

A multi-layered approach to the gender gap in patenting: Which layer best predicts women's participation?

EPOG+

**Elodie Carpentier, Intan Hamdan-Livramento
and Julio Raffo**

December 8, 2023

Motivation

- Women's underrepresentation among inventors is not only unfair to women (who are deprived of economic and reputational benefits) but also holds back the economic and technological progress of nations.
- *What prevents women from participating?* The many factors start in childhood and continue right up to the time they enter an innovation-related profession, most often unconsciously.

Barriers to women's participation (1)

Some factors are of a systematic nature, for example when they are linked to the culture and institutions of the woman's **country** of origin or residence.

- Mothers and married women are less likely to patent, and more geographically constrained (Whittington, 2011; Delgado et al., 2019), most likely because they carry the greater share of domestic and caregiving workload.
- Women are less often credited for their contribution (Jensen et al., 2018; Lissoni et al., 2013), slowing down their career progression (unconscious gender biases, on the best scenario!).

Other factors relate more specifically to the woman's **work ecosystem** and **industry**.

- Women are more likely to patent when employed in the academic sector, as opposed to corporate or government contexts (Martinez et al., 2016; Sugimoto et al., 2015; WIPO, 2023).
- Women inventors tend to cluster in chemistry-related fields, while few of them are working in electrical and mechanical engineering industries (Martinez et al., 2016). The biotechnology industry offers women greater career opportunities (Eaton, 1999).

Barriers to women's participation (2)

The characteristics of women's specific employing organizations further determine their participation.

- Becker's theory (1957) predicts that competitive pressure drives **taste-based discrimination** out of the market.
 - Supportive empirical evidence: gender employment in the banking industry (Ashenfelter and Hannan, 1986; Black and Strahan, 2001).
- Exporting manufacturing firms, that require communicating with customers outside of **regular working hours and traveling** on short notice, display higher gender wage gaps than non-exporting firms (Boler et al., 2018).
- Italian universities that opened a formal technology transfer office and established clear patent regulations display greater participation of women in patenting (Giuri et al., 2020). (**Greater transparency**)

Women do not enter the competitive sectors (**self-selection**) vs. competition affecting their ability to innovation (**impact**)

Barriers to women's participation (3)

At an even more micro level, some factors are specific to **group dynamics**.

- Women are more likely to be part of **larger research teams** than men (Moody, 2004; Martinez et al., 2016).
- All-women and women-led teams of inventors are 35% more likely to innovate in areas of women's health than all-male teams (Koning, Samila, and Ferguson, 2021).
- Academic women sometimes give up their patent inventorship rights to secure scientific authorship, especially when they are in **junior positions** or patents are privately owned (Lissoni et al., 2013; Lissoni et al., 2020).

Barriers to women's participation (end)

Finally, the socio-economic background of **individual** inventors also comes into play.

- Children from low-income families, black children and girls are less likely to become inventors. If girls were exposed to **female role models** at the rate boys are exposed to male role models, the gender gap in patenting would shrink by half (Bell et al., 2019).
- When parents have a **second-born son** instead of a daughter, they are less likely to inspire their first-born daughter to become an inventor herself (Hoisl et al., 2022).

Some factors may

- belong to **more than one layer**, or move across layers with new regulations or systemic cultural changes for instance;
- interact with factors in other layers (e.g. motherhood affects women differently across work ecosystems).

Research question

Which layer, like work ecosystem, industry, or group dynamic, is the greatest predictor of the gender gap in patenting?

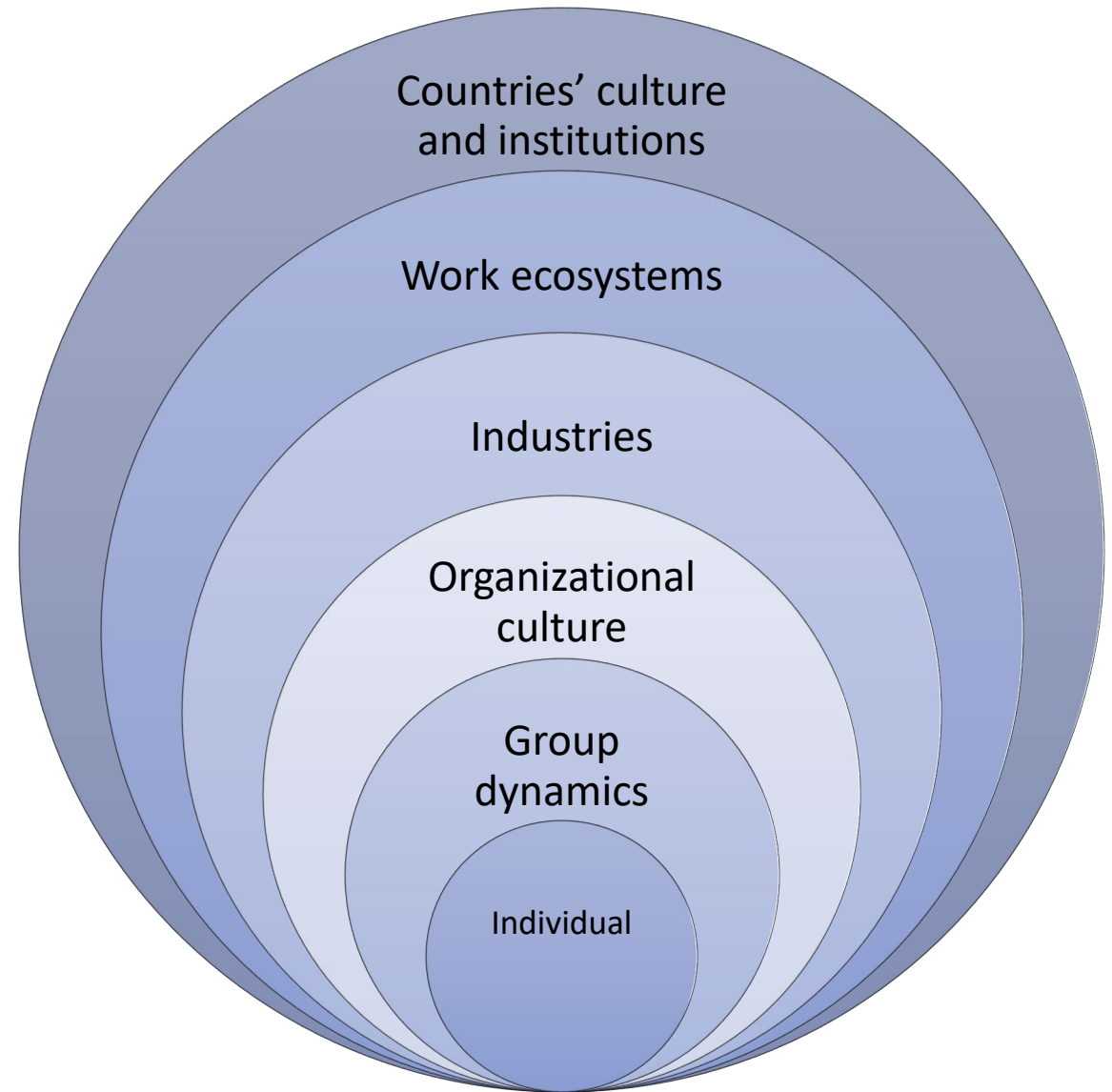


Figure 1: Six layers in the innovation gender gap.

Data

All 3,687,219 international patent applications filed through the PCT system from 1999 to 2020

1. Country of applicants and inventors
2. Applicants type (company, university, individual)
3. IPC classes as proxies for industries
4. Applicant names*
5. List of inventors as teams
6. Individual inventor names*

* Data work in progress. These two layers remain to disambiguate and merge to external data sources.

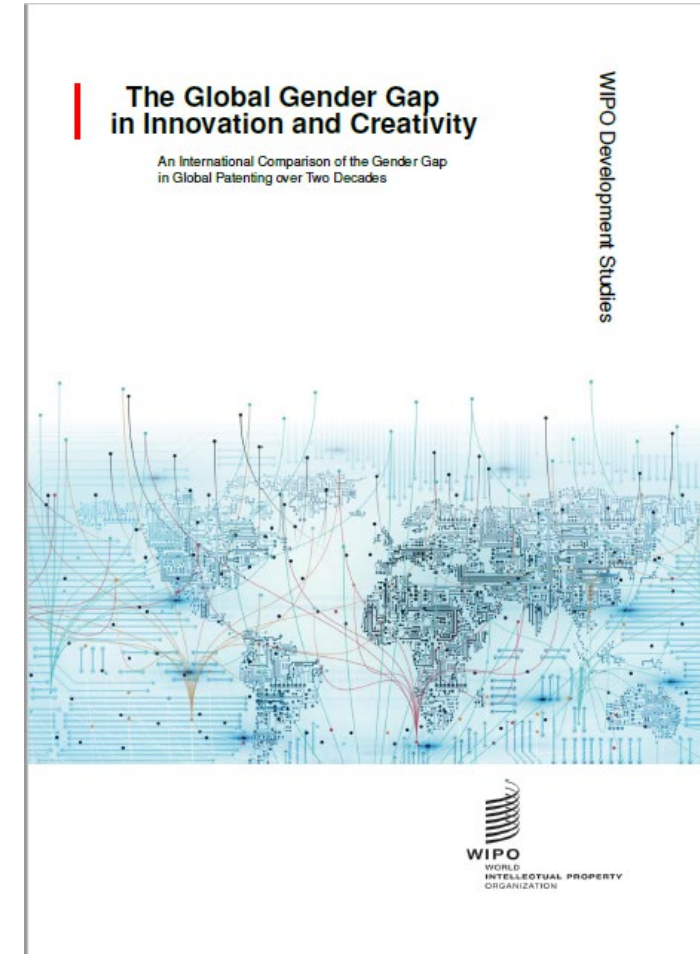
Gender assigned to 7,635,406 (96%) of inventors using WIPO's World Gender Name Dictionary 2.0 (and companion genderit STATA command).

Final dataset includes **2,987,388 patents**.

Descriptive analysis: Stylized facts

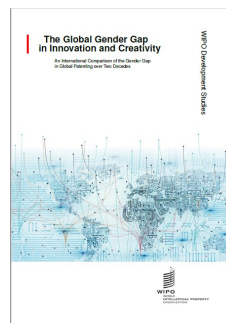
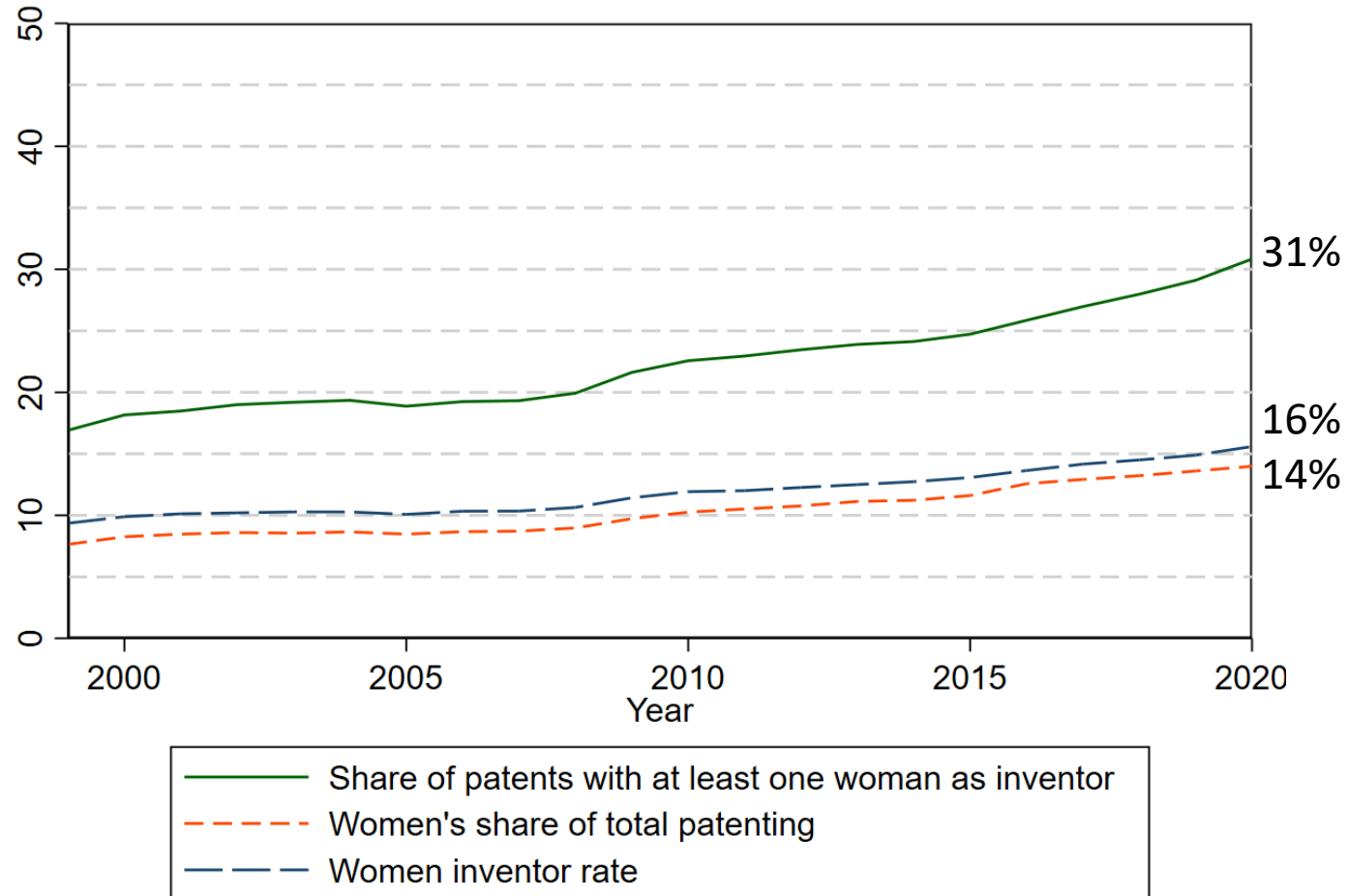
What is the share of women in patenting?

- 13.2% in Europe in 2019 (EPO, 2022)
- 12.8% in the US in 2019 (USPTO, 2020)
- 15% in Australia in 2021 (IP Australia, 2022)
- 16% in the World in 2020 (157 PCT states; WIPO, 2023)



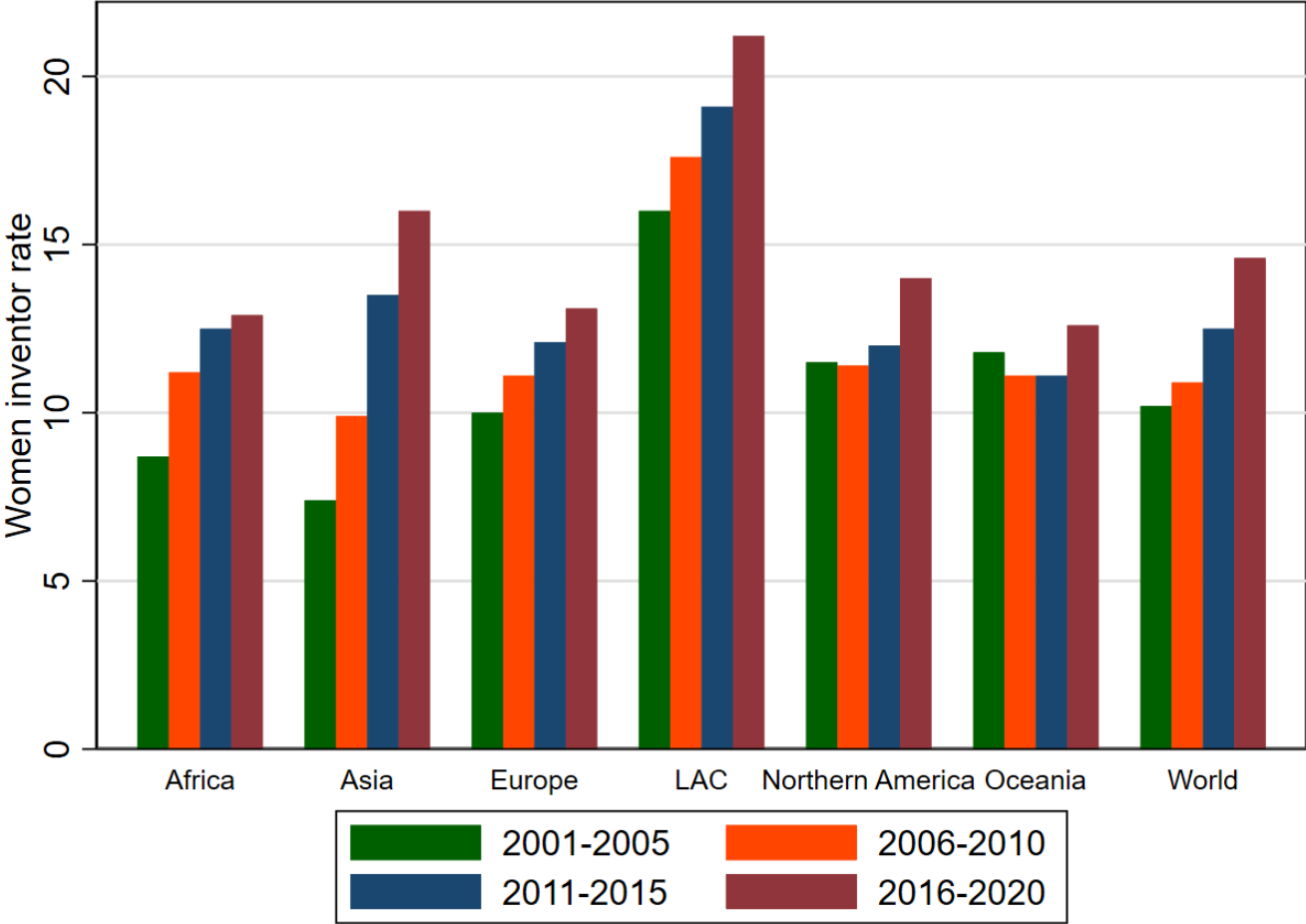
Women represent just 16% of inventors in 2020.

Figure 1: Three metrics over time



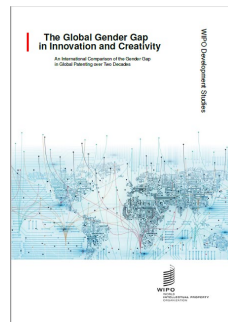
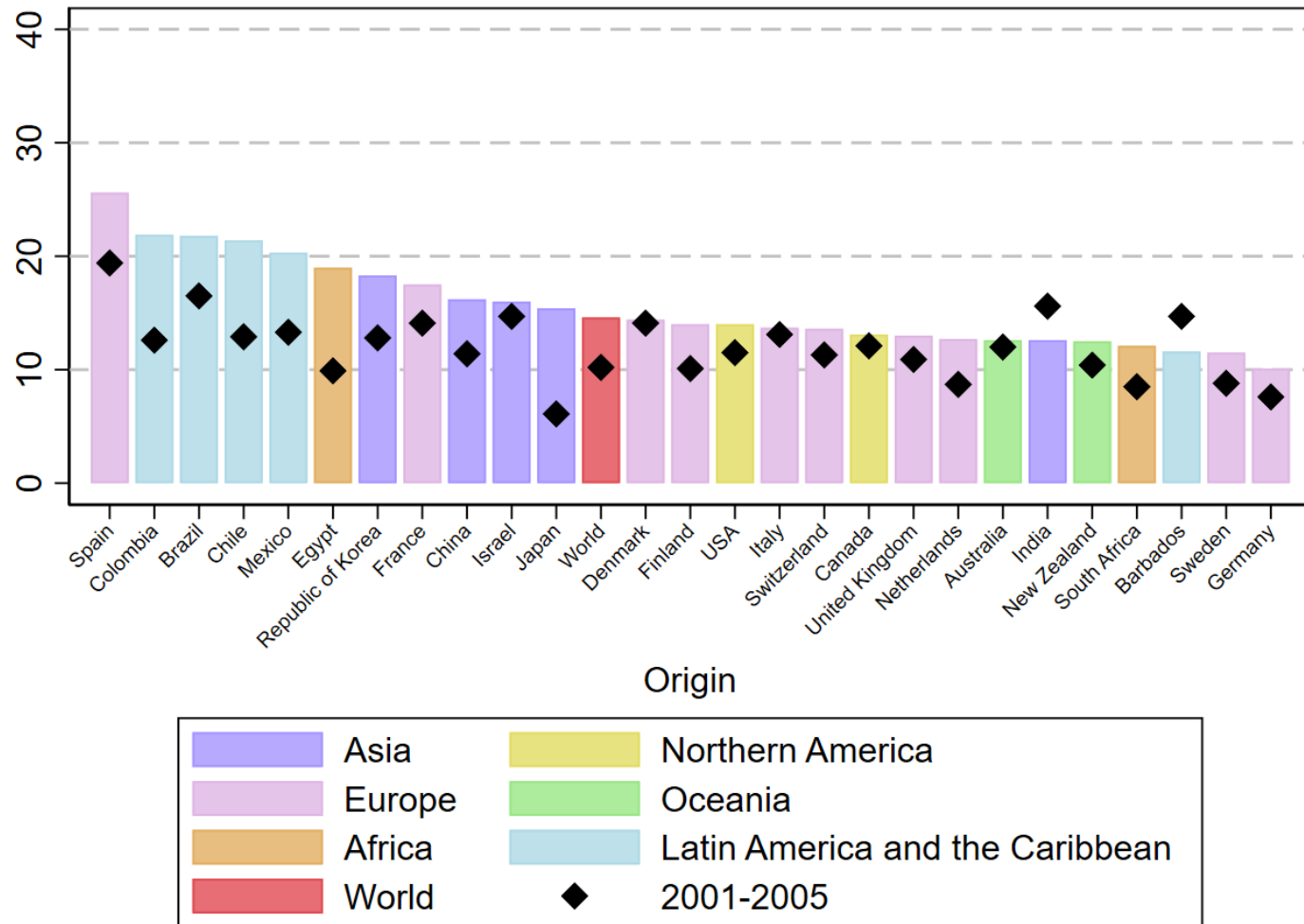
Noticeable regional differences...

Figure 2: Women inventor rate over continents and time



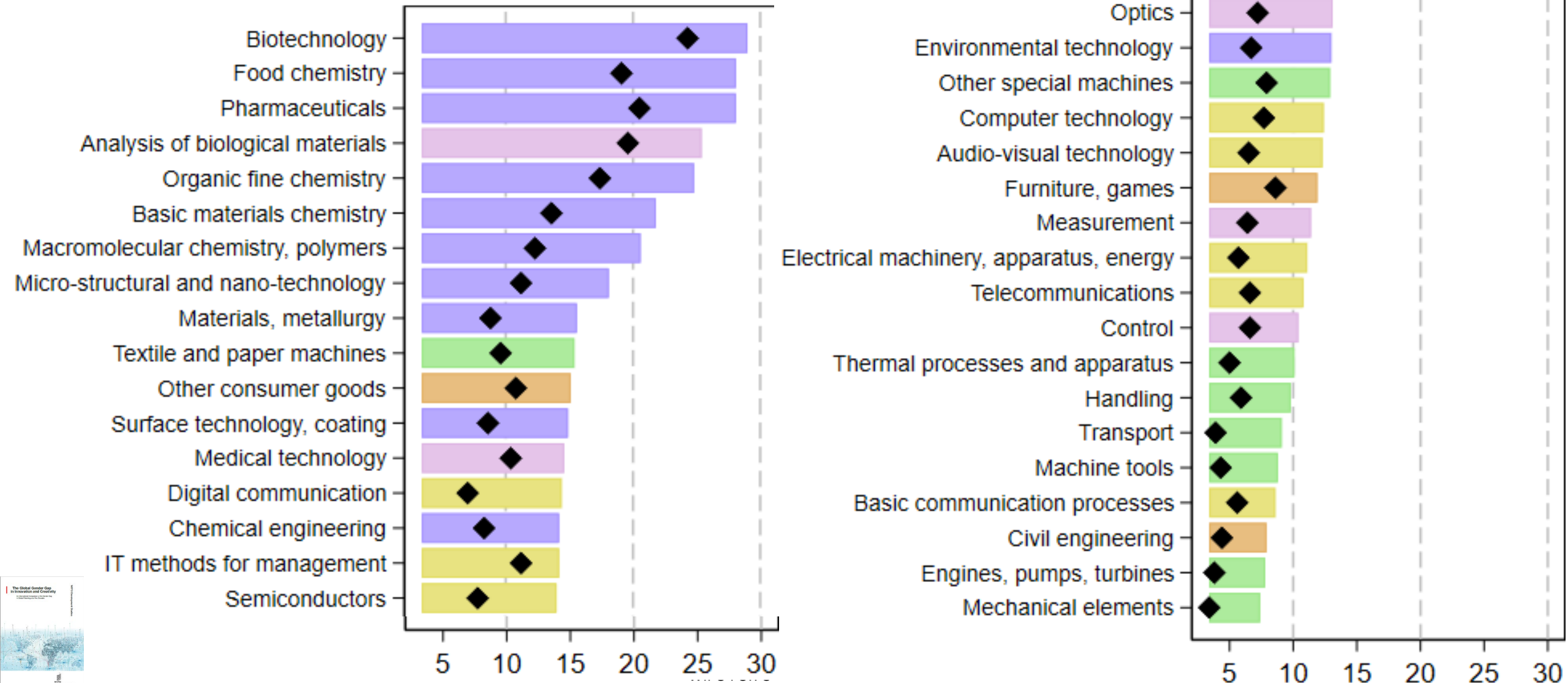
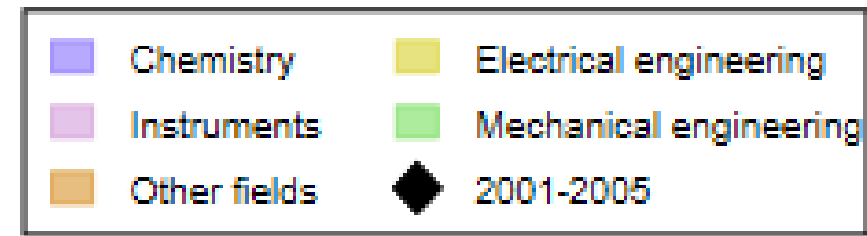
...and among countries.

Figure 3: Women inventor rate over top patenting countries (2016-2020)



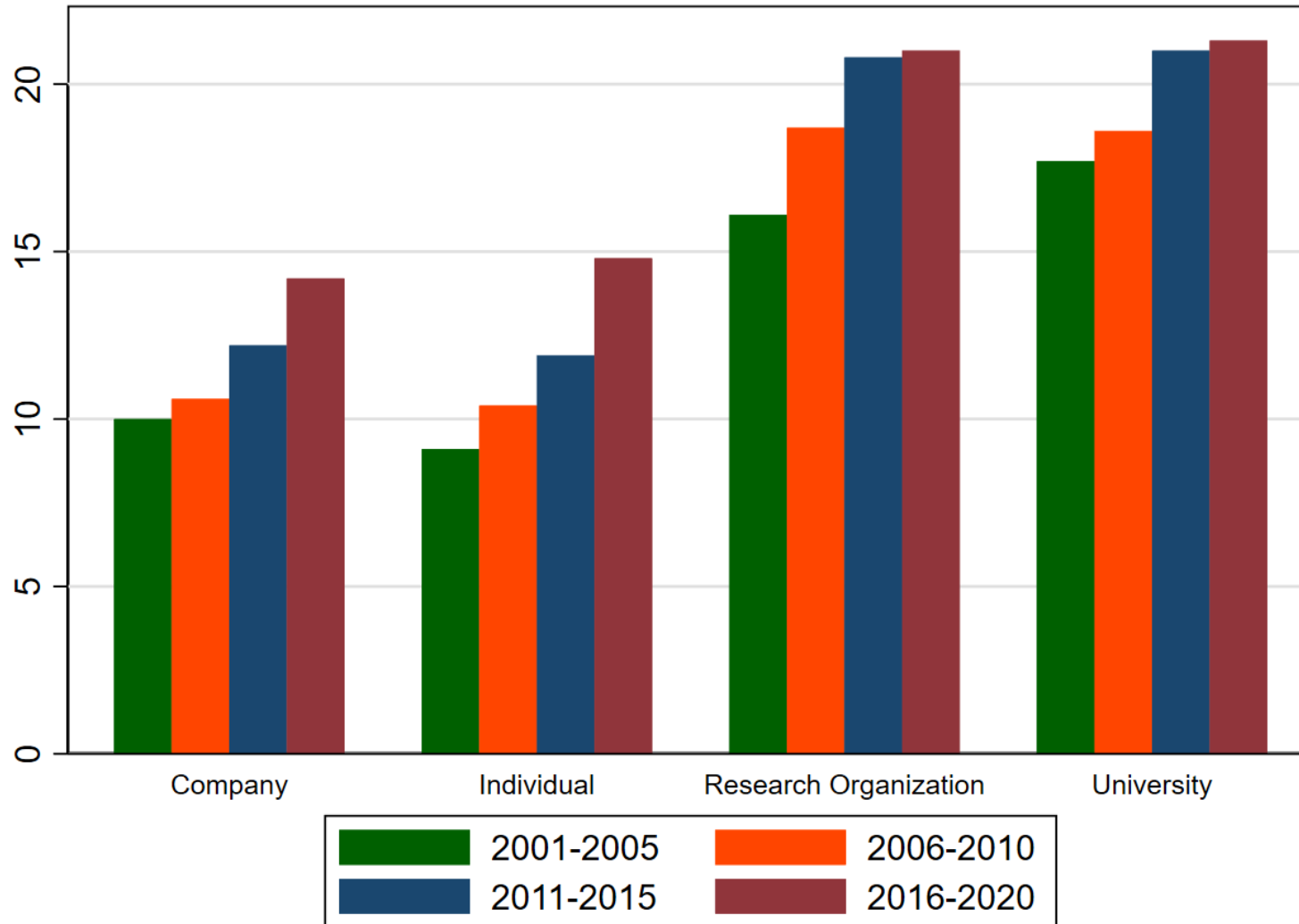
Also strong field differences

Figure 4: Women inventor rate in each industry (2016-2020)



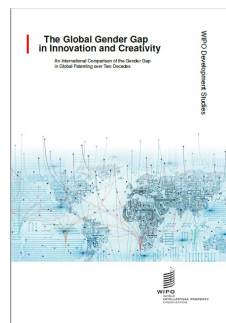
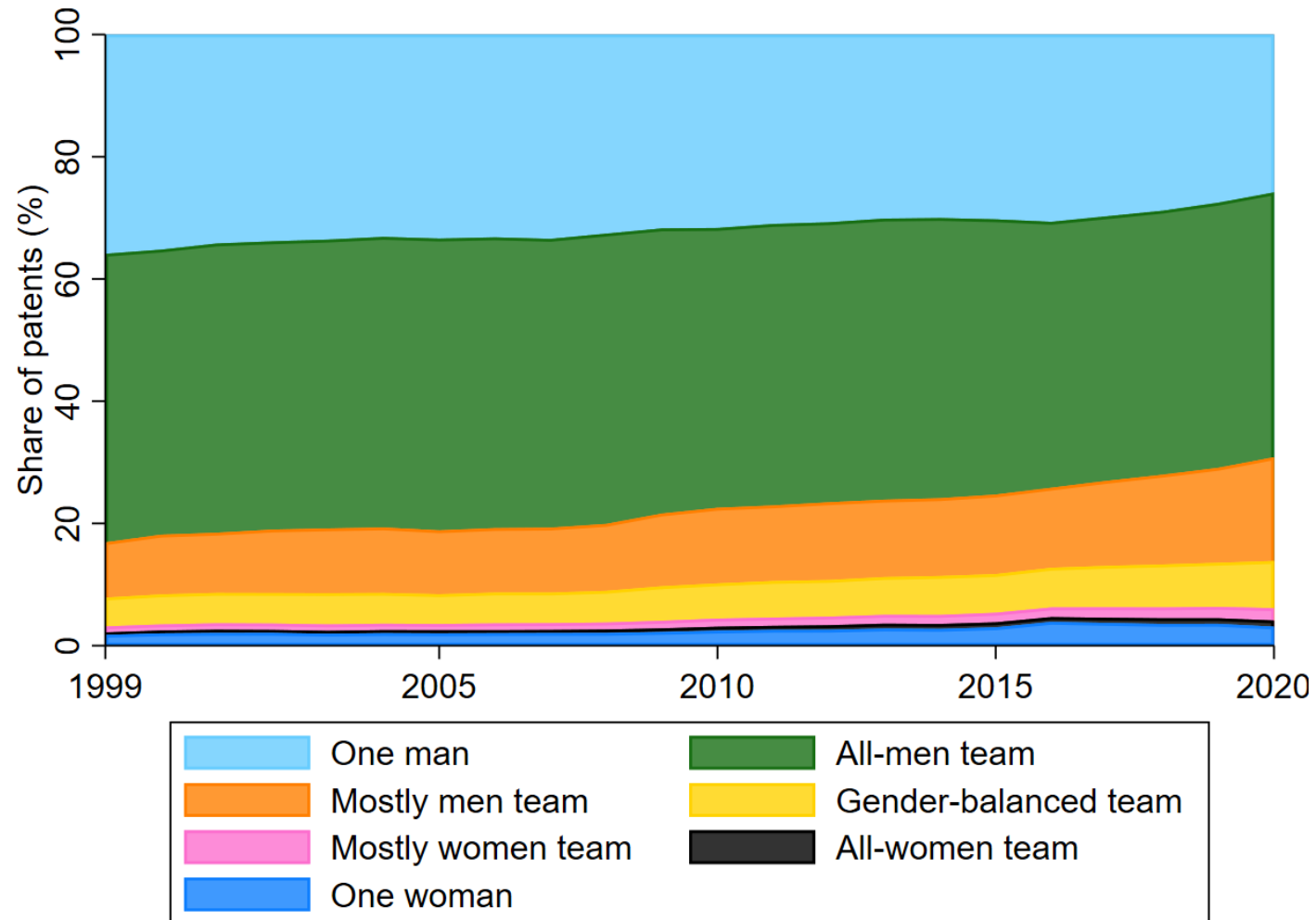
Academia is more inclusive than industry.

Figure 6: Women inventor rate over sectors



Increase in collaboration explains most of the rise in women participation

Figure 7: Gender composition of inventor teams



Methodology

We use the following set of logistic regressions at the inventor-patent level:

(1) 4 partial models: $\Pr(Woman = 1) = F(\alpha + \beta_l Layer_l + \gamma_t)$,
where $l = 1, \dots, 4$ corresponding to the 4 layers.

(2) 1 full model: $\Pr(Woman = 1) = F(\alpha + \beta_1 Country_p + \beta_2 Sector_p + \beta_3 Industry_p + \beta_4 Team_{p,i} + \gamma_t)$

Where

- $F(z) = e^z / (1 + e^z)$ is the cumulative logistic distribution
- $Woman = 1$ if the inventor on the patent is a woman, 0 otherwise
- $Country$: 30 dummies for each top patenting country and 1 for the rest of the world
- $Sector$: 4 dummies (university, company, research organization, or individual)
- $Industry$: 35 IPC classes (full counts)
- $Team$:
 - Number of inventors: categorical variable (1 person = solo, 2-3 people = small, 4-10 = medium, 11+ = large)
 - Has a female co-inventor (0/1)
- γ_t is a year-fixed effect

Cross-correlation table (selected coefficients)

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
(1) Woman (0/1)	1.00																			
(2) Woman co-inv. (0/1)	0.15	1.00																		
(3) Solo inventor	-0.05	-0.24	1.00																	
(4) Small team (2-3 p.)	-0.03	-0.20	-0.33	1.00																
(5) Middle-size team (4-10 p.)	0.05	0.32	-0.34	-0.74	1.00															
(6) Large team (11+ p.)	0.03	0.16	-0.05	-0.12	-0.12	1.00														
(7) Nb of inventors	0.07	0.42	-0.45	-0.43	0.58	0.55	1.00													
(8) Academia	0.05	0.08	-0.05	-0.01	0.05	0.00	0.05	1.00												
(9) Company	-0.03	0.00	-0.16	0.04	0.07	0.01	0.09	-0.67	1.00											
(10) Individual	-0.01	-0.08	0.27	-0.04	-0.14	-0.02	-0.16	-0.05	-0.71	1.00										
(11) Electronics	-0.03	-0.05	-0.00	0.01	-0.00	-0.02	-0.02	-0.02	0.03	-0.02	1.00									
(12) Biopharma	0.16	0.26	-0.10	-0.06	0.11	0.08	0.17	0.11	-0.05	-0.04	-0.13	1.00								
(13) Chemicals	0.03	0.06	-0.04	-0.01	0.04	0.00	0.04	-0.01	0.02	-0.02	-0.06	-0.01	1.00							
(14) Engines and transport	-0.07	-0.11	0.05	0.02	-0.05	-0.02	-0.06	-0.05	0.01	0.03	-0.02	-0.17	-0.08	1.00						
(15) Civil engineering	-0.03	-0.05	0.05	0.01	-0.04	-0.01	-0.05	-0.02	-0.03	0.06	-0.04	-0.08	-0.02	0.02	1.00					
(16) Spain	0.03	0.04	0.01	-0.02	0.02	0.01	0.02	0.09	-0.08	0.03	-0.01	0.04	-0.00	0.00	0.01	1.00				
(17) Rep. of Korea	0.02	0.02	0.03	-0.02	0.00	-0.01	-0.02	0.00	-0.03	0.04	0.02	-0.02	-0.01	-0.02	-0.01	-0.02	1.00			
(18) Germany	-0.04	-0.03	-0.02	-0.02	0.03	0.03	0.05	-0.02	0.04	-0.03	0.03	-0.01	0.05	0.13	-0.01	-0.04	-0.06	1.00		
(19) France	0.02	0.02	-0.01	0.03	-0.02	-0.02	-0.03	0.04	-0.02	-0.01	-0.01	0.02	0.00	0.05	-0.00	-0.02	-0.04	-0.08	1.00	
(20) China	0.03	0.04	0.04	-0.04	0.00	0.01	0.01	0.06	-0.06	0.03	-0.01	-0.06	-0.06	-0.04	-0.01	-0.03	-0.05	-0.11	-0.06	1.00

Regression tables

- Having a woman co-inventor increases by 7% the chances that a randomly chosen inventor is a woman.

Table 5: Random inventor's likelihood to be a woman (ME)

	(1)	(2)	(3)
	Team layer	Industry layer	Full model
Team			
Team size (ref. solo inventor)			
<i>Small team (2-3)</i>	0.0183** (2.26)		0.0182** (2.43)
<i>Medium team (4-10)</i>	0.0200* (1.91)		0.0181* (1.66)
<i>Large team (11+)</i>	0.0229** (2.26)		0.0159 (1.31)
Has a woman co-inventor	0.105*** (13.05)		0.0695*** (15.83)
Industry			
Electronics		-0.0248*** (-3.37)	-0.0205*** (-3.05)
Audio Visual		-0.0132*** (-4.88)	-0.0116*** (-3.76)
ICTs		-0.00518 (-0.87)	-0.00622 (-1.23)
Semiconductors & Optics		-0.00656*** (-3.72)	-0.00639*** (-2.81)
Instruments		0.000941 (0.28)	0.00226 (1.09)
Biopharma		0.120*** (15.49)	0.0939*** (22.57)
Chemicals		0.0238*** (8.81)	0.0229*** (8.21)
Materials		0.0152*** (3.91)	0.0121*** (5.09)
Chemistry & Environment		-0.0162*** (-4.07)	-0.0144*** (-4.19)
Engineering & Technology		-0.00103 (-0.48)	-0.000125 (-0.07)
Machines		-0.0273*** (-4.80)	-0.0210*** (-4.61)
Engines & Transport		-0.0541*** (-5.65)	-0.0454*** (-5.02)
Consumer		0.00895 (1.63)	0.0120*** (2.59)
Civil Engineering		-0.0485*** (-8.43)	-0.0433*** (-9.16)
Country FE			X
Sector FE			X
Year FE	X	X	X
Observations	7624777	7624777	7624777

Regression tables (cont'd)

Table 6: Random inventor's likelihood to be a woman (ME)

	(1)	(2)	(3)				
	Working ecosystem layer		Country layer		Full model		
Sector (ref. Academia)							
Company	-0.0756***	(-8.70)	-0.0212***	(-5.63)			
Individual	-0.0781***	(-5.45)	-0.00538	(-0.91)			
Country (ref. USA)							
Algeria		0.0679*** (98.70)	0.0771*** (15.37)				
Argentina		0.108*** (64.73)	0.0348*** (11.30)				
Australia		-0.00396*** (-7.32)	-0.000569 (-1.21)				
Barbados		0.0190*** (7.31)	0.0323*** (12.35)				
Brazil		0.0715*** (46.99)	0.0465*** (23.14)				
Canada		-0.00263*** (-19.57)	-0.000667 (-1.57)				
Chile		0.0631*** (13.53)	0.0111*** (3.55)				
China		0.0206*** (4.72)	0.0290*** (6.17)				
Colombia		0.0724*** (18.76)	0.0418*** (12.79)				
Denmark		0.0184*** (158.02)	-0.00262* (-1.82)				
Egypt		0.0216*** (24.32)	0.0157*** (3.89)				
Finland		-0.00703*** (-13.65)	0.0126*** (6.15)				
France		0.0333*** (71.04)	0.0310*** (18.73)				
Germany		-0.0356*** (-930.76)	-0.0265*** (-29.48)				
India		0.00237 (1.53)	-0.0244*** (-18.95)				
Israel		0.0253*** (192.32)	0.0166*** (65.31)				
Italy		0.0112*** (28.68)	0.0112*** (9.40)				
Japan		-0.0129*** (-12.67)	0.00477 (1.62)				
Mexico		0.0535*** (42.74)	0.0198*** (12.86)				
Morocco		0.0354*** (10.29)	0.0195*** (6.75)				
Netherlands		-0.0136*** (-60.70)	-0.00890*** (-18.99)				
New Zealand		-0.00978*** (-22.19)	-0.0116*** (-26.24)				
Rep. of Korea		0.0341*** (11.42)	0.0368*** (9.20)				
South Africa		-0.0157*** (-13.99)	-0.00360** (-2.24)				
Spain		0.106*** (104.10)	0.0550*** (19.20)				
Sweden		-0.0247*** (-1449.49)	-0.00567*** (-2.68)				
Switzerland		0.000862* (1.82)	-0.0145*** (-11.67)				
United Kingdom		-0.00625*** (-17.59)	-0.00972*** (-17.65)				
Rest of the World		0.00690 (0.63)	0.00175 (0.29)				
Team FE						X	
Industry FE						X	
Year FE	X	X	X				
Observations	7624777	7624777	7624777				

Preliminary Findings

- After regressions, ROC curves show the performance of a binary classifier with different decision thresholds.
- Industry FE improved the model performance by
 - 30% as compared to a random guess, and
 - 17% as compared to the baseline including year FE.
- Country FE and team characteristics: 15% and 5% resp.
- Work ecosystem FE: 12% and 2% resp.

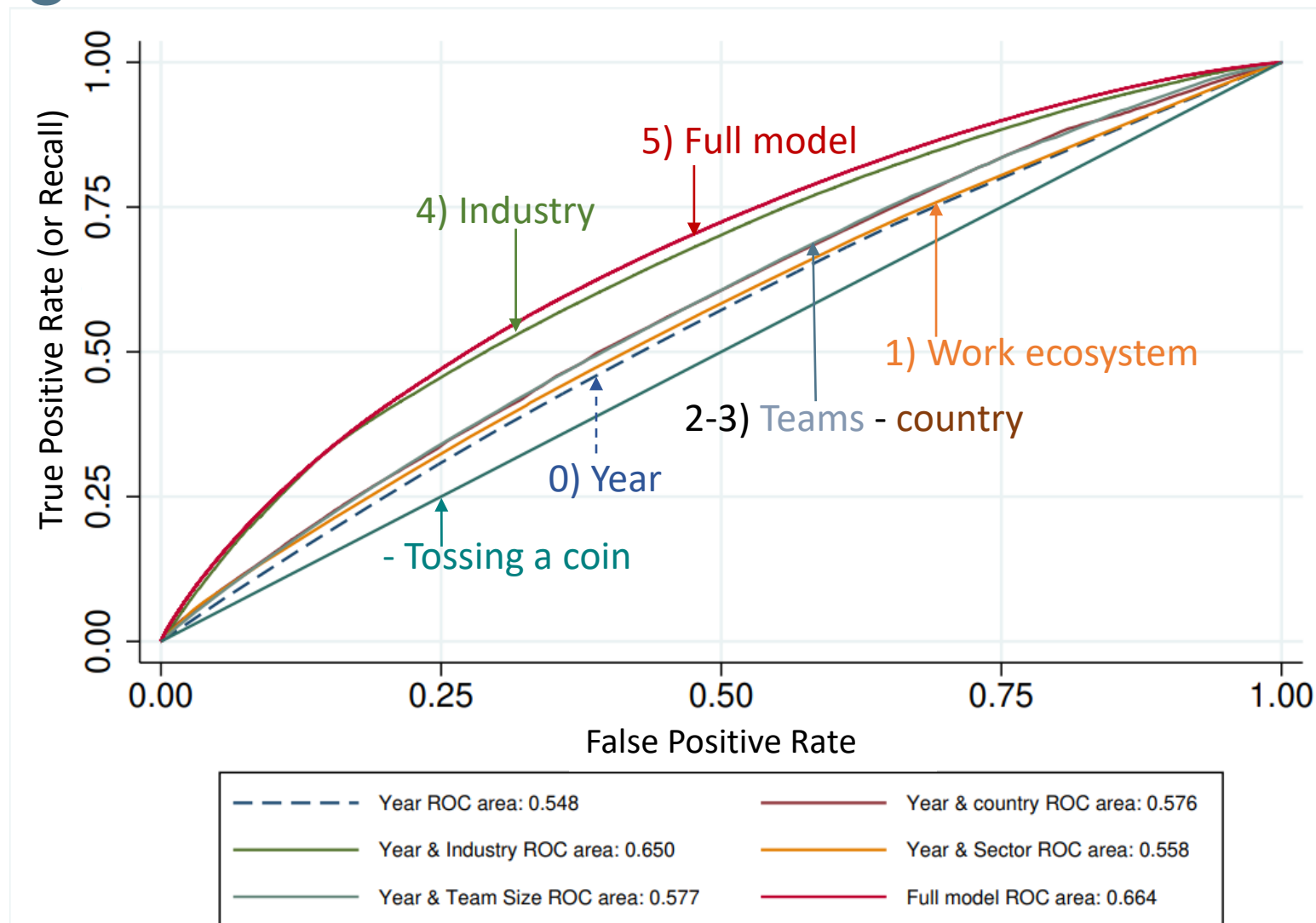


Figure 2: Predictive power of four layers in the innovation gender gap.

Preliminary Findings (cont'd)

- The full model only performs 33% better than a random guess, and 21% better than the year FE model.
- Accounting for four layers of factors, there remains a sizable unexplained gender innovation gap.

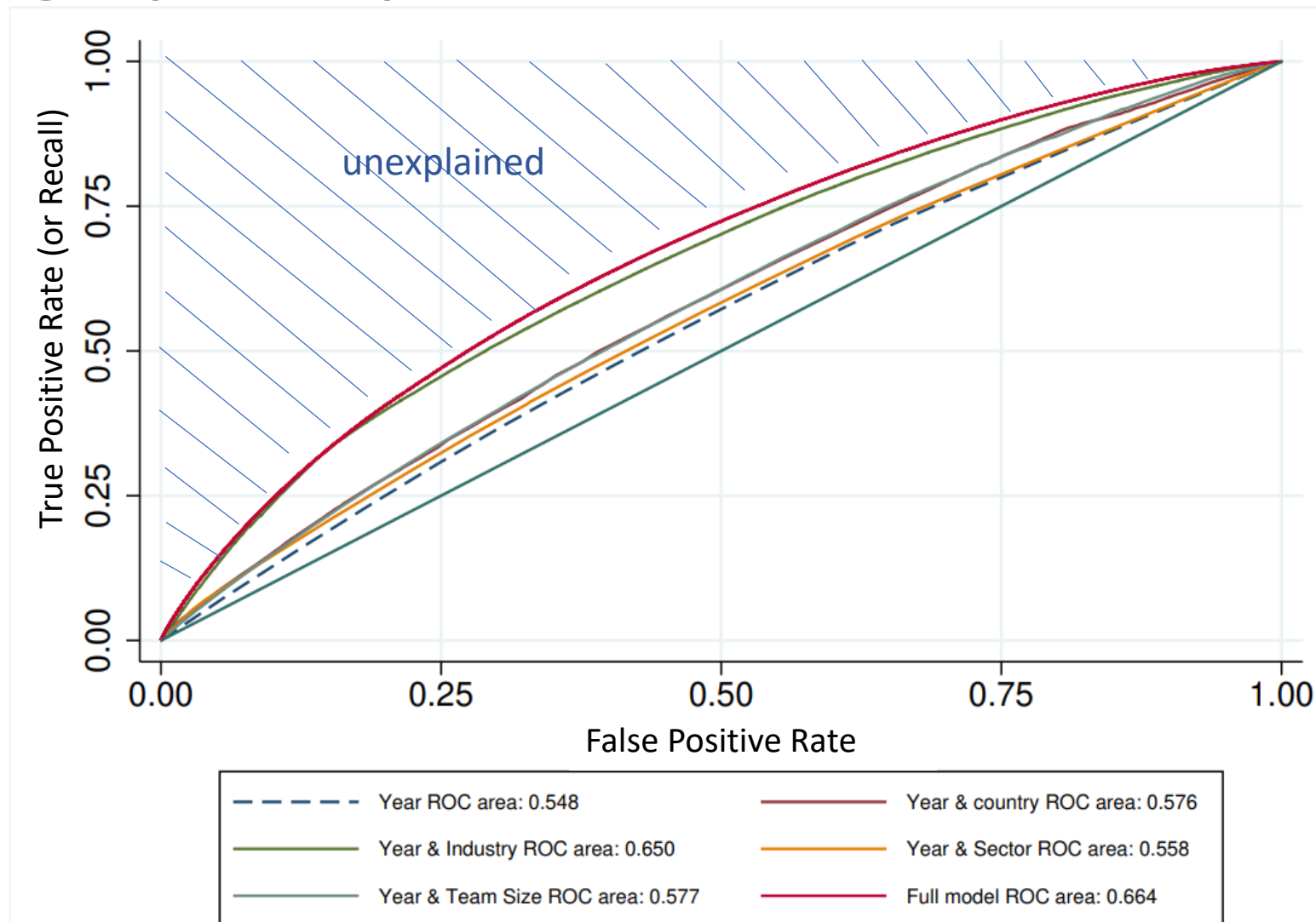


Figure 2: Predictive power of four layers in the innovation gender gap.

Discussion

Digging into the industry layer,

- The growing prevalence of increasingly gender-inclusive industries (such as ICT) may be confounded with the effect of countries' culture and institutions
 - Interaction of industry and year FE
- Educational and occupational sorting may be confounded with industries' culture and norms
 - Percentage of women among STEM graduates, Percentage of women in R&D occupations
 - Interaction of country and industry FE
- In concentrated industries, organizations' cultures and norms may be confounded with the characteristics of industries.
 - Repeat the analysis removing the most concentrated industries.

Further research

1. We are still missing two key layers: organizations, and individuals.

- Disambiguate patent applicants to include organization FE + other variables (degree of inclusivity, internationalization, organization size)

2. Imprint theory: Is it the culture of the inventors' country of origin or that of the organization's home country that matters most?

Thanks for your attention!

Any questions?



julio.raffo@wipo.int

Back Up Slides

Regression tables

- Having a woman co-inventor increases by 7% the chances that a randomly chosen inventor is a woman.

Table 5: Random inventor's likelihood to be a woman (ME)

	(1)	(2)	(3)
	Team layer	Industry layer	Full model
Team			
Nb of inventors	0.00111* (1.66)		0.000442 (0.52)
Has a woman co-inventor	0.105*** (13.05)		0.0695*** (15.83)
Industry			
Electronics		-0.0248*** (-3.37)	-0.0205*** (-3.05)
Audio Visual		-0.0132*** (-4.88)	-0.0116*** (-3.76)
ICTs		-0.00518 (-0.87)	-0.00622 (-1.23)
Semiconductors & Optics		-0.00656*** (-3.72)	-0.00639*** (-2.81)
Instruments		0.000941 (0.28)	0.00226 (1.09)
Biopharma		0.120*** (15.49)	0.0939*** (22.57)
Chemicals		0.0238*** (8.81)	0.0229*** (8.21)
Materials		0.0152*** (3.91)	0.0121*** (5.09)
Chemistry & Environment		-0.0162*** (-4.07)	-0.0144*** (-4.19)
Engineering & Technology		-0.00103 (-0.48)	-0.000125 (-0.07)
Machines		-0.0273*** (-4.80)	-0.0210*** (-4.61)
Engines & Transport		-0.0541*** (-5.65)	-0.0454*** (-5.02)
Consumer		0.00895 (1.63)	0.0120*** (2.59)
Civil Engineering		-0.0485*** (-8.43)	-0.0433*** (-9.16)
Country FE			X
Sector FE			X
Year FE	X	X	X
Observations	7624777	7624777	7624777