

Betting on Diversity – Occupational Segregation and Gender Stereotypes

Urs Fischbacher
University of Konstanz

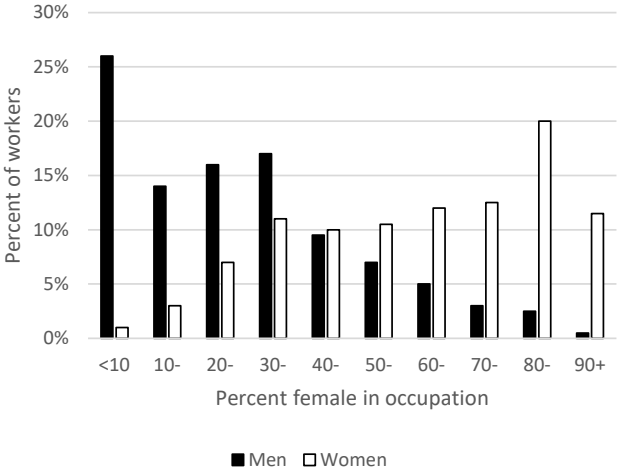
Dorothea Kübler
WZB Berlin

Robert Stüber
C-BID/NYUAD

EEA ESEM 2022

August 24 2022

Motivation: Gender segregation of jobs



Motivation: Gender segregation of jobs

- Gender segregation is costly
 - Misallocation of talent

Motivation: Gender segregation of jobs

- Gender segregation is costly
 - Misallocation of talent
 - Diversity increases innovation and growth

Motivation: Gender segregation of jobs

- Gender segregation is costly
 - Misallocation of talent
 - Diversity increases innovation and growth
 - Strengthen female labor market participation

Motivation: Gender segregation of jobs

- Gender segregation is costly
 - Misallocation of talent
 - Diversity increases innovation and growth
 - Strengthen female labor market participation
 - Reduce gender wage gap

Motivation: Gender segregation of jobs

- Gender segregation is costly
 - Misallocation of talent
 - Diversity increases innovation and growth
 - Strengthen female labor market participation
 - Reduce gender wage gap
- Many potential reasons for segregation
 - Supply-driven (e.g., gender differences in preferences)
 - Demand-driven (e.g., gender discrimination in hiring)

Motivation: Gender segregation of jobs

- Gender segregation is costly
 - Misallocation of talent
 - Diversity increases innovation and growth
 - Strengthen female labor market participation
 - Reduce gender wage gap
- Many potential reasons for segregation
 - Supply-driven (e.g., gender differences in preferences)
 - Demand-driven (e.g., gender discrimination in hiring)
- We focus on demand-sided causes of discrimination
 - Dominance of a gender in job correlates with discrimination of other
Riach and Rich (1987, 2006), Booth and Leigh (2010), Kübler et al. (2018)

Motivation: Gender segregation of jobs

- What can drive the discrimination?
 - Perceived job-specific productivity differences
 - Perceived higher performance of homogeneous groups

Motivation: Gender segregation of jobs

- What can drive the discrimination?
 - Perceived job-specific productivity differences
 - Perceived higher performance of homogeneous groups
- Use lab experiment to disentangle the two explanations by creating counterfactual situations (e.g., stereotypical male jobs performed by women)

Motivation: Gender segregation of jobs

- What can drive the discrimination?
 - Perceived job-specific productivity differences
 - Perceived higher performance of homogeneous groups
 - Use lab experiment to disentangle the two explanations by creating counterfactual situations (e.g., stereotypical male jobs performed by women)
 - Beliefs about productivity
 - Expectations about other people's productivity beliefs

Motivation: Gender segregation of jobs

- What can drive the discrimination?
 - Perceived job-specific productivity differences
 - Perceived higher performance of homogeneous groups
 - Use lab experiment to disentangle the two explanations by creating counterfactual situations (e.g., stereotypical male jobs performed by women)
 - Beliefs about productivity
 - Expectations about other people's productivity beliefs
 - Obtain additional measures of second-order beliefs

Research questions

1. Starting with a homogeneous team, how do people assess the performance of a team that *remains homogeneous* with respect to gender vs. a team that *becomes heterogeneous*?

Research questions

1. Starting with a homogeneous team, how do people assess the performance of a team that *remains homogeneous* with respect to gender vs. a team that *becomes heterogeneous*?
2. Are there differences for stereotypically male and female jobs, and for tasks for which complementarities are more or less important?

Research questions

1. Starting with a homogeneous team, how do people assess the performance of a team that *remains homogeneous* with respect to gender vs. a team that *becomes heterogeneous*?
2. Are there differences for stereotypically male and female jobs, and for tasks for which complementarities are more or less important?
3. Are potential differences in performance assessments reflected in what participants believe about the performance assessments of others?

This paper

- Two experiments
 - Lab experiment with students in Berlin and Konstanz (~ 200 participants)
 - Online experiment with personnel managers and general managers (~ 450 participants)

This paper

- Two experiments
 - Lab experiment with students in Berlin and Konstanz (~ 200 participants)
 - Online experiment with personnel managers and general managers (~ 450 participants)
- Key findings
 - Performance assessments follow task stereotypes
 - Participants expect heterogeneous teams to perform better
 - Preference for heterogeneity is underestimated

Literature

- Role of stereotypes and (distorted) beliefs for gender differences
Spencer et al. (1999), Reuben et al. (2014), Sarsons (2017), Bohren et al. (2019), Bordalo et al. (2019), Carlana (2019), Barron et al. (2020), Coffman et al. (2021), Haeckl and Kartal (2021), Sarsons et al. (2021)
→ Performance expectations about *teams* of different *gender compositions*

Literature

- Role of stereotypes and (distorted) beliefs for gender differences
Spencer et al. (1999), Reuben et al. (2014), Sarsons (2017), Bohren et al. (2019), Bordalo et al. (2019), Carlana (2019), Barron et al. (2020), Coffman et al. (2021), Haeckl and Kartal (2021), Sarsons et al. (2021)
→ Performance expectations about *teams* of different *gender compositions*
- Differences in performance of mixed and single-sex teams
Adams and Ferreira (2009), Kübler and Ivanova-Stenzel (2011), Apestequia et al. (2012), Hoogendorn et al. (2013)
→ Difference in productivity *beliefs* rather than actual productivity

Literature

- Role of stereotypes and (distorted) beliefs for gender differences
Spencer et al. (1999), Reuben et al. (2014), Sarsons (2017), Bohren et al. (2019), Bordalo et al. (2019), Carlana (2019), Barron et al. (2020), Coffman et al. (2021), Haeckl and Kartal (2021), Sarsons et al. (2021)
→ Performance expectations about *teams* of different *gender compositions*
- Differences in performance of mixed and single-sex teams
Adams and Ferreira (2009), Kübler and Ivanova-Stenzel (2011), Apesteguia et al. (2012), Hoogendorn et al. (2013)
→ Difference in productivity *beliefs* rather than actual productivity
- Laboratory experiments with varying subject pools
Gneezy et al (2009), Fréchette (2015, 2016)
→ Results robust across samples speaking to external validity

Experiment 1

Overview of design

- Team sessions to obtain reference performance
 - Teams perform a real-effort task
- Betting sessions as main interest
 - People bet on the relative performance of teams
 - People bet on bets of others

Team sessions

- Stage 1
 - Homogeneous team of four (women or men) performs a task
 - Team gets paid for its performance
 - One member is replaced by another person of same or other gender

Team sessions

- Stage 1
 - Homogeneous team of four (women or men) performs a task
 - Team gets paid for its performance
 - One member is replaced by another person of same or other gender
- Stage 2
 - Newly formed team performs similar task
 - Team gets paid for its performance

The tasks: Building shelves

- Putting together two Ikea shelves
 - Time needed to finish
 - Stereotypically male task



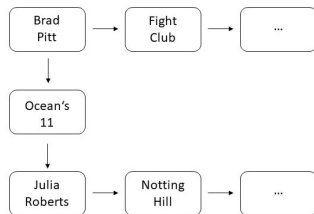
The tasks: Memory

- 12x8 - Memory game
 - Steps needed to finish
 - Stereotypically female task



The tasks: Network

- Writing down movie – actor/actress chains
 - Length of the longest chain
 - Task requiring complementarity



Betting sessions

Bets

- Bet on whether stage 2 performance is higher when the person is replaced with a woman or man
 - Receive €1.20 per correct decision
 - Six bets in total (3 tasks \times 2 types of groups, all male or all female)
 - Order of betting tasks was randomized; use first choice or all choices

Betting sessions

Bets

- Bet on whether stage 2 performance is higher when the person is replaced with a woman or man
 - Receive €1.20 per correct decision
 - Six bets in total (3 tasks \times 2 types of groups, all male or all female)
 - Order of betting tasks was randomized; use first choice or all choices

Expectations

- Bet on how many of 11 randomly chosen others bet on the man
 - Receive €1.20 per correct decision
 - Again six bets

Procedures

Team sessions

- *One* session for each combination of team composition and task
- Half of sessions run *after* the betting sessions

Procedures

Team sessions

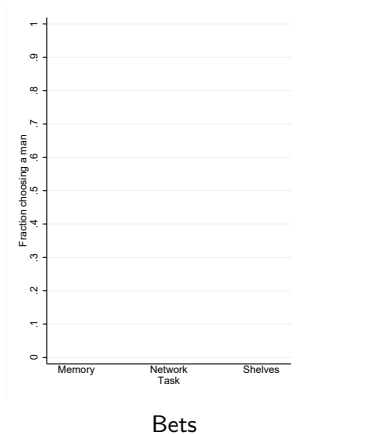
- *One* session for each combination of team composition and task
- Half of sessions run *after* the betting sessions

Betting sessions

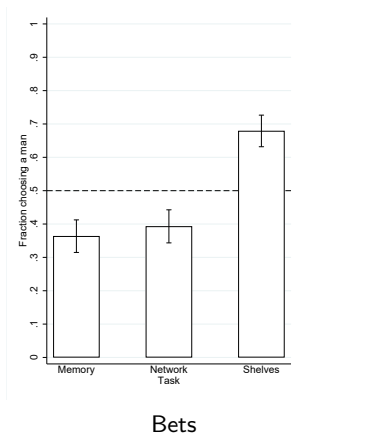
- Half of sessions run in Konstanz, half in Berlin
- 187 subjects (gender-balanced)
- \sim 1 hour, average payment of €12.64

Results

Students' bets and expectations follow task stereotypes

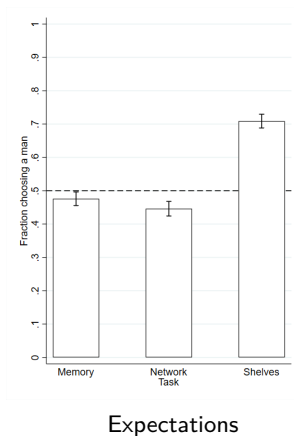
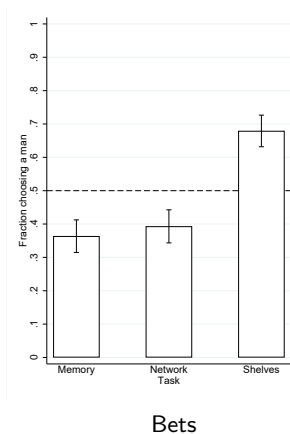


Students' bets and expectations follow task stereotypes



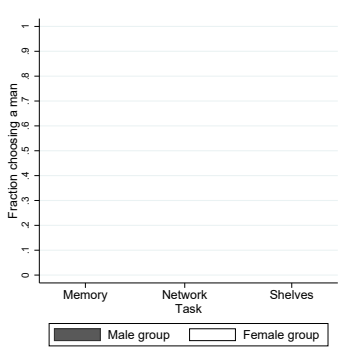
- Most evaluators select man (woman) for the male (female) task

Students' bets and expectations follow task stereotypes



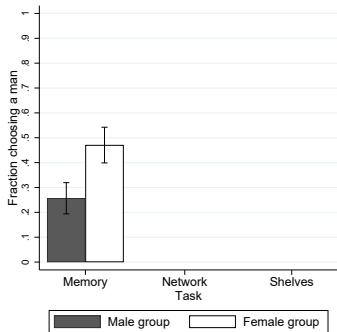
- Most evaluators select man (woman) for the male (female) task
- Difference in choices reflected in expectations about others' choices

Students bet on diverse teams



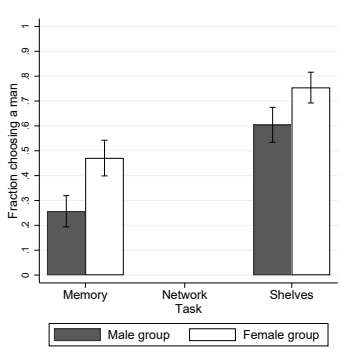
Bets

Students bet on diverse teams



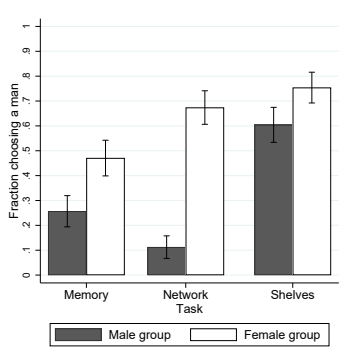
Bets

Students bet on diverse teams



Bets

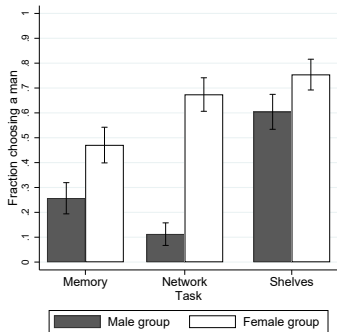
Students bet on diverse teams



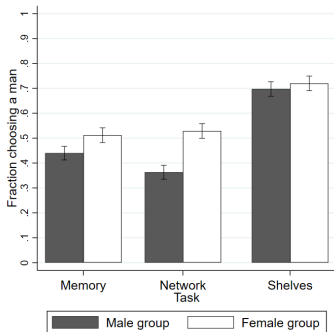
Bets

- Most evaluators predict heterogeneous teams to perform better

Students bet on diverse teams



Bets



Expectations

- Most evaluators predict heterogeneous teams to perform better
- This “preference for heterogeneity” is significantly underestimated

Experiment 2

Motivation

- Preference for heterogeneity might be driven by student sample
- Interest: understanding gender segregation by investigating productivity assessments of individuals who make hiring decisions
 - Conduct simplified version of Experiment 1 with sample of German personnel and general managers

Design and procedures

Design

- Instructions similar, but not identical, to Experiment 1 [▶ Details](#)
- In addition: measure of stereotype [▶ Details](#)

Design and procedures

Design

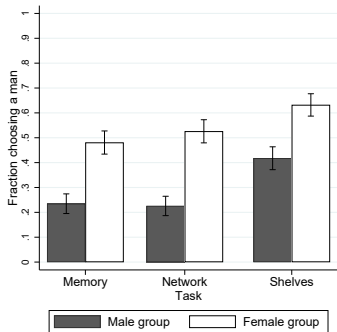
- Instructions similar, but not identical, to Experiment 1 [▶ Details](#)
- In addition: measure of stereotype [▶ Details](#)

Procedures

- Collaborate with data collecting agency *respondi*
- ~ 450 participants (gender-balanced)
- Second experiment preregistered at AEA RCT Registry

Results

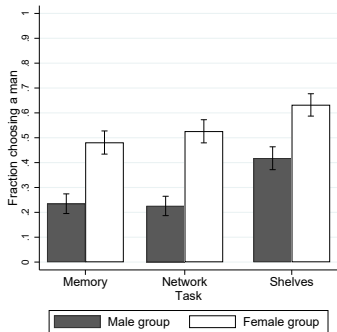
Managers bet on diverse teams



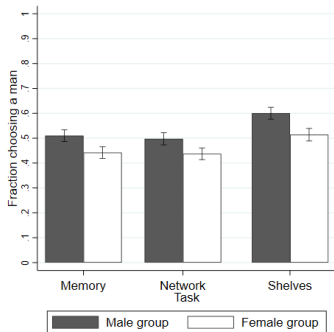
Bets

- Most evaluators predict heterogeneous teams to perform better

Managers bet on diverse teams



Bets



Expectations

- Most evaluators predict heterogeneous teams to perform better
- Participants expect others to prefer homogeneous teams

Summary

Findings (in student and personnel managers sample)

- Participants display gender stereotypes regarding the tasks
- They believe that diverse teams are more productive
- Preference for heterogeneity is underestimated (absolutely and relatively)

Conclusion

Conclusion

Policy implications

- For policies that attempt to reduce gender segregation to be successful, they need to alter stereotypes

Conclusion

Policy implications

- For policies that attempt to reduce gender segregation to be successful, they need to alter stereotypes
- Contrarily, temporary quotas might fail to permanently increase female participation when they do not successfully alter stereotypes

Conclusion

Policy implications

- For policies that attempt to reduce gender segregation to be successful, they need to alter stereotypes
- Contrarily, temporary quotas might fail to permanently increase female participation when they do not successfully alter stereotypes

Implications for literature

- By changing stereotypes, this literature may have also changed beliefs about the workforce

Thank you very much for your attention!

robert.stueber@nyu.edu

Decision screen laboratory experiment

Please make your prediction now. If your prediction is correct, you will receive 1,20 €.

- The team with four men will perform better in part 2 than the team with three men and one woman in part 2.
- The team with three men and one woman will perform better in part 2 than the team with four men in part 2.

Confirm

Decision screen online experiment

Building shelves

We asked two teams, initially consisting of four male participants, to each build two LAIVA shelves from IKEA as quickly as possible. Then, we replaced in one team one member with a man and in the other team one member with a woman.



Both teams then had the task of building two shelves of the type BILLY from IKEA in the 2nd part. Once again, the teams received a higher payout the faster they completely built the two shelves. You can find further details [here](#).

Which of the two teams needed less time to build the two shelves in part 2?

Click on the team. If your guess is correct, you will receive 120 Mingel points in addition.

Part 1: 4 Men



Part 1: 4 Men



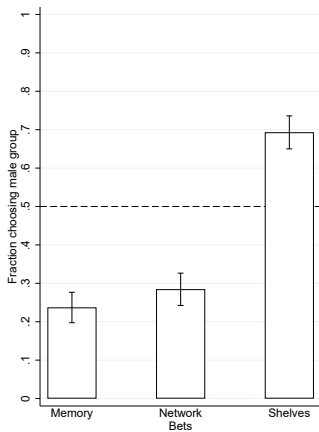
Part 2: 4 Men



Part 2: 3 Men, 1 Women



Stereotypes



Further results

Task stereotype vs. team composition

- Students better at predicting effect of stereotype than of composition

Evaluator gender

- Female and male evaluators' bets respond similarly to stereotype and composition
- Own-gender bias: share of men betting on man (10pp) higher

Classification of evaluator types

- $\sim 2/3$ s of choices in line with a preference for heterogeneity, a preference for adding a person consistent with the job stereotype, or both

Student sample vs. personnel manager sample

	Mean	Sd	p50	Min	Max
<hr/> <i>Panel A. Student sample (n = 187)</i>					
Female	0.508	0.501	1	0	1
Age	22.10	3.996	21	17	50
Education (in years)	14.73	1.545	14.50	13	18
<hr/> <i>Panel B. Sample of personnel and general managers (n = 443)</i>					
Female	0.503	0.501	1	0	1
Age	46.12	12.07	46	20	82
Education (in years)	13.98	2.698	13	9	17
Hiring decisions	2.889	1.294	3	1	5

Further results

Task stereotype vs. team composition

- Students better at predicting effect of stereotype than of composition

Evaluator gender

- Female and male evaluators' bets respond similarly to stereotype and composition
- Own-gender bias: share of men betting on man (7pp) higher

Classification of evaluator types

- $\sim 2/3$ s of choices in line with a preference for heterogeneity, a preference for adding a person consistent with the job stereotype, or both

Heterogeneity across samples

Interaction	Task		Composition
	Network	Shelves	Init. female
SamplexTask	-0.011 (0.036)	-0.149*** (0.041)	
SamplexComposition			-0.055 (0.037)

Heterogeneity analysis manager sample

Interaction	Task		Composition
	Network	Shelves	Init. female
<u>Hiring decisions</u>			
Hiring decisions×Task	-0.053 (0.041)	-0.009 (0.047)	
Hiring decisions×Composition			0.003 (0.041)
<u>Attentiveness</u>			
Attentiveness×Task	-0.024 (0.042)	-0.038 (0.047)	
Attentiveness×Composition			0.056 (0.041)
<u>Age</u>			
Age×Task	0.017 (0.041)	0.019 (0.047)	
Age×Composition			0.045 (0.040)
<u>Education</u>			
Education×Task	-0.003 (0.007)	-0.002 (0.008)	
Education×Composition			0.010 (0.007)

Future directions

- Study beliefs when teams are mixed at the outset or when teams are larger
- Allow for hierarchies within the teams
- Let evaluators come from the set of team members
- Investigate beliefs with task that are even closer to real occupations
- Investigate productivity beliefs in the field
- Analyze hiring decisions (rather than productivity beliefs)