Gender differences among innovators: a patent analysis of stars

Gender differences among innovators: a patent analysis of stars

Federico Caviggioli[a] , Alessandra Colombelli*[a] and Chiara Ravetti[a]

[a]*Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Turin, Italy*.

**Corresponding author*. *Email: alessandra.colombelli@polito.it*

PLEASE CITE USING THE MOST RECENT VERSION

https://doi.org/10.1080/10438599.2022.2065634

Gender differences among innovators: a patent analysis of stars

This article examines the gender gap in patenting activities and the predominance of male innovators among outstanding inventors, so-called “stars”. In particular, we investigate different metrics of productivity among top inventors, identified employing different definitions with respect to the quantity and quality of output. We distinguish between *prolific* inventors, with high numbers of patents registered in their name, and *high-quality* inventors, with portfolios comprising patents with large numbers of citations. Using patent data for more than 600,000 inventors, we find that star inventors differ from the pool of non-star inventors in terms of gender: while for non-star inventors being a woman constitutes a significant disadvantage, for stars it actually presents a positive association both with quantity and quality of innovative outputs. Moreover, career length constitutes a key premium for female inventors’ productivity, but with smaller magnitudes among stars. The only exception where we observe no gender differences is among inventors with large portfolios (more than five patent families): among them, women do not display any significant gap in the quality of outputs, nor does career length provide a gendered premium.

Keywords: gender gap, patents, innovation, productivity, career

Word count: 9’134

1. Introduction

Inventors are at the heart of the process of advancing human civilization, and their characteristics are key for understanding innovative dynamics in modern times. In recent years, the innovation literature has focused on diversity and gender gaps in the inventors’ population, highlighting a substantial predominance of male innovators among patent holders (Hunt et al., 2013; Lax Martinez et al., 2016; Haseltine & Chodos, 2017). Despite a positive trend in female representation, which narrowed the gender gap over time, women remain a small fraction of inventors (Heikkilä 2019; USPTO 2019) and are less likely to benefit financially from intellectual property rights (Kline et al., 2019).[1] This under-representation of women inventors increases the societal loss of potential

[1] This issue is not only problematic at the individual level: firms adopting non-patented innovations may face more difficulties in accessing financial support, especially start-ups (Caviggioli et al., 2020a; Fischer & Ringler, 2014; Caviggioli et al., 2020b, Colombelli et al. 2020).

innovations from the available talent pool (Bell et al. 2019). Conversely, diversity is desirable since it spurs the evolution of knowledge and prevents paradigmatic lock-in (Wullum Nielsen and Börjeson 2019; Stirling 2007). For gender diversity more specifically, there is already evidence that countries and organizations suffering from skill shortages might benefit from resorting to the innovative potential of women (Hoisl and Mariani, 2017; Amoroso and Audretsch, 2020; Menter, 2020; Colombelli et al. 2021).

Despite its economic and social significance, the study of female underrepresentation amongst innovators is still in its infancy. Importantly, most analyses of gender gaps in patenting do not distinguish between innovators with different levels of productivity and do not account for heterogeneities in the gender gap at the top of the distribution for so-called "star inventors", namely those individuals able to generate superior innovative outcomes (Zucker and Darby, 1997; Groysberg and Lee, 2009; Oldroyd and Morris, 2012).[2] Star inventors are rare and particularly valuable to advance societies' knowledge base, and there is already some evidence of lost 'Marie Curies' because fewer women manage to become stars (Bell et al. 2019), but to date there is no clear evidence about the difference in gender gaps among stars compared to non-star innovators.

In this context, the main contribution of this article is to examine gender gaps among inventors considering their productivity, and to shed light on how female representation and inventor-level outcomes vary among stars and non-stars. For this analysis, we develop different operational definitions of "stars" from the relevant literature (Hess and Rothaermel, 2011; Kehoe and Tzabbar, 2015), to measure patenting performance along the dimensions of *quantity* and *quality* (Call, Nyberg, and Thatcher 2015; Aguinis and O'Boyle 2014). The comparison of inventors with an average productivity versus highly

---

[2] One notable exception is the study by Bell et al. (2019:707), which shows a plot with the unconditional gender gap among highly-cited inventors, but does not perform any econometric analysis on it.

prolific and highly cited ones helps identifying more targeted policies to support women inventors, depending on the group experiencing the largest inequalities.

The analysis is based on a sample of more than 600’000 inventors over the period 1981-2010. We first examine the unconditional differences between male and female inventors according to the two definitions of quantity and quality stars. We find that, while women are under-represented with all metrics, the underrepresentation among stars is smaller when considering quality-based measures. We then perform a series of multivariate regressions,to examine the role of career length as a key mediating factor, while controlling for sectoral and geographic characteristic of inventors. It is crucial to account for the technology sector of operation due to the substantial differences in gender shares in different disciplines. Furthermore, the US offices file a large number of patents from non-residents (WIPO 2021, pp.32), thus we want to differentiate people with their main geographic interests in the US from those more tied to other countries, by using a proxy for inventors’ country of origin, the first country of patent filing (earliest priority).

Results reveal that the negative association between being a woman and patent productivity, either in terms of quantity or quality, is overturned when considering the sample of star inventors. Furthermore, any additional year of career has a small but positive impact on productivity (measured both as quantity and quality) for women with respect to male inventors. However, for the top group of high-quality inventors with large patent portfolios, there is no discernible gender difference in productivity, nor a gendered career premium.

The implications of our findings are in line with the literature on gender gaps in contexts other than patenting, which show that women’s career is penalized by family responsibilities that allow them less time flexibility and slow down career progression

(Goldin 2014; Blau and Kahn 2017b). Additionally, however, we find that some women overcome the obstacles leading to the top percentiles of patenting inventors, especially if we include smaller patent portfolios to measure quality. Our analysis does not rule out biases and discrimination against women inventors at all levels, such as stereotypes against women in sectors that are culturally associated with masculinity, like science, technology, engineering, and mathematics (STEM), and which are a major source of patents – in fact, we find evidence of large variations in gender gaps among different industrial sectors. However, the robust effect of including career length indicates that time is one of the key factors for female innovative productivity.

The rest of the article is organized as follows. Section **Errore. L'origine riferimento non è stata trovata.** examines the previous literature to develop our hypotheses. Section 3 presents the data and summary statistics, and Section 4 the empirical methodology applied. Section 5 discusses the results and finally Section 6 concludes with some policy implications and further research avenues.

2. Gender and stars: state of the art and hypotheses from the literature

The empirical literature focusing on gender differences in scientific productivity has mostly considered scientific articles as the key innovative outputs, and only to a limited extent has examined patents. We therefore consider all studies that focus on patented innovations, but also some of the literature on the production of scientific in order to inform the identification of the most productive individuals and for the mechanisms underlying gender gaps.

a. Gender gaps in patenting

Several works have recently acknowledged the under-representation of women among inventors (Hunt et al. 2013; Haseltine and Chodos 2017) with gender shares that vary over time and across office of filing (see, for example, Bell et al. 2019; Heikkilä 2019; Lax Martinez, Raffo, and Saito 2016). For example, in the US in 2016, around 20% of patents included at least one female inventor and female inventors represented 12% of total inventors (USPTO 2019). The reasons identified in the literature for this underrepresentation are multiple. Several articles found evidence of horizontal segregation - gendered differences in the field of studies and in the industrial sectors of activity (Mayer and Rathmann 2018; Wullum Nielsen and Börjeson 2019; Puuska 2010). According to studies of horizontal segregation, cultural norms and expectations push women and men towards different areas of work and study, with men more likely to be involved in activities with technical and operational aspects (Bettio and Verashchagina 2009). This cultural pressure impacts on career choices already starting in higher education and then in the labour market (Cech 2013; Charles and Bradley 2002), as seen by the under-representation of female undergrad and PhD students in STEM areas (Hunt et al. 2013; Loan and Hussain 2017; Toivanen and Väänänen 2016). Horizontal segregation in the intellectual workforce is well documented in the production of scientific articles, where women and minorities are systematically underrepresented in less technical fields (Kozlowski et al. 2022).

Moreover, gender gaps can arise in the form of vertical segregation, due to different conditions faced by men and women when trying to achieve the pinnacle of their career. Unpaid labour in the form of childbearing, childcare, housework and care of the elderly people in the family are tasks unevenly distributed between men and women, which reduce the time that women can invest in paid jobs (Blau and Kahn 2017a). Women on average report facing greater difficulties in balancing professional and family life (Stack 2004; Loan and Hussain 2017; Mayer and Rathmann 2018; Abramo, D'Angelo, and Caprasecca 2009). Related to this aspect, the difference in productivity is larger before the age of 40 than later (Kyvik and Teigen 1996). The productivity gap seems to decline with time: maternity leave has more relevant effects in earlier steps of career progression (Abramo, D'Angelo, and Caprasecca 2009; Joy 2006).

For mechanisms underlying the gender gap specifically in patenting, a masculine bias in STEM has been observed also for the filing of intellectual property rights: for instance Heikkilä, (2019) confirmed it for design rights, trademarks and utility models in Finland. There is also evidence that the gender gap in patenting could be linked to gendered tasks, such as technical design and development (Hunt et al. 2013). In this context, there is also evidence that males are more likely to produce the kind of scientific outputs that is protected by patents (Barwa and Rai 2003; Leahey and Blume 2017; Lai 2020). Even the current definition of patent law can be considered "gendered", as noticed by the study by Lai, (2020).[3]

b. Star individuals: definitions and gender gaps

Next, our article relates to the literature on scientific productivity and on the identification of star innovators. Star scientists are considered of interest since they generate superior innovation outcomes (Groysberg and Lee, 2009; Oldroyd and Morris, 2012; Zucker and Darby, 1997). Although in some cases the literature identifies negative effects in organizations due to coordination costs and conflicts related with the presence of stars inventors (Bendersky and Hays 2012; Groysberg, Polzer, and Elfenbein 2011; Swaab et al. 2014) and because hiring "stars" is often expensive (Groysberg, Polzer, and Elfenbein 2011), there is a general consensus about their overall positive impact. Their effect is not limited to a direct increase of output but also to broader support for an organization's

[3] In her analysis, patent law embodies the notion that the typical technology is something mechanical or from a male-dominated field. "This is, in turn, reflected in who is considered to be an inventor and what is deemed to be patentable. That patent law is gendered means that the law is rewarding masculine forms of invention over feminine forms." (p. 17 in Lai, 2020).

activities (Kehoe and Tzabbar 2015) and to the attraction of resources and skilled personnel (Lacetera, Cockburn, and Henderson 2004; Hess and Rothaermel 2011). They also indirectly foster the productivity of peers and collaborators thanks to learning and emulation (Lockwood and Kunda 1997).

When studying gender issues in terms of productivity of star individuals, the literature has mostly focused on scientific articles.[4] The list of studies providing evidence on the higher male research productivity in academic journals is long (e.g. Baccini, Barabesi, Cioni, & Pisani, 2014; Beaudry & Larivière, 2016; Hunter & Leahey, 2010; König, Fell, Kellnhofer, & Schui, 2015; Stack, 2004; Xie & Shauman, 1998; Akbaritabar, Casnici, & Squazzoni, 2018; Arruda, Bezerra, Neris, de Toro, & Wainera, 2009; Mayer and Rathmann, 2018). Significant differences between productivity levels of males and females were found in particular in STEM fields (Sax et al. 2002; Kretschmer and Kretschmer 2013; Abramo, D'Angelo, and Caprasecca 2009). The study of Kwiek (2016) on European data has shown that being a female academic is a strong predictor of not becoming a highly productive researcher (i.e. top 10%) but only in some countries, that is Italy and the United Kingdom, thus highlighting that the discrepancies are not only due to individual factors, but also to local cultural, socio-economic and policy conditions.

This literature shows evidence of "glass ceilings", unseen barriers that block career progress of women towards the top of a profession, despite their qualifications and skills (Kretschmer and Kretschmer 2013). In our analysis, we apply the concept of glass ceilings broadly, considering the threshold dividing star scientists from non-star ones as a potentially challenging productivity step that women may have more difficulties surpassing compared to men.

[4] There is also evidence in terms of wages that in the past decades the gender pay gap declined much more slowly at the top of the wage distribution than at the middle or bottom (Blau and Kahn 2017a), thus pointing out at some specific difficulties in closing the gap in the highest fraction of the distribution.

When dealing with star innovators and scientists, it is important to note that their identification is not only determined by the amount of output, but it is also related to its quality. For this reason, previous studies have employed citations as measure of quality, again focusing mostly on scientific publications rather than patents. The results of this literature are mixed: some analyses actually found evidence of higher citation rates for women (Borrego et al. 2010), some found a male advantage (Larivière et al. 2013), and others found no significant difference (Lerchenmueller and Sorenson 2018; Nielsen 2016).

The literature on female stars in patents is less developed but still points towards a possible gap even at the top of the distribution, in particular when considering those patents that are commercialized, largely due to women's underrepresentation in engineering and in jobs involving development and design (Hunt et al. 2013). Some evidence for the Unites States indicates that environmental factors, and particularly exposure to specific networks and to same-gender mentorship, might be key in limiting the access of female inventors to star-level patenting (Bell et al. 2019). To the best of our knowledge, however, there are to date no econometric studies that examine the presence of gender gaps among top inventors of patented innovations.

c. Hypotheses and research question

Considering these different streams of the literature regarding top inventors and the gender gap, two possible mechanisms are at play that we must consider in our analysis: on the one hand, women trying to become stars may suffer from all the limitations, glass-ceilings and constraints that make them already underrepresented in the scientific community both due to vertical and horizontal segregation. If competition and discrimination get fiercer at the top, we should observe an even larger gender gap among star inventors. On the other hand, female inventors who manage to thrive in these gendered innovation systems by overcoming all the hurdles identified in the literature are likely to be a selection of the most motivated, innovative and creative ones. If inventors at the top are a selection of the best, we may observe a reduced gender gap, or even a reversed one, where the women who make it to the top are even better than their male counterparts.

Since the evidence from the scientific articles production literature offers mixed results, and the patent literature has not yet explored this issue, we do not have a strong theoretical expectation for gender gap differences among inventors of different productivity levels. Whether the glass ceiling or the selection effect dominates among female star innovators is an empirical question, that can be influenced by the way productivity is measured and by the moderating role of career length. Our research question is thus the following: is the gender gap among top inventors different than for other scientist producing average patenting outputs? In other words, is being a woman a greater or smaller disadvantage for star innovators compared to the whole pool of innovators? To answer this question, we consider different measures of patenting productivity (capturing both quantity and quality) while controlling for the different sectors of activity, and inquire about the role of career length, which in the literature has been identified as a key factor in shaping gender gaps.

3. Data and star inventors' identification

a. Data

The main data sources for our analysis are PatentsView and PATSTAT. PatentsView is a data warehouse sourced from USPTO-provided data on granted patents. Critically for our study of individual inventors, this database includes disambiguated inventors' names from the application of an algorithm[5]. We then linked the patent level data from PatentsView to PATSTAT, the largest repository of patent data in terms of coverage and available information, maintained by the EPO with the collaboration of the main patent offices[6]. Extending data to PATSTAT provides access to additional data at patent- and family-level (see next paragraphs).

The analysis is carried out at the level of inventors, with a sample defined as follows. The initial sample includes all the inventors with at least one US granted patent filed between 2008-10[7], corresponding to 725,577 disambiguated names in PatentsView. The selected inventors are associated to around 4.3 million granted patents which are then linked to PATSTAT where further information is collected. The cut-off year of 2010 is required to allow for a sufficiently large subsequent time window to calculate quality indicators such as citations and to account for potential delays in the publication of documents. In terms of granted patents, the selected sample represents 58% of the total US activities recorded in PatentsView up to 2020.

All the selected patents are linked to their patent family through PATSTAT (2.9 million INPADOC families).[8] Patent families represent the unit of analysis that is closer to a single invention: multiple patent documents regarding the same filing, for example across different patent offices in multiple countries, are collapsed to a single unit, providing a more accurate measure of inventors' productivity (OECD Patent Statistics Manual 2009; Martínez 2011). Furthermore, country extensions provide information on the

[5] More information at www.patentsview.org (last access September 2021, data on disambiguated inventors' names from March 2020 release).
[6] More information at https://www.epo.org/searching-for-patents/business/patstat.html (last access in September 2021, version of database used in this study: fall 2019).
[7] Only "utility" patents have been considered. Withdrawn patents are included (corresponding to 0.17% of the examined granted patents).
[8] An INPADOC patent family comprises all the documents sharing directly or indirectly (e.g. via a third document) at least one priority.

geographical coverage of an invention. We also collect the earliest filing year and the IPC subclasses of each family and calculate several patentometrics following the approach described in Caviggioli et al. (2020). For each inventor it is thus possible to identify the portfolio of inventions and create portfolio level measures, as described in the next section.

With the aim to minimize potential errors in the original data, either in name disambiguation or in patent family identification, those inventors reporting a portfolio-level earliest filing date prior to 1981 (3.2%) were excluded. Inventors with no IPC codes associated to the portfolio were also eliminated (0.01%). The cleaned sample consists of a selection of 703,977 inventors active in the years 2008-10 and with a patenting history of maximum 30 years in 2010: each patent portfolio represents an inventor's cumulated inventions up to 2010.

PatentViews incorporates the identification of gender for the majority of the disambiguated inventors: details about the method employed for the identification are reported in USPTO (2019). The selected sample reports no information on gender for 9.1% of the inventors (details across sectors and categories are reported in the Appendix, in Table 12 and Table 13). The exclusion of inventors with no assigned gender brings the final sample to 640,043 individuals, of which 13.1% are female, a value in line with the literature (Bell et al. 2019; USPTO 2019).

b. Identification of star inventors

As discussed in the literature review, the identification of "stars" can entail different operationalizations of the criteria to distinguish outstanding from average performers. In general, to be a star, the individual must engage in disproportionately high accomplishments relative to most other workers in their field (Call et al., 2015; Aguinis and O'Boyle, 2014)[9]. The examined performance has been measured with different metrics ranging from productivity (Lahiri et al., 2019; Subramanian et al., 2013; Zucker et al., 2002; Kehoe and Tzabbar, 2015), to impact (Azoulay et al., 2010; Rothaermel and Hess, 2007) and, in some cases, visibility or celebrity (Oldroyd and Morris 2012).

Star individuals have been studied in several contexts[10] with particular attention to scientists/scholars (Azoulay, Zivin, and Wang 2010) and inventors (Hohberger 2016), thanks to data availability on output, namely articles and patents. Stars among scientists and scholars have been typically defined by considering either their productivity in terms of quantity of output, in most cases through the number of articles or patents, and in terms of a measure of quality of output, such as the received citations (Liu, 2014; Hohberger, 2016; Hess & Rothaermel, 2011), or a combination of the two (Kehoe & Tzabbar, 2015; Agrawal et al., 2017). Bibliometric data in patent documents make it possible to break down these two dimensions of quality (van Zeebroeck and van Pottelsberghe de la Potterie 2011; Federico Caviggioli, De Marco, et al. 2020). While counting the total number of patents by an inventor is quite straightforward, citations are usually subject to some discretionary choices: forward citations are the most commonly used measure of the technical value of a patent (van Zeebroeck and van Pottelsberghe de la Potterie 2011; Antonelli and Colombelli, 2011, 2015), and quality indicators typically considers only the citations occurring in the first five years after the filing of the patent, to account for the different time of exposure to the probability of receiving a citation (F. Caviggioli and Ughetto 2016).

In this study, we identify stars and apply both conceptual approaches: (i) prolific

[9] The extent to which achievement must be disproportional varies across studies (Call, Nyberg, and Thatcher 2015): some have used from one to three standard deviation differences (e.g. Hess and Rothaermel, 2011) , others have used the top of the examined sample, from 1 to 10% (e.g. Hohberger, 2016).

[10] For example, sport players (Chen and Garg 2018) and actors (Han and Ravid 2020).

inventors, defined by the number of patents that they produce; and (ii) high-quality inventors, involved in the creation of a large percentage of outstanding inventions that are mentioned and cited extensively by the innovation community.[11] The characteristics of each of the two are discussed below.

i. Prolific star inventors (quantity)

The first definition considers productivity in terms of quantity of outputs and is operationalized in the following way. First, the total number of granted patents was counted for all the inventors available in PatentsView (without any time restrictions). Since patent propensity is different across technological sectors, in the second step we ranked the inventors by the number of their granted patents within each of the 35 technological sectors from the WIPO concordance table. The 35 categories are based on IPC subclasses (4-digit IPC codes). Note that each patent can be associated to multiple IPC subclasses, and thus to multiple sectors. Hence, it is common that inventors are listed in more than one of the 35 sectors (73% in more than one, but only 10% in more than 5 fields).

The most prolific inventors are defined as those equal or above the 95th percentile of the distribution in their sector. Finally, we focus the analysis on the sample of inventors active in 2008-10 to avoid the comparison of inventors working in too different time periods. Since the global trend of patenting is increasing over time, in this sample of inventors active in 2008-10 we observe that prolific inventors >=95th percentile of their sector are 14.5% of our sample (Table 1). Female inventors represent 7.7% of all the stars while, among the non-stars, female are 14.0% of the sample: thus, there seems to be a further

[11] Patent quality can also be measured with generality and originality indexes, and through their geographical scope (Agostini et al. 2015; Lanjouw, Pakes, and Putnam 1998). These are however beyond the scope of this study.

under-representation of women among stars, at least in terms of averages.

Table 1 Share of outstanding inventors in terms of quantity (prolific)

Inventors	Male	Female	Total
Not star	73.5%	12.0%	*85.5%*
Stars (>=95th perc.)	13.4%	1.1%	*14.5%*
Total	*86.9%*	*13.1%*	*100%*

This definition based on quantity compares inventors with different career age and in different sectors, hence in the multivariate results reported in Section 5 we introduce career length as a relevant explanatory variable for productivity and control for sectoral fixed effects. Although the data do not report the exact start of an inventor's career, we use as a proxy the priority year of the first invention in an inventor's portfolio. The difference between the reference year 2010 and the debut year provides us with a proxy of career age (Duffy et al., 2011; König et al., 2015; Costas et al., 2015). Note that gender biases could affect the "debut" year in patenting for female inventors, which might thus occur later than for male. We report in the Appendix some summary statistics on the identification of prolific stars when discounting for the career length, which already indicate that this variable is relevant for women stars in particular (see Table 14 and Table 15).

ii. Inventors with a large share of highly-cited inventions (quality)

As an alternative approach, we define stars in terms of quality of their inventions. This approach is operationalized by considering the ratio of inventions within an inventor's portfolio that can be considered of "high quality". First of all, a single invention is considered of superior quality if it is equal or above the 95th percentile by number of citations received in the subsequent five years, with respect to a comparable set of inventions with the same priority year and in the same technological sector. Next, moving to the individual portfolio level, an inventor is considered outstanding if the share of high-quality inventions s/he generates is equal or above the 95th percentile of the focal sample. Table 2 illustrates the percentages for stars and non-stars and in terms of gender. 12.2% of the full sample is made of female non-stars and 0.9% of female high-quality stars. The gender gap seems similar in this and the previous definition of stars but, when focusing on this quality dimension, female inventors represent 14.3% of all the stars while, among the non-stars, female are 13.0% of the sample: compared to the previous results based on quantity, this finding suggests that the under-representation among stars is lower when focusing on quality.

Table 2 Share of outstanding inventors in terms of quality (highly-cited).

Inventors	**Male**	**Female**	**Total**
Not star	81.6%	12.2%	*93.8%*
Stars (>=95th perc.)	5.3%	0.9%	*6.2%*
Total	*86.9%*	*13.1%*	*100%*

However, this definition of inventor's quality is affected by portfolio size: with small portfolios, a single "hit" invention would automatically determine a high share of high-quality inventions. This is particularly relevant when considering inventors debuting in 2010 with a single "hit": they would automatically be considered stars in terms of inventions' quality. This distinction could be relevant also because women started patenting later historically, and so have smaller portfolios on average. For this reason, we replicate the previous analysis on a sample of inventors that produced at least five inventions (39% of the full sample). The 95th percentile corresponds to a share of outstanding inventions being at least 33% in the portfolio.

In this subsample of inventors with at least five inventions, high-quality inventors are 6.4% of the sample (Table 3), and female inventors represent 9.3% of all these stars, a share which is again higher than the 7.7% from the quantity-based definition. Among the non-stars, female are 9.5% of the sample. Therefore, even though this high-quality definition of stars is more restrictive, it confirms that the under-representation of women among stars when considering the quality perspective is smaller than when considering the quantity one.

Table 3 Share of outstanding inventors - only individuals with ≥5 inventions - quality.

Inventors	**Male**	**Female**	**Total**
Not star	84.6%	8.9%	*93.6%*
Stars (>=95th perc.)	5.8%	0.6%	*6.4%*
Total	*90.5%*	*9.5%*	*100%*

Appendix A3 compares the prolific and high quality definitions of stars and examines their relatively limited overlap.

c. Summary statistics on the gap between male and female star inventors

The average portfolio of patent families for female and male inventors is composed of 5.6 and 7.5 inventions respectively. Table 4**Errore. L'origine riferimento non è stata trovata.** shows that the number of female inventors is always substantially lower than males for all ranges of portfolio sizes, from innovators that only have one patent to those that have more than 100. However, the share of female innovators does not decline for larger patent portfolios, and on the contrary the second-highest share of female inventors is concentrated in large portfolios of 101 to 758 patent families. Similar patterns can be seen for patents' quantity weighted by career length (

Table 16 in the Appendix).

Table 4 Distribution of portfolio size (families per innovator).

N. of patents (range)		**M**	**F**	**% of total inventors in this range**	**% of *female* on inventors in this range**
1	1	159’720	33’761	30.2%	17.4%
2	5	200’093	2’9743	35.9%	12.9%
6	10	89’247	9’878	15.5%	10.0%
11	50	99’143	9’345	17.0%	8.6%
51	100	6’590	670	1.1%	9.2%
101	758	1’601	252	0.3%	13.6%

Next, we establish the sectoral differences in female representation among top inventors. Technological fields are derived from the WIPO IPC-Technology Concordance Table (last update in 2016). Each inventor was associated to one or more WIPO fields when it represented the largest share of families in the considered portfolio or at least 30% of portfolio. With this operationalization we are able to identify the main technological field(s) of activity of each inventor. Table 5**Errore. L'origine riferimento non è stata trovata.** shows that the share of female inventors changes across sectors, as abundantly found in the previous literature, and the variation among star inventors (prolific or high-quality ones) across the different sectors is similar. Therefore, while it is important to account for these sectoral differences, it is unlikely that stars differ greatly from non-star inventors in terms of specific sectoral patterns.

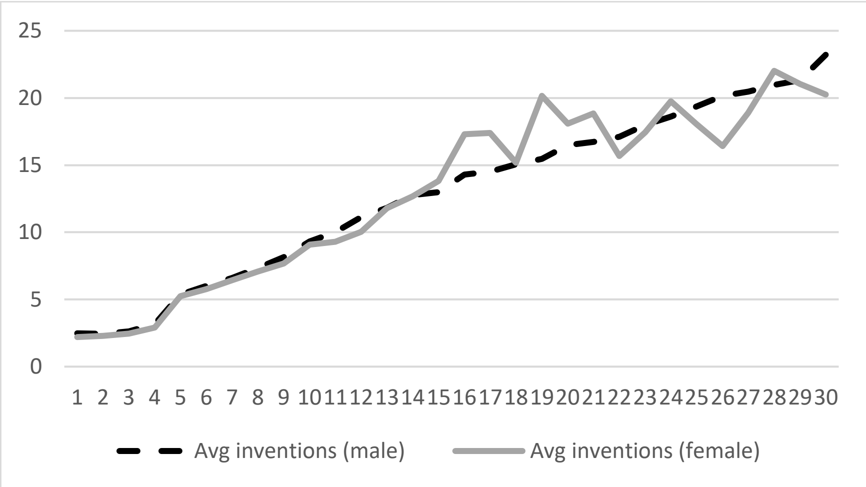
Table 5 Share of female inventors across sectors, for prolific and high-quality inventors.

WIPO field code	**WIPO field name**	**Full sample**			**Subsample of inventors with >=5 inventions**	
		N. of inventors	% of female inventors	%. of women among prolific inventors	% of female inventors	%. of women among highly-cited inventors
16	Pharmaceuticals	66,677	24.8	14.1	16.2	17.3
15	Biotechnology	56,739	24.6	15.8	15.3	16.2
14	Organic fine chemistry	73,297	22.1	12.7	15.2	17.0
18	Food chemistry	14,246	19.7	15.9	12.8	15.1
11	Analysis of biological materials	30,766	19.5	15.6	12.4	11.8
19	Basic materials chemistry	58,526	17.3	9.9	12.9	14.6
17	Macromolecular chemistry, polymers	36,475	15.7	8.3	11.8	13.4
22	Micro-structural and nano-technology	11,297	15.0	8.5	11.3	10.2
	Full sample	**640,043**	**13.1**	**7.7**	**9.5**	**9.3**
7	IT methods for management	38,369	12.4	10.3	8.4	6.1
13	Medical technology	69,200	12.4	6.2	8.1	7.6
21	Surface technology, coating	68,860	12.1	7.5	9.9	9.9
8	Semiconductors	69,893	12.1	10.3	10.7	9.1
34	Other consumer goods	35,134	11.9	8.4	7.9	7.6
20	Materials, metallurgy	35,809	11.1	7.9	8.9	9.5
6	Computer technology	196,386	10.9	7.7	8.9	6.7
9	Optics	67,200	10.8	7.4	9.4	8.3
2	Audio-visual technology	99,815	10.3	7.8	9.1	6.4
23	Chemical engineering	63,536	10.3	7.6	8.2	8.8
28	Textile and paper machines	32,719	10.2	6.9	8.1	9.3
33	Furniture, games	35,547	10.2	6.6	6.7	6.3
4	Digital communication	77,897	9.9	8.9	8.8	6.1
29	Other special machines	60,994	9.9	5.5	7.8	9.1
3	Telecommunications	87,447	9.8	8.6	8.8	6.9
1	Electrical machinery, apparatus, energy	109,544	9.6	8.1	8.3	6.6
10	Measurement	118,032	9.0	5.5	7.3	6.0
5	Basic communication processes	40,051	8.6	7.2	8.0	5.3
24	Environmental technology	27,579	8.6	4.8	6.7	8.3
12	Control	64,638	8.3	6.0	6.8	5.1
25	Handling	44,654	7.4	3.6	5.2	5.7
30	Thermal processes and apparatus	22,897	6.7	7.1	6.0	4.9
26	Machine tools	53,943	6.5	4.5	5.5	4.6
35	Civil engineering	34,847	6.5	3.3	5.4	4.6
32	Transport	56,717	5.9	2.9	4.6	5.1
27	Engines, pumps, turbines	45,514	5.9	3.9	5.0	4.9
31	Mechanical elements	51,552	5.2	2.9	4.3	5.0

Notes: the analysis on highly-cited is carried out on the subsample of those with at least five inventions. Sectors ranked in decreasing order of the percentage of female inventors.

Finally, we consider whether gender differences exist in terms of productivity along the career trajectory of male and female innovators. **Errore. L'origine riferimento non è stata trovata.** illustrates how the output of male inventors grows steadily and smoothly over time, while for female inventors the average number of patents is slightly below the one of man for the first 10 years of career (since the beginning of patenting), with a more marked negative difference in the following few years, and then with a marked increase around 15-20 years into women's career, albeit with higher volatility in average output. While a longitudinal analysis of career progression vis-à-vis patenting is beyond the scope of this study, this graph emphasizes the different career paths experienced by gender, and thus calls for careful consideration of career length in our analysis.

Figure 1 Distribution of the average number of inventions (y-axis) by career length of inventors (x-axis): dashed line for male inventors, continuous line for female inventors.



This pattern is further confirmed looking at specific career brackets (Table 6**Errore. L'origine riferimento non è stata trovata.**). Once again, there is evidence of a relatively notable increase in the patent productivity of women after 15 years of career.

Table 6 Summary statistics on productivity in terms of number of inventions broken down by career length (5-years categories).

Career length	Gender	Obs	Mean	Std. Dev.	Min	Max
Up to 5 years	male	265,875	2.972596	5.06256	1	387
	female	50,356	2.7416	5.175256	1	321
From 5 to 10 years	male	120,581	7.387872	11.24599	1	549
	female	18,584	7.05085	11.90288	1	327
From 10 to 15 years	male	75,207	11.5180	16.76582	1	758
	female	8,678	11.0436	17.68406	1	290
From 15 to 20 years	male	46,167	15.0726	20.74264	1	738
	female	3,724	17.4989	34.00939	1	595
Over 20 years	male	48,564	19.0375	22.98041	1	610
	female	2,307	18.2557	30.49225	1	713

Lastly, to further assess the presence of potential differences in the share of female and male inventors that achieve the status of “star” with our different measures of quantity and quality of innovations, we define a measure of concentration similar to a “glass ceiling” index. More precisely, “female concentration” is defined as the share of female stars among all stars, divided by the share of female inventors among all inventors. Male concentration is the equivalent for men. A concentration close to one indicates that the representation of a given gender among stars is similar to the one among inventors in general. For males, concentration among stars is close to one in all types of operationalization of stars (Table 7**Errore. L'origine riferimento non è stata trovata.**). On the contrary, for women it is definitely below one when considering the quantity dimension, confirming the under-representation, while the quality-based measure, in both examined samples, is close or above one (Table 7): suggesting that the under-representation is relatively lower when introducing the quality dimension. For comparison with the quantity-based measure, Abramo et al. (2009) found similar values for a sample of Italian academicians with full professorship: 1.14 for male and 0.57 for female.

Table 7 Share of inventors in different samples of stars.

Definition of star	**Share of women among stars**	**Reference sample**	**Female concentration**	**Male concentration**
Prolific	7.7%	Full sample (13.1% female)	0.59	1.06
Highly-cited	14.3%	Full sample (13.1% female)	1.10	0.99
Highly-cited	9.3%	Portfolios with >=5 patent families (9.5% female)	0.98	1.00

Notes: all definitions consider the 95th percentile of the distribution as a cut-off to be considered a star

4. Methodology

All summary statistics presented in the previous section are useful to characterize the unconditional distribution of women and men in the overall population of inventors and among stars, however they do not control simultaneously for the career effects on productivity, for sectoral differences, and for the large geographic predominance of US inventors. It is therefore important to analyse the gap between men and women inventors in a more rigorous framework that includes all these factors for stars and non-stars. To approach the question of gender gaps in a multivariate econometric setting, we implements a set of models based on a Poisson estimator for the quantity measure, where the dependent variable is the count of patent families (namely, portfolio size) at the inventor level *i*, and a fractional response model for the quality measure, where the dependent variable is the share of highly-cited patent families in the portfolio. The baseline model is

$$N\_p_i = e^\wedge(\beta_0 + \beta_1 Female_i + \beta_2 CareerLength_i + \beta_3\, Female_i\, x CareerLength_i + X_i\gamma)$$

Where N_p_i is the individual number of patents, with a Poisson distribution, or the share of highly cited patents in an inventor's portfolio, with a logit distribution. The independent variables of interest are the same for both models: a Female dummy for the gender of the inventor (equals one for women); Career length, a continuous variable that captures the number of years between the first patent and 2010; and the interaction of the Female dummy x Career length variables. Furthermore, we include in X individual-level controls for geographical origin, namely a dummy equal to one if the inventor shows a patent filed in the US as first earliest priority; and technological field fixed effects, dummies for the sector of activity of each inventor (note that these are not mutually exclusive, as an inventor can be active in more than one sector), in some specifications also interacted with the Female dummy. For a definition and summary of each of the variables, see Table 11 in the Appendix. While this multivariate approach allows for a more careful comparison of female and male inventors controlling for relevant variables, it is not able to capture any dynamic effects that can characterize the path-dependent evolution of productivity, since all our measures are time invariant. An interesting extension for future research could consider the evolution of productivity over time at the different stages of career development for men and women.

5. Results

The results of the multivariate regression analysis allows us to correlate more formally the fact of being female and the probability of being a star, according to the different definitions. In these regressions, the interaction of career length with female gender is key to inform our understanding of productivity for women accounting for different career spans.

a. Quantity of inventions and gender

Regarding the quantity of patents produced by an inventor, Table 8 presents the results of the Poisson regression as incident rate ratios for all inventors (first and second model) and then only for prolific stars (third model). The second model adds to the specification the interaction of all sectors and the Female dummy. The third column focuses on the subsample of star inventors according to the quantity definition, namely with a portfolio size in the top 95th percentile in any technological field.

Table 8 Poisson regression as incident rate ratios.

Model	**(1)**	**(2)**	**(3)**
Sample	Full	Full	Only prolific star inventors (>=95th percentile)
Dependent variable	Portfolio size		
Female dummy	0.8145+	0.8001+	1.1342+
	(0.0021)	(0.0033)	(0.0097)
Career length	1.0804+	1.0803+	1.0187+
	(0.0001)	(0.0001)	(0.0001)
Female dummy x Career length	1.0200+	1.0202+	1.0050+
	(0.0002)	(0.0002)	(0.0004)
Geographical origin (US=1)	0.8272+	0.8287+	0.7793+
	(0.0008)	(0.0008)	(0.0011)
Main technological field dummies	Y	Y	Y
Female dummy x main tech. field dummies		Y	Y
Observations	639860	639860	92979
PseudoR2	0.2774	0.2782	0.1636
loglik.	-3321588.3481	-3317798.5099	-874303.2559

Notes: With incident rate ratios, any result below one indicates a negative relation, above one a positive relation. The dependent variable is the quantity of patent families in an inventor's portfolio. * p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01; + p-value < 0.001

The incident rate ratios shown in the table are derived from the Poisson regression coefficients, interpreted as the difference between the log of expected counts. Formally, this can be written as $\beta = \log(\mu_x+1) - \log(\mu_x)$, where β is the regression coefficient, μ is the expected count and the subscripts represent where the predictor variable, say the Female dummy, is evaluated at 0 and 1 (implying a one unit change in the predictor variable). Since the difference of two logs is equal to the log of their quotient, $\log(\mu_x+1) - \log(\mu_x) = \log(\mu_x+1 / \mu_x)$, we can interpret the parameter estimate as the log of the ratio of expected counts (this explains the term "ratio" in incidence rate ratios). In practice, we obtain at the incidence rate ratio by exponentiating the Poisson regression coefficient. So, for example, the first coefficient for Female dummy was -0.20518, which translates into an incident rate ratio smaller than one of exp(-0.20518)= 0.8145.

These results show that, while being a woman constitutes a clear disadvantage in the pool of all inventors, it is not the case for prolific stars. The Female dummy has a significant negative impact on portfolio size when considering all inventors (it reduces the chances for an additional invention by 20% in model 2), but a significant positive effect when considering only those inventors equal or above the 95th percentile in the number of patents in their portfolio (+13% in model 3). Therefore, being a woman actually increases the number of patents in this group of stars. Career length is always positively linked to portfolio size, even though the magnitude of the effect is small: +8% of probability for any additional year to file an additional invention in the full sample (model 2) and only +2% in the sample of prolific stars (model 3). The interaction of being a woman with career length has a positive and significant effect in all cases, although small: each additional year since the first patent adds to the probability of an extra patent +2% in model 2 and +0.5% in model 3. This means that, for a woman, any year of career increases the number of patents she generates more than for a man in a statistically significant way.

b. Quality of inventions and gender

The second set of multivariate analyses focuses on the quality of the portfolios generated by each inventor. We investigated the relationship between the share of highly-cited patent families in the portfolio and gender and career length. Table 9 illustrates the results without applying any restrictions on the minimum portfolio size, while Table 10 focuses on the inventors with at least five inventions.

As before, the first and second models are tested on the full sample of inventors, adding further sectoral interactions as fixed effects in the second model, while the third one considers only the top performers in terms of highly-cited star inventors.

Table 9 Results of fractional response models using a logit for the conditional mean.

Model	**(1)**	**(2)**	**(3)**
Sample	Full	Full	Only star inventors (>=95th percentile)
Dependent variable	Share of high-quality inventions in the portfolio		
Female dummy	0.8380+	0.8587+	1.1576***
	(0.0145)	(0.0237)	(0.0608)
Career length	1.0254+	1.0255+	0.9509+
	(0.0004)	(0.0004)	(0.0009)
Female dummy X Career length	1.0248+	1.0242+	1.0058*
	(0.0015)	(0.0015)	(0.0034)
Geographical origin (US=1)	2.2644+	2.2658+	1.0815+
	(0.0160)	(0.0160)	(0.0150)
Main technological field dummies	Y	Y	Y
Female dummy x main tech. field dummies		Y	Y
Observations	639860	639860	39842
PseudoR2	0.0482	0.0484	0.0180
loglik.	-172690.3730	-172646.2176	-21702.5482

Notes: Results as odds ratios indicate a negative relationship when the coefficient is below one, a positive relationship if above one. The dependent variable is the share of high-quality inventions in the total portfolio. * p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01; + p-value < 0.001.

Once again, women's disadvantage in patenting disappears when considering the Female dummy in high-quality star inventors. In the full sample, being a female inventor is associated to a smaller proportion of outstanding inventions in the portfolio (-14% in model 2). However, being a woman increases the share of high-quality patents in an inventor's portfolio, when considering only the sample of high-quality stars (+16% in model 3).

Additional years of career show a small impact on the share of high-quality inventions: +2.5% in the full sample and -5% among stars only. This result suggests that, for high shares of high-quality inventions in the portfolio, additional years are more likely to generate non-outstanding inventions. The female premium on career length observed for portfolio size in the previous set of models (Table 8) is present also in the case of the share of high-quality patents, with similar magnitudes: +2.5% and +0.5% in model 2 and

3 respectively.

These first two sets of results indicate a robust difference between non-star women inventors and stars, and a positive role of additional years of career that give women an added benefit both for producing more patents and to increase the share of high-quality patents. However, as discussed in Section 3.b.ii, some inventors included in this latter analysis might have relatively small portfolios. Thus, it is important to consider how these dynamics change for inventors with more substantial patent portfolios. Hence, we repeat the analyses and consider only inventors with more than five inventions (Table 10). In the Appendix we provide some sensitivity tests when employing different thresholds, with three and seven inventions as minimum portfolio size (Table 21): results are similar to the ones for the sample of five inventions as minimum portfolio size.

Table 10 Results of fractional response models using a logit for the conditional mean selecting the inventors with at least five inventions in their portfolio.

Model	**(1)**	**(2)**	**(3)**
Sample	>=5 inventions	>=5 Inventions	Only star inventors among those with >=5 inventions
Dependent variable	Share of high-quality invention in the portfolio		
Female dummy	0.9454**	1.0184	1.0598
	(0.0224)	(0.0368)	(0.0504)
Career length	1.0102+	1.0104+	0.9938+
	(0.0005)	(0.0005)	(0.0006)
Female dummy X Career length	1.0036**	1.0020	0.9994
	(0.0017)	(0.0018)	(0.0024)
Geographical origin (US=1)	2.0610+	2.0603+	1.0597+
	(0.0144)	(0.0145)	(0.0096)
Main technological field dummies	Y	Y	Y
Female dummy x each main tech. field dummy		Y	Y
Observations	251481	251481	16212
PseudoR2	0.0317	0.0319	0.0012
loglik.	-68828.3078	-68818.3810	-11164.2518

Notes: Results as odds ratios indicate a negative relationship when the coefficient is below one, a positive relationship if above one. The dependent variable is the share of high-quality inventions in the total portfolio. * p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01; + p-value < 0.001

The results on this sample of inventors with at least five inventions in their portfolio show no significant gender difference either in average quality of the output or in terms of additional year of career (models 2 and 3).[12] As seen already in Table 9, the significant result that differs in the group of stars is that career length does not provide a premium, but rather a small disadvantage for high-quality stars (+1% and -0.6% in model 2 and 3 respectively), even when considering only those with portfolios of more than five inventions. However, this difference does not seem to apply to men and women differently.

There are three main implications of these results. First, despite the usual negative gap for women inventors in terms of productivity, we find that, among stars, women are more productive both in terms of quantity of patents and of the share of “hits” (high-quality patents). While the average negative productivity gap confirms the presence of a gender glass-ceiling in patenting, which prevents most women from accessing resources and opportunities to generate the same innovative output of men, this barrier can be broken by outstanding female inventors. In fact, the results for stars suggest a selection mechanism, with female inventors being on average more productive than males in terms of quantity and quality.

Secondly, when focusing on the very top of the distribution, for female star inventors with large portfolios (with at least five inventions), there is no longer any gendered difference in terms of average quality. Thus, both the negative glass ceiling of average female inventors and the positive selection effect of star women inventors disappear once we look at this sub-group of stars with large portfolios. Presumably, women are more efficient in achieving “hits” in relatively small portfolios, which is quite reasonable given that women have started patenting later historically and have been catching up with men

[12] Without the controls for female and sectoral interactions (first model of Table 10), we still have a gender gap and a career premium for women, but of lower significance and magnitude than in all pervious specifications.

only recently. In sum, (i) being a *woman innovator* on average is linked to a patenting disadvantage (glass ceiling); (ii) being a *woman star innovator* is associated with a significant advantage (selection effect with stars "breaking the glass ceiling"); and (iii) being a *woman star innovator with large portfolio* of more than 5 patents does not constitute any significant difference from male star innovators. This implies that the gender gap in patenting is not constant across types of inventors and should not be addressed uniformly by policy-makers.

The third implication of our results is that women seem to enjoy a career premium on each additional year of career since their first patent, both to expand their portfolio (quantity) and to increase its value (quality). Every additional year available to an innovator carries greater value for women than for men, ceteris paribus. Since we define career length as the time starting from the first patent, this result does not include the fact that women might take longer to even begin patenting, in the first place. Therefore, our results on the importance of career duration for women may capture only a lower bound of the effect.

6. Conclusion

In this study, we explore how different definitions of top levels of productivity in patenting relate to gender gaps in innovation. Applying measures of quantity and quality of patenting output, we observe robust evidence that star innovators are different from the overall pool of innovators: while there is a negative gender gap in patenting for women, a selection effect ensures that among the top percentiles of innovators being a woman is actually associated with more patents and a higher share of highly cited patents. This gendered effect disappears when we consider quality for star inventors with more than five patents. In that case, being a woman constitutes neither an advantage nor a disadvantage, and stars are not different from other inventors.

Furthermore, we find that career length plays a significant role in countering the disadvantage that women face in becoming innovators or in enhancing the female advantage among stars. There is generally a premium for each additional year of career after the first patent, but this premium is much larger for the overall pool of innovators and less dramatic for stars. Once again, there is no gendered premium for being a woman and having extra years of career in the group of inventors with more than five inventions. Further research could expand on these results, examining the time dimension of career progression and productivity, using longitudinal measures of individual productivity to follow patenting dynamics over time. One important caveat that can further clarify our results on career length is that our analysis does not examine the process of entry into these scientific fields and the time before the first patent as part of career length of an innovator. Therefore, if women take longer to achieve the first patent filing, this is not captured in our analysis.

Future lines of analysis should characterize further these results, for example examining how the temporal career dimension interacts with interpersonal networks, as well: patents are often the outcome of team efforts in which mentoring, peer support, competition for resources and spillovers between scientists all affect the success of individuals. While our focus of analysis is exclusively on individual inventors, evidence-based policy in support of women scientists should encourage further analysis of how gender gaps relate to the interaction between star and non-star inventors. Additionally, the sectoral and geographic dimension of patenting are not unpacked in detail in the present study, and more vertical in-depth analyses of specific sectors could provide further insights in the different disciplinary boundaries of gender gaps among stars. One important limitation of our analysis is also the focus on US granted patents, and the definition of geographic origin of inventors on the basis of the first country of patent filing. While the US remains one of the largest markets for patent production, more studies of patenting in other geographic regions are highly needed, not only in Europe but also in emerging markets like China, currently the country that grants most patents in the world (WIPO 2021, pp.32).

Overall, our results indicate the relevance of integrating a gender perspective into innovation policies. The core implications of our results derive from the finding that the selection effect for the most competitive and motivated women dominates over glass ceilings, gender discrimination and disadvantages. This is not to say that women should not be supported in their career progression, considering the possible delays that they face due to greater household and family responsibilities: in fact, among the pool of all innovators, there is still a significant gender gap. Our results indicate that the most appropriate policy strategy should account for the different directions of the gender gap identified, first and foremost by addressing the gender disadvantage in the general population of innovators, facilitating the access to resources, support and opportunities for women. Also, our results show that the glass-ceiling is broken by some of the most capable and motivated female innovators who achieve superior productivity. This result suggest that gender equality policies should focus on the overall population of inventors rather than "picking winners" or concentrating on the top of the distribution, because star female inventors already succeed at being more productive than men.

Moreover, policymakers can leverage the career effects identified in our analysis, allowing more time and long-term support for women to develop their patent portfolios and close any gender gap. Policymakers interested in more female participation among innovators and in creating a more level playing field should consider the added value from each extra year of career progression for women when designing the regulations and incentives for public support towards greater gender equality.

In parallel, to address the career-driven gender gap in patenting, innovation policies with a gender perspective could be reinforced by labour, welfare and equal opportunities policies. For example, considering that both low female participation in innovation and the delays in careers that female face can be in part explained by greater household and family responsibilities, all the measures aiming at increasing female labour market participation (like the extension of paid paternity leave for fathers and other measures to narrowing down the gender care gap as well as the increase of childcare services and family subsidies) and reducing the gender pay gap to make the inventor career more attractive to women could also be effective. In turn, also a gender perspective in education policies would reinforce all previous policies not just for stimulating more female participation in STEM studies but also to question existing prejudices and perceptions about women and men's respective family responsibilities and duties.

Acknowledgments

We are grateful for the comments and suggestions of Filippo Chiarello and the participants to the XXXII RSA AiIG conference.

Declaration of interest statement

No potential competing interest was reported by the authors.

References

Abramo, Giovanni, Ciriaco Andrea D'Angelo, and Alessandro Caprasecca. 2009. "The Contribution of Star Scientists to Overall Sex Differences in Research Productivity." *Scientometrics* 81 (1): 137–156. doi:10.1007/s11192-008-2131-7.

Agostini, L., F. Caviggioli, R. Filippini, and A. Nosella. 2015. "Does Patenting Influence SME Sales Performance? A Quantity and Quality Analysis of Patents in Northern Italy." *European Journal of Innovation Management* 18 (2). doi:10.1108/EJIM-07-2013-0071.

Agrawal, Ajay, John McHale, and Alexander Oettl. 2017. "How Stars Matter: Recruiting and Peer Effects in Evolutionary Biology." *Research Policy* 46 (4): 853–867. doi:10.1016/j.respol.2017.02.007.

Aguinis, Herman, and Ernest O'Boyle. 2014. "Star Performers in Twenty-First Century Organizations." *Personnel Psychology* 67 (2): 313–350. doi:10.1111/peps.12054.

Akbaritabar, Aliakbar, Niccolò Casnici, and Flaminio Squazzoni. 2018. "The Conundrum of Research Productivity: A Study on Sociologists in Italy." *Scientometrics* 114 (3): 859–882. doi:10.1007/s11192-017-2606-5.

Antonelli C., Colombelli A. (2011), The generation and exploitation of technological change: market value and total factor productivity, *The Journal of Technology Transfer*, 36 (4), 353-382, doi:10.1007/s10961-010-9198-z.

Antonelli C., Colombelli A. (2015), The knowledge cost function, International Journal of Production Economics, 168, 290-302, doi:10.1016/j.ijpe.2015.06.031

Arruda, Denis, Fábio Bezerra, Vânia Almeida Neris, Patricia Rocha de Toro, and Jacques Wainera. 2009. "Brazilian Computer Science Research: Gender and Regional Distributions." *Scientometrics* 79 (3): 651–665. doi:10.1007/s11192-007-1944-0.

Azoulay, Pierre, Joshua S.Graff Zivin, and Jialan Wang. 2010. "Superstar Extinction." *Quarterly Journal of Economics* 125 (2): 549–589. doi:10.1162/qjec.2010.125.2.549.

Baccini, A., L. Barabesi, M. Cioni, and C. Pisani. 2014. "Crossing the Hurdle: The Determinants of Individual Scientific Performance." *Scientometrics* 101 (3): 2035–2062. doi:10.1007/s11192-014-1395-3.

Barwa, Sharmishta, and Shirin M. Rai. 2003. "Knowledge and/as Power: A Feminist Critique of Trade Related Intellectual Property Rights." *Gender Technology And Development* 7 (1): 91–113. doi:10.1177/097185240300700105.

Beaudry, Catherine, and Vincent Larivière. 2016. "Which Gender Gap? Factors Affecting Researchers' Scientific Impact in Science and Medicine." *Research Policy* 45 (9): 1790–1817. doi:10.1016/j.respol.2016.05.009.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen. 2019. "Who Becomes an Inventor in America? The Importance of Exposure to Innovation." *Quarterly Journal of Economics* 134 (2): 647–713. doi:10.1093/qje/qjy028.

Bendersky, Corinne, and Nicholas A. Hays. 2012. "Status Conflict in Groups." *Organization Science* 23 (2): 323–340. doi:10.1287/orsc.1110.0734.

Bettio, Francesca, and Alina Verashchagina. 2009. *Gender Segregation in the Labour Market*. http://bookshop.europa.eu/is-bin/INTERSHOP.enfinity/WFS/EU-Bookshop-Site/en_GB/-/EUR/ViewPublication-Start?PublicationKey=KE7809682.

Blau, Francine D., and Lawrence M. Kahn. 2017a. “The Gender Wage Gap: Extent, Trends, and Explanations.” *Journal of Economic Literature* 55 (3): 789–865. doi:10.1257/jel.20160995.

Blau, Francine D., and Lawrence M. Kahn. 2017b. “The Gender Wage Gap: Extent, Trends, and Explanations.” *Journal of Economic Literature* 55 (3): 789–865. doi:10.1257/jel.20160995.

Borrego, Ángel, Maite Barrios, Anna Villarroya, and Candela Ollé. 2010. “Scientific Output and Impact of Postdoctoral Scientists: A Gender Perspective.” *Scientometrics* 83 (1): 93–101. doi:10.1007/s11192-009-0025-y.

Call, Matthew L., Anthony J. Nyberg, and Sherry M.B. Thatcher. 2015. “Stargazing: An Integrative Conceptual Review, Theoretical Reconciliation, and Extension for Star Employee Research.” *Journal of Applied Psychology* 100 (3): 623–640. doi:10.1037/a0039100.

Caviggioli, F., and E. Ughetto. 2016. “Buyers in the Patent Auction Market: Opening the Black Box of Patent Acquisitions by Non-Practicing Entities.” *Technological Forecasting and Social Change* 104. doi:10.1016/j.techfore.2015.11.031.

Caviggioli, Federico, Alessandra Colombelli, Antonio De Marco, and Emilio Paolucci. 2020. “How Venture Capitalists Evaluate Young Innovative Company Patent Portfolios: Empirical Evidence from Europe.” *International Journal of Entrepreneurial Behaviour and Research* 26 (4): 695–721. doi:10.1108/IJEBR-10-2018-0692.

Caviggioli, Federico, Antonio De Marco, Fabio Montobbio, and Elisa Ughetto. 2020. “The Licensing and Selling of Inventions by US Universities.” *Technological Forecasting and Social Change* 159. doi:10.1016/j.techfore.2020.120189.

Caviggioli, Federico, and Boris Forthmann. 2022. “Reach for the Stars: Disentangling Quantity and Quality of Inventors’ Productivity in a Multifaceted Latent Variable Model.” *Scientometrics*. doi:forthcoming.

Caviggioli, Federico, Giuseppe Scellato, and Elisa Ughetto. 2020. “Lenders’ Selection Capabilities, Patent Quality, and the Outcome of Patent-Backed Loans.” *Industrial and Corporate Change* 29 (1): 43–60. doi:10.1093/icc/dtz029.

Cech, Erin A. 2013. “The Self-Expressive Edge of Occupational Sex Segregation.” *American Journal of Sociology* 119 (3): 747–789. doi:10.1086/673969.

Charles, Maria, and Karen Bradley. 2002. “Equal but Separate? A Cross-National Study of Sex Segregation in Higher Education.” *American Sociological Review* 67 (4): 573–599. doi:10.2307/3088946.

Chen, John S., and Pranav Garg. 2018. “Dancing with the Stars: Benefits of a Star Employee’s Temporary Absence for Organizational Performance.” *Strategic Management Journal* 39 (5): 1239–1267. doi:10.1002/smj.2758.

Colombelli, A., Grilli, L., Minola, T., & Mrkajic, B. (2020). To what extent do young innovative companies take advantage of policy support to enact innovation appropriation mechanisms?. Research Policy, 49(10), 103797.

Colombelli, A. Grinza, E., Meliciani, V. & Rossi, M. (2021) Pulling Effects in Immigrant Entrepreneurship: Does Gender Matter?, Economic Geography, 97:1, 1-33, DOI: 10.1080/00130095.2021.1874242

Costas, Rodrigo, Tina Nane, and Vincent Larivière. 2015. "Is the Year of First Publication a Good Proxy of Scholars' Academic Age?" *Proceedings of ISSI 2015 Istanbul: 15th International Society of Scientometrics and Informetrics Conference*, 988–998.

Duffy, Ryan D., Alex Jadidian, Gregory D. Webster, and Kyle J. Sandell. 2011. "The Research Productivity of Academic Psychologists: Assessment, Trends, and Best Practice Recommendations." *Scientometrics* 89 (1): 207–227. doi:10.1007/s11192-011-0452-4.

Fischer, Timo, and Philipp Ringler. 2014. "What Patents Are Used as Collateral?-An Empirical Analysis of Patent Reassignment Data." *Journal of Business Venturing* 29 (5). Elsevier Inc.: 633–650. doi:10.1016/j.jbusvent.2014.04.002.

Forthmann, Boris, Mark Leveling, Yixiao Dong, and Denis Dumas. 2020. "Investigating the Quantity–Quality Relationship in Scientific Creativity: An Empirical Examination of Expected Residual Variance and the Tilted Funnel Hypothesis." *Scientometrics* 124 (3). Springer International Publishing: 2497–2518. doi:10.1007/s11192-020-03571-w.

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104 (4): 1091–1119. doi:10.1257/aer.104.4.1091.

Groysberg, Boris, and Linda Eling Lee. 2009. "Hiring Stars and Their Colleagues: Exploration and Exploitation in Professional Service Firms." *Organization Science* 20 (4): 740–758. doi:10.1287/orsc.1090.0430.

Groysberg, Boris, Jeffrey T. Polzer, and Hillary Anger Elfenbein. 2011. "Too Many Cooks Spoil the Broth: How High-Status Individuals Decrease Group Effectiveness." *Organization Science* 22 (3): 722–737. doi:10.1287/orsc.1100.0547.

Han, Shu, and S. Abraham Ravid. 2020. "Star Turnover and the Value of Human Capital-Evidence from Broadway Shows." *Management Science* 66 (2): 958–978. doi:10.1287/mnsc.2018.3177.

Haseltine, FlorenceP., and Mark Chodos. 2017. "'Why' vs. 'What,' or 'The Bad Penny Opera': Gender and Bias in Science." *Technology & Innovation* 18 (4): 275–279. doi:10.21300/18.4.2017.275.

Heikkilä, Jussi. 2019. "IPR Gender Gaps: A First Look at Utility Model, Design Right and Trademark Filings." *Scientometrics* 118 (3). Springer International Publishing: 869–883. doi:10.1007/s11192-018-2979-0.

Hess, Andrew M., and Frank T. Rothaermel. 2011. "When Are Assets Complementary? Star Scientists, Strategic Alliances, and Innovation in the Pharmaceutical Industry." *Strategic Management Journal* 32 (8): 895–909. doi:10.1002/smj.916.

Hohberger, Jan. 2016. "Does It Pay to Stand on the Shoulders of Giants? An Analysis of the Inventions of Star Inventors in the Biotechnology Sector." *Research Policy* 45 (3): 682–698. doi:10.1016/j.respol.2015.12.003.

Hoisl, Karin, and Myriam Mariani. 2017. "It's a Man's Job: Income and the Gender Gap in Industrial Research." *Management Science* 63 (3): 766–790. doi:10.1287/mnsc.2015.2357.

Hunt, Jennifer, Jean Philippe Garant, Hannah Herman, and David J. Munroe. 2013. "Why Are Women Underrepresented amongst Patentees?" *Research Policy* 42 (4). Elsevier B.V.: 831–843. doi:10.1016/j.respol.2012.11.004.

Hunter, Laura A., and Erin Leahey. 2010. "Parenting and Research Productivity: New Evidence and Methods." *Social Studies of Science* 40 (3): 433–451. doi:10.1177/0306312709358472.

Joy, Stephen. 2006. "What Should I Be Doing, and Where Are They Doing It? Scholarly Productivity of Academic Psychologists." *Perspectives on Psychological Science* 1 (4): 346–364. doi:10.1111/j.1745-6916.2006.00020.x.

Kehoe, Rebecca R., and Daniel Tzabbar. 2015. "Lighting the Way or Stealing the Shine? An Examination of the Duality in Star Scientists' Effects on Firm Innovative Performance." *Strategic Management Journal* 36 (5): 709–727. doi:10.1002/smj.2240.

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar. 2019. "Who Profits from Patents? Rent-Sharing at Innovative Firms*." *The Quarterly Journal of Economics* 134 (3): 1343–1404. doi:10.1093/qje/qjz011.

König, Cornelius J., Clemens B. Fell, Linus Kellnhofer, and Gabriel Schui. 2015. "Are There Gender Differences among Researchers from Industrial/Organizational Psychology?" *Scientometrics* 105 (3): 1931–1952. doi:10.1007/s11192-015-1646-y.

Kozlowski, Diego, Vincent Larivière, Cassidy R. Sugimoto, and Thema Monroe-White. 2022. "Intersectional Inequalities in Science." *Proceedings of the National Academy of Sciences* 119 (2): e2113067119. doi:10.1073/pnas.2113067119.

Kretschmer, Hildrun, and Theo Kretschmer. 2013. "Gender Bias and Explanation Models for the Phenomenon of Women's Discriminations in Research Careers." *Scientometrics* 97 (1): 25–36. doi:10.1007/s11192-013-1023-7.

Kwiek, Marek. 2016. "The European Research Elite: A Cross-National Study of Highly Productive Academics in 11 Countries." *Higher Education* 71 (3). Springer Netherlands: 379–397. doi:10.1007/s10734-015-9910-x.

Kyvik, Svein, and Mari Teigen. 1996. "Child Care, Research Collaboration, and Gender Differences in Scientific Productivity." *Science Technology and Human Values* 21 (1): 54–71. doi:10.1177/016224399602100103.

Lacetera, Nicola, Iain M. Cockburn, and Rebecca Henderson. 2004. "Do Firms Change Capabilities By Hiring New People? A Study of the Adoption of Science-Based Drug Discovery." *Advances in Strategic Management* 21: 133–159. doi:10.1016/S0742-3322(04)21005-1.

Lahiri, Amrita, Emily C. Pahnke, Michael D. Howard, and Warren Boeker. 2019. "Collaboration and Informal Hierarchy in Innovation Teams: Product Introductions in Entrepreneurial Ventures.” *Strategic Entrepreneurship Journal* 13 (3): 326–358. doi:10.1002/sej.1331.

Lai, Jessica C. 2020. “The Role of Patents as a Gendered Chameleon.” *Social and Legal Studies*. doi:10.1177/0964663920916237.

Lanjouw, Jean O., Ariel Pakes, and Jonathan Putnam. 1998. “How to Count Patents and Value Intellectual Property: The Uses of Patent Renewal and Application Data.” *Journal of Industrial Economics* 46 (4): 405–432. doi:10.1111/1467-6451.00081.

Larivière, Vincent, Chaoqun Ni, Yves Gingras, Blaise Cronin, and Cassidy R. Sugimoto. 2013. “Bibliometrics: Global Gender Disparities in Science.” *Nature* 504 (7479): 211–213.

Lax Martinez, Gema, Julio Raffo, and Kaori Saito. 2016. *Identifying the Gender of PCT Inventors*. 33. *World Intellectual Property Organization Economic Research Working Paper*.

Leahey, Erin, and Amelia Blume. 2017. “Elucidating the Process: Why Women Patent Less than Men.” In *Gender and Entrepreneurial Activity*, 151–167. doi:10.4337/9781785364747.00011.

Lerchenmueller, Marc J., and Olav Sorenson. 2018. “The Gender Gap in Early Career Transitions in the Life Sciences.” *Research Policy* 47 (6): 1007–1017. doi:10.1016/j.respol.2018.02.009.

Liu, Kun. 2014. “Human Capital, Social Collaboration, and Patent Renewal Within U.S. Pharmaceutical Firms.” *Journal of Management* 40 (2): 616–636. doi:10.1177/0149206313511117.

Loan, Fayaz Ahmad, and Manzoor Hussain. 2017. “Gender Variations in Research Productivity: Insights from Scholarly Research.” *Library Philosophy and Practice* 2017 (November).

Lockwood, Penelope, and Ziva Kunda. 1997. “Superstars and Me: Predicting the Impact of Role Models on the Self.” *Journal of Personality and Social Psychology* 73 (1): 91–103. doi:10.1037/0022-3514.73.1.91.

Martínez, Catalina. 2011. “Patent Families: When Do Different Definitions Really Matter?” *Scientometrics* 86 (1): 39–63. doi:10.1007/s11192-010-0251-3.

Mayer, Sabrina J., and Justus M.K. Rathmann. 2018. “How Does Research Productivity Relate to Gender? Analyzing Gender Differences for Multiple Publication Dimensions.” *Scientometrics* 117 (3). Springer International Publishing: 1663–1693. doi:10.1007/s11192-018-2933-1.

Nielsen, Mathias Wullum. 2016. “Gender Inequality and Research Performance: Moving beyond Individual-Meritocratic Explanations of Academic Advancement.” *Studies in Higher Education* 41 (11): 2044–2060. doi:10.1080/03075079.2015.1007945.

OECD Patent Statistics Manual. 2009. OECD Patent Statistics Manual. doi:10.1787/9789264056442-en.

Oldroyd, James B., and Shad S. Morris. 2012. “Catching Falling Stars: A Human Resource Response to Social Capital’s Detrimental Effect of Information Overload on Star Employees.” *Academy of Management Review* 37 (3): 396–418.

doi:10.5465/amr.2010.0403.

Puuska, Hanna Mari. 2010. "Effects of Scholar's Gender and Professional Position on Publishing Productivity in Different Publication Types. Analysis of a Finnish University." *Scientometrics* 82 (2): 419–437. doi:10.1007/s11192-009-0037-7.

Rothaermel, Frank T., and Andrew M. Hess. 2007. "Building Dynamic Capabilities: Innovation Driven by Individual-, Firm-, and Network-Level Effects." *Organization Science* 18 (6): 898–921. doi:10.1287/orsc.1070.0291.

Sax, Linda J., Linda Serra Hagedorn, Marisol Arredondo, and Frank A. Dicrisi. 2002. "Faculty Research Productivity: Exploring the Role of Gender and Family-Related Factors." *Research in Higher Education* 43 (4): 423–446. doi:10.1023/A:1015575616285.

Stack, Steven. 2004. "Gender, Children and Research Productivity." *Research in Higher Education* 45 (8): 891–920. doi:10.1007/s11162-004-5953-z.

Stirling, Andy. 2007. "A General Framework for Analysing Diversity in Science, Technology and Society." *Journal of the Royal Society Interface* 4 (15): 707–719. doi:10.1098/rsif.2007.0213.

Subramanian, Annapoornima M., Kwanghui Lim, and Pek Hooi Soh. 2013. "When Birds of a Feather Don't Flock Together: Different Scientists and the Roles They Play in Biotech R&D Alliances." *Research Policy* 42 (3): 595–612. doi:10.1016/j.respol.2012.12.002.

Swaab, Roderick I., Michael Schaerer, Eric M. Anicich, Richard Ronay, and Adam D. Galinsky. 2014. "The Too-Much-Talent Effect: Team Interdependence Determines When More Talent Is Too Much or Not Enough." *Psychological Science* 25 (8): 1581–1591. doi:10.1177/0956797614537280.

Toivanen, Otto, and Lotta Väänänen. 2016. "Education and Invention." *Review of Economics and Statistics* 98 (2): 382–396. doi:10.1162/REST_a_00520.

USPTO. 2019. *Progress and Potential: A Profile of Women Inventors on US Patents. In Technical Report. IP Data Highlights*. https://www.uspto.gov/sites/default/files/documents/Progress-and-Potential.pdf.

van Zeebroeck, Nicolas, and Bruno van Pottelsberghe de la Potterie. 2011. "The Vulnerability of Patent Value Determinants." *Economics of Innovation and New Technology* 20 (3): 283–308. doi:10.1080/10438591003668638.

WIPO. 2021. *World Intellectual Property Indicators 2021*.

Wullum Nielsen, Mathias, and Love Börjeson. 2019. "Gender Diversity in the Management Field: Does It Matter for Research Outcomes?" *Research Policy* 48 (7). Elsevier: 1617–1632. doi:10.1016/j.respol.2019.03.006.

Xie, Yu, and Kimberlee A. Shauman. 1998. "Sex Differences in Research Productivity: New Evidence about an Old Puzzle." *American Sociological Review* 63 (6): 847–870. doi:10.2307/2657505.

Zucker, Lynne G., and Michael R. Darby. 1997. "Individual Action and the Demand for Institutions: Star Scientists and Institutional Transformation." *American Behavioral Scientist* 40 (4): 502–513. doi:10.1177/0002764297040004012.

Zucker, Lynne G., Michael R. Darby, and Jeff S. Armstrong. 2002. "Commercializing Knowledge: University Science, Knowledge Capture, and Firm Performance in Biotechnology." *Management Science* 48 (1): 138–153. doi:10.1287/mnsc.48.1.138.14274.

Appendix

A0. Summary of key variables

Table 11 Description and summary statistics of variables used in the econometric models

Variable	**Description**	**Obs**	**Mean**	**Std. dev.**	**Min**	**Max**
Quantity	Number of patent families in the inventor's portfolio	639,860	7.25	13.6581	1	758
Quality	Inventor's portfolio share of outstanding inventions, defined as those in their top 95th percentile in the year-tech.field cohort	639,860	0.08	0.2017	0	1
Female dummy	Dummy equal 1 if the inventor is female according to the data in PatentsView	639,860	0.13	0.3371	0	1
Career length	Years since the first patent filing (= 2011 – earliest filing year)	639,860	8.18	6.8628	1	30
Geographical origin (US=1)	Dummy equal 1 if the inventor's earliest location for patent filing is in the US	639,860	0.44	0.4967	0	1

A1. Missing information on gender

Table 12 reports the distribution of missing data about gender across technological fields.

Table 12 Distribution of missing data about gender across technological fields

WIPO field code	WIPO field	% of missing gender
-	Full sample	9.1
1	Electrical machinery, apparatus, energy	9.26
2	Audio-visual technology	9.61
3	Telecommunications	9.97
4	Digital communication	10.63
5	Basic communication processes	10.79
6	Computer technology	9.37
7	IT methods for management	5.91
8	Semiconductors	11.50
9	Optics	9.26
10	Measurement	7.82
11	Analysis of biological materials	7.99
12	Control	6.88
13	Medical technology	5.89
14	Organic fine chemistry	10.29
15	Biotechnology	9.51
16	Pharmaceuticals	9.85
17	Macromolecular chemistry, polymers	8.73
18	Food chemistry	7.20
19	Basic materials chemistry	8.69
20	Materials, metallurgy	8.78
21	Surface technology, coating	8.55
22	Micro-structural and nano-technology	11.22
23	Chemical engineering	7.49
24	Environmental technology	6.43
25	Handling	4.60
26	Machine tools	5.84
27	Engines, pumps, turbines	5.62
28	Textile and paper machines	6.21
29	Other special machines	6.30
30	Thermal processes and apparatus	7.03
31	Mechanical elements	5.06
32	Transport	4.75
33	Furniture, games	5.08
34	Other consumer goods	6.21
35	Civil engineering	5.02

Table 13 shows descriptive statistics on the portfolio size (number of patent families) for female, male and inventors with missing gender data. With respect to the identified prolific inventors, missing data on gender represent 6.9% of stars and 9.4% of non-stars.

Table 13 Average portfolio of patent families for the selected sample

Gender	Mean	Std. Dev.	Min	Max
female	5.645	13.279	1	713
male	7.491	13.704	1	758
missing	7.289	15.260	1	1041

A2. Prolific star inventors accounting for career length

The definition of prolific stars can be modified by weighting for career length the number of patents in an inventor's portfolio. In fact, the examined inventors had different time windows to produce their stock of patents. Inventors have different debut year, identified as the priority year of their first invention: this serves to compute the career age and discount the count of patents by such a weight. With this alternative definition, individual inventors are defined as prolific stars both with the simple count of patents and with the patent count divided by career length in less than 9% of cases, while they identify different inventors in around 20% of cases (Table 14). The majority of inventors (more than 70%) is not prolific with either definition.[13]

Table 14 Overlap between the baseline definition of prolific inventors and a career-weighted one. Definition 1 is the count of patents; definition 2 is the count of patents divided by career length (years since first patent).

Inventors	Definition 2: Not a star	Definition 2: Star	Total
Definition 1: Not a star	71.3%	14.5%	*85.8%*
Definition 1: Star	5.5%	8.7%	*14.2%*
Total	76.8%	23.2%	100.0%

When looking at women inventors only, we see that the adjustment for career length increases the identified share of prolific female stars from 8.5% to 22%, with an overlap of only 6% of cases between the two definitions (Table 15). Therefore, we take this as indication that it is important to examine more formally in the regression analysis the role of career length in the representation of star innovators and for the analysis of gender, in particular.

[13] The sample of prolific stars when weighting for career length is even more dense of prolific inventors due to the inflation of those debuting in the last year: 23% of innovators are prolific (>=95th percentile in their sector). Across fields the average number of patents per year ranges between 0.605 (field 29, "Other special machines") and 1.268 (field 8, "Semiconductors"): inventors with at least one filing in 2010 are very likely to be considered prolific.

Table 15 Overlap between the baseline definition of prolific inventors and a career-weighted one in the sample of female inventors. Definition 1 is the count of patents; definition 2 is the count of patents divided by career length (years since first patent).

Inventors	Definition 2: Not a star	Definition 2: Star	Total
Definition 1: Not a star	75.5%	16.0%	91.5%
Definition 1: Star	2.6%	6.0%	8.5%
Total	78.0%	22.0%	100.0%

Table 16 Distribution of portfolio size weighted by career length (average yearly number of patent families since debut year).

Range		M	F	% of total inventors in this range	% of female on inventors in this range
>0	0.25	99’149	17’093	18.2%	14.7%
>0.25	0.5	183’119	29’890	33.3%	14.0%
>0.5	1	141’937	19’197	25.2%	11.9%
>1	1.5	46’297	5’455	8.1%	10.5%
>1.5	2	32’401	4’253	5.7%	11.6%
>2	758	53’491	7’761	9.6%	12.7%

A3. Overlap of prolific and high-quality star inventors (including small portfolios)

In this Appendix we establish how much our definitions capture similar types of star inventors. The quality measure (“hit ratio”) is not expected to have a high correlation with individual portfolio size (Forthmann et al. 2020; Caviggioli and Forthmann 2022), since the ability to produce a lot of patents is mostly related to the number of high-quality inventions but not to their portfolio share. Our data confirm such evidence from the literature: the two definitions of quantity and quality performance do not overlap much in the identification of stars (Table 17**Errore. L'origine riferimento non è stata trovata.**), with less than 3% of the sample being a star under both definitions.

Table 17 Overlap of stars’ definitions (quantity vs. quality) - subsample of inventors with at least five inventions.

Overlap Quantity/Quality	High-quality: not a star	High-quality: star	Total
Prolific: not a star	61.1%	3.7%	64.7%
Prolific: star	32.5%	2.8%	35.3%
Total	93.6%	6.4%	100.0%

Notes: “Prolific stars” are top 5 percentile inventors by family portfolio size. “Highly-cited stars” are the inventors having a portfolio share of high-quality inventions equal or above 95th percentile.

The same analysis broken down by gender is reported in the Table 19 for the subsample of female inventors, and Table 20 for the male inventors. The results show that 2.2% of female inventors is considered a star under both definitions, compared to 2.8% of male inventors. Clearly, women face more difficulties than their male counterparts in becoming stars, even if they already have more than five inventions.

Whether we use the threshold of quality with more than five inventions or not does not make much of a difference, since the overlap is indeed quite limited: Table 18 reports the results for the high-quality definition without any portfolio size thresholds. The criterion for the identification of inventors with an outstanding share of high-quality inventions seems stricter than the one for being prolific.

Table 18 Overlap of the two types of stars' definitions (quantity versus quality) on the full sample of inventors.

Overlap Quantity/Quality	*High- quality: not a star*	*High-quality: star*	*Total*
Prolific (def. 1): not a star	79.9%	5.5%	85.5%
Prolific (def. 1): Star	13.8%	0.7%	14.5%
Total	93.8%	6.2%	100.0%

Notes: "Prolific stars" are top 5 percentile inventors by family portfolio size. "Highly-cited stars" are the inventors having a portfolio share of high-quality inventions equal or above 95th percentile.

The following tables illustrate how the overlap between our definitions based on quantity and on quality vary by gender in the sample of inventors with at least 5 inventions (Table 19 for the subsample of female inventors, Table 20 for the male inventors). In this sample of inventors with at least 5 inventions, 27.8% of women are defined as prolific stars, versus 36.1% of men. The difference for highly-cited stars is instead smaller: 6.3% of the female inventors are high-quality, versus 6.5% of males. Finally, 2.2% of female inventors is considered a star under both definitions, compared to 2.8% of male inventors.

Table 19 Overlap of the two types of stars on the subsample of female inventors with at least five inventions.

Overlap Quantity/Quality	*High- quality: not a star*	*High-quality: star*	*Total*
Prolific: not a star	68.1%	4.1%	72.2%
Prolific: star	25.5%	2.2%	27.8%
Total	93.7%	6.3%	100.0%

Notes: "Prolific stars" are top 5 percentile inventors by family portfolio size. "Highly-cited stars" are the inventors having a portfolio share of high-quality inventions equal or above 95th percentile.

Table 20 Overlap of the two types of stars on the subsample of male inventors with at least five inventions.

Overlap Quantity/Quality	*High- quality: not a star*	*High-quality: star*	*Total*
Prolific: not a star	60.3%	3.6%	63.9%
Prolific: star	33.2%	2.8%	36.1%
Total	93.5%	6.5%	100.0%

Notes: "Prolific stars" are top 5 percentile inventors by family portfolio size. "Highly-cited stars" are the inventors having a portfolio share of high-quality inventions equal or above 95th percentile.

A4. Sensitivity analysis on the minimum number of patent families for the sample used to test the quality metrics

Table 21 Results of fractional response models using a logit for the conditional mean selecting the inventors with at least three/seven inventions in their portfolio

Model	(1)	(2)	(3)	(4)	(5)	(6)
Sample	>=3 Inv.	>=7 Inv.	>=3 Inv.	>=7 Inv.	>=3 inv. and star by high-quality	>=7 inv. and star by high-quality
Dependent variable	Share of high-quality inventions in the portfolio					
Female dummy	0.9405***	0.9481**	1.0455	0.9911	1.1297**	1.0420
	(0.0194)	(0.0256)	(0.0327)	(0.0409)	(0.0543)	(0.0687)
Career length	1.0147+	1.0074+	1.0149+	1.0075+	0.9858+	0.9919+
	(0.0004)	(0.0005)	(0.0004)	(0.0005)	(0.0007)	(0.0008)
Female dummy X Career length	1.0066+	1.0015	1.0052***	0.9999	0.9995	0.9998
	(0.0016)	(0.0019)	(0.0016)	(0.0020)	(0.0026)	(0.0032)
Geographical origin (US=1)	2.1347+	2.0000+	2.1349+	1.9989+	1.0828+	1.0552+
	(0.0141)	(0.0151)	(0.0141)	(0.0151)	(0.0107)	(0.0116)
Main technological field dummies	Y	Y	Y	Y	Y	Y
Female dum. x each main tech. field dummy			Y	Y	Y	Y
Observations	356554	188924	356554	188924	18125	9814
PseudoR2	0.0345	0.0298	0.0348	0.0299	0.0032	0.0015
loglik.	-97543.28	-51738.75	-97521.91	-51729.96	-12463.32	-6759.36

Notes: Results as odds ratios indicate a negative relationship when the coefficient is below one, a positive relationship if above one. The dependent variable is the share of high-quality inventions in the total portfolio. * p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01; + p-value < 0.001