

Pudas, Atte

Working Paper

Why are there so few female inventors?

Helsinki GSE Discussion Papers, No. 51

Provided in Cooperation with:

Helsinki Graduate School of Economics

Suggested Citation: Pudas, Atte (2026) : Why are there so few female inventors?, Helsinki GSE Discussion Papers, No. 51, ISBN 978-952-7543-50-4, Helsinki Graduate School of Economics, Helsinki

This Version is available at:

<https://hdl.handle.net/10419/335199.2>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

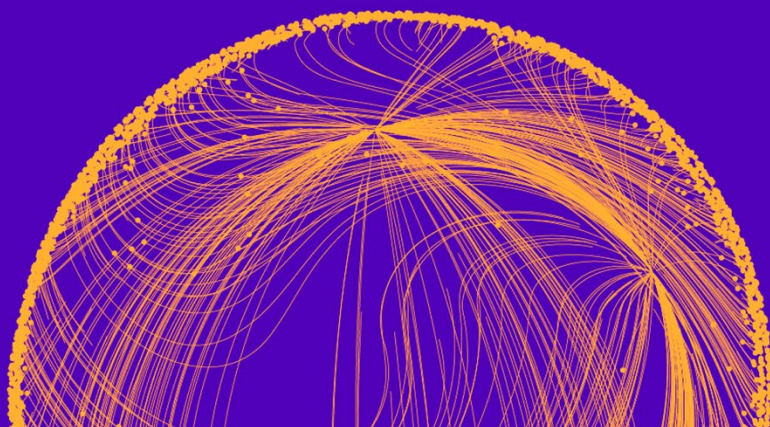
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

HELSINKI GSE DISCUSSION PAPERS 51 · 2025

Why Are There So Few Female Inventors?

Atte Pudas



HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI



A! Aalto University

Helsinki GSE Discussion Papers

Helsinki GSE Discussion Papers 51 · 2025

Atte Pudas

Why Are There So Few Female Inventors?

ISBN 978-952-7543-50-4 (PDF)

ISSN 2954-1492

Helsinki GSE Discussion Papers:

<https://www.helsinkigse.fi/discussion-papers>

Helsinki Graduate School of Economics

PO BOX 21210

FI-00076 AALTO

FINLAND

Helsinki, Revised August 2026

Why Are There So Few Female Inventors?*

Atte Pudas[†]

August 21, 2026

Abstract

Only a small fraction of inventors are women. Using Finnish administrative data linked to patent records, I show that educational and occupational sorting explain roughly half of the gender gap in patenting. The remaining half arises because women patent less than men in nearly all occupations, even after controlling for education, high school grades, and employing firm. Career effects of parenthood are a key mechanism: using a matched-control event-study design, I find that mothers' annual patenting rates decline by 78% after their first childbirth, with partial long-term recovery, whereas fathers' patenting rates increase persistently by roughly 10%.

Keywords: gender gap, decomposition, innovation, inventors, occupations, parenthood, patenting

JEL Codes: O31, J16, J13, J24

*I am particularly thankful for Otto Toivanen, Kristiina Huttunen, and Jarkko Harju for their invaluable comments and suggestions. I also thank Olof Ejermo, Ari Hyytinen, H.C. Kongsted, Myriam Mariani, Karin Hoisl, Ioana Igna, Valentina Tartari, Peter Holmström, Tuomas Matikka, and Toni Juuti for very helpful discussions and comments. I also thank the participants of workshops and conferences at the University of Bristol, WIPO, Helsinki GSE, the Annual Meeting of the Finnish Economic Association, and the Nordic PhD Workshop in the Economics of Education for their valuable comments and insights. Finally, I thank Eero Nurmi for his camaraderie in collecting and linking the inventor data. All errors are mine. I am grateful to Yrjö Jahnsson Foundation for funding (grants no. 20217375, 20227559, 20237688, 20247829, and 20257971). This research was also funded by FORTE grant no. 2021-01552.3, Women in science and technology: career impediments and the child penalty (WINST).

[†]Aalto University & Helsinki GSE; Email: atte.pudas@aalto.fi; Personal website: www.attepudas.com

1 Introduction

Over recent decades, gender gaps have narrowed across several aspects, including employment, hours worked, and earnings (Goldin, 2006; Olivetti and Petrongolo, 2016). Women’s representation in STEM (science, technology, engineering, and mathematics) has also increased. Nonetheless, sizable gaps remain. In 2017, women accounted for 34% of tertiary graduates in STEM fields globally (Schmuck, 2017), and by 2024, the share of women in STEM occupations stood at only 28% (Kali Pal et al., 2024). In innovation, the gender divide is even more pronounced: the share of female patent inventors rose from less than 5% in 1980 to only 13% in Europe and the U.S. by 2019 (Toole et al., 2020; EPO, 2022). Thus, a clear disparity persists between the number of women with competencies to innovate and women’s representation among inventors, suggesting crucial differences in the career trajectories of potential male and female innovators.

This paper makes the following key contributions to understanding these differences. First, I decompose the gender gap in patenting into two parts: the part attributable to educational and occupational sorting, and the unexplained part, reflecting differences between equally qualified women and men working in the same finely defined occupations. While the sorting channel has been extensively studied, the more puzzling differences within occupations remain poorly understood. This paper provides the first comprehensive analysis of these differences across a complete set of occupations. Second, I shed light on the underlying mechanisms by presenting new population-level evidence on the differential long-term effects of parenthood on women’s and men’s patenting activity and careers in patent-intensive occupations, using a matched-control event-study design. My analysis draws on detailed Finnish administrative data linked to European Patent Office patent application records.

Why women remain underrepresented among inventors is an important economic question for two main reasons. First, some women with high inventive potential may never become inventors, resulting in the so-called “lost Marie Curies” (Bell et al., 2019; Hoisl et al., 2022). A higher number of female inventors could, therefore, increase not only the quantity but also

the quality of inventions, spurring growth through an improved allocation of talent (Hsieh et al., 2019; Brouillette, forthcoming). Second, the gender imbalance in innovation shapes what is invented and who benefits from it, as female innovators are more likely to address needs specific to women (Koning et al., 2021; Einiö et al., Forthcoming).

Moreover, the gender gap in innovation is closely tied to broader gender differences in the labor market, such as gender segregation in education and occupations, particularly in STEM (e.g., Kahn and Ginther, 2017; Speer, 2023), and women’s higher exit rates from science and engineering careers (Preston, 1994; Hunt, 2016). Inventors also earn substantial returns from patenting (Toivanen and Väänänen, 2012; Aghion et al., 2018; Kline et al., 2019), linking women’s underrepresentation in innovation to gender differences in earnings and earnings growth (e.g., Bertrand et al., 2010; Goldin, 2014; Blau and Kahn, 2017; Goldin et al., 2017; Hospido et al., 2022).

My findings are as follows. Women’s annual probability of filing a patent application is 88% lower than men’s. Education and occupation fixed effects explain roughly 50% of this gap, reflecting the fact that men are more likely to sort into patent-intensive fields of education and occupations, such as many engineering professions.¹ The underexplored part of the gap — the remaining 50% — stems from women patenting less than men within the education–occupation cells.² I perform a threefold Oaxaca–Blinder decomposition to highlight the complementary nature of the sorting and within-cell disparities. The interaction term in the decomposition accounts for 58% of the total gender gap, indicating that the gap is strongly driven by men being disproportionately represented in fields and occupations where they also patent at higher rates than women. These findings carry an important policy implication: although encouraging female participation in STEM and patent-intensive occupations is essential for expanding the pool of potential inventors and fostering innovation,

¹Consistent with previous studies linking STEM participation to inventive activity (e.g., Hunt et al., 2013; Toivanen and Väänänen, 2016; Hoisl et al., 2022).

²I distinguish track-specific education and occupations at a granular level by classifying fields of study at the 3-digit level (ISCED-2011), education levels at the 2-digit level (e.g., separating academic and vocational secondary education), and occupations at the most detailed 4- or 5-digit level.

such policies alone are insufficient, as substantial within-occupation gender disparities persist.

To illustrate the pervasiveness of these disparities, I estimate the gender patenting gaps separately for each narrowly defined occupation. About 55% of occupations with any patenting activity have had only male inventors, and just 5% have had only female inventors. Among the remaining 40% of occupations in which both women and men have patented, women patent at lower rates in 84% of cases, and the largest gaps appear in the most patent-intensive occupations. The estimated within-occupation gender gaps remain largely unexplained after controlling for age, education, high school grades, and firm fixed effects, indicating that differences in potential experience, competencies, general ability, and employers' innovation capacity cannot fully account for them. For example, among research and development managers — the most inventor-intensive occupation — women are roughly 2.8 percentage points (76%) less likely to patent than men, whose annual patenting rate is 3.7%. After accounting for the covariates listed above, 1.5 percentage points (54%) of the gap remains unexplained.

Multiple mechanisms may underlie these gender differences. Career disruptions due to childrearing may be particularly detrimental to inventive productivity, given its reliance on up-to-date expertise and collaboration. I study the effects of parenthood on patenting activity using an event-study design that compares the patenting trajectories of parents with those of closely matched childless individuals. I find that women's annual patenting rates decline by roughly 78% after their first childbirth, followed by a slow recovery. However, the recovery appears to remain imperfect: the average gap is still around -20% more than 15 years after the first childbirth, although the estimate is not statistically significant. Cumulatively, over the 15 years following the first childbirth, mothers lose roughly six years' worth of patenting relative to their pre-childbirth annual patenting rate. The decline in patenting is associated with mothers shifting into less patent-intensive occupations. However, the observed occupational shifts are not able to explain the decline in mothers' patenting, suggesting that parenthood affects mothers' inventive productivity primarily through adjustments within

occupations. These adjustments may include, for instance, reduced discretionary effort devoted to innovation or shifts towards less inventive tasks and roles that are easier to combine with family life.

In contrast, men’s annual patenting rates rise persistently by around 10% on average after becoming fathers, resulting in a 15-year cumulative increase of 1.6 times their pre-parenthood annual rate. Part of this fatherhood premium appears to reflect fathers’ higher long-term employment rates relative to the matched childless men. Importantly, I also find substantial gender patenting gaps between future parents already before they have children, as well as between childless women and men. Even after controlling for age, education, and occupation, the gender patenting gap among individuals without children still amounts to almost 50% of the gap among parents. Taken together, my results indicate that the unequal career consequences of parenthood are a key driver of the gender gap in patenting, but cannot fully explain it.

Finally, while this paper focuses on the extensive margin of patenting, I also document a gender gap in patent grant rates of about 4.5 percentage points (7%), conditional on filing, consistent with prior U.S.-based evidence on gendered frictions and biases in the patenting process ([Jensen et al., 2018](#); [Aneja et al., 2024](#); [Gavrilova-Zoutman and Juranek, 2025](#); [Pairolero et al., 2025](#)). Differences in grant rates may contribute to the broader disparities in innovation: if potential female inventors anticipate a lower likelihood of success, they may be discouraged from innovating in the first place. Grant rate differences may also reflect variation in application quality. However, I examine a range of patent quality indicators and find mixed evidence of gender differences: female inventors’ patents receive, on average, fewer citations than those by male inventors, whereas measures such as patent family size, technological scope, originality, and radicalness indicate that they are highly comparable in quality. Furthermore, interpreting the observed differences in citations is complicated by the possibility that patents by female inventors may be undercited relative to their quality and relevance ([Hochberg et al., Forthcoming](#); [Li and Liu, Forthcoming](#)).

Prior work and contributions. The first main contribution of this paper is to provide the first comprehensive analysis of the within-occupation gender differences in patenting and their role in perpetuating the underrepresentation of women in innovation. Prior literature emphasizes the gendered effects of early-life exposure to innovation and role models, which disproportionately favor boys over girls (Bell et al., 2019; Hoisl et al., 2022; Aghion et al., 2023; Breda et al., 2023), as well as women’s underrepresentation in STEM education (Jung and Ejermo, 2014; Toivanen and Väänänen, 2016; Aghion et al., 2017; Bianchi and Giorcelli, 2020) as key drivers of the gender gap in innovation. Later stages of potential inventors’ careers, beyond education, have received less attention. A notable exception is Hunt et al. (2013), who document a strong link between the gender patenting gap and the underrepresentation of STEM-educated women in patent-intensive job tasks. A handful of papers also document gender disparities in patenting within a single, specific occupational group, namely, academic scientists (Thursby and Thursby, 2005; Whittington and Smith-Doerr, 2005; Ding et al., 2006; Whittington and Smith-Doerr, 2008; Whittington, 2011; Ceci et al., 2014). Understanding the points in potential inventors’ careers at which the gender gap in innovation emerges is essential for effectively designing and targeting policies to strengthen women’s participation in innovation.

My second main contribution is to provide new population-level evidence on the long-term effects of parenthood on inventing activity. Prior evidence on parenthood’s impact on patenting is scarce, with notable exceptions including Whittington (2011) and Kim and Moser (2021), who provide suggestive evidence that motherhood is associated with declines in inventive productivity. In addition, in a contemporaneous independent study, Mueller et al. (2026) examine the more short-term effects of parenthood on patenting using German data and find adverse effects on women’s inventive productivity that are consistent with my findings, albeit smaller in magnitude. I contribute to the extensive literature on the career costs of children (e.g., Angelov et al., 2016; Lundborg et al., 2017; Kleven et al., 2019a,b; Kim and Moser, 2025) by showing that the adverse effects of parenthood experienced by mothers

extend to inventive productivity, beyond what can be explained by changes in labor market attachment and occupational mobility. My findings on the increases in fathers’ patenting rates also complement the often-overlooked evidence of “fatherhood premia” in labor market outcomes (Lundberg and Rose, 2000; Correll et al., 2007; Killewald, 2013; Goldin et al., 2024).

My findings highlight the key role of parenthood in driving the gender gap in patenting and contribute to the growing body of evidence on the mechanisms underlying gender disparities in innovation, including gender differences in monetary incentives (Toivanen and Väänänen, 2012; Hoisl and Mariani, 2017), gendered within-firm frictions and barriers in the innovation processes (Chien and Grennan, 2024), gender differences in risk and competition aversion (e.g., Gneezy et al., 2003; Niederle and Vesterlund, 2007; Booth and Nolen, 2012), gender biases and discrimination in crediting inventors (Ross et al., 2022), women’s lower patent grant rates (Jensen et al., 2018; Aneja et al., 2024; Gavrilova-Zoutman and Juranek, 2025; Pairolero et al., 2025), and female inventors’ lower recognition in terms of patent citations (Hochberg et al., Forthcoming; Li and Liu, Forthcoming). My results on mothers’ shifts into less-inventive occupations also contribute to the limited evidence on how parenthood differentially shapes occupational choices by gender (Tannenbaum and Timpe, 2025; Liu, 2025).

2 Data and Institutional Context

I use individual-year level administrative data on Finland’s permanent resident population from the FOLK basic, FOLK degree, and FOLK employment modules, covering the period 1987–2019 (Statistics Finland, 2026a,c,d). The data, provided by Statistics Finland, include detailed information on variables such as education level and field, occupation, and employer firm, and all datasets can be merged using encrypted individual identifiers. Occupation codes are available annually from 2004 through 2018; accordingly, the decomposition analysis in Section 5 is based on data from this period. Other relevant variables are available for all years. Variable values are measured at the end of each year. I also use data on high

school matriculation examination results for the years 1967–2017 ([Statistics Finland, 2026e](#)), allowing me to observe which final exams individuals took in high school and the grades they received.³ In addition, I use data that link children born in 1970 or later to their parents ([Statistics Finland, 2026b](#)), allowing me to track individuals’ parental status and the timing of their children’s births. For the analysis of the impacts of parenthood in Section 6, I will utilize the full panel length 1987–2019.

Throughout the paper, I categorize occupations at the most detailed four-digit or five-digit level. In addition, I distinguish fields of study at the 3-digit level (compatible with ISCED-2011) and educational levels at the 2-digit level (e.g., differentiating between high school and vocational secondary education). When I use the term ‘college’, I refer to all tertiary-level degrees, including doctoral degrees. In this paper, STEM refers to the following broad fields of study: (i) natural sciences, mathematics, and statistics; (ii) information and communication technologies; (iii) engineering, manufacturing, and construction; and (iv) three subfields of health: medicine, pharmacy, medical diagnostics, and treatment technology.

Furthermore, I use patent application data from the European Patent Office (EPO), obtained from PATSTAT, to identify Finnish patent inventors and track their patenting activity. The data include all inventors listed on EPO patent applications filed between 1978 and 2019, with a corresponding address in Finland. The inventors are linked to the register data using information on their names, addresses, and their employers’ names and addresses (see Supplemental Appendix B for details). We successfully link an individual in the register data to about 90% of inventor–patent observations, with no evidence of systematic gender differences in linkage success. I define an inventor as an individual who is named on at least one patent application, regardless of whether the application is ultimately granted.⁴

³See Supplemental Appendix A for details of the school system in Finland.

⁴Since I define inventorship based on all patent applications, both rejected and granted, the main results are not affected by the lower patent grant rates of female inventors (as documented by [Jensen et al., 2018](#); [Aneja et al., 2024](#); [Gavrilova-Zoutman and Juranek, 2025](#); [Pairolero et al., 2025](#)). I present evidence of gender differences in patent grant rates among Finnish inventors in Section 5.

The decision to use Finnish data stems from Finland being one of the few countries where patent data can be linked to individual-level register data. However, Finland also provides an intriguing context for studying gender differences in patenting, as the institutional environment itself does not inherently support such disparities. As a Nordic welfare state with globally high levels of gender equality (e.g., [Hausmann et al., 2009](#)) and an education system that is fundamentally free of charge up to and including the doctoral level, the barriers to education and career opportunities are low. In addition, discrimination based on gender is prohibited by law. However, notable gender disparities still persist, including strong segregation in both educational fields and occupations, as well as a clear gender pay gap ([Statistics Finland, 2021](#)). In addition, Finnish women’s careers are disproportionately affected by parenthood compared to men’s, with long-term “child penalties” broadly similar to those in other Nordic countries ([Kleven et al., 2019a](#); [Sieppi and Pehkonen, 2019](#)).

3 Inventiveness of Occupations

I refer to my baseline measure of occupational inventiveness as ‘inventor intensity’, defined as the average annual number of patenting workers in a given occupation divided by the average total number of workers in that occupation. Thus, inventor intensity captures the average annual rate at which individuals in a particular occupation are listed as inventors on patent applications. I distinguish occupations using the most specific occupation code level available (See Supplemental Appendix [C](#) for the details).

In Section [5](#), I analyze the gender patenting gaps across all occupations involved in patenting. Before that, it is useful to examine the characteristics of the most patenting-active occupations. Table [1](#) displays the occupations that fall within the top 5% of the inventor intensity distribution.⁵ These occupations are predominantly related to engineering

⁵I exclude three small occupations with fewer than four unique inventors during 2004–2018 from Table [1](#). In these cases, measured inventor intensity arguably reflects the patenting activity of a very small number of specific individuals rather than occupational inventiveness. For completeness, however, I impose no such restrictions in the analysis in Section [5](#).

and information technologies, as well as chemistry, pharmacy, and medicine, underscoring the strong link between STEM fields and patenting. The table displays both the overall and gender-specific inventor intensities for each occupation. In addition, the two rightmost columns present the absolute and relative gender gaps in inventor intensity, showing that women are less likely to patent in 90% of the most inventor-intensive occupations. On average, for every thousand workers, there are 7.55 fewer female inventors than male inventors, which translates to women patenting 53% less frequently in these occupations. Over 80% of these occupations are male-dominated (see Supplemental Appendix Table [A.1](#) for workforce composition), but notable exceptions exist, such as “chemists” and “pharmacologists, pathologists, and related professionals”, where women are the majority.

Table 1 here

It is important to note that, due to the low number of female inventors, the inventor intensity of most occupations is dominated by the patenting activity of men. Women also tend to patent more in certain technology sectors, such as chemistry, compared to fields like mechanical engineering ([EPO, 2022](#)). Nevertheless, it seems that the most inventor-intensive occupations are relatively similar for women and men. Supplemental Appendix Table [A.2](#) compares the pooled ranking of occupations by their inventor intensity and the placement of these occupations in the corresponding rankings of women and men. For instance, the twenty occupations with the highest female patenting activity all fall within the top 5% of occupations ranked by overall inventor intensity.

4 Gender Inventor Gap and Supply of Potential Inventors

Figure [1](#) characterizes the gender differences in the likelihood of becoming an inventor and the supply of potential inventors. Panel A shows the probability of becoming an inventor for the full birth cohorts of 1962–1982, as well as for three subpopulations of STEM college graduates: (i) all STEM graduates, (ii) those employed at least once in a top 5% inventor-

intensive occupation, and (iii) those who pursued a career in these occupations for at least five consecutive years. Panel B displays the share of women in these samples.

Figure 1 here

Going from left to right in Figure 1 and increasing the characteristics associated with patenting, the following key patterns emerge: (i) the probability of becoming an inventor increases substantially for both women and men; (ii) the relative increases are larger for women, and the female-to-male inventorship likelihood ratio increases from 0.15 to 0.66; (iii) the absolute gender gap, however, increases from about 12 to 48 inventors per thousand individuals, and importantly; (iv) the share of women decreases from 49% to just 16%.⁶

These findings highlight the two main factors driving the shortage of female inventors. First, the low representation of women among potential inventors with a high probability of patenting, specifically, STEM-educated individuals employed in inventor-intensive occupations. Second, even among these potential inventors, women are only about two-thirds as likely as men to become inventors. More generally, the first factor reflects gender segregation in education and occupations, while the second illustrates differences between women and men with similar education and occupations. In the next section, I decompose the gender gap in patenting into the contributions of these two factors.

5 Decomposing the Gender Gap in Patenting

As the baseline decomposition approach, I estimate a simple gender gap equation using panel data for the prime working age population aged 25–65 over the 2004–2018 period:⁷

⁶Being an inventor is rare: in total, there are 1,617 female inventors and 11,046 male inventors in the 1962–1982 birth cohorts (see Supplemental Appendix Table A.3).

⁷In the main estimation sample, I exclude observations with missing occupation or firm information, which may arise from unemployment, full-time studying, staying home with children, retirement, or other non-employment reasons. As a robustness check, I re-estimate the model, including these observations and assigning them to “missing” categories. These observations comprise roughly 31% of the expanded sample. The results are reported in Supplemental Appendix Table A.4. Annual patenting rates decline to 0.013% for women and 0.110% for men, reflecting the rarity of patenting outside employment. Relaxing the sample does not affect the main conclusions.

$$y_{it} = \alpha + \gamma Woman_i + \lambda_t + \mathbf{X}'_{it}\beta + \varepsilon_{it}, \quad (1)$$

where for individual i in calendar year t , the outcome variable y_{it} is defined as either (i) an indicator for filing a patent application in a given year (including both granted and rejected applications), or (ii) an indicator for filing a patent application that is later granted.⁸ The former indicator captures individual-level inventive activity, while the latter additionally reflects potential gender disparities in the patent examination process and variation in patent application quality. Calendar year fixed effects are denoted by λ_t . Depending on the specification, $\mathbf{X}'_{it}$ include baseline controls — indicator variables for age, native language (Finnish, Swedish, or other), and employment status at the end of the year (employed versus other) — as well as fixed effects for education level by field, occupation, employer firm, and high school matriculation exam grades in mathematics and the mother tongue, which serve as proxies for general ability.⁹ The constant term is denoted by α , and ε_{it} is the error term. I estimate the model using ordinary least squares (OLS), with standard errors clustered at the individual level.

The gender gap estimate γ captures the portion of the raw gender difference in annual patenting probabilities that is not explained by differences in covariates, such as the allocation of women and men across fields of study and occupations. Note that the coefficient γ is equivalent to the unexplained component in a corresponding “pooled” twofold Oaxaca–Blinder decomposition, where the coefficients from a regression pooling women and men, as in Equation (1), serve as the reference coefficient structure.¹⁰

⁸I assign the patent grant indicator to the filing year, rather than the grant year, to capture the inventor’s occupation at the time the invention was finalized. At the EPO, the typical lag between filing and grant is three to five years (EPO, 2025a). I observe the grant status as of 2023. This allows me to capture grant outcomes for most applications filed through 2018, the final year of the estimation sample.

⁹I control for indicator variables capturing matriculation examination grades in mathematics (basic and advanced levels separately) and in mother tongue and literature (a compulsory exam for all high school students). I include categories for individuals who did not take the mathematics exam or did not complete high school to retain them in the sample. While high school grades are imperfect proxies for ability, achieving a high grade in advanced mathematics is positively associated with the likelihood of patenting, even after controlling for occupation and education (estimates available upon request).

¹⁰The Oaxaca–Blinder decomposition dates back to the seminal contributions of [Blinder \(1973\)](#) and

Table 2, Panel A, reports the estimated gender gap, $\hat{\gamma}$, in the annual probability of filing a patent application. I scale the coefficients by 100 for readability. The gender-gap coefficient in Column 1 indicates that women’s patenting rate is 0.133 percentage points (pp.) lower than men’s. The small magnitude reflects the rarity of patenting: the constant term implies an average annual filing probability of 0.151% for men. Relative to this baseline, women’s annual patenting probability (0.018%) is 88% lower. Adding the baseline controls and fixed effects for education and occupation to the specification absorbs roughly 50% of the raw gender gap, as $\hat{\gamma}$ declines to -0.0671 (Col. 2). The education and occupation fixed effects absorb the portion of the gender patenting gap that arises from men being disproportionately concentrated in fields of education and occupations with higher patenting activity, such as many engineering fields and professions. The remaining unexplained gender gap reflects differences between men and women within the same education-occupation cells. The pooled twofold Oaxaca–Blinder decomposition corresponding to Column 2 in Table 2, reported in Supplemental Appendix Table A.5, indicates that occupational sorting accounts for 82% of the explained part of the gender gap, while differences in education contribute roughly 15%.

Table 2 here

Including firm fixed effects in the regression has only a modest impact on the gender gap estimate once education and occupation are controlled for.¹¹ Firm fixed effects account for the possibility that equally qualified women and men may sort into firms with different levels

Oaxaca (1973). The pooled twofold Oaxaca–Blinder decomposition of a mean gender gap R (e.g., gender gap in patenting probability) is given by $R = (\bar{X}_M - \bar{X}_W)' \hat{\beta}^* + \bar{X}_M' (\hat{\beta}_M - \hat{\beta}^*) + \bar{X}_W' (\hat{\beta}^* - \hat{\beta}_W)$, where $\bar{X}_g$ are mean covariate values for gender $g \in \{W, M\}$, $\hat{\beta}_g$ are coefficients for the covariates from gender-specific regressions of the outcome (e.g., patenting dummy) on the covariates, and $\hat{\beta}^*$ are coefficients from a pooled regression combining women and men and including a gender dummy. The first term of the decomposition captures the part of the gap explained by differences in covariate means (e.g., differences in the allocation of women and men across occupations). The remaining two terms capture the unexplained part, driven by differences in coefficients (e.g., different patenting rates for women and men within occupations). Supplemental Appendix Section D.2 provides the derivations and shows that the unexplained component is captured by the coefficient on the gender dummy in the pooled regression, such as γ in Equation (1). Table A.5 reports the results of the pooled twofold Oaxaca–Blinder decomposition corresponding to Table 2, Column (2).

¹¹Between 2004 and 2018, at least one inventor filed a patent in 3,353 firms, representing 0.6% of all observed firm IDs. On average, 458 firms per year had at least one inventor patenting (min 398, max 502), while an average of 181 firms per year had at least two inventors patenting (min 164, max 194).

of innovation capacity and activity.¹² The estimated gender gap becomes slightly larger when firm effects are included (-0.0734 in Col. 3 versus -0.0671 in Col. 2), consistent with women being more, rather than less, concentrated in patent-active firms. Overall, the results indicate that within-firm differences are more salient than sorting to firms in explaining the gender patenting gap within education–occupation cells. Furthermore, including indicator variables for high school grades does not notably alter the gender gap estimate once education and occupation are controlled for (-0.0646 in Col. 4 versus -0.0671 in Col. 2). This suggests that the gender patenting gap within education-occupation cells is not explained by differences in general ability among those sorted into these groups.

I complement the main regression results presented in Table 2, Column 2, with a corresponding threefold Oaxaca–Blinder decomposition to obtain a more nuanced understanding of the joint contributions of between- and within-occupation gender differences. I present the decomposition using the women’s coefficient structure as the reference point. Then, the decomposition of the raw gender patenting gap R is given by (Supplemental Appendix D.3 presents the derivations):

$$\begin{aligned} R &= \mathbb{E}(Y_M) - \mathbb{E}(Y_W) \\ &= (\mathbb{E}(X_M) - \mathbb{E}(X_W))' \beta_W + \mathbb{E}(X_W)' (\beta_M - \beta_W) + (\mathbb{E}(X_M) - \mathbb{E}(X_W))' (\beta_M - \beta_W), \end{aligned} \quad (2)$$

where Y_g is the outcome variable for gender $g \in \{M, W\}$. Denoting covariate sample averages with $\bar{X}_g$ and the estimated gender-specific coefficient vectors on the covariates with $\hat{\beta}_g$, the decomposition can be estimated as:

$$\begin{aligned} R &= \underbrace{(\bar{X}_M - \bar{X}_W)' \hat{\beta}_W}_{\equiv Q \text{ ("endowments")}} + \underbrace{\bar{X}_W' (\hat{\beta}_M - \hat{\beta}_W)}_{\equiv C \text{ ("coefficients")}} + \underbrace{(\bar{X}_M - \bar{X}_W)' (\hat{\beta}_M - \hat{\beta}_W)}_{\equiv I \text{ ("interaction")}}. \end{aligned} \quad (3)$$

The raw gap is decomposed into three components: endowments (Q), coefficients (C),

¹²A sometimes overlooked fact is that company employees produce the majority of patents. Between 2004 and 2018, 91% of Finnish inventors listed on EPO patent applications were company employees in the filing year, while only 5% were entrepreneurs.

and interaction (I). Using women as the reference group frames the decomposition from their perspective. Q measures the portion of the gender gap explained by differences in observed characteristics, evaluated at women’s coefficient values — that is, the expected change in women’s mean outcome if women had men’s characteristics (e.g., the same allocation of occupations), holding women’s coefficients fixed. C captures the part attributable to differences in coefficients — the expected change in women’s mean outcome if women had men’s coefficients (e.g., similar patenting rates within occupations), holding women’s characteristics fixed. I represents the interaction term, accounting for the simultaneous presence of differences in both characteristics and “returns” to these characteristics.

The results of the decomposition are reported in Table 3.¹³ Looking at Column 1.1, the raw gender gap (R) in patent-filing probability (reported as men – women) is again 0.133 pp. Differences in covariates (Q) account for 0.033 pp. (25%) of this gap, with occupation fixed effects contributing the vast majority, 0.0305 pp. (91.5% of Q). Differences in coefficients (C) account for 0.022 pp. (16.7%) of the raw gap.¹⁴ The interaction term (I) is the largest component, accounting for 0.078 pp. (58.3%) of the raw gap, with occupation fixed effects contributing the most, 0.060 pp. (77.3% of I). Thus, the majority of the gap can be attributed to men receiving higher “returns” on the characteristics they are also more likely to possess. For example, men are both overrepresented in engineering occupations and simultaneously exhibit substantially higher patenting rates within these occupations.

Table 3 here

Differences in grant rates and patent quality. Differences in patenting success may contribute to the gender gap in innovation participation if potential female inventors anticipate a lower likelihood of successfully obtaining a patent. Table 2, Panel B, reports the

¹³The underlying gender-specific regressions in the decomposition use the same sample and follow the specification in Table 2, Column 2, explaining the patent-filing dummy with baseline controls, education-level-by-field fixed effects, and occupation fixed effects.

¹⁴The decomposition of the “coefficients” part into contributions from different coefficient groups is meaningfully interpretable only for variables with a natural zero point (Jann, 2008), and is therefore omitted.

gender gap estimates for the probability of filing a patent application that is later granted. At the EPO, about 60% of applications are granted on average (EPO, 2025b). Consistently, the raw annual probability of filing at least one patent application that is eventually granted is 0.093% for men (constant in Col. 1), which is roughly 61.5% of their raw filing probability reported in Panel A. For women, the results indicate a raw probability of filing a later-granted patent of 0.0103% ($0.0930 - 0.0827$), which is roughly 57% of their raw filing probability obtained from Panel A. Thus, the results indicate a roughly 4.5 pp. (7%) gender gap in the patent grant rates. This gap is explained by the covariates to a similar extent as the gender gap in filing rates in Panel A. The Oaxaca-Blinder decomposition results for the later-granted patent filings, reported in Columns 2.1 and 2.2 of Table 3, also lead to conclusions that closely mirror those for the overall filing probability discussed above.

Differences in grant rates may reflect differences in the quality of applications. I further examine gender differences in grant rates and multiple measures of patent quality drawn from the prior literature using the full inventor-patent data spanning 1987–2019. The analysis, results, and discussions are presented in Supplemental Appendix E and Table A.6. In summary, patent applications by female inventors have significantly lower grant rates and receive fewer citations than those by male inventors, even after controlling for multiple application- and inventor-level characteristics. Notably, female inventors are more than 50% less likely to have a “breakthrough” patent, defined as a patent in the top 1% of received citations within an EPO filing-year–technology-field cohort.

These results align with the prior evidence from the U.S. and are not necessarily straightforward to interpret. While lower grant rates may reflect differences in application quality, they may also be driven by female inventors’ lower persistence following initial rejection (Aneja et al., 2024), inadequate access to support in the application process (Pairolero et al., 2025), and gender biases among patent examiners (Jensen et al., 2018; Gavrilova-Zoutman and Juranek, 2025). In addition, while citations reflect technological impact and patents’ economic value (e.g., Trajtenberg, 1990; Harhoff et al., 1999; Hall et al., 2005), patents from

female inventors may also be undercited relative to comparable patents by male inventors (Hochberg et al., Forthcoming; Li and Liu, Forthcoming).

Moreover, I find that for other proxies of patent quality — patent family size and variables that capture patents’ technological scope, originality, and radicalness — the raw differences favor female inventors, and once the control variables are included, the gender differences effectively disappear. Composite indices of patent quality likewise indicate only very modest gender differences. Taken together, my findings on gender differences in patent quality are mixed but indicate modest rather than large differences.¹⁵

Gender patenting gaps across occupations. The results above reveal a large average gender gap in patenting within occupations. To analyze the pervasiveness of these within-occupation differences, I estimate the regression model in Equation (1) separately for each occupation in which both men and women have engaged in patenting. The data include 462 occupation titles with any patenting activity: 185 have had both male and female inventors, 253 have seen patenting exclusively by men, and only 24 exclusively by women.¹⁶ I estimate the model both without controls and with fixed effects for calendar year, age, native language, education level by field, and high school grades. The former specification captures the raw gender gaps, whereas the latter estimates the ‘penalties’ associated with being a woman relative to men with the same occupation, potential experience, education, and ability, as proxied by high school performance.

Figure 2 displays the estimated gender gap coefficients for each occupation.¹⁷ Panel

¹⁵My estimates for patent citations and patent family size are consistent in sign and broadly similar in magnitude to those of Hoisl and Mariani (2017), who analyze a sample of inventors and patents across 23 countries, albeit using somewhat different regression specifications. They conclude that female inventors’ patents do not appear to be lower in quality than those of their male counterparts.

¹⁶Since I observe occupations only at the end of the year, there may be some measurement error in linking occupations to patenting. Some inventors may file patents before changing occupation within the same year, leading to misassignment. Using data from the previous year would reduce this error but also introduce another issue for individuals who first switch occupations and then patent within the same year. This limitation is unlikely to substantially affect the results, as occupations are relatively stable: over the 15 years from 2004 to 2018, the average inventor held 3.7 different occupation titles. Nonetheless, this measurement issue may partly explain the large number of occupations with very low patenting activity, which may also reflect rare cases of individuals patenting independently while working in unrelated jobs.

¹⁷I exclude a single occupation from the figure with an average workforce size below 30, as the estimate’s

A displays the raw gender gaps. Panel B displays the estimates from the specification with the controls. The horizontal axis orders the occupations by male workers’ raw annual probability of patenting, providing a natural benchmark for interpreting the magnitude of the estimates.¹⁸ Occupations are categorized by their main occupational group: managers, professionals, technicians and associate professionals, and other workers.¹⁹

Figure 2 here

The estimates highlight the following key findings. First, patenting activity is highly concentrated in a small subset of occupations, with most occupations exhibiting low rates of patenting. The majority of occupations with high patenting activity are classified as professionals, except for the second-rightmost estimate, which corresponds to R&D managers. Second, women are less likely than men to patent in a clear majority of occupations. Furthermore, the coefficients in panel B of Figure 2 indicate that, particularly at higher levels of patenting activity, the gender gaps remain largely statistically significant even after controlling for worker characteristics: of the 185 occupations, 156 (84%) have a negative gender gap coefficient (82 of which are statistically significant at the 5% level, 44 also at the 0.1% level), while the remaining 29 have a positive point estimate (none of which is statistically significant at the 5% level).

Third, comparing the estimates with men’s patenting probabilities on the horizontal axis shows that the gender gaps are sizeable. For example, the rightmost coefficient of -2.75 pp. in Figure 2, Panel A, corresponds to chemists, among whom men have an annual probability of patenting of roughly 4%. This implies that female chemists have approximately

wide confidence intervals would reduce the readability of the figures.

¹⁸Supplemental Appendix Figures A.1 and A.2 present the occupation-specific gender gap estimates ordered by total inventor intensity and by the share of men in each occupation, respectively. Total inventor intensity closely mirrors men’s patenting probability. When ordered by the share of male workers, no clear pattern in gender patenting gaps emerges. Occupations with the largest gaps (and the highest patenting activity) tend to have a high share of male workers. However, even in patenting-active occupations dominated by women, the gender gap coefficients are systematically negative or indistinguishable from zero.

¹⁹The main occupational groups are based on the first digit of the occupation code. Occupations classified as “managers” or “professionals” correspond to ISCO skill level 4, while “technicians and associate professionals” fall under skill level 3. The “other” group includes occupations at skill levels 1 and 2.

69% lower probability of patenting than male chemists. After accounting for the individual characteristics, about 2.5 pp. (91%) of this gap remains unexplained.

Fourth, the negative relationship between the estimated gender gaps and men’s patenting probabilities is near linear, indicating that although women are also more likely to patent in occupations where men patent at higher rates, men retain a relatively stable proportional advantage. Supplemental Appendix Figure A.3 further illustrates this point by displaying the gender gap coefficients by occupation, scaled by men’s patenting probability. Among the most inventor-intensive 5% of occupations, the relative gender gap averages around -50%.

Finally, to account for sorting into firms, I estimate a specification that includes firm fixed effects in addition to the individual-level controls, separately for each occupation. Supplemental Appendix Figure A.4 shows that the gender gap coefficients remain either essentially unchanged or shift modestly toward zero. The broader pattern persists, but a larger share of estimates are insignificant: of the 185 occupations, 141 have negative point estimates (55 of which are statistically significant at the 5% level, 31 also at the 0.1% level), while the remaining 44 have positive coefficients (none of which are significant at the 5% level).

6 Parenthood and the Gender Gap in Patenting

In general, parenthood has a large and persistent negative impact on women’s labor market outcomes, such as earnings, hours worked, labor force participation, and wage rates (e.g., Angelov et al., 2016; Lundborg et al., 2017; Kleven et al., 2019a,b). In contrast, men are generally either virtually unaffected or experience what is often referred to as a “fatherhood premium.” Multiple papers have documented such a premium in career outcomes, such as earnings, arguing that it may arise from employers’ favorable assessments of fathers as workers, as well as from men’s heightened motivation and social pressure to work harder to provide for their families (Korenman and Neumark, 1991; Lundberg and Rose, 2000; Correll et al., 2007; Killewald, 2013; Goldin et al., 2024).

Career interruptions due to child-rearing may have particularly negative consequences for

inventive productivity, given its reliance on up-to-date expertise and team-based collaboration. Career breaks may weaken professional networks and diminish individuals’ engagement with the latest developments in their field. In addition, mothers may seek tasks that offer predictable hours and flexible arrangements, such as part-time or remote work, which are easier to manage alongside parental responsibilities (Goldin, 2014; Goldin and Katz, 2016; Wiswall and Zafar, 2018; Kleven et al., 2019b).²⁰ Inventive roles may not easily accommodate such arrangements, as they may require extended and unpredictable hours — for example, during time-sensitive experiments in a wet lab.

To examine how parenthood affects inventive productivity, I estimate a matched-control event-study model separately for women and men, comparing trends in patenting activity between parents and comparable childless individuals.²¹ Although the decision to have children is endogenous, a child’s arrival may generate changes in parents’ patenting activity that are orthogonal to unobserved determinants that affect patenting more smoothly over time. For instance, women who choose to have children may differ in their inherent preferences for pursuing innovation relative to women who choose to remain childless; however, such differences would be expected to shape patenting activity more smoothly over time rather than producing sudden shifts at the time of childbirth. A potential threat to identification is that the timing of the first childbirth is not random. For instance, women may choose to have children after their inventive productivity has begun to decline. Reassuringly, I find no evidence of such patterns: neither the raw means nor the event-study estimates in Figure 3

²⁰Time-use surveys conducted in Finland (e.g., Statistics Finland, 2009) and across various European countries (Craig and Mullan, 2011; García-Mainar et al., 2011) indicate that mothers devote more time than fathers to childcare and related housework, such as cleaning.

²¹Using matched childless individuals as the control group for parents (following, e.g., Casarico and Lattanzio, 2023; Kim and Moser, 2025) enables me to construct and transparently display benchmark outcome trajectories for parents, capturing, in particular, gender-specific age trends. Alternative control group options include not-yet-parents or using fathers as a control group for mothers. While using not-yet-parents could help mitigate concerns about unobserved differences between parents and the benchmark group, it would confine benchmark observations to the period before the first childbirth, which, particularly for mothers, restricts the observations to relatively young ages. Using fathers as a control group for mothers would be problematic in the presence of substantial gender differences in patenting trends, as observed in Figure 3, Panel A. Moreover, having a control group of childless men for fathers makes it possible to analyze the dynamics of patenting around the birth of a first child for men as well, which is important for a comprehensive understanding of how parenthood shapes men’s and women’s careers differently.

indicate a pre-birth decline in patenting.

My analysis uses panel data from 1987–2019 on individuals born in 1950 or later, observed at ages 19–65, who had their first child between 1991 and 2019.²² One year before the birth of their first child, I match each parent to a childless individual exactly on calendar year, gender, age, 2-digit education level, 3-digit field of study, mother tongue (Finnish, Swedish, or other), and employment status (employed or other). Then, within the exactly matched strata, each parent is paired with the nearest neighbor childless individual based on a propensity score estimated using a logit model. The logit model includes 3-digit industry and region-of-residence indicators measured in the matching year, as well as employment status and taxable income in each of the four years preceding the parents’ first childbirth.²³ Given the smaller number of childless individuals than parents, I allow each childless individual to be matched with multiple parents across different matching years.²⁴ About 73% of mothers and 79% of fathers are successfully matched, corresponding to roughly 534,000 matched mother–childless woman pairs and 578,000 father–childless man pairs.

Using this matched sample, I estimate the following equation separately for men and women:

$$w_{iyt} = \sum_{\substack{h \in T \\ h \neq -1}} \mathbb{1}\{t = h\} \cdot \left(\phi_t + \delta_t \text{Parent}_i \right) + \alpha_i + \rho_y + \xi_{a(i,y)} + \varepsilon_{iyt}, \quad (4)$$

where w_{iyt} denotes the outcome variable (e.g., an indicator for filing a patent application) for individual i in calendar year y and event year t (years from the first childbirth). For the

²²The sample is limited to individuals born in 1950 or later because childbirth data are only available from 1970 onward, which introduces increasing measurement error in child count for earlier cohorts. Using treatment cohorts from 1991 onward ensures that a pre-parenthood period of at least four years is observable for all individuals and can be used for matching.

²³Industry classification at the 3-digit level distinguishes between relatively specific sub-industries. For example, 2-digit industry “manufacture of chemicals and chemical products” includes 3-digit categories such as “manufacture of basic chemicals”, “manufacture of pesticides and other agrochemical products”, and “manufacture of pharmaceuticals, medicinal chemicals and botanical products”.

²⁴Within a matching year, the matching is without replacement, to limit the number of duplicate benchmark individuals. The presence of duplicate individual observations is accounted for when clustering standard errors. Supplemental Appendix Table A.8 displays the number of duplicate benchmark individuals in the estimation sample.

childless benchmark individuals, t is defined as years from the birth of the matched parent’s first child.²⁵ $Parent_i$ is an indicator for ever having children in the data. The specification includes individual fixed effects α_i , which absorb unobserved individual heterogeneity that is constant over time, calendar year fixed effects ρ_y , which control for business cycle effects, and 3-year age bins $\xi_{a(i,y)}$, which control for life-cycle effects. The coefficients ϕ_t capture the average within-individual outcome trajectory of the benchmark individuals relative to the reference year ($t = -1$), net of controls. The coefficients of interest, δ_t , capture the dynamic differences in the outcome trajectories between parents and their matched childless counterparts. The set T encompasses the event window from $t = -5$ to $t = 15$, as well as binned endpoint periods for $t \leq -6$ and $t \geq 16$. I apply two-way clustering at the individual and match levels to account for dependencies between benchmark individual duplicates and within matched parent–childless individual pairs.

Figure 3, Panel A, displays the raw annual probabilities of filing a patent application in the estimation sample. Figure 3, Panel B, reports the event-study estimates ($\hat{\delta}_t$) for women and men. I scale the estimates by 100 for readability. The pre-treatment trends for mothers and childless women evolve in a parallel manner, supporting the validity of the empirical strategy. Following the first childbirth, mothers’ patenting rates begin to decline. The decrease occurs gradually over the first two years, consistent with the fact that many patent applications filed during this period likely stem from innovation efforts undertaken before parenthood. The decline in patenting activity is considerable: the average annual decrease from childbirth to 15 years afterward is 0.009 pp., with the largest drop of roughly 0.018 pp. occurring two years after the first childbirth. Relative to the reference-year ($t = -1$) patenting probability for mothers (0.023%), these correspond to roughly 39% and

²⁵ As discussed by e.g., [Schmieder et al. \(2023\)](#) and [Wing et al. \(2024\)](#), when I select an explicit benchmark individual for each parent and track them relative to the same event year, I compare the outcome trajectories of parents only to those of matched non-parents, thereby avoiding comparisons between earlier- and later-treated individuals and the resulting complications that arise in two-way fixed effects models (as shown by [De Chaisemartin and d’Haultfoeuille, 2020](#); [Goodman-Bacon, 2021](#)). In addition, using one-to-one matching ensures a constant parent-to-childless individual ratio of one across all treatment cohorts, thereby avoiding the complications from varying treated-to-control ratios noted by [Wing et al. \(2024\)](#).

78% decreases, respectively, as shown by Figure 4. The upward trajectory of the estimates beginning around two years after childbirth suggests that mothers’ patenting rates gradually recover. However, the negative long-term estimates, although statistically insignificant at the 5% level, suggest that the recovery may remain incomplete: the rightmost estimate in Figure 4 shows that the average gap remains around -20% even more than 15 years after the first childbirth.

Figures 3 and 4 here

In contrast, fathers’ patenting trajectories initially parallel those of childless men, with no noticeable divergence until about 3–5 years after the birth of the first child. At that point, fathers’ patenting rates rise significantly. On average, fathers’ annual patenting increases by 0.014 pp. relative to the matched childless men over the 15 years following childbirth, corresponding to a 10% increase relative to their reference-year mean. The largest increases occur 5 to 9 years after childbirth, amounting to roughly a 20% rise. Notably, the average 15-year increase in patenting for fathers relative to their reference-year mean is roughly four times smaller than the corresponding decline for mothers.

Given the potentially multi-year gap between the start of an innovation project and patent filing, the lagged increases in patenting are consistent with fathers experiencing a fatherhood premium in their innovation productivity after their child’s birth. As noted above, this premium may reflect both employers’ favorable perceptions of fathers as workers and men’s increased motivation and pressure to increase effort and advance in their careers after having children, potentially leading to a higher proportion of fathers in inventive roles and greater patenting activity. Furthermore, the differing timing of changes in mothers’ and fathers’ patenting activity is consistent with differences in the underlying mechanisms: for mothers, childbirth causes an immediate career disruption, whereas for fathers, status- and motivation-related changes may emerge more gradually.

The shifts in mothers’ and fathers’ patenting activity accumulate over time. Setting aside estimation uncertainty for a moment, the cumulative sum of the point estimates in Figure 3,

Panel B, suggests that mothers experience an overall decline in patenting of close to 0.15 pp. during the 15 years following their first childbirth, which is 6.3 times their reference-period patenting rate, as illustrated by Supplemental Appendix Figure A.5. In other words, the 15-year cumulative loss for mothers is equivalent to roughly six years of patenting at their pre-parenthood patenting rate. By contrast, fathers exhibit a cumulative 15-year increase of roughly 0.23 pp., 1.6 times their baseline patenting rate.

These findings suggest that the asymmetric effects of parenthood play a pivotal role in perpetuating the gender gap in innovation. However, as Figure 3, Panel A, shows, future mothers’ patenting rates already lag behind future fathers’ before parenthood, and women patent clearly less than men even among childless individuals. The regression analysis in Supplemental Appendix F reinforces this point, showing that even after controlling for age, education, high school grades, and occupation, the patenting gap among individuals without children amounts to roughly 48% of the gap observed between mothers and fathers. Parenthood thus amplifies gender disparities in patenting, but cannot fully explain them.

Disentangling mechanisms. Part of the drop in mothers’ patenting rates may be driven by a nearly 20 pp. decline in their employment rate after the first childbirth, with a slow recovery.²⁶ However, mothers may also become less innovation-active conditional on remaining employed, both due to transitions to less inventive occupations and to reduced patenting while remaining in the same occupation. Similarly, part of the estimated fatherhood pre-

²⁶Supplemental Appendix Figure A.6 displays the raw employment rates in the estimation sample. Matching on pre-parenthood employment variables is crucial for constructing a credible control group of childless individuals in the presence of selection into parenthood. A related concern, however, is regression toward the mean (RTM): the employment rates of childless controls may drift toward a lower average in the post period once they are no longer required to match the positively selected future parents. Some RTM may indeed be present in the estimation sample, as the employment rates of, in particular, childless men drop by around 5 pp. at $t = 0$. However, several pieces of evidence suggest that RTM is unlikely to explain my findings on patenting rates. First, employment among childless men declines at $t = 0$ and then stabilizes, whereas the estimated “fatherhood premium” in patenting rates emerges only later, around five years after the first child. Moreover, the raw patterns in Figure 3, Panel A, indicate that the fatherhood premium is driven by increases in fathers’ patenting rates, rather than declines among childless men. Finally, the estimated fatherhood premium persists, although it declines, when the sample is restricted to employed individuals (see Supplemental Appendix Figures A.7 and A.8). Overall, the event-study estimates of patenting rates appear to be driven primarily by changes in outcomes among parents, rather than by shifts in the control group.

mium may reflect fathers’ higher employment rates relative to childless men.

To disentangle these channels, I first re-estimate parents’ patenting responses using a sample that includes only individual-year observations in which the individual is employed.²⁷ Supplementary Appendix Figure A.7 displays raw patenting rates conditional on employment status, while Figure A.8 presents the event-study estimates comparing the unrestricted sample with the sample restricted to employed individuals. Consistent with the fact that fewer than 3% of patents are filed in years of non-employment, mean patenting rates increase. For mothers, the event-study estimates change only slightly. Thus, the decline in patenting does not appear to be driven by reduced labor market attachment, but rather by a fall in patenting activity among those who remain employed. For fathers, the positive estimates are attenuated, consistent with part of the fatherhood premium, particularly in the long run, reflecting higher employment.

To assess whether the decline in mothers’ patenting rates is driven by a shift toward less-inventive occupations, I re-estimate the regression model using occupational inventor intensity as the outcome.²⁸ Figure 5 shows that, conditional on working, mothers’ occupational inventor-intensiveness gradually decreases after the first childbirth, relative to the occupations of the matched childless women. Relative to mothers’ reference-year mean (0.5 inventors per 1000 workers), the decline in occupational inventiveness 10 years after the first child is around 18%.²⁹ For fathers, if anything, the occupational inventor intensity increases modestly compared to childless men, by around 2% relative to the reference-year mean (1.4

²⁷An individual’s employment status is determined by whether they earned wages in the last week of the year. This exercise should be viewed as descriptive, as employment is a potential outcome of parenthood.

²⁸As described in Section 3, inventor intensity is defined as the mean number of inventors per thousand workers within each narrowly defined occupation over 2004–2018, and is therefore time-invariant within occupations. Inventor intensity is very low in most occupations: the threshold to be included in the top 5% is 3.8 inventors per 1000 workers (see Table 1). Moreover, for roughly half of all occupations, the measure equals zero, indicating that no workers in those occupations have patented. Occupation codes are observed in 1990, 1993, 1995, 2000, and annually from 2004 to 2018; accordingly, the estimation sample in Figures 5 and A.9 is restricted to these years. The raw means of parents’ and benchmark individuals’ occupations’ inventor-intensity within the estimation window are shown in Supplemental Appendix Figure A.9.

²⁹Similarly, relative to their reference-year mean, mothers’ likelihood of holding a top 5% inventor-intensive occupation decreases by roughly 20% by 10 years after the first child and remains at that level thereafter (see Supplemental Appendix Figure A.10).

inventors per 1,000 workers). These results suggest that, unlike fathers, mothers sort into less patent-active occupations following the birth of their first child.

Figure 5 here

However, the extent to which these occupational changes explain the motherhood penalty in patenting appears limited. I estimate a reweighted version of the event-study model in which the mothers and fathers are weighted to exactly match the occupational distributions of the matched childless women and men, respectively, in each event period t .³⁰ Supplemental Appendix Figure A.11 presents the results, comparing the estimates between the weighted and the corresponding unweighted specifications.³¹ The dynamic coefficients for mothers and fathers change minimally after balancing the occupational distribution. This stability indicates that although mothers do move into less inventor-intensive occupations, the transitions of the observed magnitude are not able to explain the decline in their patenting. Most of the divergence in inventive activity appears to occur within occupations, potentially through changes in tasks, roles, projects, discretionary effort devoted to innovation, working hours, or other within-occupation adjustments.³²

Robustness and complementary outcomes. I test the robustness of the results to an alternative event-study specification, a sample restricted to fully observed parent cohorts, and a relaxed matching criterion. In addition, I present results for three complementary outcome

³⁰I construct the occupation–period-specific weights separately for mothers and fathers. First, I compute the share of parents and childless benchmark individuals in each occupation o and event period t by gender g . The weights for parents of gender g in given o and t are then defined as the ratio of the gender-specific childless share to the parent share. Intuitively, if a relatively larger (smaller) share of benchmark individuals than parents has occupation o in period t , the observations of parents in that occupation and period are upweighted (downweighted). For fully balanced occupations that, e.g., 1% of both fathers and childless men have in a given period, the weight equals 1. Benchmark individuals always have a weight of 1, and the weights are normalized so that the mean is 1 within each period.

³¹Including occupation fixed effects in the regression specification instead of weighting yields very similar results (available upon request). The estimation sample is restricted to employed individuals in years with observed occupation codes. Because occupation is itself a potential outcome of parenthood, this exercise is purely descriptive.

³²Unfortunately, I do not observe working hours.

variables that capture the patenting intensive margin, grant rates, and patent quality, from which the main results abstract.

First, I modify the specification by replacing individual fixed effects with a parent-status indicator and substituting single-year age fixed effects for the 3-year age-bin dummies. This specification uses within-group variation rather than relying solely on within-individual variation, and allows age effects to be modeled more flexibly. As such, the model more closely follows traditional child-penalty specifications (see, e.g., [Kleven et al., 2019b](#)). The estimated parental response patterns remain robust (see Supplemental Appendix Figure [A.12](#)).

Second, not all first-time parent cohorts from 1991–2019 are observed over the full estimation window $t \in [-5, 15]$, because the panel is limited to the years 1987–2019.³³ To address concerns that changes in sample composition over the estimation window may affect the results, I re-estimate the main specification using only always-observed cohorts, namely parents whose first child was born between 1992 and 2004. The main patterns and conclusions remain robust, though the estimates are less precise, reflecting the reduced sample size (see Supplemental Appendix Figure [A.13](#)).

Third, I relax the matching procedure by matching parents and childless control individuals exactly on gender, age, broad 1-digit education level, broad 1-digit field of study, mother tongue, and employment status in the matching year. This relaxed criterion allows me to match more than 99% of parents in the 1991–2019 parent cohorts, yielding true population-level estimates. The estimates based on this sample largely mirror the main findings (see Supplemental Appendix Figures [A.14](#)). However, the point estimates suggest that the decline in mothers’ patenting activity at $t = 2$ is somewhat smaller, roughly 65%.

Supplemental Appendix Figures [A.15](#) and [A.16](#) show the estimated impacts of parenthood using the number of annual patent filings as the outcome. The estimated patterns change only modestly relative to the main results using a filing indicator; inventors submit a single patent application per year in 70% of cases. However, the long-term point estimates are

³³Supplemental Appendix Table [A.7](#) illustrates when each cohort is observed within the estimation window.

closer to zero for men, suggesting that although fathers maintain higher patenting frequency relative to childless men, the advantage on the intensive margin does not persist.

Supplemental Appendix Figures A.17 and A.18 present results using an indicator for filing a patent application that is eventually granted as the outcome. Relative to their reference-year mean, point estimates suggest mothers experience a broadly similar decline in granted patent filing rates as in total filing rates. For fathers, the advantage over childless men is attenuated compared to total filing rates, suggesting that the fatherhood premium does not extend to patenting success.

Finally, Supplemental Appendix Figure A.19 compares the main estimates against those obtained using a patent-filing indicator weighted by filing-year–technology-sector adjusted received citations.³⁴ To facilitate comparisons, the figure presents estimates scaled by parents’ reference-year means for each respective outcome. The citation-weighted indicator takes the value 0 for no patents, 1 for filings with no forward citations, and 1 plus the number of adjusted citations for filings with citations.³⁵ Then, for instance, if the motherhood penalty is larger with the citation-weighted indicator, this suggests that women not only patent less but also produce lower-impact patents after childbirth. However, Figure A.19 shows no systematic differences in the estimates for mothers with and without reweighting. For fathers, the estimated premiums appear to systematically decrease after reweighting, although the differences are small. Nevertheless, the attenuation under citation weighting together with the weaker premium in granted filings suggests that fathers’ higher filing propensity is not matched by a proportional rise in patent quality.

³⁴Received citations reflect invention’s technological impact and correlate with economic value (Trajtenberg, 1990; Harhoff et al., 1999; Hall et al., 2005). I adjust received citations for filing year and technology sector to account for differences in citation accumulation time and sectoral citation intensity. Specifically, I estimate a Poisson fixed-effects model in which I explain the number of citations received by Finnish inventors’ patent applications — measured as simple DOCDB family-to-family citations as of 2023 — with technology-sector and filing-year fixed effects. I then use the model to predict the expected number of citations for each sector–year cell and divide observed citations by these predictions. The resulting measure equals 0 for uncited patents, falls between 0 and 1 for patents cited less than predicted, and exceeds 1 for patents cited more than predicted.

³⁵If the inventor files multiple patents in the same year, the filing dummy is weighted with the mean of the adjusted citations of those filings.

7 Discussion on Other Potential Mechanisms

My results indicate the presence of additional relevant mechanisms within occupations and firms underlying the gender gap in innovation, beyond the career consequences of parenthood. Prior literature suggests several potential ones. Financial incentives to pursue innovation may vary by gender. [Bell et al. \(2019\)](#) report a gender income gap of roughly 14% among inventors. [Hoisl and Mariani \(2017\)](#) document a comparable gap, with more than half of it unexplained by a rich set of inventor, job, and patent characteristics. Patent grants generate monetary returns for inventors ([Toivanen and Väänänen, 2012](#); [Kline et al., 2019](#)), linking the gender gap in patent grant rates to lower expected returns to patenting for women. In addition, [Toivanen and Väänänen \(2012\)](#) present evidence suggesting that female inventors capture lower returns to patent grants than male inventors.

In addition, even within the same occupation, women may perform different tasks than men ([Fana et al., 2021](#); [Pető and Reizer, 2021](#)).³⁶ Previous research indicates that these task differences help explain occupational gender disparities, such as wage gaps ([Stinebrickner et al., 2018](#); [Bizopoulou, 2019](#)). Using survey data on U.S. college graduates' job characteristics, [Hunt et al. \(2013\)](#) shows that women's underrepresentation in patent-intensive job tasks is a key factor behind the gender gap in patenting.³⁷

Women may encounter greater friction and barriers to innovation than men. [Chien and Grennan \(2024\)](#) present evidence that female engineers at high-tech firms are substantially less likely to have their early-stage inventions advanced to the patent application stage than their male counterparts. Management practices and corporate culture seem to play a crucial role in this disparity: women report receiving inadequate support from management during the inventing process, as well as a lack of collaborators and advisors to evaluate the patent-worthiness of their ideas. Relatedly, [Aneja et al. \(2024\)](#) find that majority-female inventor

³⁶Parenthood may affect the types of tasks that women and men choose or are assigned to ([Tannenbaum and Timpe, 2025](#); [Schulz and Zoch, 2025](#)).

³⁷However, [Hunt et al. \(2013\)](#) do not distinguish between occupations and job tasks, and do not study the allocation of tasks within occupations.

teams are more likely to abandon patent applications after an initial rejection, which the authors primarily attribute to resource constraints and institutional factors, a view supported also by the findings of [Pairolero et al. \(2025\)](#).

There may also be behavioral explanations for the gender patenting gap, such as risk and competition aversion. Patenting and research tasks are arguably competitive and involve a significant risk of failure, and typically take place in male-dominated fields and occupations. Literature indicates that women are less willing to enter competitive environments than men ([Niederle and Vesterlund, 2007](#); [Flory et al., 2015](#)) and may perform worse in competition ([Gneezy et al., 2003](#); [Gneezy and Rustichini, 2004](#)), especially against men ([Gneezy et al., 2003](#); [Niederle and Vesterlund, 2011](#)). Moreover, women may have lower preferences for risk-taking than men, especially in mixed-sex environments ([Booth and Nolen, 2012](#)).³⁸ While sorting into fields and occupations likely attenuates gender differences in risk tolerance and competitiveness (e.g., [Buser et al., 2014](#); [Fouarge et al., 2014](#); [Dillon, 2018](#)), some within-group differences may persist.

Finally, women may also face gender biases, discrimination, and exclusion, making it more difficult for them to access inventor roles and receive recognition for their contributions. [Ross et al. \(2022\)](#) document a 58% gap between men and women in the likelihood of being named as an inventor on patents produced by their research teams, and a 13% gap in the probability of being credited as an author on academic publications, both measured after controlling for team members' job titles, tenure in the team, and team fixed effects. Furthermore, recent evidence suggests that female inventors' patents are under-recognized, as measured by the number of citations they receive, relative to their relevance ([Hochberg et al., Forthcoming](#); [Li and Liu, Forthcoming](#)).

³⁸A substantial body of literature generally concludes that women tend to be more risk-averse than men (see reviews in, e.g., [Croson and Gneezy, 2009](#); [Charness and Gneezy, 2012](#), and their references). However, a less unanimous interpretation of the literature has also been made ([Nelson, 2015](#)).

8 Conclusions

In this paper, I examine the forces driving the persistent gender divide in innovation, using European Patent Office records linked to Finnish administrative data. The raw difference is stark: women’s patenting rate is 88% lower than men’s. I show that roughly half of this gap stems from educational and occupational gender segregation: men are more likely to sort into patent-intensive domains, such as most engineering professions. The other half reflects the more puzzling differences in patenting activity between women and men with similar educational backgrounds working within the same narrowly defined occupations. Women are less likely to patent in nearly every occupation, and accounting for differences in potential experience, qualifications, general ability proxied by high school grades, and sorting across firms with different innovation capacities leaves these within-occupation gaps largely unexplained.

The asymmetric effects of parenthood are a key mechanism: using a matched-control event-study design, I show that women’s patenting activity drops by approximately 78% following the first childbirth, with evidence of partial recovery over the long term. The cumulative 15-year decline for mothers is equivalent to roughly six years of patenting at their pre-parenthood rate. Parenthood appears to affect mothers’ inventive productivity primarily through adjustments within occupations, rather than through movements to less patent-intensive occupations. These adjustments may include, for example, shifts to non-inventive tasks and roles, or a reduction in discretionary effort devoted to innovation. Fathers, in contrast, exhibit a sustained annual premium in their patenting rates of roughly 10%. However, I also find considerable gender differences in the patenting rates among future parents before they have children, as well as among childless individuals. Together, my findings indicate that the differential impacts of parenthood are a central driver of the gender gap in innovation, though they cannot fully account for it.

The implications of my results are double-edged. On the one hand, they highlight the importance of policies that encourage women to pursue male-dominated, patent-intensive

fields, such as many STEM fields, to reduce gender disparities in patenting and unlock women’s untapped inventive potential. Such policies could involve exposing students in primary and secondary education to female role models through school visits or mentoring programs, as well as implementing initiatives that make STEM fields more inclusive and appealing to women. On the other hand, my results also indicate that these measures, while valuable, are insufficient on their own, as substantial and pervasive disparities persist within occupations and firms.

My findings further suggest that policies designed to reduce the disproportionate career penalties of parenthood and promote a more equitable division of family responsibilities could increase women’s inventive productivity and strengthen their representation among inventors. Policymakers and employers seeking to foster innovation should consider whether aspects of working conditions, workplace culture, and innovation processes are unnecessarily difficult to combine with family life — or otherwise reinforce gender disparities — and result in an inefficient use of women’s innovation potential. Further investigating why women continue to lag behind men in patenting, even after entering patent-intensive occupations and firms, remains a promising avenue for future research.

References

- AGHION, P., U. AKCIGIT, A. HYYTINEN, AND O. TOIVANEN (2017): “The Social Origins of Inventors,” Working Paper 24110, National Bureau of Economic Research.
- (2018): “On the Returns to Invention Within Firms: Evidence from Finland,” in *AEA Papers and Proceedings*, vol. 108, 208–212.
- (2023): “Parental Education and Invention: The Finnish Enigma,” *International Economic Review*, 64, 453–490.
- ANEJA, A., O. RESHEF, AND G. SUBRAMANI (2024): “Attrition and the Gender Patenting Gap,” *Review of Economics and Statistics*, 1–31.
- ANGELOV, N., P. JOHANSSON, AND E. LINDAHL (2016): “Parenthood and the Gender Gap in Pay,” *Journal of Labor Economics*, 34, 545–579.
- 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.

- BERTRAND, M., C. GOLDIN, AND L. F. KATZ (2010): “Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors,” *American Economic Journal: Applied Economics*, 2, 228–255.
- BIANCHI, N. AND M. GIORCELLI (2020): “Scientific Education and Innovation: from Technical Diplomas to University STEM Degrees,” *Journal of the European Economic Association*, 18, 2608–2646.
- BIZOPOULOU, A. (2019): “Job Tasks and Gender Wage Gaps within Occupations,” VATT Institute for Economic Research Working Papers 124.
- BLAU, F. D. AND L. M. KAHN (2017): “The Gender Wage Gap: Extent, Trends, and Explanations,” *Journal of Economic Literature*, 55, 789–865.
- BLINDER, A. S. (1973): “Wage Discrimination: Reduced Form and Structural Estimates,” *Journal of Human resources*, 436–455.
- BOOTH, A. L. AND P. NOLEN (2012): “Gender Differences in Risk Behaviour: Does Nurture Matter?” *The Economic Journal*, 122, F56–F78.
- BREDA, T., J. GRENET, M. MONNET, AND C. VAN EFFENTERRE (2023): “How Effective Are Female Role Models in Steering Girls Towards STEM? Evidence From French High Schools,” *The Economic Journal*, 133, 1773–1809.
- BROUILLETTE, J.-F. (forthcoming): “The Lost Marie Curies and Foregone Economic Growth,” *American Economic Journal: Macroeconomics*.
- BUSER, T., M. NIEDERLE, AND H. OOSTERBEEK (2014): “Gender, Competitiveness, and Career Choices,” *The Quarterly Journal of Economics*, 129, 1409–1447.
- CASARICO, A. AND S. LATTANZIO (2023): “Behind the Child Penalty: Understanding What Contributes to the Labour Market Costs of Motherhood,” *Journal of Population Economics*, 36, 1489–1511.
- CECI, S. J., D. K. GINTHER, S. KAHN, AND W. M. WILLIAMS (2014): “Women in Academic Science: A Changing Landscape,” *Psychological Science in the Public Interest*, 15, 75–141.
- CHARNESS, G. AND U. GNEEZY (2012): “Strong Evidence for Gender Differences in Risk Taking,” *Journal of Economic Behavior & Organization*, 83, 50–58.
- CHIEN, C. V. AND J. GRENNAN (2024): “Unpacking the Innovator-Inventor Gap: Evidence from Engineers,” Available at SSRN: <http://dx.doi.org/10.2139/ssrn.4721522>.
- CORRELL, S., S. BENARD, AND I. PAIK (2007): “Getting a Job: Is There a Motherhood Penalty?” *American Journal of Sociology*, 112, 1297–1338.
- CRAIG, L. AND K. MULLAN (2011): “How Mothers and Fathers Share Childcare: A Cross-National Time-Use Comparison,” *American sociological Review*, 76, 834–861.

- CROSON, R. AND U. GNEEZY (2009): “Gender Differences in Preferences,” *Journal of Economic Literature*, 47, 448–474.
- DE CHAISEMARTIN, C. AND X. D’HAULTFOEUILLE (2020): “Two-Way Fixed Effects Estimators With Heterogeneous Treatment Effects,” *American Economic Review*, 110, 2964–2996.
- DILLON, E. W. (2018): “Risk and Return Trade-Offs in Lifetime Earnings,” *Journal of Labor Economics*, 36, 981–1021.
- DING, W. W., F. MURRAY, AND T. E. STUART (2006): “Gender Differences in Patenting in the Academic Life Sciences,” *Science*, 313, 665–667.
- EINIÖ, E., J. FENG, AND X. JARAVEL (Forthcoming): “Social Push and the Direction of Innovation,” *American Economic Review*.
- EPO (2022): “Women’s Participation in Inventive Activity: Evidence from EPO Data,” Report, European Patent Office (EPO).
- (2025a): “How Long Does the Grant Procedure Take?” Accessed: 2025-04-16; <https://www.epo.org/en/service-support/faq/applying-patent/procedure/how-long-does-grant-procedure-take>.
- (2025b): “What to Expect from the Patent Process,” Accessed: 2025-04-16; <https://www.epo.org/en/new-to-patents/what-to-expect>.
- FANA, M., D. VILLANI, AND M. BISELLO (2021): “Mind the Task: Evidence on Persistent Gender Gaps at the Workplace,” JRC Working Papers Series on Labour, Education and Technology 2021/03.
- FLORY, J. A., A. LEIBBRANDT, AND J. A. LIST (2015): “Do Competitive Workplaces Deter Female Workers? A Large-Scale Natural Field Experiment on Job Entry Decisions,” *The Review of Economic Studies*, 82, 122–155.
- FOUARGE, D., B. KRIECHEL, AND T. DOHMEN (2014): “Occupational Sorting of School Graduates: The Role of Economic Preferences,” *Journal of Economic Behavior & Organization*, 106, 335–351.
- GARCÍA-MAINAR, I., J. A. MOLINA, AND V. M. MONTUENGA (2011): “Gender Differences in Childcare: Time Allocation in Five European Countries,” *Feminist Economics*, 17, 119–150.
- GAVRILOVA-ZOUTMAN, E. AND S. JURANEK (2025): “The Role of Reviewer Characteristics on the Diversity of Successful Applicants,” *Organization Science*, 36, 1726–1744.
- GNEEZY, U., M. NIEDERLE, AND A. RUSTICHINI (2003): “Performance in Competitive Environments: Gender Differences,” *The Quarterly Journal of Economics*, 118, 1049–1074.
- GNEEZY, U. AND A. RUSTICHINI (2004): “Gender and Competition at a Young Age,” *American Economic Review*, 94, 377–381.

- GOLDIN, C. (2006): “The Quiet Revolution That Transformed Women’s Employment, Education, and Family,” *American Economic Review*, 96, 1–21.
- (2014): “A Grand Gender Convergence: Its Last Chapter,” *American Economic Review*, 104, 1091–1119.
- GOLDIN, C. AND L. F. KATZ (2016): “A Most Egalitarian Profession: Pharmacy and the Evolution of a Family-Friendly Occupation,” *Journal of Labor Economics*, 34, 705–746.
- GOLDIN, C., S. P. KERR, AND C. OLIVETTI (2024): “The Parental Pay Gap Over the Life Cycle: Children, Jobs, and Labor Supply,” *Journal of Economic Dynamics and Control*, 169, 104963.
- GOLDIN, C., S. P. KERR, C. OLIVETTI, AND E. BARTH (2017): “The Expanding Gender Earnings Gap: Evidence From the Lehd-2000 Census,” *American Economic Review*, 107, 110–114.
- GOODMAN-BACON, A. (2021): “Difference-in-Differences with Variation in Treatment Timing,” *Journal of Econometrics*, 225, 254–277.
- HALL, B. H., A. JAFFE, AND M. TRAJTENBERG (2005): “Market Value and Patent Citations,” *The RAND Journal of Economics*, 36, 16–38.
- HARHOFF, D., F. NARIN, F. M. SCHERER, AND K. VOPEL (1999): “Citation Frequency and the Value of Patented Inventions,” *Review of Economics and statistics*, 81, 511–515.
- HAUSMANN, R., L. D. TYSON, AND S. ZAHIDI (2009): “The Global Gender Gap Report 2009,” World Economic Forum.
- HOCHBERG, Y. V., A. KAKHBOD, P. LI, AND K. SACHDEVA (Forthcoming): “Are Patents With Female Inventors Under-Cited? Evidence From Text Estimation,” *Journal of Financial Economics*.
- HOISL, K., H. C. KONGSTED, AND M. MARIANI (2022): “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.
- HOSPIDO, L., L. LAEVEN, AND A. LAMO (2022): “The Gender Promotion Gap: Evidence From Central Banking,” *Review of Economics and Statistics*, 104, 981–996.
- HSIEH, C.-T., E. HURST, C. I. JONES, AND P. J. KLENOW (2019): “The Allocation of Talent and U.S. Economic Growth,” *Econometrica*, 87, 1439–1474.
- HUNT, J. (2016): “Why Do Women Leave Science and Engineering?” *ILR Review*, 69, 199–226.
- HUNT, J., J.-P. GARANT, H. HERMAN, AND D. J. MUNROE (2013): “Why Are Women Underrepresented Amongst Patentees?” *Research Policy*, 42, 831–843.

- JANN, B. (2008): “The Blinder–Oaxaca Decomposition for Linear Regression Models,” *The stata journal*, 8, 453–479.
- JENSEN, K., B. KOVÁCS, AND O. SORENSON (2018): “Gender Differences In Obtaining And Maintaining Patent Rights,” *Nature biotechnology*, 36, 307–309.
- JUNG, T. AND O. EJERMO (2014): “Demographic Patterns and Trends in Patenting: Gender, age, and Education of Inventors,” *Technological Forecasting and Social Change*, 86, 110–124.
- KAHN, S. AND D. GINTHER (2017): “Women and STEM,” Working Paper w23525, National Bureau of Economic Research.
- KALI PAL, K., K. PIAGET, AND S. ZAHIDI (2024): “Global Gender Gap Report 2024,” World Economic Forum.
- KILLEWALD, A. (2013): “A Reconsideration of the Fatherhood Premium: Marriage, Coresidence, Biology, and Fathers’ Wages,” *American sociological review*, 78, 96–116.
- KIM, S. AND P. MOSER (2021): “Women in Science. Lessons From the Baby Boom,” Working Paper w29436, National Bureau of Economic Research.
- (2025): “Women in Science. Lessons From the Baby Boom,” *Econometrica*, 93, 1521–1560.
- KLEVEN, H., C. LANDAIS, J. POSCH, A. STEINHAUER, AND J. ZWEIMÜLLER (2019a): “Child Penalties Across Countries: Evidence and Explanations,” *AEA Papers and Proceedings*, 109, 122–126.
- KLEVEN, H., C. LANDAIS, AND J. E. SØGAARD (2019b): “Children and Gender Inequality: Evidence from Denmark,” *American Economic Journal: Applied Economics*, 11, 181–209.
- KLINE, P., N. PETKOVA, H. WILLIAMS, AND O. ZIDAR (2019): “Who Profits from Patents? Rent-Sharing at Innovative Firms,” *The Quarterly Journal of Economics*, 134, 1343–1404.
- 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.
- KORENMAN, S. AND D. NEUMARK (1991): “Does Marriage Really Make Men More Productive?” *Journal of Human Resources*, 282–307.
- LI, Z. AND S. LIU (Forthcoming): “Gender Difference in Innovation Recognition: A Textual Analysis Approach,” *Journal of Labor Economics*.
- LIU, A. Y. (2025): “Parenthood and Occupational Mobility,” *Economic Inquiry*, 63, 504–522.

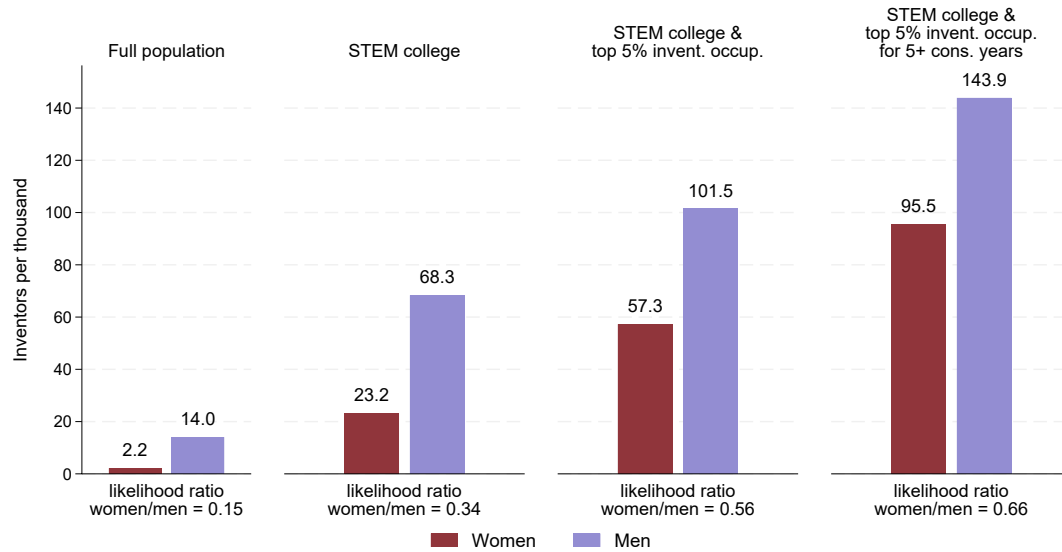
- LUNDBERG, S. AND E. ROSE (2000): “Parenthood and the Earnings of Married Men and Women,” *Labour Economics*, 7, 689–710.
- LUNDBORG, P., E. PLUG, AND A. W. RASMUSSEN (2017): “Can Women Have Children and a Career? IV Evidence From IVF Treatments,” *American Economic Review*, 107, 1611–1637.
- MUELLER, C., S. OBERNBERGER, AND A. SEIDLITZ (2026): “The Missing Female Inventors: Family Formation and the Gender Gap in Innovation,” SSRN Working Paper 5527321, SSRN, available at SSRN.
- NELSON, J. A. (2015): “Are Women Really More Risk-Averse Than Men? A Re-analysis of the Literature Using Expanded Methods,” *Journal of Economic Surveys*, 29, 566–585.
- NIEDERLE, M. AND L. VESTERLUND (2007): “Do Women Shy Away from Competition? Do Men Compete Too Much?” *The Quarterly Journal of Economics*, 122, 1067–1101.
- (2011): “Gender and Competition,” *Annual Review of Economics*, 3, 601–630.
- OAXACA, R. (1973): “Male-Female Wage Differentials in Urban Labor Markets,” *International economic review*, 693–709.
- OLIVETTI, C. AND B. PETRONGOLO (2016): “The Evolution of Gender Gaps in Industrialized Countries,” *Annual Review of Economics*, 8, 405–434.
- PAIROLERO, N. A., A. A. TOOLE, P.-A. PAPPAS, C. A. DE GRAZIA, AND M. H. TEODORESCU (2025): “Closing the Gender Gap in Patenting: Evidence From a Randomized Control Trial at the USPTO,” *American Economic Journal: Economic Policy*, 17, 281–310.
- PETÓ, R. AND B. REIZER (2021): “Gender Differences in the Skill Content of Jobs,” *Journal of Population Economics*, 34, 825–864.
- PRESTON, A. E. (1994): “Why Have All the Women Gone? A Study of the Exit of Women from the Science and Engineering Professions,” *The American Economic Review*, 84, 1446–1462.
- ROSS, M. B., B. M. GLENNON, R. MURCIANO-GOROFF, E. G. BERKES, B. A. WEINBERG, AND J. I. LANE (2022): “Women Are Credited Less in Science Than Men,” *Nature*, 608, 135–145.
- SCHMIEDER, J. F., T. VON WACHTER, AND J. HEINING (2023): “The Costs of Job Displacement Over the Business Cycle and Its Sources: Evidence From Germany,” *American Economic Review*, 113, 1208–1254.
- SCHMUCK, C. (2017): *Women in STEM Disciplines: The Yfactor 2016 Global Report on Gender in Science, Technology, Engineering and Mathematics*, Cham: Springer International Publishing, 1 ed.

- SCHULZ, W. AND G. ZOCH (2025): “The Job Task Penalty for Motherhood,” *Journal of Marriage and Family*.
- SIEPPI, A. AND J. PEHKONEN (2019): “Parenthood and Gender Inequality: Population-based Evidence on the Child Penalty in Finland,” *Economics Letters*, 182, 5–9.
- SPEER, J. D. (2023): “Bye bye Ms. American Sci: Women and the leaky STEM pipeline,” *Economics of Education Review*, 93, 102371.
- SQUICCIARINI, M., H. DERNIS, AND C. CRISCUOLO (2013): “Measuring Patent Quality: Indicators of Technological and Economic Value,” OECD Science, Technology and Industry Working Papers 2013/03, Organisation for Economic Co-operation and Development (OECD), Paris.
- STATISTICS FINLAND (2009): “Official Statistics of Finland (OSF): Time Use [e-publication]. Time use changes through the 2000s,” Referred: 1.10.2025. Access method: http://stat.fi/til/akay/2009/05/akay_2009_05_2011-12-15_tie_001_en.html.
- (2021): “Gender Equality in Finland 2021,” Available at <https://urn.fi/URN:ISBN:978-952-244-670-1>, Statistics Finland.
- (2026a): “FOLK Basic Data,” https://taika.stat.fi/en/aineistokuvaus.html#!?dataid=FOLK_19872023_jua_perus24_001.xml (Accessed January 8, 2026).
- (2026b): “FOLK Child-Parents Data,” https://taika.stat.fi/en/aineistokuvaus.html#!?dataid=FOLK_19702022_jua_lapsv24_001.xml (Accessed January 8, 2026).
- (2026c): “FOLK Degree Data,” https://taika.stat.fi/en/aineistokuvaus.html#!?dataid=FOLK_19872022_jua_tutk24_001.xml (Accessed January 8, 2026).
- (2026d): “FOLK Employment Data,” https://taika.stat.fi/en/aineistokuvaus.html#!?dataid=FOLK_19872022_jua_tkt25_001.xml (Accessed January 8, 2026).
- (2026e): “FOLK Matriculation Examination Data,” https://taika.stat.fi/en/aineistokuvaus.html#!?dataid=EDUC_19902025_jua_ytl_001.xml (Accessed January 8, 2026).
- STINEBRICKNER, T. R., R. STINEBRICKNER, AND P. J. SULLIVAN (2018): “Job Tasks and the Gender Wage Gap among College Graduates,” Working Paper 24790, National Bureau of Economic Research.
- TANNENBAUM, D. AND B. TIMPE (2025): “Job Tasks, Leadership, and the Gender Gap after Parenthood,” *AEA Papers and Proceedings*, 115, 271–75.
- 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.

- TOIVANEN, O. AND L. VÄÄNÄNEN (2012): “Returns to Inventors,” *Review of Economics and Statistics*, 94, 1173–1190.
- (2016): “Education and Invention,” *Review of Economics and Statistics*, 98, 382–396.
- TOOLE, A. A., M. J. SAKSENA, C. A. W. DEGRAZIA, K. P. BLACK, F. LISSONI, E. MIGUELEZ, AND G. TARASCONI (2020): “Progress and Potential: 2020 Update on U.S. Women Inventor-Patentees,” [IP Data Highlights 4](#), USPTO, Office Of The Chief Economist.
- TRAJTENBERG, M. (1990): “A Penny for Your Quotes: Patent Citations and the Value of Innovations,” *The Rand journal of economics*, 172–187.
- WHITTINGTON, K. B. (2011): “Mothers of Invention? Gender, Motherhood, and New Dimensions of Productivity in the Science Profession,” *Work and Occupations*, 38, 417–456.
- 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.
- (2008): “Women Inventors in Context: Disparities in Patenting Across Academia and Industry,” *Gender & Society*, 22, 194–218.
- WING, C., S. M. FREEDMAN, AND A. HOLLINGSWORTH (2024): “Stacked Difference-in-Differences,” *NBER Working Paper 32054*.
- WISWALL, M. AND B. ZAFAR (2018): “Preference for the Workplace, Investment in Human Capital, and Gender,” *The Quarterly Journal of Economics*, 133, 457–507.

Figures and Tables

(A) Probability of Becoming an Inventor



(B) Share of Women

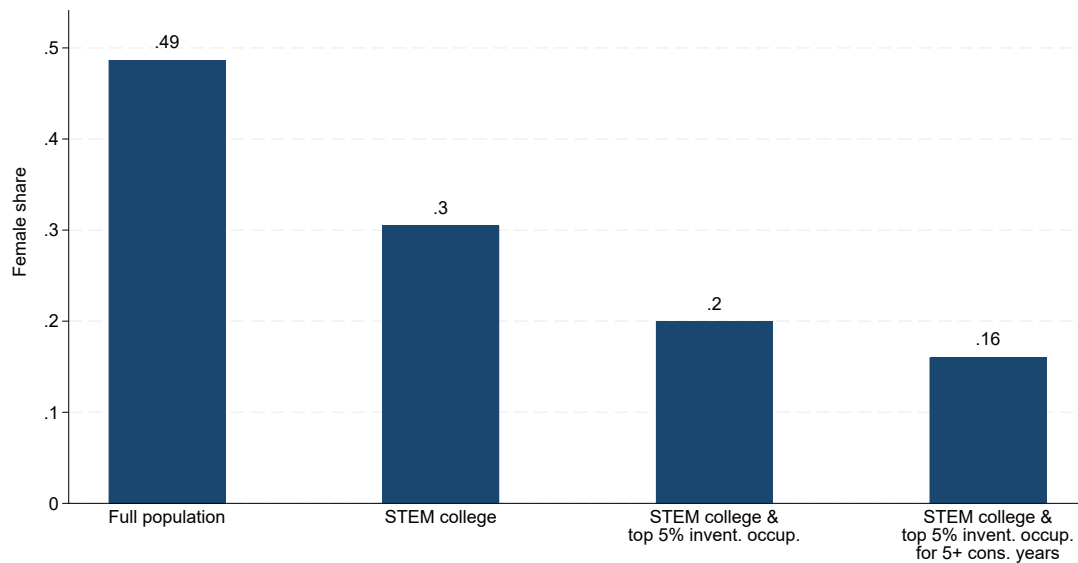


Figure 1: Gender Inventor Gap and Supply of Potential Female Inventors

Notes. Panel A displays the share of female and male inventors among all women and men, respectively, for the full population of birth cohorts from 1962 to 1982 and for the three subsamples of STEM college graduates. Inventor shares are expressed as the number of inventors per thousand individuals. The female-to-male likelihood ratios are displayed below each bar for every sample. Panel B displays the share of women in the samples. The number of individuals and inventors in the samples is reported in Supplemental Appendix Table A.3. Inventorship is determined using EPO patent application data from 1978 to 2019. Occupations are observed between 2004 and 2018.

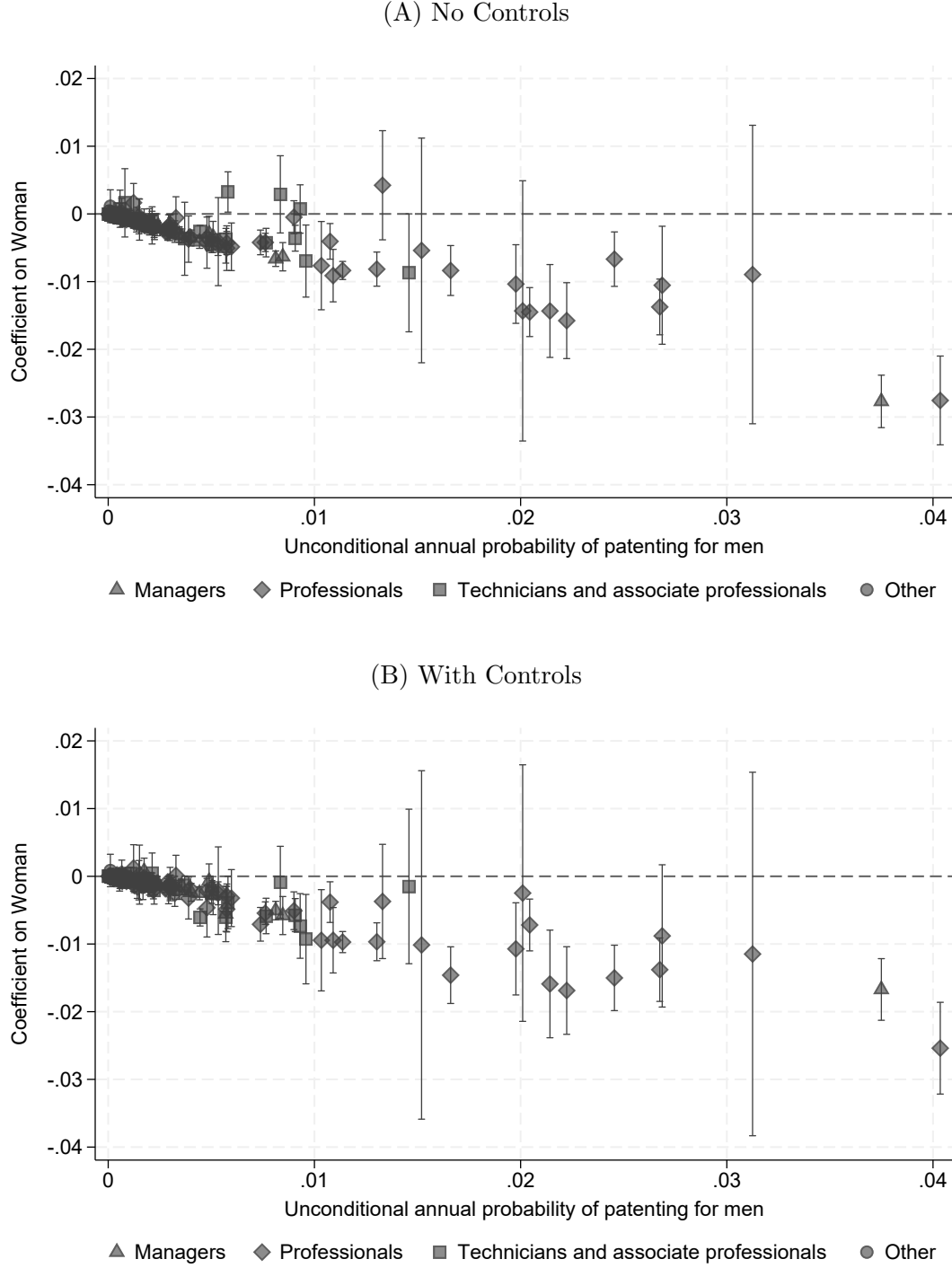


Figure 2: Gender Patenting Gaps by Occupation

Notes. This figure presents the occupation-specific gender gap coefficients, along with their 95% confidence intervals. The coefficients are estimated with the regression model in Equation (1) separately for each occupation. The estimation sample comprises individuals observed in 2004–2018 at ages 25–65. Panel A displays the raw gender gaps. Panel B displays estimates from a specification that includes fixed effects for calendar year, age, native language, high school grades, and education level by field. The horizontal axis orders occupations by the patenting probability of male workers.

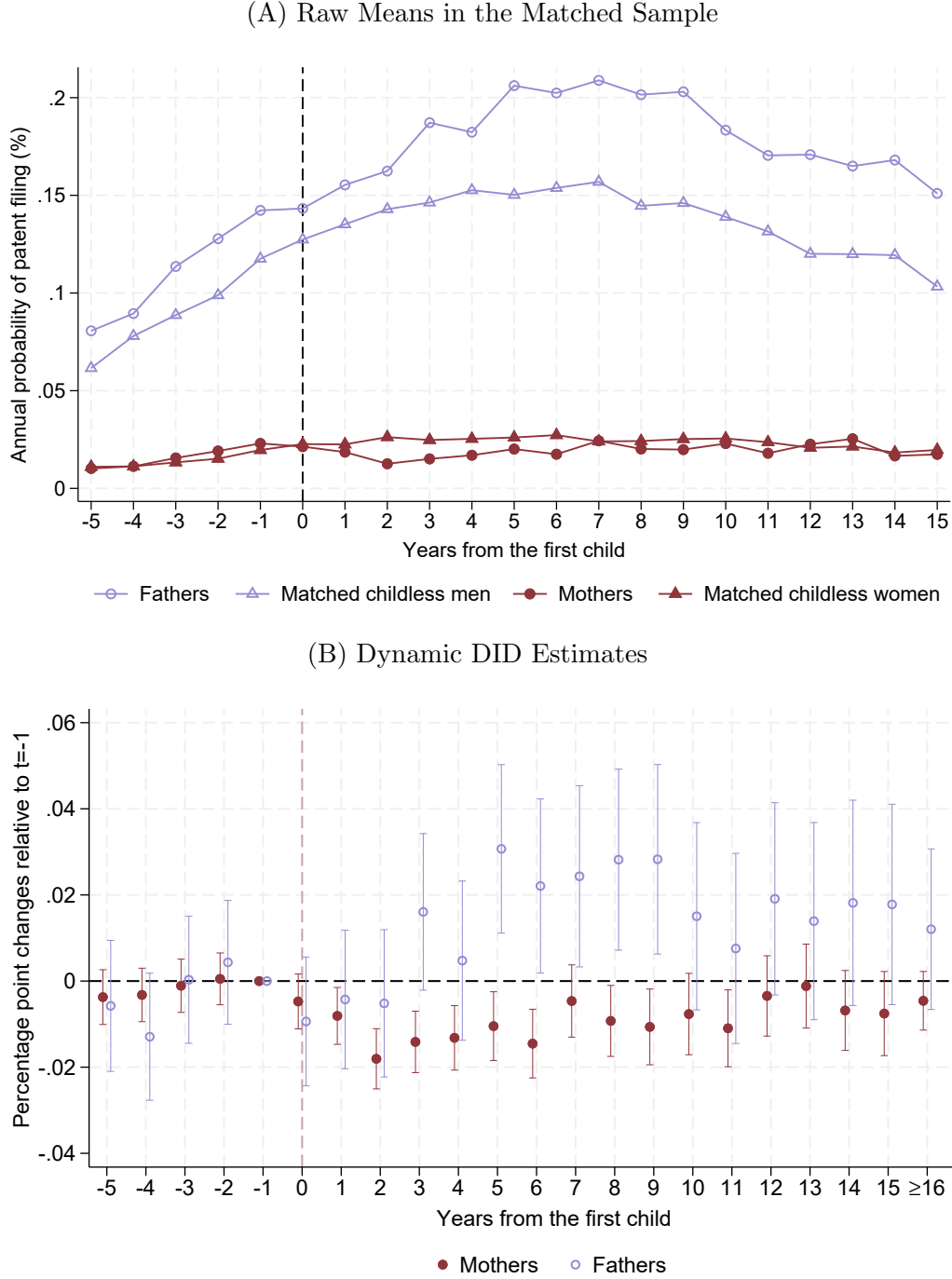


Figure 3: Changes in Annual Patenting Rates Around the First Childbirth

Notes. Panel A displays the raw probabilities of patenting for parents and matched benchmark individuals in the estimation window, expressed as percentages. Panel B plots point estimates and 95% confidence intervals, estimated separately for women and men using the dynamic DID model in Eq. (4). For readability, the estimates are scaled by 100 and should be interpreted in percentage points. The outcome variable is an indicator for filing a patent application in a given year. The model includes calendar-year, individual, and 3-year age-bin fixed effects. The sample comprises matched parent–childless pairs from cohorts born in 1950 or later, observed from 1987 to 2019. Standard errors are clustered at the individual and the matched-pair levels.

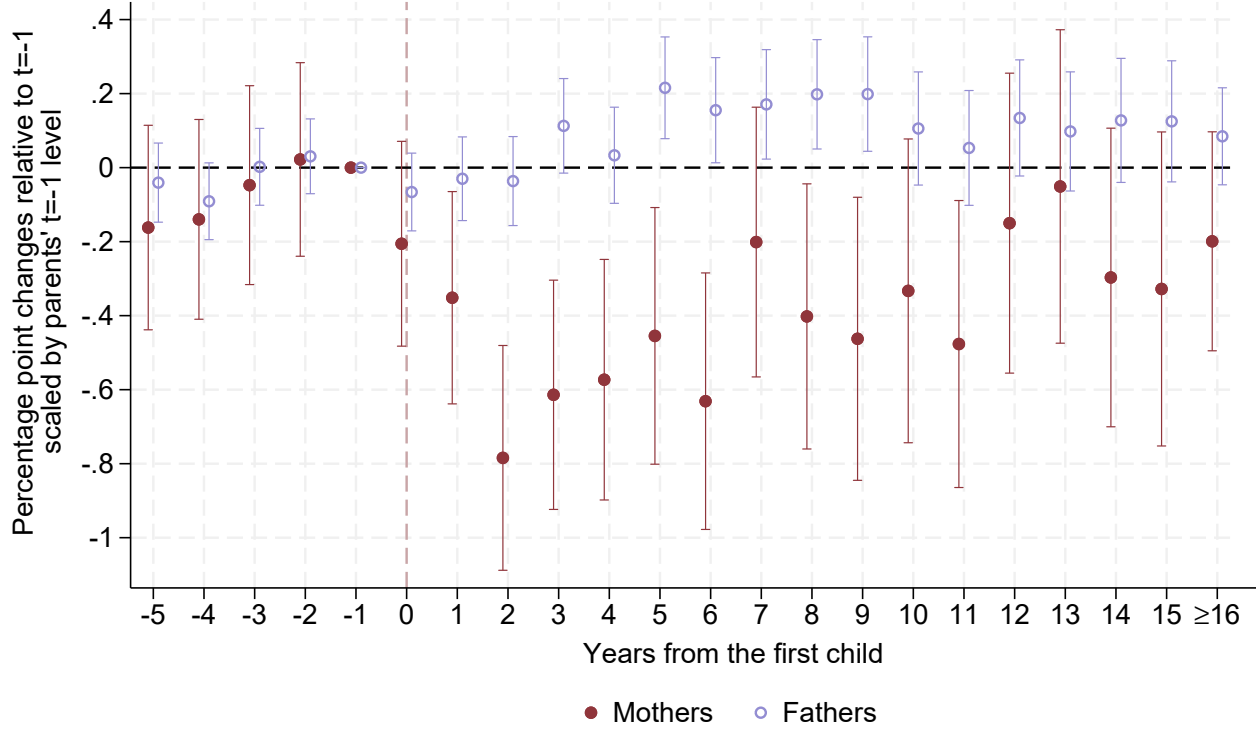


Figure 4: Scaled Dynamic DID Estimates from Figure 3, Panel B

Notes. This figure displays the dynamic DID estimates from Figure 3, Panel B, scaled by the parents' reference year ($t = -1$) patenting rate. The outcome variable is an indicator for filing a patent application in a given year. The dynamic DID estimates are estimated separately for women and men using the model in Eq. (4). The model includes calendar-year, individual, and 3-year age-bin fixed effects. The estimation sample consists of the matched parent-childless individual pairs from cohorts born in 1950 or later, observed from 1987 to 2019. Standard errors are clustered at the individual and the matched-pair levels.

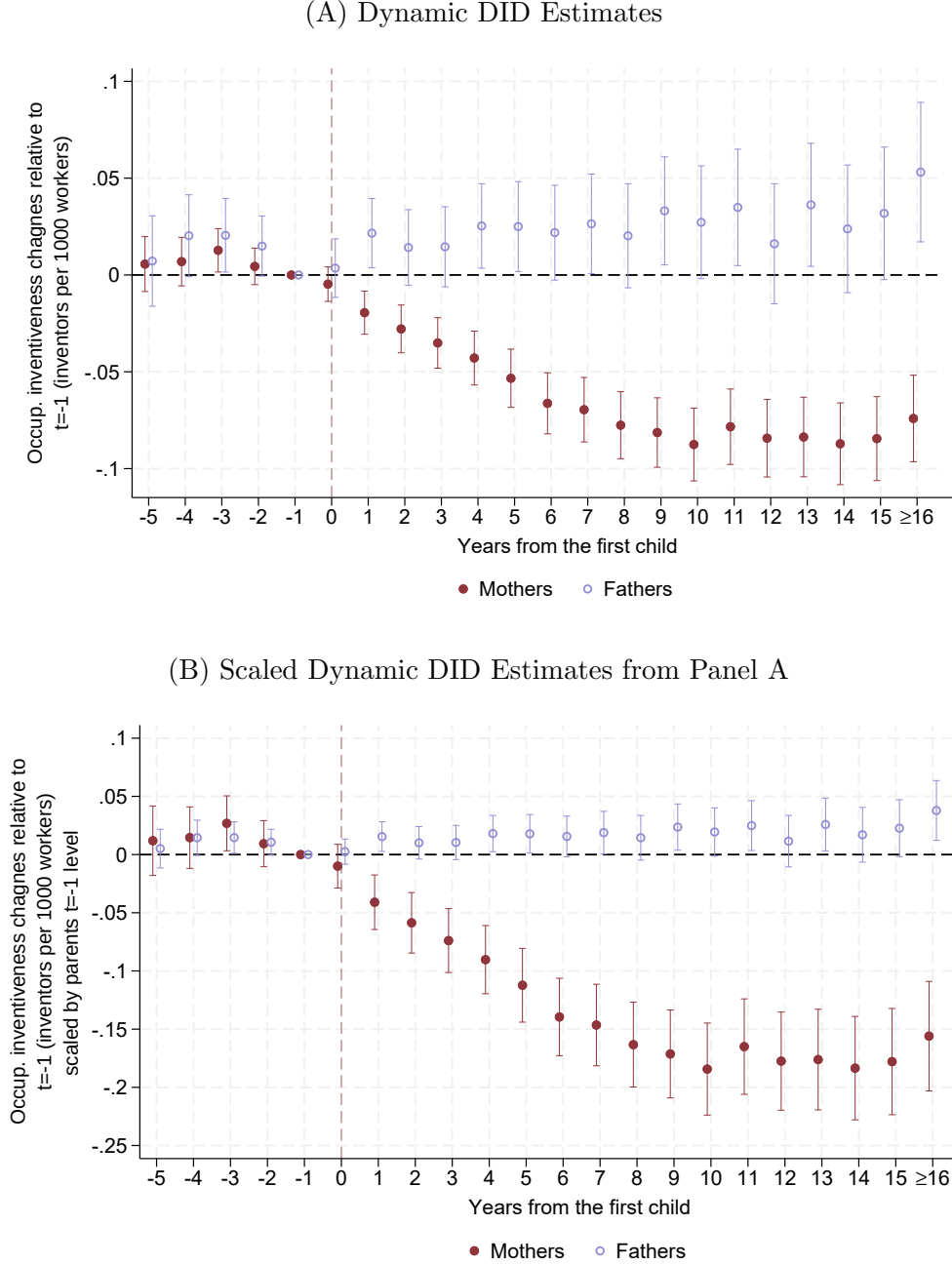


Figure 5: Changes in Occupational Inventiveness Around the First Childbirth:
Inventors per 1000 Workers

Notes. The outcome variable is the mean number of inventors per thousand workers in the occupation held by the individual in a given year. The outcome is an occupation-specific mean calculated using information over the years 2004–2018, as described in Section 3. Panel A plots point estimates and 95% confidence intervals, estimated separately for women and men using the dynamic DID model in Eq. (4). Panel B scales the estimates from Panel A by the parents’ reference year ($t = -1$) outcome level. The model includes calendar-year, individual, and 3-year age-bin fixed effects. The sample comprises matched parent–childless pairs from cohorts born in 1950 or later, observed from 1987 to 2019. Standard errors are clustered at the individual and the matched-pair levels.

Table 1: Top 5% Inventor-Intensive Occupations

	Mean total inventors per 1,000 workers (I)	Mean female inventors per 1,000 female workers (I_w)	Mean male inventors per 1,000 male workers (I_m)	Women less likely to patent (0/1)	Absolute gender gap: $I_w - I_m$	Relative gender gap: $(I_w - I_m)/I_m$
1. Research and development managers	31.5	9.8	37.5	1	-27.7	-0.74
2. Telecommunications researchers	25.4	16.3	26.9	1	-10.5	-0.39
3. Chemists	25.3	12.8	40.4	1	-27.6	-0.68
4. Pharmacologists, pathologists and related professionals	25.0	22.3	31.3	1	-9.0	-0.29
5. Electronics and telecommunications engineers	24.9	13.0	26.7	1	-13.7	-0.51
6. Chemical engineers	22.6	17.9	24.6	1	-6.7	-0.27
7. Information technology researchers	18.8	7.1	21.4	1	-14.3	-0.67
8. Physicists and astronomers	18.6	6.5	22.2	1	-15.8	-0.71
9. Telecommunications engineers	18.2	9.4	19.8	1	-10.4	-0.52
10. Professors	16.6	5.9	20.4	1	-14.5	-0.71
11. Electrical engineers	15.9	8.2	16.6	1	-8.4	-0.50
12. Social science and related professionals	15.8	0.0	24.5	1	-24.5	-1.00
13. Mining engineers, metallurgists and related professionals	14.0	17.5	13.3	0	4.2	0.32
14. Physical and engineering science associate professionals	12.6	5.9	14.6	1	-8.7	-0.60
15. Electronics engineers	12.0	4.9	13.0	1	-8.2	-0.63
16. Health professionals (except nursing)	11.9	9.8	15.2	1	-5.4	-0.35
17. Mechanical engineers	10.8	3.0	11.4	1	-8.4	-0.73
18. Chemical engineering technicians	9.5	10.0	9.3	0	0.7	0.08
19. Mining and metallurgical technicians	9.0	11.2	8.3	0	2.9	0.35
20. Architects, engineers and related professionals not elsewhere classified	8.9	8.5	9.0	1	-0.5	-0.05
21. Health professionals (except nursing) not elsewhere classified	8.8	5.8	20.1	1	-14.3	-0.71
22. Electronics and telecommunications technicians	8.4	5.5	9.1	1	-3.6	-0.40
23. Biologists, botanists, zoologists and related professionals	8.2	6.7	10.8	1	-4.1	-0.38
24. Production and operations managers in manufacturing	7.9	2.1	8.5	1	-6.3	-0.75
25. Manufacturing managers	7.3	1.5	8.1	1	-6.6	-0.81
26. Electronics engineering technicians	6.9	3.4	7.7	1	-4.2	-0.55
27. Engineering professionals not elsewhere classified	6.6	3.4	7.6	1	-4.2	-0.55
28. Wood processing and chemical engineering technicians	6.6	9.0	5.8	0	3.2	0.56
29. Industrial and production engineers	6.5	3.2	7.4	1	-4.2	-0.57
30. Chemical and physical science technicians	6.3	2.6	9.6	1	-7.0	-0.73
31. Environmental engineers	6.2	2.7	10.3	1	-7.6	-0.74
32. Product and garment designers	5.7	1.8	10.9	1	-9.1	-0.84
33. Electrical engineering technicians	5.4	2.0	5.7	1	-3.7	-0.65
34. Computer network professionals	5.1	1.1	6.0	1	-4.9	-0.81
35. Managers of small enterprises in manufacturing	5.0	0.6	5.7	1	-5.1	-0.89
36. Other specialist managers not elsewhere classified	4.4	1.7	5.1	1	-3.4	-0.67
37. Mechanical engineering technicians	4.3	1.8	4.5	1	-2.6	-0.59
38. Policy and planning managers	4.2	2.0	4.9	1	-2.9	-0.60
39. Managers of small enterprises of business services enterprises	3.9	0.8	5.0	1	-4.2	-0.84
40. Information and communications technology service managers	3.9	0.3	4.4	1	-4.1	-0.93
41. Assistants and part-time lecturers (university)	3.8	1.4	5.8	1	-4.3	-0.75
Column average	11.53	6.34	13.88	0.90	-7.55	-0.53

Notes. This table describes the inventor intensity of occupations that fall in the top 5% of the inventor-intensity distribution. The means are calculated using the years 2004–2018 and workers aged 25–65. See Supplementary Appendix C for calculation details and Supplementary Appendix Table A.1 for occupation sizes and gender compositions.

Table 2: Gender Patenting Gap

	(1)	(2)	(3)	(4)
<i>Panel A. Outcome = patent filing dummy</i>				
Woman	-0.1333*** (0.0020)	-0.0671*** (0.0023)	-0.0734*** (0.0026)	-0.0646*** (0.0024)
% of Woman in Column (1) explained		50%	45%	52%
Constant term	0.1513*** (0.0019)	0.1183*** (0.0017)	0.1214*** (0.0018)	0.1082*** (0.0020)
Adj. R^2	0.0005	0.0202	0.0466	0.0203
<i>Panel B. Outcome = later-granted patent filing dummy</i>				
Woman	-0.0827*** (0.0015)	-0.0405*** (0.0017)	-0.0455*** (0.0019)	-0.0393*** (0.0017)
% of Woman in Column (1) explained		51%	45%	53%
Constant term	0.0930*** (0.0014)	0.0719*** (0.0013)	0.0744*** (0.0013)	0.0655*** (0.0015)
Adj. R^2	0.0003	0.0129	0.0310	0.0130
Baseline Controls	-	Yes	Yes	Yes
Education level $\times$ field FE	-	Yes	Yes	Yes
Occupation FE	-	Yes	Yes	Yes
Firm FE	-	-	Yes	-
High school grade FE	-	-	-	Yes

Notes. This table provides the estimated gender gap coefficients $\hat{\gamma}$ from regression specifications described in Equation (1), estimated with OLS. The coefficients and standard errors are multiplied by 100 and rounded to four decimal places. The dependent variable in Panel A is a patent filing indicator. In Panel B, the outcome is a patent grant indicator, defined as 1 in the year the later-granted patent is filed and 0 otherwise. The estimation sample is a panel of individuals observed between 2004 and 2018, aged 25–65. The baseline controls include indicator variables for calendar year, age, native language, and employment. Educational fields are classified at the 3-digit level (ISCED-2011), and educational levels at the 2-digit level. Occupations are categorized at the most detailed 4-digit or 5-digit level. Standard errors are clustered at the individual level. The number of observations in each regression is 30,275,290. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3: Threefold Oaxaca–Blinder Decomposition (Female Reference)

<i>Outcome:</i>	(1.1) <i>Patent filing</i>	(1.2)	(2.1) <i>Granted patent filing</i>	(2.2)
	Coefficient	SE	Coefficient	SE
<i>Panel A. Decomposition</i>				
Male Mean	0.1513***	(0.0019)	0.0930***	(0.0014)
Female Mean	0.0180***	(0.0006)	0.0102***	(0.0004)
Raw gap (R)	0.1333***	(0.0020)	0.0827***	(0.0015)
Endowments (Q)	0.0333***	(0.0023)	0.0200***	(0.0016)
<i>Contributions</i>				
Baseline controls	0.0005***	(0.0001)	0.0003***	(0.0001)
Education level $\times$ field FEs	0.0023	(0.0021)	0.0022	(0.0015)
Occupation FEs	0.0305***	(0.0021)	0.0175***	(0.0016)
Coefficients (C)	0.0222***	(0.0017)	0.0131***	(0.0012)
Interaction (I)	0.0778***	(0.0031)	0.0496***	(0.0022)
<i>Contributions</i>				
Baseline controls	0.0025***	(0.0002)	0.0014***	(0.0001)
Education level $\times$ field FEs	0.0152***	(0.0036)	0.0095***	(0.0025)
Occupation FEs	0.0601***	(0.0035)	0.0387***	(0.0025)
<i>Panel B. Relative contributions</i>				
<i>% of the raw gap</i>				
Endowments	25.0		24.2	
Coefficients	16.7		15.8	
Interaction	58.3		60.0	
<i>Relative contributions to total endowments (%)</i>				
Baseline controls	1.6		1.5	
Education level $\times$ field FEs	6.9		11.0	
Occupation FEs	91.5		87.5	
<i>Relative contributions to total interaction (%)</i>				
Baseline controls	3.2		2.8	
Education level $\times$ field FEs	19.5		19.2	
Occupation FEs	77.3		78.0	

Notes. This table displays the estimated threefold Oaxaca–Blinder decomposition as in Equation (3). All coefficients and standard errors are multiplied by 100 and rounded to four decimal places. The decomposition is implemented using the *oaxaca* package in Stata (Jann, 2008). The estimation sample is a panel of individuals observed between 2004 and 2018, aged 25–65. In Columns 1.1 and 1.2, the dependent variable is a patent filing indicator. In Columns 2.1 and 2.2, the outcome is a patent grant indicator, defined as 1 in the year the later-granted patent is filed and 0 otherwise. The baseline controls include indicator variables for calendar year, age, native language, and employment. Educational fields are classified at the 3-digit level (ISCED-2011), and educational levels at the 2-digit level. Occupations are categorized at the most detailed 4-digit or 5-digit level. Standard errors are clustered at the individual level. Number of male observations is 15,198,437; Number of female observations is 15,076,853. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

ONLINE APPENDIX

Why Are There So Few Female Inventors?

Atte Pudas[†]

August 21, 2026

Contents of Online Appendix

A School System in Finland	2
B Linking Inventors to Register Data	3
C Occupational Inventor-Intensity	6
D Oaxaca–Blinder Decomposition	7
D.1 Setup	7
D.2 Twofold Oaxaca–Blinder Decomposition	8
D.3 Threefold Oaxaca–Blinder Decomposition	11
E Gender, Patent Grant Rates, and Quality Measures	12
F Parenthood and Patenting: Descriptive Regression Analysis	16
G Additional Figures	21
H Additional Tables	40

[†]Aalto University & Helsinki GSE; Email: atte.pudas@aalto.fi

A School System in Finland

Education is fundamentally free of charge in Finland. The school system starts with compulsory basic education at age 7. After nine grades, individuals can apply to high school or vocational education. High school studies culminate in the matriculation examination, which is a standardized nationwide set of separate subject-specific final exams. The final examinations are graded on a 7-point scale, with the lowest grade indicating a failing score.¹ The grades are based on students' performance relative to all other test takers nationwide (e.g., the top 5% receive the highest grade). The final exams include subjects such as mother tongue and literature, mathematics (basic or advanced level), foreign languages, humanities such as history, and natural sciences such as physics or chemistry. Only the final exam for mother tongue and literature has always been mandatory. Otherwise, the students can choose, with some restrictions, which exams to take. To be eligible for a final exam, a student must have completed all compulsory courses in the subject. Additionally, performing well on the exam usually necessitates knowledge gained from elective courses. Consequently, the final exams students choose to take are closely aligned with the courses they took, and their performance on these exams should strongly reflect their overall learning throughout high school.

After completing secondary education, individuals can apply to universities and universities of applied sciences, which I will collectively refer to as 'college' in this paper. During the years studied in this paper, most applicants to specific degree programs were required to pass a program-specific entrance exam. Applicants were then ranked based on their performance in this exam and prior academic achievements, typically their matriculation exam grades in program-related subjects. Entry into many degree programs, including many STEM fields, is competitive. Thus, students aiming for an engineering degree, for example, benefit from excelling in subjects like advanced mathematics, physics, and chemistry in high school.

¹The grades are (from lowest to highest) as follows: I, A, B, C, M, E, and L. Originally, the grading system included only four grades, I, A, C and L. Grades B and M were introduced in 1970, and E in 1996.

B Linking Inventors to Register Data

I obtain data on Finnish patent inventors and applicants from the OECD REGPAT Database (Maraut et al., 2008). The database fully derives the European Patent Office’s (EPO) Worldwide Statistical Patent Database (PATSTAT Global).

Inventors were linked to Finnish register data by Statistics Finland (SF) using the information available in the patent applications: inventors’ full names, corresponding addresses (including postal codes, street names, and numbers), patent applicant firm names and addresses, and the year of the patent application. The primary challenges in linking inventors arise from two factors: (i) common names, which correspond to a large number of individuals in the register data, and (ii) significant variations in the format of names and addresses between patent applications and the register data. As a result, effective linking requires cleaning the name and address data and interpreting correct matches based on information about addresses and employers, in addition to the inventor names.

I combine the results of two matching efforts to create a dataset of Finnish inventors in EPO patent applications filed between 1978 and 2019, linked to register data. Each linked inventor is assigned a unique encrypted ID number, which connects them to all individual-level datasets of SF.

The first matching was conducted in 2013, in which inventors from patent applications filed between 1978 and 2012 were linked to the register data by a civil servant at SF. The resulting matched inventor data has been used in several published papers (Aghion et al., 2017, 2018, 2022, 2023). This process used the aforementioned patent information along with corresponding non-encrypted name and address data from SF’s databases. Essentially, the civil servant cleaned the raw names and addresses in the patent data and used them to exact matching. Various combinations of matching variables were employed, with the matching year also sometimes differing from the patent application year (e.g., matching an inventor in a 2005 patent with the firm name and address data from 2004 or 2006). The list of all variable combinations and the corresponding matching rates can be provided upon request.

The second round of matching was carried out in 2024 for patent applications filed between 1978 and 2019. The matching was performed by an employee of SF using a matching algorithm developed for this purpose (Berge, 2023). Comprehensive documentation and all corresponding R-code for the algorithm are publicly available in GitHub [lrberge/linking-winst](https://github.com/lrberge/linking-winst). The algorithm utilizes the same information from the inventor and register data as the first matching, but uses more advanced name cleaning and Bayesian match quality assessment algorithms. The algorithm is designed to be transparent and stringent in determining which matches are considered accurate.

I combine the results from the two matching efforts for the overlapping period 1978–2012. The combined matching successfully identifies a match in the register data for 93% of the inventor-patent observations in the overlapping period. The two rounds of matching appear to be complementary. For 63% of the matched inventor-patent observations, both matching rounds identified the same individual in the register data; 21% were linked solely by the old matching, and 14% exclusively by the new matching. Importantly, only 1.3% of the matches are such that the two matching rounds link different individuals in the register data to the same inventor-patent observation (in which case at least one of the matches must be a false positive). In these cases, I retain the result from the more recent matching.

Several factors likely account for the differences between the results of the first and second matching rounds. First, for the last application cohorts in the first matching round, some patent applications had not yet been published in 2013 and were therefore not included in the patent dataset, meaning they were only observed in the second matching round.² Second, during the initial matching round, the civil servant developed several combinations of matching variables, some of which were less stringent than those applied by the more recent algorithm. The civil servant also utilized his knowledge of Finnish naming conventions, including the fact that many streets have both Finnish and Swedish names. Third, the thorough name data cleaning in the more recent matching algorithm may have enabled

²After the expiry of 18 months from the date of filing or the earliest priority date, the EPO applications are published as soon as possible (EPO, 2026).

the identification of matches that were previously missed. Fourth, the register data may have been supplemented and refined retrospectively. Given these considerations, and in the absence of evidence indicating a substantial presence of false positive matches, I combine the results from both matching rounds.

In total, for EPO patent applications filed between 1978 and 2019, the combined matching successfully identifies a person in the register data for 90% (79,446) of all inventor-patent observations with a reported correspondence address in Finland, corresponding to 22,557 unique matched Finnish inventors. The match rates may be slightly downward-biased because some inventors may list a correspondence address in Finland (e.g., a company address) without actually residing there permanently, in which case they do not appear in Finnish register data. At the patent application level, all inventors were matched in roughly 80% of the patent applications; in 14% of the patents, some, but not all, inventors were matched; and in the remaining 6% of patent applications, no inventor matches were found.

Directly evaluating match rates by gender is not possible, as gender is not recorded in patent data and is only observed after successful linkage to register data. In general, we are not aware of any reason to expect the name- and address-based matching to perform differently by gender. However, one potential source of differential match quality is surname changes following marriage, which are substantially more common among women in Finland than among men. That said, any resulting bias is likely to be small. For name changes after marriage to affect the matching, a specific sequence of events must occur: an inventor would need to file a patent application before marriage in the same year, report their pre-marriage surname in the application, and then appear under a different surname in the register data in the same year. We expect such cases to be rare in practice.

If name changes were causing a substantial number of female inventors to go unmatched, we would expect to observe a missing mass of patents and inventors in women's marriage years relative to men in the linked data. However, we do not observe such a pattern. The patents filed in the year of marriage constitute approximately 2.4% of all men's patent filings

between 1987 and 2019. For female inventors, the corresponding share is very similar, 2.3%. When calculated based on the number of inventors who filed a patent each year, the marriage-year shares are 2.3% for men and 2.6% for women. Thus, filing a patent and getting married in the same year is rare, and we observe no notable gender differences in its frequency in the linked data. In conclusion, we find no evidence of systematic bias in linking female inventors to the register data.

C Occupational Inventor-Intensity

I define the inventor intensity of occupations as follows. For each year and occupation, I first count the number of individuals listed as inventors on at least one patent application filed in the given year. So, importantly, I do not calculate the annual number of persons in a given occupation who have *ever* patented, but the number of persons who file a patent while having the given occupation. Therefore, the same inventor can be counted multiple times for multiple occupations if she patents in several years and occupations. Then, I calculate the mean annual number of these inventors for each occupation over the years 2004–2018, the years for which the occupation and patenting data are available annually. To calculate the inventor intensity, I then divide the mean number of inventors by the corresponding mean number of all workers in the given occupation.

Essentially, I distinguish occupations at the finest 4-digit level (ISCO-compatible) or national 5-digit level when available. More precisely, I distinguish occupations by their name (at the finest occupation level) in the data. I use the occupation names as they are more stable in the data across years than occupation codes — the occupation classification was updated in 2010, and as a result, some occupations received new codes (while the occupation name did not necessarily change), and some occupations were combined into bigger units or divided into smaller units with new codes. However, also the occupation names of effectively the same occupations may have changed. Therefore, the number of years each occupation name appears in the data varies.

If I wanted to follow individuals' occupations over time, I would have to use a crosswalk between the two occupation classifications. However, for the purposes of the analyses in this paper, it is not a threat if the effectively same occupation appears twice with two different names. Also, even with a crosswalk between the two classifications, it is not clear how to allocate workers to, for example, two occupations in the old (new) classification that are combined into a single occupation in the new (old) classification.

D Oaxaca–Blinder Decomposition

D.1 Setup

The Oaxaca–Blinder decomposition is a counterfactual decomposition method widely used to analyze mean differences between groups (dating back to [Blinder, 1973](#); [Oaxaca, 1973](#)). Following [Jann \(2008\)](#), consider two groups, M (men) and W (women). Let Y denote the outcome (e.g., a patent filing indicator), X the vector of covariates including a constant, and ε the error term. Define a linear model:

$$Y_g = X_g' \beta_g + \varepsilon_g, \quad g \in \{M, W\}. \quad (\text{A.1})$$

Assume $\mathbb{E}[\varepsilon_g] = 0$ and $\mathbb{E}[\beta_g] = \beta_g$. Then, taking expectations gives

$$\mathbb{E}(Y_g) = \mathbb{E}(X_g)' \beta_g. \quad (\text{A.2})$$

Thus, the mean gender difference in the outcomes can be expressed as

$$R = \mathbb{E}(Y_M) - \mathbb{E}(Y_W) = \mathbb{E}(X_M)' \beta_M - \mathbb{E}(X_W)' \beta_W. \quad (\text{A.3})$$

D.2 Twofold Oaxaca–Blinder Decomposition

Introducing a reference coefficient vector β^* , we can add and subtract the terms $\mathbb{E}(X_M)'\beta^*$ and $\mathbb{E}(X_W)'\beta^*$ to Eq. (A.3) and reorganize:

$$\begin{aligned} R &= \mathbb{E}(X_M)'\beta_M - \mathbb{E}(X_W)'\beta_W \\ &= \mathbb{E}(X_M)'\beta_M - \mathbb{E}(X_W)'\beta_W + \mathbb{E}(X_M)'\beta^* - \mathbb{E}(X_M)'\beta^* + \mathbb{E}(X_W)'\beta^* - \mathbb{E}(X_W)'\beta^* \quad (\text{A.4}) \\ &= (\mathbb{E}(X_M) - \mathbb{E}(X_W))'\beta^* + \mathbb{E}(X_M)'(\beta_M - \beta^*) + \mathbb{E}(X_W)'(\beta^* - \beta_W) \end{aligned}$$

Replacing the expected value of the covariates $\mathbb{E}(X_g)$ by the sample average $\bar{X}_g$, and denoting the estimated coefficient vectors by $\hat{\beta}$, the decomposition can be estimated as

$$R = \underbrace{(\bar{X}_M - \bar{X}_W)'\hat{\beta}^*}_{\equiv E \text{ ("explained")}} + \underbrace{\bar{X}_M'(\hat{\beta}_M - \hat{\beta}^*) + \bar{X}_W'(\hat{\beta}^* - \hat{\beta}_W)}_{\equiv U \text{ ("unexplained")}}. \quad (\text{A.5})$$

This expression represents the twofold Oaxaca–Blinder decomposition, where the first component E captures the part of the gender gap explained by differences in covariate means evaluated at the reference coefficient vector β^* , and the second component U captures the unexplained component due to gender differences in coefficients (and potential effects of unobserved variables).

The choice of the reference coefficient structure affects the decomposition results. For example, using male coefficients as the reference structure ($\beta^* \equiv \beta_M$) yields different results than using female coefficients ($\beta^* \equiv \beta_W$). To define a single meaningful reference structure, I adopt a “pooled” approach, in which β^* is estimated with a regression that pools men and women and includes a gender indicator variable:

$$Y = \beta_0 + X'\beta^* + \gamma\mathbf{1}\{g = W\} + \varepsilon. \quad (\text{A.6})$$

Note that this specification is similar to the gender gap regression in Eq. (1). The underlying assumption of the pooled approach is that the inclusion of the gender indicator

sufficiently removes systematic gender-specific effects from the estimated coefficients, so that β^* represents a coefficient structure that is applicable to both men and women. Compared with using male or female coefficients as the counterfactual, the pooled reference structure arguably provides a more balanced assessment of the portion of the gender gap attributable to gender differences in covariates. The use of pooled regression coefficients as a “nondiscriminatory” reference has been advocated by, e.g., [Neumark \(1988\)](#) and [Fortin \(2008\)](#). See also [Jann \(2008\)](#) and [Fortin et al. \(2011\)](#) for further discussions.

Note that the coefficient γ from the pooled regression in Eq. (A.6) directly captures the unexplained part of the pooled twofold Oaxaca-Blinder decomposition. To see this, use Eq. (A.6) to express the mean gender gap:

$$\begin{aligned}
R &= \mathbb{E}(Y_M) - \mathbb{E}(Y_W) \\
&= \mathbb{E}(\beta_0 + X'_M \beta^* + \varepsilon) - \mathbb{E}(\beta_0 + X'_W \beta^* + \gamma + \varepsilon) \\
&= \beta_0 + \mathbb{E}(X_M)' \beta^* - \beta_0 - \mathbb{E}(X_W)' \beta^* - \gamma \\
&= (\mathbb{E}(X_M) - \mathbb{E}(X_W))' \beta^* - \gamma,
\end{aligned} \tag{A.7}$$

which can be estimated as:

$$(\bar{X}_M - \bar{X}_W)' \hat{\beta}^* - \hat{\gamma}. \tag{A.8}$$

Now $-\hat{\gamma}$ corresponds to the unexplained component U in Eq. (A.5). To directly see this, compare the gender gap estimate of -0.0671 in Table 2, Panel A, Column 2, to the total unexplained estimate of 0.0671 in the pooled twofold decomposition presented in Table A.5, Column 1.1.³ The gender gap regression reported in Table 2, Panel A, Column 2, and the underlying regressions of the decomposition in Table A.5, Columns 1.1 and 1.2, are based on the same sample and include an identical set of covariates. In both cases, the patent-filing indicator is explained by baseline controls (year, age, language, and main activity fixed effects) and education-level-by-field and occupation fixed effects. As a result, the gender gap

³The sign of the regression coefficient depends on whether the gender indicator is defined to take the value of 1 for women or for men.

regression and the pooled twofold decomposition yield equivalent estimates.

Table A.5, Columns 1.1 and 1.2, show that the raw patenting rate is 0.1333 percentage points higher for men, and roughly half (0.0662 percentage points) of this gap can be explained by differences in observable characteristics (E), while the remaining half (0.0671 percentage points) is unexplained (U). The explained component is driven primarily by differences in occupational sorting, which alone accounts for 0.0554 percentage points (or 82.2%) of it. Differences in education contribute a more modest share (0.0102 percentage points, or 15.4%), while baseline controls play only a minor role. Using the later-granted patent filing dummy as the outcome, the decomposition results reported in Table A.5, Columns 2.1 and 2.2, yield conclusions that closely align with those for the overall patent filing gender gap discussed above.

The explained component in the twofold decomposition can be divided into contributions from different covariate groups, as the total explained component is a sum of the individual contributions:

$$E = (\bar{X}_M - \bar{X}_W)' \hat{\beta}^* = (\bar{X}_{1M} - \bar{X}_{1W}) \hat{\beta}_1^* + (\bar{X}_{2M} - \bar{X}_{2W}) \hat{\beta}_2^* + \dots, \quad (\text{A.9})$$

where $\bar{X}_{g1}, \bar{X}_{g2}, \dots$ denote the mean values of the individual covariates for gender g , and $\hat{\beta}_{g1}, \hat{\beta}_{g2}, \dots$ are the corresponding estimated coefficients. Each term in the sum represents the contribution of the gender difference in a specific covariate: the first term captures the differences in $\bar{X}_1$, the second term captures differences in $\bar{X}_2$, and so on, and these terms can be grouped. The unexplained part can be similarly decomposed. However, unlike the explained component, the decomposition of the unexplained part into contributions from different coefficient groups is only meaningfully interpretable for variables with a natural zero point (Jann, 2008). Thus, the detailed decomposition of the unexplained part is not reported in Table A.5.

D.3 Threefold Oaxaca–Blinder Decomposition

To derive the threefold Oaxaca-Blinder decomposition, consider the mean gender gap R in Eq. (A.3). Adding and subtracting the term $\mathbb{E}(X_M)'\beta_W$ and reorganizing yields:

$$\begin{aligned} R &= \mathbb{E}(X_M)'\beta_M - \mathbb{E}(X_W)'\beta_W \\ &= \mathbb{E}(X_M)'\beta_M - \mathbb{E}(X_W)'\beta_W + \mathbb{E}(X_M)'\beta_W - \mathbb{E}(X_M)'\beta_W \\ &= (\mathbb{E}(X_M) - \mathbb{E}(X_W))'\beta_W + \mathbb{E}(X_M)'(\beta_M - \beta_W) \end{aligned} \quad (\text{A.10})$$

Further adding and subtracting $\mathbb{E}(X_W)'(\beta_M - \beta_W)$ and reorganizing yields:

$$\begin{aligned} R &= (\mathbb{E}(X_M) - \mathbb{E}(X_W))'\beta_W + \mathbb{E}(X_M)'(\beta_M - \beta_W) + \mathbb{E}(X_W)'(\beta_M - \beta_W) - \mathbb{E}(X_W)'(\beta_M - \beta_W) \\ &= (\mathbb{E}(X_M) - \mathbb{E}(X_W))'\beta_W + \mathbb{E}(X_W)'(\beta_M - \beta_W) + (\mathbb{E}(X_M) - \mathbb{E}(X_W))'(\beta_M - \beta_W). \end{aligned} \quad (\text{A.11})$$

Replacing population expectations with sample averages $\bar{X}_g$ and denoting the estimated coefficient vectors with $\hat{\beta}_g$, the decomposition can be estimated as

$$R = \underbrace{(\bar{X}_M - \bar{X}_W)'\hat{\beta}_W}_{\equiv Q \text{ ("endowments")}} + \underbrace{\bar{X}_W'(\hat{\beta}_M - \hat{\beta}_W)}_{\equiv C \text{ ("coefficients")}} + \underbrace{(\bar{X}_M - \bar{X}_W)'(\hat{\beta}_M - \hat{\beta}_W)}_{\equiv I \text{ ("interaction")}}. \quad (\text{A.12})$$

This is the threefold Oaxaca-Blinder decomposition, using women as the reference group.⁴

Here, Q measures the portion of the gender gap explained by differences in observed characteristics between men and women, evaluated at women's coefficient values. In other words, the expected change in women's mean outcome if women had men's predictor levels, keeping their coefficients constant. Component C captures the part of the gender gap attributable to differences in coefficients, that is, the expected change in women's mean outcome if women had men's coefficients, keeping women's predictor levels fixed. I represents the interaction term, accounting for the simultaneous presence of differences in both characteristics and coefficients.

⁴The decomposition can also be reversed to reflect men's viewpoint:
 $R = (\bar{X}_M - \bar{X}_W)'\hat{\beta}_M + \bar{X}_M'(\hat{\beta}_M - \hat{\beta}_W) + (\bar{X}_M - \bar{X}_W)'(\hat{\beta}_M - \hat{\beta}_W)$

E Gender, Patent Grant Rates, and Quality Measures

Measuring patent quality and value is a notoriously challenging task. I examine gender differences in patent quality across multiple measures proposed in the existing literature, using inventor–patent-level data on all European Patent Office (EPO) applications filed by Finnish inventors between 1987 and 2019. I begin the analysis by studying patent grant rates, using the applications’ grant status as of 2023. In addition, I investigate several complementary measures intended to reflect the multiple facets of patent quality and value.

Since inventions are typically developed in teams — about 69% of patent applications in the estimation sample list multiple inventors — the same patent may be linked to multiple inventors. The median inventor team size in the data is 2, and the mean is 2.57.⁵ In addition, some inventors contribute to many patents: while the median inventor files just one application over the sample period, the mean is 3.55. To account for the resulting dependence structure, I use two-way clustered standard errors at the patent application and inventor levels in the regression model presented below.

As the baseline measure, I use the number of received family-to-family citations, defined as the number of distinct DOCDB simple patent families that, by 2023, cite at least one publication or application belonging to the DOCDB simple family of the focal application. A patent family refers to the collection of jurisdictions in which protection for the same invention has been sought through related patent applications. Thus, family-to-family citations effectively capture invention-to-invention citations. Citations reflect the invention’s impact on subsequent innovation and also serve as a proxy for the patent’s economic value (e.g., [Trajtenberg, 1990](#); [Harhoff et al., 1999](#); [Hall et al., 2005](#)). In the regression analysis, I control for filing year, technology sector (5 categories), and the inventor’s 3-digit employer industry to account for business cycles, time since filing, and sectoral differences in citation

⁵Team size counts all inventors, including foreign inventors not linked to the register data. When restricted to inventors linked to Finnish register data (which determines the size of patent clusters in the regression model), multiple inventors appear in 65% of patent applications, the median number of inventors is 2, and the mean is 2.1.

accumulation and grant rates.

In addition, I use patent family size, measured as the size of the simple DOCDB family of the focal patent application. Patent family size reflects the geographical scope of the sought patent protection for the invention. Larger patent families are associated with higher (potential) economic value of the patent (e.g., [Harhoff et al., 2003](#)).

Furthermore, I use several patent-quality indicators provided by the OECD, constructed from data covering the full universe of EPO patent documents. The data on quality indicators is available online for researchers ([OECD, 2026](#)). See [Squicciarini et al. \(2013\)](#) for detailed definitions and discussions on these indicators, as well as related references.

First, I use an indicator for a "breakthrough" invention that equals 1 if the patent belongs to the top 1% of the forward-citation (received citations) distribution within its filing-year–technology-field cohort.⁶ Only citations up to 5 years after patent publication are considered.

Second, I examine patent scope, a measure of technological breadth proxied by the number of distinct 4-digit International Patent Classification (IPC) subclasses assigned to the invention. A greater number of distinct 4-digit IPC classes indicates a broader technological scope and, consequently, higher potential technological and economic value of the patent ([Lerner, 1994](#)).

Third, I employ an originality index, which measures the breadth of the technology fields on which the focal patent relies, as indicated by the IPC classes in the backward citations (prior patents the focal patent cites). The index is defined as follows. For focal patent p , let s_{pj} be the share of its backward citations that fall into 4-digit IPC class j . The index is given by $\text{Originality}_p = 1 - \sum_{j=1}^{n_p} s_{pj}^2$, where n_p is the total number of IPC classes represented among the patents cited by p . The index approaches 0 when all citations lie in a single class and increases toward 1 as citations are spread more evenly across many classes. Higher originality values indicate that the focal patent draws on knowledge from a more diverse set of technology fields, which may indicate a more original invention ([Trajtenberg et al., 1997](#)).

⁶There are 35 technology fields in the EPO's patent classification.

Fourth, I use a radicalness index that captures the extent to which a patent draws on technologies outside its own IPC classes, and thus the extent to which the invention departs from existing inventions (Shane, 2001). Radicalness is strongly positively correlated with patent originality as defined above (Squicciarini et al., 2013). Radicalness is defined as the share of distinct 4-digit IPC classes appearing in the focal patent’s backward citations that are not among the IPC classes assigned to the focal patent itself; that is, the number of distinct cited IPC classes outside the focal patent’s own classes divided by the total number of distinct IPC classes in its backward citations. The index takes values between 0 and 1. An important caveat, however, is that the measure is backward-looking: it reflects the technological breadth of the prior art on which the patent builds, but does not indicate whether the focal patent is itself distinctive relative to other contemporaneous inventions.

Finally, I integrate two composite indicators into the analysis, introduced in Squicciarini et al. (2013) and building on Lanjouw and Schankerman (2004): Patent Quality Index 4 and 6. These composite measures are experimental and should be interpreted with caution, but they are useful in summarizing the gender differences in patent quality. Also importantly, these indices are constructed only for granted patents. The indices are linear combinations of the respective components, with each component receiving equal weight. All components of the indices are normalized within the filing year and technological field. The Quality Index 4 combines forward citations (within five years from publication), patent family size, number of patent claims, and the generality index.⁷ The Quality Index 6 extends this set by additionally including the number of backward citations and the grant lag index, which captures the duration of the examination process. A higher index value indicates a higher composite quality of the patent application.

Using these patent quality indicators, I estimate the following gender gap regression specification:

⁷Patent claims define the legal scope of protection, as enforceable rights extend only to what is specified in them. The number and content of claims, therefore, determine the breadth of patent protection.

$$y_{ip} = \alpha + \gamma Woman_i + \lambda_t + \mathbf{X}'_p \sigma + \mathbf{X}'_{ip} \beta + \varepsilon_{ip}, \quad (\text{A.13})$$

where y_{ip} is the outcome variable for inventor i and patent application j . I denote the constant term with α , $Woman_i$ is a female indicator, $\mathbf{X}'_p$ includes application-level controls, $\mathbf{X}'_{ip}$ includes inventor-level controls, and ε_{ip} is the error term. The application-level controls include filing-year fixed effects, broad technology-sector indicators (5 categories), and inventor-team size indicators (binning teams of sizes 6–9 and ≥ 10).⁸ The inventor-level controls (measured in the patent’s filing year) include age fixed effects, 2-digit education level by 3-digit field-of-education fixed effects, and 3-digit employer industry fixed effects. I estimate the model with the OLS. For forward citation count and patent family size, I also present results from a corresponding Poisson regression model.

Table A.6 reports the results across the outcomes specified above. The table reports the gender gap estimate $\hat{\gamma}$ (also as a percentage of the male inventors’ raw outcome mean) and the constant term. Column (1) shows raw mean gender differences, Column (2) adds patent-level controls, and Column (3) further adds inventor-level controls. Thus, the coefficients for the constant term in Column (1) always show the raw outcome mean for the male inventors.

The results in Panels 1–4 indicate that female inventors have systematically lower patent grant probabilities and receive fewer citations than male inventors. The grant rate is 8.8% lower for female inventors, and about half of this difference persists after adding the control variables. Applications by female inventors receive 5.2% fewer citations on average than those by male inventors, but the raw difference is not statistically significant. However, conditional on the patent- and inventor-level controls, the estimated gap more than doubles and becomes strongly significant. This indicates that, compared with male inventors with similar potential experience and educational background in the same industry, female inventors’ patents filed in the same year receive fewer citations on average. The lower citations are also

⁸The five broad technology sectors are: (i) electrical engineering, (ii) instruments (includes e.g., measurement and analytics technologies); (iii) chemistry, (iv) mechanical engineering, and (v) other fields (EPO, 2025).

reflected in the differences in the probability of filing a “breakthrough” patent. Notably, the unconditional probability of a breakthrough patent is roughly 54% lower for female inventors compared to male inventors. Furthermore, conditioning on the covariates only widens the gap.

However, for the other proxies of patent quality in Panels 5–10, the gender gap is reversed: female inventors’ patents have, on average, larger patent family size, broader technological scope, and are more original and radical. The raw means are 7% to 24% higher for female inventors, and while controlling for patent application characteristics reduces these differences, they generally remain significantly positive. However, once inventor-level controls are included, the gender differences effectively disappear. Finally, in Panels 11 and 12, the composite quality indices indicate differences of less than 2% between female and male inventors’ granted patents once the control variables are included in the regression.

In conclusion, the results on gender differences in patent quality are mixed. On the one hand, in grant rates and citations, female inventors have poorer outcomes. On the other hand, across other quality measures, the average female inventor’s patent performs better than the average patent by male inventors, although the differences disappear once application- and inventor-level control variables are added to the regression specification. Interpreting differences in grant rates and citations is further complicated by prior evidence suggesting gender differences in the patent examination process that are not directly related to underlying patent quality ([Jensen et al., 2018](#); [Gavrilova-Zoutman and Juranek, 2025](#); [Aneja et al., 2024](#); [Pairolero et al., 2025](#)), as well as gender biases in patent citation patterns ([Hochberg et al., Forthcoming](#); [Li and Liu, Forthcoming](#)).

F Parenthood and Patenting: Descriptive Regression Analysis

I describe the association between parenthood and patenting activity for women and men using a simple regression. I use panel data on the full population of individuals born in 1950 or later, observed between ages 25 and 65 in the years 2004–2018, to estimate the following

equation:⁹

$$\begin{aligned} Patent_{iy} = & \alpha + \beta_0 Woman_i + \beta_1 HasChidlren_{iy} + \beta_2 HasChidlren_{iy} \times Woman_i \\ & + \lambda_y + \mathbf{X}'_{iy} \delta + \varepsilon_{iy}, \end{aligned} \quad (\text{A.14})$$

where *HasChidlren* is a indicator variable for whether individual i in calendar year y has any children. Otherwise, the model is similar to Equation (1). The parameters of interest are β_0 , which reflects the mean gender gap (women – men) in the annual probability of patenting among individuals who do not have children (but may have them later); β_1 , which reflects the association between having children and patenting among men (fathers – childless men); and β_2 , which indicates how this association differs for women. In addition, the following sums are informative: $\beta_1 + \beta_2$ captures the association between having children and patenting among women (mothers – childless women), and $\beta_0 + \beta_2$ captures the gender gap in patenting among parents (mothers – fathers).

The results are presented in Table A.9. The coefficients are scaled by 100 and interpreted as percentage point differences. Column (1) reports the raw gender gap estimate of -0.1030 percentage points, which equals -88% compared to men’s raw mean indicated by the constant term). Column (2) introduces the dummy for having children, interacted with the female dummy, along with the baseline control variables (dummy variables for calendar year, age, native language, and employment). The magnitude of the coefficient $\hat{\beta}_0$ decreases substantially from column (1) to column (2). However, a strong negative association remains, indicating that being a woman is negatively associated with patenting activity among individuals without children. The coefficient $\hat{\beta}_1$ in Column (2) indicates that having children is positively associated with men’s patenting rates, while for women, the sum $\hat{\beta}_1 + \hat{\beta}_2$ indicates a negative association.

In column (3), the inclusion of fixed effects for education, occupation, and high school

⁹The sample is limited to individuals born in 1950 or later because childbirth data are only available from 1970 onward, which introduces increasing measurement error in child count for earlier cohorts. The restriction on years reflects the availability of occupation data.

grades reduces the magnitude of these estimates by roughly half, but the overall patterns remain unchanged. The coefficient $\hat{\beta}_0$ shows that women without children are 0.0315 percentage points less likely to patent than childless men with the same age, education, high school grades, and occupation. The sum $\hat{\beta}_1 + \hat{\beta}_2$ indicates that, compared to women without children, mothers have an additional ‘penalty’ in patenting of 0.0072 percentage points. Furthermore, the coefficient $\hat{\beta}_1$ indicates that fathers patent at a rate 0.0268 percentage points higher than comparable childless men. Driven by the negative association for mothers and the positive association for fathers, the difference between fathers and mothers is -0.0655 percentage points (the sum $\hat{\beta}_0 + \hat{\beta}_2$). The magnitude of the estimated gender patenting gap among childless individuals amounts to roughly 48% of the gap between mothers and fathers ($\hat{\beta}_0$ of -0.0315 versus the sum $\hat{\beta}_0 + \hat{\beta}_2$ of -0.0655).

References for Appendices

- AGHION, P., U. AKCIGIT, A. HYYTINEN, AND O. TOIVANEN (2017): “The Social Origins of Inventors,” Working Paper 24110, National Bureau of Economic Research.
- (2018): “On the Returns to Invention Within Firms: Evidence from Finland,” in *AEA Papers and Proceedings*, vol. 108, 208–212.
- (2022): “A Year Older, A Year Wiser (and Farther from Frontier): Invention Rents and Human Capital Depreciation,” *Review of Economics and Statistics*, 1–31.
- (2023): “Parental Education and Invention: The Finnish Enigma,” *International Economic Review*, 64, 453–490.
- ANEJA, A., O. RESHEF, AND G. SUBRAMANI (2024): “Attrition and the Gender Patenting Gap,” *Review of Economics and Statistics*, 1–31.
- BERGE, L. R. (2023): “An algorithm to match Swedish registry data to European patents data for the WINST-project,” https://github.com/lrberge/linking_winst.
- BLINDER, A. S. (1973): “Wage Discrimination: Reduced Form and Structural Estimates,” *Journal of Human resources*, 436–455.

- EPO (2025): “European patent applications — All technical fields,” <https://www.epo.org/en/about-us/statistics/technology-dashboard-2025/statistics-and-indicators/european-patent-applications/all-technical-fields>, accessed: 2026-05-22.
- (2026): “Guidelines for Examination in the European Patent Office, Part A, Chapter IV, 1. Publication of application,” https://www.epo.org/en/legal/guidelines-epc/2026/a_vi_1.html, accessed on 2026-05-25.
- FORTIN, N., T. LEMIEUX, AND S. FIRPO (2011): “Decomposition Methods in Economics,” in *Handbook of labor economics*, Elsevier, vol. 4, 1–102.
- FORTIN, N. M. (2008): “The Gender Wage Gap Among Young Adults in the United States: The Importance of Money Versus People,” *Journal of human resources*, 43, 884–918.
- GAVRILOVA-ZOUTMAN, E. AND S. JURANEK (2025): “The Role of Reviewer Characteristics on the Diversity of Successful Applicants,” *Organization Science*, 36, 1726–1744.
- HALL, B. H., A. JAFFE, AND M. TRAJTENBERG (2005): “Market Value and Patent Citations,” *The RAND Journal of Economics*, 36, 16–38.
- HARHOFF, D., F. NARIN, F. M. SCHERER, AND K. VOPEL (1999): “Citation Frequency and the Value of Patented Inventions,” *Review of Economics and statistics*, 81, 511–515.
- HARHOFF, D., F. M. SCHERER, AND K. VOPEL (2003): “Citations, Family Size, Opposition and the Value of Patent Rights,” *Research policy*, 32, 1343–1363.
- HOCHBERG, Y. V., A. KAKHBOD, P. LI, AND K. SACHDEVA (Forthcoming): “Are Patents With Female Inventors Under-Cited? Evidence From Text Estimation,” *Journal of Financial Economics*.
- JANN, B. (2008): “The Blinder–Oaxaca Decomposition for Linear Regression Models,” *The stata journal*, 8, 453–479.
- JENSEN, K., B. KOVÁCS, AND O. SORENSON (2018): “Gender Differences In Obtaining And Maintaining Patent Rights,” *Nature biotechnology*, 36, 307–309.
- LANJOUW, J. O. AND M. SCHANKERMAN (2004): “Patent Quality and Research Productivity: Measuring Innovation With Multiple Indicators,” *The Economic Journal*, 114, 441–465.

- LERNER, J. (1994): “The Importance of Patent Scope: An Empirical Analysis,” *The RAND Journal of Economics*, 319–333.
- LI, Z. AND S. LIU (Forthcoming): “Gender Difference in Innovation Recognition: A Textual Analysis Approach,” *Journal of Labor Economics*.
- MARAUT, S., H. DERNIS, C. WEBB, V. SPIEZIA, AND D. GUELLEC (2008): “The OECD REGPAT Database,” OECD Science, Technology and Industry Working Papers 2008/02.
- NEUMARK, D. (1988): “Employers’ Discriminatory Behavior and the Estimation of Wage Discrimination,” *Journal of Human Resources*, 23, 279–295.
- OAXACA, R. (1973): “Male-Female Wage Differentials in Urban Labor Markets,” *International economic review*, 693–709.
- OECD (2026): “Intellectual Property Statistics,” <https://www.oecd.org/en/data/datasets/intellectual-property-statistics.html>, accessed 2026-06-02.
- PAIROLERO, N. A., A. A. TOOLE, P.-A. PAPPAS, C. A. DE GRAZIA, AND M. H. TEODORESCU (2025): “Closing the Gender Gap in Patenting: Evidence From a Randomized Control Trial at the USPTO,” *American Economic Journal: Economic Policy*, 17, 281–310.
- SHANE, S. (2001): “Technological Opportunities and New Firm Creation,” *Management Science*, 47, 205–220.
- SQUICCIARINI, M., H. DERNIS, AND C. CRISCUOLO (2013): “Measuring Patent Quality: Indicators of Technological and Economic Value,” OECD Science, Technology and Industry Working Papers 2013/03, Organisation for Economic Co-operation and Development (OECD), Paris.
- TRAJTENBERG, M. (1990): “A Penny for Your Quotes: Patent Citations and the Value of Innovations,” *The Rand journal of economics*, 172–187.
- TRAJTENBERG, M., R. HENDERSON, AND A. JAFFE (1997): “University Versus Corporate Patents: A Window on the Basicness of Invention,” *Economics of Innovation and New Technology*, 5, 19–50.

G Additional Figures

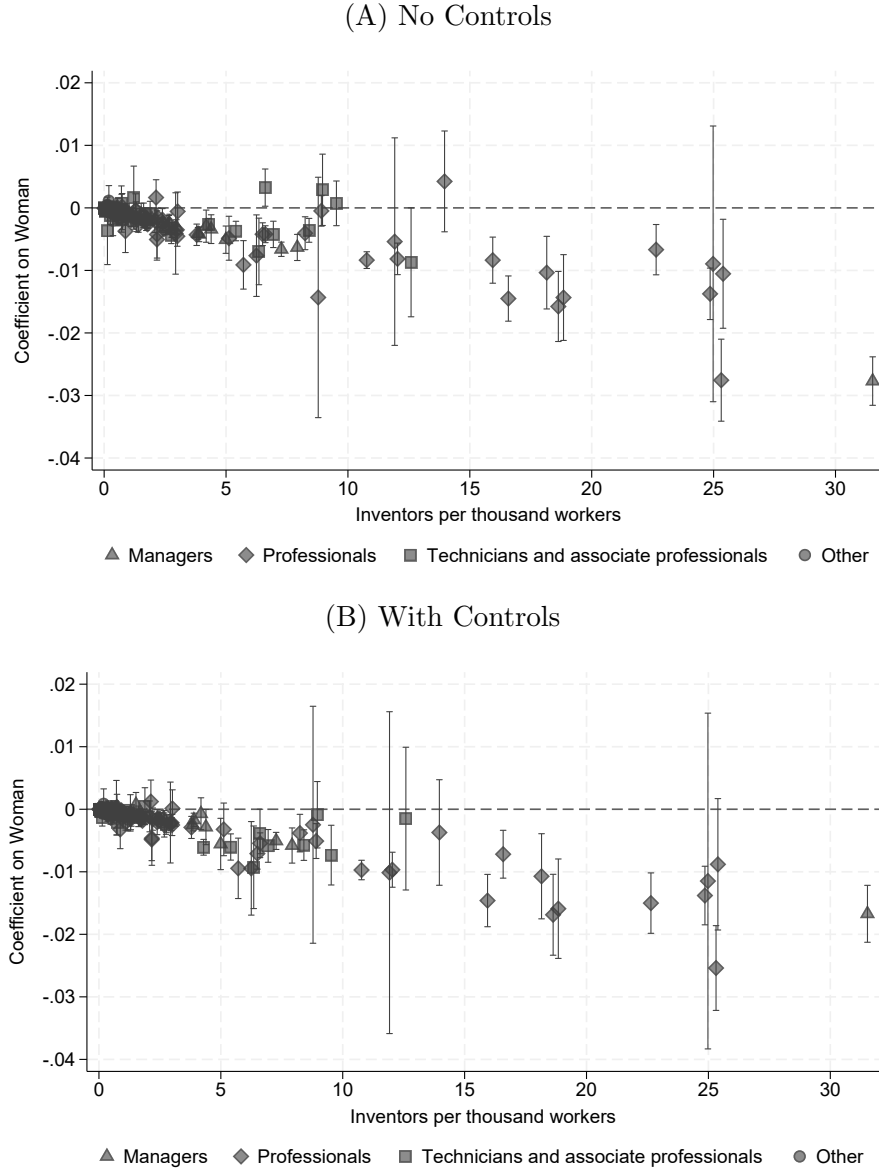


Figure A.1: Gender Patenting Gaps by Occupation, Sorted by Inventor Intensity

Notes. This figure shows the estimated gender patenting gaps and their 95% confidence intervals by occupation, ordering the occupations by the inventor intensity. The coefficients are estimated with a model described in Equation (1), separately for each occupation. Figure A.1A shows the estimates from regressions without control variables. Figure A.1B shows the estimates from regressions that control for indicator variables for year, age, native language, education level by field, and high school matriculation examination grades for mathematics and mother tongue. The estimation sample is a panel of individuals observed during the years 2004–2018 at ages 25–65. I exclude a single occupation with an average workforce size below 30 from the figure, as the estimate’s wide confidence intervals would reduce the figure’s readability. Standard errors are clustered at the individual level.

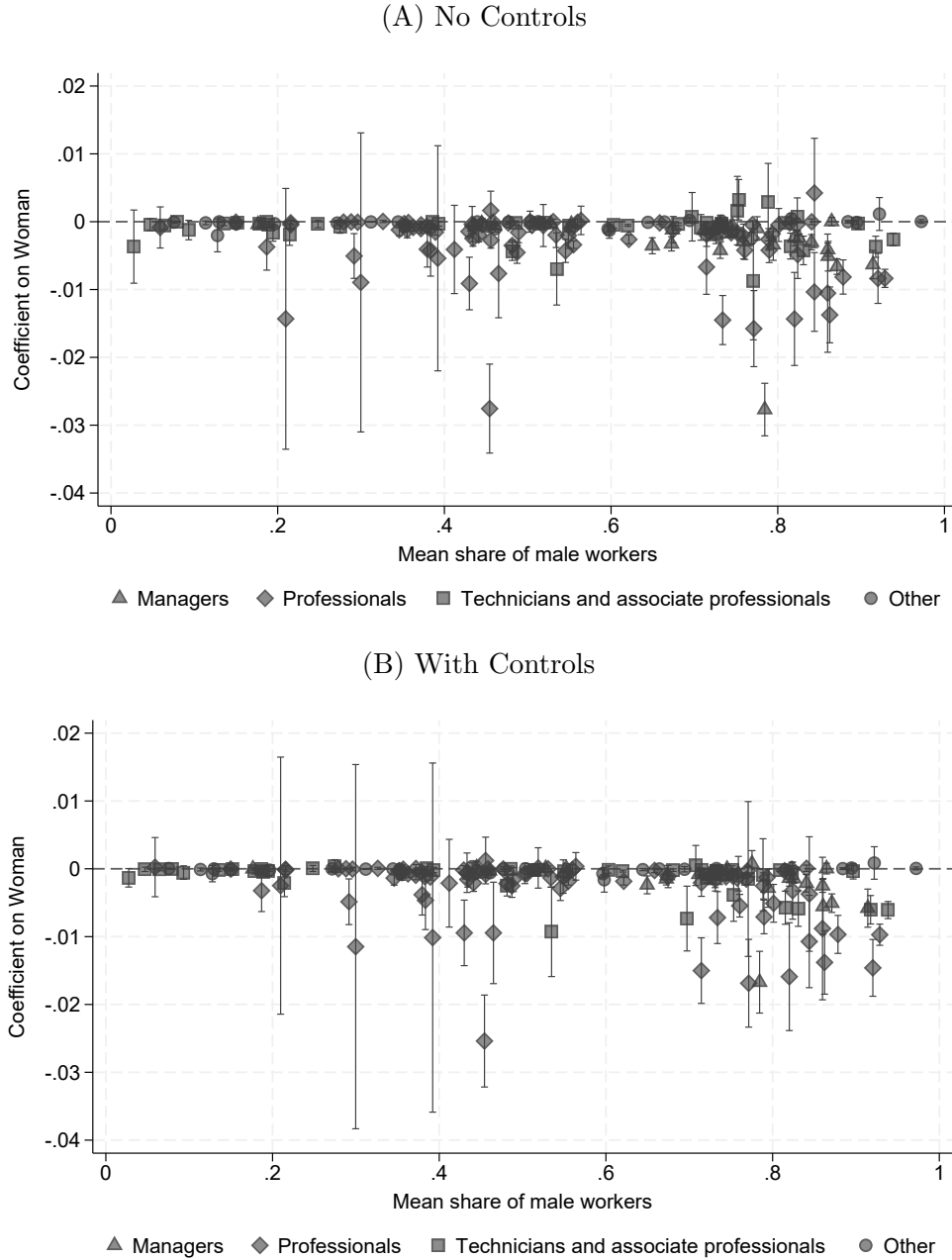


Figure A.2: Gender Patenting Gaps by Occupation, Sorted by Male Share

Notes. This figure shows the estimated gender patenting gaps and their 95% confidence intervals by occupation, ordering the occupations by the average share of male workers. The coefficients are estimated with a model described in Equation (1), separately for each occupation. Figure A.2A shows the estimates from regressions without control variables. Figure A.2B shows the estimates from regressions that control for indicator variables for year, age, native language, education level by field, and high school matriculation examination grades for mathematics and mother tongue. The estimation sample is a panel of individuals observed during the years 2004–2018 at ages 25–65. I exclude a single occupation with an average workforce size below 30 from the figure, as the estimate’s wide confidence intervals would reduce the figure’s readability. Standard errors are clustered at the individual level.

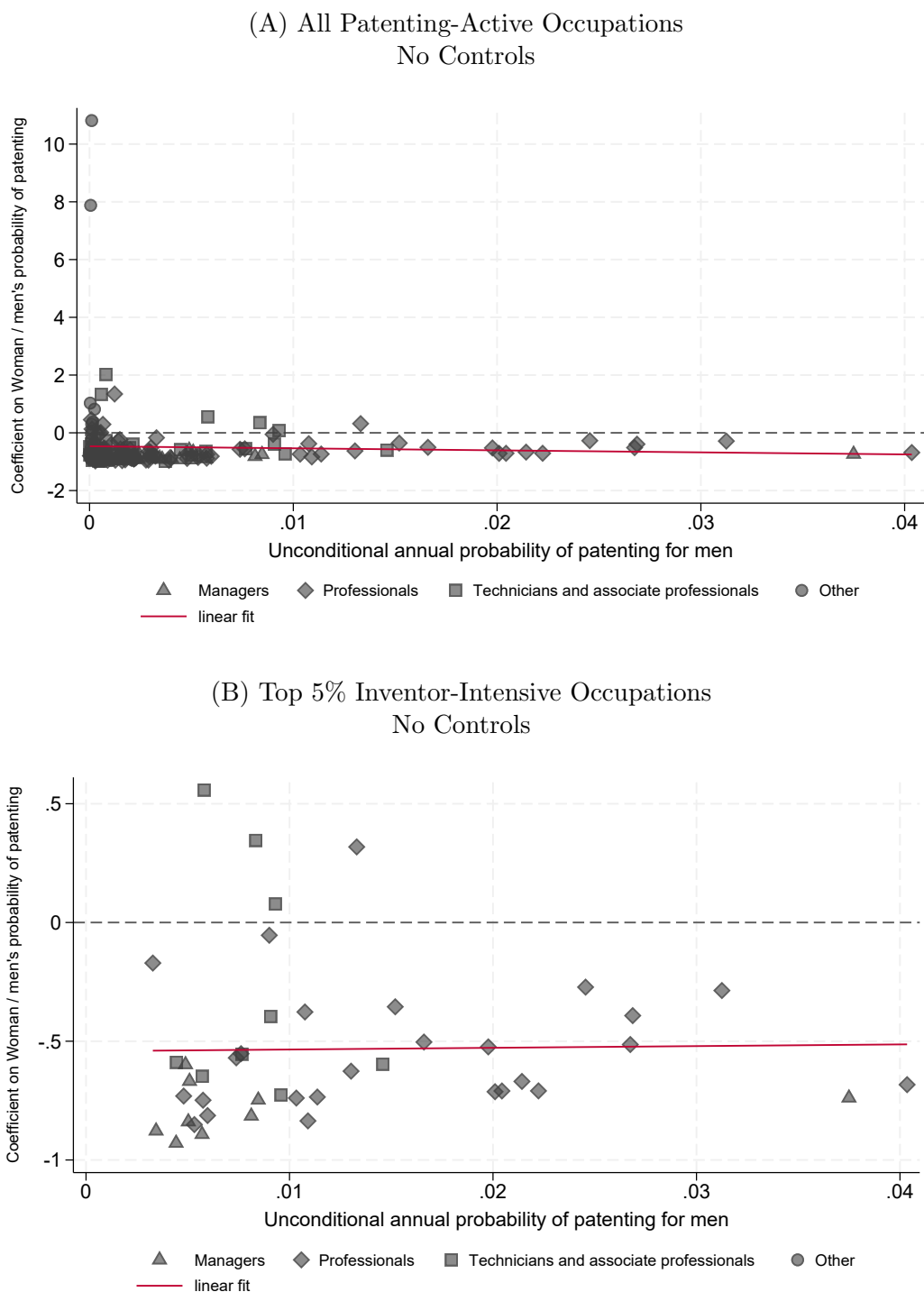


Figure A.3: Gender Patenting Gaps by Occupation, Scaled by Men's Patenting Probability

Notes. This figure displays the by-occupation unconditional gender patenting gaps from Figure 2, Panel A, divided by men's probability of patenting (which is displayed on the x-axis). Figure A.3A displays the scaled estimates for all occupations with both male and female patenting activity, and Figure A.3B zooms into the top 5% most inventor-intensive occupations.

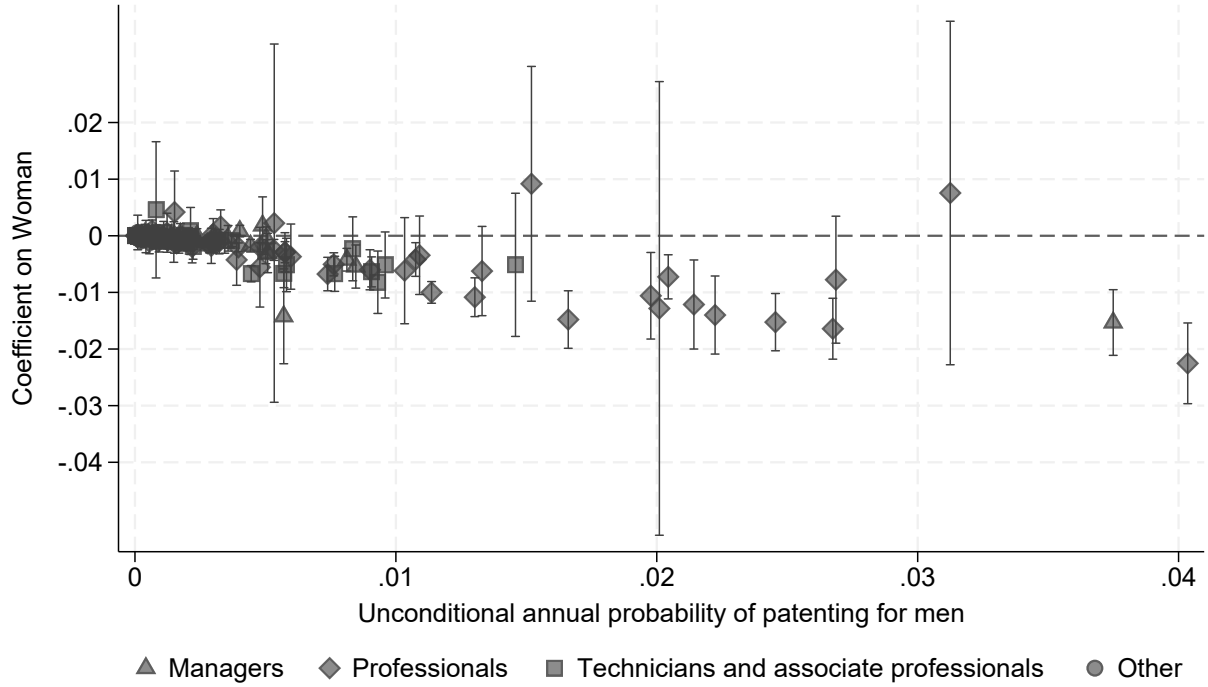
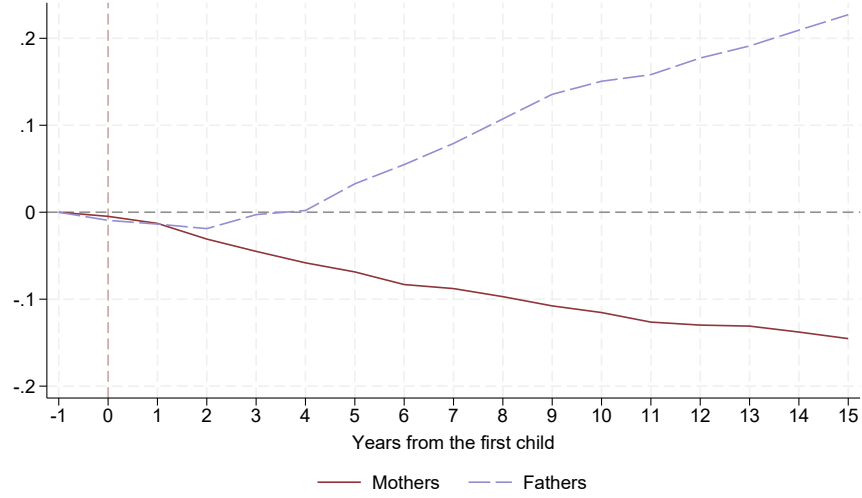


Figure A.4: Gender Patenting Gaps by Occupation, Controlling for Firm FEs

Notes. This figure shows the estimated gender patenting gaps and their 95% confidence intervals by occupation. The coefficients are estimated with a model described in Equation (1), separately for each occupation. The regression specification includes indicator variables for year, age, native language, education level by field, high school matriculation examination grades for mathematics and mother tongue, and firm. The estimation sample is a panel of individuals observed during the years 2004–2018 at ages 25–65. I exclude a single occupation from the figure with an average workforce size below 30, as the estimate’s wide confidence intervals would reduce the readability of the figure. Standard errors are clustered at the individual level.

(A) Cumulative Sum of the Point Estimates in Figure 3, Panel B



(B) Cumulative Sum of the Point Estimates in Figure 3, Panel B
Scaled by Reference Year ($t = -1$) Mean Patenting Rates

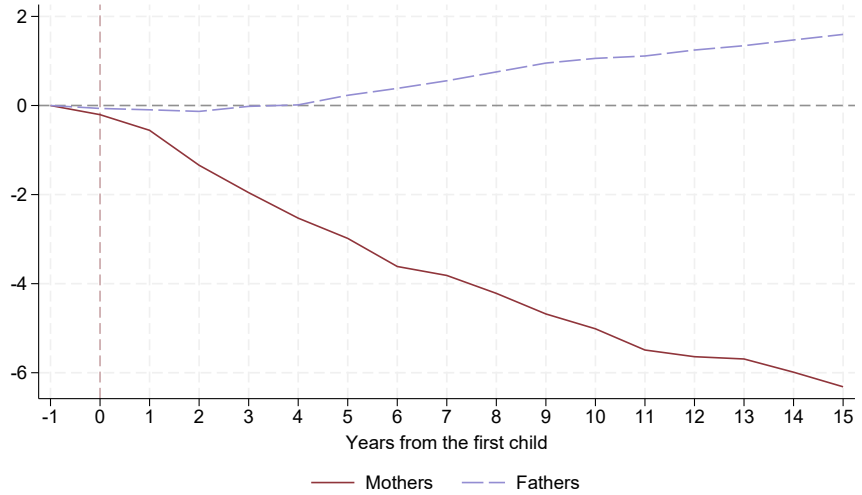


Figure A.5: Cumulative Changes in Patenting Rates Around First Childbirth

Notes. Figure A.5A displays the cumulative sums of the point estimates in Figure 3, Panel B, reported in percentage points, illustrating the cumulative impact of parenthood on fathers' and mothers' patenting rates. Figure A.5B scales the cumulative estimates for mothers and fathers by their patenting rates in the reference period ($t = -1$).

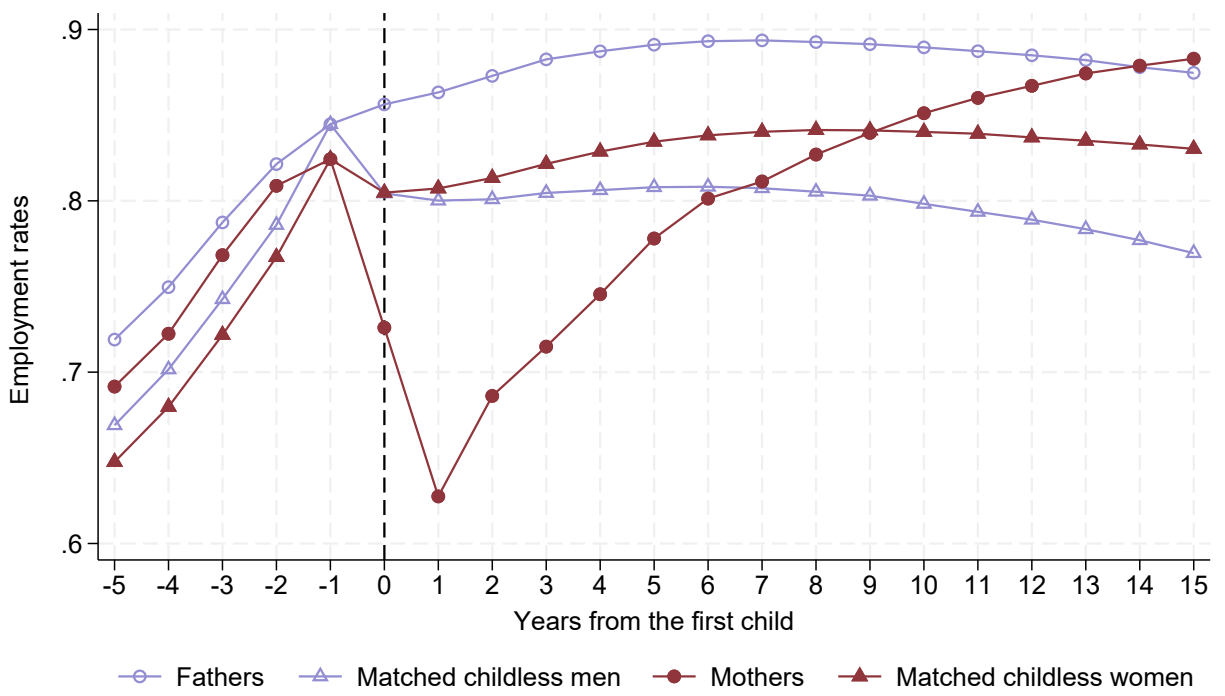


Figure A.6: Raw Employment Rates in the Matched Sample

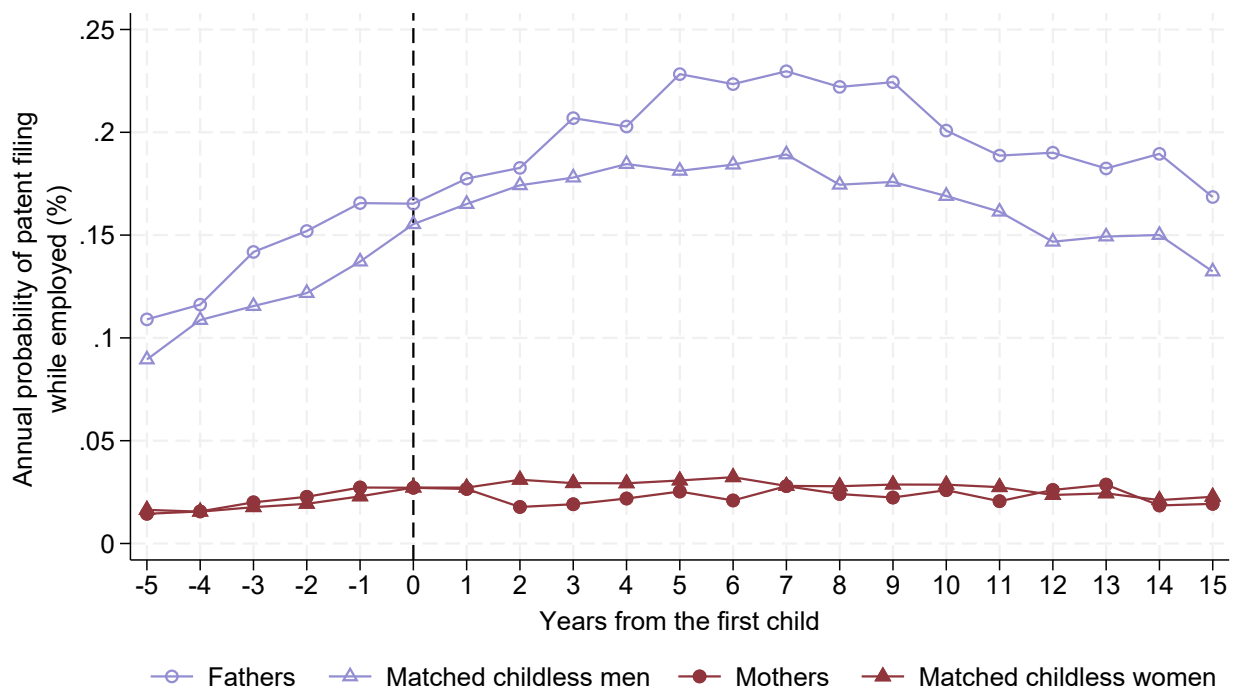


Figure A.7: Raw Patenting Rates Around First Childbirth, Conditional on Employment

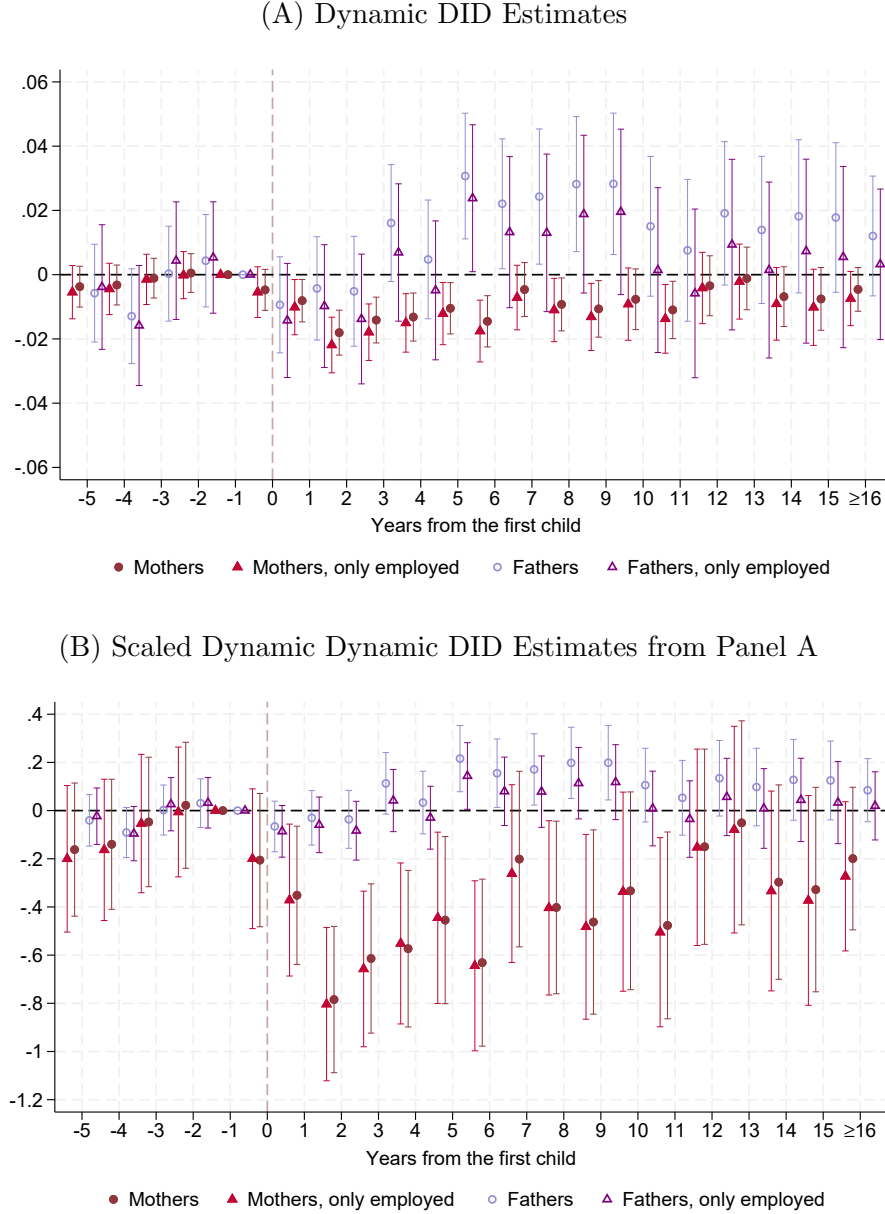


Figure A.8: Patenting Rates Around First Childbirth, All vs. Employed

Notes. Panel A plots point estimates and 95% confidence intervals, estimated separately for women and men using the dynamic DID model in Eq. (4). The figure compares estimates from the unrestricted sample (“mothers” and “fathers” in the legend) with those from a restricted sample that includes only individual-year observations in which the individual is employed (“mothers, only employed” and “fathers, only employed”). For readability, the estimates in Panel A are scaled by 100. Panel B scales the estimates in Panel A by parents’ reference-year ($t = -1$) patenting rates. The estimation sample comprises matched parent–childless pairs from treatment cohorts 1991–2019, born in 1950 or later, observed from 1987 to 2019. The outcome variable is an indicator for filing a patent application in a given year. The model includes calendar-year, individual, and 3-year age-bin fixed effects. Standard errors are clustered at the individual and the matched-pair levels.

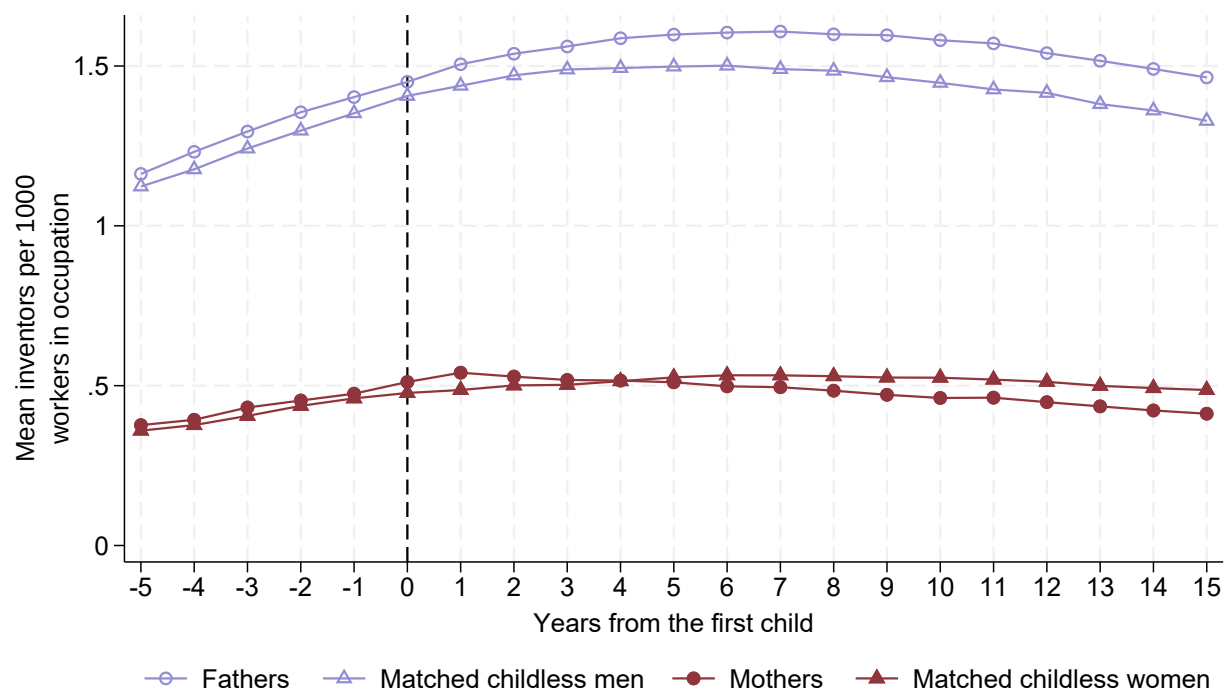


Figure A.9: Mean Occupational Inventiveness: Inventors per 1000 Workers

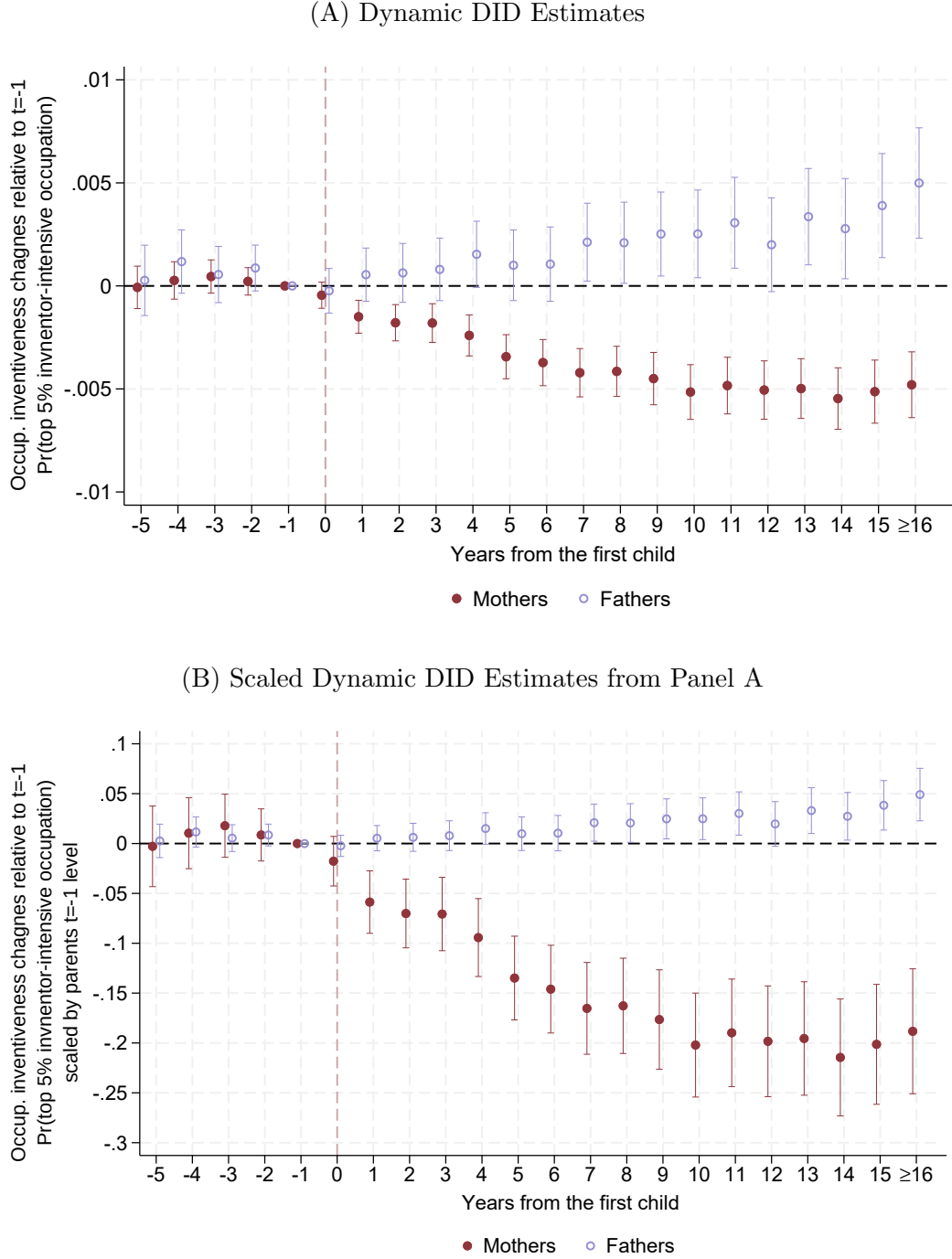


Figure A.10: Changes in Occupational Inventiveness Around First Childbirth:
Pr(Top 5% Inventor-Intensive Occupation)

Notes. The outcome variable is an indicator for working in an occupation that falls within the top 5% in the inventor-intensity distribution (as defined in Section 3). Panel A plots point estimates and 95% confidence intervals, estimated separately for women and men using the dynamic DID model in Eq. (4). Panel B scales the estimates from Panel A by the parents' reference year ($t = -1$) outcome level. The model includes calendar-year, individual, and 3-year age-bin fixed effects. The sample comprises matched parent-childless pairs from cohorts born in 1950 or later, observed from 1987 to 2019. Standard errors are clustered at the individual and the matched-pair levels.

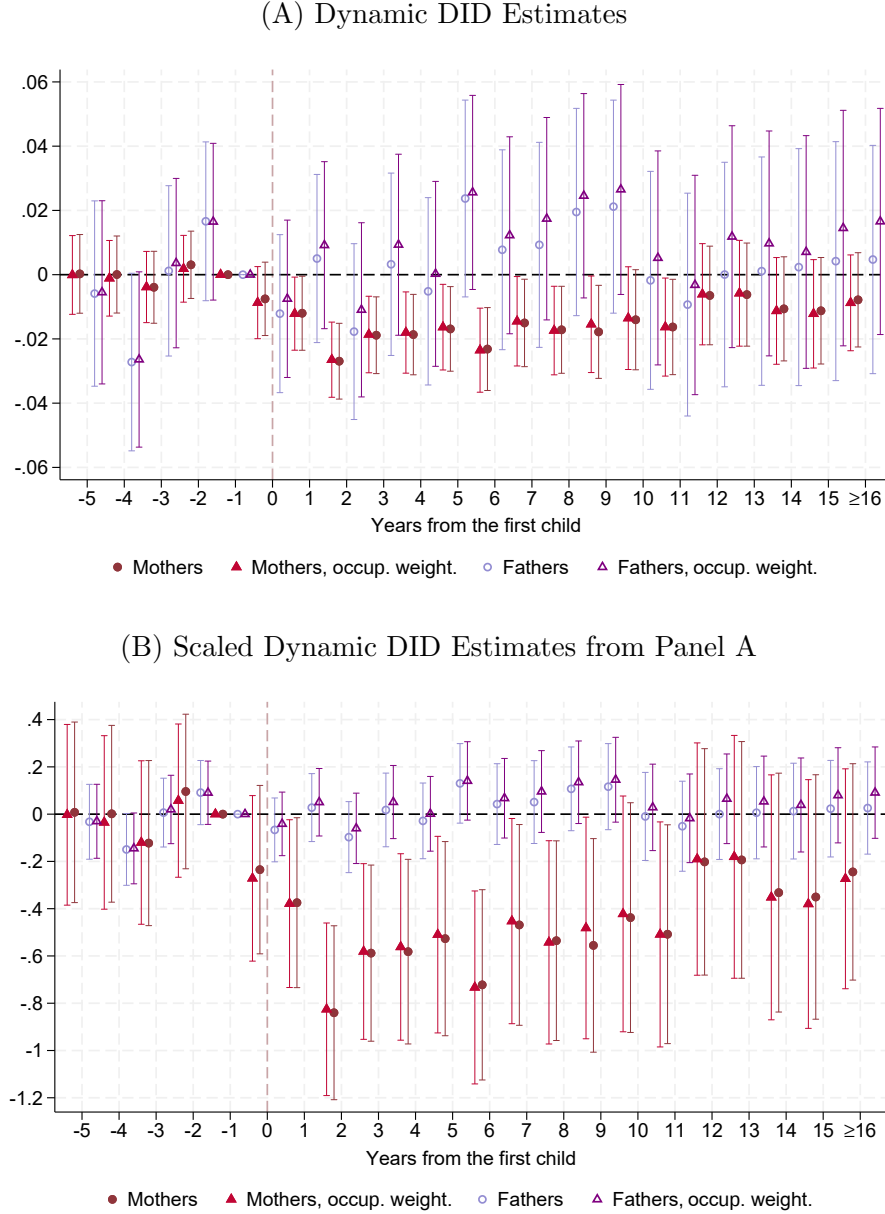


Figure A.11: Patenting Rates Around First Childbirth:
Occupation-Weighted vs. Unweighted

Notes. Panel A plots point estimates and 95% confidence intervals, estimated separately for women and men using the dynamic difference-in-differences model in Eq. (4), and compares results from the non-weighted and weighted regression specifications. For readability, the estimates in Panel A are scaled by 100. Panel B scales the estimates in Panel A by parents' reference-year ($t = -1$) patenting rates. The estimation sample comprises matched parent-childless pairs from cohorts born in 1950 or later, observed in the years 1990, 1993, 1995, 2000, and 2004–2018, when occupation codes are available. The outcome variable is an indicator for filing a patent application in a given year. The regression model includes calendar-year, individual, and 3-year age-bin fixed effects. Standard errors are clustered at the individual and the matched-pair levels.

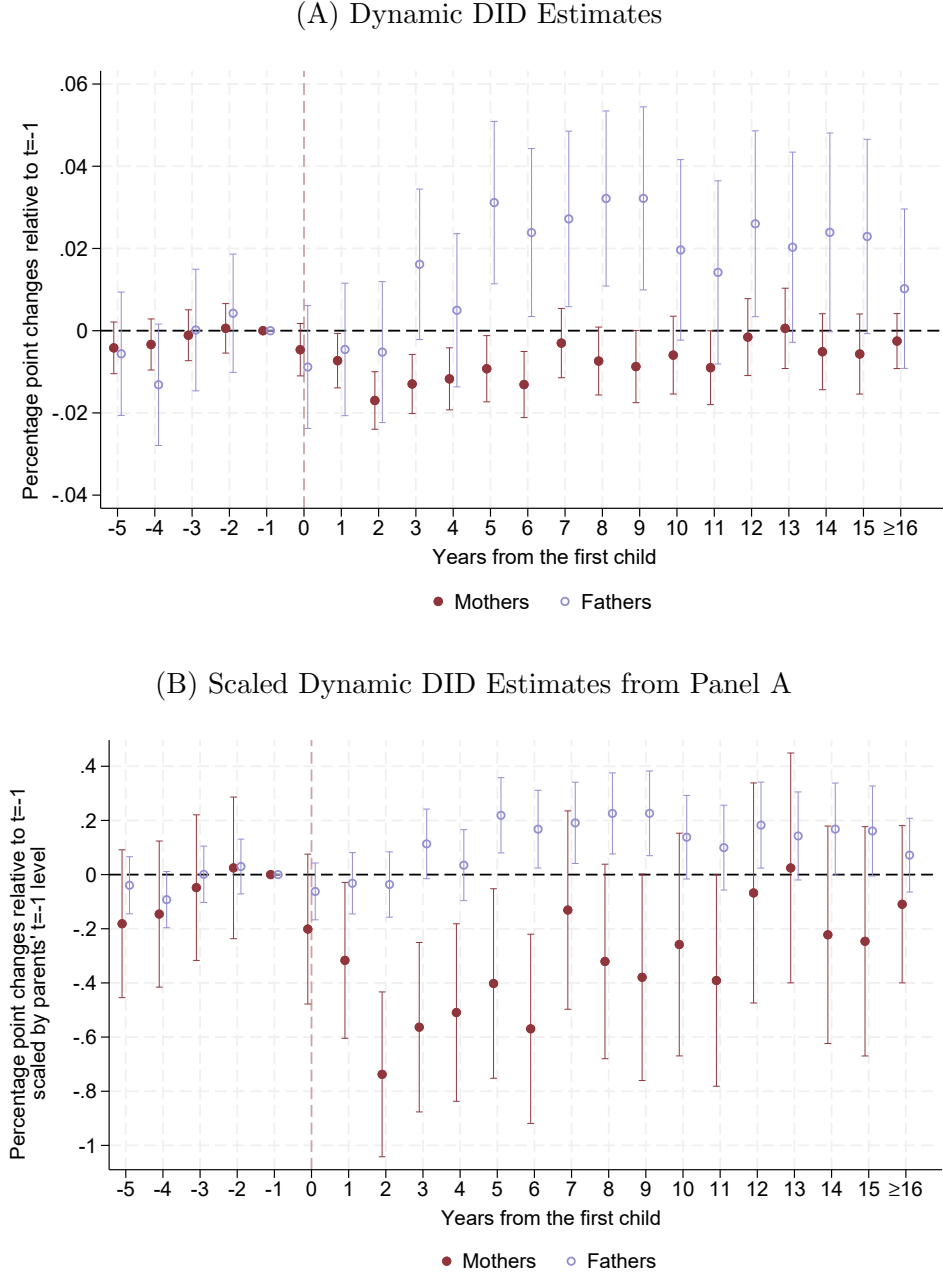


Figure A.12: Patenting Rates Around First Childbirth:
Alternative Regression Specification

Panel A plots point estimates and 95% confidence intervals, estimated separately for women and men using a dynamic DID model including a parent-status dummy, calendar year fixed effects, and age fixed effects. For readability, the estimates in Panel A are scaled by 100. Panel B scales the estimates in Panel A by parents' reference-year ($t = -1$) patenting rates. The outcome variable is an indicator for filing a patent application in a given year. Standard errors are clustered at the individual and the matched-pair levels.

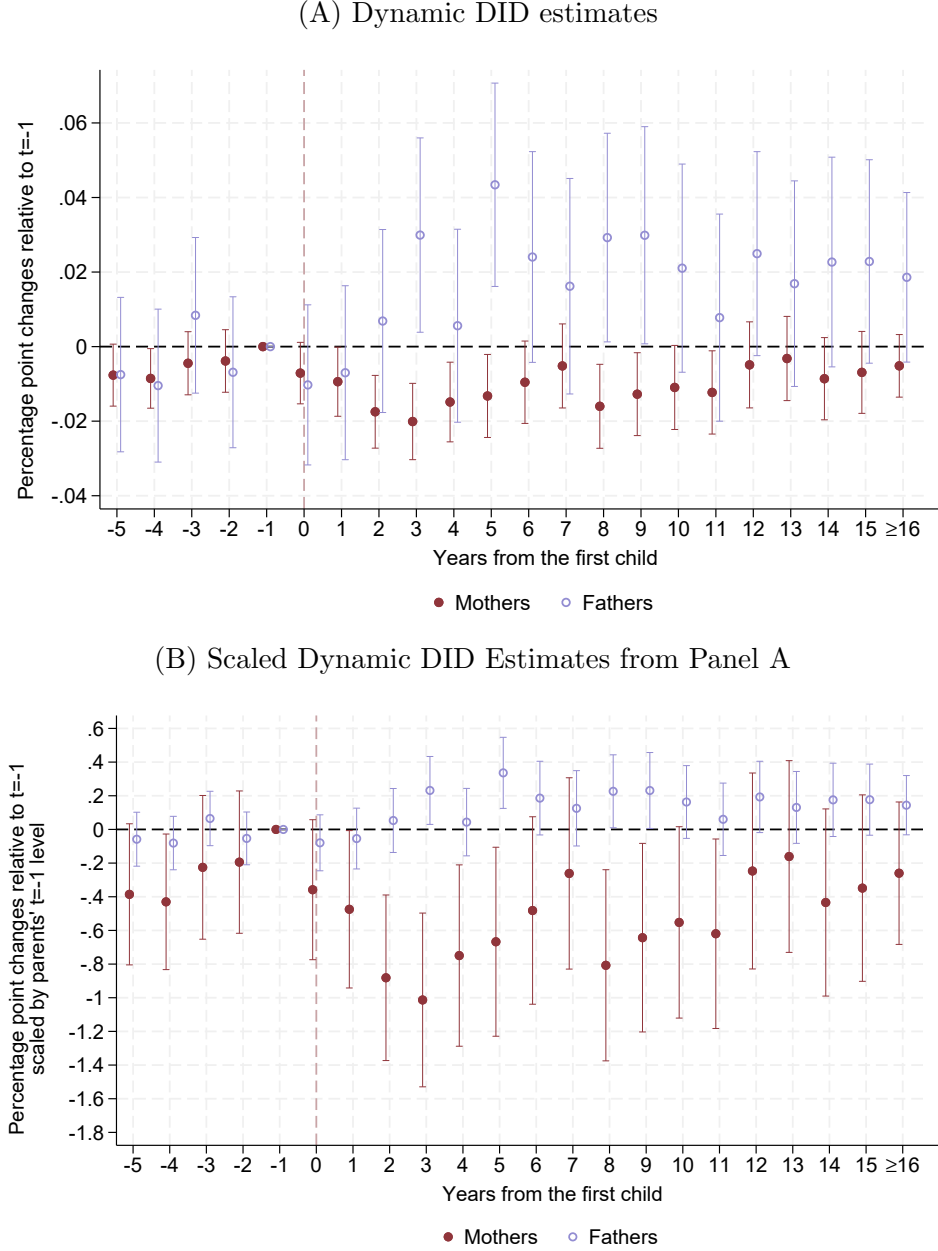


Figure A.13: Changes in Annual Patenting Rates Around First Childbirth:
Sample Trimmed to Treatment Cohorts 1992–2004

Notes. Panel A plots point estimates and 95% confidence intervals, estimated separately for women and men using a dynamic DID model including individual, calendar year, and 3-year age-bin fixed effects. For readability, the estimates in Panel A are scaled by 100. Panel B scales the estimates in Panel A by parents' reference-year ($t = -1$) patenting rates. The outcome variable is an indicator for filing a patent application in a given year. Standard errors are clustered at the individual and the matched-pair levels. The sample comprises matched parent–childless pairs from cohorts born in 1950 or later, observed from 1987 to 2019. The sample is further restricted to parents who had their first child between 1992 and 2004, and thereby can be observed throughout the estimation window $t \in [-5, 15]$. Parents are matched with childless individuals as described in Section 6.

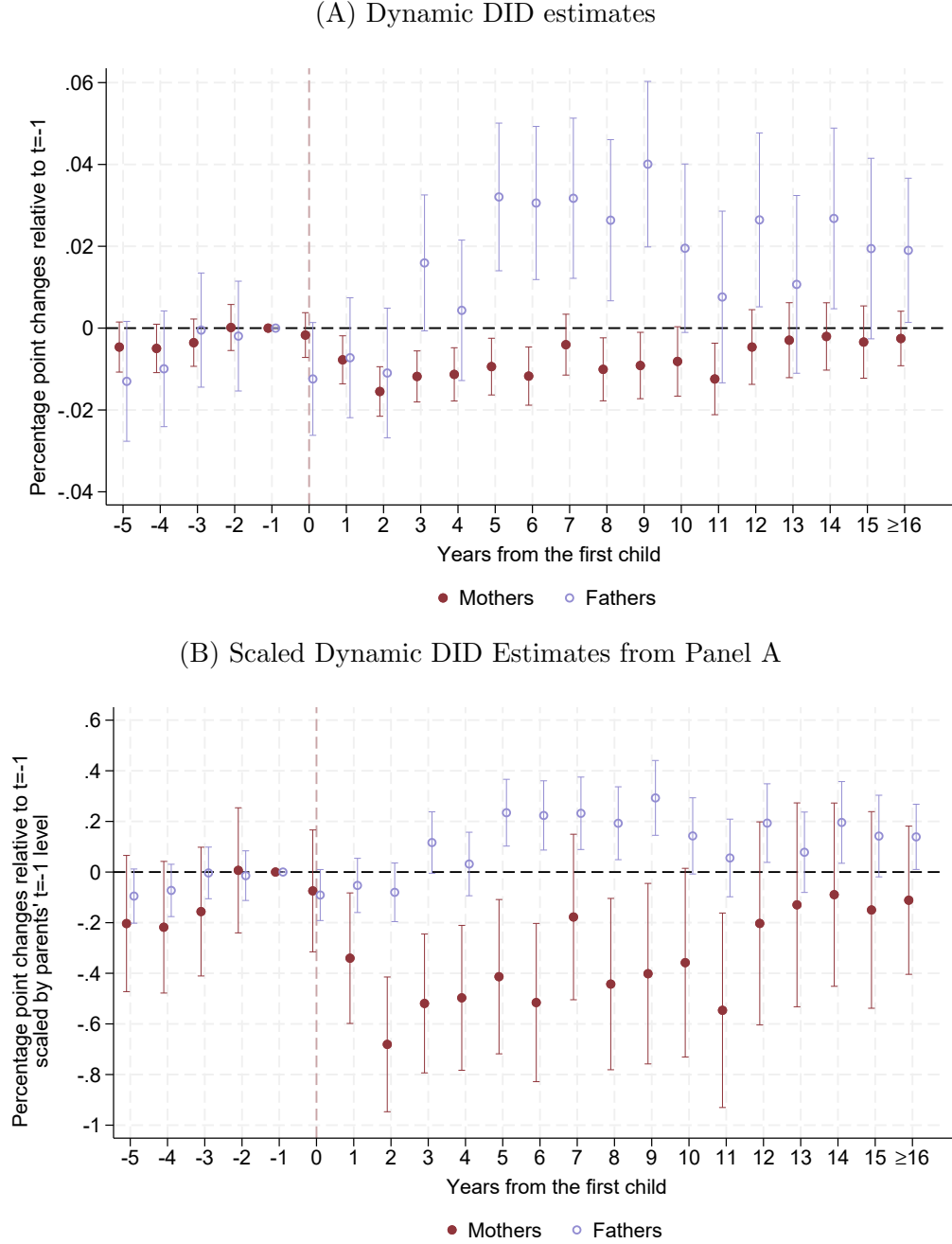


Figure A.14: Changes in Annual Patenting Rates Around First Childbirth:
Relaxed Matching

Notes. Panel A plots point estimates and 95% confidence intervals, estimated separately for women and men using a dynamic DID model specified in Eq. (4). For readability, the estimates in Panel A are scaled by 100. Panel B scales the estimates in Panel A by parents' reference-year ($t = -1$) mean patenting rates. The outcome variable is an indicator for filing a patent application in a given year. Standard errors are clustered at the individual and the matched-pair levels. The sample comprises matched parent-childless pairs from cohorts born in 1950 or later, observed from 1987 to 2019. Parents are matched with childless individuals one year before their first child's birth. The matching is exact by gender, age, mother tongue, education level, and field (at the 1-digit level), and employment status.

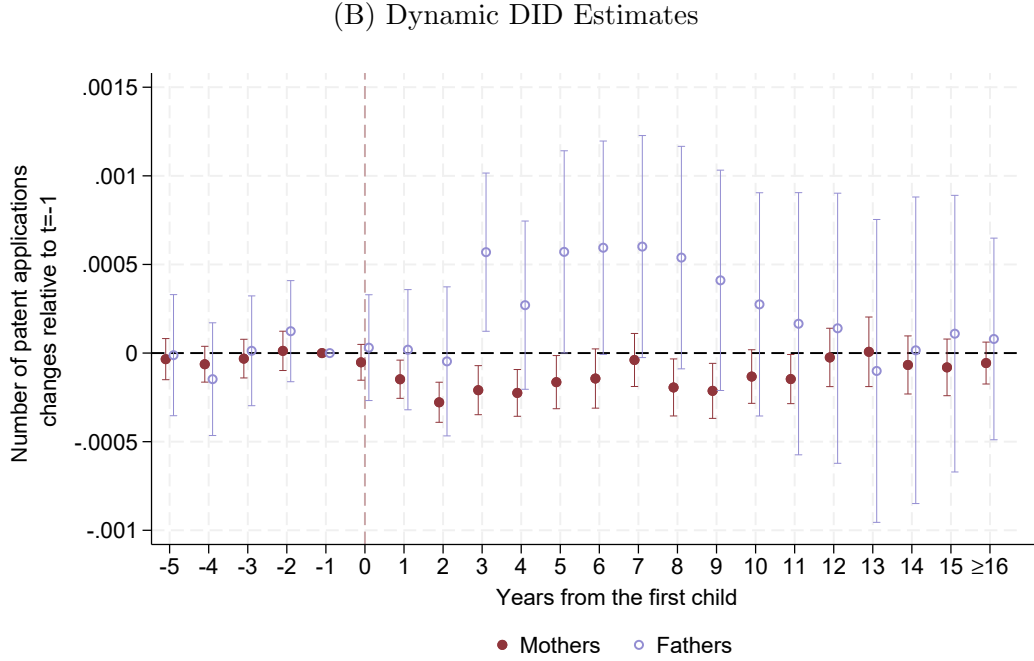
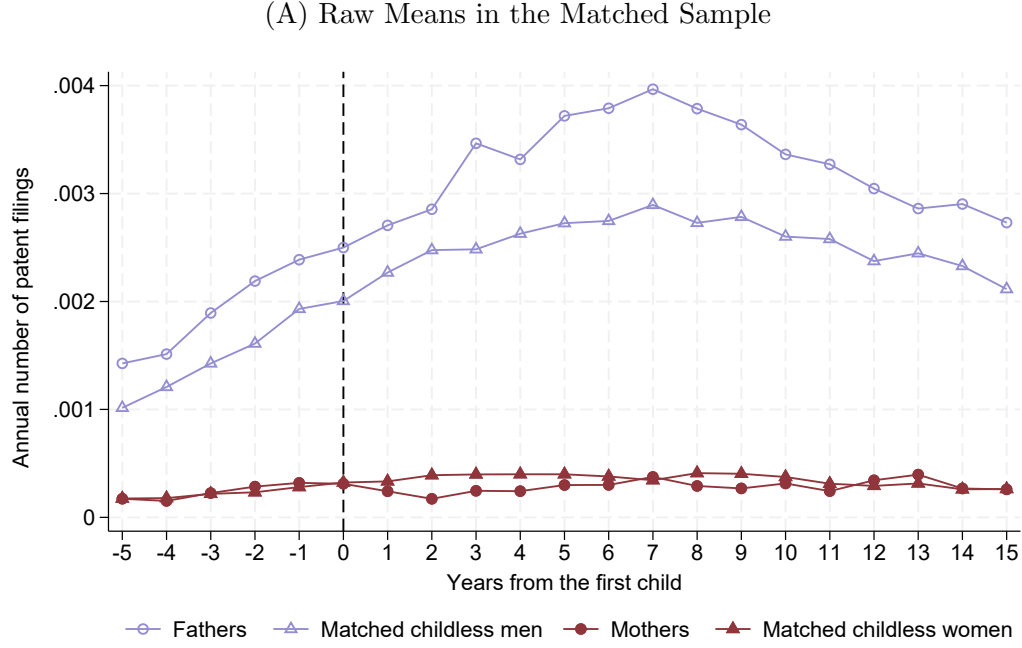


Figure A.15: Changes in the Number of Patent Filings Around First Childbirth

Notes. Panel A displays the raw probabilities of patenting for parents and matched benchmark individuals in the estimation window, expressed as percentages. Panel B plots point estimates and 95% confidence intervals, estimated separately for women and men using the dynamic DID model in Eq. (4). The outcome variable is the number of patent filings in a given year. The model includes calendar-year, individual, and 3-year age-bin fixed effects. The sample comprises matched parent-childless pairs from cohorts born in 1950 or later, observed from 1987 to 2019. Standard errors are clustered at the individual and the matched-pair levels.

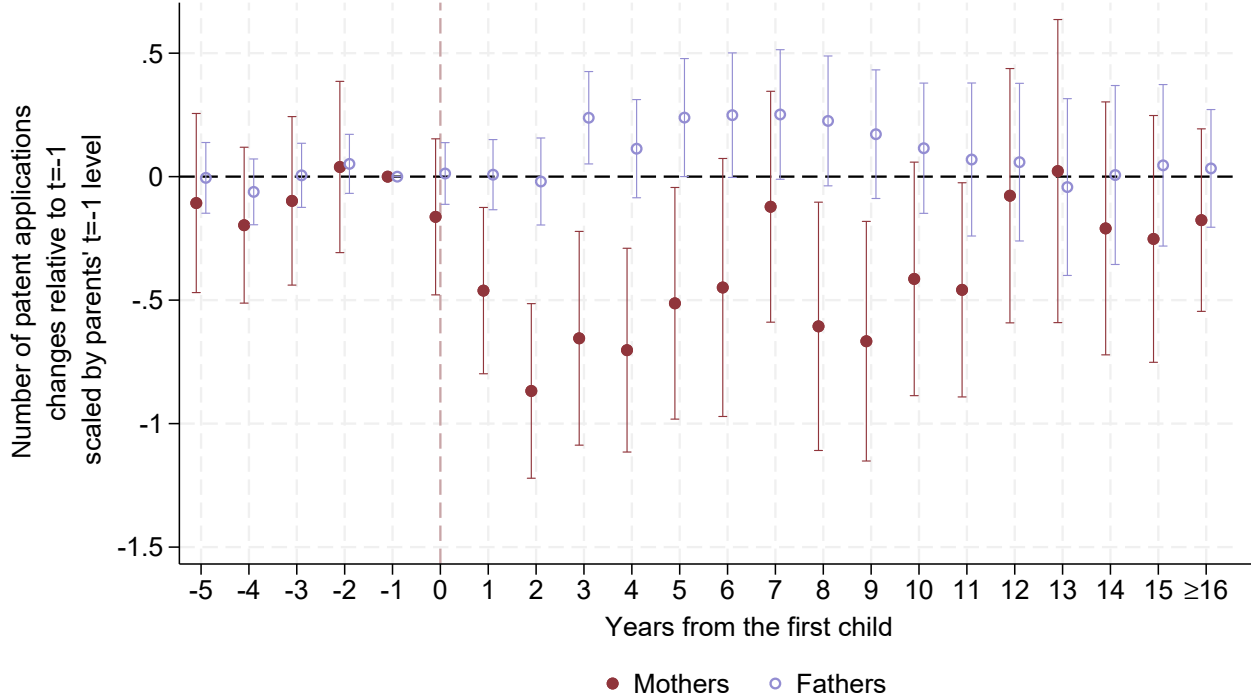


Figure A.16: Scaled Dynamic DID Estimates from Figure A.15B, Panel B

Notes. This figure displays the dynamic DID estimates from Figure 3, Panel B, scaled by the parents' reference year ($t = -1$) patenting rate. The outcome variable is the number of patent filings in a given year. The dynamic DID estimates are estimated separately for women and men using the model in Eq. (4). The model includes calendar-year, individual, and 3-year age-bin fixed effects. The estimation sample consists of the matched parent-childless individual pairs from cohorts born in 1950 or later, observed from 1987 to 2019. Standard errors are clustered at the individual and the matched-pair levels.

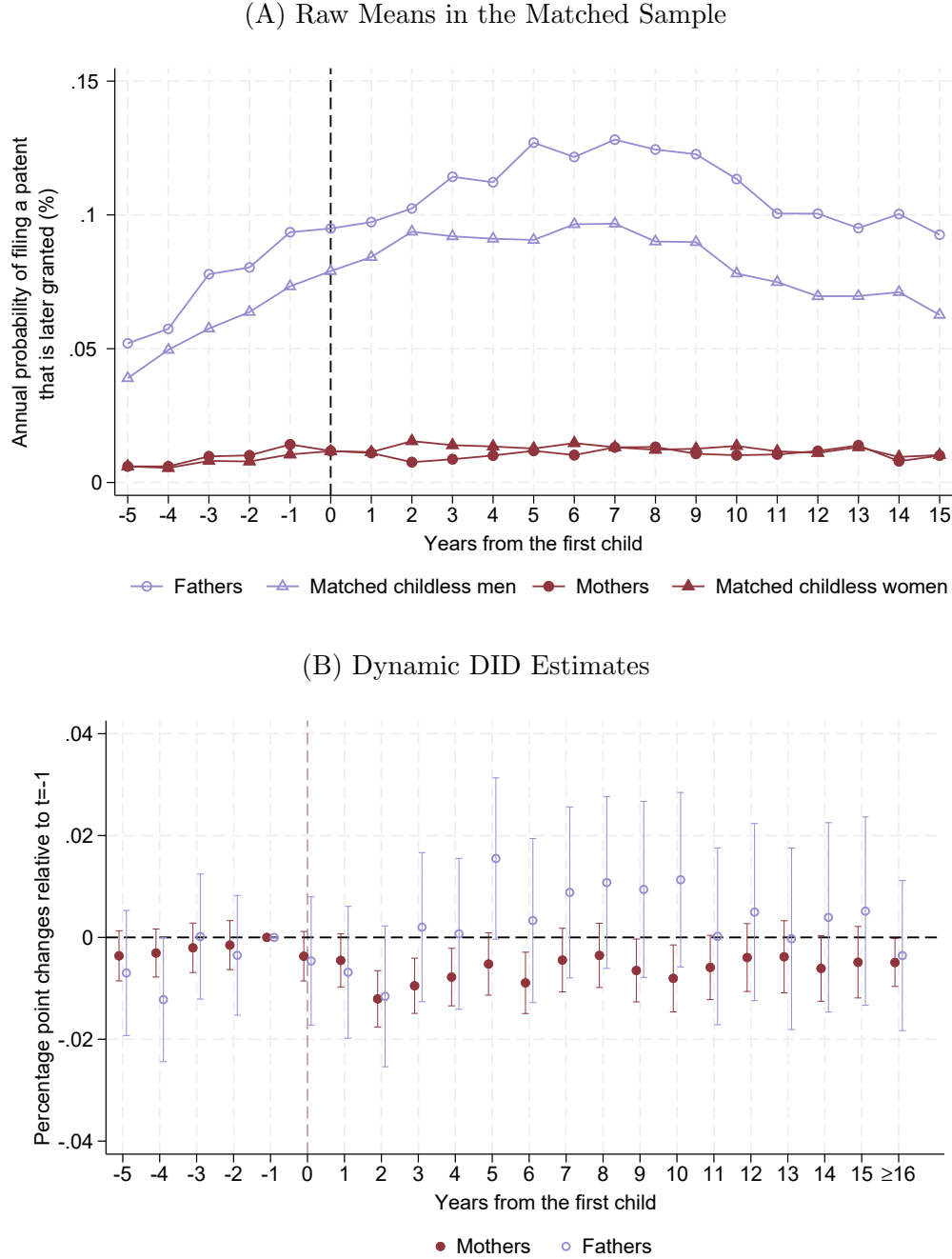


Figure A.17: Changes in Annual Filing Rates of Subsequently Granted Patent Applications Around First Childbirth

Notes. Panel A displays the raw probabilities of filing a later-granted patent application for parents and matched benchmark individuals across the estimation window, expressed as percentages. Panel B plots point estimates and 95% confidence intervals, estimated separately for women and men using the dynamic DID model in Eq. (4). For readability, the estimates are scaled by 100. The outcome variable is the number of patent filings in a given year. The model includes calendar-year, individual, and 3-year age-bin fixed effects. The sample comprises matched parent-childless pairs from cohorts born in 1950 or later, observed from 1987 to 2019. Standard errors are clustered at the individual and the matched-pair levels.

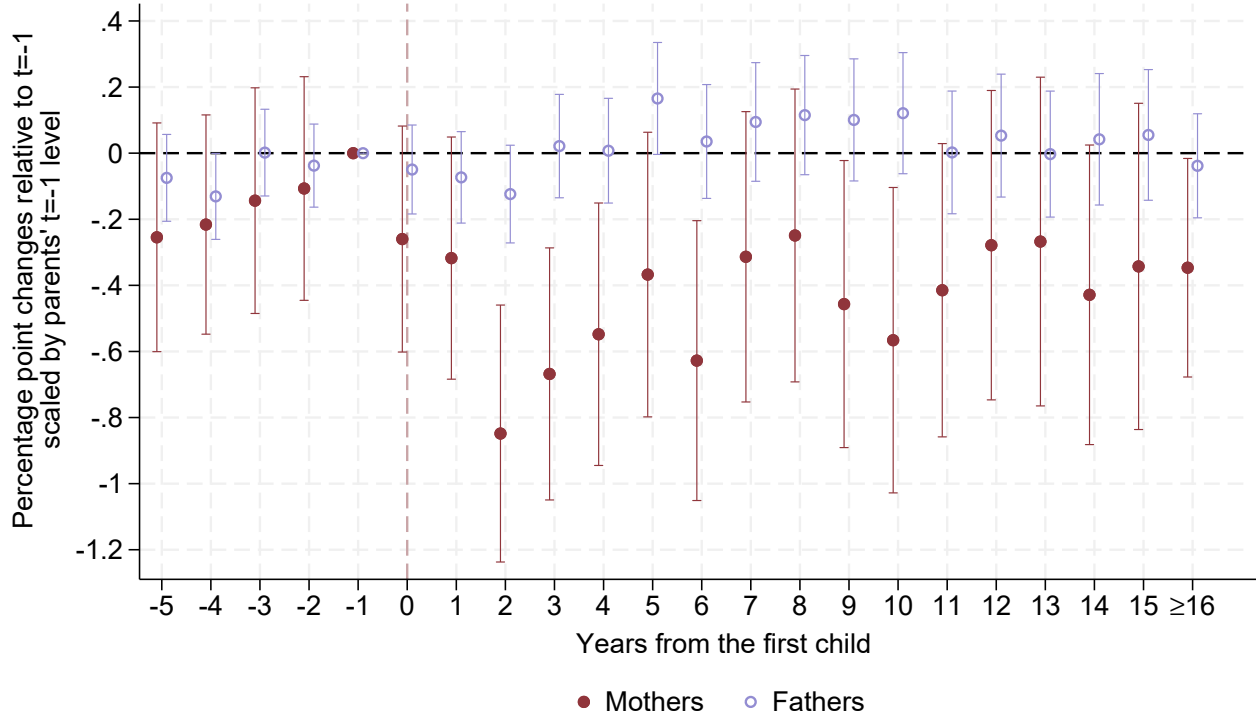


Figure A.18: Scaled Dynamic DID Estimates from Figure A.17, Panel B

Notes. This figure displays the dynamic DID estimates from Figure A.17, Panel B, scaled by the parents' reference year ($t = -1$) patenting rate. The outcome variable is the number of patent filings in a given year. The dynamic DID estimates are estimated separately for women and men using the model in Eq. (4). The model includes calendar-year, individual, and 3-year age-bin fixed effects. The estimation sample consists of the matched parent-childless individual pairs from cohorts born in 1950 or later, observed from 1987 to 2019. Standard errors are clustered at the individual and the matched-pair levels.

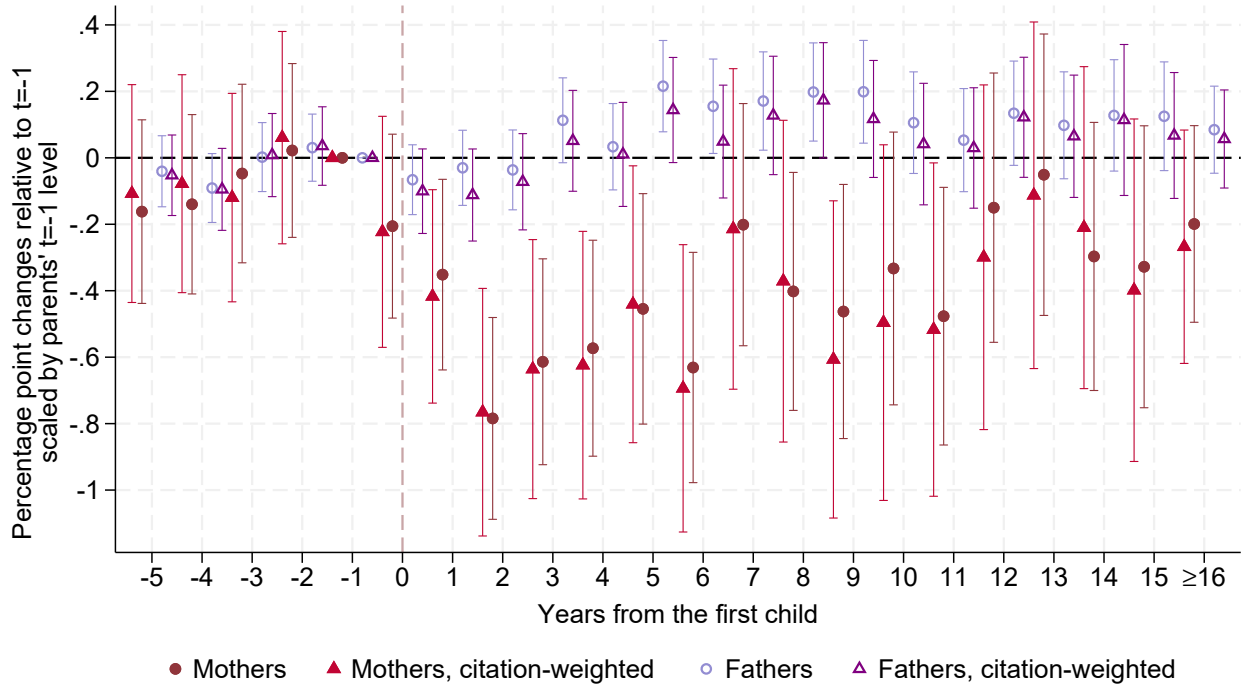


Figure A.19: Patenting Rates Around First Childbirth:
Citation-Weighted vs. Unweighted

Notes. Panel A plots point estimates and 95% confidence intervals, estimated separately for women and men using the dynamic difference-in-differences model in Eq. (4), and compares results from the non-weighted and citation-weighted regression specifications. The estimates are scaled parents' reference-year ($t = -1$) mean patenting rates, and thus represent percentage changes relative to that baseline. The outcome variable is an indicator for filing a patent application in a given year. The regression model includes calendar-year, individual, and 3-year age-bin fixed effects. Standard errors are clustered at the individual and the matched-pair levels.

H Additional Tables

Table A.1: Top 5% Inventor-Intensive Occupations: Size and Workforce Composition

	Mean N workers	Female workers share	Male share over 50% (0/1)	Mean N inventors	Female inventors share	Any female inventors (0/1)
1. Research and development managers	4,574	0.22	1	144.2	0.07	1
2. Telecommunications researchers	3,550	0.14	1	90.1	0.09	1
3. Chemists	1,757	0.55	0	44.5	0.28	1
4. Pharmacologists, pathologists and related professionals	214	0.70	0	5.3	0.62	1
5. Electronics and telecommunications engineers	11,083	0.14	1	275.5	0.07	1
6. Chemical engineers	4,325	0.29	1	97.9	0.23	1
7. Information technology researchers	882	0.18	1	16.6	0.07	1
8. Physicists and astronomers	991	0.23	1	18.5	0.08	1
9. Telecommunications engineers	1,059	0.16	1	19.2	0.08	1
10. Professors	2,404	0.27	1	39.9	0.10	1
11. Electrical engineers	7,097	0.08	1	113.1	0.04	1
12. Social science and related professionals	76	0.36	1	1.2	0.00	0
13. Mining engineers, metallurgists and related professionals	1,170	0.16	1	16.3	0.20	1
14. Physical and engineering science associate professionals	445	0.23	1	5.6	0.11	1
15. Electronics engineers	5,253	0.12	1	63.2	0.05	1
16. Health professionals (except nursing)	140	0.61	0	1.7	0.50	1
17. Mechanical engineers	14,550	0.07	1	156.7	0.02	1
18. Chemical engineering technicians	3,218	0.30	1	30.7	0.32	1
19. Mining and metallurgical technicians	1,094	0.21	1	9.8	0.27	1
20. Architects, engineers and related professionals not elsewhere classified	10,232	0.20	1	91.2	0.19	1
21. Health professionals (except nursing) not elsewhere classified	475	0.79	0	4.2	0.52	1
22. Electronics and telecommunications technicians	10,864	0.18	1	91.3	0.12	1
23. Biologists, botanists, zoologists and related professionals	2,580	0.62	0	21.3	0.50	1
24. Production and operations managers in manufacturing	5,449	0.09	1	43.2	0.02	1
25. Manufacturing managers	9,737	0.13	1	70.8	0.03	1
26. Electronics engineering technicians	4,034	0.17	1	28.0	0.08	1
27. Engineering professionals not elsewhere classified	8,570	0.24	1	56.7	0.12	1
28. Wood processing and chemical engineering technicians	5,525	0.25	1	36.5	0.34	1
29. Industrial and production engineers	4,325	0.21	1	28.1	0.10	1
30. Chemical and physical science technicians	273	0.47	1	1.7	0.19	1
31. Environmental engineers	462	0.53	0	2.9	0.23	1
32. Product and garment designers	1,303	0.57	0	7.4	0.18	1
33. Electrical engineering technicians	5,229	0.08	1	28.3	0.03	1
34. Computer network professionals	635	0.18	1	3.3	0.04	1
35. Managers of small enterprises in manufacturing	1,936	0.14	1	9.7	0.02	1
36. Other specialist managers not elsewhere classified	2,886	0.21	1	12.7	0.08	1
37. Mechanical engineering technicians	17,144	0.06	1	73.6	0.03	1
38. Policy and planning managers	1,407	0.24	1	5.9	0.11	1
39. Managers of small enterprises of business services enterprises	6,121	0.27	1	23.8	0.06	1
40. Information and communications technology service managers	5,048	0.14	1	19.4	0.01	1
41. Assistants and part-time lecturers (university)	2,869	0.45	1	10.9	0.17	1
Column average	4,170	0.27	0.83	44.4	0.16	0.98

Notes. This table extends Table 1 and displays the numbers of workers and inventors, and the share of women in these groups, for the top 5% inventor-intensive occupations. The means are calculated using the years 2004–2018 and workers aged 25–65.

Table A.2: Top 5% Inventor-Intensive Occupations: Gender Ranking Comparison

Occupation name	Pooled ranking	Placement in women's ranking	Placement in men's ranking
Research and development managers	①	⑩	②
Telecommunications researchers	②	④	④
Chemists	③	⑥	①
Pharmacologists, pathologists and related professionals	④	①	③
Electronics and telecommunications engineers	⑤	⑤	⑤
Chemical engineers	⑥	②	⑥
Information technology researchers	⑦	15	⑨
Physicists and astronomers	⑧	17	⑧
Telecommunications engineers	⑨	11	12
Professors	⑩	18	⑩
Electrical engineers	11	14	13
Social science and related professionals	12	-	⑦
Mining engineers, metallurgists and related professionals	13	③	16
Physical and engineering science associate professionals	14	19	15
Electronics engineers	15	22	17
Health professionals (except nursing)	16	⑨	14
Mechanical engineers	17	26	18
Chemical engineering technicians	18	⑧	23
Mining and metallurgical technicians	19	⑦	27
Architects, engineers and related professionals not elsewhere classified	20	13	25
Health professionals (except nursing) not elsewhere classified	21	20	11
Electronics and telecommunications technicians	22	21	24
Biologists, botanists, zoologists and related professionals	23	16	20
Production and operations managers in manufacturing	24	31	26
Manufacturing managers	25	37	28
Electronics engineering technicians	26	23	29
Engineering professionals not elsewhere classified	27	24	30
Wood processing and chemical engineering technicians	28	12	33
Industrial and production engineers	29	25	31
Chemical and physical science technicians	30	30	22
Environmental engineers	31	29	21
Product and garment designers	32	35	19
Electrical engineering technicians	33	32	36
Computer network professionals	34	44	32
Managers of small enterprises in manufacturing	35	63	37
Other specialist managers not elsewhere classified	36	36	41
Mechanical engineering technicians	37	34	48
Policy and planning managers	38	33	44
Managers of small enterprises of business services enterprises	39	50	42
Information and communications technology service managers	40	81	49
Assistants and part-time lecturers (university)	41	38	34

Notes. This table compares the position of each top 5% inventor-intensive occupation in the pooled ranking with its respective placement in the rankings for women and men. The ranking for women is based on the ratio of female inventors to female workers, with an analogous metric applied for men. An empty cell indicates that there are no female inventors in the given occupation. The top 10 inventor-intensive occupations are highlighted in green with circles, ranks 11–20 in yellow with squares, and ranks 21–30 in orange with diamonds.

Table A.3: Sample Sizes and Likelihood Ratios for Birth Cohorts 1962–1982 in Figure 1

	Women	Men
<i>Full population</i>		
N all	746,059	786,877
N inventors	1,617	11,046
Inventors per thousand	2.2	14.0
Likelihood ratio women/men		0.15
<i>STEM college</i>		
N all	55,421	126,351
N inventors	1,283	8,633
Inventors per thousand	23.2	68.3
Likelihood ratio women/men		0.34
<i>STEM college & top 5% invent. occup.</i>		
N all	19,797	79,432
N inventors	1,134	8,062
Inventors per thousand	57.3	101.5
Likelihood ratio women/men		0.56
<i>STEM college & top 5% invent. occup for 5+ cons. years</i>		
N all	8,368	43,898
N inventors	799	8,062
Inventors per thousand	95.5	143.9
Likelihood ratio women/men		0.66

Notes: This table presents the sample sizes, inventors per thousand, and the corresponding likelihood ratios for the full population of birth cohorts from 1962 to 1982, as well as for the three STEM college graduate subsamples shown in Figure 1. Inventorship is determined using EPO patent application data from 1978 to 2019. Occupations are observed between 2004 and 2018.

Table A.4: Gender Patenting Gap:
Including Observations with Missing Occupation or Firm Information

	(1)	(2)	(3)	(4)
<i>Panel A. Outcome = patent filing dummy</i>				
Woman	-0.0967*** (0.0014)	-0.0458*** (0.0015)	-0.0473*** (0.0016)	-0.0437*** (0.0015)
% of Woman in Column (1) explained		53%	51%	55%
Constant term	0.1098*** (0.0013)	0.0845*** (0.0011)	0.0769*** (0.0013)	0.0853*** (0.0012)
Adj. R^2	0.0004	0.0186	0.0439	0.0188
<i>Panel B. Outcome = granted patent filing dummy</i>				
Woman	-0.0597*** (0.0010)	-0.0272*** (0.0011)	-0.0288*** (0.0011)	-0.0261*** (0.0011)
% of Woman in Column (1) explained		54%	52%	56%
Constant term	0.0671*** (0.0010)	0.0510*** (0.0008)	0.0518*** (0.0009)	0.0462*** (0.0010)
Adj. R^2	0.0002	0.0119	0.0293	0.0120
Baseline Controls	-	Yes	Yes	Yes
Education level $\times$ field FE	-	Yes	Yes	Yes
Occupation FE	-	Yes	Yes	Yes
Firm FE	-	-	Yes	-
High school grade FE	-	-	-	Yes

Notes. This table provides the estimated gender gap coefficients $\hat{\gamma}$ from regression specifications described in equation (1), estimated with OLS. All coefficients and standard errors are multiplied by 100 and rounded to four decimal places. The estimation sample is a panel of individuals observed between 2004 and 2018, aged 25–65. Observations with missing firm or occupation information are included in the estimation sample and assigned to “missing” categories. The dependent variable in Panel A is a patent filing indicator. In Panel B, the outcome is a patent grant indicator, defined as 1 in the year the later-granted patent is filed and 0 otherwise. The baseline controls include fixed effects for calendar year, age, and native language, as well as an employment dummy. Educational fields are classified at the 3-digit level (ISCED-2011) and education levels at the 2-digit level. Occupations are categorized at the most detailed 4-digit or 5-digit level. Standard errors are clustered at the individual level. The number of observations in each regression is 43,684,360.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.5: Twofold Pooled Oaxaca–Blinder Decomposition of the Gender Patenting Gap, Corresponding to Table 2, Column (2)

<i>Outcome:</i>	(1.1) <i>Patent filing</i>	(1.2)	(2.1) <i>Granted patent filing</i>	(2.2)
	Coefficient	SE	Coefficient	SE
<i>Panel A: Decomposition</i>				
Male mean	0.1513***	(0.0019)	0.0930***	(0.0014)
Female mean	0.0180***	(0.0006)	0.0102***	(0.0004)
Raw gap (R)	0.1333***	(0.0020)	0.0827***	(0.0015)
Explained (E)	0.0662***	(0.0017)	0.0422***	(0.0013)
<i>Contributions:</i>				
Baseline controls	0.0016***	(0.0001)	0.0008***	(0.0001)
Education level $\times$ field FEs	0.0102***	(0.0015)	0.0072***	(0.0011)
Occupation FEs	0.0544***	(0.0015)	0.0342***	(0.0011)
Unexplained (U)	0.0671***	(0.0023)	0.0405***	(0.0017)
<i>Panel B: Relative contributions</i>				
<i>% of the raw gap</i>				
Explained	49.7		51.0	
Unexplained	50.3		49.0	
<i>Relative contributions to the total explained (%)</i>				
Baseline controls	2.4		2.0	
Education level $\times$ field FEs	15.4		17.1	
Occupation FEs	82.2		80.9	

Notes: This table displays the estimated pooled twofold Oaxaca–Blinder decomposition as in Equation (A.5). All coefficients and standard errors are multiplied by 100 and rounded to four decimal places. The decomposition was implemented using the *oaxaca* package in Stata (Jann, 2008). Fixed effects are normalized to ensure that the results are invariant to the choice of omitted categories. The estimation sample is a panel of individuals observed between 2004 and 2018, aged 25–65. In Columns (1.1) and (1.2), the dependent variable is a patent filing indicator. In Columns (2.1) and (2.2), the outcome is a patent grant indicator, defined as 1 in the year the later-granted patent is filed and 0 otherwise. The baseline controls include indicator variables for calendar year, age, native language, and employment. Educational fields are classified at the 3-digit level (ISCED-2011), and educational levels at the 2-digit level. Occupations are categorized at the most detailed 4-digit or 5-digit level. Standard errors are clustered at the individual level. The pooled sample comprises 30,275,290 observations, of which 15,198,437 correspond to men and 15,076,853 to women. $p < 0.1$. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A.6: Gender Differences in Patent Characteristics

	(1)	(2)	(3)
Specification:	Raw difference	+ Patent controls	+ Individual controls
1. Patent granted (0/1)			
Woman	-0.0498*** (0.0108)	-0.0507*** (0.0106)	-0.0250*** (0.0095)
Constant	0.5643*** (0.0044)	0.5643*** (0.0037)	0.5620*** (0.0034)
Woman as % of Col. (1) Constant	-8.8%	-9.0%	-4.4%
Observations	72,302	72,302	72,302
2. Number of received family-to-family citations			
Woman	-0.9910 (0.8168)	-1.1071 (0.7793)	-2.8689*** (0.7973)
Constant	19.0652*** (0.4214)	19.0759*** (0.3748)	19.2386*** (0.3603)
Woman as % of Col. (1) Constant	-5.2%	-5.8%	-15.0%
Observations	72,302	72,302	72,302
3. Number of received family-to-family citations (Poisson model)			
Woman	-0.0535 (0.0447)	-0.0436 (0.0394)	-0.1231*** (0.0380)
Constant	2.9480*** (0.0221)	3.1909*** (0.0184)	3.2554** (0.0170)
Woman in %	-5.2%	-4.3%	-11.6%

Table continues on the next page

Observations	72,294	72,294	72,294
4. Breakthrough (0/1)			
Woman	-0.0027*** (0.0010)	-0.0037*** (0.0011)	-0.0044*** (0.0012)
Constant	0.0050*** (0.0007)	0.0051*** (0.0007)	0.0051*** (0.0006)
Woman as % of Col. (1) Constant	-54.0%	-74.0%	-88.0%
Observations	72,302	72,302	72,302
5. Patent family size			
Woman	1.6832*** (0.2081)	0.3483* (0.1842)	0.0304 (0.1700)
Constant	6.9814*** (0.0763)	7.1047*** (0.0709)	7.1341*** (0.0647)
Woman as % of Col. (1) Constant	24.1%	5.0%	0.4%
Observations	72,302	72,302	72,302
6. Patent family size (Poisson model)			
Woman	0.2160*** (0.0247)	0.0394* (0.0221)	0.0026 (0.0198)
Constant	1.9433*** (0.0109)	1.9952*** (0.0100)	2.0221*** (0.0090)
Woman in %	24.1%	4.0%	0.3%
Observations	72,302	72,302	72,302
7. Patent scope			
Woman	0.3329***	0.1117***	0.0169

Table continues on the next page

	(0.0362)	(0.0316)	(0.0276)
Constant	1.9482***	1.9686***	1.9774***
	(0.0113)	(0.0107)	(0.0097)
Woman as % of Col. (1) Constant	17.1%	5.7%	0.9%
Observations	72,291	72,291	72,291

9. Originality

Woman	0.0485***	0.0176***	0.0010
	(0.0048)	(0.0046)	(0.0042)
Constant	0.7141***	0.7169***	0.7184***
	(0.0018)	(0.0017)	(0.0016)
Woman as % of Col. (1) Constant	6.8%	2.5%	0.1%
Observations	70,511	70,511	70,511

10. Radicalness

Woman	0.0273***	0.0142**	0.0030
	(0.0065)	(0.0062)	(0.0059)
Constant	0.3265***	0.3277***	0.3287***
	(0.0025)	(0.0025)	(0.0023)
Woman as % of Col. (1) Constant	8.4%	4.4%	0.9%
Observations	72,291	72,291	72,291

11. Quality Index 4

Woman	0.0190***	0.0013	-0.0030
	(0.0043)	(0.0042)	(0.0034)
Constant	0.2782***	0.2797***	0.2801***
	(0.0015)	(0.0015)	(0.0013)

Table continues on the next page

Woman as % of Col. (1) Constant	6.8%	0.5%	-1.1%
Observations	42,535	42,535	42,535
12. Quality Index 6			
Woman	0.0001	-0.0067**	-0.0063**
	(0.0032)	(0.0031)	(0.0025)
Constant	0.3208***	0.3214***	0.3214***
	(0.0011)	(0.0011)	(0.0009)
Woman as % of Col. (1) Constant	0.0%	-2.1%	-2.0%
Observations	42,535	42,535	42,535
Patent controls	No	Yes	Yes
Individual controls	No	No	Yes

Notes. This table reports the coefficients for the female indicator across patent-quality measures. The outcomes and the regression model are defined in Supplemental Appendix E. Column (1) shows raw gender differences; Column (2) adds patent application-level controls: filing-year fixed effects, broad technology-sector indicators (5 categories), and inventor-team size indicators; Column (3) further adds inventor-level controls: age fixed effects, 2-digit education level by 3-digit field-of-education fixed effects, and 3-digit employer industry fixed effects. The Poisson model coefficients are converted to percentage changes as $100 * (e^{\gamma} - 1)$, where γ denotes the estimated coefficient from the Poisson model. Standard errors, clustered two-way at the patent application and inventor levels, are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.7: Observation Window by Treatment Cohort

Cohort	Event Time (t)																				
	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1991		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1992	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1993	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1994	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1995	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1996	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1997	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1998	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
1999	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2000	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2001	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2002	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2003	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2004	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
2005	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		
2006	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		
2007	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x			
2008	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
2009	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					
2010	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x						
2011	x	x	x	x	x	x	x	x	x	x	x	x	x								
2012	x	x	x	x	x	x	x	x	x	x	x	x									
2013	x	x	x	x	x	x	x	x	x	x	x										
2014	x	x	x	x	x	x	x	x	x	x											
2015	x	x	x	x	x	x	x	x	x												
2016	x	x	x	x	x	x	x	x													
2017	x	x	x	x	x	x	x														
2018	x	x	x	x	x	x															
2019	x	x	x	x	x																

Notes. This table shows the observed event-time periods within the estimation window $t \in [-5, 15]$ for the treatment cohorts (year of the first childbirth) from 1991 to 2019 in the panel data covering 1987 to 2019. An 'x' indicates that a given treatment cohort is observed at the corresponding event period t (years from the first childbirth). Treatment cohorts 1992–2004 are observed throughout the estimation window.

Table A.8: Number of Benchmark Individual Copies in the Matched Estimation Sample

(1) Number of copies	(2) Number of unique benchmark individuals	(3) Number of duplicates	(4) Total number of benchmark individuals
1	254,528	0	254,528
2	124,629	124,629	249,258
3	70,690	141,380	212,070
4	39,421	118,263	157,684
5	20,686	82,744	103,430
6	10,349	51,745	62,094
7	5,133	30,798	35,931
8	2,446	17,122	19,568
9	1,136	9,088	10,224
10	492	4,428	4,920
11	202	2,020	2,222
12	60	660	720
13	14	168	182
14	2	26	28
Total	529,788	583,071	1,112,859

Notes. This table shows the number of benchmark individual copies in the matched parent-childless individual estimation sample. Column 1 categorizes the benchmark individuals by copy count; Column 2 shows the number of unique benchmark individuals in each group; Column 3 shows the number of benchmark individual duplicates in each group; and Column 4 shows the total (unique + duplicates).

Table A.9: Gender Patenting Gap and Parenthood

Variable	Coef.	(1)	(2)	(3)
		Outcome: patent dummy		
<i>Woman</i>	$\hat{\beta}_0$	-0.1030*** (0.0015)	-0.0680*** (0.0018)	-0.0315*** (0.0019)
<i>HasChildren</i>	$\hat{\beta}_1$		0.0479*** (0.0027)	0.0268*** (0.0026)
<i>HasChildren</i> $\times$ <i>Woman</i>	$\hat{\beta}_2$		-0.0642*** (0.0029)	-0.0340*** (0.0027)
<i>Constant</i>		0.1173*** (0.0015)	0.0919*** (0.0018)	0.0693*** (0.0019)
	$\hat{\beta}_1 + \hat{\beta}_2$		-0.0163*** (0.0011)	-0.0072*** (0.0011)
	$\hat{\beta}_0 + \hat{\beta}_2$		-0.1323*** (0.0024)	-0.0655*** (0.0024)
Adj. R^2		0.0004	0.0007	0.193
Observations		39,932,367	39,932,367	39,932,367
Baseline controls		-	Yes	Yes
Education level $\times$ field FE		-	-	Yes
Occupation FE		-	-	Yes
High school grades FE		-	-	Yes

Notes. This table presents results from three different specifications of Equation (A.14), estimated with OLS. All coefficients and standard errors are multiplied by 100 and rounded to four decimal places. The dependent variable in each regression is a patent filing indicator. The baseline controls include indicator variables for calendar year, age, native language, and employment. Educational fields are classified at the 3-digit level (ISCED-2011), and educational levels at the 2-digit level. Occupations are categorized at the most detailed 4-digit or 5-digit level. The estimation sample is a panel of individuals born in 1950 or later, observed in 1995, 2000, and 2004–2018 at ages 18–65. I retain observations with missing occupation information and assign them to a “missing” category. Robust standard errors are clustered at the individual level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$