Gendered Teacher Feedback, Students' Math Performance and Enrollment Outcomes

A Text Mining Approach

Pauline Charousset
Paris School of Economics

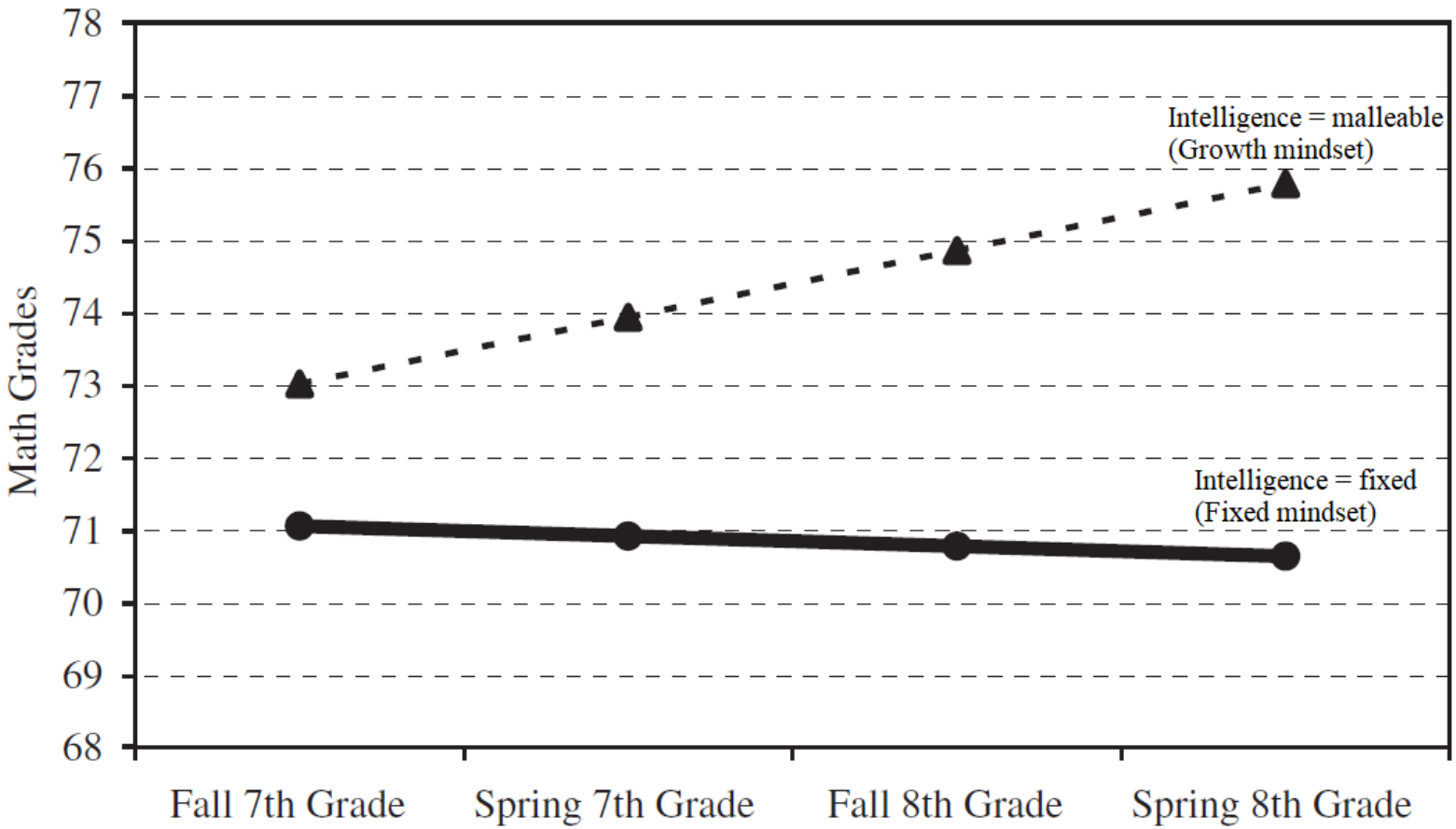
Marion Monnet
Ined

AEA-ASSA, January 2022

Motivation

1. Students viewing intelligence as malleable (growth mindset) perform better at school, persevere more and set higher goals

   Blackwell et.al, 2007; Bettinger et.al, 2018; Huillery et.al, 2021



Source: Blackwell et.al, 2007

Motivation

1. Students viewing intelligence as malleable (growth mindset) perform better at school, persevere more and set higher goals

   Blackwell et.al, 2007; Bettinger et.al, 2018; Huillery et.al, 2021

Motivation

1. Students viewing intelligence as malleable (growth mindset) perform better at school, persevere more and set higher goals

   Blackwell et.al, 2007; Bettinger et.al, 2018; Huillery et.al, 2021

2. Teachers play a key role in shaping students' mindset through their everyday classroom interactions namely through the feedback provided

   Alan & Ertac, 2018; Alan, Boneva et.al, 2019

Motivation

1. Students viewing intelligence as malleable (growth mindset) perform better at school, persevere more and set higher goals

   Blackwell et.al, 2007; Bettinger et.al, 2018; Huillery et.al, 2021

2. Teachers play a key role in shaping students' mindset through their everyday classroom interactions namely through the feedback provided

   Alan & Ertac, 2018; Alan, Boneva et.al, 2019

3. Male and female students do not react to the different types of feedback in the same manner

$\Rightarrow$ females perform better and are more motivated when given a growth-mindset type of feedback

Coprus & Lepper, 2007; Good et.al, 2012

Research questions and contributions

- **Research questions:**
  1. Do math teachers provide different feedback to equally able G-12 Science track female and male students?

Research questions and contributions

- **Research questions:**
  1. Do math teachers provide different feedback to equally able G-12 Science track female and male students?
  2. If so, what type of feedback do they use for one sex over the other?

Research questions and contributions

- **Research questions:**
  1. Do math teachers provide different feedback to equally able G-12 Science track female and male students?
  2. If so, what type of feedback do they use for one sex over the other?
  3. What is the impact of a differentiated feedback on school performance, higher education choices and enrollment?

Research questions and contributions

- **Research questions:**
  1. Do math teachers provide different feedback to equally able G-12 Science track female and male students?
  2. If so, what type of feedback do they use for one sex over the other?
  3. What is the impact of a differentiated feedback on school performance, higher education choices and enrollment?
- **Contributions:**
  1. New index to measure the degree of teachers' gender differentiation using textual data
  2. Using comprehensive and non-experimental data
  3. Analyse the effect of gendered feedback on students' choices (rank-order lists)
  4. Policy relevant (awareness-raising campaigns)

Outline

1. Motivation
2. Research questions and contributions
3. Data and descriptive stats
4. Teacher gendered vocabulary index (GDV)
5. Effects on performance and enrollment outcomes

Data sources

- **Higher education application data 2012-2017:**
  - Grade-12 students' school reports (grades + written feedback) $\Rightarrow$ teacher GDV index
  - Rank-order lists of higher ed. programs $\Rightarrow$ outcome variables
  - Students' and teachers' characteristics

Data sources

- **Higher education application data 2012-2017:**
  - Grade-12 students' school reports (grades + written feedback) $\Rightarrow$ teacher GDV index
  - Rank-order lists of higher ed. programs $\Rightarrow$ outcome variables
  - Students' and teachers' characteristics
- **National exams data 2012-2017:**
  - End-of middle school exam (DNB): control for initial math ability
  - End-of high school exam (*Baccalauréat*): outcome variable

Data sources

- **Higher education application data 2012-2017:**
  - Grade-12 students' school reports (grades + written feedback) $\Rightarrow$ teacher GDV index
  - Rank-order lists of higher ed. programs $\Rightarrow$ outcome variables
  - Students' and teachers' characteristics
- **National exams data 2012-2017:**
  - End-of middle school exam (DNB): control for initial math ability
  - End-of high school exam (*Baccalauréat*): outcome variable
- **Higher education enrollment data 2012-2017:**
  - Enrollment after high school graduation $\Rightarrow$ outcome variable

Summary Statistics - G12 students Science Track

Students' socioeconomic characteristics

	All	Boys	Girls
Demographics			
Female student	0.47		
Age (years)	18.09	18.12	18.06
Scholarshio student	0.13	0.12	0.14
High SES	0.43	0.44	0.41
Medium-high SES	0.16	0.16	0.16
Medium-low SES	0.24	0.24	0.25
Low SES	0.17	0.16	0.18
Nb. of observations	691,093	368,922	322,171

Summary Statistics - G12 students Science Track

School performance and electives

	All	Boys	Girls
School performance (end-of middle school)			
Rank at DNB: math	50.28	52.18	48.13
Rank at DNB: French	50.33	44.69	56.73
Elective course in G-12 science track			
Maths	0.23	0.27	0.19
Physics-Chemistry	0.26	0.27	0.25
Earth & Life Science	0.37	0.26	0.50
Engineering & Info	0.13	0.20	0.06
Nb. of observations	691,093	368,922	322,171

Summary Statistics

Grade-12 math teachers' characteristics

Characteristics	
Male math teacher	0.58
Share of head teacher at least once	0.53
Number of teacher observations	3.70
Average number of classes per year	1.09
Average number of students per class	28.02
Nb. of teachers	6,772

Outline

1. Motivation
2. Research questions and contributions
3. Data and descriptive stats
4. Teacher gendered vocabulary index (GDV)
5. Effects on performance and enrollment outcomes

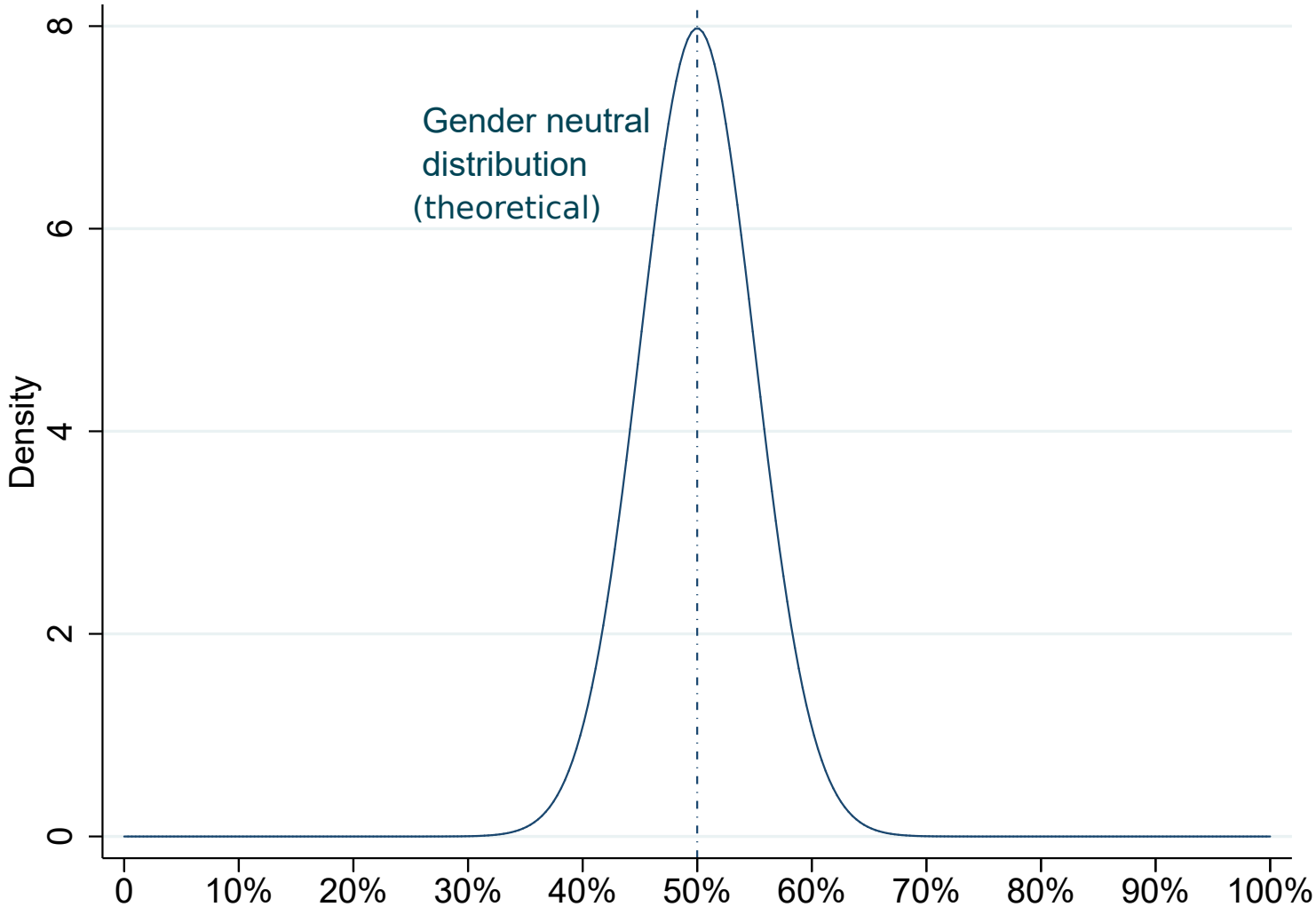
Measuring teacher GDV - Intuition

- **Purpose of the GDV index:**
  - Measuring the degree of gender differentiation in the vocabulary used by math teachers in their written feedback:
    i) depending on student's gender
    ii) with equal math aptitudes

Measuring teacher GDV - Intuition

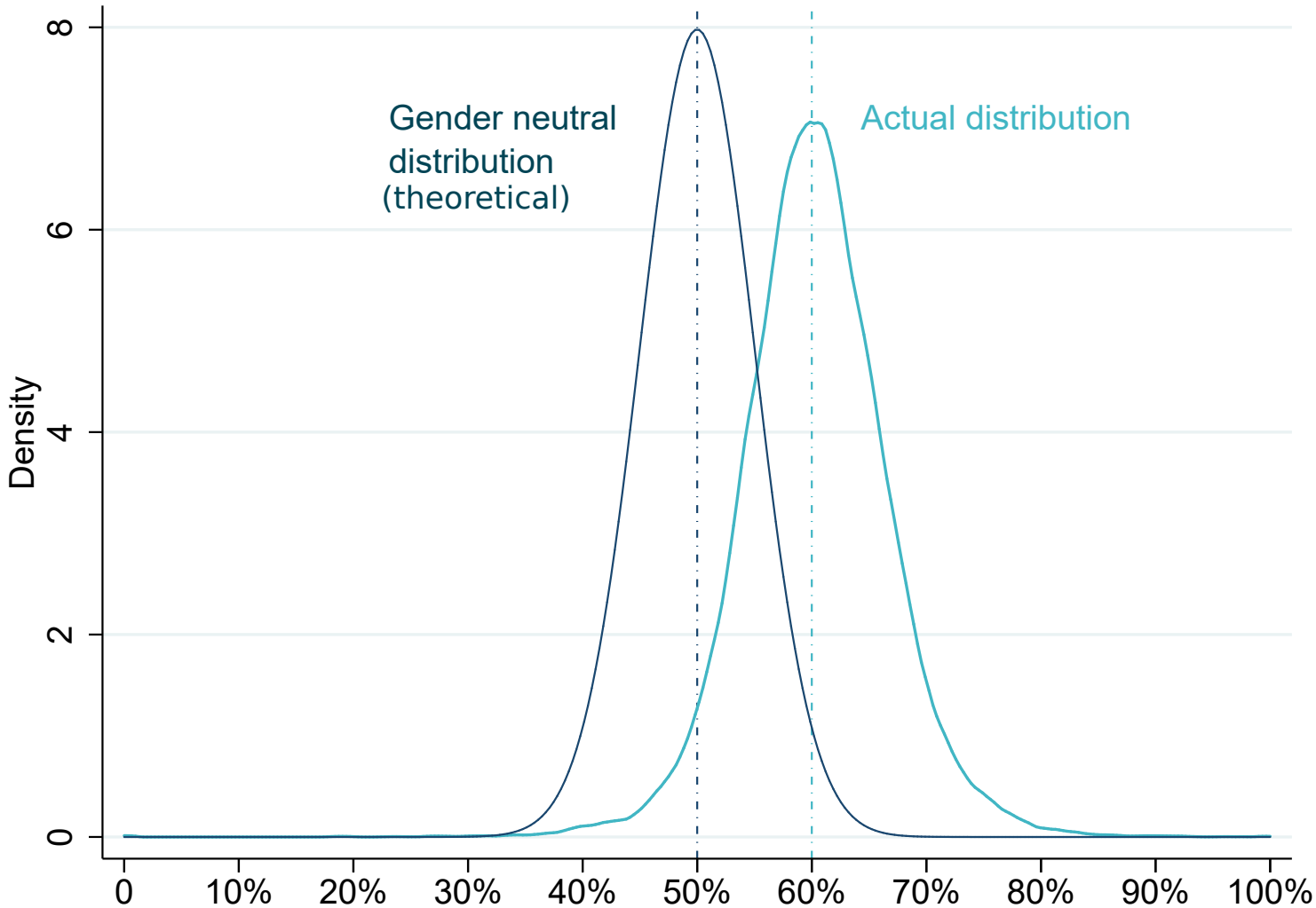
- **Purpose of the GDV index:**
  - Measuring the degree of gender differentiation in the vocabulary used by math teachers in their written feedback:
    i) depending on student's gender
    ii) with equal math aptitudes
- **Measuring teacher GDV:** Details Appendix
  - Step 1: Predict students' gender based on the words used in the written feedback
  - Step 2: Compute the quality of the prediction for each teacher
  - Metric used: accuracy $\Rightarrow$ % of students correctly classified (actual sex = predicted sex)
  - The higher the accuracy $\Rightarrow$ the more the teacher uses a vocabulary specific to males and females

Distribution of teachers' GDV



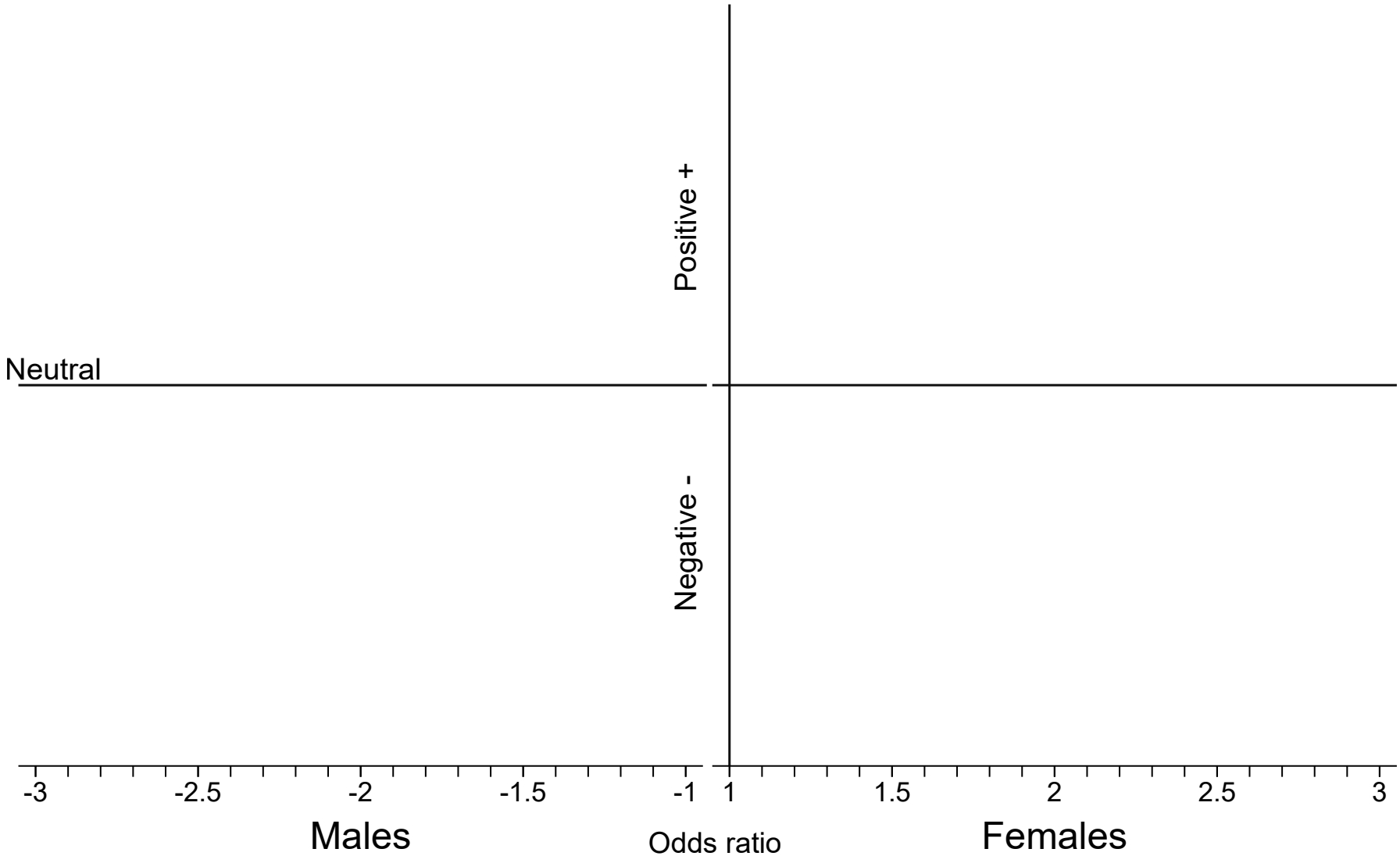
If the vocabulary used is gender-neutral + assuming the sample is made of 50% of male and 50% females, the average accuracy should be 50% (= random guessing).

Distribution of teachers' GDV

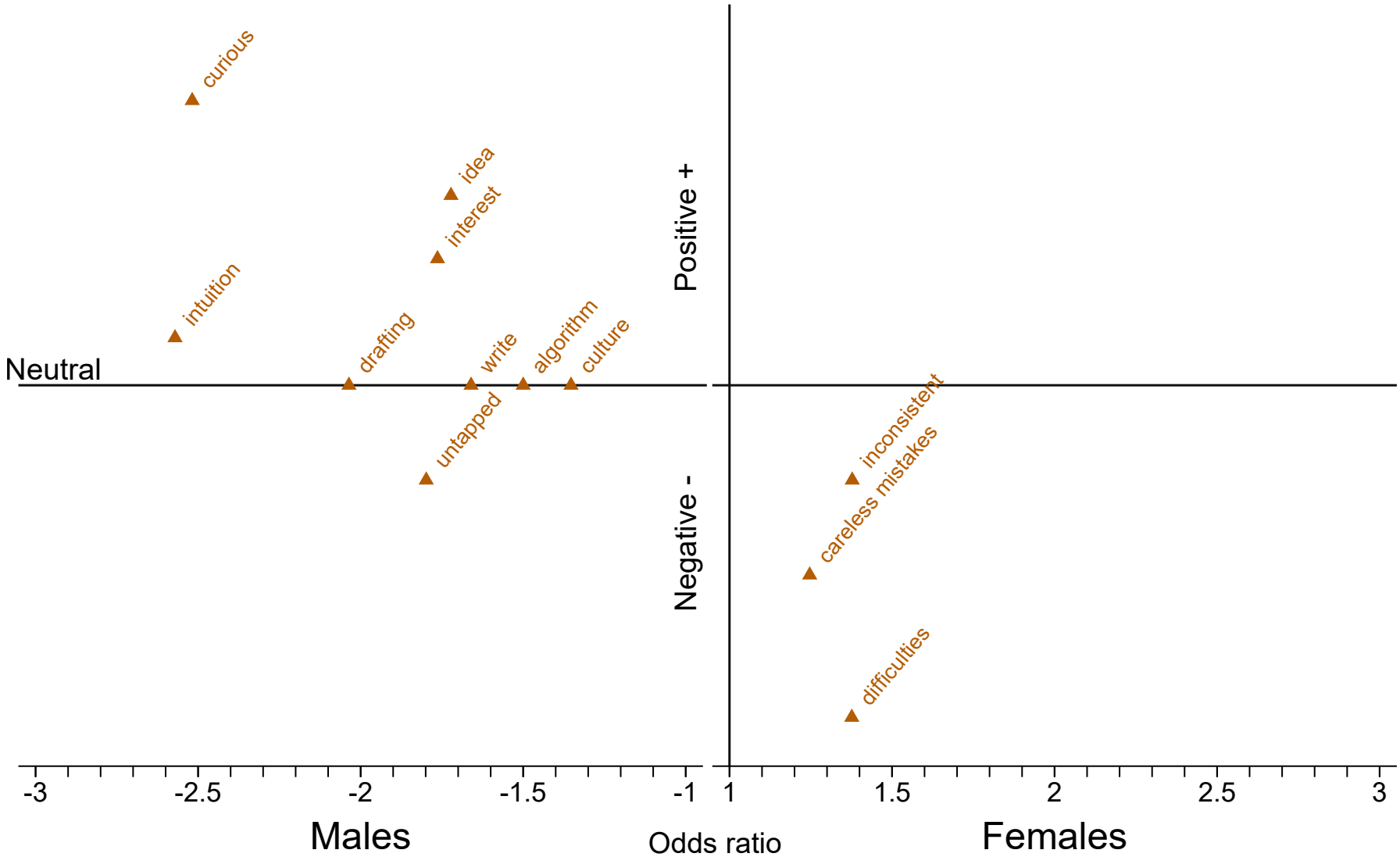


The proportion of correctly classified students is 60% on average → evidence of gendered vocabulary.

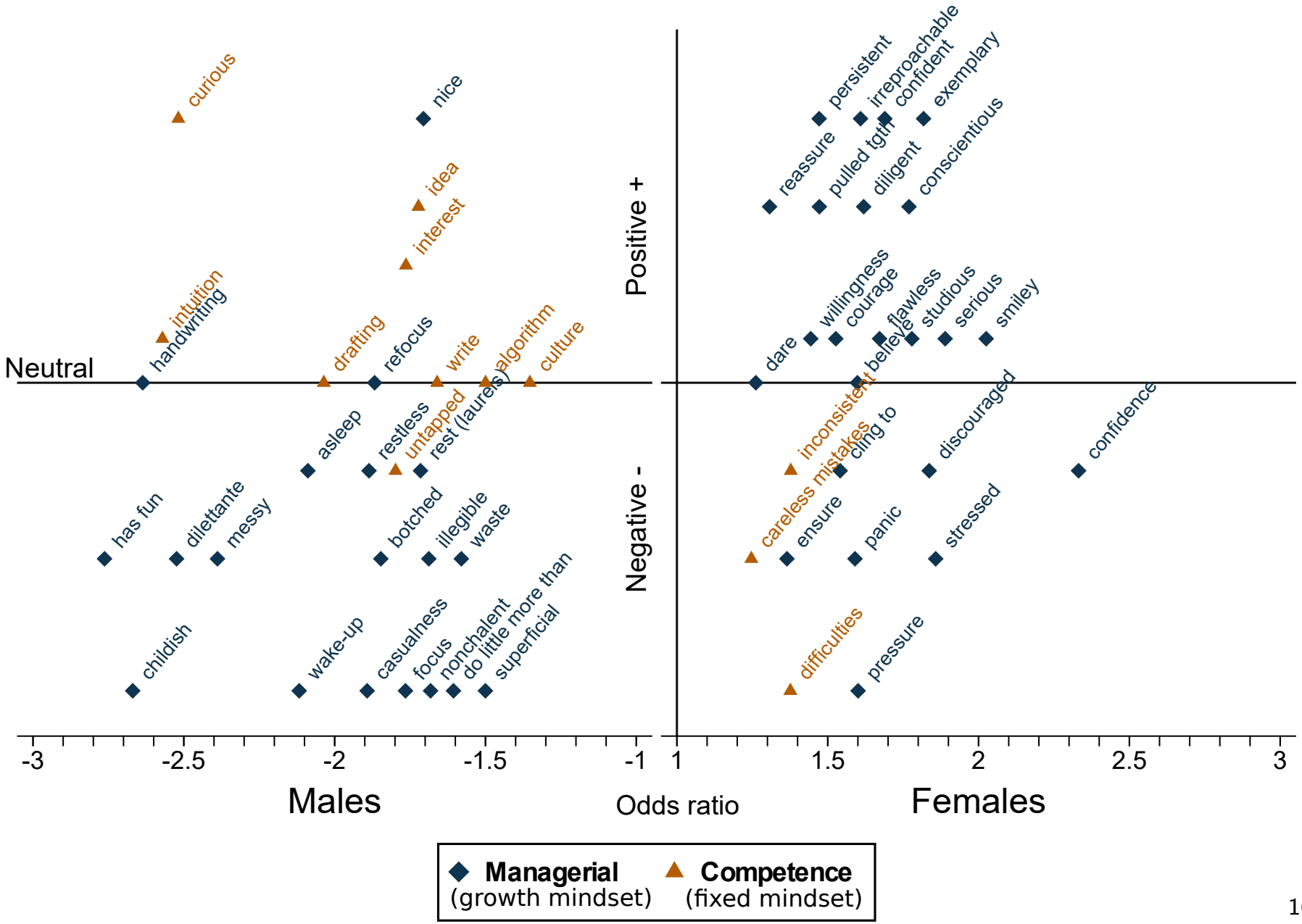
Classification of the top 30 gender predictors



Classification of the top 30 gender predictors



Classification of the top 30 gender predictors



Outline

1. Motivation
2. Research questions and contributions
3. Data and descriptive stats
4. Teacher gendered vocabulary index (GDV)
5. Effects on performance and enrollment outcomes

Identification strategy

- **Identifying the effect of teacher GDV:**
  - Compare students from the same high school s, same elective course e but exposed to teachers with varying GDV
  - The within high school comparison allows to control for some unobservable characteristics
  - More specifically, we estimate the following equation:

$$Y_{isjet} = \alpha + \beta_1 GDV_{jc} + \gamma_s + \eta_e + \delta_t + \epsilon_{isjet} \qquad (1)$$

Identification strategy

- **Identifying the effect of teacher GDV:**
  - Compare students from the same high school s, same elective course e but exposed to teachers with varying GDV
  - The within high school comparison allows to control for some unobservable characteristics
  - More specifically, we estimate the following equation:

$$Y_{isjet} = \alpha + \beta_1 GDV_{jc} + \gamma_s + \eta_e + \delta_t + \epsilon_{isjet} \quad (1)$$

- **Identifying assumptions:**
  - Relies on the quasi random allocation of teachers to classes conditional on elective courses
  - Statistical tests provide support for the validity of the design:
    1. "Balancing test" between GDV and students' X Table
    2. χ^2 test of random allocation Table

Effects of teacher GDV - Rank at Baccalauréat

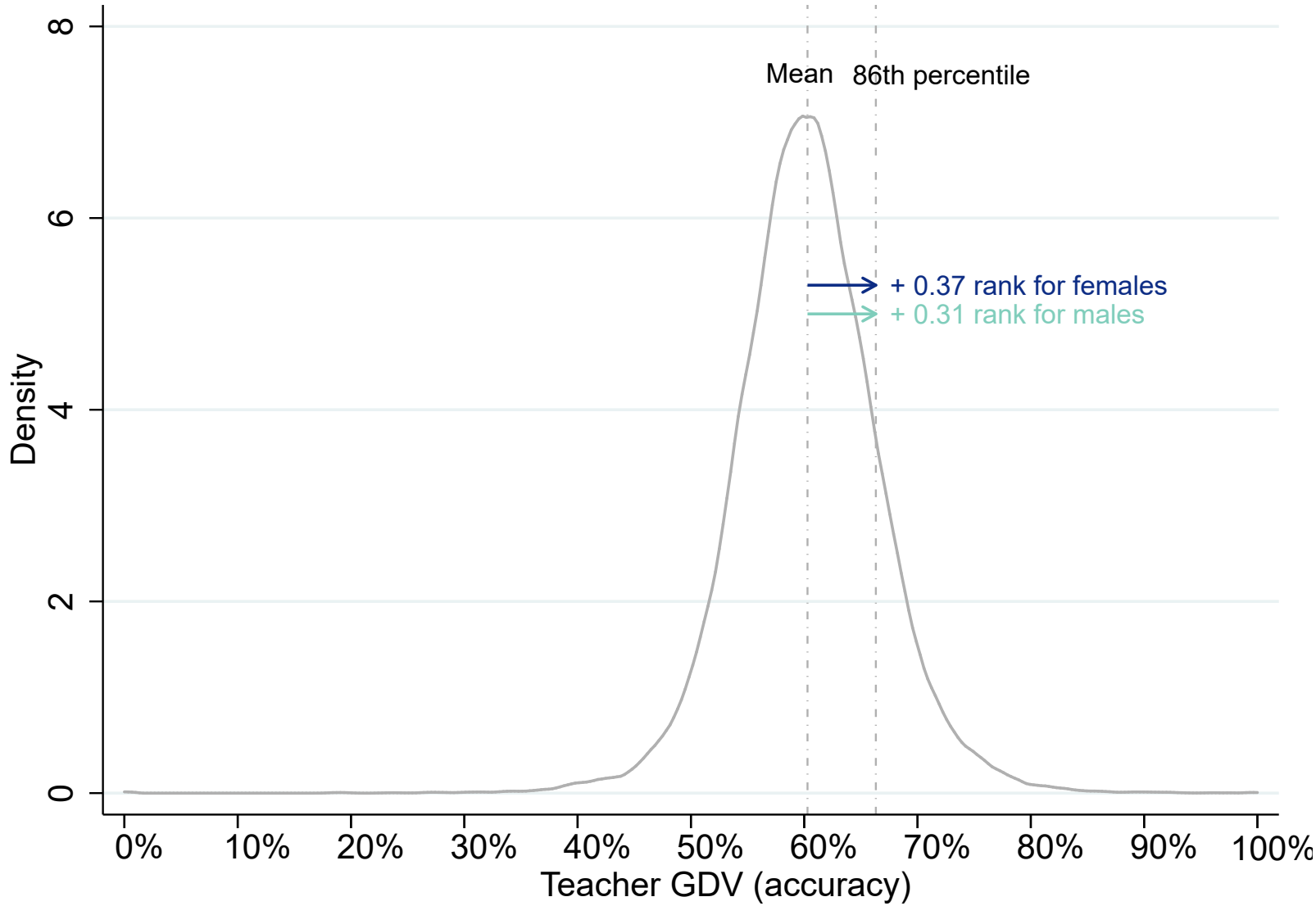
- **Average effects on rank at Baccalauréat:**

	All	Males	Females
Rank at *baccalauréat*: math	0.3351***	0.3113***	0.3735***
	(0.0668)	(0.0747)	(0.0811)
Rank at *baccalauréat*: philo	-0.0235	-0.0381	-0.0093
	(0.0646)	(0.0746)	(0.0786)
Nb. observations	649,105	345,201	303,904

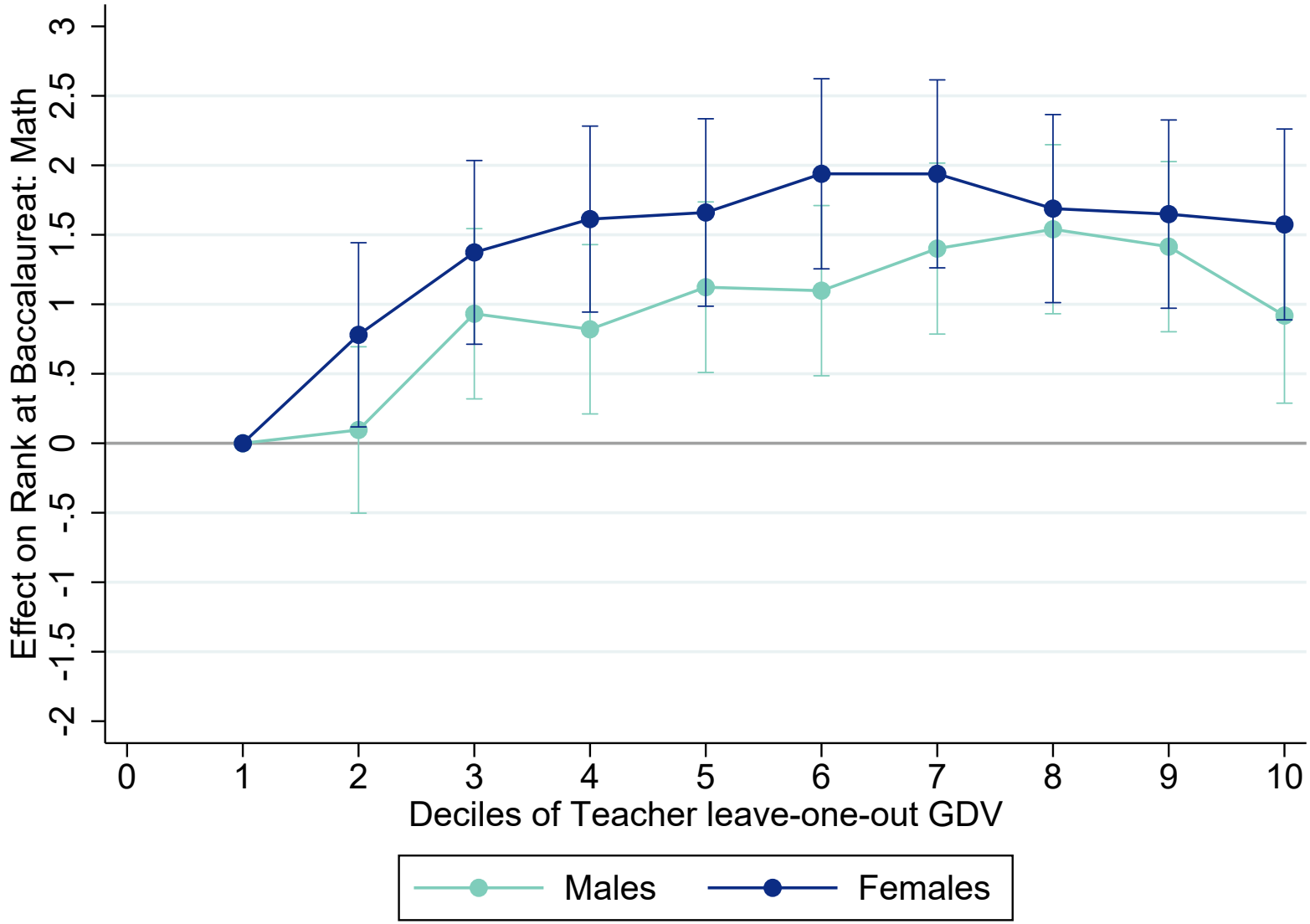
N.B: ranks are percentile ranks where 0 is for the lowest performing student and 100 the highest performing one

  - Being exposed to a teacher with a one-standard deviation higher GDV increases the rank at Baccalauréat by 0.37 for females and by 0.31 fore males

Graphical representation GDV effects



Effects on Baccalauréat - GDV deciles



Note: Compared to females exposed to the bottom 10% of the GDV distribution, females students exposed to teachers from the 4^{th} decile or above see their rank increase by 1.5 to 2 ranks

Effects of teacher GDV - Choices and enrollment

- **Effects on rank-order lists and enrollment:**
  - No effect on the probability to top rank a scientific program in the rank-order list Graph
  - No effect on enrollment Graph
  - No heterogenous effects
- **Mecanisms for the effects on baccalauréat:**
  - More in the paper !!

Conclusion

- On average, G12 math teachers provide different feedback to their male and female students
- They are more likely to mention the positive attitude and efforts for female students . . .
- . . . as well as the unruly behaviour and the intellectual skills of their male students (to a lesser extent)
- Students exposed to high-GDV teachers perform better at the Baccalauréat exam
- But this does not translate into different higher education outcomes

Thanks for your attention !

marion.monnet@ined.fr
pauline.charousset@gmail.com

APPENDIX

Measuring teacher GDV - Estimation steps

1. **Transforming written feedback into a database:**
   - Cleaning feedback:
     - neutralize words' gender (e.g. "*sérieuse*" and "*sérieux*" $\rightarrow$ "*sérieux*");
     - remove first names and very frequent words;
     - replace words by their root (e.g. "*sérieux*" and "*sérieusement*" $\rightarrow$ "*sérieux*")

Measuring teacher GDV - Estimation steps

1. **Transforming written feedback into a database:**
   - Cleaning feedback:
     - neutralize words' gender (e.g. "*sérieuse*" and "*sérieux*" $\rightarrow$ "*sérieux*");
     - remove first names and very frequent words;
     - replace words by their root (e.g. "*sérieux*" and "*sérieusement*" $\rightarrow$ "*sérieux*")

   $\Rightarrow$ Gives the vocabulary (**W**) used in math feedbacks (**W** $\simeq$ 1600 words)

   - The vocabulary is transformed into a matrix: 1 column = 1 word and 1 row = 1 student Matrix
   - Column = 1 if the word is contained in the student's feedback

Measuring teacher GDV - Estimation steps

2. Identifiying the best gender predictors:

- We assume a logistic form for $P(Female_i = 1|W_i)$:

$$P(Female_i = 1|W_i) = \frac{exp(\alpha W_i)}{1 + exp(\alpha W_i)} \quad \forall i \tag{2}$$

- α coefficients are obtained by minimizing the (penalized) log-likelihood (Logistic-LASSO)
- The α coefficients are interpreted as odds-ratios: *"The word Y is X times more likely to appear in a female's feedback than in a male's feedback."*
- The model is fitted on a training sample, balanced by gender and math performance
- The 10 best males' and females' feedback predictors : Graph

Measuring teacher GDV - Estimation steps

3. **Predicting students' gender for each teacher:**
   - Use the model fitted in step 2 to predict students' gender on the (balanced) test sample:

$$\widehat{Sex_i} = \widehat{P}(Female_i = 1|W_i) = \frac{exp(\hat{\alpha} W_i)}{1 + exp(\hat{\alpha} W_i)} \qquad (3)$$

Measuring teacher GDV - Estimation steps

3. **Predicting students' gender for each teacher:**
   - Use the model fitted in step 2 to predict students' gender on the (balanced) test sample:

$$\widehat{Sex_i} = \widehat{P}(Female_i = 1|W_i) = \frac{exp(\hat{\alpha} W_i)}{1 + exp(\hat{\alpha} W_i)} \tag{3}$$

4. **Compute teachers' gender-differenciated vocabulary:**
   - For each teacher j and class c, we compute two GDV indices:
     1. GDV = % of correctly predicted cases (accuracy) :

$$GDV_{jc} = \frac{1}{N_{jc}} \sum_{i=1}^{N_{jc}} \mathbb{1}\{Sex_i = \widehat{Sex_i}\} \times 100 \quad \forall j, c \tag{4}$$

Measuring teacher GDV - Estimation steps

4. Compute teachers' gender-differenciated vocabulary (cont'd):

- For each teacher j and class c, we compute two GDV indices:
  2. Leave-one-out GDV : Class c gets the average GDV measured in other teacher j's classes:

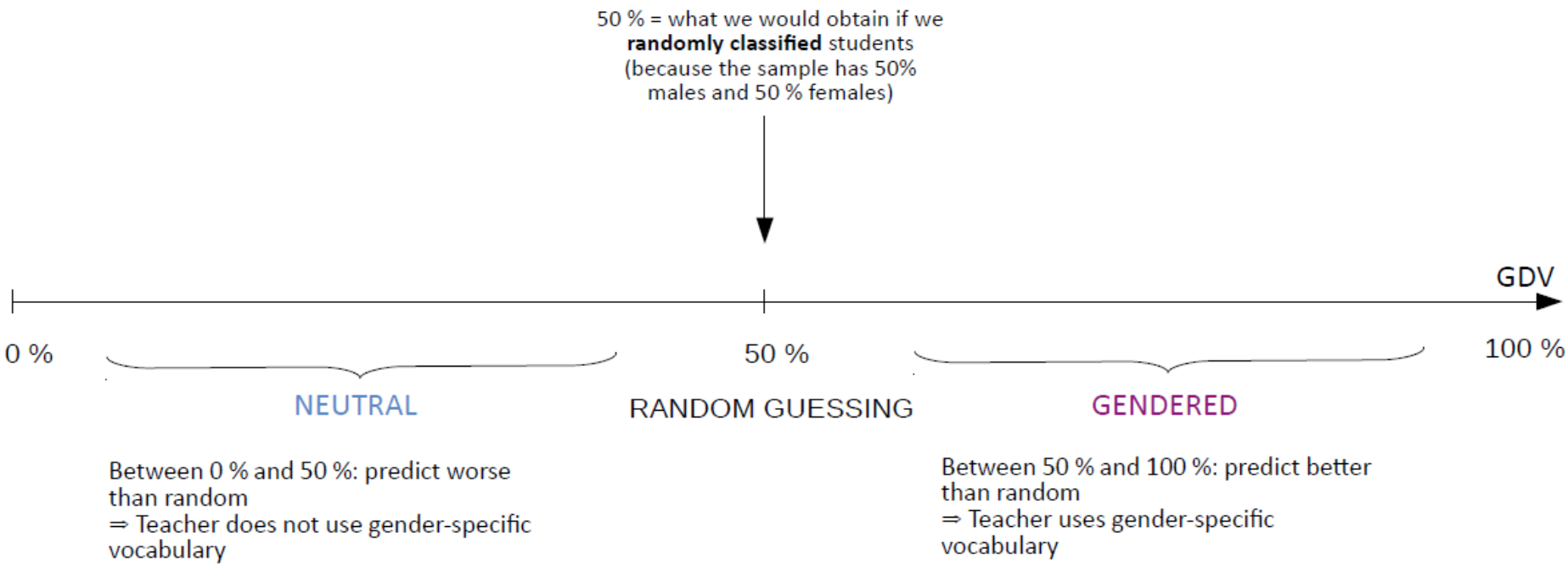
$$GDV_{j\setminus c} = \frac{1}{N_j - 1} \sum_{c' \neq c} GDV_{jc'} \quad \forall j, c \tag{5}$$

$\Rightarrow$ ensures that GDV $\perp$ unobserved class-characteristics

$\Rightarrow$ removes noise

Interpretation of teacher GDV indices

- Both GDV indices range between 0 and 100
- The higher the GDV, the more the teacher uses the gender-specific vocabulary identified in step 2.



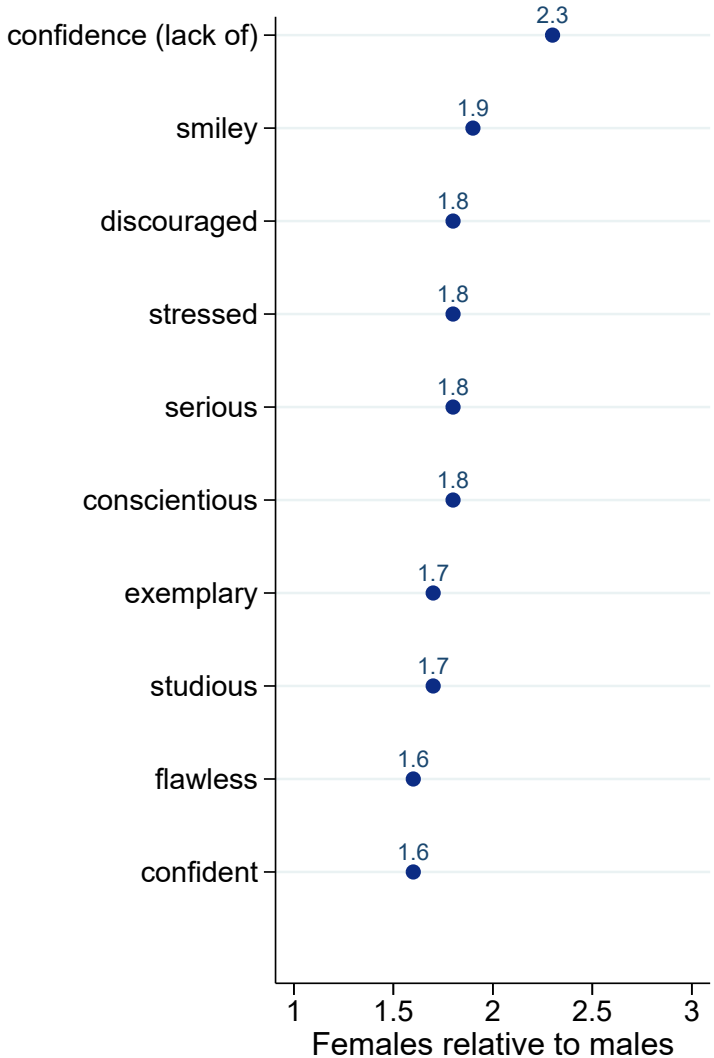
The vocabulary matrix

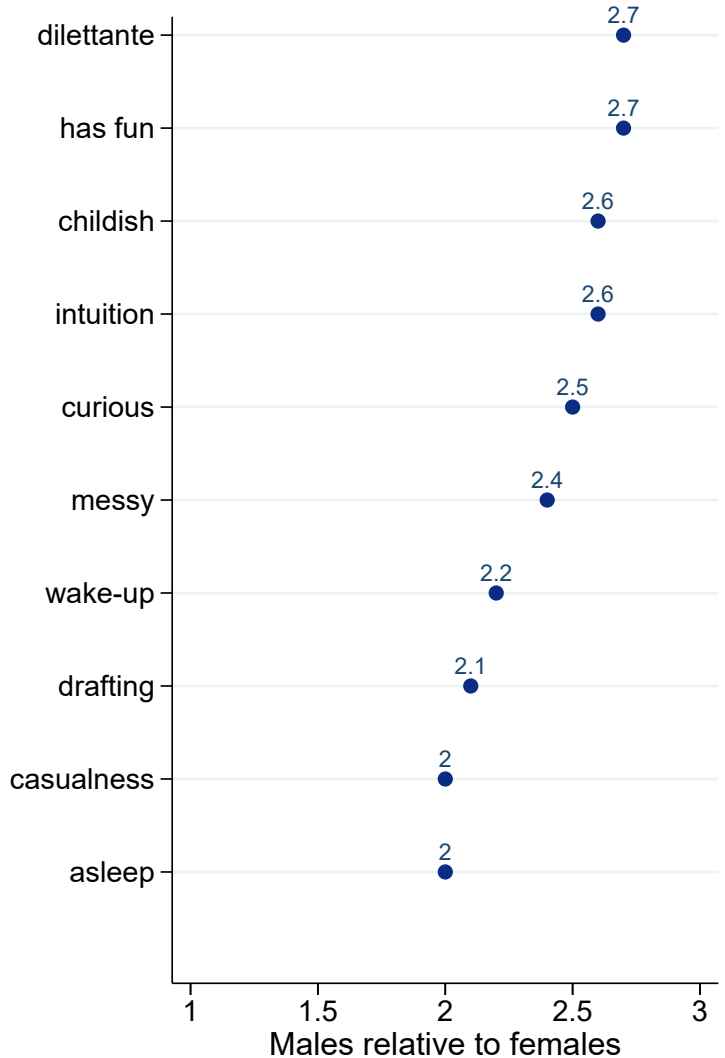
		Vocabulary							
Student	**Sex**	results	alarming	good	work	effort	messy	exam	...
Student 1	**1**	1	1	0	0	0	1	0	...
Student 2	**0**	0	0	1	1	1	0	1	...
Student 3	**0**	1	0	1	1	1	0	1	...

Reading: Student 1 is a female (sex=1) and has the words "results", "alarming" and "messy" in her feedback. Back

The 10 best males' and females' predictors

Retour

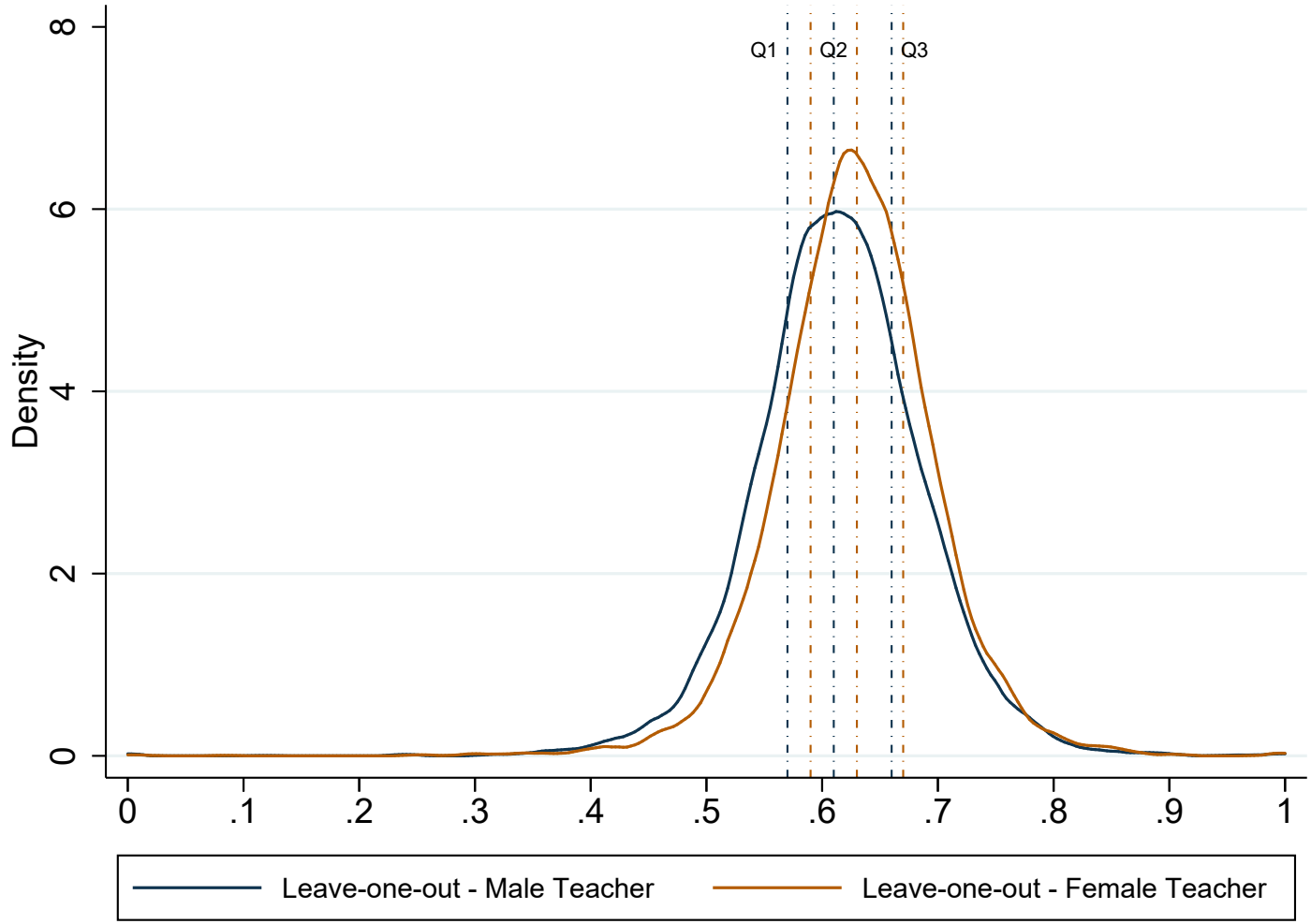




Reading: The word "confidence" is 2.5 times more likely to appear in a females' rather than in a males' feedback.

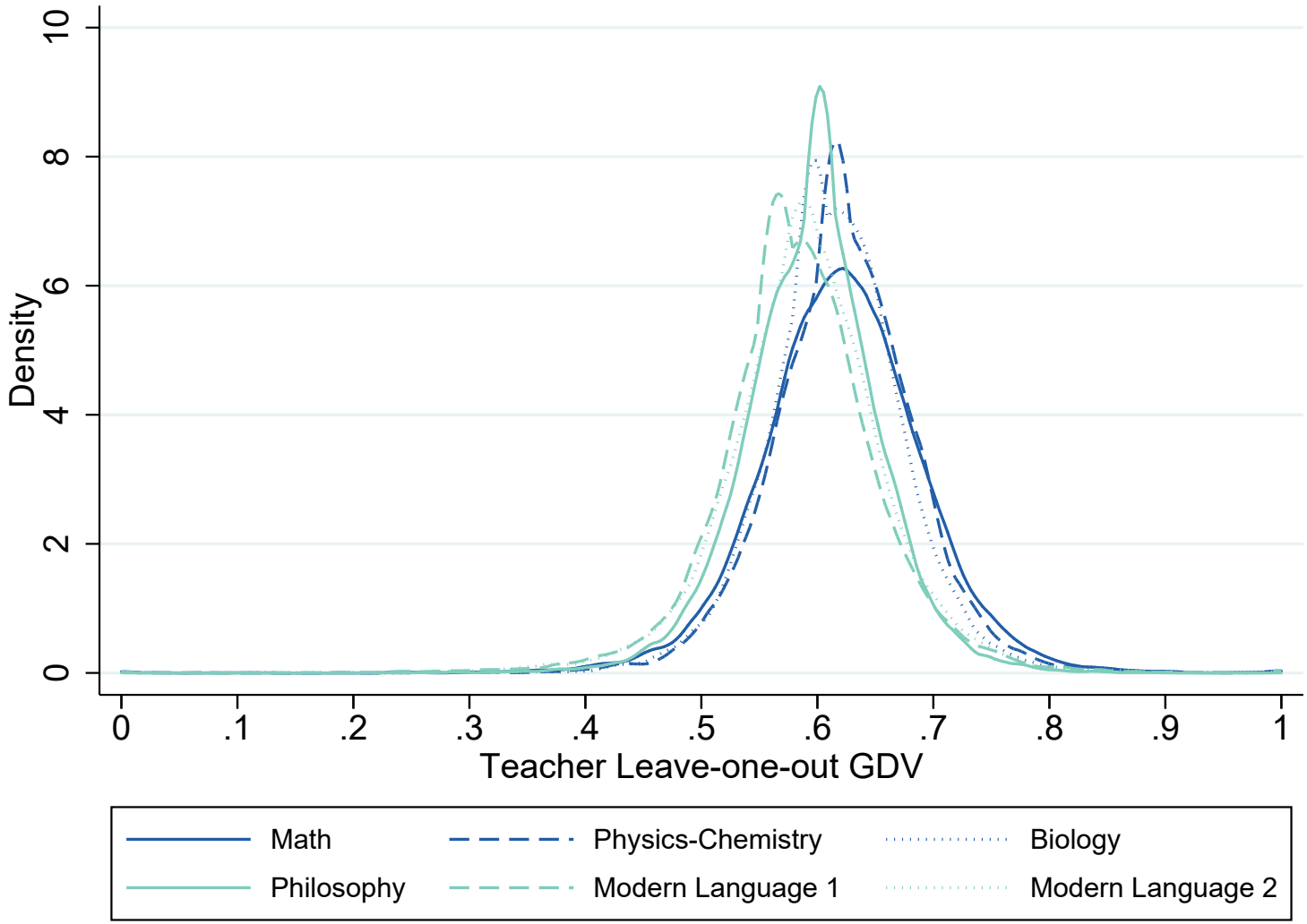
Distribution of GDV - By teacher's gender

Retour



Distribution of GDV - By core subjects

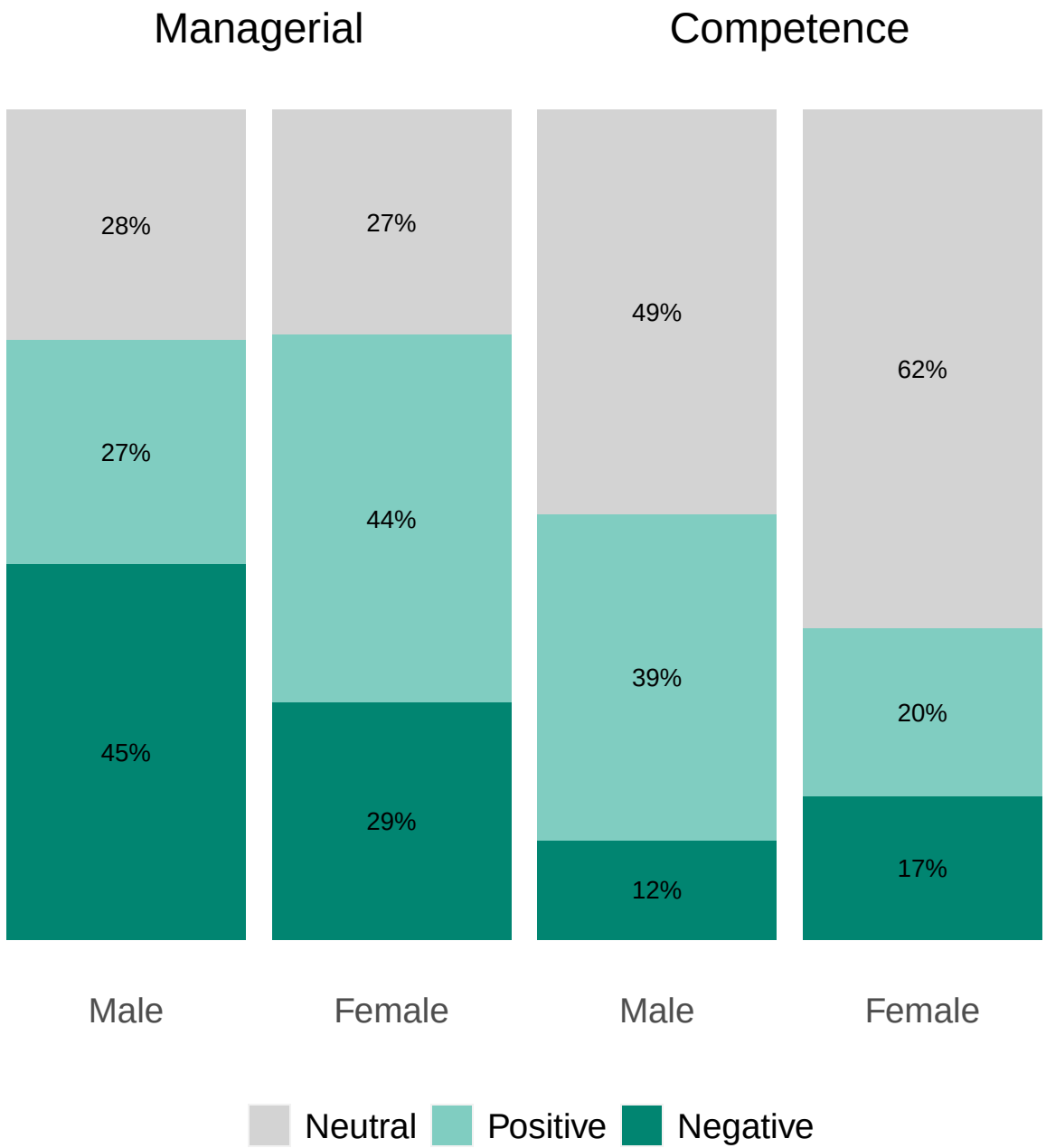
Retour



Note: Steps 1 to 3 have been repeated separately for each subject.

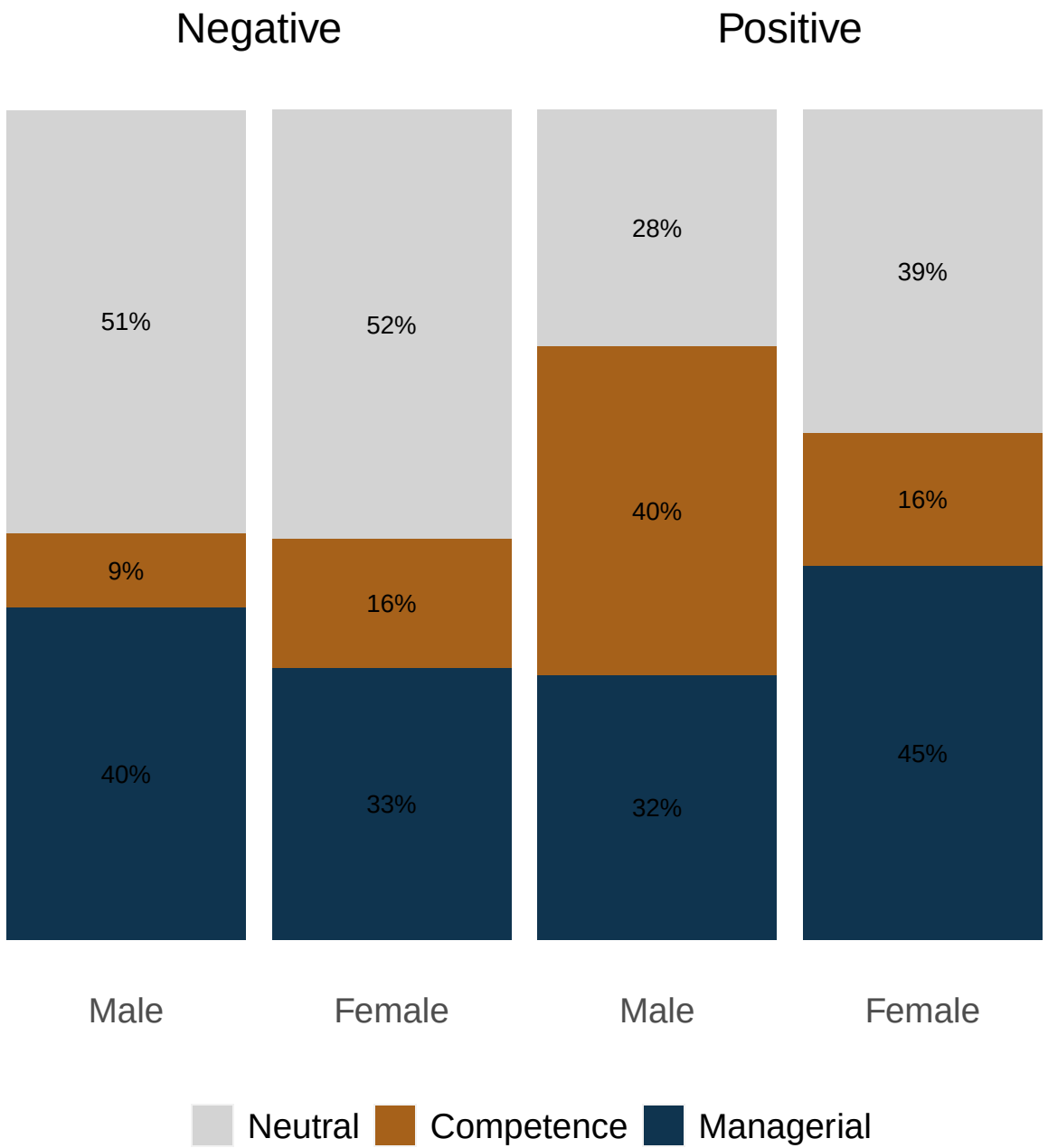
Conditional % of positive vs. negative predictors

Back



Conditional % of managerial vs. competence

Back



Balancing test: GDV on students' characteristics

	Coeff.	S.e	p-value
Female student	-0.0038	0.0028	0.1747
Age (years)	-0.0050 *	0.0026	0.0567
Free lunch student	0.0052	0.0037	0.1575
Foreign student	0.0034	0.0086	0.6929
High SES	-0.2377	0.2682	0.3755
Medium-high SES	-0.2365	0.2682	0.3780
Medium-low SES	-0.2413	0.2680	0.3681
Low SES	-0.2385	0.2683	0.3739
Rank at DNB: math	-0.0000	0.0000	0.5614
Rank at DNB: French	0.0000	0.0001	0.8591
Rank at *baccalauréat*: French (written)	0.0000	0.0001	0.4050
Rank at *baccalauréat*: French (oral)	0.0001	0.0001	0.1510
High school, elective, year FE	Yes		
F-stat (p-value)	1.22	(0.264)	
Nb. of observations	573,600		

Back

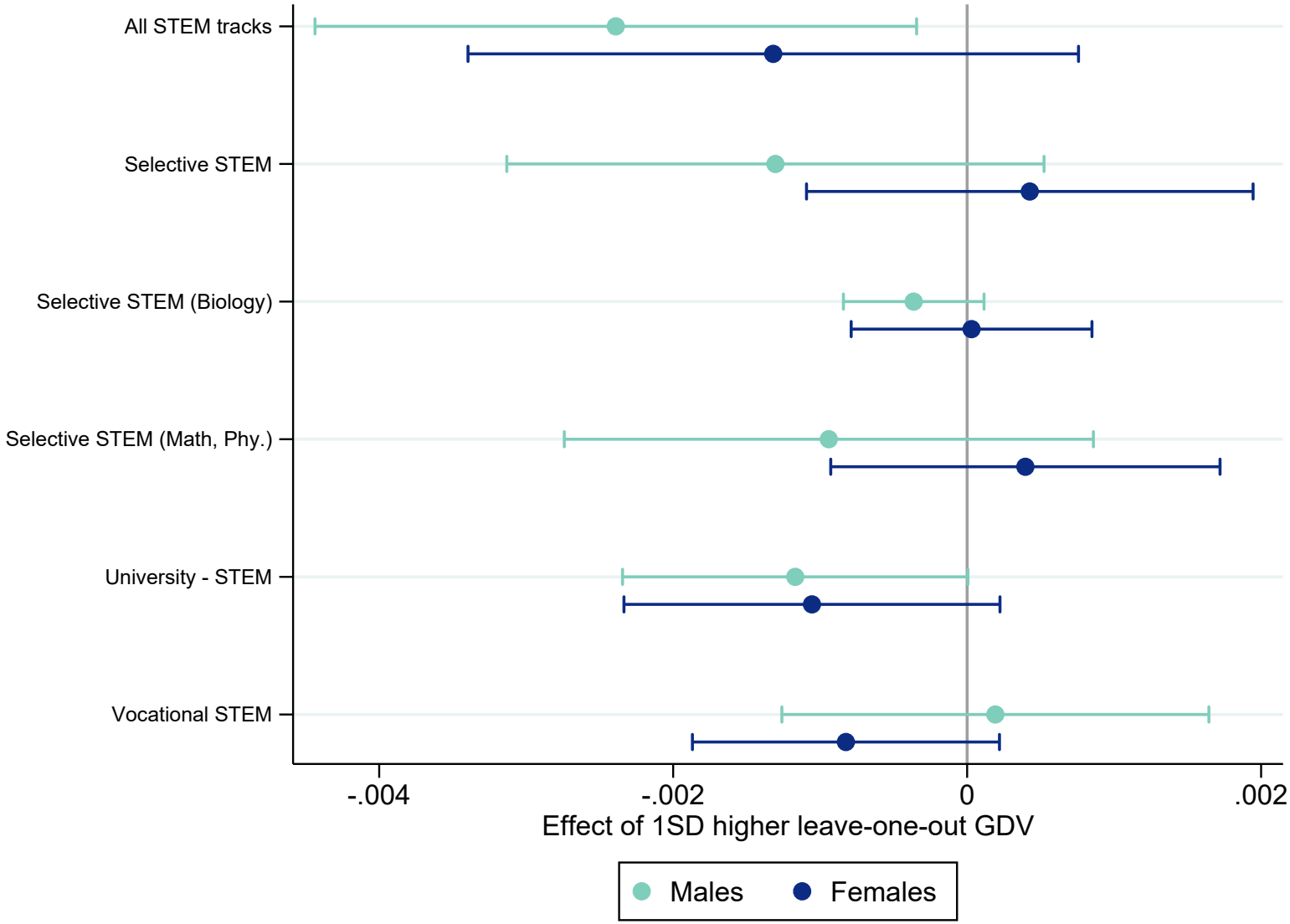
χ^2 tests of random allocation

	Share of significant at	
	5%	1%
Female	11.32	3.37
Age (years)	8.04	2.60
Free-lunch	5.08	1.29
Foreign student	4.05	1.36
High SES	6.72	1.44
Medium-high SES	4.46	0.78
Medium-low SES	5.27	0.93
Low SES	5.53	1.15
Rank at DNB: Math	6.12	1.21
Rank at DNB: French	6.82	1.34
Rank at Baccalaureat: French (written)	7.29	1.58
Rank at Baccalaureat: French (oral)	7.18	1.41

back

Effet of teacher GDV on first wishes

Back



Note: Being exposed to a 1SD higher GDV-teacher reduces the probability of top-ranking a STEM program by 0.2 percentage point for male students.

Effet of teacher GDV on enrollment

Back