The Benefits of Gender-Inclusive Innovation: Evidence from US Patents

Joel Waldfogel
University of Minnesota, NBER, ZEW

December 9, 2024

PRELIMINARY DRAFT - DO NOT CITE OR CIRCULATE

Abstract

Despite concern about undiscovered Einsteins and lost Curies, there is little direct evidence of the value of including more groups of people in innovative activity. Between 1976 and 2020 the share of new US patents with at least one female inventor rose from under 10 to over 25 percent. How has this affected the value of innovations? I begin by exploring descriptive questions. First, a growing number of women is pursuing innovation-relevant education. Second, additional patents appear to "expand the market" by attracting additional citations, rather than simply diverting citations from others. Third, growing numbers of female inventors and patents appear not to displace male inventors or patents. Fourth, as the female-invented share of patents grows across vintages, the share of citations attracted by the female-invented patents grows nearly proportionally, suggesting that the additional female-inventor patents are valuable. I then construct an equilibrium model of demand for, and supply of, patents, which allows me to simulate a counterfactual environment without the female influx. The female influx raises the female share of patents from roughly 10 to 20 percent of available patents, raising overall inventor surplus by 2.7 percent (1.2 if I allow for endogenous male inventor entry). Female inventor surplus rises by 4.6 (2.8), while male inventor surplus rises by 2.1 (0.7) percent.

1 Introduction

Women have long lagged behind men in science, engineering, and innovative activity: Women account for relatively few of the inventors listed on patents, and women account for relatively few authors in many creative industries. Given the importance of innovation for growth, there is a concern that reliance on only a subset of human innovative potential has hindered progress. There is evidence that greater female exposure to innovation would lead to more innovation (Bell et al., 2019; Cook, 2011; Hoisl et al., 2023) and more generally that better allocation of workers to tasks could raise the output (Hsieh et al., 2019). Innovation is responsible for the majority of growth in living standards, so non-inclusive innovation could therefore slow growth (Solow, 1957).

Despite great concern above innovation forgone due to non-inclusive participation in inventive activity, there is little direct evidence on the value of the additional innovation that inclusiveness might deliver. The female share of patent inventors has been rising (Frietsch et al., 2009; Ding et al., 2006; Hunt et al., 2013; Kim and Moser, 2020; Toole et al., 2021), and this paper makes use of this shift to directly quantify the benefit of more inclusive innovation. As a growing share of patents have some female inventors, do the "female patents" account for a growing share of the value of new inventive activity. In short – at least as a descriptive matter – as the female share of patents has risen, has the share of citations to female patents also risen?[1]

More inclusive innovation could be socially beneficial for a variety of reasons. First, growing female participation could deliver rewards for the creation of innovation more broadly. Second, female participation in innovation may change the direction of innovative activity (Koning et al., 2021), as women may invent in areas neglected by male inventors. Finally, there is the question pursued here: drawing from a larger pool of potential inventors might

[1] I follow the literature in using patent citations to proxy their value (Hall et al., 2005; Kogan et al., 2017).

raise the value of innovations of interest to diverse sorts of end users.

Given that most patents are not valuable, it is not obvious that an influx of new patent inventors – and patents – would deliver valuable innovations. However, when innovation quality is unpredictable, an increase in the number of products or innovations can be valuable (Aguiar and Waldfogel, 2018); and it is well-known that patent value is unpredictable at the time of invention (Scherer and Harhoff, 2000). This makes it plausible to expect that a female influx into patenting would deliver valuable inventions.

This study takes gender as its context, and gender is important. Not only has women's participation in innovation traditionally lagged men's, women are a large group, making up half the population. But gender is not the only dimension of interest. Other groups of people, including ethnic and racial minorities, as well as residents of low-income countries, have also participated in innovative activity at relatively low rates. Those dimensions of inclusion also merit attention. I study gender here in part because the link between names and gender make the question researchable.

The paper proceeds in five sections after the introduction. Section 2 describes background on both the relevant literatures on female participation in the economy generally and as patent inventors in particular. Section 3 presents a simple theory, and an equilibrium model, of the benefits of greater inclusivity in innovation when the realized quality of the innovations is unpredictable. The "demand" side consists of inventors choosing to cite existing patents. The supply side is the decision by male and female inventors to create new patents, which I term "entry of patents." Quality is unpredictable, and patents enter if they expect citations above an empirically determined citation threshold. Section 4 describes the US patent data used the study. These include patents granted between 1976 and 2020, along with the gender of the inventors on those patents, their technology category, and the citations those patents have received from subsequent US patents. Section 5 presents descriptive results, in four parts. First, I document the growing number of women pursuing innovation-relevant

education as well as a growing share of patents with female inventors. Second, I show evidence that additional patents "expand the market" by attracting additional citations, rather than simply diverting citations from others. Third, growing numbers of female-invented patents appear not to displace male-inventor patents. Fourth, I present direct evidence suggesting the welfare benefit from the female influx: as the female-invented share of patents grows across vintages, the share of citations attracted by the female-invented patents grows nearly proportionally. Section 6 presents welfare estimates. First, I estimate a structural nested logit model of inventor demand for patents to cite. Second, I measure the predictabiity of patents' realized quality, and I combine these ingredients to solve an equilibrium entry model. I estimate that the female influx into patenting has raised male and female inventors' analogue to consumer surplus by 0.67 and 2.78 percent, respectively.

2 Background

For a variety of reasons, women's labor force participation has grew substantially during the 20^{th} century (Fernández, 2013; Greenwood et al., 2005; Costa, 2000; Goldin, 2006). Contributing factors include new technologies reducing the burden of home activity, such as washing machines and birth control. Changing social attitudes have also effectively reduced women's opportunity costs. While women made up 20 percent of the US labor force in 1920, their share grew to 30 percent by 1950 and has since 2000 been over 45 percent.[2]

Still, female participation in innovative activity has lagged men's; and it is well known that women appear as patent inventors at lower rates than men (Hunt et al., 2013; Frietsch et al., 2009; Ding et al., 2006; Martínez et al., 2016; Hoisl and Mariani, 2017; Fechner and Shapanka, 2018; Thursby and Thursby, 2005; Whittington and Smith-Doerr, 2005; Murray and Graham, 2007). Various studies explore reasons such as child-rearing (Kim and Moser,

[2]See `https://www.dol.gov/agencies/wb/about/history`, `https://www.bls.gov/opub/mlr/2002/05/art2full.pdf`.

2020) and gender gaps in science commercialization Koffi and Marx (2021).

For a long time, technologically innovative professions did not include women. In 1940, less than one percent of US engineers were women, and the share remained at one percent for the following two decades (Pidgeon, 1954). Perhaps not surprisingly, women were also absent from PhD programs in technical fields. In 1960, 6 of 959 PhDs US in engineering were awarded to women. In 1949, women earned 175 of 52,246 engineering bachelor's degrees, 15 of 4,496 engineering master's degrees, and just 1 of 417 engineering doctorates. Over the next few decades, women's participation in technical fields expanded. By 1980, women earned 10 percent of engineering bachelor's degrees, 8 percent of engineering master's degrees, and 4 percent of engineering doctorates.

The concern about lost Einsteins and Curies is accompanied by evidence that some innovative potential is untapped. Greater exposure to innovation appears to make women more likely to become inventors (Bell et al., 2019; Cook, 2011; Hoisl et al., 2023; Aghion et al., 2017). Moreover, increasing female inclusion in innovative activity could have various benefits including a changed direction of innovation (Koning et al., 2021). Increased female participation in book authorship appears to have delivered substantial growth in the number of new books valued by diverse readers (Waldfogel, 2023). Outside of the creative industries, however, little is known about the effects of more inclusive participation in the sort of inventive activity reflected on patents.

3 Theory: the female influx and its effect

3.1 Motivation

As we will see below, there has been a female influx into patent invention. How would we expect such an influx to affect the value of patents? The answer depends on both the nature of selection into inventive activity and the ex ante predictability of invention quality.

In the usual way of thinking about entry, people compare their expected benefits to costs and take action (become an inventor, file a patent) if the expected benefit exceeds the cost. Historically, fewer women than men have become inventors; and women have participated in the labor force less than men generally. Low female participation can be rationalized by higher implicit costs of female participation. These higher costs could reflect gatekeeper bias, or they might arise from social pressures to engage in nonmarket activity. Either way, women may have faced higher entry costs than men, and the female influx is consistent with a falling female participation threshold.

As the cost threshold falls, what sorts of female inventors might enter? To the extent that inventor, or invention, quality is predictable prior to investment, we might expect the quality of marginal inventors to fall. If invention quality were completely predictable at the time an inventor were contemplating undertaking costly effort, then relaxation of a participation cost threshold from T to $T^{'}$ would bring marginal inventions with expected – and realized – value between T and $T^{'}$. The value of a female influx in such an environment would be positive but small.

It is well understood, however, that invention quality is unpredictable (Scherer and Harhoff, 2000). At the extreme, if it were entirely unpredictable, then marginal inventions would be as good, on average, as inframarginal inventions. This is not the only mechanism that could lead to innovations from a female influx to have high value. It is possible that social conventions inhibiting female participation have historically decoupled selection into inventive activity from expected benefits. In that case, a female influx would not be expected to bring inventors of lower quality than inframarginal female inventors, even if innovation quality were predictable. Both because of unpredictability and beneficial selection, the female influx could deliver highly valuable innovations. With unpredictable innovation quality, the welfare effect of an influx of innovators – and innovations – is similar to the welfare effect of a large number of additional products (Aguiar and Waldfogel, 2018).

The foregoing suggests a natural way to seek evidence on the value of a female influx. As the female share of patents has risen, has the share of value contributed by female patents also grown? Addressing these questions is aided by explicit models of demand (for inferring welfare) and supply/entry (for characterizing possible equilibrium responses to the counterfactual elimination of the female patent influx.

3.2 Model

My approach makes an analogy to demand and entry. I treat inventors as buyers of inputs into the creation of patents. The inputs that they buy are the citations that they make to extant patents. The demand side of the model is inventors making choices to cite those existing patents. The supply, or entry, side of the model is the decision of inventors to create inventions, which manifest themselves as patents.

3.2.1 Demand

Inventors draw on existing patents to create new ones. While there is no explicit price of citing existing patents, there is implicitly a positive price, as inventors cite finite numbers of patents. I model inventors as making a series of discrete choices of whether to cite patents, and I use plain logit in the text to ease description. The utility to inventor i of citing patent j is given by

$$u_{ij} = x_j\beta + \xi_j + \epsilon_{ij},$$

where x_j includes characteristics of patent j (technology category, whether it has female inventors, vintage, etc.), ξ_j contains aspect of patent j's quality that are not observed by the researcher, and ϵ_{ij} is an extreme-value error. The utility of the outside good, not citing a patent, is given by $u_{i0} = 0$. Define the quality of patent j as $\delta_j \equiv x_j\beta + \xi_j$.

If M is market size, and q_j is the number of citations to patent j, then $\delta_j = \ln(s_j) - \ln(s_0)$,

where $s_0 = 1 - \Sigma_{j \in J} s_j$, and

$$s_j = \frac{q_j}{M} = \frac{e^{\delta_j}}{1 + \Sigma_{i \in J} e^{\delta_j}},$$

where J is the set of extant patents. Then "inventor surplus" – the analog to consumer surplus – is proportional to

$$\ln(1 + \Sigma_{j \in J} e^{\delta_j}).$$

It is useful to distinguish between male- and female-inventor patents. Define J_f and J_m as female-inventor and male-inventor patents. Then female-inventor patents as a group receive the following share

$$s^f = \Sigma_{j \in J_f} s_j = \frac{\Sigma_{j \in J_f} e^{\delta_j}}{1 + \Sigma_{i \in J_f} e^{\delta_j} + \Sigma_{i \in J_m} e^{\delta_j}}.$$

There is an analogous measure for the male-inventor patents.

3.3 Supply

An influx of patents from female inventors brings a set of quality draws, or δ_j's. The effect of additional female patents on the value of available patents depends on how the distribution of additional δ_js relates to the distribution of existing qualities. It is well understood that realized quality is an unpredictable random variable Scherer and Harhoff (2000); an important question for the welfare effect of additional patents is whether patent quality is predictable at the time of entry or, more accurately, how predictable it is.

Suppose, initially, that quality is completely unpredictable. Then, in expectation, δ_j does not vary with the number of entering patents. Additional patents would be as good, on average, as existing patents. Suppose there are N_f female-inventor patents. Then define $D_f \equiv \Sigma_{j \in J_f} e^{\delta_j} = N_f e^{\bar{\delta f}}$. As a result, the market share for the marginal entrant is the same as the market share of the average entrant. It is given by

$$s^f = \frac{e^{\bar{\delta f}}}{1 + \Sigma_{i \in J_f} e^{\delta_j} + \Sigma_{i \in J_m} e^{\delta_j}}.$$

If patent quality is entirely unpredictable, then adding new patents should substantially increase the value of patents. Moreover, simulating an environment without the female influx can be accomplished by simply removing female-inventor patents from the choice set at random. If N_f^0 is the counterfactual number of female-inventor patents absent the female influx, then the proportionate increase in CS from the influx is given by

$$\frac{CS}{CS^0} = \frac{\ln(1 + N_f e^{\bar{\delta f}} + N_m e^{\bar{\delta m}})}{\ln(1 + N_f^0 e^{\bar{\delta f}} + N_m e^{\bar{\delta m}})}.$$

It is also possible that patent quality is somewhat predictable at the time of invention, or that

$$\delta_j = z_j \gamma + \varepsilon_j,$$

where $\delta_j^{'} \equiv z_j \gamma$ is a prediction of quality at the time of invention/filing, and ε_j is the quality realization error. In that case it is natural to assume that patents would have entered in order of expected quality. Indeed, it would be a Nash equilibrium to have all patents with expected value above some threshold to enter. In that case, because investment costs are sunk, no already-existing patents wish to exit. And because an additional patent would have expected value below the threshold, its inventor would not create it.

With partial predictability, D_f would not longer be linear in expectation. Instead, since better patents would enter earlier in the order, there would be a nonlinear function relating the number of entering patents to the sum of the exponentiated qualities $D_f = D_f(N_f)$, with $D_f^{'} > 0$ and $D_f^{''} < 0$. The degree of predictability would determine the exact shape. Greater predictability would give rise to curves that would have greater initial slope. Complete unpredictability would give $D_f^{'} > 0$ and $D_f^{''} = 0$.

More generally, the proportionate increase in CS from the influx is given by

$$\frac{CS}{CS^0} = \frac{\ln(1 + D_f(N_f)) + D_m(N_m)}{\ln(1 + D_f(N_f^0) + D_m(N_m))}.$$

Implementing this requires suitable functional forms. One natural alternative is $D(N) = aN^b$. When $b = 1$, this is consistent with no predictability, and $\bar{\delta} = a$. When patent quality is somewhat predictable, then $b < 1$.

3.3.1 Equilibrium and counterfactual solution

First, I need to calculate the expected citations for the marginal entrant in the status quo, separately for male and female-invented patents. Call these K_f and K_m . I calculate $K_f = \frac{e^{\Delta_f}}{1+D_f(N_f)+D_m(N_m)}$ where $\Delta_f = D_f(N_f) - D_f(N_f - 1)$. That is, Δ_f is the expected quality of the marginal entering female patent. Analogously, $K_m = \frac{e^{\Delta_m}}{1+D_f(N_f)+D_m(N_m)}$.

Define functions $s_f(N_f, N_m)$ and $s_m(N_f, N_m)$, which (when weighted by market size) show the expected numbers of citations received by the marginal entering female and male patent, respectively. In the status quo, equilibrium arises when the expected "revenue" (citations) to the marginal patent covers costs, but one more does not: $Ms_f(N_f, N_m) > K_f > Ms_f(N_f + 1, N_m)$ and $Ms_m(N_f, N_m) > K_m > s_m(N_f, N_m + 1)$.

Removing the female influx means reducing N_f to N_f^0 (which I implement below as bringing the female share of patents to its relative share in 1976). The raises the return to additional female patents, but they are held fixed for exogenous reasons (reflecting the forces keeping women out of science in the past).

The male return also rises, and I allow male entry to equilibrate – rise to N_m^1 – so that $Ms_m(N_f^0, N_m^1) > K_m > s_m(N_f^0, N_m^1 + 1)$.

Then the inventor-side welfare effect of the female influx into patenting is given by

$$\frac{CS}{CS^0} = \frac{\ln(1 + D_f(N_f) + D_m(N_m))}{\ln(1 + D_f(N_f^0) + D_m(N_m^1))}.$$

Allowing endogenous male entry response is conservative in that raises the value of patents available in the no-female-influx counterfactual.

3.4 Empirical implications

The welfare effects of a female patent influx depend on three broad things: First, it depends on the predictability of quality at entry which, in turn, determines how good additional patents are in relation to existing patents. Second, welfare effects depend on the extent of substitutability of existing for new patents. Finally, welfare effects depend on whether the influx delivers net new patents or displaces patents that would otherwise come into existence. The questions can all be explored descriptively as a means to prepare for modeling. Prior to estimating an explicit structural model (in Section 6), I explore a number of descriptive questions in search of evidence on the welfare effects of the female influx.

3.4.1 The influx and the share of citations garnered by female-inventor patents $\frac{\partial s_f}{\partial n_f}$

One of the empirical questions below is whether the influx of female patents has brought valuable patents to the patent stock. The framework above makes clear how this might work. For concreteness, suppose we parameterize $D_f = aN_f^b$ and $D_m = cN_m^d$. How would an increase in the female share of patents ($n_f \equiv {}^{N_f}/_{N_f+N_m}$)affect the share of citations garnered by female-invented patents (s_f)? This will depend on the relative values of the parameters $\{a, b, c, d\}$, but we can start with a case in which $b = d = 0$, and $a = \bar{\delta^f} = c = \bar{\delta^m}$. In this case

$$\frac{\partial s_f}{\partial n_f} = 1.$$

More generally, it depends on all of the parameters. Still, it is interesting to measure

$$\frac{\partial s_f}{\partial n_f} = 1.$$

A high value indicates that a female influx is valuable.

One of the descriptive empirical questions pursued below is the measurement of $\frac{\partial s_f}{\partial n_f}$.

3.4.2 Market expansion

The welfare effect of an influx of female patents depends on the extent to which additional patents are substitutes for existing patents, as well as what would have happened to entry absent the female influx. Growth in the number of female patents, if they are valuable, should "expand the market." That is, growth in female-invented patents should make extant patents more valuable and should raise usage.

A challenge with documenting this is that citation practice may also change over calendar time (not just across vintages with different numbers of female-invented patents). A natural way to test is to make comparisons across not only vintages but also categories. Do categories with larger growth in patents generally (and female patents as well) experience larger growth in citation to those patents?

3.4.3 Displacement

If inventors add patents in expectation of receiving benefits, and if patents are substitutes for one another, then it is possible that inventive activity by one group displaces inventive activity by other groups. It is also possible that the demand for innovative activity is sufficiently high that growth in female inventor activity does not displace male invention activity. Below I explore whether the grow

4 Data

The data for this study consist of 6,347,032 patents granted in the US between 1976 and 2022, along with all of their citations to other US patents granted over the same period. The data include 92.2 million citations to 5,047,238 patents receiving citations. For each citing and cited patent, I know its WIPO technology category, its grant year, and the number of its inventors who are women. I classify patents into 35 technology categories according to their first listed WIPO category.[3] The data are obtained from the USPTO.[4] Of the cited patents, 8.27 percent have mostly female inventors (half of more of the inventors), and 17.59 percent have at least one female inventor. Including the patents receiving no citations, the patents are cited an average of 13.9 times with a median of 4. The mostly-female patents receive 6.68 percent of these citations, while the any-female inventor patents receive 16.80 percent of the citations. Figure 3 shows the number of patents by WIPO category. Most categories have over 100,000 patents. The right panel shows the female inventor shares by category. Using the any-female definition, these vary between under 10 percent (civil engineering, mechanical elements) and nearly 40 percent (biotechnology, pharmaceuticals). Female inventor shares are highest in pharmaceuticals and biotechnology, and this may reflect the relatively high female shares of biology and health PhDs granted. Female inventor shares are lowest in categories related to engineering, mirroring the pattern for PhDs granted.

Figure 4 shows the evolution of the female-invented patents granted over time. Of the patents granted in 1976, under 5 percent are mostly female, and this measure rises to 12

[3]The resulting categories are analysis of biological materials; audio-visual technology; basic communication processes; basic materials chemistry; biotechnology; chemical engineering; civil engineering; computer technology; control digital communication; electrical machinery, apparatus, energy; engines, pumps, turbines; environmental technology; food chemistry; furniture, games; handling; IT methods for management; machine tools; macromolecular chemistry, polymers; materials, metallurgy; measurement; mechanical elements; medical technology; micro-structural and nano-technology; optics; organic fine chemistry; other consumer goods; other special machines; pharmaceuticals; semiconductors; surface technology, coating; telecommunications; textile and paper machines; thermal processes and apparatus; and transport.

[4]See `https://patentsview.org/`.

percent by 2020. About 7 percent of the 1976 patents have at least one female inventor, and the share of patents with at least one female inventor rises past 25 percent by 2020. This increase in the female inventor share appears across virtually all technology categories. Figure 5 shows the growth rate of the female share based on regressions of the log female share on vintages. Using the any-female measure, the female inventor share rises by over 3 percent per year in nearly half of categories and by more than 2 percent per year in all but one of the remainder.

While male and female patents have different average citations, these differences are swamped by the variability across patents in the number of citations received. Figure 6 illustrates the similarly substantial variability in realized patent quality with the realized distributions of eventual sample citations (by 2020) to male and female patents granted in 2000. The inter-quartile ranges for both all-male-inventor and any-female-inventor patents both run from 6 to 33.

Table 2 shows patterns of citation by gender of citing and cited patents, and male and female-invented patents have different tendencies to cite female-inventor patents. Of the 71.8 million citations from all-male-inventor patents, 15.3 percent are to patents with at least one female inventor. Of the 20.5 million citations from patents with at least one female inventor, 22.0 percent are to patents with at least one female inventor. In the bottom panel (based on the mostly female definition), 6.4 percent of citations from male patents are citations to female patents, while 11.2 percent of citations from mostly female patents are to mostly female patents. There is therefore a fairly substantial difference in the tendencies of male and female inventors to cite female-invented patents. Of course, the cross tabulation alone cannot determine whether these patterns are due to gender or to, say, the mix of technology patterns by gender.

4.1 Measures and notation

Documenting the impact of the female influx into patent invention on the value of resulting patents requires a measure of patent value. For this I follow the literature (Hall et al., 2005; Kogan et al., 2017) in using the citations to patents as a proxy for their value. In particular, I measure the relative value of a set of cited patents by the share of citations that those patents receive. For example, I assume that the share of citations to female-invented patents, among the total patents granted in some vintage, reflects the share of the value of those patents attributable to the female-invented patents.[5]

Some notation is helpful for developing specific hypotheses about the effect of a female influx. Each *cited* patent has three characteristics of interest here: its vintage (grant year) v, the gender of its inventor(s) (g), and its WIPO technology category (w). Each *citing* patents also has these three characteristics. I refer to the grant year of the citing patent using t rather than v to avoid confusion. I employ a notation that distinguishes the *citing* patents (which I view as the demand side) from the *cited* patents (the supply side): I use subscripts to define characteristics of the cited patents, and I use superscripts to define characteristics of the citing patents. I use Q to denote numbers of citations and N for numbers of patents.

For example, N_v is the number of patents granted in year v, N_{vw} is the number of patents in technology category w granted in year v, and N_{fvw} is the is the number of patents in technology category w granted in year v to patents I classify as having female inventors. I define the female-invented share of patents from a group with lower-case n. For example, $n_{fvw} = \frac{N_{fvw}}{N_{vw}}$. In words, n_{fvw} is the female-invented share of patents granted in year v and technology class w.

I define Q_{vw} as the number of citations to patents granted in vintage v and technology

[5] Jaffe and De Rassenfosse (2019) presents a thoughtful discussion of the use of citations to measure patent quality. Many of the challenges relate to inferring the value of specific patents from citations. Here, by contrast, I am trying to measure the share of value attached to a subset of patents from the aggregate share of citations that those patents receive. I can therefore rely on the law of large numbers for averaging out the errors in the way that citations reflect the value of each particular patent.

category w, and I define Q_{fvw} as the citations to patents in technology category w and vintage v with female inventors. Then s_{fv} is the share of patent citations garnered by female patents among those in vintage v, so $s_{fv} = \frac{Q_{fv}}{Q_v}$. Measures of citations are more complicated than measures of the numbers of patents, as the number of citations to patents in a cited group will vary across citing years. That is, the number of patents granted in, say, 1980 remains the same during subsequent years. The numbers of citations to the 1980 patents, by contrast, vary over time. For example, patents granted in year t cite patents granted in year v, so that Q^t_{vw} is the citations from patents granted in year t to patents granted in year v and in category w. I define Q^t_{fvw} as the citations to the female-invented patents among them, so that $s^t_{fvw} = \frac{Q^t_{fvw}}{Q^t_{vw}}$.

Citing patents vary not just by their grant years t but also by the gender of their inventors and the citing patent's technology categories. We may gain insight into the effects of the female influx by examining potentially varying tendencies for disparate inventors – by gender or technology class – to cite female-inventor patents. Accordingly, it is useful to define measures like the year-t female-inventor tendency to cite female-inventor patents from vintage v. This is given by $s^{tf}_{fv} = \frac{Q^{tf}_{fv}}{Qv^{tf}}$. Similarly, it is of interest to see how the tendency to cite female-inventor patents varies across citing patents (and inventors) in different technology categories. For example, of the citations from patents granted in year-t and technology class w to patents from year v and technology class w, s^{tw}_{fvw} is the share of the citations to female-invented patents (from year v and class w). That is, $s^{tw}_{fvw} = \frac{Q^{tw}_{fvw}}{Q^{tw}_{vw}}$.

Below, I will explore the effect of the female influx into invention using the empirical relationship between s_f and n_f across various groups of patents, as well as for subgroups of citing inventors.

5 Descriptive analysis

5.1 The supply of female inventors and the female patent influx

By the early 1980s women were taking the Graduate Record Exam (GRE) in greater numbers than men. In 1982, 71,845 of the US test takers were men, while 95,235 were women. After 2000, the majority – over 60 percent – of test takers were women. Although average scores for women are lower than averages for men, there is substantial dispersion. The Educational Testing Service reports GRE means and standard deviations by gender for US test takers. Assuming normality, one can calculate the numbers of men and women with scores above any absolute level. As Figure 1 shows, the influx of women into test taking roughly doubled the number of US test takers receiving a quantitative score over 700. Shifting tendencies to go to college and to pursue graduate education provided a rightward shift in the supply curve for technical talent.

Growing female educational attainment has set the stage for growing participation in innovative activity. In addition to growing participation in the economy, women have attained more education over time: The female share of bachelor's degrees granted grew from 15 percent in 1870, 34 percent in 1920, 43 percent in 1970, and 58 percent in 2020.[6] This trend extends to graduate education, including the sort of education potentially leading to innovative activity. Figure 2 shows how the female shares of new PhDs have evolved over the past half century, as reported by the US Department of Education.[7] The female share of PhD grantees varies across fields but has risen in all of them since 1960. At the high end, the female shares are over 75 percent in psychology, nearly 70 percent in education, and roughly 55 percent in the arts. Among the fields more directly relevant to patents, the female share exceeds 50 percent in health and biology and by 2018 had reached roughly 20-25 percent

[6] See Table 156, `https://nces.ed.gov/pubs91/91660.pdf` and `https://nces.ed.gov/programs/digest/d21/tables/dt21_318.10.asp?current=yes`.

[7] See Table 352, `https://nces.ed.gov/programs/digest/2021menu_tables.asp`.

in math, computer science, and engineering. It is clear that the cadre of doctorate holders, including those in STEM fields, is growing more female over time. This raises the possibility of harnessing the innovative potential of a previously less engaged group.

The growing numbers of prepared women has given rise to the growing share of female-authored patents, reported above in Figure 4.

5.2 Displacement?

Even if new patents from female inventors garner citations, it is possible that the additional female-invented patents displace male-inventor patents that would otherwise have come into existence. At the same time, there may be sufficient demand for inventive ability that the female-induced growth in the number of highly qualified people (based on GRE scores and PhDs) need not displace male inventive activity. To explore this, I regress the number of male patents on female patents in each vintage. First, I do this by vintage v.

$$n_{fv} = \theta_0 + \theta_1 n_{mv} + \epsilon_v$$

Second, I estimate the relationship between male and female patenting within vintages and technology categories, using both measures of female patents and the following estimating equation:

$$n_{fvw} = \gamma_0 + \gamma_1 n_{mvw} + \mu_v + \mu_w + \epsilon_v.$$

See Table 5. All specifications yield positive coefficients. I do not take this to mean that female patenting stimulates male patenting; rather, I suspect that both male and female patenting are driving by changing promise of innovation in different areas. Still, there is no evidence of displacement. Both male and female patenting are growing.

I undertake two kinds of counterfactuals below. First, compare a status quo vintage

containing the female influx with a counterfactual environment that removes the female patent influx. Second, I remove the female influx and allow for an equilibrium expansion of male-invented patents. It is not clear to me a priori which is the most reasonable exercise to undertake, but they may together provide bounds on a reasonable estimate.

5.3 Market expansion

The welfare effect of an influx of patents depends in part on whether existing patents serve as substitutes for the new ones. This is the patent analogue to the usual question of "business stealing" vs "market expansion." If new patents were close substitutes for existing patents, then the benefit of new patents would be small, and growth in the number of extant patents would not induce more use of patents overall. We can get some evidence on this by examining citation patterns over time and across patents as the number of patents grows. It is possible that citation patterns change over time, so the cross-year "experiment" is potentially uninformative.

Still, there is differential growth in the number of patents available to be cited across different technologies. This allows me to ask whether citations from patents granted in year t to extant patents in category w (Q_w^t) grow more quickly in categories where the number of extant patents (N_w^t) is growing faster. To this end I run regressions of the form:

$$Q_w^t = \lambda N_w^t + \mu_w + \mu^t + \varepsilon_w^t,$$

where the μ terms are fixed effects for cited technology category w and citing year t, and ε_w^t is an error term.

Table 6 reports variants on this regression. Columns (1) and (2) use levels for Q and N, while columns (3) and (4) use logs. The columns labelled "OLS" lack fixed effects, while the columns labelled "FE" include w and t fixed effects. All of the estimates of the λ

coefficients are positive, indicating that growth in extant patents in a technology category attracts growth in citations. This is at least suggestive of market expansion which, in turn, indicates that an influx of patents provides something not substituted for by what is already present.

5.4 Are female-inventor patents as cited as male-invented patents?

My goal is to measure the impact of the female influx into patenting on the value of the female-invented patents. To this end, I seek to measure $\frac{\partial s^f}{\partial n^f}$ – which I term "α" – across groups of patents, for example by vintage, or by vintage and technology category. Essentially, this measures how useful are marginal patents in relation to inframarginal female patents.

The *cited* patent data can be distinguished by inventor gender, vintage v, and technology w. Hence, the female-invented share of patents (n_f) can be disaggregated by vintage as n_{fv} or by vintage and technology category as n_{fvw}. I seek to measure how the tendency of *citing* patents to cite female-invented patents, or s_f varies with n_f. Like the cited patents themselves, the flows of citations from citing to cited patents can be disaggregated by cited patent characteristics. That is, I can measure the shares of citations to female-invented patents of vintage v (s_{fv}) or the shares of citations to female-invented patents among those in technology category w (s_{fvw}). Then I can estimate α using a regression such as:

$$s_{fv} = \alpha n_{fv} + \epsilon_v. \tag{1}$$

This approach identifies the coefficient α from the cross-vintage relationship between the vintage's female inventor share (n_{fv}) and the tendency for subsequent patents (and inventors) to cite the female patents among those in each vintage.

This approach is vulnerable the concern that vintage quality varies over time in a way that is correlated with, by not caused by, the changing female share of patent inventors.

Disaggregating to $v \times w$ allows inclusion of vintage and technology category fixed effects, via:

$$s_{fvw} = \alpha n_{fvw} + \mu_v + \mu_w + \epsilon_v. \tag{2}$$

Equations 1 and provide my basic approaches for measuring the effect of the female influx on the overall tendency to cite female-invented patents.

Because I know the inventor gender, technology category, and grant year of the citing patents, I can also estimate the coefficient α separately by these dimensions. I can estimate separate coefficients by citing patent inventor gender g $(g = m, f)$ using:

$$s^g_{fvw} = \alpha^g n_{fvw} + \mu^g_v + \mu^g_w + \epsilon^g_v, \tag{3}$$

which, given the inclusion of $g \times v$ and $g \times w$ fixed effects, is effectively citing patent inventor gender-specific regressions. A comparision of estimated α^f and α^m indicates the relative usefulness of the female influx to female vs male inventors of the citing patents.

Similarly, I explore the potentially different αs across citing patent inventor technology categories using:

$$s^w_{fvw} = \alpha^w n_{fvw} + \mu^w_v + \mu^w_w + \epsilon^w_v, \tag{4}$$

and I explore whether the the usefulness of female-invented patents to citing patent inventors changes over time using:

$$s^t_{fvw} = \alpha^t n_{fvw} + \mu^t_v + \mu^t_w + \epsilon^t_v. \tag{5}$$

Finally, I explore the evolution of α by gender and time using:

$$s^{gt}_{fvw} = \alpha^{gt} n_{fvw} + \mu^{gt}_v + \mu^{gt}_w + \epsilon^{gt}_v. \tag{6}$$

5.5 The overall value of the female influx

Table 3 reports regressions of f_f on n_f. The first two columns implement Equation **??** using the any-female and mostly-female definitions, respectively, and delivers estimates of α of 1.081 (se = 0.251) and 0.0670 (0.0336). The latter two columns use the data disaggregated to vintage by technology category, and the analogous estimates of α are 1.056 (0.0326) and 1.039 (0.0791). In all but one of the specifications, I cannot reject that $\alpha = 1$. This is consistent with the female influx delivering patents that are as valuable, on average, as the pre-existing female-authored patents.

5.6 The female influx and heterogeneous downstream inventors

Results above indicate that the female influx into inventive activity delivers patents that are useful to other inventors. I can go further by measuring the impact of the influx of various kinds of inventors, differing by gender and technology category. The goal of these regressions is to measure how useful the female influx is to male vs female inventors, and to inventors whose patents are in various technology categories.

5.6.1 α over time

By estimating Equation 5, we can explore whether the value of the female influx has varied to citing inventors over time. As Figure 7 shows, α is roughly stable over time.

5.6.2 Citing inventor gender

Table 4 reports estimates of variants of Equation 3. The first two columns using the data aggregated to vintage and gender of citing patent inventors. The coefficients for female inventors are substantially higher than for male inventors. This pattern continues in the latter two columns, which use data aggregated to vintage and technology category of cited

patents and citer gender. The pattern of coefficients indicates that the female inventor influx delivers a greater benefit to downstream female inventors than to male inventors.

Figure 8 shows the α coefficients by citing patent inventor gender over time. The female coefficients are persistently above the male coefficients, but neither series has an apparent trend.

5.6.3 Citing inventor technology category

Figure 9 presents the α estimates by technology category based on Equation 4. Inventors in different categories derive different degrees of benefit from the female influx. The coefficients are high, using both female measures, for "furniture, games" and for "textile and paper machines" and systematically low for nanotechnology.

5.7 Other measures of female success outcomes

As the female share of patent inventors rises, what happens to other measures of female success in science and engineering, such as major scientific honors? I have annual inductees into the National Academies of Science, Medicine, and Engineering, as well as annual honorees with the US National Medal of Science. See Figure 10. Each of these shares rises from zero, or nearly zero, in 1960 to about a quarter in 2020 for the National Medal of Science and the National Academy of Engineering and beyond 40 percent for the National Academies of Science and Medicine. To the extent that awards are similarly reflective of accomplishment for male and female researchers, the growing female shares of award recipients corroborate the growing contribution to innovation documented in the patent data.

Table 7 shows the relationship between the female share of inductees for each outcome and the share of patents with at least one female inventor. Each of the coefficients is positive and significant: As the female share of inventors has risen, so has the female share of recipients of these elite honors. These results indicate the female influx also affects right-tail success

outcomes.

6 Measuring welfare effects

The evidence about is clearly suggestive of welfare benefits from the female influx into patenting. What the exercises above lack is a method for quantifying welfare. I am interested in measuring the welfare benefit delivered by the female influx into patenting. This, in turn, requires both a way of valuing the "choice set" of extant patents to inventors, as well as specification of the counterfactual choice set of patents that would exist absent the female influx.

I make an analogy between the decision to cite a patent and a consumer's product choice problem, with a few adjustments. First, an inventor makes multiple choices to cite patents, as patents tend to cite more than one patent. Second, unlike a consumer who faces a price for purchasing a product, an inventor faces no explicit price for citing a patent. I assume that the price of citing patents is equal across patents.

The exposition that follows uses the plain logit model for clarity, but the model that I implement uses the nested logit. An inventor i filing a patent makes a series of discrete choices to cite one of the J_t extant patents during year t. In particular, for each of the instances in which an inventor cites a preceding patent, the inventor i chooses a patent j to maximize utility based on characteristics of the cited patent j as well as an extreme value error.

$$u_{ij} = x_j\beta + \xi_j + \epsilon_{ij} \equiv \delta_j + \epsilon_{ij}$$

This gives rise to expressions for the market share for each cited patent:

$$s_j = \frac{e^{\delta_j}}{1 + \Sigma\delta_j}.$$

Moreover, this setup allows for welfare analysis of the choice set of extant patents to cite. The set of extant patents is changed by the female influx into inventive activity, and I compare the status quo citeable patent choice set in 2020 to a hypothetical choice set with fewer female-invented patents. The consumer surplus for a choice set of patents is given by

$$CS = \frac{M}{\alpha} \ln(1 + \Sigma e^{\delta_j}).$$

Here, α is normally a price coefficient in the utility function, and M is market size. We have no price, so we cannot estimate the dollar value of inventors' consumer surplus. We can, however, measure the proportionate change in CS from a change in the choice set of available patents.

One way to simulate the absence of a female influx is to reduce the female share of patents in each vintage to the female-invented share of patents for the first year in the sample, 1976. This allows the number of female-invented patents granted to grow over time, but it keeps the female share of each cohort at the 1976 level. It is further necessary to decide which female patents to remove from the choice set. If patent quality were entirely unpredictable, then I could remove female patents at random. Yet, the evidence above (that $\frac{\partial s}{\partial n}$

6.1 Demand

Define Q_{jt} = citations to patent j during year t, M_t = market size in year t, which I define as the number of patents granted in year t, times a scalar (100) representing the number of possible citations.

Then $s_{jt} = \frac{Q_{jt}}{M_t}$, $s_t = \Sigma_{j \in t} s_{jt}$, $s_{0t} = 1 - s_t$, and the inside share is given by $s_{jt|in} = \frac{s_{jt}}{s_t}$.

Then I can estimate the nested logit model as:

$$\ln s_{jt} - \ln s_{0t} = x_{jt}\beta + \sigma \ln s_{jt|in} + \xi_{jt}$$

as in Berry (1994). The parameter σ reflects the degree of substitutability across patents. The vector x_j contains patent characteristics such as vintage, the gender of the inventor, and the technology category. Estimation of this model requires an instrument for the inside share. Natural candidates include the number of available patents to cite. Define N^t as the number of extant patents, granted up to year t.

I am interested in potentially different substitution patterns for male vs female inventors. Accordingly I estimate the model separately for male and female inventors. For example, I can define each of the shares s according to citations from female-invented citing patents. Table 8 reports OLS and IV estimates of the parameters, as well as first stages.

Table 9, separately by citing inventor gender. Using OLS, the male substitution parameter is 0.928, while the female parameter is 0.783. The IV estimates indicate lower substitutability: 0.523 for male inventors and 0.343 for female inventors. Lower substitution parameters indicate that growth in patents expands the market more. That is, they raise total citations rather than just diverting citations from other patents.

6.2 Supply/entry

The supply side of the model depends on the predictability of quality at entry. Patents are assumed to enter in order of expected quality. I implement this by predicting realized quality. To this end, I regress realized quality of a patent cohort on characteristics of the patents, and their inventors, known at the time of entry. That is, I regress δ_j on measures of the numbers of patents, and patent citations, received by inventors on the current pattern, as well as technology category indicators. The regression's R-squared is roughly 10 percent.

Because there are some significant coefficients – and because R-squared is not zero – there is some predictability; but the predictability is far from complete. The predicted quality of each patent is then used to sort patents into an entry order.

To solve the entry model, I need empirical measures of expected $D_f(N_f)$ and $D_m(N_m)$. I obtain $D_f(N_f)$ as follows. I sort the female-invented patents in the 2020 choice set according to predicted quality δ_j'. Using this ordering – and illustrating with plain logit – the realized $d_f(N_f) = \Sigma_{j \in J_f}^{N_f} e^{\delta_j}$.

I obtain my measured of the expected sum $D_f(N_f)$ with the following parameterization. I regress the vector of $\ln(d_f(n_f)$ on $\ln(n_f)$ and a constant. The resulting two parameters then allow me to parameterize $D_f(N_f) = \alpha_f N_f^{\beta_f}$. I do the analogous exercise for male-inventor patents.

I simulate an environment without the female influx in two ways. First, I simply remove female-invented patents from each vintage (and technology category) to bring the ratio of female to male patents in the vintage to its 1976 level. That is, I remove female-invented patents without adding male patents. This is a reasonable characterization if we expect no displacement relationship between male and female-invented patents.

Second, I allow for a male inventor entry response. Using my estimated $D(N)$ functions, I solve for the number of male patents needed, absent the female influx, to bring the expected citations of the marginal male patent to the expected citations for the marginal entering male patents in the presence of the female influx. This equilibrium approach will generate smaller welfare effects of the female influx, as the counterfactual environment without female patents has some compensation in additional male patents.

Table 10 reports results from the nested logit simulations with imperfect predictability. Relative to a world without the female influx, the influx raises overall inventor CS by 2.7 percent. Female inventors gain 4.6, while male inventors gain 2.1 percent. Citations to female-invented patents are much higher (38.6 percent) with the influx, while citations to

male patents are 3.9 percent lower. When I allow for a counterfactual male inventor patent response to the elimination of the female influx, CS effects are muted. The influx raises overall CS by 1.2 percent (2.8 percent for female inventors, and 0.7 percent for male inventors). It is important to note that in these simulations, the influx raises the female share of patents from about 10 percent to about 20. A continued growth in female participation could of course raise welfare more.

7 Conclusion

Large segments of the population have traditionally been absent from innovation activity, raising a concern that society has missed out on valuable innovations because of insufficient inclusivity. Between 1976 and 2020 the share of patents with at least one female inventor has risen from under a tenth to over a quarter, making it possible to measure the impact of more inclusive innovation on the value of the additional innovations. I document growing numbers of women prepared to innovated in technological fields and I document that as the share of patents in a vintage and technology category cell have risen, the share of the citations received by female-invented patents in the cell has risen nearly proportionally. Despite a persistent gender differential in the tendency to cite female-inventor patents, the absence of evidence that new female-invented patents displace additional male patents, indicates that the female influx into patenting has delivered valuable innovations. Direct welfare measurement shows that the "consumer surplus" enjoyed by inventors is between 1.2 and 2.7 percent and even higher from the perspective of female inventors, who make greater use of female-invented patents.

References

Aghion, P., U. Akcigit, A. Hyytinen, and O. Toivanen (2017): "The social origins of inventors," Tech. rep., National Bureau of Economic Research.

Aguiar, L. and J. Waldfogel (2018): "Quality predictability and the welfare benefits from new products: Evidence from the digitization of recorded music," *Journal of Political Economy*, 126, 492–524.

Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen (2019): "Who becomes an inventor in America? The importance of exposure to innovation," *The Quarterly Journal of Economics*, 134, 647–713.

Berry, S. T. (1994): "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, 242–262.

Cook, L. D. (2011): "Inventing social capital: Evidence from African American inventors, 1843–1930," *Explorations in Economic History*, 48, 507–518.

Costa, D. L. (2000): "From mill town to board room: The rise of women's paid labor," *Journal of Economic Perspectives*, 14, 101–122.

Ding, W. W., F. Murray, and T. E. Stuart (2006): "Gender differences in patenting in the academic life sciences," *science*, 313, 665–667.

Fechner, H. and M. S. Shapanka (2018): "Closing diversity gaps in innovation: Gender, race, and income disparities in patenting and commercialization of inventions," *Technology & Innovation*, 19, 727–734.

Fernández, R. (2013): "Cultural change as learning: The evolution of female labor force participation over a century," *American Economic Review*, 103, 472–500.

Frietsch, R., I. Haller, M. Funken-Vrohlings, and H. Grupp (2009): "Gender-specific patterns in patenting and publishing," *Research policy*, 38, 590–599.

Goldin, C. (2006): "The quiet revolution that transformed women's employment, education, and family," *American economic review*, 96, 1–21.

Greenwood, J., A. Seshadri, and M. Yorukoglu (2005): "Engines of liberation," *The Review of Economic Studies*, 72, 109–133.

Hall, B. H., A. Jaffe, and M. Trajtenberg (2005): "Market value and patent citations," *RAND Journal of economics*, 16–38.

Hoisl, K., H. C. Kongsted, and M. Mariani (2023): "Lost Marie Curies: Parental impact on the probability of becoming an inventor," *Management Science*, 69, 1714–1738.

Hoisl, K. and M. Mariani (2017): "It's a man's job: Income and the gender gap in industrial research," *Management Science*, 63, 766–790.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2019): "The allocation of talent and us economic growth," *Econometrica*, 87, 1439–1474.

Hunt, J., J.-P. Garant, H. Herman, and D. J. Munroe (2013): "Why are women underrepresented amongst patentees?" *Research Policy*, 42, 831–843.

Jaffe, A. B. and G. De Rassenfosse (2019): "Patent citation data in social science research: Overview and best practices," *Research handbook on the economics of intellectual property law*, 20–46.

Kim, S. and P. Moser (2020): "Women in science: Lessons from the Baby Boom," .

Koffi, M. and M. Marx (2021): "Cassatts in the Attic," .

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017): "Technological innovation, resource allocation, and growth," *The Quarterly Journal of Economics*, 132, 665–712.

Koning, R., S. Samila, and J.-P. Ferguson (2021): "Who do we invent for? Patents by women focus more on women's health, but few women get to invent," *Science*, 372, 1345–1348.

Martínez, G. L., J. Raffo, K. Saito, et al. (2016): *Identifying the gender of PCT inventors*, vol. 33, WIPO.

Murray, F. and L. Graham (2007): "Buying science and selling science: Gender differences in the market for commercial science," *Industrial and Corporate Change*, 16, 657–689.

Pidgeon, M. E. (1954): *Changes in women's occupations, 1940-1950*, 253-256, US Department of Labor, Employment Standards Administration, Women's Bureau.

Scherer, F. M. and D. Harhoff (2000): "Technology policy for a world of skew-distributed outcomes," *Research policy*, 29, 559–566.

Solow, R. M. (1957): "Technical change and the aggregate production function," *The Review of Economics and Statistics*, 312–320.

Thursby, J. G. and M. C. Thursby (2005): "Gender patterns of research and licensing activity of science and engineering faculty," *The Journal of Technology Transfer*, 30, 343–353.

Toole, A. A., M. J. Saksena, C. A. Degrazia, K. P. Black, F. Lissoni, E. Miguelez, G. Tarasconi, et al. (2021): "Progress and Potential: 2020 update on US women inventor-patentees," Tech. rep.

Waldfogel, J. (2023): "The Welfare Effect of Gender-Inclusive Intellectual Property Creation: Evidence from Books," Tech. rep., National Bureau of Economic Research.

Whittington, K. B. and L. Smith-Doerr (2005): "Gender and commercial science: Women's patenting in the life sciences," *The Journal of Technology Transfer*, 30, 355–370.

8 Figures and Tables

Figure 1: Quantitative GRE ≥ 700

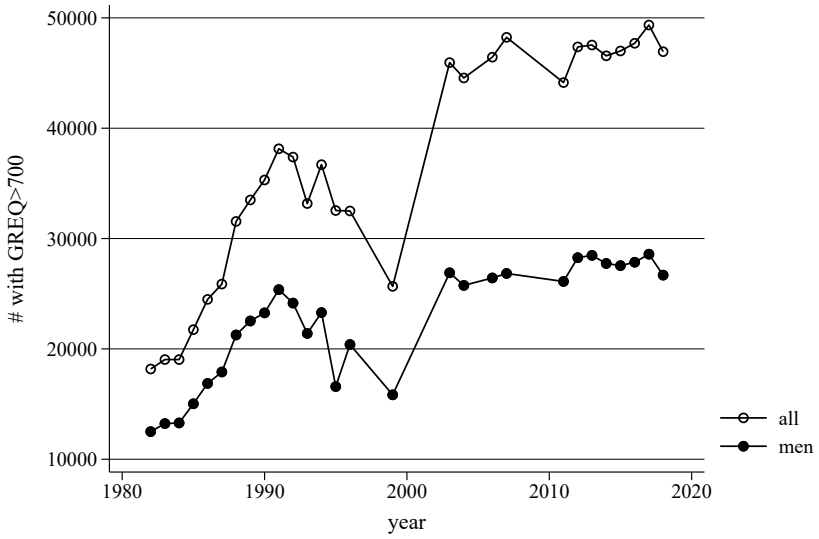


Figure 2: Female share of PhDs granted

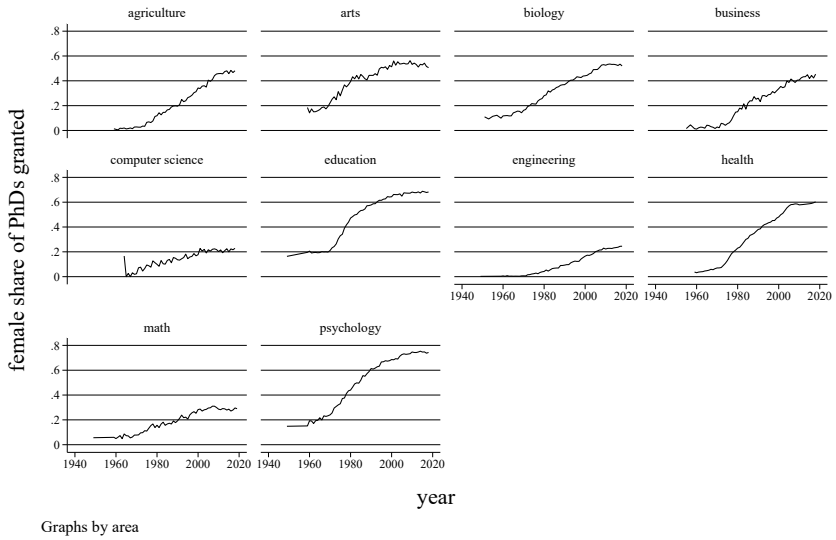
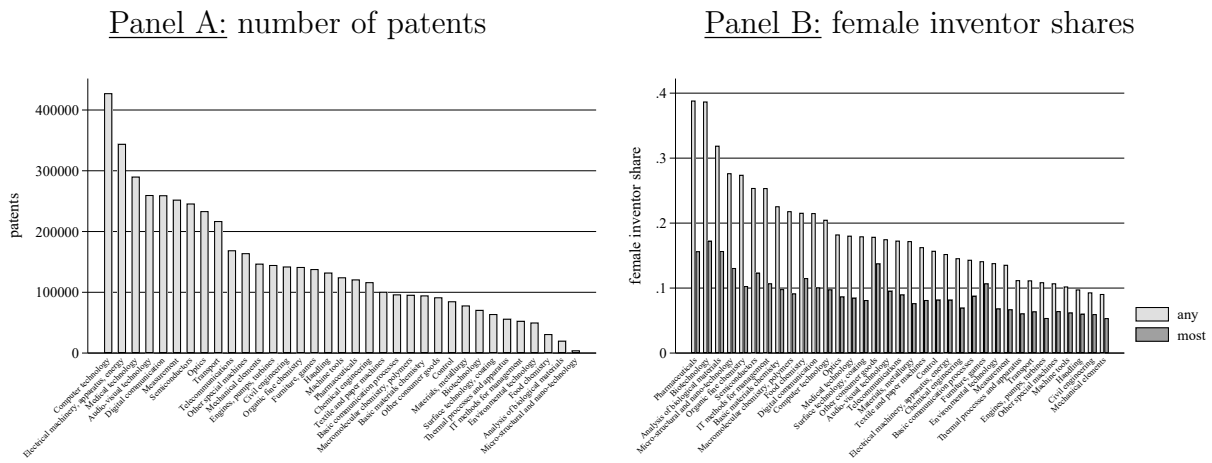


Figure 3: WIPO category information



Note:

Figure 4: Female-invented patent shares

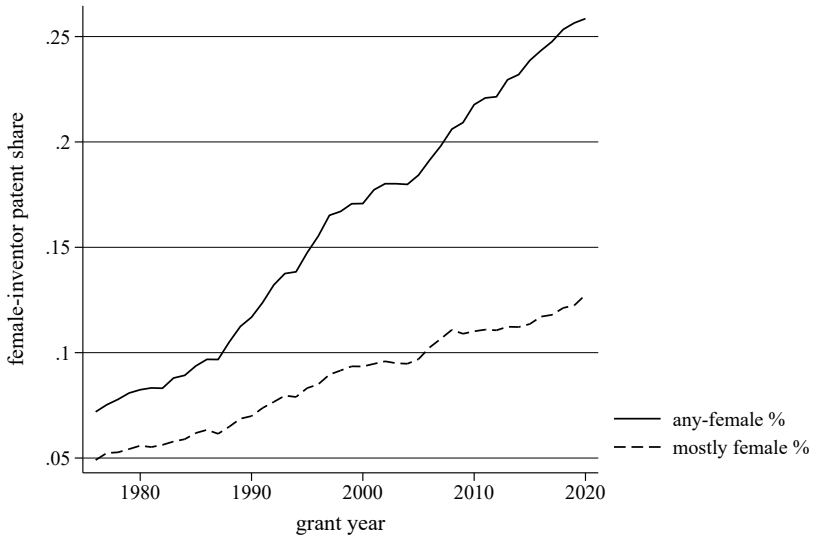
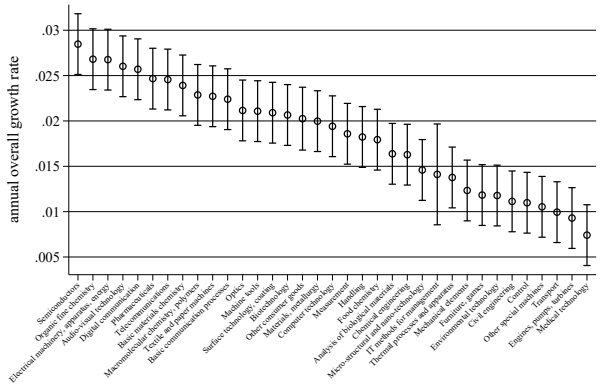
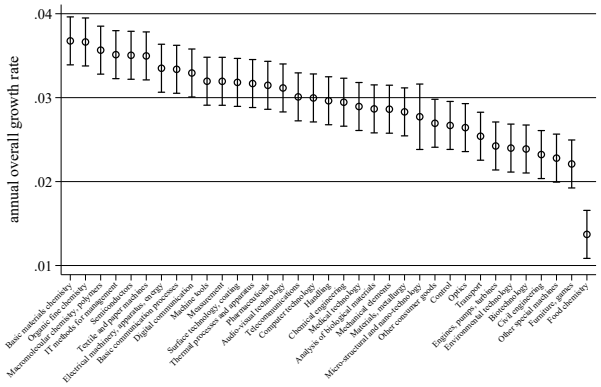


Figure 5: Female share growth rate by WIPO category

Panel A: mostly female



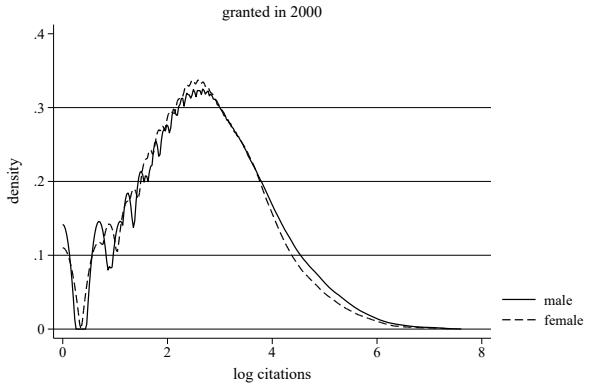
Panel B: any female



Note: Category-specific coefficients from regressions of log female inventor share, among patents from a vintage, on vintage.

Figure 6: Patent citation success distribution

Panel A: mostly female



Panel B: any female

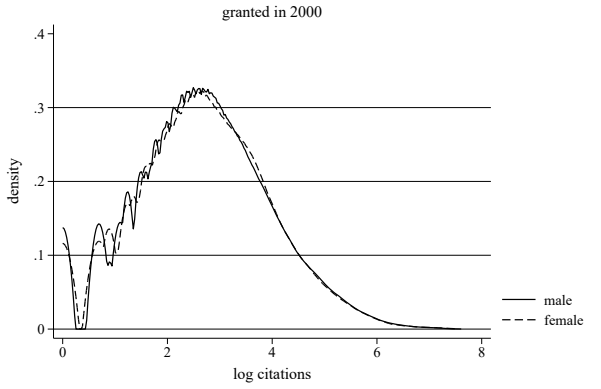
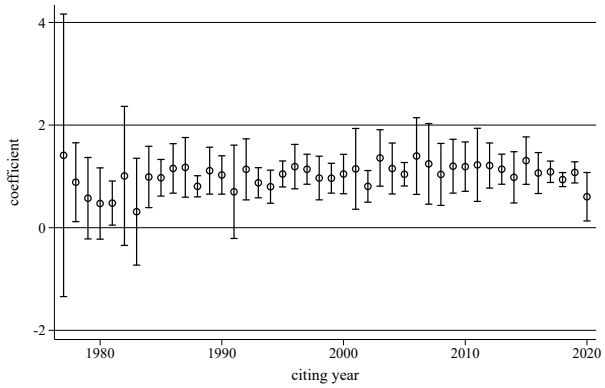


Figure 7: α over time

Panel A: mostly female



Panel B: any female

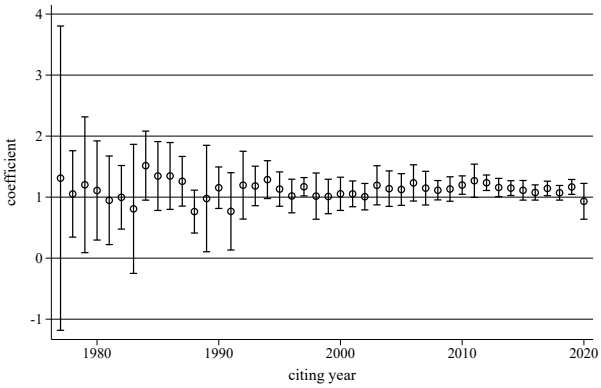
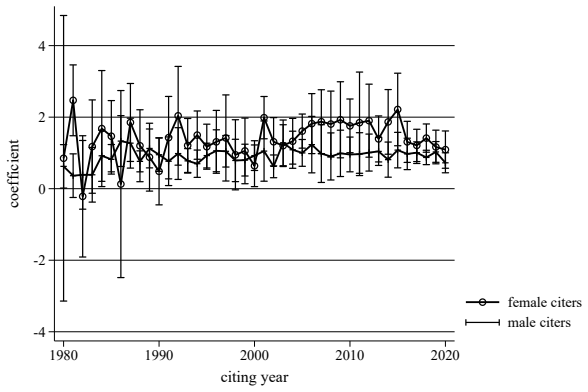


Figure 8: α over time by citing patent inventor gender

Panel A: mostly female



Panel B: any female

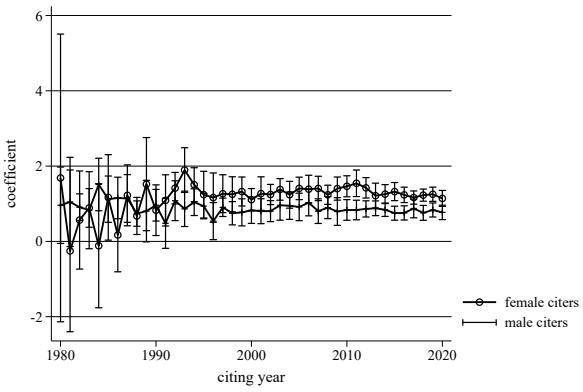


Figure 9: Influx effect by user/inventor tech category

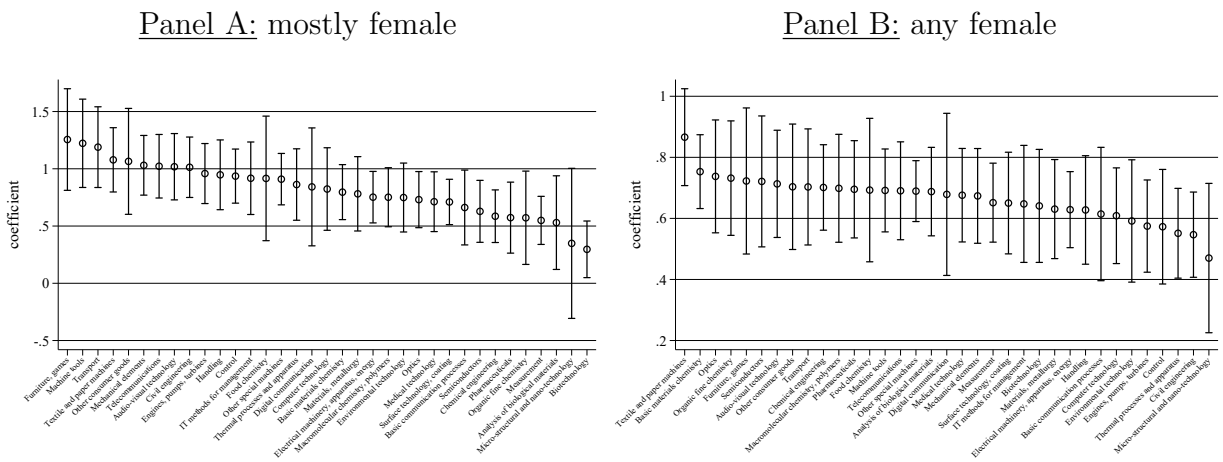


Figure 10: Female share of science award recipients

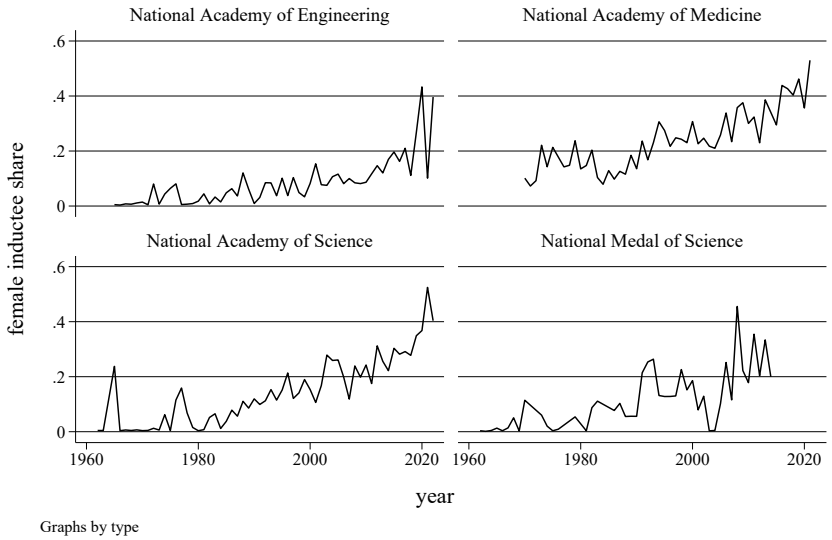


Table 1: Sample description

dataset	female-invented %	# patents per cell	citations per cell	% of citations to female-invented patents	# cells
v	15.98%	111,890.8	1,988,526.0	16.69%	45
vw	17.19%	3,215.2	57,141.1	17.67%	1,566
g_vw			28,588.8	19.40%	3,130
w_vw			1,932.5	18.50%	46,263
gw_vw			1,082.2	19.89%	82,520
t_vw			2,521.6	14.43%	35,483
gt_vw			1,280.0	15.88%	69,889
wt_vw			157.0	19.98%	555,528
gwt_vw			108.8	22.70%	791,347

Notes: Datasets are aggregations of the underlying patent and citation data by citing and cited patent characteristics. Cited patent characteristics are vintage v and technology category w. Citing patent characteristics are inventor gender g, technology category w, and year of citing patent t. Datasets are denoted by x_y, where x is citing patent characteristics and y is cited patent characteristics. The table uses the any-female definition of female-invented patents.

Table 2: Cross-gender citations

ANY		cited:		
		male	female	total
citing:	male	84.66%	15.34%	71,750,755
	female	77.96%	22.04%	20,450,175
	total	83.17%	16.83%	92,200,930

MOST		cited:		
		male	female	total
citing:	male	93.62%	6.38%	86,300,214
	female	88.78%	11.22%	5,900,716
	total	93.31%	6.69%	92,200,930

Notes: In the top panel a citation's patent has a female inventor if any of its inventors are women. In the bottom panel a citation's patent has a female inventor if half or more of its inventors are women. A row shows the percent of citations from, say, male-inventor citing patents that are to male and female-invented patents.

Table 3: Regression of female citation share on female patent share

	(1) any: v	(2) most: v	(3) any: v,w	(4) most: v,w
female-inventor share	1.059***	0.822***	0.988***	0.868***
	(0.0170)	(0.0334)	(0.0700)	(0.0797)
Observations	45	45	1722	1715
$\overline{R^2}$	0.989	0.932	0.852	0.682

Notes: Dependent variable is the share of citations garnered by female-inventor patents. Columns (1) and (2) aggregate to the patent vintage. Columns (3) and (4) aggregate to cited patent vintage (v) $\times$ technology category (w) and includes v and w fixed effects. Columns (1) and (3) use the any-female definition, and columns (2) and (4) use the mostly-female definition.

Table 4: Use of female patents by inventor gender

	(1) any: v	(2) most: v	(3) any: v,w	(4) most: v,w
0.df_ing $\times$ female share of patents	0.781***		0.616***	
	(0.0524)		(0.0915)	
1.df_ing $\times$ female share of patents	1.641***		1.284***	
	(0.0524)		(0.0755)	
dmostf_ing=0 $\times$ female share of patents		0.647***		0.574***
		(0.102)		(0.128)
dmostf_ing=1 $\times$ female share of patents		2.171***		1.818***
		(0.102)		(0.154)
Observations	90	90	3435	3417
$\overline{R^2}$	0.943	0.900	0.788	0.643

Notes: Dependent variable is the share of citations garnered by female-inventor patents. Column (1) aggregates to the patent vintage. Column (2) aggregates to cited patent vintage (v) $\times$ technology category (w) and includes v and w fixed effects. Column (3) aggregates to citing patent year t and cited patent vintage and technology category (w), or to $t \times v \times w$ and includes t, v, and w fixed effects. Column (4) uses the same data as column (3) but includes fixed effects for $t \times v$ and $t \times w$. These regressions use the mostly-female definition.

Table 5: Displacement of male patents

	(1) any	(2) any	(3) most	(4) most
female patents	2.360***	2.931***	6.237***	6.236***
	(0.146)	(0.0292)	(0.237)	(0.0604)
Observations	45	1722	45	1715
$\overline{R^2}$	0.856	0.952	0.940	0.952

Notes: Dependent variable is the number of patents created by male inventors, while the independent variable is the number created by female inventors. The columns labelled "most" define female patents as those with half or more female inventors. The columns labelled "any" define a patent as female if any of its inventors are female. The independent variable is the female share of patents granted. Columns (2) and (4) use data by grant year and WIPO technology category and include grant year and technology category fixed effects.

Table 6: Patent growth and market expansion

	(1) OLS	(2) FE	(3) OLS	(4) FE
cat w patents extant in t	1.310***	1.651***		
	(0.0309)	(0.0495)		
log cat w patents extant in t			1.239***	1.136***
			(0.0135)	(0.0365)
Observations	1567	1567	1567	1567
$\overline{R^2}$	0.535	0.698	0.844	0.940

Notes: Dependent variables are the year-t citations to patents in a technology category. The independent variable of interest is the number extant patents in the category. The data are citing year $\times$ technology category. The first two columns use levels, and the last two use logs. Columns labelled "FE" include technology category and citing year fixed effects.

Table 7: Female award recipients and female patenting

	(1) NAE	(2) NAM	(3) NAS	(4) NMS
female patent share	0.962***	1.461***	1.427***	1.250***
	(0.135)	(0.130)	(0.117)	(0.312)
Observations	45	44	45	34
$\overline{R^2}$	0.531	0.745	0.771	0.313

Notes: Dependent variables are the shares of inductees in a year, while the independent variable is share of patents with at least one female inventor granted during a year. The columns represent the National Academy of Engineering, the National Academy of Medicine, the National Academy of Science, and the National Medal of Science.

Table 8: Logit demand estimates

	(1) OLS	(2) 1st stg	(3) IV
lsin	0.866***		0.581***
	(0.0000449)		(0.000103)
(mean) df_ed	0.000879***		0.00495***
	(0.0000907)		(0.000165)
N		0.000000380***	
		(1.05e-08)	
N_sq		-2.24e-13***	
		(2.81e-15)	
lN		-1.209***	
		(0.00418)	
Observations	30023659	31130714	30023659
$\overline{R^2}$	0.952	0.380	0.844

Notes: Columns (1) reports the uninstrumented demand model. Columns (2) reports the IV estimate, and column (3) reports IV. N refers to the number of patents available to be cited during the observation's year.

Table 9: Logit demand estimates by inventor gender

	(1) OLS:male	(2) OLS:female	(3) 1st stg:male	(4) 1st stg:female	(5) IV:male	(6) IV:female
lsin0	0.885***				0.640***	
	(0.0000417)				(0.0000953)	
(mean) df_ed	0.00135***	-0.00892***			-0.000458***	0.0124***
	(0.0000823)	(0.000197)			(0.000147)	(0.000284)
lsin1		0.716***				0.432***
		(0.000142)				(0.000168)
N			-0.000000412***	0.00000233***		
			(1.05e-08)	(1.55e-08)		
N_sq			2.46e-14***	-7.60e-13***		
			(2.83e-15)	(4.14e-15)		
lN			-0.885***	-2.350***		
			(0.00420)	(0.00628)		
Observations	26050184	9726302	27009083	10057029	26050184	9726302
$\overline{R^2}$	0.965	0.868	0.379	0.567	0.887	0.721

Notes: Columns (1) and (2) report uninstrumented demand models for male and female inventors, respectively. Columns (3) and (4) report first-stage regressions, and columns (5) and (6) report the IV estimates. N refers to the number of patents available to be cited during the observation's year.

Table 10: Welfare estimates

	inventor surplus			citations ("revenue")			patents (entry)		
% change status quo rel to:	overall	female	male	overall	female	male	overall	male	female
status quo	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
no female influx	2.69	4.58	2.08	2.64	38.58	-3.78	8.68	0.00	68.81
no FI, endogenous male entry	1.17	2.78	0.67	1.15	41.62	-5.79	3.15	-5.79	68.81

Notes: The table shows the percent change in the columns outcome for the status quo relative to either a) environment with no female influx and no endogenous male entry response, or b) no female influx, *with* a male entry response. The estimates are based on the nested logit demand model and an entry model with imperfect predictability.