Introduction	Setting and Data	Methodology	Results	Conclusion	Appendix
0000	0000	O	00000	O	000

Do Students and Parents Value Same-Gender Teachers? Evidence from an Online Tutor Marketplace

Ini Umosen

University of California, Berkeley

AEA Annual Meeting January 5, 2024

Introduction ●000	Setting and Data	Methodology O	Results 00000	Conclusion O	Appendix 000
Introduct	ion				

The tendency for individuals to associate with others of the same gender (gender homophily), has been shown to be a key component in the formation of social ties

Introduction ●000	Setting and Data	Methodology O	Results 00000	Conclusion O	Appendix 000
Introduc	tion				

- The tendency for individuals to associate with others of the same gender (gender homophily), has been shown to be a key component in the formation of social ties
- Research Questions: Does this phenomenon extend to the classroom? (i.e. do students prefer to work with teachers of the same gender?) Are there sources of heterogeneity such as by academic subject?

Introduction ○●○○	Setting and Data 0000	Methodology ○	Results 00000	Conclusion O	Appendix 000

Measure gender homophily in student-tutor matches on a large online tutor marketplace

Introduction	Setting and Data	Methodology	Results	Conclusion	Appendix
○●○○	0000	O	00000	O	000

Measure gender homophily in student-tutor matches on a large online tutor marketplace

Female tutors 14pp more likely to match with a female student

Introduction	Setting and Data	Methodology	Results	Conclusion	Appendix
○●○○	0000	O	00000	O	000

Measure gender homophily in student-tutor matches on a large online tutor marketplace

- Female tutors 14pp more likely to match with a female student
- Gender matching is especially strong when:
 - There is more parental involvement
 - The subject is math

Introduction	Setting and Data	Methodology	Results	Conclusion	Appendix
○●○○	0000	O	00000	O	000

Measure gender homophily in student-tutor matches on a large online tutor marketplace

- Female tutors 14pp more likely to match with a female student
- Gender matching is especially strong when:
 - There is more parental involvement
 - The subject is math
- Suggestive evidence that the Covid-19 pandemic increased gender matching on the platform

Previous Literature & Contributions

Student preferences for same-gender

Teachers: Carrington et al. (2007), Chen, Zhang, and Wang (2021), Lahelma (2006), McGrath and Sinclair (2013), Skelton et al. (2013)

F

Previous Literature & Contributions

Student preferences for same-gender

Teachers: Carrington et al. (2007), Chen, Zhang, and Wang (2021), Lahelma (2006), McGrath and Sinclair (2013), Skelton et al. (2013)

- Mixed, largely qualitative or survey-based
- Contribution: Use large-scale data on real student-teacher matches

F

Previous Literature & Contributions

Student preferences for same-gender

Teachers: Carrington et al. (2007), Chen, Zhang, and Wang (2021), Lahelma (2006), McGrath and Sinclair (2013), Skelton et al. (2013)

- Mixed, largely qualitative or survey-based
- Contribution: Use large-scale data on real student-teacher matches
- Mentors: Gallen and Wasserman (2023)
 - Analyze observational data from an online mentoring platform and a hypothetical choice experiment

Previous Literature & Contributions

Student preferences for same-gender

Teachers: Carrington et al. (2007), Chen, Zhang, and Wang (2021), Lahelma (2006), McGrath and Sinclair (2013), Skelton et al. (2013)

- Mixed, largely qualitative or survey-based
- Contribution: Use large-scale data on real student-teacher matches

Mentors: Gallen and Wasserman (2023)

- Analyze observational data from an online mentoring platform and a hypothetical choice experiment
- Contributions: Observe a broader range of instruction levels.
 Financial incentives to reveal true preferences. Can tease out student vs parent preferences

Introduction	Setting and Data	Methodology ⊙	Results 00000	Conclusion O	Appendix 000
Related	Literature				

Introduction 000●	Setting and Data	Methodology O	Results 00000	Conclusion O	Appendix 000
Related	Literature				

 Test scores/grades: Ammermueller and Dolton (2006), Antecol, Eren, and Ozbeklik (2015), Bhattacharya et al. (2022), Canann and Mouganie (2023), Dee (2007), Gong, Lu, and Song (2018), Hoffman and Oreopoulous (2009), Holmund and Sund (2008), Parades (2014), Winters et al. (2013)

Introduction 000●	Setting and Data	Methodology O	Results 00000	Conclusion O	Appendix 000
Related	Literature				

- Test scores/grades: Ammermueller and Dolton (2006), Antecol, Eren, and Ozbeklik (2015), Bhattacharya et al. (2022), Canann and Mouganie (2023), Dee (2007), Gong, Lu, and Song (2018), Hoffman and Oreopoulous (2009), Holmund and Sund (2008), Parades (2014), Winters et al. (2013)
- Major choice, interest in STEM: Bettinger and Long (2005), Carrell, Page, and West (2010), Chen, Sonnet, Sadler (2019), Cho (2012), Porter and Serra (2020), Price (2010), Sansone (2017), Sevilla (2023)

Introduction 000●	Setting and Data	Methodology O	Results 00000	Conclusion O	Appendix 000
Related	Literature				

- Test scores/grades: Ammermueller and Dolton (2006), Antecol, Eren, and Ozbeklik (2015), Bhattacharya et al. (2022), Canann and Mouganie (2023), Dee (2007), Gong, Lu, and Song (2018), Hoffman and Oreopoulous (2009), Holmund and Sund (2008), Parades (2014), Winters et al. (2013)
- Major choice, interest in STEM: Bettinger and Long (2005), Carrell, Page, and West (2010), Chen, Sonnet, Sadler (2019), Cho (2012), Porter and Serra (2020), Price (2010), Sansone (2017), Sevilla (2023)
- Student evaluations of teaching: Boring and Philippe (2021)

troduction	Setting and Data	Methodology	Results	Conclusion	Appendix
000	●000	○	00000	O	000

 Analyze data from Wyzant.com, a marketplace that facilitates matches between students and freelance tutors across a variety of academic subjects

troduction	Setting and Data	Methodology	Results	Conclusion	Appendix
	●000	⊙	00000	O	000

- Analyze data from Wyzant.com, a marketplace that facilitates matches between students and freelance tutors across a variety of academic subjects
- Scrape the public profiles of freelance tutors which include reviews left by students and parents

roduction	Setting and Data	Methodology	Results	Conclusion	Appendix
	●000	○	00000	O	000

- Analyze data from Wyzant.com, a marketplace that facilitates matches between students and freelance tutors across a variety of academic subjects
- Scrape the public profiles of freelance tutors which include reviews left by students and parents
- $\blacksquare \to$ Measure gender homophily in the set of student-tutor matches that result in a review



Hourly Rate: \$95

Response time: 3 hours

To make a positive difference and be a part of their academic success is

High School Chemistry College General Chemistry

Alcebra 1

S Hourly Fate: \$95

(ii) Lesson cancellation: 24 hours notice required

Sun Midnight Midnight Mon Midnight Midnight

ed Subjects	Business GRC
	Elementary Edu Elementary Math
	Hameschool Chemistry, SAT IN Algebra 2, Calcul
	Conditional Corra

Algebra 2, Calculus, Geometry,

Examples of Expertise

Solve has provided examples of their subject expertise by answering 1 Meeters 1 examples submitted

O OUTSTON Find the numbers whose

Decease you are looking for 2

Ratings and Reviews

Chemistry, Algebra 1, Algebra 2,

Chemistry, SAT Methy Algebra 1.

Algebra 2, Calculus, GED, Geometry

Chemistry, Anatomy

Semmer

***** 50 159 04000

10.11	
4 star	
3 star	
2 silar	
1 9544	

Exceptional tutor; gives above and beyond?

prepares, communicates with my deaphter's teachers, and prepares short

Sphie is patient and knowledgeable and has helped my daughter

Introduction	Setting and Data	Methodology	Results	Conclusion	Appendix
0000	00●0	O	00000	O	000

Assigning Proxy for Gender

Tutors:

Use first name + Genderize.io

Students:

- Method 1: Review mentions gendered, child-related keywords (e.g. "daughter", "son", "niece")
- Method 2: Name associated with review + Genderize.io if no child keywords used

Knowledgable and patient tutor!

Jane helped mydaughter prepare for an Algebra 1 test and was very helpful! Mydaughter did well on the assignment. I recommend her highly. Great and quick communication.

-Halley 1 lesson with Jane

Introductio	Setting and Data 000●	Methodology O	Results 00000	Conclusion O	Appendix 000
	 <u> </u>				

Descriptive Statistics

- Assign student gender to 89% of reviews, and assign a gender to 88% of tutors
- Analysis sample of 6,463 tutors with 250,000 total reviews

Female Tutors

Male Tutors

Statistic	Mean	St. Dev.	Min	Max	Statistic	Mean	St. Dev.	Min	Max
Female tutor	1.000	0.000	1	1	Female tutor	0.000	0.000	0	0
Female students	0.609	0.319	0.000	1.000	Female students	0.476	0.303	0.000	1.000
Hourly rate	66.231	40.501	16	1,001	Hourly rate	73.675	58.785	11	1,001
5-star rating share	0.957	0.080	0.000	1.000	5-star rating share	0.957	0.079	0.000	1.000
Available hours	50.306	34.281	0	168	Available hours	62.524	36.383	0	168
No. of hours tutored	582.523	1,195.415	1	15,086	No. of hours tutored	689.733	1,319.922	1	13,883
No. of ratings	170.565	356.296	1	4,523	No. of ratings	205.448	389.282	1	5,654
No. of reviews	33.775	74.552	1	1,220	No. of reviews	43.083	89.104	1	1,937
Teaches math	0.482	0.500	0	1	Teaches math	0.680	0.467	0	1
Math review share	0.176	0.277	0.000	1.000	Math review share	0.227	0.292	0.000	1.000
Background check	0.463	0.499	0	1	Background check	0.552	0.497	0	1
Age	36.861	12.084	7	78	Age	37.062	12.057	1.500	88.000
Smile	0.874	0.332	0	1	Smile	0.708	0.455	0	1
Eyeglasses	0.233	0.423	0	1	Eyeglasses	0.316	0.465	0	1
N				3,091	N				3,372

Note: Data collected from the online tutor marketplace Wyzant during fall 2023.

Introduction	Setting and Data	Methodology	Results	Conclusion	Appendix
0000	0000	●	00000	O	000

Measuring Homophily

$StudentShareFemale_{i} = \alpha + \beta TutorFemale_{i} + X_{i} + \epsilon_{i}$ (1)

Introduction	Setting and Data	Methodology	Results	Conclusion	Appendix
0000		•	00000	O	000

Measuring Homophily

 $StudentShareFemale_{i} = \alpha + \beta TutorFemale_{i} + X_{i} + \epsilon_{i}$ (1)

- β is the *relative homophily* statistic
- X': Tutor profile page and review-level attributes

Covariates

Results 00000

Relative Gender Homophily Results

		Dependent variable:					
		Female	e Share of St	tudents			
	(1)	(2)	(3)	(4)	(5)		
Female tutor	0.134*** (0.008)	0.164*** (0.009)	0.143*** (0.010)	0.144*** (0.011)	0.145*** (0.011)		
Constant	0.476*** (0.005)	-4.901 (3.801)	-4.215 (4.108)	-9.033 (5.649)	-8.624 (5.796)		
Tutor & review attributes Word embedding PC Facial landmarks Endogenous variables	(****)	Yes	Yes Yes	Yes Yes Yes	Yes Yes Yes Yes		
Observations Adjusted R ²	6,463 0.044	6,463 0.093	6,463 0.111	6,463 0.111	6,463 0.112		
Note			*n<0.1	· **n<0.05·	***n<0.01		

*p<0.1; **p<0.05; ***p<0.01

Data extracted from Wyzant.com in fall 2023.

F

Tutor Preferences for Gender Matching

 Data from the Ask an Expert forum, a section of the platform where tutors asynchronously respond to student questions

		Dependent variable:					
		Share femal	e of students				
	(1)	(2)	(3)	(4)			
Female tutor	0.027*** (0.008)	0.036*** (0.009)	0.040*** (0.009)	0.023 (0.016)			
Constant	0.636*** (0.005)	0.631*** (0.007)	14.657*** (3.345)	10.636 (7.849)			
Question attributes Tutor attributes Word embedding PC Facial landmarks			Yes	Yes Yes Yes Yes			
Observations Adjusted R ²	25,141 0.001	19,163 0.001	19,163 0.004	19,163 0.012			
Note:	*p<0.1; **p<0.05; ***p<0.01						

Standard errors are clustered at the tutor level. Columns (2) - (4) are estimated on the sample of tutor answers with no missing question or tutor-level covariates.

Relative Gender Homophily Results: Heterogeneity

	Dependent variable:			
		Share Female St	udents	
	Parent	Tutor subject	Review subject	
	(1)	(2)	(3)	
Female tutor	0.085*** (0.016)	0.149 (0.103)	0.040 (0.036)	
Female tutor x Parent review	0.149*** (0.027)			
Female tutor x Math		0.052*** (0.018)	0.186*** (0.044)	
Female tutor x Science		0.004 (0.017)	0.046 (0.053)	
Female tutor x Other		-0.040 (0.103)	0.102** (0.041)	
Constant	-8.786 (5.575)	-9.122 (5.630)	-8.604 (5.618)	
Observations Adjusted R ²	6,463 0.117	6,463 0.112	6,463 0.114	
Note:		*p<0.1; **p<	<0.05; *** p<0.01	

English/writing is the leave-out group in columns (2) and (3).

Results 000●0

C o ion

Appendix 000

Covid-19 Pandemic: Rise in Number of Tutors



Relative Gender Homophily Results: Covid-19 Pandemic

		Dependent variable:					
		Share Female Students					
	Unbalar	Unbalanced Panel Balanced Panel					
	No controls	With controls	No controls	With controls			
	(1)	(2)	(3)	(4)			
Female tutor	0.115***	0.115***	0.121***	0.095***			
	(0.010)	(0.012)	(0.020)	(0.027)			
Post2020	-0.007	-0.014*	-0.025	-0.027			
	(800.0)	(0.008)	(0.016)	(0.017)			
Female tutor × Post2020	0.030**	0.029**	0.039	0.046*			
	(0.012)	(0.012)	(0.026)	(0.028)			
Constant	0.483***	-4.443	0.497***	-9.639			
	(0.007)	(2.771)	(0.011)	(9.938)			
Observations	17,163	17,163	1,948	1,948			
Adjusted R ²	0.031	0.078	0.052	0.156			
Note:	*p<0.1; **p<0.05; ***p<0.01						

Standard errors clustered at the tutor level.



Introduction 0000	Setting and Data	Methodology O	Results 00000	Conclusion	Appendix 000

 Find evidence of strong preferences for same-gender teachers on a large online tutor marketplace

Introduction 0000	Setting and Data	Methodology ⊙	Results 00000	Conclusion	Appendix 000

- Find evidence of strong preferences for same-gender teachers on a large online tutor marketplace
- Parental involvement associated with more gender matching

Introduction	Setting and Data	Methodology	Results	Conclusion	Appendix
0000		O	00000	•	000

- Find evidence of strong preferences for same-gender teachers on a large online tutor marketplace
- Parental involvement associated with more gender matching
- Variation by subject
 - Math stands out
 - Results conflict with a STEM vs Non-STEM framing

Introduction	Setting and Data	Methodology	Results	Conclusion	Appendix
0000		O	00000	•	000

- Find evidence of strong preferences for same-gender teachers on a large online tutor marketplace
- Parental involvement associated with more gender matching
- Variation by subject
 - Math stands out
 - Results conflict with a STEM vs Non-STEM framing
- Future work:
 - Partner with a platform
 - Hypothetical (or real) choice experiment

Email: iumosen@berkeley.edu

Introduction 0000	Setting and Data	Methodology ⊙	Results 00000	Conclusion O	Appendix ●00

Covariates

- Tutor attributes: Background check indicator, rank in search results, approved subject indicators, whether the tutor has an introductory video, is approved to use IXL, has a cancellation policy, has additional hourly rate details, ever publicly commented on a review, the number of alphanumeric characters in the tutor's: biography, explanation of subject qualifications, and education section, the number of answers posted on the Ask an Expert forum, the tutor's state of residence, and
- Review attributes: Whether the review was written by a parent, mentions subject-specific keywords, and the year the review was left.
- Word embeddings PC: 100 principal components extracted from 786 dimensional word embeddings themselves extracted from tutor-written biographies, subject qualification descriptions, and education summaries
- Facial landmarks: Series of variables extracted from tutor profile photos using Amazon Rekognition's facial analysis capabilities.
- Endogenous attributes: Log of hourly rate, number of hours tutored, number of ratings received, share of ratings that are 5 stars



Introduction	Setting and Data	Methodology	Results	Conclusion	Appendix
0000		O	00000	O	0●0

Covid-19 Event Study: Unbalanced Panel



◀ Back

ntroduction	Setting and Data	Methodology O	Results 00000	Conclusion O	Appendix 00●

Covid-19 Event Study: Balanced Panel



Back