Empirical Economics

Case-based seminar I - Research with OLS

Teacher: Andrew Proctor
andrew.proctor@phdstudent.hhs.se

September 14, 2017

Outline

Today's main paper

The assigned paper for today is:

> "Does autonomy over teacher hiring affect student math and science achievement?"
>
> Youngran Kim
>
> *Education Economics* (2017)

Introduction

What is the general topic that Kim is interested in?

Introduction

What is the general topic that Kim is interested in?

- Kim's analysis is related to economic literature considering how reforms to decentralization decision-making affects student achievement.

What is the specific research question addressed in this analysis?

Introduction

What is the general topic that Kim is interested in?

- Kim's analysis is related to economic literature considering how reforms to decentralization decision-making affects student achievement.

What is the specific research question addressed in this analysis?

- For South Korean middle school students in 1995, how did greater private school flexibility in teacher hiring decisions affect student test scores?

Source of 'Identification'

What are the basic features of the Kim's dataset that she claims allows her to estimate the causal effect of hiring autonomy?

Source of 'Identification'

What are the basic features of the Kim's dataset that she claims allows her to estimate the causal effect of hiring autonomy?

1. *Claimed* randomization of middle school students into schools (public or private) conditional on school district.

Source of 'Identification'

What are the basic features of the Kim's dataset that she claims allows her to estimate the causal effect of hiring autonomy?

1. *Claimed* randomization of middle school students into schools (public or private) conditional on school district.
2. Kim *argues* that the only significant difference between public and private schooling organization & operation is in teaching hiring autonomy.

Source of 'Identification'

What are the basic features of the Kim's dataset that she claims allows her to estimate the causal effect of hiring autonomy?

1. *Claimed* randomization of middle school students into schools (public or private) conditional on school district.
2. Kim *argues* that the only significant difference between public and private schooling organization & operation is in teaching hiring autonomy.
    - Curriculum taught in both instances is the national curriculum
    - Tuition and teacher salaries are the same in both systems.
    - Teacher hiring is at the province level for public schools, school level for private schools.
    - Teacher contracts less flexible in public schools.

Data

- **Dataset:** South Korean educational data from 1995 'Third International Mathematics and Science Study' (TIMMS)
  - A three-stage (school, class, student) random sample of 13-year-old students.
- **Variables of analysis:**
  - Outcome: math and science scores from TIMSS tests.
  - Explanatory variable: inferred 'private' indicator
  - Control variables:
    - *Indicators for* 'parent graduated college,' female, grade, 'disadvantaged,' and 'outskirts' school location.
    - *Quantitative*: class size, index of home resources.
    - School district *fixed effects*.

Table 1. Descriptive statistics by school type (school-level data).

Variable	Public	Private
Student characteristics		
% disadvantaged backgrounds	8.88	7.93
% parents no primary education	3.16	4.44
% one parent	6.17	5.99
% attended preschool	45.22	49.95
% 1st language other than Korean	0.50	1.31
% learning problems	2.27	1.17
% health problems	0.37	1.22
% nutrition problems	0.53	0.61
School characteristics		
Teacher student ratio	27.96	29.25
Average class size	50.49	51.10
% schools located in city	60.49	52.18
% boys-only schools	26.46	43.17
% girls-only schools	24.12	26.54
Teacher characteristics		
% part-time teachers	0.06	2.41
Years of teaching	11.45	13.24
% male	46.64	57.80
% master degree	13.59	17.10
% experience >3 years	88.22	88.29
Number of schools	57	69

Note: School Level Weight is applied.

Table 2. Descriptive statistics by school type (student-level data).

Variable	Public	Private
Math scores	576.46	576.69
	(80.06)	(77.59)
Science scores	538.02	537.61
	(80.07)	(77.53)
% female	46.44	41.18
% seventh grader	50.26	48.73
# book at home	38.20	38.23
Home resources	7.95	8.02
% father with college degree	27.79	29.29
% mother with college degree	13.33	16.07
Number of students	2213	2674

Note: Student Level Weight is applied. Standard deviations are reported in parentheses.

Balance Test

To further check whether random assignment 'worked,' Kim also does what's called a *balance test*.

- A balance test uses regression to estimate whether there is a nonzero relationship between the randomized 'treatment' and observable characteristics about individuals.
- Usually, this is done by running a series of regressions, each using the randomized variable as an explanatory variable to explain an observable characteristic.
  - For instance $female_i = \alpha + \beta_1 private_i + e_i$, if we consider $private_i$ to be random.
- Could also run a regression with the randomized variable as the outcome variable and all the observables as the explanatory variables, then run an F-test on the observables.
  - Kim does something like this second method.

Balance Test

Table 3. Covariate balance and verification of random assignment (student-level data).

Variable	Private	SE
Math scores	0.227	−2423
Science scores	−0.407	(2.351)
% female	−0.0525***	(0.0147)
% seventh grader	−0.0153	(0.0148)
# book at home	0.0334	(0.440)
Home resources	0.0745*	(0.0412)
% father with college degree	0.015	(0.0139)
% mother with college degree	0.0273**	(0.0107)
Observation	4887	

Note: Standard errors are reported in brackets. The regression is weighted by Student Level Weight.
*Statistical significance at the 10% level.
** Statistical significance at the 5% level.
***Statistical significance at the 1% level.

Balance Test ctd

Table 4. Covariate balance and verification of random assignment (school-level data).

Variable	Private	SE
Student characteristics		
% disadvantaged backgrounds	−0.952	(2.358)
% parents no primary education	1.279	(2.092)
% one parent	−0.176	(0.968)
% attended preschool	4.725	(6.110)
% 1st language other than Korean	0.815	(0.898)
% learning problems	−1.049*	(0.606)
% health problems	0.144	(0.245)
% nutrition problems	0.0825	(0.234)
School characteristics		
Teacher student ratio	1.287	(0.847)
Average class size	0.610	(0.763)
% schools located in city	−0.0830	(0.113)
% boys-only schools	0.167*	(0.0988)
% girls-only schools	0.0242	(0.0849)
Teacher characteristics		
% part-time teachers	0.0235	(0.0145)
Years of teaching	1.795	(1.244)
% male	0.112*	(0.0622)
% master degree	0.0351	(0.0336)
% experience >3 years	0.000709	(0.0354)
Observation	126	

Note: Standard errors are reported in brackets. The regression is weighted by School Level Weight.
*Statistical significance at the 10% level.

Revising Kim's Balance Test

What is the difference in the balance test Kim uses and method 2 I suggested?

Revising Kim's Balance Test

What is the difference in the balance test Kim uses and method 2 I suggested?

- The second method uses an F-test to test against the null joint hypothesis that *ALL* the observables have no effect.
- Kim similar puts them in all in a single regression, but instead looks only at the p-values for t-tests (individual hypothesis).

Revising Kim's Balance Test ctd

What happens when you perform a t-test on a parameter estimate when there are irrelevant variables in the model?

Revising Kim's Balance Test ctd

What happens when you perform a t-test on a parameter estimate when there are irrelevant variables in the model?

- You increase the variance of the parameter estimate relative to the true model, making the t-statistic smaller and increasing the probability of a Type II error.
- This is why you run an F-test if you're running a 'big' regression. Otherwise you should run a bunch of simple regressions for each observable characteristic.

About random assignment

Do the descriptive statistics look balanced?

About random assignment

Do the descriptive statistics look balanced?

- Throughout the descriptive statistics and 'balance' test section, there appeared to be some consistent types of differences between public and private schools.
- In particular, students in private schools are modestly, but consistently socioeconomically 'better off' than students in public schools
- Private schools are also substantially more likely to be located in cities and to be 'boys-only.'

About random assignment ctd

Is the incomplete randomization a problem? How does Kim address concerns about randomness of private school assignment?

About random assignment ctd

Is the incomplete randomization a problem? How does Kim address concerns about randomness of private school assignment?

- Actually, most of the reported differences might not be a problem at all, because the descriptive statistics and balance test don't correspond to the randomization Kim uses!

About random assignment ctd

Is the incomplete randomization a problem? How does Kim address concerns about randomness of private school assignment?

- Actually, most of the reported differences might not be a problem at all, because the descriptive statistics and balance test don't correspond to the randomization Kim uses!
  - *How so?*

About random assignment ctd

Is the incomplete randomization a problem? How does Kim address concerns about randomness of private school assignment?

- Actually, most of the reported differences might not be a problem at all, because the descriptive statistics and balance test don't correspond to the randomization Kim uses!
  - *How so?* Kim argues for randomness of private school assignment *conditional* on school district.
- *How should descriptive stats and balance tests be presented?*
  - Need to present descriptive statistics in terms of conditional means ($\Longrightarrow$ need to show the difference in statistics rather than a side by side comparison)
  - Need to include district indicators in balance test.

About random assignment ctd

- For the 'boys-only schools' issue, Kim controls for gender of student (which determines eligibility for 'boys only' private schools)
  - This essentially randomizes assignment, but does not address the fact that private schools differ in other ways than just teacher hiring (ie gender composition of peers).

Table 5. Effect of private school education on student achievement.

	Math			
Variables	(1)	(2)	(3)	(4)
Private	1.092	1.386	2.045	2.318
	(4.450)	(4.499)	(4.3016)	(5.365)
Female		−14.91***	−16.24***	−16.15***
		(3.513)	(3.014)	(3.517)
Seventh grader		−21.28***	−23.95***	−23.71***
		(2.361)	(2.407)	(2.462)
# books at home			1.521***	1.474***
			(0.0916)	(0.0966)
Home resources			7.667***	8.276***
			(0.952)	(0.955)
Father with college degree			14.17**	13.72***
			(3.453)	(3.761)
Mother with college degree			6.826*	7.668*
			(3.788)	(4.113)
% students with low SES				−0.128
				(0.334)
Teacher student ratio				0.360
				(2.148)
Average class size				0.829
				(2.117)
District fixed effect	√	√	√	√
Observations	4887	4882	4268	3898
R-squared	0.084	0.104	0.233	0.233

Table 5. Effect of private schoo

	Science			
Variables	(1)	(2)	(3)	(4)
Private	0.297	0.987	1.030	1.548
	(5.458)	(5.238)	(4.856)	(6.638)
Female		−26.03***	−27.04***	−26.36***
		(4.383)	(3.868)	(5.054)
Seventh grader		−25.19***	−27.41***	−27.05***
		(2.028)	(2.047)	(2.102)
# books at home			1.564***	1.517***
			(0.0944)	(0.0993)
Home resources			5.013***	5.239***
			(1.163)	(1.261)
Father with college degree			10.79***	10.91***
			(3.238)	(3.540)
Mother with college degree			7.005	7.618
			(4.756)	(5.231)
% students with low SES				0.281
				(0.405)
Teacher student ratio				−0.393
				(1.640)
Average class size				−0.0649
				(3.159)
District fixed effect	√	√	√	√
Observations	4887	4882	4779	4367
R-squared	0.067	0.104	0.251	0.251

Concluding thoughts

- Because of the way Kim presents results, it is difficult to assess whether or not the South Korean education policy at this time actually randomized private school placement.
- Assuming the randomization was successful, it seems there is little effect on private school placement on the middle school students' test scores.

Concluding thoughts ctd

- Because the descriptive statistics make it seem uncertain whether or not private and public schools *really* only differed in teaching hiring autonomy, it seems better to interpret estimated effects for private schooling as a whole in this setting.
- Control variables which are not randomized do not become exogenous just because one of the variables has been randomized.
    - The coefficients for the randomized variable are still unbiased and consistent, but the coefficients for these controls that may be subject to omitted variable bias should not be interpreted causally.

Second paper:

"Separate and Unequal in the Labor Market: Human Capital and the Jim Crow Wage Gap"

Celeste K. Carruthers and Marianne H. Wanamaker

Journal of Labor Economics, vol. 35, no. 3 (2017)

Context

- From the late 19^{th} century until the mid-1960s, black persons in America were systematically discriminated against in almost every facet of civic life by a system of legislation known as the 'Jim Crow' laws.
- The mendacious fiction that the U.S. Supreme Court used to justify this segregated system in *Plessy v. Ferguson* (1896) was the notion that providing 'separate but equal' facilities for black persons did not violate the U.S. Constitution.
- In practice, the 'separate' part was embraced, but not the 'equal.' Facilities for black persons and communities were much worse than their white neighbors.

Research Topic

- **General topic**
  - Carruthers and Wanamaker are interested in how much this *unequal* system of provisioning, specifically in education, contributed to labor market inequality between black and white persons.
- **Specific research question**:
  - How much can difference in schooling resources (i.e. spending) for white and black persons during basic education account for the black-white wage gap among residents of the Southern U.S. in 1940.

Data

- **Dataset:** Public Use Sample of the 1940 US Census and various state & federal publications of schooling statistics
  - Use Census data to identify labor market outcomes for young working men (18-25) in 10 southern states in 140.
  - Use educational financing statistics to develop estimates of educational resources, using a combination of information on expenditures, staffing, teacher salaries, and term length.
    - This proxies for educational quality.

Variables of Analysis

- Outcome: racial gap in either:
  - weekly/annual wages, occupation score, or weeks worked
- Explanatory variable:
  - School quality measure: avg standardized score for available school resources metrics, for each county-year combo).
  - Years of schooling (in some specifications).
- Control variables:
  - *Indicators for:* age.
  - *Quantitative*: County % urban population, per capita retail sales, crop value, & manufacture value added.
  - State / county *fixed effects*.

Authors' identification argument

- Conditional on going to school in the same county, racial differences in educational resources were likely to be politically determined rather than based on expectations about workers' future labor market outcomes.
- Including state (or county) indicators control for any fixed general characteristics of the region, potentially including other determinants of the racial wage gap.
    - For instance, according to the authors, the 'degree of discrimination' in the state or county.

Comments on identification argument

- It's hard to get good identification on more distant historical phenomenons.
- That racial difference in educational investment may've been largely political isn't too reassuring from an OVB standpoint.
  - The same prejudicial perspectives that determine lower schooling investment was likely related to attitudes and practices limiting black persons in other ways too.
- State or county indicators can control for fixed characteristics of the region that affect both white and black persons, but more likely sources of OVB are things that affect black and white persons in a region deferentially - like job discrimination.
  - The author's arguments that these fixed effects control for the discrimination environment is unconvincing.

Regression specification

$$\ln Y_{icra} = \alpha + \delta \text{Black}_i + \beta X_{icra} + \varepsilon_{icra}, \quad (1)$$

where Y_{icra} is the labor market outcome of interest for individual i educated in county c residing currently in county r of age a. In this setting, Y_{icra} measures one of four labor market outcomes: weekly wages, occupation score, annual wages, or weeks worked. Black_i is a binary indicator, and the estimated wage gap is negative when black respondents have lower (conditional) labor market outcomes than whites. In all specifications, we cluster standard errors by the 1940 county of residence.

Summary Statistics

	All Black (1)	All White (2)	Baseline Sample Black (3)	Baseline Samp White (4)
Individual:				
Average annual wage income (1939) in natural log	5.42	5.93	5.42	5.93
Average weekly wage (1939) in natural log	1.87	2.40	1.87	2.40
% Reporting	58.8	56.0	100.0	100.0
Occupational score[a] in natural log	6.99	7.35	6.99	7.35
% Reporting	87.0	82.2	97.5	96.4
Average weeks worked (1939)	40.9	40.8	39.0	38.8
Unemployment rate[b]	9.2	9.5	9.9	8.9
If unemployed, duration (continuous weeks)[b]	38.8	45.4	35.4	43.8
Labor force participation rate[a]	88.9	85.0	98.2	97.5
Highest grade completed	5.6	8.9	5.6	8.9
School quality index (standardized (0,1))	−.55	.48	−.50	.55

Summary Statistics

State of residence in 1940:				
Alabama	12.8	9.0	12.3	8.7
Arkansas	5.6	6.5	4.0	6.0
Georgia	15.5	9.6	18.1	9.7
Kentucky	2.7	12.2	3.0	10.9
Louisiana	11.4	7.5	12.4	7.5
Mississippi	6.6	2.2	3.8	2.2
North Carolina	14.3	12.6	14.3	12.7
South Carolina	13.9	5.6	14.0	6.2
Tennessee	5.8	11.6	6.4	11.0
Texas	11.6	23.3	11.8	25.0
County of residence:				
Percent rural	68.4	67.2	64.2	62.0
Per capita manufacturing value ($1,940)	69.6	73.6	80.1	86.7
Per capita retail sales ($1,940)	.19	.20	.21	.22
Per capita crop value ($1,940)	59.3	53.0	53.4	46.3
Number of observations	5,423	14,849	3,141	8,253

Table 4
Estimates of Black-White Labor Market Outcome Gaps

	Outcome							
	ln(Weekly Wage)				ln(Occupation Score)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Unconditional gap (SE)		−.529 (.024)				−.359 (.016)		
Black-white gap	−.490	−.315	−.181	−.191	−.334	−.203	−.160	−.168
	(.022)	(.022)	(.031)	(.032)	(.014)	(.015)	(.021)	(.022)
Contribution of school quality			−.140	−.011			−.044	.136
			(.022)	(.046)			(.016)	(.041)
Contribution of educational attainment			−.168	−.164			−.129	−.129
			(.010)	(.011)			(.007)	(.008)
Contribution of interaction				−.123				−.174
				(.043)				(.038)
Age and county controls?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Years of schooling?	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School quality?	No	No	Yes	Yes	No	No	Yes	Yes
Interacted HC controls?	No	No	No	Yes	No	No	No	Yes
N	11,394	11,394	11,394	11,394	11,021	11,021	11,021	11,021
Adjusted R^2	.24	.29	.30	.30	.18	.25	.25	.26

	Outcome							
	ln(Annual Wages)				ln(Weeks Worked)			
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Unconditional gap (SE)		−.513 (.027)				.016 (.014)		
Black-white gap	−.467	−.265	−.116	−.137	.023	.050	.065	.054
	(.024)	(.027)	(.037)	(.038)	(.014)	(.016)	(.024)	(.024)
Contribution of school quality			−.155	.080			−.014	.092
			(.026)	(.050)			(.018)	(.043)
Contribution of educational attainment			−.196	−.184			−.027	−.020
			(.013)	(.014)			(.007)	(.010)
Contribution of interaction				−.226				−.102
				(.046)				(.040)
Age and county controls?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Years of schooling?	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School quality?	No	No	Yes	Yes	No	No	Yes	Yes
Interacted HC controls?	No	No	No	Yes	No	No	No	Yes
N	11,394	11,394	11,394	11,394	11,394	11,394	11,394	11,394
Adjusted R^2	.22	.26	.27	.28	.04	.05	.05	.06

Table 5
Returns to Human Capital in Equation (1)

	ln(Weekly Wage)	ln(Occupation Score)	ln(Annual Wage)	ln(Weeks Worked)
Marginal effect of school quality	.137	.046	.162	.025
	(.048)	(.033)	(.057)	(.043)
Marginal effect of educational attainment	.054	.040	.058	.004
	(.006)	(.004)	(.008)	(.006)

NOTE.—Authors' calculations are from 1940 IPUMS data (Ruggles et al. 2010) and annual reports of state education departments. The table contains estimated marginal effects for school quality or educational attainment, evaluated at the mean. Bootstrapped standard errors (from 1,000 replications within 10% random subsamples) are in parentheses.

Robustness checks

Caruthers and Wanamaker try to make their analysis more *robust* by addressing possible limitations in their model:

- Correlation between (lack of) educational resources during schooling and discrimination in the labor market as an adult.
  - These two processes seem a priori to very likely be correlated.
  - To strengthen their results, Caruthers and Wanamaker rerun their regressions only for individuals who had moved in the last five years (a proxy for working in a different county than where one was educated).
- Also try to address potential omitted variable bias from cognitive ability by developing a partially imputed ability measure from military (AFQT) testing.

Estimates of Black-White Labor Market Outcome Gaps, Excluding Nonmigrant Blacks

	Outcome								
	ln(Weekly Wage)			ln(Occupation Score)			ln(Annual Wages)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Black-white gap	−.411	−.428	−.192	−.360	−.354	−.276	−.415	−.439	−.06
	(.048)	(.055)	(.072)	(.038)	(.036)	(.050)	(.065)	(.072)	(.09
Age and county controls?	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Interacted HC controls?	No	No	Yes	No	No	Yes	No	No	Yes
N	8,464	8,464	8,464	8,166	8,166	8,166	8,464	8,464	8,46
Adjusted R^2	.01	.17	.26	.01	.12	.23	.00	.20	.27

NOTE.—See table 4 for estimates notes and definitions. Black males who did not migrate between counties over the period 1935–40 are e
errors are in parentheses. See further discussion in Sec. IV.B.

Concluding thoughts

- Because of the limited amount of data available for this issue, there is necessarily a lot of potential OVB that could be going on.
- Fixed region effects are likely able to account for some OVB, but many relevant variables are likely to have a differential effect within region.

Concluding thoughts ctd

- Certainly the same critique from Kim (2017) applies here too about control variables.
  - Selection of controls such as % urban, and value of manufacturing or farming sectors in an area is likely related to other local characteristics concerning labor market outcomes for black people.
  - When you can't cleanly get random variation, there is often this dilemma of needing to control for something that would otherwise be OVB, but when you do, you add other OVB because the control is not exogenous.
- Overall, I think this is a decent example of both the application of simple regression and the pitfalls of causal interpreation using it.