female PhDs have a 21% lower likelihood of being matched with TIs than male PhDs and, when matched, are 23% less likely than male PhDs to become new inventors.

Exploring whether these patterns hold across technology classes, we find that the distribution of New Inventor-PhDs across classes for the 25-universities versus the TIs is similar (Table S4 versus S7), suggesting that the differences between TIs and other faculty advisors cannot be attributed entirely to field. Importantly, we can compare the *Biological and Biomedical* PhDs – a field well-represented by female PhDs that accounts for about 25% of all the PhDs trained by the universities and their TIs – to the *Drugs and Medical* inventor data (Fig. 2B and Table S9). Our results hold for this field, and the gender differential probability of becoming new inventors gets worse at universities overall (53% lower rate for females versus 22% lower rate for females with TIs). There could be gender-related subfield differences among students (and advisors) affecting our findings. We can account for this in the rest of our analysis as we shift our focus to individual TIs and their students.

The Role of Female versus Male Top Inventor Faculty Advisors in Inclusion

We ask whether top inventor advisors differ by gender in their shares of female PhDs and their rates of transformation of female PhDs into new inventors. These questions are grounded in prior findings that homophily or cultural similarity may play a role in the formation of startup, research, and inventor teams (*18*, *29*, *30*), especially among under-represented groups (*31*). Common research interests also matter for the advisor-advisee matching (*26*); and to the extent that there are gender differences in research topic preferences (*16*), this may influence same-gender advisor-advisee matching and potentially co-patenting rates. In terms of publications, recent work shows that female PhDs with female advisors publish more papers during their studies than females with male advisors, but fewer papers than males with female advisors (*4*). Overall, the results in prior work are mixed and focus on publications and particular universities and fields.

Our empirical analysis for this section is primarily at the TI level (i.e., the unit of observation is a top inventor faculty advisor in a particular university) to control for key observable attributes of advisors such as gender, university, field, and their pool of PhD graduates (Table S10 and Section S5). In our sample, Female Top Inventor advisors (FTIs) represent only 8% of all TIs (68 versus 808). Nonetheless, they are similar to Male Top Inventor advisors (MTIs) in their average number of PhD graduates 1995-2015 (16.2 versus 15.5) and patents (14.2 versus 16.9). FTIs and MTIs also have a similar share of New Inventor-PhDs among their advisees (27.4% versus 28.4% (unweighted) and 21.5% versus 22.6% (weighted by TI count of advisees), and this holds by advisor field too (Fig. S2). This suggests that FTIs and MTIs are similarly engaged in patentable research and in training New Inventor-PhDs.

To account for differences among students, we also estimate models at the student-TI level: the unit of observation is a PhD advisee of a TI, totaling 13,602 observations since a few students have multiple advisors (Table S10 and Section S5). These models allow us to control for the graduation year and the thesis topic of the students. We expect advisees' research topics to be related to those of their advisors. Yet there could be gender differences in students' research topic preferences (even within-advisor) affecting the probability of patenting. To identify the PhD Thesis Topic, we transform the thesis title of the advisees into topics that are included as fixed effects (*32*).

First, we examine whether there are differences between male and female TIs in their *Female % of PhDs* (Fig. 3A). We find that female PhDs are more likely to be trained by female than male TIs. The *Female % of PhDs* is statistically higher ($p < 0.01$) for FTIs than MTIs: the gap ranges from 7.3 to 9.4 pp across models. The gap is 8.9 pp (33.6% versus 24.7%) in the TI-level model that controls for a TI's main technology and university. It declines to 7.3 pp (32.2% versus 25.0%) in the student-TI level model that adds PhD Graduation Year and PhD Thesis Topic. This suggests that the higher female share of PhDs for FTIs versus MTIs cannot simply be explained by student differences in research interest. Given that the *Female % of PhDs* at the 25-universities is 30.1% (Fig. 2), female PhDs seem over-represented among FTIs (32-34%) and under-represented among MTIs (25%).

Next, we estimate whether TIs differ by gender in the % of New Inventor-PhDs who are women. Fig. 3B shows that the *Female % of New Inventor-PhDs* is statistically higher ($p < 0.01$) for FTIs than MTIs: the estimated gap is 9.7 pp (30.4% versus 20.7%) in the TI level model, and 8.8 pp in the student-TI level model that adds PhD Thesis Topic. However, the female-versus-male advisor gap becomes insignificant after controlling for the *Female % of PhDs* of each advisor (Table S14). Thus, the *relative* rate of transformation of female-versus-male advisees into new inventors is independent of the gender of the advisor. This rate is less than one (lower probability for female advisees) since female and male TIs have an estimated *Female % of New Inventor-PhDs* lower than their *Female % of PhDs* (Fig. 3).

Lastly, to assess the magnitude of this differential probability of first-patenting (and to confirm that it is independent of the advisor gender), we estimate the rates of transformation of female and male PhD advisees into new inventors (*Probability of New Inventor-PhDs*) and compare these rates for female versus male TIs (See Fig. 4 and Table S15).

We find (Fig. 4A) that the rate at which all TIs transform female PhDs into New Inventor-PhDs is statistically lower ($p < 0.01$) than for male PhDs: the estimated gap ranges from -4 to -8 pp (17-26% lower rate). In the TI level model with technology and university controls the gap is -7.6 pp (a 26% lower rate for women: 21.6% versus 29.2% probability). The estimated gap remains significant but of lower magnitude at -4.9 pp (17% lower rate) if we include TI fixed effects. Therefore, our findings hold within-advisor: they are not driven by female students matching to advisors who care less about patentable research topics. Similarly, the gap remains at -4.5 pp (19% lower rate: 19.3% versus 23.7% probability) in the student-TI level model that controls for TI, PhD Graduation Year and PhD Thesis Topic. This suggests that the gap is not dependent upon female students choosing less patentable topics. Importantly, female and male TIs have a similar rate of engaging their female PhDs into first-patent, and the female-vs-male advisee probability gap does not depend on the advisor gender across models (Fig. 4B). The estimated gender gap holds for most TI technology classes, and are unrelated to the advisor gender (Fig. 5 and S3).

Put into context, if we eliminate the male-female student differential probability of becoming new inventors with their top inventor advisors (i.e., remove the lowest estimated gap of 4 pp keeping constant the number of male NI-PhDs), there would be a 10% increase in the number of female New Inventor-PhDs in the 25-universities improving the *Female % of New Inventor-PhDs* from 21% to 23%. Given the importance of expanding engagement in patenting, we discuss potential interventions to improve the presence of all, but especially female, New Inventor-PhDs.

Discussion and Conclusion

In today's top 25 U.S. universities (by patenting), STEM PhD students increasingly become new inventors, and they start on their journey of contributing to commercial science. Top inventor faculty advisors (TIs) play a key role in training these PhDs to be new inventors by co-patenting: the chances of patenting for students trained by a TI are about ten times higher than with other advisors. However, female PhD students (supervised by TIs and especially by other faculty) have a lower probability of becoming new inventors relative to their male counterparts. Our analysis of advisees of TIs shows that this differential probability remains even after accounting for university, field, advisor gender, and thesis topic.

Our examination of channels for the inclusion (or exclusion) of female STEM PhD students as new inventors allows us to highlight those settings and individuals who could serve as catalysts for greater gender inclusion at the earliest stages of women's PhD careers. Students are most likely to learn commercial science with top inventor advisors. Thus, one channel for creating more female new inventors at universities would be to increase the number of female PhDs who get trained by TIs. On average these TIs are less likely to have female PhDs than universities overall. Given the much higher rate at which TI advisees are likely to patent, increasing the number of female PhDs working with TIs would be a critical step. Potential interventions would require a better understanding of the advisor-advisee matching process (*18, 26*).

Beyond providing incentives for all TIs to train more female PhDs it is important to consider a second channel for increasing female New Inventor-PhDs: increase the number of female TIs (who constitute only 8% of TIs). This channel will be important if homophily (*18*) or gender differences in research interest (*16*) affects same-gender advisor-advisee match. On average, female TIs have a significantly higher share of female PhDs than male TIs even after controlling for thesis topic (consistent with prior findings in biomedical labs (*18*)). By training more female PhDs, they also have a significantly higher share of female New Inventor-PhDs in their patents than do male TIs. Thus, an indirect outcome of programs to encourage female faculty to engage in high levels of patenting (e.g., through better access to technology transfer offices (*23, 33*)) will be more female New Inventor-PhDs.

Lastly, we have shown that, regardless of advisor gender, there seems to be a "leaky pipeline" of female inventors-to-be: both female and male top inventor advisors have a rate of transformation of their female PhDs into new inventors (co-patenting) significantly lower than for male PhDs. Thus, a third channel to improve overall entry into patenting by female PhDs is to increase female participation in patenting. This channel (and potential interventions) is poorly understood. The likely factors for the gender difference in the probability of becoming New inventor-PhDs fall into two categories (*33*): “supply-side” factors (differences between female and male advisees in patenting preferences, self-assessment of skills, or resources) and “demand-side” factors (possible faculty gender bias and lower patenting opportunities for female advisees).

On the supply-side, our findings are not driven by female advisees matching with TIs who care less about patenting because the gender gap holds in our within-advisor analysis. Moreover, we have accounted for the fact that female advisees may select into research topics that are less patentable than those of males by controlling for thesis topic. That said, there are limitations to our approach. Female advisees may have lower preferences for patenting than males per se (regardless of topic), affecting their engagement in patenting. We think that this is less likely in our setting because the advisors (versus their PhD students) would assess the patentability of a project. However, students’ patenting preferences together with biased self-assessment of their own skills and contributions (*34*, *35*), may lead them to select out of patenting efforts or to advocate less strongly for their contributions. Interestingly, recent studies show that female PhDs publish less than their male counterparts during their training (*4*). This suggests that the gender gap in New Inventor-PhDs may not simply be driven by women having higher preferences for publishing versus patenting. Overall, while differential first-patenting rates are unlikely to be a purely supply-side phenomenon, it is important to consider the potential value of interventions (before and during doctoral training) that provide female role models, highlighting the contributions of female top inventors. In addition, prior work shows that barriers to accessing commercialization resources (versus gender-specific patenting preferences per se) help to explain the lower patenting rate of female faculty (*23*, *33*). Thus, a potential intervention might be to ensure equal access to commercialization resources for female and male students. These are all important areas for further research to foster gender-balanced inventorship during PhD training.

On the demand-side, our findings are consistent with the possibility that women’s innovation skills and contributions are somewhat under-valued by male and female advisors (*20*, *36*). Recent work shows that female graduates, postdocs and faculty are less likely to be listed as co-authors in publications and patents generated by their research teams than their male counterparts (*20*). Qualitative studies find that women’s higher rate of unacknowledged contributions relate to their work being less visible to senior researchers (*20*). Relatedly, studies have found that faculty’s gender biases favor male students when they apply for the position of lab manager (*36*), and (in a related but different context) that investors prefer pitches presented by male over female entrepreneurs (*37*). Thus, interventions that help advisors to consider the contributions and potential of all their students are welcome and could shift outcomes towards parity.

The outlook for universities contributing to human capital, invention and patenting in the U.S. economy is positive. Yet to be more relevant to the innovation economy, universities might take additional steps to provide opportunities for all PhD students to engage in commercial science. Our results highlight the chance for additional new inventors to enter the economy during graduate training. A critical element of this opportunity is to ensure that it fosters gender parity. This is a key consideration for top inventor faculty and other advisors and could make significant inroads into inclusive innovation. By offering further training for PhDs, especially women, to become new inventors, we would create a cumulative advantage across their careers and improve innovation throughout the economy (*11–13*, *10*). Early first-patenting may influence multiple research and commercialization outcomes like long-term publication and patenting productivity (*9*, *10*, *19*), the ability to create high potential startups based on the patented inventions (*38*), and insights into university-industry effective collaborations (*23*). The examination of the potential long-term effects of early inventorship will be the focus of our follow-on scholarship.

Materials and Methods

We combine USPTO, ProQuest, and IPEDS datasets to compare the population of STEM PhDs to the New Inventor PhD students for the 25-universities and their top inventor advisors. These datasets are available from their source websites. We implement two sample validation exercises: 1) using the MIT roster of students, faculty, and inventors and 2) using a representative random sample of New Inventors who were searched online (See SI Materials and Methods). The university-patent assignee bridge and our name-gender match algorithm will be available on the corresponding author’s website. Other relevant codes are available from the authors upon request.

Acknowledgements

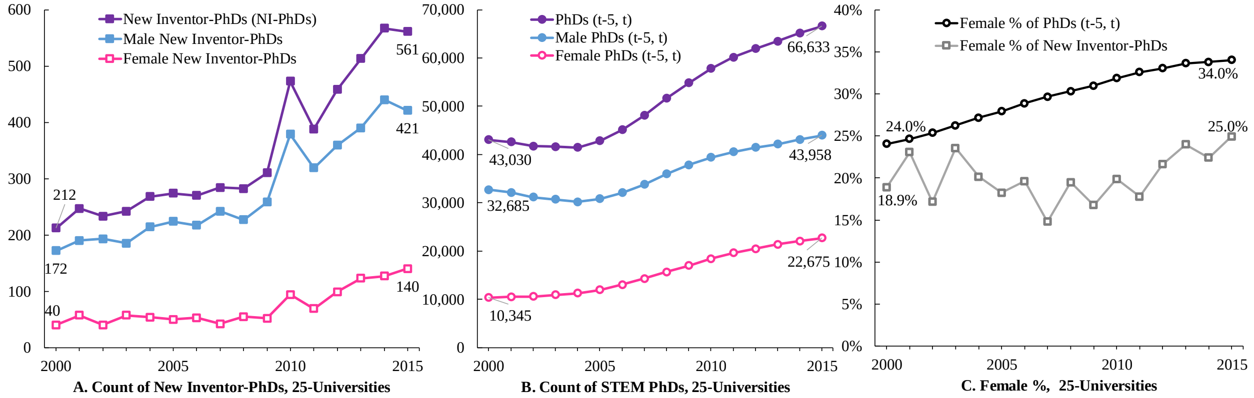
We thank (especially) Luca Gius, Peter Favaloro, Rich Bryden, Western Bonime, and Cory Ventres-Pake for their research assistance with the data collection and visualization. We thank Lydia Snover and Abdou Khadre Seck and MIT Institutional Research for access to MIT data. We are grateful for comments by the editor and the anonymous reviewers, and by Maryann Feldman, Michael Cima, Adam Jaffe, Shulamit Kahn, Myriam Mariani, Paula Stephan, Patrick Gaule, Karin Hoisl, Valerie Karplus, Erin Scott, Stefan Sorg, and Don Sull. **Funding:** This project has been funded by a National Science Foundation, Science of Science and Innovation Policy Grant (Award #1757344).

References

1. B. F. Jones, The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder? *Review of Economic Studies*. **76**, 283–317 (2009).
2. P. Giuri, M. Mariani, S. Brusoni, G. Crespi, D. Francoz, A. Gambardella, W. Garcia-Fontes, A. Geuna, R. Gonzales, D. Harhoff, K. Hoisl, C. Le Bas, A. Luzzi, L. Magazzini, L. Nesta, Ö. Nomaler, N. Palomeras, P. Patel, M. Romanelli, B. Verspagen, Inventors and invention processes in Europe: Results from the PatVal-EU survey. *Research Policy*. **36**, 1107–1127 (2007).
3. M. Delgado, F. Murray, Mapping the Regions, Organizations, and Individuals That Drive Inclusion in the Innovation Economy. *Entrepreneurship and Innovation Policy and the Economy*. **1**, 67–101 (2022).
4. M. Pezzoni, J. Mairesse, P. Stephan, J. Lane, Gender and the Publication Output of Graduate Students: A Case Study. *PLoS ONE*. **11**, e0145146 (2016).
5. Survey of Doctoral Recipients, 2015. National Science Foundation. https://ncsesdata.nsf.gov/doctoratework/2015/html/SDR2015_DST_12_1.html.
6. Y. Xue, R. Larson, STEM crisis or STEM surplus? Yes and yes. *MLR* (2015), doi:10.21916/mlr.2015.14.
7. E. Islam, J. Zein, Inventor CEOs. *Journal of Financial Economics*. **135**, 505–527 (2020).
8. F. Murray, Innovation as co-evolution of scientific and technological networks: exploring tissue engineering. *Research Policy*. **31**, 1389–1403 (2002).
9. P. Azoulay, W. Ding, T. Stuart, The Impact of Academic Patenting on the Rate, Quality, and Direction of (Public) Research Output. *The Journal of Industrial Economics*. **57**, 637–676 (2009).
10. K. B. Whittington, Mothers of Invention?: Gender, Motherhood, and New Dimensions of Productivity in the Science Profession. *Work and Occupations*. **38**, 417–456 (2011).
11. R. K. Merton, The Matthew Effect in Science: The reward and communication systems of science are considered. *Science*. **159**, 56–63 (1968).
12. P. Allison, J. Stewart, Productivity Differences Among Scientists: Evidence for Accumulative Advantage. *American Sociological Review*. **39**, 596–606 (1974).
13. T. A. DiPrete, G. M. Eirich, Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments. *Annu. Rev. Sociol.* **32**, 271–297 (2006).
14. 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).
15. L. Cook, Unequal Opportunity: The Innovation Gap in Pink and Black, "Unequal Opportunity: The Innovation Gap in Pink and Black" in *Does America need more innovators?*, M. H. Wisnioski, E. S. Hintz, M. S. Kleine, Eds. (The MIT Press, Cambridge, MA, 2019), pp. 221–249.
16. R. Koning, S. Samila, J.-P. Ferguson, Who do we invent for? Patents by women focus more on women's health, but few women get to invent. *Science*. **372**, 1345–1348 (2021).
17. P. Gaule, M. Piacentini, An advisor like me? Advisor gender and post-graduate careers in science. *Research Policy*. **47**, 805–813 (2018).
18. J. M. Sheltzer, J. C. Smith, Elite male faculty in the life sciences employ fewer women. *Proceedings of the National Academy of Sciences*. **111**, 10107–10112 (2014).
19. J. Huang, A. J. Gates, R. Sinatra, A.-L. Barabási, Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 4609–4616 (2020).
20. M. B. Ross, B. M. Glennon, R. Murciano-Goroff, E. G. Berkes, B. A. Weinberg, J. I. Lane, Women are credited less in science than men. *Nature*. **608**, 135–145 (2022).
21. S. J. Ceci, D. K. Ginther, S. Kahn, W. M. Williams, Women in Academic Science: A Changing Landscape. *Psychol Sci Public Interest*. **15**, 75–141 (2014).
22. 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).

23. W. W. Ding, F. Murray, T. E. Stuart, Gender Differences in Patenting in the Academic Life Sciences. *Science*. **313**, 665–667 (2006).
24. J. Hunt, J.-P. Garant, H. Herman, D. J. Munroe, Why are women underrepresented amongst patentees? *Research Policy*. **42**, 831–843 (2013).
25. K. Jensen, B. Kovács, O. Sorenson, Gender differences in obtaining and maintaining patent rights. *Nat Biotechnol*. **36**, 307–309 (2018).
26. P. Azoulay, C. C. Liu, T. E. Stuart, Social Influence Given (Partially) Deliberate Matching: Career Imprints in the Creation of Academic Entrepreneurs. *American Journal of Sociology*. **122**, 1223–1271 (2017).
27. Materials and Methods are available as Supporting Information Appendix (SI).
28. B. Hall, A. Jaffe, M. Trajtenberg, "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools" (w8498, National Bureau of Economic Research, Cambridge, MA, 2001), , doi:10.3386/w8498.
29. 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).
30. K. B. Whittington, "A tie is a tie? Gender and network positioning in life science inventor collaboration." *Research Policy*. **47**, 511–526 (2018).
31. M. McPherson, L. Smith-Lovin, J. M. Cook, Birds of a Feather: Homophily in Social Networks. *Annu. Rev. Sociol.* **27**, 415–444 (2001).
32. M. Gentzkow, B. Kelly, M. Taddy, Text as Data. *Journal of Economic Literature*. **57**, 535–574 (2019).
33. F. Murray, L. Graham, Buying science and selling science: gender differences in the market for commercial science. *Industrial and Corporate Change*. **16**, 657–689 (2007).
34. M. F. Fox, P. E. Stephan, Careers of Young Scientists: Preferences, Prospects and Realities by Gender and Field. *Social Studies of Science*. **31**, 109–122 (2001).
35. S. J. Correll, Gender and the Career Choice Process: The Role of Biased Self-Assessments. *American Journal of Sociology*. **106**, 1691–1730 (2001).
36. C. A. Moss-Racusin, J. F. Dovidio, V. L. Brescoll, M. J. Graham, J. Handelsman, Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*. **109**, 16474–16479 (2012).
37. A. W. Brooks, L. Huang, S. W. Kearney, F. E. Murray, Investors prefer entrepreneurial ventures pitched by attractive men. *Proceedings of the National Academy of Sciences*. **111**, 4427–4431 (2014).
38. J. Guzman, S. Stern, Where is Silicon Valley? *Science*. **347**, 606–609 (2015).

Figures and Tables



A. Count of New Inventor-PhDs, 25-Universities

B. Count of STEM PhDs, 25-Universities

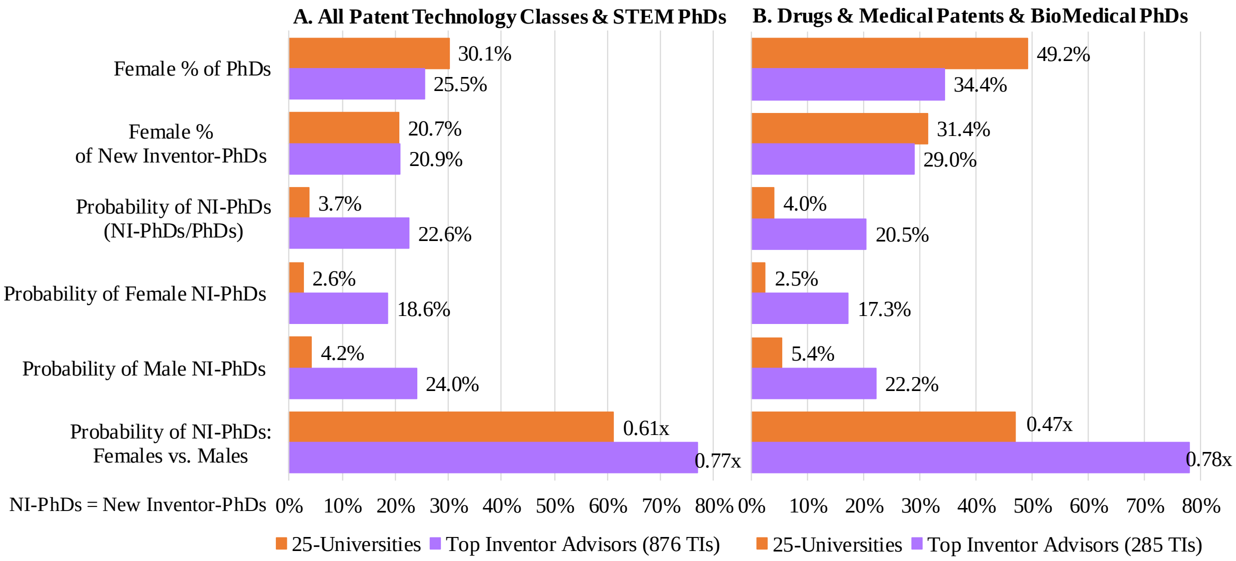
C. Female %, 25-Universities

	Annual Growth	** p < .01
NI-PhDs	$6.7\%^{**}$	
Male NI-PhDs	$6.5\%^{**}$	
Female NI-PhDs	$7.6\%^{**}$	

	Annual Growth	** p < .01
PhDs	$3.7\%^{**}$	
Male PhDs	$2.7\%^{**}$	
Female PhDs	$6.1\%^{**}$	

	Annual Change	Years-to-Parity
Female % of PhDs	$0.71\%^{**}$	(21, 22, 24)
Female % of NI-PhDs	0.19%	(49, 133, NA)

Figure 1. Trends in New Inventor-PhDs and STEM PhDs by Gender. Fig. A shows the count of New Inventor-PhD students in patents granted in yeart *t* who graduated in the period (t-5, t) (e.g., NI-PhDs in 2015 granted patents who graduated during 2010-2015). This time lag assumes that (i) patents are filed 3-year prior to be granted and (ii) PhD students file their first-patent 3-years prior to graduation to 2-years after graduation (See Section S4A). Similarly, Fig. B shows each year *t* the count of STEM PhDs who graduated in the period (t-5, t). In Fig. A-B, the estimated Annual Growth is the slope in the annual trends of log(Y). Fig. C plots the *Female % of New Inventor-PhDs* in patents granted in year t who graduated in (t-5, t). (e.g., 25% in 2015: 140 women out of 561 NI-PhDs); and the *Female % of PhDs* granted in (t-5, t). The estimated Annual Change is the slope in the annual trends. Using the slope, we compute the Years-to-Parity (50% score) and 95% confidence intervals. The trends in Fig. 1 are robust to using all NI-PhDs (i.e., those filing first-patent 6-years prior to graduation to 2-years after graduation). NI-PhDs data (USPTO and ProQuest) and STEM PhDs data (IPEDS). See Table S5 for the STEM definition.



	STEM PhDs	New Inventor-PhDs
25-Universities	184,865	6,847
876 TIs	13,286	3,008

	BioMedical PhDs	New Inventor-PhDs
25-Universities	47,683	1,896
285 TIs	3,675	754

Figure 2. STEM PhDs and New Inventor-PhDs by Gender: 25-Universities versus Top Inventor Advisors. Fig. 2A compares the pool of STEM PhD graduates (1995-2015) and New Inventor PhD students (NI-PhDs) in patents (2000-2015) by gender for the 25-universities and their Top Inventor advisors (TIs) (See Table S8). Sample of 876 TIs who have at least one NI-PhD. We identify TIs' PhD advisees and those who became NI-PhDs in the TIs' patents. The STEM PhD data for the universities come from IPEDS (See Table S5 for the STEM definition). We validate that TIs and their PhDs are in our STEM fields (Section S2). Fig. 2B compares Biological & Biomedical Science ('BioMedical') PhDs to NI-PhDs in Drugs & Medical patents for the 25-universities and the set of 285 TIs whose main technology is Drugs & Medical (See Table S9).

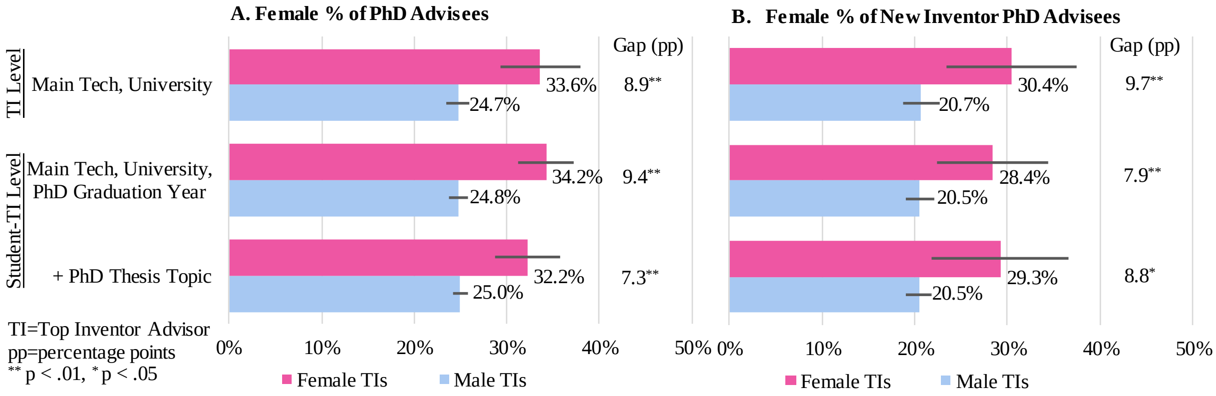
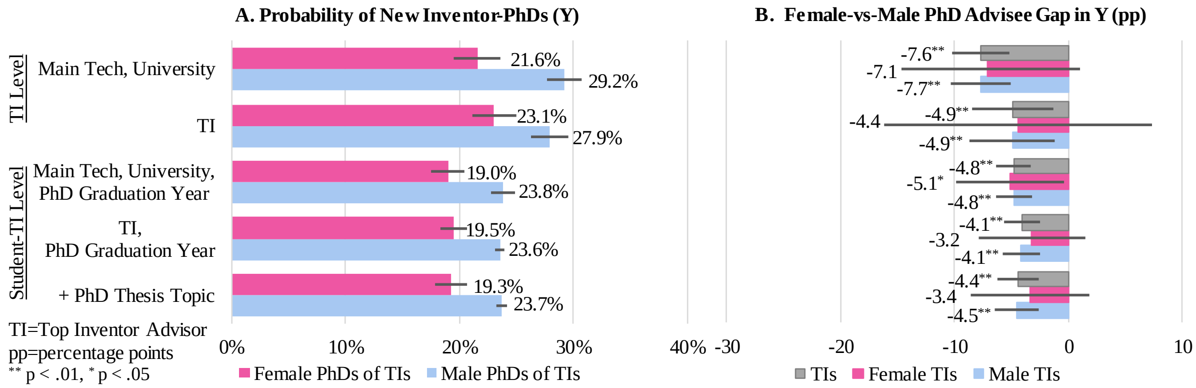


Figure 3. Comparison of Female versus Male Top Inventor Advisors in their Female % of PhDs and Female % of New Inventor-PhDs. Graphs show the estimated *Female % of PhDs* (A) and *Female % of New Inventor-PhDs* (B) for female TIs (FTIs), male TIs (MTIs), and the FTIs-vs-MTIs gap in percentage points (OLS estimates and the 95% confidence intervals). See Tables S11-S12 and Section S6. In the TI level models, the unit of observation is each TI *i* in the 25-universities (N = 876; Equation 2): Y_i = female share of PhD advisees of a TI (1995-2015 graduates) (A) and Y_i = female share of New Inventor PhD students (NI-PhDs) in a TI's (2000-2015) patents (B). Y is a function of TI gender and the indicated controls (fixed effects): (1) TI Main Tech class and University. In the Student-TI level models, Y_{si} = Female PhD (1 = Yes, 0 = No) for each student *s* of a TI (N = 13,602; Equation 3) in (A) and Y_{si} = Female NI-PhD (1 = Yes, 0 = No) for each new inventor (N = 3,068; Equation 4) in (B). In these models, a TI is weighted based on his/her count of PhDs (A) or count of NI-PhDs (B). The estimates control for (2) TI Main Tech, University, and PhD Graduation Year, and (3) also adds PhD Thesis Topic. The estimated FTIs-vs-MTIs gaps are robust to adding other TI attributes and using LASSO in (3).



TI=Top Inventor Advisor
pp=percentage points
** p < .01, * p < .05

Figure 4. Probability of Becoming New Inventors for Female versus Male PhD Students of Top Inventor Advisors. Graphs show the estimated *Probability of New Inventor-PhDs* (Y) for female and male PhDs of all TIs (A) and the female-vs-male advisee gap in Y for students of (all, female and male) TIs (B). (OLS estimates and the 95% confidence intervals). Sample of 1995-2015 graduates and 2000-2015 TI patents. See Table S15 and Section S7. In the TI level models (Equation 6), Y_{ig} = share of new inventors among PhD advisees (by advisee gender *g*) for each advisor *i* (N = 1,615). Y is a function of advisor gender (Female TI indicator), advisee gender (Female PhD indicator), their interaction, and the indicated controls: (1) TI Main Tech class and University fixed effects (FEs); or (2) TI FEs. In the Student-TI level models, the unit of observation is a PhD student *s* of a TI *i* (N = 13,602), Y_{si} = New Inventor-PhD (1 = Yes, 0 = No) (Equations 7-8). These models weigh a TI based on the count of advisees. The estimates control for: (3) TI Main Tech, University, and PhD Graduation Year FEs; (4) TI FEs and PhD Graduation Year and (5) also adds PhD Thesis Topic FEs. Fig. B shows that the gender gap is similar for students of female and male TIs (i.e., the coefficient of Female TI×Female PhD is insignificant). Findings are robust to: broadening the definition of NI-PhD (including students patenting without their advisors); focusing on alternative pools of PhD graduates; using LASSO in (5); and controlling for other TI attributes (Tables S15-S16).

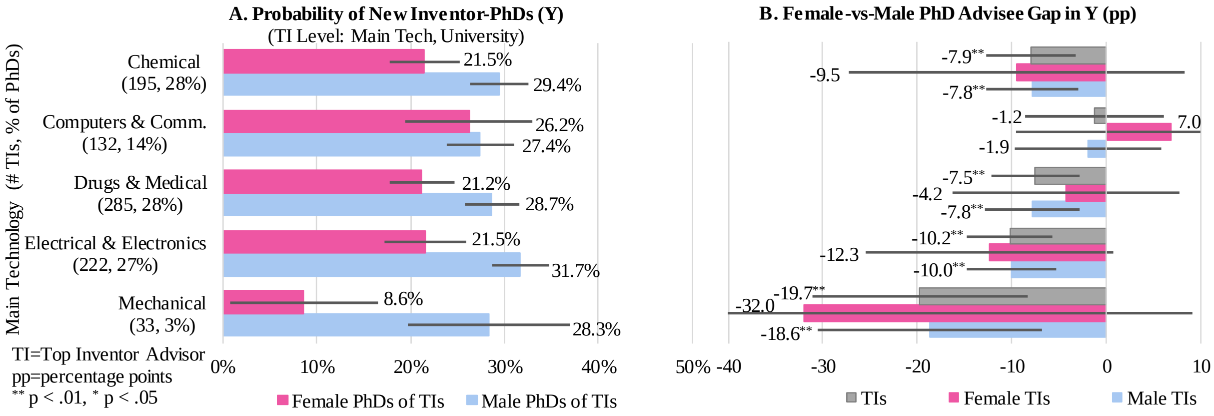


Figure 5. Probability of Becoming New Inventors for Female versus Male PhD Students of Top Inventor Advisors: Estimates across Technology Classes. For each field (TI main technology), the graphs show the estimated *Probability of New Inventor-PhDs* (Y) for female and male PhDs (A) and the female-vs-male advisee gap in Y for students of (all, female, and male) TIs (B). (OLS estimates and the 95% confidence intervals). Sample of 1995-2015 graduates and 2000-2015 TI patents. TI level model (N = 1,615): Y_{ig} = share of new inventors among PhD advisees (by advisee gender *g*) for each advisor *i*. Y is a function of a 3-way interaction of TI gender, advisee gender, and TI Main Tech (six classes) controlling for university fixed effects. The figure shows the estimated Y across five classes that account for 867 TIs and 99% of the advisees (Table S7). While the magnitude of the gender gap differs across technologies, these differences are not statistically significant (Fig. 5B). The results in Fig. 5 are robust to the Student-TI level model (Fig. S3).