Department of Economics and Statistics "Cognetti de Martiis"
Campus Luigi Einaudi, Lungo Dora Siena 100/A, 10153 Torino (Italy)
www.est.unito.it

Dipartimento
Economia e Statistica
Cognetti de Martiis

Working Paper Series

17/22

PEERS AND STARS: THE ROLE OF GENDER AMONG COINVENTORS

FEDERICO CAVIGGIOLI, ALESSANDRA COLOMBELLI and CHIARA RAVETTI

Bureau of Research on Innovation, Complexity and Knowledge

UNIVERSITÀ DEGLI STUDI DI TORINO

The Department of Economics and Statistics "Cognetti de Martiis" publishes research papers authored by members and guests of the Department and of its research centers. ISSN: 2039-4004

Peers and stars: the role of gender among coinventors

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

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

**Corresponding author:* *federico.caviggioli@polito.it*

Abstract

This article examines the role of gender in patent inventors' collaborations and individual productivity. We study how the time needed by an inventor to eventually become a "star" relates to their portfolio of female and male coinventors, characterised in terms of gender, career seniority and productivity. Our empirical analysis applies different survival models to a sample of almost 100k inventors debuting in 2000 and all their patenting peers, followed over a period of 20 years. We find that being female and having female coinventors is correlated to a longer time to become star and that is not a matter of homopily. Seniority is also correlated to a longer time, while having a star among coinventors to a shorter time, in particular for female inventors. These findings confirm the presence of a relevant gender bias and suggest a potential beneficial mentoring/role model mechanism with stars being a strong catalyst of other stars, especially when among women.

Keywords: patents, innovation, gender, star inventors, homophily

1 Introduction

Contemporary innovation is fundamentally an interactive and cooperative activity. Knowledge networks enable the transfer and recombination of complex information, ideas and competences, and current scientific breakthroughs are deeply influenced by collective dynamics within teams and organizations (Wuchty, Jones, and Uzzi 2007; Singh and Fleming 2010; Guan and Liu 2016). As individuals specialize in narrower knowledge areas, and as the knowledge frontier keeps shifting, the returns to collaboration increase (Jones 2009; Agrawal, Goldfarb, and Teodoridis 2016). Focusing exclusively on individual productivity in innovation systems neglects a key driver of knowledge creation and creates a distorted view of intellectual human capital, with no appreciation of social spillovers and synergies (Grigoriou and Rothaermel 2014). The composition and configuration of teams is thus an important driver of innovative scientific outputs and, in this context, group heterogeneity and diversity can be an important source of creativity (Perry-Smith and Shalley 2014; J. Wang et al. 2019).

Heterogeneity in scientific research teams can be measured along multiple dimensions. Many studies have considered the background knowledge or expertise of team members and some have recently focused on their demographic characteristics, such as age, gender, ethnicity or nationality (Hall et al. 2018; Vakili and Kaplan 2021). The analyses of such dimensions can unveil underrepresentation of certain categories, as in the case of gender gaps in patented innovations (Hunt et al. 2013; Lax Martinez, Raffo, and Saito 2016; Haseltine and Chodos 2017; Groysberg and Lee 2009; Oldroyd and Morris 2012). The underrepresentation of female inventors entails a critical loss for society of potential innovations and progress from the overall talent pool (Bell et al. 2019; Shannon et al. 2019).

Previous studies linking gender to intellectual productivity - and patenting, in particular - suggest that *homophily* (innovators matching with other inventors with their same characteristics, such as same gender) is expected to boost productivity, especially for women (Frietsch et al. 2009; Meng 2016; Y. Wang et al. 2020; Whittington 2018). This result should account for the fact that women represent a minority in patented innovations (USPTO 2019), and it is hard for them to partner with other female colleagues, and to find more experienced female mentors and role models. At the same time, gender diversity in innovative sectors is desirable as it encourages new knowledge creation (Tshetshema and Chan 2020) and prevents intellectual lock-in (Wullum Nielsen and Börjeson 2019). Thus, it is unclear whether women's collaborations can actually enhance female inventors' career and more empirical evidence is needed on the specific collaborative patterns that could foster women's success in patented innovation.

This article focuses on the role of gender and seniority/expertise in individual traits in collaborations within innovation activities, measured through co-patenting over time. We examine how individual and team-related characteristics shape the chances of success for an inventor. To measure the accomplishments of an inventor, we consider "star" individuals with an outstanding productivity in terms of number of generated patents (Zucker and Darby 1997; Groysberg and Lee 2009; Oldroyd and Morris 2012; Kehoe and Tzabbar 2015; Liu, Mihm, and Sosa 2018). We analyse how the achievement of an exceptional level of productivity is related to the team-level characteristics of an inventor, with respect to the presence of coinventors of the same gender, their seniority and experience, and whether some of them are already stars.

We apply different survival models to a representative sample of almost 40’000 inventors, over a period of 10 years. We find that being a female inventor is on average correlated with a longer time frame to become a prolific star innovator. Moreover, a higher share of female coinventors delays the occurrence of becoming a star (although the effect is not statistically significant in all the models). However, having collaborators of the opposite gender is associated with a shorter time to reach the status of star: diversity is linked to more successful patenting trajectories. The seniority of coinventors does not seem to matter in and of itself but having experts of the same gender in the patenting group (expertise and homophily) reduces the time to become stars. Finally, if there are stars among coinventors, this seems to favour the speed towards success, especially if the stars are female.

The rest of the article is organized as follows. Section 2 examines positions our research within the previous literature, to develop our research questions and hypotheses regarding gender and collaborations in relation to star inventors. Section 3 presents the data and some relevant summary statistics, and Section **Errore. L'origine riferimento non è stata trovata.**4, the empirical methodology applied. Section 5 discusses the results and concludes with some policy implications and further research avenues.

2 Literature and research questions

Our work contributes to the literature on gender gaps in patenting, as well as to the evidence on peer effects and team characteristics in the generation of innovations. We discuss our contribution to each of these literatures in turn.

2.1 Gender gap in patenting

The existing evidence on individual inventors’ characteristics identifies a substantial gender gap in patenting, with a stark predominance of male innovators, despite a recent historical trend of increasing female representation (Heikkilä 2019; USPTO 2019). Among patents granted by the European Patent Office (EPO), for instance, the share of female inventors increased between 1978 and 2019 from 1.2% to 8.9% (Tahmooresnejad and Turkina 2022). In the US, the percentage of female inventors is slightly higher, but still constituting a minority, with around 13% of inventors being women (USPTO 2020). The patenting gender gap negatively impacts the innovation system both at the individual level, as some female inventors cannot fully benefit financially from their discoveries through intellectual property rights (Kline et al. 2019): patent applications by women inventors are more likely to be rejected and female granted patents have a smaller fraction of their claims allowed and are less likely to be maintained by their assignees than those of men (Jensen, Kovács, and Sorenson 2018). Moreover, at an aggregate level, the gender gap creates “lost Marie Curies” that could contribute to the global advancement of science and technology (Bell et al. 2019).

Technological fields where patenting is most relevant are typically related to STEM (Science, Technology, Engineering and Mathematics) disciplines, which already at the educations level suffer from a significant gender gap (Legewie and DiPrete 2014) and are further characterized by ‘leaky pipelines’, with the representation of women in science dropping at each career stage (Schmuck 2017). Senior, experienced women inventors with already developed successful careers are therefore likely to be rare in most patenting sectors.

Areas with a low proportion of women, such as STEM disciplines, are expected to be characterized by strong discrimination against women, but even when equal shares are reached, biases can still persist (Begeny et al. 2020). In such a context, it is important to improve the understanding of the team level mechanisms as leverages towards equality.

In scientific publications, women are underrepresented at prestigious journals and in articles attracting the highest number of citation (Bendels et al. 2018). Not only are women underrepresented, but the gender gap seem related to the size of the co-authorship: the gender-specific differences in citation rates increase the more authors contribute to an article (Bendels et al. 2018). Moreover, patents to which women contributed are associated with a higher number of coinventors (Sugimoto et al. 2015). It is therefore key to study gender gaps in tandem with the number of co-authors, controlling for the specific sectors, and unpacking possible mechanisms and channels behind these gaps.

2.2 Peer effects: gender homophily, diversity and mentorship

In patenting as in all scientific endeavours requiring team efforts, seniority, authorship position, collaboration and team configuration are all highly interlinked variables (Larivière et al. 2013). Multiple empirical studies have shown that the network dynamics of male and female collaborations differ, both in the nature of the interactions and in the innovative outputs they generate (Szell and Thurner 2013). The positive influence of a central network position is greater for male inventors than female inventors (Tahmooresnejad and Turkina 2022). In academic research output, women seem to co-author more often with the same people, and a higher fraction of their co-authors collaborate with each other (Ductor, Goyal, and Prummer 2021). Women benefit from collaborating with women, and are more likely to collaborate with women, but both men and women collaborate with mostly men (Whittington 2018).

While the networks' literature is key to understand women's positioning and strategies in social networks, it usually relies on the analysis of one specific network, for instance one discipline, to map all relations and ties arising within that networks' participants. We take a different approach and look at the portfolio of coinventors of all innovators active in a period in the US, spanning multiple disciplines. There is a need for large-scale studies on gender-related differences (Tahmooresnejad and Turkina 2022). We move beyond classic network metrics (centrality, brokerage, clustering) which have been studied extensively elsewhere (Bellotti et al. 2022; Tahmooresnejad and Turkina 2022; Whittington 2018) and focus instead on specific relational variables capturing the seniority, experience and previous success of co-patenting inventors, to capture potential role model effects. For scientific publications, it has already been shown that co-authoring with a top scientist can have long lasting productivity impacts on their peers (Li et al. 2019). We consider peer effects among similar people (driven for instance by homophily) and hierarchical effects with coinventors with high experience or who are already highly successful stars. Regarding peer effects, it is often assumed that similar people are more likely to form social ties. This phenomenon, referred to as social *homophily*, has been documented in several academic disciplines and is typically based on socially salient characteristics, such as gender, ethnicity, religion, social class, and leads to social segregation (Reme et al. 2022). Previous literature has examined the characteristics of an inventor's team in the creation of new ideas and inventions (C. Wang et al. 2014; Guan and Liu 2016; Grigoriou and Rothaermel 2017; Zhang, Wang, and Duan 2020). A number of recent studies has looked more specifically at the role of gender diversity, finding that gender homophily is quite common, as co-innovators tend to prefer matches with other inventors of the same gender (Frietsch et al. 2009; Meng 2016; Whittington 2018; Y. Wang et al. 2020).
Moreover, in terms of underlying mechanisms, the evidence for the Unites States indicates that environmental factors, and particularly exposure to specific networks and to same-gender mentorship, might be key in determining the access of female inventors

to star-level patenting (Bell et al. 2019). Due to the gender gap in these technological fields, female inventors could find it harder than men to work in teams with experienced female mentors that could act as role models. A special role in teams could be played by star inventors, the extremely prolific ones. Even though they could be associated to negative effects in organizations due to coordination costs and conflicts with collaborators (Bendersky and Hays 2012; Groysberg, Polzer, and Elfenbein 2011; Swaab et al. 2014), star scientists provide directly with extremely superior innovation output, and indirectly with support to an organization’s activities (Kehoe and Tzabbar 2015) and to the attraction of resources and skilled personnel (Lacetera, Cockburn, and Henderson 2004; Hess and Rothaermel 2011). In particular, their presence fosters the productivity of peers thanks to learning and emulation (Lockwood and Kunda 1997) and they can act as relational pivots to foster further innovation (Grigoriou and Rothaermel 2014). In this sense, they could represent a role model for their collaborators.

Even though gender gaps in innovation can be clearly detrimental, the understanding of female underrepresentation amongst innovators is still in its early stages, especially in the context of star inventors. To date there is limited evidence on whether the relationship with stars is somewhat gendered (Caviggioli, Colombelli and Ravetti 2022), and if gender matching in a team is related to the probability that an individual becomes a star. Given the scarce representation of women among patenting scientists, and the social expectations about team behaviour of women (Motro, Spoelma, and Ellis 2021), the effect of male and female stars in a group of innovators is not theoretically obvious.

We can expect that, due to homophily, the highly represented male stars may support particularly other male inventor, thus creating a further gender gap. However, alternatively, if diversity fosters innovation and junior women inventors can bring new ideas that complement the expertise of a star, this could reduce gender gaps and favour a rapid achievement of the status of star even for women. These different mechanisms need to be disentangled empirically looking at which channel prevails in the data.

Our main research question stemming from these different literatures is thus whether the gender characteristics of the group of collaborators play a role in the pathway towards success of an individual inventor. The teams of coinventors are described in terms of gender differences, applied also to experience and star status, and in relation to the own characteristics of the individual inventor. We hypothesise that the relationship might be different depending on male or female homophily pairings and capture possible role model effect for stars, hence we test these different characteristics of inventors in separate specifications.

3 Data

3.1 Data sources and sample selection

Data were collected from PatentsView, a data warehouse that contains data on granted patents at the USPTO. Critically for our study of individuals, this database includes disambiguated inventors' names and gender identification from the application of an algorithm (USPTO 2019).

With the aim to consider comparable careers, we collected the data of the inventors having their debut in patenting in year 2000: these are in total 104,102 individuals (2% of the whole population in PatentsView). The debut year is defined as the earliest application year in the portfolio of granted patents belonging to each inventor. Although this definition is not an exact measure of the actual career start of the inventor, it is a reliable proxy (Duffy et al. 2011; König et al. 2015; Costas, Nane, and Larivière 2015). Data are examined up to year 2019, i.e., 20 years of career are considered.

For 5.0% of the selected sample, the algorithm employed in PatentsView was not able to associate the gender. Our investigation excludes individuals with no gender and thus considers 98,940 subjects, 14.4% are female.

For each patent in the selected inventors' portfolios, we traced and analysed all the coinventors: we identified their gender, the debut year in patenting as a measure of seniority, and whether they are star inventors or not. The issue of missing gender affects also the coinventors. For this reason, our main analyses will focus on the full sample, including those inventors having one or more peers with no gender. In addition, we will report in the Appendix the results of the analyses when excluding inventors having at least one coinventor with no identified gender.

3.2 Operationalization of star inventors

The identification of "stars" can apply different criteria to distinguish outstanding from average performers. In this study, stars are the "prolific" inventors, defined by their cumulative number of granted USPTO patents, in a sector and year. In any year and technological field, we count the total number of patents for all inventors available in PatentsView up to that date. We highlight that this approach is not limited to the inventors in our sample, but it includes the whole population: the ranking is defined on all the inventors, despite they are included or not in the examined sample.

Since patent propensity is different across technological sectors, we rank inventors by the number of patents within the 35 technological sectors from WIPO concordance table (4-digit IPC codes). Each patent can be associated to multiple IPC subclasses and sectors. The star inventors are defined as the most prolific when their patent portfolio size is above the 95th percentile of the distribution in their sector from 1868 to the considered year. The variable is thus operationalized as a dummy equal to one if the productivity of an inventor is above 95^{th} percentile of all the inventors in the same technological field in a certain year.

There are 15,076 stars, 15.2% of the sample. This is more than the expected 5% by construction, due to increasing trend in patenting and the ranking on the whole population of inventors (older inventors, which might be inactive, are kept in the ranking). Female inventors are 10.3% of the stars.

The share of inventors that never reach the status of star in the examined 20-years career is equal to 84% and 89% for male and female respectively. Figure 1 shows the share of inventors becoming star in each gendered sub-sample: the relative percentage of female becoming stars is smaller than of male.

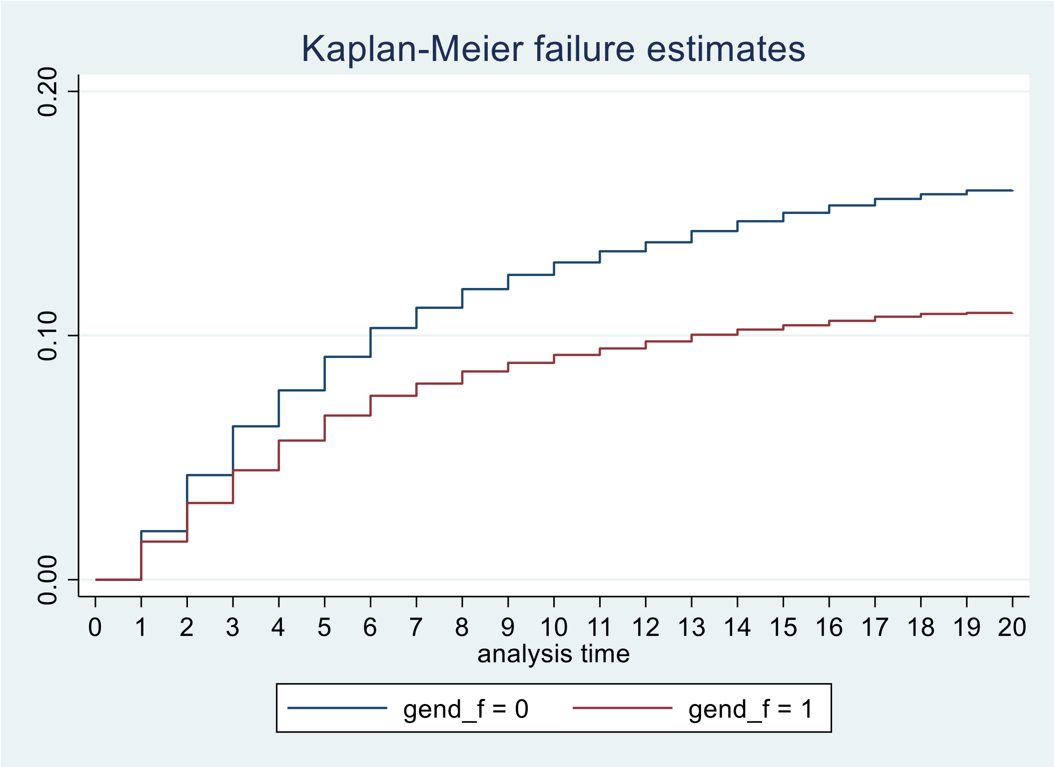

Figure 1 Kaplan-Meier failure estimates: share of inventors in each gender expected to reach the status of star in the examined years (year 1 on x-axis is the debut year - 2000)

Focusing on the sub-sample of stars, 12.6% reach this status in their debut year. In total, 57.2% of male stars and 61.5% of female stars have become so by year five (Figure 2). In this sub-sample, the difference time between male and female shows that male inventors become stars with an average delay of 0.43 years: although the t-test on the difference with female inventors reports a significant p-value, since the unit of analysis is in years, this difference cannot be considered strong. These preliminary basic statistics suggest the presence of additional hurdles for women to become stars, but there is weak evidence that those that are able to overcome the difficulties or are extremely outstanding reach the star-status faster than male.

Figure 2 share of star inventors reaching the status of star for the first time (y-axis) per number of years since the debut in patenting (x-axis). The graph captures only scientists who eventually become stars in our sample.

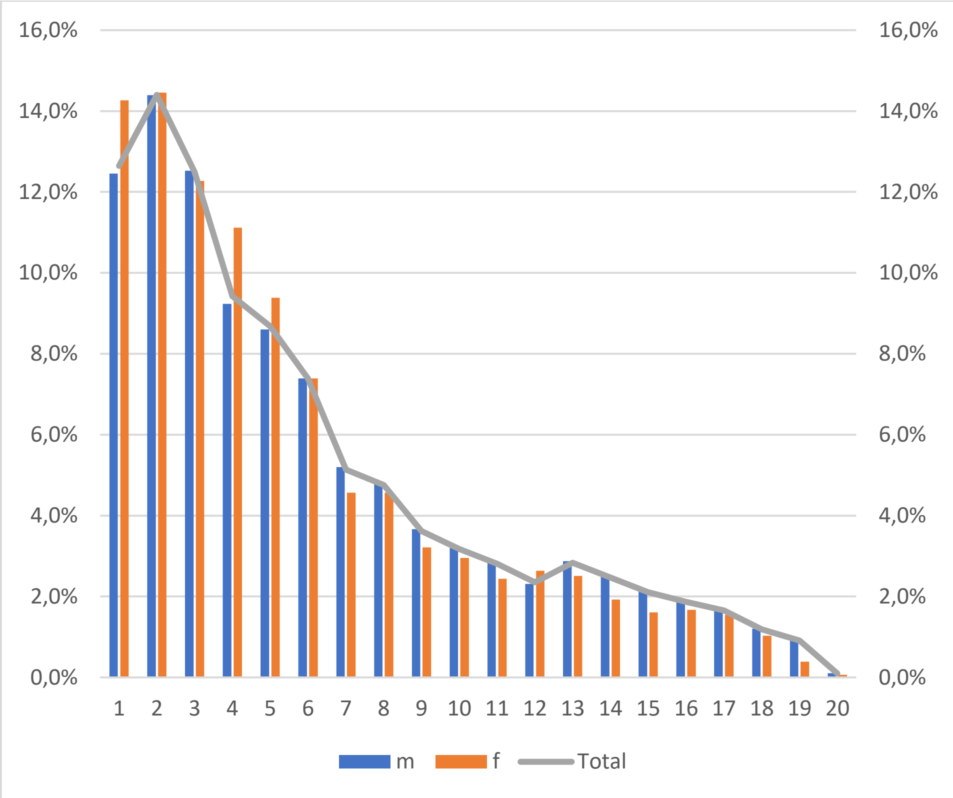

Figure 2 provides also support to the identification strategy of stars. Our sample covers a time window of twenty years after the first filing for any inventor: if an inventor would "become a star" after the analysed time span, it is not identified. However, the data distribution suggests that inventors are outstandingly prolific in the first part of their career: we do not identify a significantly increasing number of stars in the right tail of Figure 2. Hence, we believe that the number of inventors becoming stars after the last examined year is marginal. Our truncation after twenty years should not significantly bias the identification of stars in the sample.

As suggested by the literature on gendered studies and technological fields (Mayer and Rathmann 2018; Wullum Nielsen and Börjeson 2019; Puuska 2010), there are differences across areas of innovations. Our data (Table 1) confirm the presence of heterogeneity across sectors according to multiple dimensions. The fields where the presence of female is the highest are "Pharmaceuticals", "Biotechnology" and "Organic fine chemistry"; the lowest in "Mechanical elements", "Engines, pumps, turbines" and "Civil engineering". In general, there is a high correlation (0.98) between the share of female inventors and the share of female star. In terms of time to become stars, female are particularly faster in the "Thermal processes and apparatus", "Machine tools" (both male dominated) and "Other consumer goods"; male are faster in "Environmental technology", "Mechanical elements” and “Handling”. The correlation analysis at the field level between the presence of female star inventors and the average delay in becoming a star shows a value of 0.27.

Table 1 Statistics across technological field (WIPO concordance table): share of female on total inventors; share of female on star inventors; average time to become a star (in years); average delay to become star of male inventors with respect to female (year difference).

WIPO code	**Description**	**Female inv. (on tot. inv.)**	**Female stars (on stars)**	**Avg time to become star**	**Delay of male star vs female**
	Full sample	14.4%	10.3%	5.03	0.43
01	Electrical machinery, apparatus, energy	10.4%	8.8%	6.54	0.02
02	Audio-visual technology	11.9%	10.4%	6.08	0.24
03	Telecommunications	11.6%	10.1%	5.21	0.86
04	Digital communication	11.3%	10.2%	4.74	0.37
05	Basic communication processes	9.4%	9.2%	6.24	0.85
06	Computer technology	12.2%	9.9%	5.11	0.76
07	IT methods for management	13.5%	11.3%	3.72	1.17
08	Semiconductors	13.9%	10.8%	6.08	0.39
09	Optics	13.1%	10.7%	6.52	0.66
10	Measurement	10.5%	8.3%	6.46	0.31
11	Analysis of biological materials	21.4%	15.8%	7.34	0.81
12	Control	10.3%	8.9%	5.65	1.32
13	Medical technology	14.7%	9.9%	6.64	0.62
14	Organic fine chemistry	23.2%	16.6%	7.22	0.77
15	Biotechnology	25.8%	16.7%	7.48	1.18
16	Pharmaceuticals	25.9%	16.8%	7.00	0.79
17	Macromolecular chemistry, polymers	19.3%	14.2%	7.26	0.48
18	Food chemistry	22.2%	13.2%	7.32	0.48
19	Basic materials chemistry	19.2%	14.3%	7.41	0.01
20	Materials, metallurgy	13.0%	10.8%	6.82	0.49
21	Surface technology, coating	13.6%	11.1%	7.10	-0.13
22	Micro-structural and nano-technology	14.5%	12.4%	6.39	1.09
23	Chemical engineering	12.2%	9.1%	6.73	0.41
24	Environmental technology	10.2%	7.7%	6.46	-1.45
25	Handling	9.3%	7.3%	6.80	-0.44
26	Machine tools	8.1%	6.6%	6.87	1.87
27	Engines, pumps, turbines	7.7%	6.4%	6.10	0.55
28	Textile and paper machines	14.8%	11.8%	6.28	1.02
29	Other special machines	11.6%	9.6%	7.18	1.15
30	Thermal processes and apparatus	8.9%	7.0%	6.73	2.25
31	Mechanical elements	6.4%	5.7%	6.70	-1.11
32	Transport	8.4%	6.5%	6.15	0.85
33	Furniture, games	14.4%	8.4%	6.21	0.16
34	Other consumer goods	16.9%	10.5%	6.62	1.67
35	Civil engineering	7.9%	7.0%	6.72	0.38

3.3 Characteristics of coinventors

The key contribution of our study is to move beyond the individual characteristics of an inventor and relate them to the other members of the patenting team. In each year and for each focal inventor, we look at the past patenting activities and calculate the cumulative number of female and male coinventors[1] and a set of descriptors of the team in terms of gender, seniority and being stars. The descriptors are computed as share on the total number of peers up to the examined year.

The measures on gender are built as share of female (male) inventors. This approach is useful to directly investigate female scientists as a specific subgroup in the innovation system - given their scarce representation and the potential biases they may face, especially in more masculine fields, but also provides the possibility to investigate the overall gender pairing/homophily, considering that the match of male-male inventors is much more common than female-female and the heterophilic ones.

The career age of coinventors is employed to determine the seniority of peers as cumulative collaborations up to the considered year. Coinventors are considered "senior" when they have more than ten years of experience, measured from the patenting debut. For each coinventor we also determine whether s/he is a star in any given year. While seniority and star-status might theoretically be correlated, we have shown previously that many successful scientists become stars rapidly, and therefore they might be quite different from older inventors that have cumulated experience over the years and may be more mentoring figures. Finally, the gender dimension is added to the analyses on seniority and stars: we computed the share of female (male) and of same-gender peers having seniority and being stars on the cumulated number of coinventors.

Table 2 reports the descriptive statistics of the main regressors included in the panel of the econometric specifications. A single failure approach is employed to investigate data (further details in the next section about the survival model): this means that once an inventor reaches the status of star (better than 95% of her/his comparable peers), s/he is out of the sample. The correlation matrix is reported in the appendix (Table 5).

[1] If the coinventor "Jane Smith" appears in two patents, she is counted twice.

Table 2 Panel level summary statistics of data organized with a single failure approach (N= 1,768,245)

Variable	**Mean**	**S.D.**	**Min**	**25th Perc.**	**Median**	**75th Perc.**	**Max**
Female dummy	0.15	0.36	0	0	0	0	1
Nr. of coinventors	5.93	8.94	0	1	3	7	655
Share of fem. coinv.	0.09	0.19	0	0	0	0.11	1
Share of male coinv.	0.74	0.37	0	0.6	1	1	1
Share of same-gender coinv.	0.66	0.41	0	0.22	0.88	1	1
Share of coinv. with >10yrs exp.	0.19	0.27	0	0	0	0.33	1
Share of fem. coinv. with >10yrs exp.	0.01	0.06	0	0	0	0	1
Share of male. coinv. with >10yrs exp.	0.18	0.26	0	0	0	0.33	1
Share of same-gend. coinv. with >10yrs exp.	0.16	0.25	0	0	0	0.25	1
Share of star coinventors	0.24	0.31	0	0	0	0.44	1
Share of fem. star coinv.	0.02	0.08	0	0	0	0	1
Share of male star coinv.	0.22	0.3	0	0	0	0.39	1
Share of same-gend. star coinv.	0.19	0.29	0	0	0	0.33	1

Additional summary statistics are reported in the appendix (Table 6) and provide details for the sub-samples of male and female inventors only.

4 Empirical Model

We employed a survival analysis to investigate the correlations between the time to become a star for the first time and the gender of the focal inventors and the characteristics of their coinventors. Note that this approach includes the observations of those individuals who become stars on their debut year. Our single-failure model excludes 210,555 observations of inventors being stars also after their first event of becoming a star. 1,768,245 observations represent the total analysis time at risk for 98,940 subjects and 15,076 failures (first time to become star). Models report standard errors that allow for intragroup correlation (clustering on each inventor) and assume a loglogistic distribution (selected according to the results of the AIC – BIC tests).

The baseline models define the dependent variable as the individual time to reach the status of star, with a maximum likelihood estimation for parametric regression survival-time model. The focal inventor is described through the regressor *Female dummy* (equal to one for women) and the set of control dummies identifying the main technological field of activity up to the examined year (note that these variables are not mutually exclusive, as an inventor can be active in more than one sector). The rest of the independent variables describe the group of coinventors of each focal inventor up to the examined year: the share of female (male) coinventors; the share of senior coinventors and the share of female (male) senior peers; the share of star coinventors and the share of female (male) star peers. The models control for the number of coinventors.

The results (Table 3) show that, unsurprisingly, being a female inventor is correlated to a longer time to become a star inventor. Moreover, the share of female coinventors is correlated to an increase in the time to become a star, while the share of male coinventors is not statistically significant. These two results point at significant disadvantages for women, both individually and in groups.

Having senior coinventors is significantly related to a longer time to become a star. This is especially true when the seniors are female rather than male: the coefficient for female senior colleagues is statistically larger than for male peers[2]. The presence of experienced colleagues in the team is associated to a delay in the achievement of star status, especially when there are experienced female peers.

On the contrary, having stars among coinventors is related to a shorter time to become stars and this is confirmed even when highlighting gender, with no significant difference between female and male stars in the team. Thus, once female peers are outstanding, they are not different from male ones in the relationship with the chances of becoming star for the focal inventor: the most productive and talented women that manage to achieve peak success seem to operate in cooperative partnerships that also benefit all other members of the team. Or, when an individual reach the status of star, gender is less likely to be a hindering factor and the behavioural equality is spread also to non-star team members.

[2] Coefficient comparisons are made with the Stata commands "suest" (seemingly unrelated estimation) and "test".

Table 3 Maximum likelihood estimation for parametric regression survival-time models (streg in Stata 17). Multiple-record per individual and single-failure as first time to become a star (above 95th percentile in the reference cohort). The survival models employed loglogistic distribution.

Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Female dummy	0.195***	0.179***	0.195***	0.197***	0.193***	0.196***	0.190***	0.217***	0.167***
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.029)	(0.032)	(0.030)
Share of fem. coinv.		0.302***							
		(0.070)							
Share of male coinv.			0.006						
			(0.038)						
Share of coinv. with >10yrs exp.				0.105***					
				(0.040)					
Share of fem coinv. with >10yrs exp.					0.602**				
					(0.263)				
Share of male coinv. with >10yrs exp.						0.083**			
						(0.041)			
Share of star coinventors							-1.420***		
							(0.027)		
Share of fem. star coinv.								-1.285***	
								(0.076)	
Share of male star coinv.									-1.327***
									(0.027)
Nr. of coinventors	-0.091***	-0.091***	-0.091***	-0.091***	-0.091***	-0.091***	-0.070***	-0.089***	-0.074***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
wipo dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
Constant	3.674***	3.651***	3.669***	3.650***	3.667***	3.656***	4.028***	3.687***	3.979***
	(0.027)	(0.027)	(0.038)	(0.028)	(0.027)	(0.028)	(0.028)	(0.027)	(0.028)
lngamma	-0.413***	-0.413***	-0.413***	-0.414***	-0.414***	-0.414***	-0.478***	-0.418***	-0.464***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Obs	280657	280657	280657	280657	280657	280657	280657	280657	280657
Log lik.	-25993.427	-25981.071	-25993.414	-25989.799	-25988.637	-25991.243	-24722.332	-25889.549	-24962.519
chi-squared	4950.007	4973.638	4988.806	4956.922	4959.090	4954.325	8351.010	5302.334	7697.574

Table 4 Summary of the results of the parametric regression survival-time models (streg in Stata 17) on the sub-samples of male and female inventors. Results of the tests on the difference between each pair of coefficients. Note: each row is extracted from a different model. Details on the models are reported in the appendix.

Variable	Sub-sample: Male inventors only	Sub-sample: Female inventors only	Test of difference Prob > chi2
Share of homo. coinv.	0.023	0.282*	**0.127**
Share of hete. coinv.	0.299***	-0.058	**0.009**
Test of difference Prob > chi2	**0.007**	**0.175**	
Share of homo. coinv. with >10yrs exp.	0.101**	0.872*	**0.143**
Share of hete. coinv. with >10yrs exp.	0.568*	-0.053	**0.051**
Test of difference Prob > chi2	**0.115**	**0.086**	
Share of homo. star coinv.	-1.313***	-1.665***	**0.021**
Share of hete. star coinv.	-1.193***	-1.417***	**0.053**
Test of difference Prob > chi2	**0.184**	**0.166**	

Table 4 shows a summary of the results from the different models tested on the sub-samples of male and female inventors (for the complete set of model results see the appendix: Table 7 and Table 8). This analysis highlights the potential presence of correlations connected to homophily and heterophily.

Concerning the gender of coinventors, the results confirms that it is not the perspective of homophily /heterophily in the team composition to play a significant role, but the presence of female coinventors, which is associated to a longer time to become star.

When considering the presence of senior peers, again the results seem to suggest that it is not a matter a diversified gender composition in the team, but rather seniority per se representing the slowing factor to become star. Nevertheless, the presence of male senior peers is not correlated to the time to become star for female.

When focusing on the presence of star collaborators, the findings are more robust: female inventors are those who benefit more from the presence of stars in the team, especially when they are female themselves. Mentorship, in particular among female peers, could be behind this result. Combining this finding with the one about seniority suggests that the presence of star inventors is able to cancel the gender bias with respect to their non-star peers: the inclusion of outstanding inventors in a team compensates the detrimental effects of gender bias, more when female non-stars are paired with female stars.

In the Appendix we report the results of some robustness tests that analyse the sub-samples of stars only, of individuals with at least one coinventor and of specific sectors with a high (low) presence of female innovators.

The results for the sub-sample limited to stars confirm the preliminary statistics of Figure 2: being female is no longer associated to a significantly longer time to star-status, nor the coinventors' gender is significant. This result suggests that the discriminatory findings from the full sample might be particularly driven by those cases that are never able to become stars, a confirmation of gender bias towards female inventors. Concerning the seniority of peers, this is not significant in general, but it is slightly so when focusing on male seniors. The presence of stars among coinventors is again associated to a shorten time to stars.

Table 10 in the appendix provides the results on the sample of individuals having at least one coinventor (83%). This test excludes solo inventors for which the data on coinventors were forced to zero and are more likely to be individuals inventing as a hobby or outside the boundaries of companies where innovation is a needed driver for growth. Previous findings are confirmed in all cases but on the share of male coinventors which is accelerating the time to become a star. Although different from previous models, it goes in the same direction and adds to the evidence of a gender gap.

Table 11 and Table 12 of the Appendix show the results on the subsamples of inventors in fields with a high and low presence of female individuals respectively. Baseline results are confirmed with the coefficients on seniority that are less robust. Note that "Pharmaceuticals" is the field with the largest share of female inventors, but this value is 25.9%, very far from an equal representation. Hence, even where female inventors are more represented, the bias and the positive relationship of collaborating with stars are still evident.

Finally, the baseline analyses are replicated considering only individuals having coinventors with identified gender, i.e., no missing values (Table 14): the models confirm the previous results on all the dimension except for seniority where the delaying coefficient is mostly driven by male coinventors, while no significant result come from female seniors.

5 Conclusion

In this study we examine the role of gender in interactions among top inventors, depending also on seniority and reciprocal success. Our findings show that gender pairings and the relationship to other prolific inventors matter. First and foremost, in alignment with the rest of the literature on gender gaps, we find that being a female inventor is associated with a longer time to become a star innovator. This result is further confirmed by the finding on coinventors: having a female inventor as patenting colleague is linked with longer time to become a star.

Seniority of coinventors has a general delaying relation with the time to become star, except for women having many senior male coinventors. Instead, having stars among coinventors always reduces the time to become stars, independently of the gender of the inventor and the star collaborators, especially if the stars are female. The presence of stars among conventions is a strong catalyst of other stars, even for the minority of women starts (and irrespective of the gender of the focal inventor).

Overall, our results highlight the relevance of integrating a nuanced gender perspective into innovation policies. First, as our findings indicate that female inventors face greater delays in becoming stars, the most appropriate policy strategy should address the gender disadvantage in the general population of innovators, facilitating the access to resources, support, and opportunities for women. Moreover, policymakers and companies interested in supporting a faster generation of stars, should support gender diversity in inventors' teams and provide incentives for mentorship from highly productive experts of the field to create large teams around them. The benefits of this supervision are going to apply especially to female inventors, irrespectively of the gender of the top scientist. The incentives for such actions could be embedded within the practices of R&D teams in large companies, as well as in the design of regulations and programmes by public bodies supporting innovation, for both private and public support towards greater gender equality.

References

Agrawal, Ajay, Avi Goldfarb, and Florenta Teodoridis. 2016. ‘Understanding the Changing Structure of Scientific Inquiry’. *American Economic Journal: Applied Economics* 8 (1): 100–128. https://doi.org/10.1257/app.20140135.

Begeny, C. T., M. K. Ryan, C. A. Moss-Racusin, and G. Ravetz. 2020. ‘In Some Professions, Women Have Become Well Represented, yet Gender Bias Persists—Perpetuated by Those Who Think It Is Not Happening’. *Science Advances* 6 (26): eaba7814. https://doi.org/10.1126/sciadv.aba7814.

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. https://doi.org/10.1093/qje/qjy028.

Bellotti, Elisa, Dominika Czerniawska, Martin G. Everett, and Luigi Guadalupi. 2022. ‘Gender Inequalities in Research Funding: Unequal Network Configurations, or Unequal Network Returns?’ *Social Networks* 70 (July): 138–51. https://doi.org/10.1016/j.socnet.2021.12.007.

Bendels, Michael H. K., Ruth Müller, Doerthe Brueggmann, and David A. Groneberg. 2018. ‘Gender Disparities in High-Quality Research Revealed by Nature Index Journals’. Edited by Sergi Lozano. *PLOS ONE* 13 (1): e0189136. https://doi.org/10.1371/journal.pone.0189136.

Bendersky, Corinne, and Nicholas A. Hays. 2012. ‘Status Conflict in Groups’. *Organization Science* 23 (2): 323–40. https://doi.org/10.1287/orsc.1110.0734.

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*, 988–98.

Ductor, Lorenzo, Sanjeev Goyal, and Anja Prummer. 2021. ‘Gender and Collaboration’. *The Review of Economics and Statistics*, October, 1–40. https://doi.org/10.1162/rest_a_01113.

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): 207–27. https://doi.org/10.1007/s11192-011-0452-4.

Frietsch, Rainer, Inna Haller, Melanie Funken-Vrohlings, and Hariolf Grupp. 2009. ‘Gender-Specific Patterns in Patenting and Publishing’. *Research Policy* 38 (4): 590–99. https://doi.org/10.1016/j.respol.2009.01.019.

Grigoriou, Konstantinos, and Frank T. Rothaermel. 2014. ‘Structural Microfoundations of Innovation: The Role of Relational Stars’. *Journal of Management* 40 (2): 586–615. https://doi.org/10.1177/0149206313513612.

———. 2017. ‘Organizing for Knowledge Generation: Internal Knowledge Networks and the Contingent Effect of External Knowledge Sourcing’. *Strategic Management Journal* 38 (2): 395–414. https://doi.org/10.1002/smj.2489.

Groysberg, Boris, and Linda Eling Lee. 2009. ‘Hiring Stars and Their Colleagues: Exploration and Exploitation in Professional Service Firms’. *Organization Science* 20 (4): 740–58. https://doi.org/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): 722–37. https://doi.org/10.1287/orsc.1100.0547.

Guan, Jiancheng, and Na Liu. 2016. ‘Exploitative and Exploratory Innovations in Knowledge Network and Collaboration Network: A Patent Analysis in the Technological Field of Nano-Energy’. *Research Policy* 45 (1): 97–112. https://doi.org/10.1016/j.respol.2015.08.002.

Hall, Kara L., Amanda L. Vogel, Grace C. Huang, Katrina J. Serrano, Elise L. Rice, Sophia P. Tsakraklides, and Stephen M. Fiore. 2018. ‘The Science of Team Science: A Review of the Empirical Evidence and Research Gaps on Collaboration in Science.’ *American Psychologist* 73 (4): 532–48. https://doi.org/10.1037/amp0000319.

Haseltine, FlorenceP., and Mark Chodos. 2017. '"Why" vs. "What," or "The Bad Penny Opera": Gender and Bias in Science'. *Technology & Innovation* 18 (4): 275–79. https://doi.org/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): 869–83. https://doi.org/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. https://doi.org/10.1002/smj.916.

Hunt, Jennifer, Jean Philippe Garant, Hannah Herman, and David J. Munroe. 2013. 'Why Are Women Underrepresented amongst Patentees?' *Research Policy* 42 (4): 831–43. https://doi.org/10.1016/j.respol.2012.11.004.

Jensen, Kyle, Balázs Kovács, and Olav Sorenson. 2018. 'Gender Differences in Obtaining and Maintaining Patent Rights'. *Nature Biotechnology* 36 (4): 307–9. https://doi.org/10.1038/nbt.4120.

Jones, Benjamin F. 2009. 'The Burden of Knowledge and the "Death of the Renaissance Man": Is Innovation Getting Harder?' *Review of Economic Studies* 76 (1): 283–317. https://doi.org/10.1111/j.1467-937X.2008.00531.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): 709–27. https://doi.org/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. https://doi.org/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): 1931–52. https://doi.org/10.1007/s11192-015-1646-y.

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: 133–59. https://doi.org/10.1016/S0742-3322(04)21005-1.

Larivière, Vincent, Chaoqun Ni, Yves Gingras, Blaise Cronin, and Cassidy R. Sugimoto. 2013. 'Bibliometrics: Global Gender Disparities in Science'. *Nature* 504 (7479): 211–13. https://doi.org/10.1038/504211a.

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

Legewie, Joscha, and Thomas A. DiPrete. 2014. 'The High School Environment and the Gender Gap in Science and Engineering'. *Sociology of Education* 87 (4): 259–80. https://doi.org/10.1177/0038040714547770.

Li, Weihua, Tomaso Aste, Fabio Caccioli, and Giacomo Livan. 2019. 'Early Coauthorship with Top Scientists Predicts Success in Academic Careers'. *Nature Communications* 10 (1): 5170. https://doi.org/10.1038/s41467-019-13130-4.

Liu, Haibo, Jürgen Mihm, and Manuel E. Sosa. 2018. 'Where Do Stars Come From? The Role of Star vs. Nonstar Collaborators in Creative Settings'. *Organization Science* 29 (6): 1149–69. https://doi.org/10.1287/orsc.2018.1223.

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. https://doi.org/10.1037/0022-3514.73.1.91.

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): 1663–93. https://doi.org/10.1007/s11192-018-2933-1.

Meng, Yu. 2016. 'Collaboration Patterns and Patenting: Exploring Gender Distinctions'. *Research Policy* 45 (1): 56–67. https://doi.org/10.1016/j.respol.2015.07.004.

Motro, Daphna, Trevor M. Spoelma, and Aleksander P. J. Ellis. 2021. 'Incivility and Creativity in Teams: Examining the Role of Perpetrator Gender.' *Journal of Applied Psychology* 106 (4): 560–81. https://doi.org/10.1037/apl0000757.

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. https://doi.org/10.5465/amr.2010.0403.

Perry-Smith, Jill E., and Christina E. Shalley. 2014. 'A Social Composition View of Team Creativity: The Role of Member Nationality-Heterogeneous Ties Outside of the Team'. *Organization Science* 25 (5): 1434–52. https://doi.org/10.1287/orsc.2014.0912.

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): 419–37. https://doi.org/10.1007/s11192-009-0037-7.

Reme, Bjørn-Atle, Andreas Kotsadam, Johannes Bjelland, Pål Roe Sundsøy, and Jo Thori Lind. 2022. 'Quantifying Social Segregation in Large-Scale Networks'. *Scientific Reports* 12 (1): 6474. https://doi.org/10.1038/s41598-022-10273-1.

Schmuck, Claudine. 2017. *Women in STEM Disciplines*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-41658-8.

Shannon, Geordan, Melanie Jansen, Kate Williams, Carlos Cáceres, Angelica Motta, Aloyce Odhiambo, Alie Eleveld, and Jenevieve Mannell. 2019. 'Gender Equality in Science, Medicine, and Global Health: Where Are We at and Why Does It Matter?' *The Lancet* 393 (10171): 560–69. https://doi.org/10.1016/S0140-6736(18)33135-0.

Singh, Jasjit, and Lee Fleming. 2010. 'Lone Inventors as Sources of Breakthroughs: Myth or Reality?' *Management Science* 56 (1): 41–56. https://doi.org/10.1287/mnsc.1090.1072.

Sugimoto, Cassidy R., Chaoqun Ni, Jevin D. West, and Vincent Larivière. 2015. 'The Academic Advantage: Gender Disparities in Patenting'. Edited by Alejandro Raul Hernandez Montoya. *PLOS ONE* 10 (5): e0128000. https://doi.org/10.1371/journal.pone.0128000.

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): 1581–91. https://doi.org/10.1177/0956797614537280.

Szell, Michael, and Stefan Thurner. 2013. 'How Women Organize Social Networks Different from Men'. *Scientific Reports* 3 (1): 1214. https://doi.org/10.1038/srep01214.

Tahmooresnejad, Leila, and Ekaterina Turkina. 2022. 'Female Inventors over Time: Factors Affecting Female Inventors' Innovation Performance'. *Journal of Informetrics* 16 (1): 101256. https://doi.org/10.1016/j.joi.2022.101256.

Tshetshema, Caspar T., and Kai-Ying Chan. 2020. 'A Systematic Literature Review of the Relationship between Demographic Diversity and Innovation Performance at Team-Level'. *Technology Analysis & Strategic Management* 32 (8): 955–67. https://doi.org/10.1080/09537325.2020.1730783.

USPTO. 2019. 'Progress and Potential: A Profile of Women Inventors on US Patents. In Technical Report'. *IP Data Highlights*.

———. 2020. 'Progress and Potential: 2020 Update on U.S. Women Inventor-Patentees'. Office of the chief economist IP data highlights. https://www.uspto.gov/ip-policy/economic-research/publications/reports/progress-potential.

Vakili, Keyvan, and Sarah Kaplan. 2021. 'Organizing for Innovation: A Contingency View on Innovative Team Configuration'. *Strategic Management Journal* 42 (6): 1159–83. https://doi.org/10.1002/smj.3264.

Wang, Chunlei, Simon Rodan, Mark Fruin, and Xiaoyan Xu. 2014. 'Knowledge Networks, Collaboration Networks, and Exploratory Innovation'. *Academy of Management Journal* 57 (2): 484–514. https://doi.org/10.5465/amj.2011.0917.

Wang, J., Grand H.-L. Cheng, Tingting Chen, and Kwok Leung. 2019. 'Team Creativity/Innovation in Culturally Diverse Teams: A Meta-analysis'. *Journal of Organizational Behavior* 40 (6): 693–708. https://doi.org/10.1002/job.2362.

Wang, Yukai, Zhongkai Yang, Lanjian Liu, and Xianwen Wang. 2020. 'Gender Bias in Patenting Process'. *Journal of Informetrics* 14 (3). https://doi.org/10.1016/j.joi.2020.101046.
Whittington, Kjersten Bunker. 2018. '"A Tie Is a Tie? Gender and Network Positioning in Life Science Inventor Collaboration"'. *Research Policy* 47 (2): 511–26. https://doi.org/10.1016/j.respol.2017.12.006.
Wuchty, Stefan, Benjamin F. Jones, and Brian Uzzi. 2007. 'The Increasing Dominance of Teams in Production of Knowledge'. *Science* 316 (5827): 1036–39. https://doi.org/10.1126/science.1136099.
Wullum Nielsen, Mathias, and Love Börjeson. 2019. 'Gender Diversity in the Management Field: Does It Matter for Research Outcomes?' *Research Policy* 48 (7): 1617–32. https://doi.org/10.1016/j.respol.2019.03.006.
Zhang, Gupeng, Xiao Wang, and Hongbo Duan. 2020. 'Obscure but Important: Examining the Indirect Effects of Alliance Networks in Exploratory and Exploitative Innovation Paradigms'. *Scientometrics* 124 (3): 1745–64. https://doi.org/10.1007/s11192-020-03586-3.
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): 502–13. https://doi.org/10.1177/0002764297040004012.

6 Appendix

6.1 Additional statistics and robustness tests

Table 5 Correlation matrix (N=1,768,245)

	Variable	1	2	3	4	5	6	7	8	9	10
1	Star (dummy = 1)	1.000									
2	Female dummy	-0.010	1.000								
3	Nr. of coinventors	0.233	0.022	1.000							
4	Share of fem. coinv.	-0.002	0.174	0.084	1.000						
5	Share of male coinv.	0.021	-0.525	0.178	-0.204	1.000					
6	Share of coinv. with >10yrs exp.	0.017	-0.005	0.113	-0.078	0.230	1.000				
7	Share of fem. coinv.with >10yrs exp.	0.001	0.053	0.041	0.318	-0.065	0.183	1.000			
8	Share of male coinv. with >10yrs exp.	0.018	-0.251	0.095	-0.132	0.410	0.884	0.030	1.000		
9	Share of star coinventors	0.279	0.010	0.765	0.006	0.109	0.226	0.052	0.197	1.000	
10	Share of fem. star coinv.	0.018	0.058	0.069	0.350	-0.072	0.047	0.421	-0.015	0.148	1.000
11	Share of male star coinv.	0.061	-0.229	0.155	-0.112	0.382	0.457	0.005	0.545	0.417	0.064

Table 6 Summary statistics for the sub-samples of male and female inventors in the last examined year.

Variables cumulated up to 2019	**Sub-sample**	
	Male inventors only	Female inventors only
Sample size	84,710	14,230
Portfolio size	5.97	4.63
Nr. of coinventors	17.77	15.93
Share of fem. Coinv.	0.11	0.18
Share of male Coinv.	0.80	
Share of coinv. With >10yrs exp.	0.26	0.23
Share of fem. Coinv. With >10yrs exp.	0.02	0.02
Share of male Coinv. With >10yrs exp.	0.25	
Share of star coinventors	0.33	0.31
Share of fem. star coinv.	0.03	0.04
Share of male star coinv.	0.30	

Table 7 Maximum likelihood estimation for parametric regression survival-time models (streg in Stata 17). Multiple-record per individual and single-failure as first time to become a star (above 95th percentile in the reference cohort). The survival models employed loglogistic distribution. Sample of male inventors only

Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Share of fem. coinv.	0.299***							
	(0.076)							
Share of male coinv.		0.023						
		(0.040)						
Share of coinv. with >10yrs exp.			0.121***					
			(0.043)					
Share of fem coinv. with >10yrs exp.				0.568*				
				(0.292)				
Share of male coinv. with >10yrs exp.					0.101**			
					(0.043)			
Share of star coinventors						-1.391***		
						(0.028)		
Share of fem. star coinv.							-1.193***	
							(0.085)	
Share of male star coinv.								-1.313***
								(0.029)
Nr. of coinventors	-0.094***	-0.094***	-0.094***	-0.094***	-0.094***	-0.072***	-0.092***	-0.076***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
wipo dummies	y	y	y	y	y	y	y	y
Constant	3.679***	3.684***	3.675***	3.696***	3.680***	4.037***	3.715***	3.992***
	(0.029)	(0.040)	(0.030)	(0.029)	(0.030)	(0.030)	(0.029)	(0.030)
lngamma	-0.412***	-0.411***	-0.413***	-0.412***	-0.412***	-0.475***	-0.415***	-0.463***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Obs	245447	245447	245447	245447	245447	245447	245447	245447
Log lik.	-23076.810	-23086.734	-23082.681	-23083.230	-23084.039	-22018.892	-23015.297	-22199.158
chi-squared	4318.047	4335.933	4311.176	4310.999	4308.099	7089.245	4544.326	6608.587

Table 8 Maximum likelihood estimation for parametric regression survival-time models (streg in Stata 17). Multiple-record per individual and single-failure as first time to become a star (above 95th percentile in the reference cohort). The survival models employed loglogistic distribution. Sample of female inventors only

Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Share of fem. coinv.	0.282*							
	(0.164)							
Share of male coinv.		-0.058						
		(0.113)						
Share of coinv. with >10yrs exp.			-0.001					
			(0.120)					
Share of fem coinv. with >10yrs exp.				0.872*				
				(0.512)				
Share of male coinv. with >10yrs exp.					-0.053			
					(0.123)			
Share of star coinventors						-1.617***		
						(0.078)		
Share of fem. star coinv.							-1.665***	
							(0.155)	
Share of male star coinv.								-1.417***
								(0.081)
Nr. of coinventors	-0.076***	-0.075***	-0.075***	-0.075***	-0.075***	-0.060***	-0.071***	-0.065***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)
wipo dummies	y	y	y	y	y	y	y	y
Constant	3.620***	3.700***	3.659***	3.646***	3.669***	4.178***	3.698***	4.064***
	(0.080)	(0.114)	(0.082)	(0.078)	(0.082)	(0.087)	(0.078)	(0.086)
lngamma	-0.433***	-0.432***	-0.433***	-0.434***	-0.432***	-0.504***	-0.447***	-0.477***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Obs	35210	35210	35210	35210	35210	35210	35210	35210
Log lik.	-2868.735	-2870.476	-2870.634	-2869.214	-2870.543	-2672.374	-2834.142	-2735.254
chi-squared	758.106	756.541	752.386	757.783	752.506	1323.507	904.889	1140.816

Table 9 Maximum likelihood estimation for parametric regression survival-time models (streg in Stata 17). Multiple-record per individual and single-failure as first time to become a star (above 95th percentile in the reference cohort). The survival models employed loglogistic distribution. Sample of star inventors only

Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Female dummy	-0.030	-0.031	-0.032	-0.030	-0.030	-0.030	-0.015	-0.019
	(0.028)	(0.028)	(0.027)	(0.028)	(0.028)	(0.028)	(0.023)	(0.027)
Share of fem. coinv.		0.011						
		(0.055)						
Share of male coinv.			-0.035					
			(0.032)					
Share of coinv. with >10yrs exp.				0.062				
				(0.039)				
Share of fem coinv. with >10yrs exp.					-0.048			
					(0.132)			
Share of male coinv. with >10yrs exp.						0.070*		
						(0.040)		
Share of star coinventors							-0.499***	
							(0.024)	
Share of fem. star coinv.								-0.258***
								(0.058)
Share of male star coinv.								
Nr. of coinventors	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	0.002***	-0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
wipo dummies	y	y	y	y	y	y	y	y
Constant	0.280***	0.279***	0.309***	0.269***	0.281***	0.269***	0.597***	0.292***
	(0.020)	(0.020)	(0.033)	(0.021)	(0.020)	(0.021)	(0.023)	(0.020)
lngamma	-1.597***	-1.597***	-1.597***	-1.597***	-1.597***	-1.597***	-1.635***	-1.598***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)
Obs	15076	15076	15076	15076	15076	15076	15076	15076
Log lik.	6039.858	6039.889	6040.697	6041.900	6039.947	6042.290	6325.033	6052.151
chi-squared	2033.212	2033.962	2040.961	2058.429	2034.588	2059.454	3833.775	2138.278

Table 10 Maximum likelihood estimation for parametric regression survival-time models (streg in Stata 17). Multiple-record per individual and single-failure as first time to become a star (above 95th percentile in the reference cohort). The survival models employed loglogistic distribution. Sample of inventors only having always at least one coinventor.

Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Female dummy	0.210***	0.194***	0.195***	0.211***	0.208***	0.211***	0.205***	0.235***	0.177***
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.031)	(-0.034)	(0.032)
Share of fem. coinv.		0.291***							
		(0.075)							
Share of male coinv.			-0.254***						
			(0.064)						
Share of coinv. with >10yrs exp.				0.059					
				(0.043)					
Share of fem coinv. with >10yrs exp.					0.628**				
					(0.291)				
Share of male coinv. with >10yrs exp.						0.034			
						(0.044)			
Share of star coinventors							-1.647***		
							(0.029)		
Share of fem. star coinv.								-1.357***	
								(0.079)	
Share of male star coinv.									-1.533***
									(0.030)
Nr. of coinventors	-0.094***	-0.094***	-0.094***	-0.094***	-0.094***	-0.094***	-0.073***	-0.092***	-0.078***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
wipo dummies	3.754***	3.730***	3.978***	3.739***	3.747***	3.746***	4.247***	3.770***	4.181***
Constant	(0.031)	(0.031)	(0.066)	(0.032)	(0.031)	(0.032)	(0.033)	(0.031)	(0.033)
	-0.402***	-0.402***	-0.402***	-0.403***	-0.403***	-0.402***	-0.486***	-0.408***	-0.468***
lngamma	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)
	239286	239286	239286	239286	239286	239286	239286	239286	239286
Obs	-22959.864	-22950.038	-22949.901	-22958.907	-22955.439	-22959.564	-21526.367	-22856.334	-21799.829
Log lik.	4366.189	4392.950	4383.131	4372.287	4373.813	4369.128	7607.057	4700.738	6939.781
chi-squared	0.210***	0.194***	0.195***	0.211***	0.208***	0.211***	0.205***	0.235***	0.177***

Table 11 Maximum likelihood estimation for parametric regression survival-time models (streg in Stata 17). Multiple-record per individual and single-failure as first time to become a star (above 95th percentile in the reference cohort). The survival models employed loglogistic distribution. Sample of inventors in fields where the presence of female is high (>19% of inventors): Analysis of biological materials, Organic fine chemistry, Biotechnology, Pharmaceuticals, Macromolecular chemistry, polymers, Food chemistry, Basic materials chemistry.

Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Female dummy	0.288***	0.276***	0.286***	0.288***	0.283***	0.288***	0.272***	0.304***	0.259***
	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.060)	(0.054)	(0.059)	(0.055)
Share of fem. coinv.		0.292**							
		(0.134)							
Share of male coinv.			-0.057						
			(0.101)						
Share of coinv. with >10yrs exp.				0.049					
				(0.093)					
Share of fem coinv. with >10yrs exp.					0.859*				
					(0.456)				
Share of male coinv. with >10yrs exp.						-0.004			
						(0.095)			
Share of star coinventors							-1.686***		
							(0.064)		
Share of fem. star coinv.								-1.545***	
								(0.158)	
Share of male star coinv.									-1.539***
									(0.066)
Nr. of coinventors	-0.077***	-0.077***	-0.077***	-0.077***	-0.077***	-0.077***	-0.061***	-0.075***	-0.065***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)
wipo dummies	4.022***	3.986***	4.066***	4.008***	4.006***	4.023***	4.481***	4.046***	4.398***
Constant	(0.067)	(0.068)	(0.102)	(0.070)	(0.067)	(0.070)	(0.071)	(0.067)	(0.070)
	-0.399***	-0.400***	-0.399***	-0.400***	-0.401***	-0.399***	-0.481***	-0.408***	-0.460***
lngamma	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
	82491	82491	82491	82491	82491	82491	82491	82491	82491
Obs	-6159.413	-6156.446	-6159.220	-6159.261	-6156.586	-6159.412	-5838.524	-6124.741	-5922.712
Log lik.	848.291	851.671	853.624	850.455	855.605	849.915	1409.093	902.515	1311.405
chi-squared	0.288***	0.276***	0.286***	0.288***	0.283***	0.288***	0.272***	0.304***	0.259***

Table 12 Maximum likelihood estimation for parametric regression survival-time models (streg in Stata 17). Multiple-record per individual and single-failure as first time to become a star (above 95th percentile in the reference cohort). The survival models employed loglogistic distribution. Sample of inventors in fields where the presence of female is low (<10% of inventors): Handling, Machine tools, Engines, pumps, turbines, Thermal processes and apparatus, Mechanical elements, Transport, Civil engineering.

Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Female dummy	0.120*	0.099	0.121**	0.120**	0.117*	0.120*	0.122**	0.136**	0.098*
	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.061)	(0.057)	(0.061)	(0.058)
Share of fem. coinv.		0.405***							
		(0.128)							
Share of male coinv.			0.018						
			(0.058)						
Share of coinv. with >10yrs exp.				0.023					
				(0.057)					
Share of fem coinv. with >10yrs exp.					0.606				
					(0.378)				
Share of male coinv. with >10yrs exp.						0.001			
						(0.058)			
Share of star coinventors							-1.238***		
							(0.042)		
Share of fem. star coinv.								-1.100***	
								(0.163)	
Share of male star coinv.									-1.182***
									(0.042)
Nr. of coinventors	-0.096***	-0.096***	-0.096***	-0.096***	-0.096***	-0.096***	-0.074***	-0.094***	-0.077***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
wipo dummies	3.556***	3.531***	3.542***	3.551***	3.550***	3.556***	3.836***	3.569***	3.803***
Constant	(0.040)	(0.040)	(0.056)	(0.041)	(0.040)	(0.041)	(0.042)	(0.040)	(0.042)
	-0.446***	-0.447***	-0.446***	-0.447***	-0.447***	-0.446***	-0.502***	-0.449***	-0.494***
lngamma	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
	105831	105831	105831	105831	105831	105831	105831	105831	105831
Obs	-10133.448	-10126.175	-10133.394	-10133.366	-10131.417	-10133.448	-9754.945	-10112.103	-9808.249
Log lik.	1311.943	1310.446	1327.479	1312.975	1316.054	1314.243	2132.662	1343.636	2028.342
chi-squared	0.120*	0.099	0.121**	0.120**	0.117*	0.120*	0.122**	0.136**	0.098*

6.2 Tests on the subsample of inventors having coinventors with identified gender

Gender in PatentsView data is not always identified: 5% in our starting sample of inventors debuting in 2000. We perform additional analyses on the subsample of inventors having only coinventors with identified gender: 80,881 inventors are included (78% of the population of inventors debuting in 2000; 82% of the main sample studied in this article).

This second approach introduces some bias with respect to larger teams (the solo inventors are in both the main and this sample), especially when involving teams with multiple nationalities. Indirectly, it affects the relative representation of stars in the sample. We have no elements to conjecture a gender polarization (i.e., missing gender data should not be biased towards one gender). Nonetheless, data reported in Table 13 shows that the sample with no missing gender among coinventors underestimate the presence of stars (as expected due to the average smaller team size in this second sample) and, slightly, of female stars.

Table 13 Samples

Sample	**Inventors debuting in 2000 with identified gender**	**Inventors debuting in 2000 with identified gender AND having all the coinventors with identified gender**
Size	98,940	80,881
Perc. on population of inventors debuting in 2000	95.0	77.7
Female inventors (perc.)	14.4	14.1
Avg. team size of coinv. (up to 2019)	17.51	9.67
Stars (perc. on sample)	15.2	10.6
Female stars (perc. on stars)	10.3	9.4
Delay for female to become star (in years)	-0.43 T-test is significant	-0.723 T-test is significant

Table 14 Maximum likelihood estimation for parametric regression survival-time models (streg in Stata 17). Multiple-record per individual and single-failure as first time to become a star (above 95th percentile in the reference cohort). The survival models employed loglogistic distribution. Sample of inventors having collaborators with identified gender only.

Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Female dummy	0.142***	0.130***	0.153***	0.143***	0.141***	0.143***	0.151***	0.163***	0.131***
	(0.040)	(0.040)	(0.040)	(0.039)	(0.039)	(0.039)	(0.036)	(0.040)	(0.037)
Share of fem. coinv.		0.210***							
		(0.081)							
Share of male coinv.			0.184***						
			(0.042)						
Share of coinv. with >10yrs exp.				0.158***					
				(0.049)					
Share of fem coinv. with >10yrs exp.					0.419				
					(0.314)				
Share of male coinv. with >10yrs exp.						0.146***			
						(0.050)			
Share of star coinventors							-1.302***		
							(0.032)		
Share of fem. star coinv.								-1.268***	
								(0.092)	
Share of male star coinv.									-1.236***
									(0.033)
Nr. of coinventors	-0.103***	-0.104***	-0.105***	-0.103***	-0.103***	-0.103***	-0.079***	-0.101***	-0.083***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
wipo dummies	3.764***	3.750***	3.622***	3.728***	3.760***	3.732***	4.040***	3.771***	4.018***
Constant	(0.033)	(0.033)	(0.045)	(0.034)	(0.033)	(0.034)	(0.034)	(0.032)	(0.034)
	-0.414***	-0.414***	-0.416***	-0.417***	-0.415***	-0.417***	-0.472***	-0.420***	-0.462***
lngamma	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
	195338	195338	195338	195338	195338	195338	195338	195338	195338
Obs	-16388.132	-16383.928	-16378.495	-16382.423	-16386.497	-16383.449	-15672.336	-16323.073	-15775.326
Log lik.	4122.369	4111.825	4080.127	4143.800	4131.127	4138.445	5859.261	4310.775	5505.828
chi-squared	0.142***	0.130***	0.153***	0.143***	0.141***	0.143***	0.151***	0.163***	0.131***