

# Fair but Imperfect – Discovering a Disparate Impact in a Procedurally Fair Hiring Process

**Dominique Meurs**

Université Paris Nanterre; INED, Paris

**Patrick A. Puhani**

Leibniz Universität Hannover; IZA, Bonn

**Séminaire Insee – 24 mars 2017**

## Background and Motivation

Griggs vs. Duke Power Company U.S. Supreme Court Decision of March 8, 1971

1955: the company added the requirement of a high school diploma for its higher paid jobs.

“The court ruled that the company’s employment requirements did not pertain to applicants’ ability to perform the job, and so was discriminating against African-American employees, even though the company had not intended it to do so...The facts of this case demonstrate the inadequacy of broad and general testing devices, as well as the infirmity of using diplomas or degrees as fixed measures of capability”

## Research Questions of this Project

- **Procedural fairness**: are there systematic differences between (non-anonymous) oral exam results and anonymous written results by gender?
- **Disparate impact/functional discrimination**: does trainee performance differ by gender conditional on entry contest scores within the pool of hired workers?
- Data come from a real hiring process in the French public sector, the “concours des IRA”
- **Contributions**: (1) adding to audit literature, (2) real world hiring process, (3) post-hiring outcomes

## Preview of Results – Procedural Discrimination

- Contrary to most previous literature, we cannot detect any statistically significant evidence for procedural discrimination against women (cf. Neumark, Bank and Van Nort, 1996 on restaurants; Goldin and Rouse, 2000, on orchestra musicians; Bertrand and Mullainathan, 2004, on service jobs)
- Women receive even slightly better marks than men in the oral exam

## Preview of Results – Disparate Impacts

- However, hired women do better than men, even conditional on hiring test scores
- This suggests that the tests for hiring may contain a gender bias in that women perform better in the training program than men conditional on the test scores that determine hiring into the training program/job
- In that sense, the entry exams are biased, i.e. have a disparate impact, although they would pass a test similar to a traditional audit study
- The results exhibit time trends

# Outline

1. Le Concours - Recruitment in the French Public Sector
2. Data and Descriptive Statistics
3. Oral Exam Results Conditional on Written Results and Other Characteristics
4. Job Performance Conditional on Being Hired
5. Conclusions

# 1. Entry Exams for Regional Administration Institutes

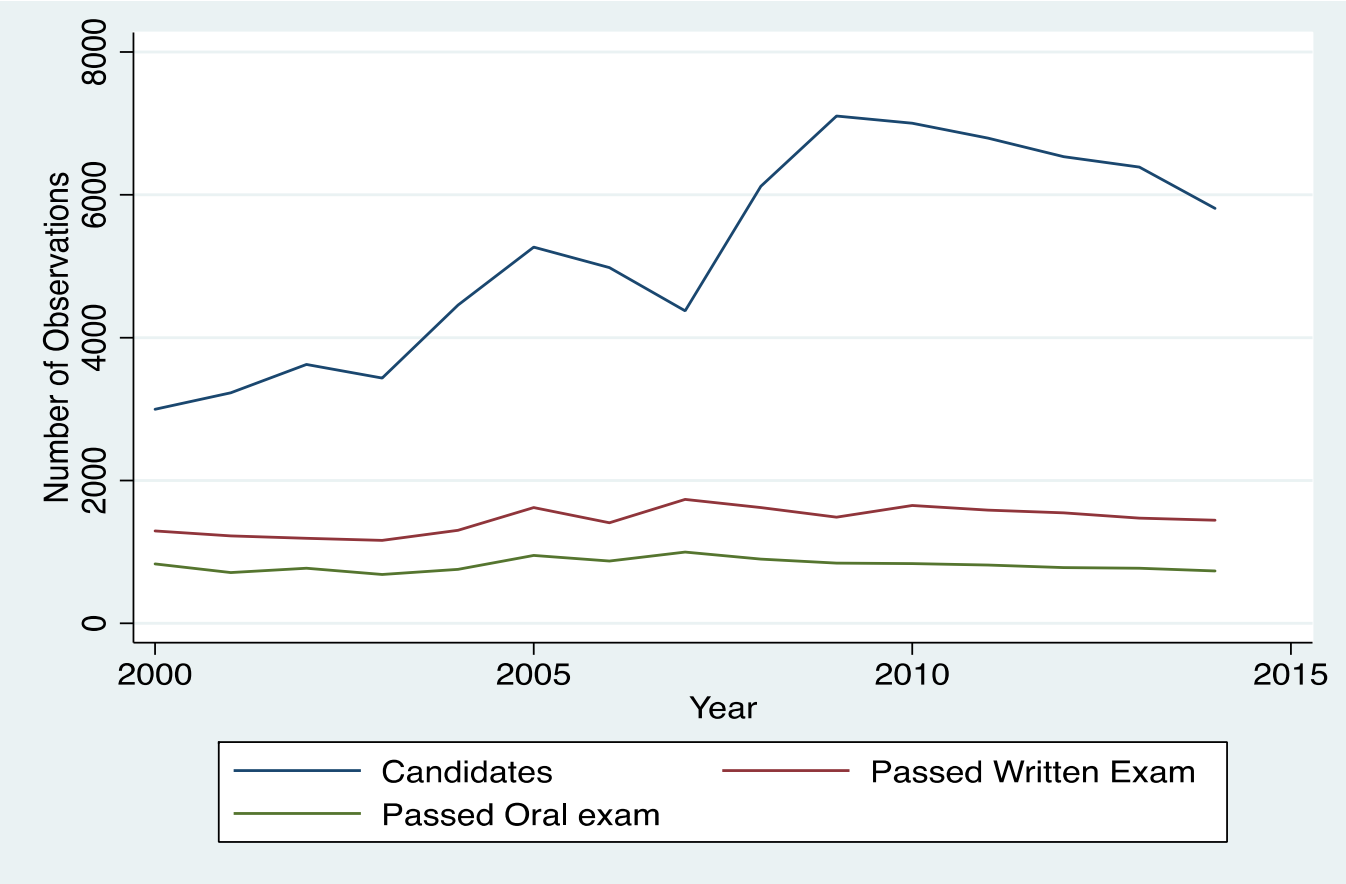
- Ecole nationale d'administration (ENA) recruits about 80 persons per year into the top management level (attended by many French presidents and other important figures in France)
- Instituts régionaux d'administration (IRA) are one level below that in the hierarchy, and recruit about 600-700 people per year
- When hired, start a one-year trainee program at one of 5 regional centers (Bastia, Lille, Lyon, Metz, Nantes) and receive a stipend of about €1500
- The only way to get this job is through the entrance exam

# 1. Entry Exams for Regional Administration Institutes

- Exams: there are about 20 examiners per regional training center, who first correct the written exams anonymously. Out of these 20 people, about two-thirds or more are chosen to build juries of three people for the oral examinations
- Each three-person jury must consist of at least one man and one woman.
- Each jury is also supposed to be balanced by professional status, work experience and type of career.
- All jury members are civil servants at a hierarchical position equal or higher than the positions the job candidates will take
- Jury members get a one-day interview training, are sensitized on discrimination and have to document the candidates' evaluation



# 1. Number of Candidates and Candidates Having Passed the Written and Oral Steps



# 1. Exam Structure Over Time

	2007	2008	2009	2010	2011	2012
<i>Internal</i>						
Dossier [factor 4]	x	x	x	x	x	x
Ten Short Questions [factor 5, since 2008: 4]	x					
Oral [factor 4]	xx	x	x	x	x	x
Language [factor 1]	o	o	o	o	o	o
<i>External</i>						
Essay on Common Culture [factor 4]	x	x	x	x	x	x
Ten Short Questions [factor 5, since 2008: 4]	x	x	x	x	x	x
Summary Essay on Adm. Doc. [factor 3]	x					
Oral [factor 4]	xx	x	x	x	x	x
Language [factor 1]	x	x	x	x	x	x

# Outline

1. Le Concours - Recruitment in the French Public Sector

2. Data and Descriptive Statistics

3. Oral Exam Results Condition on Written Results and  
Other Characteristics

4. Job Performance Conditional on Being Hired

5. Conclusions

## 2. Data

- **Administrative data** on the *concours*/contest of 2007/2008 until 2012/2013), contains only test results (written [February] and oral [June]) and gender; NOT immigrant status or other background characteristics (education, parental background)
- 8,838 participants in sample of candidates who passed the written exam, of which 5,193 are women and 3,645 are men
- ignore candidates from the “third-exam track” with at least 5 years’ work experience, smaller sample
- For the post-hiring job-related trainee scores, we have 3,609 candidates, of which 2,174 were women and 1,435 were men

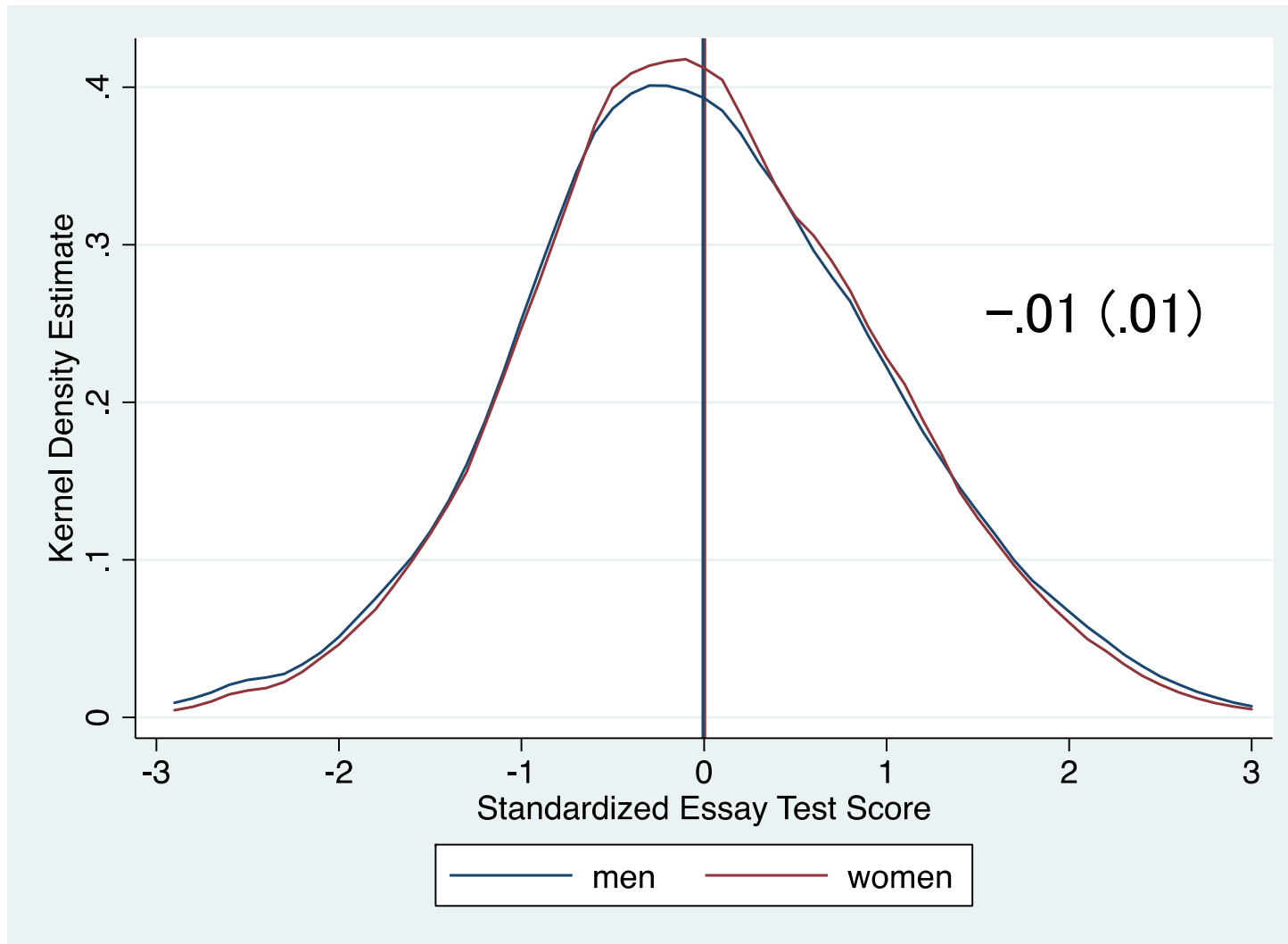
## 2. Data – Administrative Publication in Candidate Types

	2007	2008	2009	2010	2011	2012
Failed the written exam	2,615	4,496	5,614	5,343	5,209	4,990
Left after passing the written	183	121	120	141	109	79
Failed the oral exam	785	605	516	676	661	685
Left after Passing the oral	67	328	242	228	239	191
Passed and accepted the Job	688	578	609	603	584	599
Total number of candidates	4,338	6,128	7,101	6,991	6,802	6,544

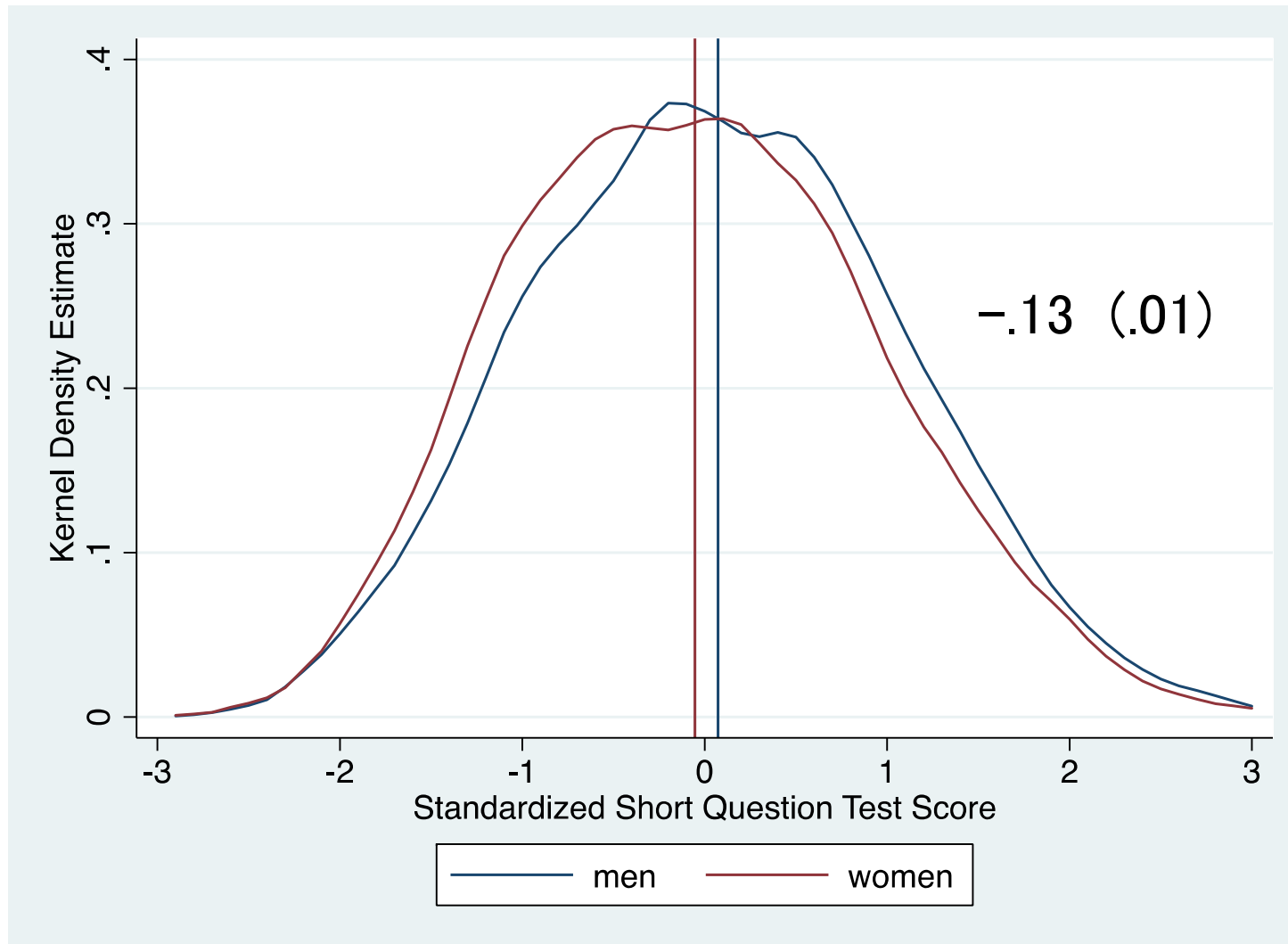
## 2. Sample Means

	Women	Observations	Men	Observations
Passed written exam	0.25	22,953	0.27	14,906
Passed oral exam	0.57	5,193	0.54	3,645
Average Essay/Dossier score	1.11	5,193	1.07	3,645
Average Short Question score	0.92	3,174	1.00	2,665
Average Summary Essay score	0.71	554	0.55	407
Average Oral score	0.01	5,193	-0.02	3,645
Average Language score	0.04	3,879	-0.06	3,037
Internal	0.45	5,193	0.34	3,645
<i>Training Center</i>				
Bastia	0.20	5,193	0.20	3,645
Lille	0.20	5,193	0.20	3,645
Lyon	0.21	5,193	0.19	3,645
Metz	0.18	5,193	0.20	3,645
Nantes	0.20	5,193	0.21	3,645
<i>Year</i>				
2007	0.16	5,193	0.19	3,645
2008	0.18	5,193	0.16	3,645
2009	0.15	5,193	0.16	3,645
2010	0.17	5,193	0.17	3,645
2011	0.18	5,193	0.15	3,645
2012	0.16	5,193	0.17	3,645

## 2. Male and Female Essay Scores

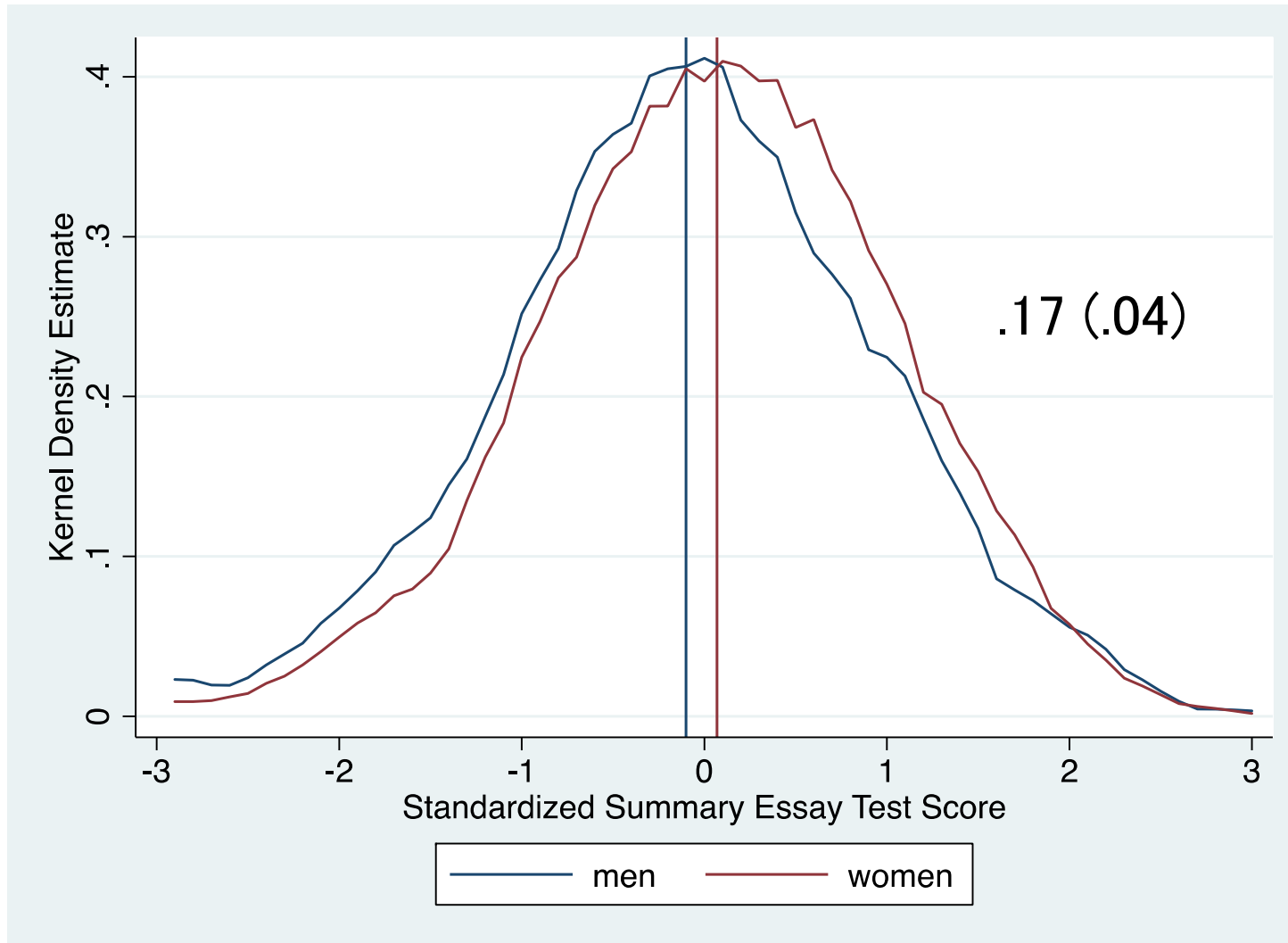


## 2. Male and Female Short Question Scores

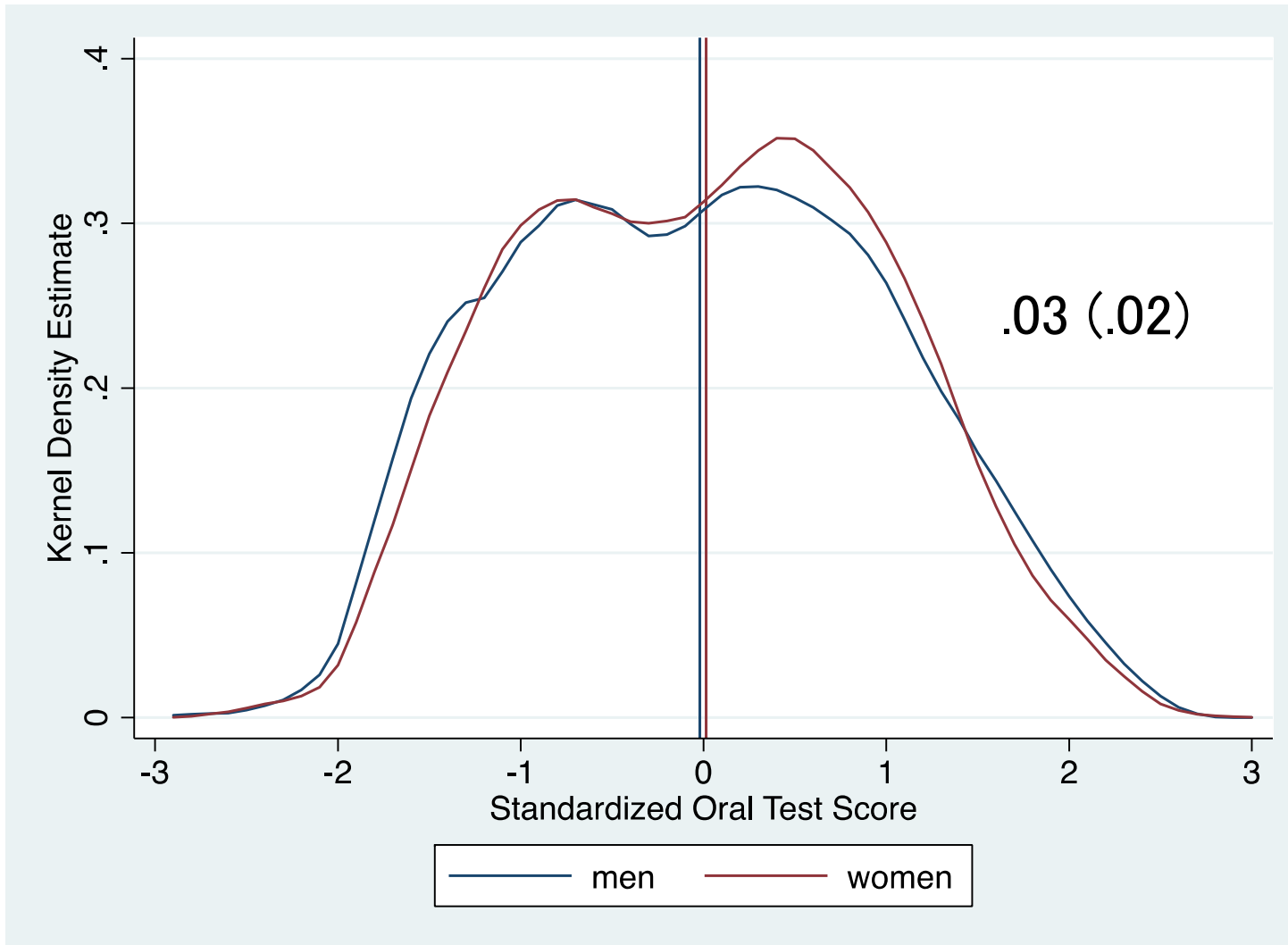




## 2. Male and Female Summary Essay Scores



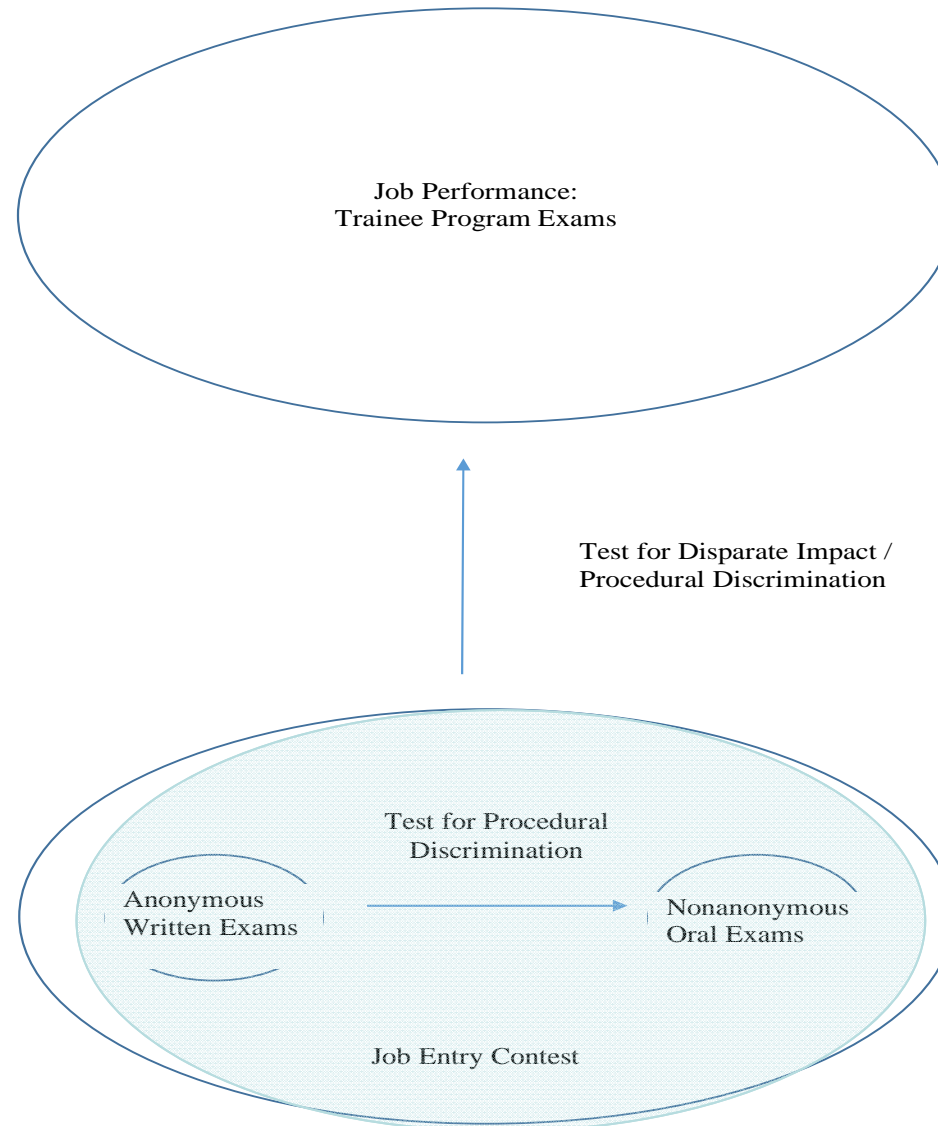
## 2. Male and Female Oral Scores



# Outline

1. Le Concours - Recruitment in the French Public Sector
2. Data and Descriptive Statistics
3. Oral Exam Results Condition on Written Results and Other Characteristics
4. Job Performance Conditional on Being Hired
5. Conclusions

### 3. Structure of Empirical Analysis



### 3. Simple Approach to Measure Procedural Discrimination

$$E[O_1 | Female] < E[O_1 | Male]$$

- But maybe

$$E[O_1 | Female, WT_1] = E[O_1 | Male, WT_1]$$

$$resultoral_i = \tau Min_i + \rho_1 essay_i + \rho_2 shortquestions_i + \beta' x_i + \varepsilon_i$$

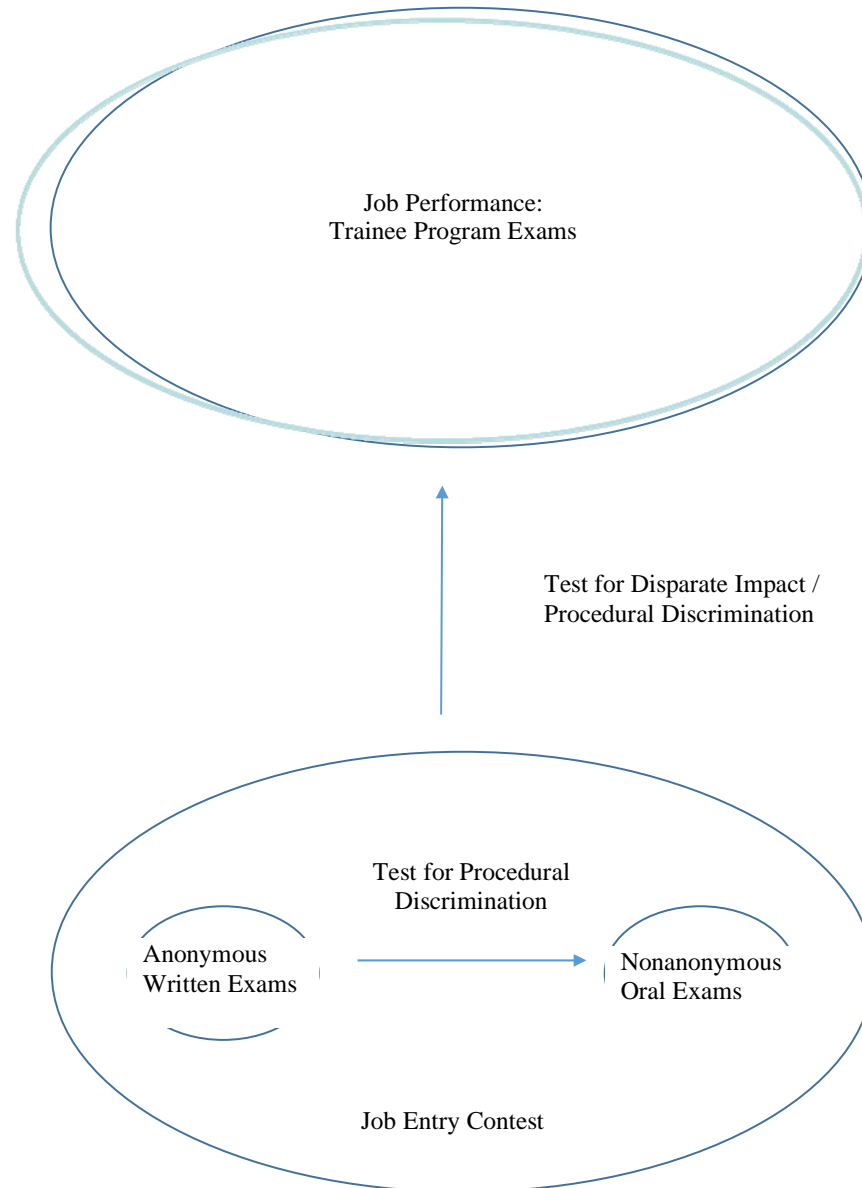
### 3. Regressions Oral Results

	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)
Female	0.03	(0.02)	0.05***	(0.02)	0.07***	(0.02)	0.07***	(0.02)
Score Essay/Dossier							0.17***	(0.03)
Score Short Questions							0.21***	(0.04)
Score Summary Essay							0.15***	(0.04)
Internal candidate			-0.14***	(0.02)	0.19***	(0.05)	0.30***	(0.06)
<i>Training Center (Lyon)</i>								
Bastia			0.21***	(0.04)	0.61***	(0.08)	0.63***	(0.08)
Lille			0.07***	(0.03)	0.31***	(0.08)	0.20***	(0.08)
Metz			0.19***	(0.04)	0.15	(0.09)	0.16*	(0.08)
Nantes			0.09***	(0.03)	0.42***	(0.08)	0.32***	(0.08)
<i>Year (2007)</i>								
2008			0.00	(0.04)	0.34***	(0.10)	0.33***	(0.13)
2009			0.01	(0.04)	0.34***	(0.09)	0.31***	(0.13)
2010			0.01	(0.04)	0.24***	(0.09)	0.28***	(0.12)
2011			0.00	(0.04)	0.47***	(0.09)	0.49***	(0.13)
2012			0.01	(0.04)	0.46***	(0.09)	0.18	(0.14)
Year and internal interactions	no		no		yes		yes	
Year and training center interactions	no		no		yes		yes	
Year and exam score interactions	no		no		no		yes	
Constant	-0.02	(0.02)	-0.09***	(0.04)	-0.40***	(0.07)	-0.76***	(0.08)
R2	0.00		0.01		0.04		0.06	
Observations	8,838		8,838		8,838		8,838	

# Outline

1. Le Concours - Recruitment in the French Public Sector
2. Data and Descriptive Statistics
3. Oral Exam Results Condition on Written Results and Other Characteristics
4. Job Performance Conditional on Being Hired
5. Conclusions

# 4. Structure of Empirical Analysis





## 4. Post-Recruitment Tests

- After being hired, candidates receive a one-year training and get tested after each semester
- The total score weights the first-semester results with one third and the second-semester results with two-thirds
- At the end of the one-year training period, workers are being ranked and allocated to full-blown jobs

## 4. Post-Recruitment Tests

- Only written exam of post-recruitment tests anonymous
- Others potentially subject to procedural discrimination bias
- Under the assumption that they are fair, we would expect women and men to have similar conditional scores (if variance of skill similar)
- Under the assumption that they are fair, we can also derive results on possible biases in the setup of the *concours*/contest

#### 4. Post-Recruitment Tests: 1st Semester

- Written 5h exam based on a set of documents on an administrative problem combining legal, organizational, financial and communication issues [5/20]
- Internship mark based on supervisor's judgement and candidate's internship report [7/20]
- Oral 20min exam based on internship; testing administrative management skills [5/20]
- Defense of an administrative group report: 40min oral exam, grade awarded at the individual level [3/20]

## 4. Post-Recruitment Tests: 2nd Semester

- Written 5h exam consisting of solving one or several practical administrative cases [5/20]
- Internship mark based on supervisor's judgement and candidate's internship report, similar to 1st semester [5/20]
- Oral 30min exam on the field of specialization [5/20]
- Oral foreign language exam: 20min discussion on a professional topic [1/20]

## 4. Testing for a disparate impact

Correlations between trainee performance scores and the results of the job entry contest

<b>Test Score</b>	First Semester	Second Semester
<i>Women and Men</i>		
Overall Total	-	0.28***
Total	0.25***	0.24***
Written	0.12***	0.10***
Internship	0.12***	0.10***
Oral	0.29***	0.21***
Defense/Language	0.015	0.32***
<i>Women</i>		
Overall Total	-	0.34***
Total	0.30***	0.30***
Written	0.18***	0.13***
Internship	0.16***	0.13***
Oral	0.32***	0.24***
Defense/Language	0.01	0.35***
<i>Men</i>		
Overall Total	-	0.21***
Total	0.19***	0.30***
Written	0.06**	0.08***
Internship	0.09***	0.09***
Oral	0.24***	0.17***
Defense/Language	0.03	0.27***

## 4. Estimating Equation for Disparate Impact

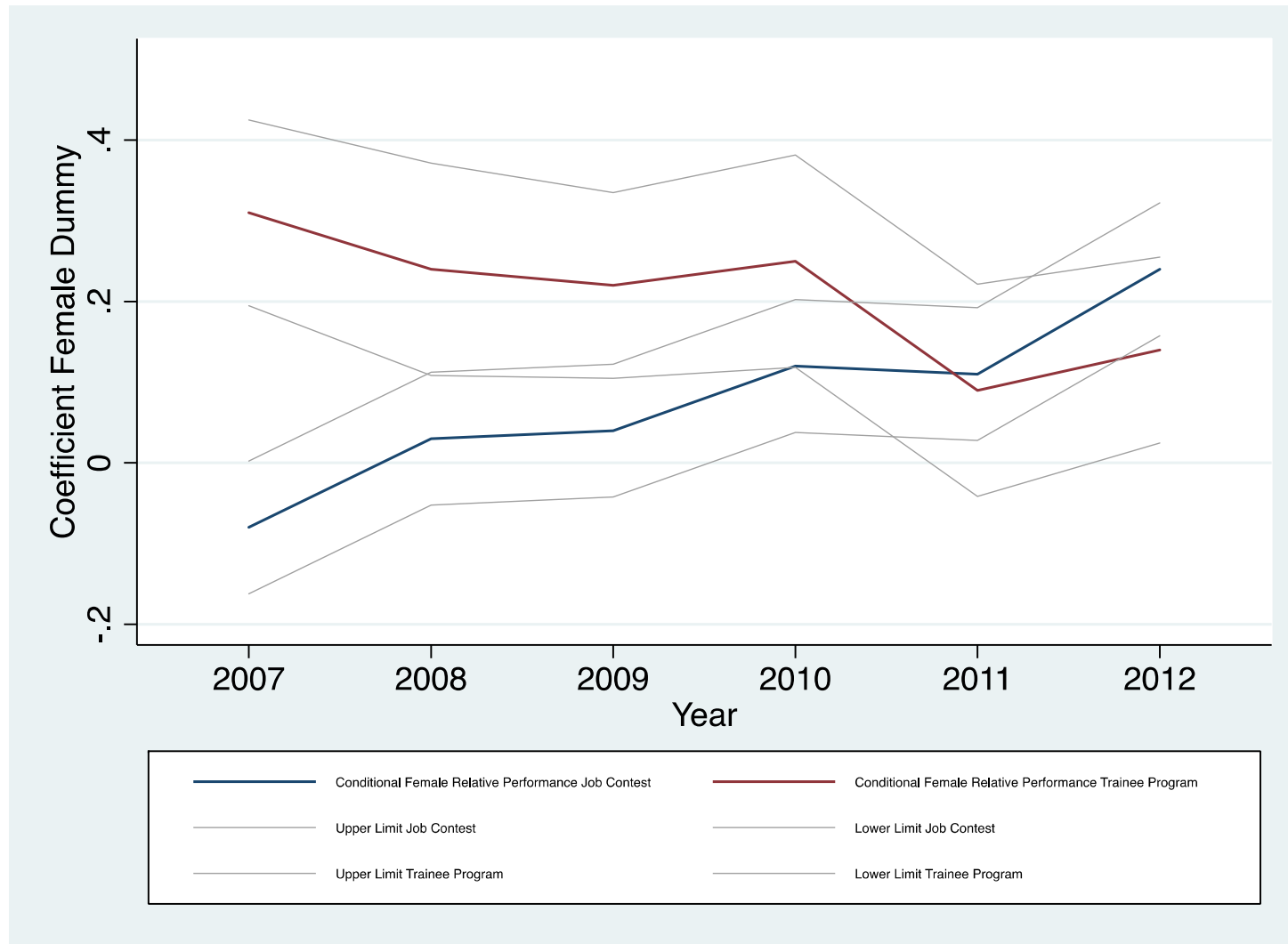
*traineescore<sub>i</sub>*

$$= \tau \textit{female}_i + \rho \textit{entrycontestscore}_i + \beta x_i + \varepsilon_i$$

## 4. Post-Recruitment Performance Women Versus Men; n= 3,609

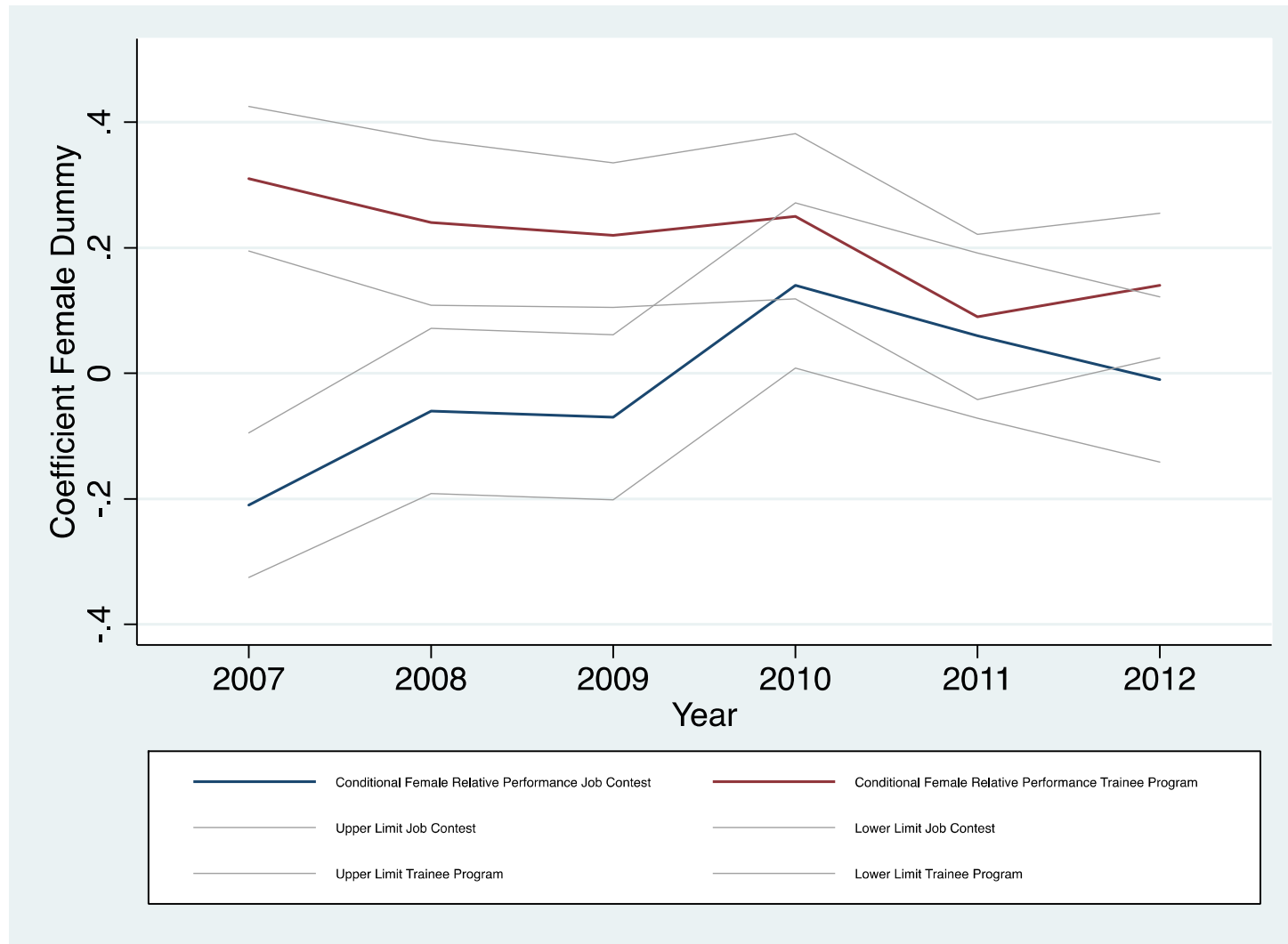
Dependent Variable	No controls		Controlling for year, training center and internal/external		Controlling for job entry contest scores (also interacted by year)	
	Coeff.	(s.e.)	Coeff.	(s.e.)	Coeff.	(s.e.)
Total Trainee Score	0.16***	(0.03)	0.19***	(0.03)	0.21***	(0.03)
Total First Semester	0.11***	(0.03)	0.15***	(0.03)	0.16***	(0.03)
Total Second Semester	0.20***	(0.03)	0.20***	(0.03)	0.21***	(0.03)
Written First Semester	0.15***	(0.03)	0.16***	(0.03)	0.17***	(0.03)
Written Second Semester	0.20***	(0.03)	0.20***	(0.03)	0.21***	(0.03)
Internship First Semester	0.17***	(0.03)	0.18***	(0.03)	0.19***	(0.03)
Internship Second Semester	0.15***	(0.03)	0.16**	(0.03)	0.17***	(0.03)
Oral First Semester	-0.04	(0.03)	0.01	(0.03)	0.02	(0.03)
Oral Second Semester	0.03	(0.03)	0.06*	(0.03)	0.07**	(0.03)
Defense First Semester	-0.01	(0.03)	-0.01	(0.03)	0.00	(0.03)
Language Second Semester	-0.01	(0.03)	0.01	(0.03)	0.06*	(0.03)

## 4. Estimates by Contest Year – Candidate Sample





## 4. Estimates by Contest Year – Trainee Sample



# Outline

1. Le Concours - Recruitment in the French Public Sector
2. Data and Descriptive Statistics
3. Oral Exam Results Condition on Written Results and Other Characteristics
4. Job Performance Conditional on Being Hired
5. Conclusions

## 5. Conclusions

- Conditional on anonymous written tests and on other human capital characteristics, the oral exam relevant for the hiring decision seems to display no bias against women (no procedural discrimination)
- However, the characteristics valued in the oral exam seem different from the ones driving productivity “on the job“ (during the trainee program) and women are doing better in the post-training tests than men conditional on previous test scores; in that sense, there is a gender bias (disparate impact, functional discrimination) in these test
- Changing the design of the entry contest (written and oral) reduced the estimated magnitude of the disparate impact