

Gender Gaps in Access to Medical Intern Positions: The Role of Competition

M. Díez-Rituerto , J. Gardeazabal, N. Iriberry, P. Rey-Biel

University of the Basque Country (UPV/EHU), and ESADE Business School (U. Ramón Lull)

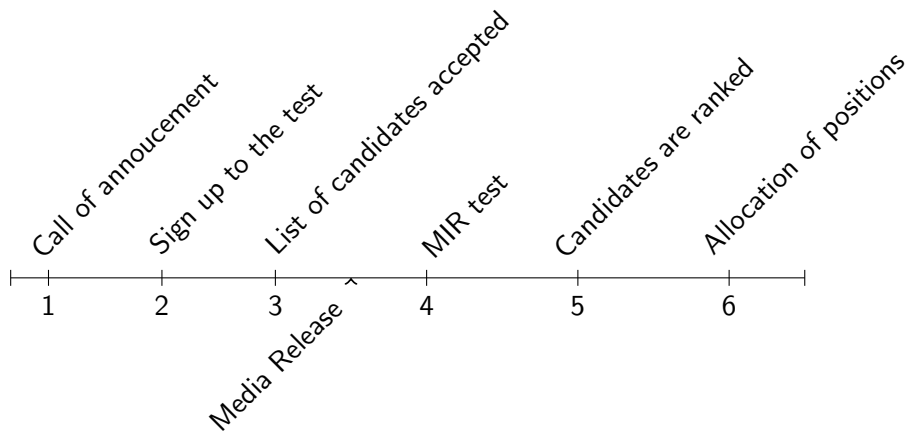
August 31, 2023

- **Labour matching:** skills and qualifications of a worker is compared with the requirements of a particular job vacancy.
- Some **labour markets** are **regulated** by very particular selection processes where employers announce job positions and job seekers are ranked according to some selection processes.
 - **Efficiency and fairness** are two important goals in regulated markets.
- Our setting: **allocation of medical intern positions (MIR) in Spain.**

Medico Interno Residente (MIR)

- **MIR system:** It is a centralized system by which graduates in Medicine are assigned to an intern position.
- When graduated in Medicine (5-6 years degree) in order to work in a hospital (public or private) graduates have to become specialist doctors.
- The only way to become a specialist doctor is accessing a MIR position (3-5 years intern, “learning by doing”)
 - Access to this training programme is regulated by a multiple-choice test (**MIR test**)

Institutional Setting



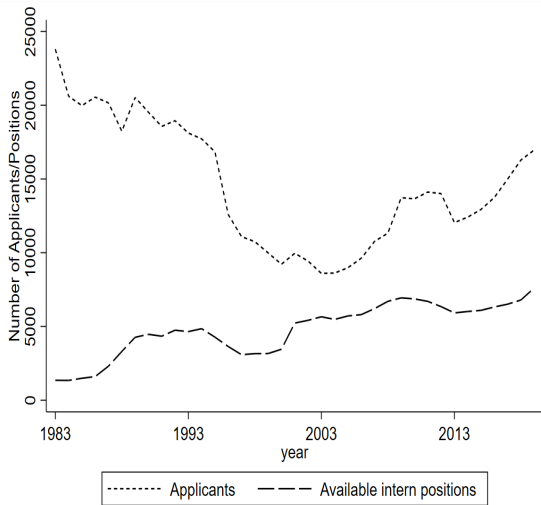
Is Spain different?

- **USA:** medical students are matched into residency programs according to the preferences of both the students and the programs
- **France, Denmark:** demand side is shut down as hospitals cannot choose interns and supply of graduates in medicine is centrally planned according to the health system demand for doctors (Amer-Mestre and Charpin (2021) and Fadlon et al. (2020))
- **Spain:** demand side is shut down as hospitals cannot choose interns and supply of interns has shown a great variation over the past forty years...
 - **Number of candidates has historically been much higher than the number of training positions**

Is Spain different?

Number of candidates and positions over the years:

- Quite a remarkable variation over the past 40 years.
- Source of exogenous variation we exploit



Why is this a gender paper?

- Mismatch between number of candidates and positions creates a **competition problem**: candidates have to compete for available intern positions.
- **Gender differences** at the presence of **competition**:
 - Women underperform compared to men under competitive incentive schemes (Gneezy et al.(2003)), specially when stakes are high (Azmat et al.(2016))
 - Evidence from the field (mostly from educational settings): Jurajda and München (2011), Ors et al. (2013), Cornwell et al. (2013) and Iriberry and Rey-Biel (2019)

The paper's contribution in a nutshell

We study **gender gaps** in the outcomes of **the selection process** in the **medical profession** in Spain over the past 40 years, through the **lens of competition**.

We find that **the more competitive the process, the higher underperformance of women** compared to men.

Our contribution:

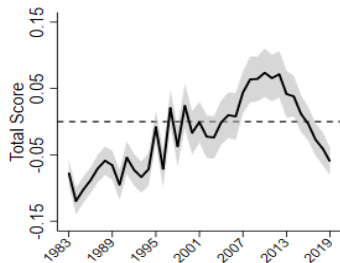
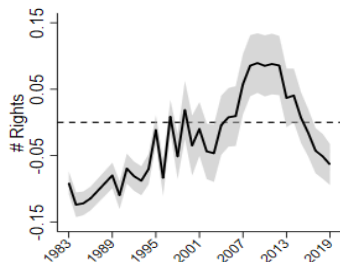
- Unique setting: high stakes, highly skilled pool of individuals, Medical profession has experienced a big feminization process.
- Labor setting in contrast to the existing literature using field data (mostly from educational settings). Nice complement to Flory et al.(2015)
- High quality data on universe of candidates and their performance.

- **Universe of candidates** (500.000) that have taken the MIR test since its creation (1983-2019 period)
 - Focused on Spanish candidates (male and female) for comparability purposes: **around 428.000 candidates**
- **Performance data:** GPA from Medicine degree, number of questions answered/right/wrong, test score, total score, and probability of getting a position.

Gender Performance at the MIR

Gender gaps in performance are observed along the period:

- Inverted U-shape evolution of gender gaps.
- Same time pattern across all outcome variables.



How Competitive is the MIR?

The MIR system is a competitive assignment process...

- Number of applicants has always exceeded the number of intern positions.

... but it shows great variation on its competitiveness level over the past 40 years.

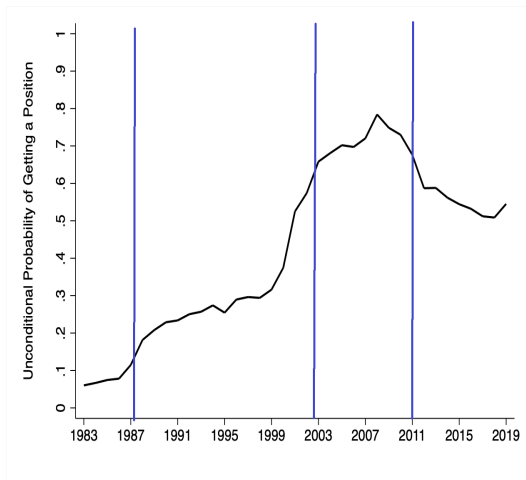
Unconditional probability of ending up with a position

- It is defined as the ratio of the number of intern positions to the number of applicants
- It measures the **inverse of competitiveness**, meaning that the higher this probability the less competitive the process

How Competitive is the MIR?

Four periods of competitiveness:

- 1 Very high: 1983-1988
- 2 High: 1989-2002
- 3 Low: 2003-2011
- 4 Medium: 2012-2019



Very interestingly, the evolution of the competition level of the MIR test and the observed gender gaps, share a **very similar inverted U-shape pattern**.

Considering the variation on competitiveness level along the period, **could the evolution on the observed gender gaps on MIR test's performance be explained through the lens of competition?**

We hypothesize that the competitiveness level might be the driving factor of the observed gender gaps:

- We replace any time-varying component of our model (year fixed effects) with the competitiveness level and try to pin down the relationship.

Results:

- 1 Gender gaps by periods of different competitiveness
- 2 Gender gaps by (inverse) competitiveness
- 3 Robustness tests: external confounders, changes in test's format, differences in ability, reverse causality

Gender Gaps by Periods of Different Competitiveness

Table 1: Gender Gap by Periods that Differ in their Competitiveness Level

	(1)	(2)	(3)	(4)	(5)
	# Answers	# Rights	Test Score	Total Score	Position
Female	0.0203** (0.0083)	0.0436*** (0.0077)	0.0393*** (0.0062)	0.0374*** (0.0058)	0.0744*** (0.0038)
Comp. Medium	0.0086 (0.0093)	0.0264*** (0.0086)	0.0143** (0.0067)	0.0202*** (0.0064)	-0.0825*** (0.0043)
Female x Comp.Med.	-0.0304*** (0.0111)	-0.0513*** (0.0102)	-0.0415*** (0.0080)	-0.0387*** (0.0076)	-0.0415*** (0.0052)
Comp. High	0.1657*** (0.0073)	0.0417*** (0.0070)	-0.1887*** (0.0060)	-0.1764*** (0.0056)	-0.3633*** (0.0036)
Female x Comp. High	-0.1059*** (0.0088)	-0.1149*** (0.0085)	-0.0861*** (0.0073)	-0.0826*** (0.0068)	-0.0880*** (0.0043)
Comp. Very High	0.2017*** (0.0075)	0.0903*** (0.0071)	-0.2021*** (0.0061)	-0.1961*** (0.0056)	-0.5750*** (0.0034)
Female x Comp.Very High	-0.1045*** (0.0092)	-0.1558*** (0.0089)	-0.1358*** (0.0077)	-0.1238*** (0.0072)	-0.0979*** (0.0041)
GPA	0.1299*** (0.0014)	0.3470*** (0.0017)	0.3628*** (0.0016)	0.4838*** (0.0015)	0.1423*** (0.0007)
Constant	0.0703*** (0.0070)	0.1494*** (0.0065)	0.3789*** (0.0052)	0.3564*** (0.0049)	0.6801*** (0.0032)
Observations	392,990	392,990	428,453	428,453	428,453
R-squared	0.0368	0.1594	0.1942	0.3180	0.3007

Gender Gaps by Competitiveness

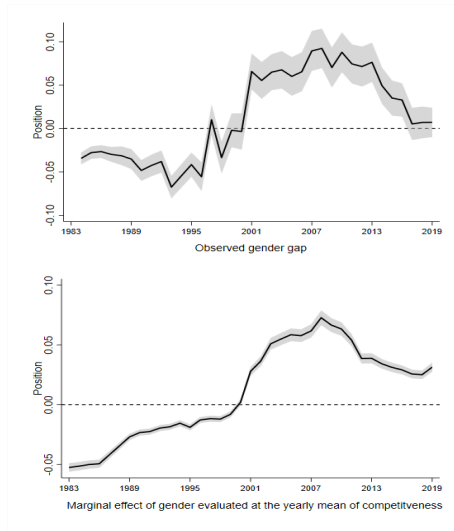
Table 2: Gender Gaps and the Inverse of Competitiveness

	(1)	(2)	(3)	(4)	(5)
	# Answers	# Rights	Test Score	Total Score	Position
Female	-0.1089*** (0.0037)	-0.1375*** (0.0043)	-0.1275*** (0.0043)	-0.1165*** (0.0040)	-0.0590*** (0.0019)
Inverse Comp.	-0.3952*** (0.0095)	-0.1745*** (0.0092)	0.3617*** (0.0081)	0.3537*** (0.0076)	0.9773*** (0.0044)
Female x Inv. Comp.	0.1644*** (0.0120)	0.2527*** (0.0117)	0.2490*** (0.0104)	0.2276*** (0.0097)	0.1654*** (0.0056)
GPA	0.1301*** (0.0014)	0.3474*** (0.0017)	0.3632*** (0.0016)	0.4842*** (0.0015)	0.1422*** (0.0007)
Constant	0.3265*** (0.0026)	0.2512*** (0.0030)	0.1251*** (0.0031)	0.1131*** (0.0029)	0.0316*** (0.0013)
Observations	392,990	392,990	428,453	428,453	428,453
R-squared	0.0385	0.1592	0.1925	0.3163	0.3210

Main finding

Comparison of **observed gender gaps** (upper panel) with the **marginal effect of gender from Table 2** (bottom panel):

- **Astonishing resemblance** between figures.
- Just accounting for **competition seems to be enough to capture most of the time variation in the observed gender gaps.**



- What about other potential changes over the 40 year period... **Are external confounders driving our results?**
No! Controlling for proportion of women, proportion of foreigners, unemployment, gender attitudes towards female labour participation results remain.
- MIR's test format has changed over time ... **Are changes in the format driving our results?**
No! Controlling for changes in the format, the inverse of competitiveness remains significant and of same sign.

- The ability of the pool of students taking the exam each year might have evolved differently by gender . . . **Has the ability of female test takers changed over the past 40 years?**

No! The distribution of GPA scores of men and women over the past 40 years has been constant.

- **Reverse causality can be ruled out.** Candidates are able to infer each year's level of competitiveness before they seat at the test.

The paper in a nutshell

We study **gender gaps in the outcomes of the selection process in the medical profession** in Spain over the past 40 years **through the lens of competition**.

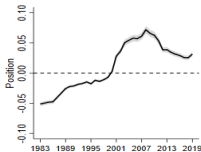
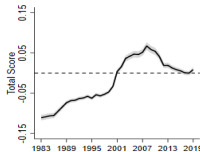
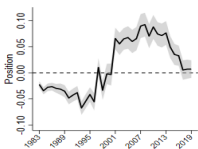
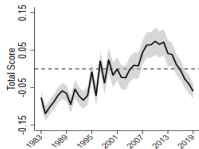
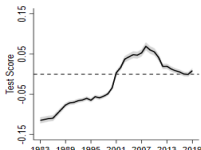
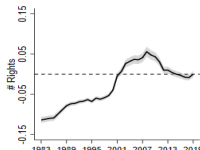
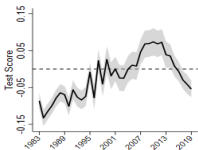
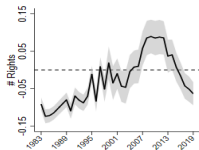
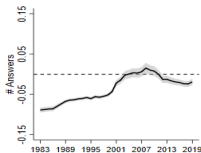
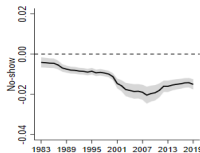
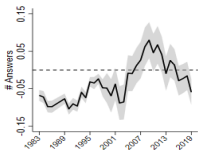
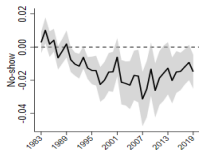
We find that:

- 1 The **observed gender gaps** is **associated with the time pattern** observed in changes **in competition**.
- 2 **Effects are substantial:** probability of gaining a position (ranging from negative 7% up to positive 9% depending on the period)

Policy implications

- Efficiency and fairness are important criteria when evaluating regulated selection processes
- Collateral damage of excessive competition in certain selection process for specific groups, such as women
- These results call for a more centralized planning of the number of positions/Medicine graduates according to the system needs

Back-up slides: Gender Gaps



Back-up slides: Alternative explanations

Table 6: Changes in the MIR Test

Domain	Start	End	Description
No. Questions and Duration	1983	1993	Test with 250 questions and 4.5 hours long
No. Questions and Duration	1994	2008	Test with 250 questions and 5 hours long
No. Questions and Duration	2009	2018	Test with 225 questions and 5 hours long
No. Questions and Duration	2019	2019	Test with 175 questions and 4 hours long
No. Alternative Answers	1983	2014	5 alternative answers per question
No. Alternative Answers	2015	2019	4 alternative answers per question
Weight of Test Score	1983	2009	Test accounts for 75% of total score
Weight of Test Score	2010	2019	Test accounts for 90% of total score
Min. Score	1983	2012	No lower threshold in place
Min. Score	2012	2013	30% of the best 10 test takers' score
Min. Score	2013	2019	35% of the best 10 test takers' score

Back-up slides: Alternative explanations

Table 7: Gender Gaps and Changes in the MIR Test's Format

	(1)	(2)	(3)	(4)	(5)	(6)
	No-show	# Answers	# Rights	Test Score	Total Score	Position
Female	-0.0017 (0.0016)	-0.0914*** (0.0041)	-0.1200*** (0.0045)	-0.1184*** (0.0046)	-0.1076*** (0.0042)	-0.0600*** (0.0021)
Inverse Comp.	0.0926*** (0.0057)	-0.2486*** (0.0160)	-0.1420*** (0.0167)	0.2424*** (0.0159)	0.2297*** (0.0148)	0.9329*** (0.0087)
Female × Inverse Comp.	-0.0129* (0.0072)	0.0705*** (0.0202)	0.1478*** (0.0210)	0.1698*** (0.0198)	0.1501*** (0.0185)	0.1619*** (0.0109)
GPA	-0.0369*** (0.0005)	0.1303*** (0.0014)	0.3475*** (0.0017)	0.3633*** (0.0016)	0.4844*** (0.0015)	0.1422*** (0.0077)
No. Questions (1994)	-0.0063*** (0.0020)	-0.0608*** (0.0049)	-0.0637*** (0.0062)	-0.0173*** (0.0065)	-0.0132** (0.0061)	0.0173*** (0.0034)
No. Questions (1994) × Female	-0.0091*** (0.0026)	0.0148** (0.0064)	0.0311*** (0.0080)	0.0325*** (0.0083)	0.0301*** (0.0078)	-0.0021 (0.0044)
No. Questions (2009)	0.0090 (0.0072)	-0.0060 (0.0224)	0.1518*** (0.0208)	0.1904*** (0.0168)	0.1735*** (0.0157)	-0.1029*** (0.0105)
No. Questions (2009) × Female	0.0071 (0.0087)	0.1036*** (0.0261)	0.0676*** (0.0244)	0.0279 (0.0198)	0.0294 (0.0184)	0.0112 (0.0125)
No. Questions (2019)	0.0037 (0.0049)	-0.0023 (0.0165)	-0.0449*** (0.0151)	-0.0685*** (0.0109)	-0.0882*** (0.0102)	-0.0212*** (0.0080)
No. Questions (2019) × Female	-0.0022 (0.0059)	-0.0470** (0.0203)	-0.0390** (0.0180)	-0.0391*** (0.0131)	-0.0466*** (0.0122)	-0.0156 (0.0099)
No. Alt. Answers (2015)	-0.0346*** (0.0050)	-0.0151 (0.0162)	0.0394*** (0.0150)	0.0371*** (0.0117)	0.0385*** (0.0115)	0.0165** (0.0076)
No. Alt. Answers (2015) × Female	0.0031 (0.0060)	-0.0200 (0.0193)	-0.0594*** (0.0177)	-0.0479*** (0.0138)	-0.0492*** (0.0135)	-0.0342*** (0.0092)
Weight Score (2010)	0.0142* (0.0084)	-0.0178 (0.0263)	-0.0315 (0.0242)	-0.0284 (0.0194)	-0.0107 (0.0184)	0.0824*** (0.0121)
Weight Score (2010) × Female	-0.0095 (0.0101)	-0.0197 (0.0306)	0.0038 (0.0283)	0.0097 (0.0226)	0.0120 (0.0215)	0.0183 (0.0144)
Min. Threshold (2012)	0.0349*** (0.0087)	-0.0287 (0.0264)	-0.0992*** (0.0243)	-0.0326* (0.0193)	-0.0332* (0.0190)	0.0217* (0.0122)
Min. Threshold (2012) × Female	0.0051 (0.0103)	-0.0074 (0.0306)	0.0164 (0.0282)	0.0212 (0.0225)	0.0193 (0.0221)	0.0089 (0.0144)
Min. Threshold (2013)	-0.0364*** (0.0084)	0.0231 (0.0253)	0.0176 (0.0235)	-0.0120 (0.0187)	-0.0118 (0.0184)	0.0340*** (0.0118)
Min. Threshold (2013) × Female	-0.0006 (0.0100)	-0.0308 (0.0295)	-0.0447 (0.0274)	-0.0330 (0.0217)	-0.0294 (0.0214)	-0.0077 (0.0141)
Constant	0.0618*** (0.0011)	0.3171*** (0.0030)	0.2588*** (0.0033)	0.1448*** (0.0033)	0.1327*** (0.0031)	0.0336*** (0.0015)
Observations	467,977	392,990	392,990	428,453	428,453	428,453
R-squared	0.0222	0.0399	0.1618	0.1961	0.3199	0.3232

Back-up slides: Alternative explanations

