

# An Empirical Analysis of Teacher Assignment Problems

Bryan Graham, UC Berkeley

Geert Ridder, USC

Petra Thiemann, USC Dornsife INET

Gema Zamarro, U Arkansas

March 10, 2016

Loyola Marymount University, Department of Economics

# Matching problems

- ▶ Production inputs
- ▶ Firm supply networks
- ▶ Marriage markets
- ▶ School choice
- ▶ Medical residency

# Matching problems

1. How/under which conditions do agents form matches?

(Galichon and Hsieh, 2015; Hsieh 2014; Menzel 2015)

2. **Match output and allocation effects**

(Graham, 2011; Graham, Imbens and Ridder, 2007, 2010, 2014, 2016)

3. Mechanisms to implement a desired allocation

(Roth, 1984, Roth, Sonmez, and Unver 2004, Abdulkadiroglu, Pathak and Roth, 2009)

# This paper

- ▶ Teacher-classroom assignment  
(Clotfelter, Ladd, and Vigdor, 2005, Kalogrides and Loeb 2013, Dieterle, Guarino, Reckase, and Wooldridge 2012)
- ▶ Match output
- ▶ Allocation effects
- ▶ Method: Graham, Imbens, and Ridder, 2007
- ▶ Data: Measures of Effective Teaching Study

# Example: Teachers and classrooms

- ▶ Objective: maximize average test scores
- ▶ Constraints:
  - ▶ available teachers (teacher quality)
  - ▶ available classrooms (average student quality)
- ▶ Choice: Teacher-classroom match/assignment

## Teacher quality

low	25
middle	50
high	25

100

## Classroom quality

low	30
middle	30
high	40

100

# Focal assignments

1. Status quo
2. Random assignment
3. Positive sorting
4. Negative sorting
5. Optimal assignment

# Education production function

$$y_c = g(w_c, x_c) + \epsilon_c \quad (1)$$

$y_c$  average test scores in classroom  $c$

$w_c$  quality of teacher in classroom  $c$

$x_c$  quality of classroom  $c$

$g(\cdot)$  function, to be specified

$\epsilon_c$  error term

$$y_c = \sum_k \sum_l \gamma_{kl} I(w_c = k) * I(x_c = l) + \epsilon_c \quad (2)$$

$k$  level of teacher quality  $k \in \{low, middle, high\}$

$l$  level of classroom quality  $l \in \{low, middle, high\}$

$I(\cdot)$  indicator function

# Design

1. Schools sign up for the study
2. Teachers sign up for the study
3. School year 1 begins
4. Researchers collect data on teacher quality and student outcomes
5. School year 1 ends
6. Principal selects relevant classrooms
7. Researchers assign teachers randomly to classrooms
8. School year 2 begins
9. Researchers collect data on teacher quality and student outcomes
10. School year 2 ends



Table: Design

Section/ classroom	School	Subject	Grade	Randomization group
1 2	A	ELA	9	1
3 4	A	Math	7	2
5 6 7	B	Math	5	3

Table: Student compliance to the randomization scheme

Compliance	Percent
< 20%	0.5
20- 29%	2.1
30-39%	7.7
40-49%	8.0
50-59%	6.9
60-69%	15.2
70-79%	14.9
80-89%	21.3
90-99%	2.9
100%	20.5

Based on 8,575 observations and 376 classrooms.

# Sample

- ▶ English Language Arts (ELA) (376 sections)
- ▶ at least 50% compliance (drop 69 sections)
- ▶ variation in teacher quality within randomization group (drop 130 sections)
- ▶ Grades 4-9
- ▶ 5 out of 6 school districts  
Charlotte-Mecklenburg, Dallas, Denver, Hillsborough County,  
Memphis City, New York City
- ▶ 54 schools
- ▶ 67 randomization groups
- ▶ 177 teachers/sections
- ▶ 4,077 students

Table: Descriptive statistics

	Mean	SD	Min.	Max.
Teacher variables				
Quality score	0.06	0.83	-2.07	1.63
Low (tercile 1)	29%			
Middle (tercile 2)	41%			
High (tercile 3)	29%			
Student variables				
Avg. ELA score year 1	0.19	0.60	-1.27	1.52
Low (tercile 1)	30%			
Middle (tercile 2)	33%			
High (tercile 3)	36%			
Outcome				
Avg. ELA score year 2	0.19	0.57	-1.05	1.55
Instrumental variable				
Avg. ELA score year 1 (assigned)	0.18	0.64	-1.60	1.53

# Education production function

$$y_{cr} = \sum_k \sum_l \gamma_{kl} I(w_{cr} = k) * I(x_{cr} = l) + \delta_r + \epsilon_{cr} \quad (3)$$

$y_{cr}$  average test scores in classroom  $c$

$w_{cr}$  quality of teacher in classroom  $c$

$x_{cr}$  quality of classroom  $c$

$\delta_r$  randomization group fixed effects

$\epsilon_{cr}$  error term

$k$  level of teacher quality  $k \in \{low, middle, high\}$

$l$  level of classroom quality  $l \in \{low, middle, high\}$

$I(.)$  indicator function

Instrument  $I(w_{cr} = k) * I(x_{cr} = l)$  with  $I(w_{cr} = k) * I(x_{cr}^{assign} = l)$

Table: Output (average grades)

	<u>Classroom</u>		
	low	middle	high
<u>Teacher</u>			
low	-0.35	0.25	0.62
middle	-0.31	0.04	0.76
high	0.29	0.18	0.68

Table: Status quo

	Classroom		
	low	middle	high
Teacher			
low	0.16	0.07	0.06
middle	0.11	0.13	0.17
high	0.04	0.12	0.14

Table: Random

	Classroom		
	low	middle	high
Teacher			
low	0.09	0.10	0.11
middle	0.13	0.14	0.15
high	0.09	0.10	0.11

Table: Positive sorting

	Classroom		
	low	middle	high
Teacher			
low	0.29	0	0
middle	0.01	0.33	0.07
high	0	0	0.29

Table: Negative sorting

	Classroom		
	low	middle	high
Teacher			
low	0	0	0.29
middle	0.01	0.33	0.07
high	0.29	0	0

Table: Optimal allocation

	Classroom		
	low	middle	high
Teacher			
low	0	0.29	0
middle	0.01	0.04	0.37
high	0.29	0	0



# Estimation and inference

- ▶ Outcome for each allocation: Weighted average of the production function estimates
- ▶ Inference procedure: Randomization inference
  - ▶ Uncertainty about the production function estimates
  - ▶ Specific form of production function could be due to chance
  - ▶ want to know whether effects are “exceptional”
  - ▶ exact p-value (Fisher)
  - ▶ Sharp null: No effect of reallocations (i.e. no complementarities)
  - ▶ Construct the randomization distribution under the null
  - ▶ Based on counterfactual assignments
  - ▶ E.g. 60 randomization blocks with two classroom teacher pairs –  $2^{60}$  possible assignments
  - ▶ Generate random subset of these allocations
  - ▶ Estimate production function and reallocation effects
  - ▶ Compare randomization distribution of reallocation effects to actual effects

Table: Full sample, unrestricted reallocation

	Sample mean	Reallocation distribution mean	p-value
<i>Average outcome</i>			
Status quo (SQ)	0.19	-	-
<i>Reallocation effects</i>			
Random - SQ	0.03	0.00	0.00
Positive sorting - SQ	-0.07	-0.04	0.19
Negative sorting - SQ	0.10	-0.04	0.00
Optimal - SQ	0.21	-0.03	0.00

# Outlook

- ▶ Restricted allocations (only within district)
- ▶ Relax constraints (e.g. allow for improvement of teacher quality)
- ▶ Continuous measure for teacher and student quality

Thank you!

Outcome: Testscores	OLS	IV
Teacher quality middle	-0.003 (0.056)	0.018 0.061
Teacher quality high	0.274*** (0.093)	0.677* 0.400
Classroom quality middle	0.831*** (0.087)	0.506*** 0.156
Classroom quality high	0.831*** (0.087)	0.883*** 0.108
Teacher middle X classroom middle	-0.062 (0.096)	-0.194 0.144
Teacher middle X classroom high	0.062 (0.080)	0.100 0.094
Teacher high X classroom middle	-0.246** (0.124)	-0.690 0.506
Teacher high X classroom high	-0.195 (0.127)	-0.658 0.400
Randomization group fixed effects	Yes	Yes
p-value of F-test:		
joint significance of interaction terms	0.192	0.061
Obs.	4,077	4,077

Table: Three largest districts, unrestricted reallocation

	Sample mean	Reallocation distribution mean	p-value
<i>Average outcome</i>			
Status quo (SQ)	0.24	-	-
<i>Reallocation effects</i>			
Random - SQ	0.05	0.00	0.00
Positive sorting - SQ	-0.07	-0.04	0.19
Negative sorting - SQ	0.19	-0.04	0.00
Optimal - SQ	0.29	-0.02	0.00