
HOW LARGE ARE TEACHER EFFECTS?

by

Larry V. Hedges

Northwestern University

(Presented at New York University 2/2/09)

RATIONALE

- Most of us believe that teachers have a substantial effect on student performance
 - But is this belief true?
 - If so, are differences in teacher effects large enough to be of policy significance?
-

APPROACHES TO THE STUDY OF TEACHER EFFECTS

- Education production function studies
 - Studies of variation accounted for by teachers
-

EDUCATION PRODUCTION FUNCTION STUDIES

- Examine Association Between Specific Teacher Characteristics and Student Achievement
 - Control for Student Background
 - Social Class
 - Previous Achievement
-

VARIATION IN TEACHER EFFECTS STUDIES

- Examine Variation in Student Achievement Gain Due to Differences in Teacher Effectiveness
 - Teacher Effects are Represented by Teacher Specific Dummies not Specific Characteristics
 - Change in R^2 (ΔR^2) Associated with Teacher Effects
-

Study	Sample	Outcome	Grade	ΔR^2	ΔR
Armour, et al. 1976 ¹	LA Blacks	Reading	6	0.14	0.37
Armour, et al. 1976 ¹	LA Mexican	Reading	6	0.07	0.26
Goldhaber & Brewer, 1997 ²	NELS	Math	10	0.12	0.35
Hanushek, 1971 ³	White, manual	SAT	3	0.10	0.31
Hanushek, 1971 ³	White,nonmanual	SAT	3	0.09	0.29
Hanushek, 1971 ³	Mexican, manual	SAT	3	0.09	0.30
Hanushek, 1971 ⁴	White, manual	SAT	2	0.12	0.34
Hanushek, 1971 ⁴	White,nonmanual	SAT	2	0.13	0.35
Hanushek, 1971 ⁴	Mexican, manual	SAT	2	0.12	0.35
Hanushek, 1992 ⁵	Gary, IN	Vocabulary	2-6	0.16	0.40
Hanushek, 1992 ⁵	Gary, IN	Reading	2-6	0.10	0.32
Murnane & Phillips,1981 ⁶	Mid City Blacks	Vocabulary	3	0.10	0.32
Murnane & Phillips,1981 ⁷	Mid City Blacks	Vocabulary	4	0.21	0.46
Murnane & Phillips,1981 ⁸	Mid City Blacks	Vocabulary	5	0.16	0.40
Murnane & Phillips,1981 ⁹	Mid City Blacks	Vocabulary	6	0.21	0.46
Rivkin, et al., 2001 ¹⁰	Texas		4-6	>0.01	>0.11
Rowan, et al., 2002 ¹¹	Prospects	Reading	3-6	.03-.13	.02 - .36
Rowan, et al., 2002 ¹¹	Prospects	Math	3-6	.06-.13	.24 - .36

FINDINGS

- Production Function Studies Controversial
 - Variation Studies Find Substantial Variation Accounted for by Teachers
 - Difficult to Interpret the Relation Between Teacher Characteristics and Student Achievement
-

LIMITATIONS

- Teacher Effects Can be Confounded with Unmeasured Student, Family Characteristics
 - Difficult to Make Inferences About the Causal Direction of the Association Between Teacher Characteristics and Student Achievement
-

HOW CAN THESE LIMITATIONS BE ELIMINATED?

- Randomly Assign Students to Classrooms
 - Differences Between Students in Different Classes is not Systematic (by Chance)
 - Randomly Assign Teachers to Classrooms
 - No Differential Assignment of Teachers to Classrooms
-

THE PRESENT STUDY

Tennessee Class Size Experiment

- Uses Data from a Large-Scale Randomized Experiment
 - Examines Variation in Student Achievement Between Classrooms Within Schools Using Multi-Level Models
 - Determines how Large these Differences Are
-

PROJECT STAR

- (Almost) Unique Experiment in Education
 - Random Assignment of *Both* Teachers and Students in Classrooms within Each School
 - Broad Range of Schools
-

VARIABLES

- Outcome Variables:
 - Standardized Mathematics and Reading Scores
 - Covariates:
 - Student Gender, Race, Social Class, Previous Achievement, and Treatment (Class Size)
-

ANALYSIS

- Achievement Status: Cross-Sectional Analysis for Grades K to 3
 - Achievement Gains: Cross-Sectional Analysis for Grades 1 to 3
 - Use Three Level Hierarchical Linear Models (HLM)
 - Accounts for the Nesting Structure of the Data
 - Decomposes Variation in Achievement into Three Categories: Within-Classroom, Between-Classrooms, Between-Schools
-

ANALYSIS MODEL

Level 1 (Student)

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}Pre + \beta_{2jk}Fem + \beta_{3jk}SES + \beta_{4jk}Min + \varepsilon_{ijk}$$

Level 2 (Classroom)

$$\beta_{0jk} = \pi_{00k} + \pi_{01k}Small + \pi_{02k}Aide + \xi_{0jk}$$

Level 3 (School)

$$\pi_{00k} = \gamma_{000} + \eta_{00k}$$

Variance components

$$\varepsilon \sim N(0, \sigma^2), \quad \xi \sim N(0, \tau^2), \quad \eta \sim N(0, \omega^2)$$

ANALYSIS MODEL

Level 1 (Student)

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}Pre + \beta_{2jk}Fem + \beta_{3jk}SES + \beta_{4jk}Min + \varepsilon_{ijk}$$

Level 2 (Classroom)

$$\beta_{0jk} = \pi_{00k} + \pi_{01k}Small + \pi_{02k}Aide + \xi_{0jk}$$

Level 3 (School)

$$\pi_{00k} = \gamma_{000} + \eta_{00k}$$

Variance components

$$\varepsilon \sim N(0, \sigma^2), \quad \xi \sim N(0, \tau^2), \quad \eta \sim N(0, \omega^2)$$



Teacher Effect

VALIDITY OF THE EXPERIMENT

- Was Randomization into Treatment Groups Effective?
 - Students, Teachers
 - Was Randomization within Treatment Groups Effective?
 - Students
 - Are the Results Generalizable?
-

Table 2

P-values for tests of the difference across classes on pre-assignment student characteristics

Student Characteristic	Kindergarten	Grade 1	Grade 2	Grade 3					
SES	0.99	0.94	0.98	0.98					
Ethnicity	0.83	0.85	0.87	0.81					
Age	0.23	0.09	0.25	0.10					

Table 3						
Numbers of schools in Analysis of Teacher Effects having Various Num						
	Number of classes per school					
Grade	3	(%)	4 or more	(%)	5 or more	(%)
Achievement Gains						
1	22	29.0	54	71.0	28	36.8
2	16	21.6	58	78.4	28	37.8
3	16	21.3	59	78.7	29	38.7
Achievement Status						
Kindergarten	29	36.7	50	63.3	24	30.3
1	22	28.9	54	71.1	29	38.2
2	13	17.6	61	82.4	28	37.8
3	15	20.0	60	80.0	30	40.0

Table 4. Comparison of Whole STAR Sample and Two Subsamples on Important Characteristics

	Achievement Status			Achievement Gains		
	Whole Sample	4 or more classes	5 or more classes	Whole Sample	4 or more classes	5 or more classes
Grade						
<u>Kindergarten</u>						
Percent Minority	32.7%	34.5%	42.2%			
Percent Low SES	48.1%	47.7%	54.4%			
Average Mathematics Achievement (SAT)	485.7	485.8	483.8			
Average Reading Achievement (SAT)	436.8	436.9	437.4			
Percent of Urban Schools	31.0%	31.4%	41.9%			
Percent of Suburban Schools	21.7%	23.2%	12.7%			
Percent of Rural Schools	47.3%	45.4%	45.4%			
Days Absent from School	10.3	9.9	10.0			
Percent of Teachers with Graduate Degree	35.3%	34.7%	31.2%			
Percent of Teachers with > 3 Years of Experience	80.6%	78.8%	77.4%			
N	5766	4239	2409			
<u>First</u>						
Percent Minority	33.9%	32.2%	42.3%	30.3%	28.3%	37.3%
Percent Low SES	49.2%	45.9%	51.6%	44.9%	41.4%	46.4%
Average Mathematics Achievement (SAT)	531.0	531.8	530.1	535.2	536.4	535.0
Average Reading Achievement (SAT)	520.9	522.7	519.9	527.4	529.7	528.1
Percent of Urban Schools	29.9%	26.1%	39.6%	28.7%	25.4%	36.9%
Percent of Suburban Schools	23.9%	27.3%	19.6%	19.5%	22.4%	15.7%
Percent of Rural Schools	46.3%	46.6%	40.9%	51.8%	52.2%	47.3%
Days Absent from School	7.5	7.5	7.5	7.3	7.3	7.3
Percent of Teachers with Graduate Degree	34.7%	34.2%	28.6%	35.0%	34.5%	28.3%
Percent of Teachers with > 3 Years of Experience	79.5%	80.0%	80.7%	80.9%	81.8%	83.6%
N	6377	5118	3206	4045	3193	2003

Second

Percent Minority	34.2%	32.9%	42.4%	30.6%	28.1%	38.9%
Percent Low SES	46.7%	44.5%	47.8%	42.4%	39.7%	44.7%
Average Mathematics Achievement (SAT)	581.1	581.8	579.5	584.7	586.1	583.4
Average Reading Achievement (SAT)	584.4	585.6	581.9	588.8	590.6	586.7
Percent of Urban Schools	27.9%	24.6%	31.5%	25.3%	21.0%	29.3%
Percent of Suburban Schools	25.4%	28.2%	28.8%	23.7%	25.8%	27.0%
Percent of Rural Schools	46.8%	47.2%	39.7%	51.0%	53.2%	43.7%
Percent of Teachers with Graduate Degree	36.9%	36.4%	30.3%	36.2%	35.5%	28.5%
Percent of Teachers with > 3 Years of Experience	86.0%	87.5%	86.0%	85.7%	87.1%	86.4%
N	5968	5186	2805	4525	3828	2120

Third

Percent Minority	33.6%	32.3%	40.8%	31.6%	30.1%	39.1%
Percent Low SES	47.1%	44.7%	49.1%	43.3%	40.8%	46.2%
Average Mathematics Achievement (SAT)	617.9	617.9	614.5	620.6	620.6	616.9
Average Reading Achievement (SAT)	615.6	615.9	613.5	618.0	618.5	615.7
Percent of Urban Schools	27.8%	23.2%	33.9%	26.6%	22.3%	32.5%
Percent of Suburban Schools	25.1%	29.5%	24.2%	23.5%	27.8%	23.2%
Percent of Rural Schools	47.0%	47.3%	49.4%	49.9%	50.0%	44.3%
Days Absent from School	6.7	6.7	6.6	6.5	6.5	6.4
Percent of Teachers with Graduate Degree	44.3%	43.6%	42.4%	44.3%	43.9%	42.6%
Percent of Teachers with > 3 Years of Experience	87.9%	89.3%	90.9%	87.4%	89.0%	90.7%
N	5903	5032	2809	4627	3924	2220

HOW LARGE ARE TEACHER EFFECTS?

- Mathematics: Teacher Effects are $>10\%$ of the Achievement Variation
 - Reading: Teacher Effects are $< 10\%$ of the Achievement Variation
 - Achievement Status and Gains Analyses Produced Comparable Estimates
-

Table 5. Variance Decomposition by Grade

Grade	Mathematics			Reading		
	Within Classroom	Between Classroom	Between School	Within Classroom	Between Classroom	Between School
Achievement Gains						
<u>Unconditional Model</u>						
First	0.696	0.148*	0.198*	0.746	0.092*	0.209*
Second	0.729	0.139*	0.178*	0.740	0.096*	0.163*
Third	0.739	0.123*	0.168*	0.793	0.090*	0.121*
<u>Full Model</u>						
First	0.397	0.128*	0.090*	0.439	0.066*	0.097*
Second	0.323	0.135*	0.044*	0.312	0.068*	0.026*
Third	0.312	0.123*	0.048*	0.317	0.074*	0.019*
Achievement Status						
<u>Unconditional Model</u>						
Kindergarten	0.709	0.126*	0.165*	0.724	0.114*	0.166*
First	0.698	0.131*	0.177*	0.734	0.084*	0.184*
Second	0.736	0.125*	0.169*	0.751	0.098*	0.152*
Third	0.736	0.115*	0.155*	0.800	0.088*	0.115*
<u>Full Model</u>						
Kindergarten	0.663	0.113*	0.155*	0.675	0.100*	0.142*
First	0.647	0.110*	0.106*	0.677	0.065*	0.096*
Second	0.673	0.108*	0.096*	0.696	0.078*	0.063*
Third	0.700	0.104*	0.082*	0.748	0.075*	0.041*

Note: The unconditional models include only intercepts at each level.

The full models includes student level covariates and treatment type at the classroom level. Because of slight differences in the sample that was standardized to have unit variance, the variances do not sum to unity.

* $p < 0.05$

SENSITIVITY ANALYSES

- Can Teacher Effects be Explained by Variation in Actual Class Size?
 - Can Teacher Effects be Explained by Teacher Experience and Education?
 - Analyses within Treatment Groups
-

Table 6. Three Level HLM Fixed Effects Estimates for Teacher Experience and Education

Grade	Mathematics		Reading	
	Experience	Education	Experience	Education
Achievement Gains				
First	-0.023	0.028	0.074	-0.013
Second	0.089	-0.010	0.149*	0.006
Third	0.189*	0.093*	0.058	0.045
Achievement Status				
Kindergarten	0.045	-0.022	0.015	-0.0003
First	0.022	0.041	0.069	0.004
Second	0.014	-0.043	0.142*	-0.042
Third	0.079	0.091**	0.032	0.058

* $p < 0.05$, ** $p < 0.10$

DO TEACHER EFFECTS VARY BY SCHOOL SES OR STUDENT SES?

- Teacher Effects are More Pronounced in Low SES Schools
 - No Evidence that Classroom and School Specific SES Effects Influence Teacher Effect Estimates
-

Table 7. Variance Decomposition by Grade by Low and High SES Schools: Achievement Gains

Grade	Low SES Schools			High SES Schools		
	Within Classroom	Between Classroom	Between School	Within Classroom	Between Classroom	Between School
Mathematics Achievement Gains						
<u>Unconditional Model</u>						
First	0.546	0.186*	0.113*	0.723	0.137*	0.112*
Second	0.613	0.252*	0.175*	0.765	0.083*	0.157*
Third	0.618	0.199*	0.067*	0.806	0.075*	0.089*
<u>Full Model</u>						
First	0.387	0.139*	0.120*	0.387	0.099*	0.024*
Second	0.312	0.159* ^a	0.0003	0.323	0.096* ^a	0.065*
Third	0.275	0.179* ^a	0.024	0.334	0.103* ^a	0.025*
Reading Achievement Gains						
<u>Unconditional Model</u>						
First	0.534	0.123*	0.099*	0.774	0.064*	0.139*
Second	0.569	0.158*	0.191*	0.810	0.049*	0.059*
Third	0.651	0.199*	0.033*	0.833	0.007*	0.067*
<u>Full Model</u>						
First	0.386	0.098* ^a	0.099*	0.440	0.049* ^a	0.036*
Second	0.244	0.079*	0.019*	0.322	0.049*	0.013*
Third	0.297	0.140* ^b	0.004	0.342	0.038* ^b	0.013*

* p < 0.05

^a The p value of the z statistic is < 0.10

^b The p value of the z statistic is < 0.05

Table 8.
Variance Decomposition by Grade by Low and High SES Schools: Achievement Status

Grade	Low SES Schools			High SES Schools		
	Within Classroom	Between Classroom	Between School	Within Classroom	Between Classroom	Between School
Mathematics Achievement						
<u>Unconditional Model</u>						
Kindergarten	0.651	0.176*	0.296*	0.750	0.048*	0.107*
First	0.565	0.167*	0.110*	0.728	0.104*	0.113*
Second	0.595	0.191*	0.161*	0.761	0.081*	0.138*
Third	0.599	0.183*	0.064*	0.796	0.061*	0.085*
<u>Full Model</u>						
Kindergarten	0.639	0.157* ^b	0.285*	0.704	0.051* ^b	0.095*
First	0.549	0.146* ^a	0.094*	0.680	0.077* ^a	0.115*
Second	0.578	0.159* ^b	0.116*	0.735	0.064* ^b	0.145*
Third	0.576	0.165* ^b	0.062*	0.763	0.057* ^b	0.077*
Reading Achievement						
<u>Unconditional Model</u>						
Kindergarten	0.464	0.224*	0.151*	0.903	0.033*	0.138*
First	0.521	0.111*	0.089*	0.774	0.042*	0.146*
Second	0.561	0.156*	0.129*	0.848	0.045*	0.056*
Third	0.672	0.171*	0.042*	0.839	0.015*	0.063*
<u>Full Model</u>						
Kindergarten	0.451	0.209* ^b	0.151*	0.861	0.034* ^b	0.119*
First	0.489	0.100* ^b	0.046*	0.726	0.037* ^b	0.132*
Second	0.534	0.109* ^b	0.111*	0.804	0.036* ^b	0.062*
Third	0.644	0.154* ^b	0.025	0.792	0.012* ^b	0.055*

* $p < 0.05$

^a The p value of the z statistic is < 0.10

^b The p value of the z statistic is < 0.05

Table 9. Variance Components Indicating Variation of SES Effect Between Classrooms
(Within Treatment Types within Schools) and Between Schools

Grade	Mathematics			Reading		
	SES Effect	Teacher or School Effect (SES fixed)	Teacher or School Effect (SES random)	SES Effect	Teacher or School Effect (SES fixed)	Teacher or School Effect (SES random)
Achievement Gains						
<u>First</u>						
Between Classrooms	0.020*	0.128*	0.127*	0.004*	0.066*	0.066*
Between Schools	0.015*	0.090*	0.089*	0.019*	0.097*	0.095*
<u>Second</u>						
Between Classrooms	0.008*	0.135*	0.134*	0.0005*	0.068*	0.069*
Between Schools	0.0002	0.044*	0.044*	0.009*	0.026*	0.027*
<u>Third</u>						
Between Classrooms	0.018*	0.123*	0.121*	0.005*	0.074*	0.075*
Between Schools	0.00006	0.048*	0.048*	0.003	0.019*	0.020*
Achievement Status						
<u>Kindergarten</u>						
Between Classrooms	0.009*	0.113*	0.113*	0.005*	0.100*	0.104*
Between Schools	0.017	0.155*	0.154*	0.021	0.142*	0.137*
<u>First</u>						
Between Classrooms	0.028*	0.110*	0.107*	0.034*	0.065*	0.062*
Between Schools	0.042*	0.106*	0.108*	0.052*	0.096*	0.098*
<u>Second</u>						
Between Classrooms	0.001*	0.108*	0.109*	0.002*	0.078*	0.078*
Between Schools	0.003	0.096*	0.097*	0.010*	0.063*	0.068*
<u>Third</u>						
Between Classrooms	0.039*	0.104*	0.099*	0.015*	0.075*	0.068*
Between Schools	0.009	0.082*	0.082*	0.010	0.041*	0.042*

* $p < 0.05$

RESULTS & CONCLUSION

- Teacher Effects are Large Enough to be Important
 - Teacher Effects are Typically Larger than School Effects
 - Teacher Effects are Larger in Mathematics than in Reading
 - Teacher Effects are Larger in Low SES Schools than in High SES Schools
-

WE MAY HAVE *UNDERESTIMATED* TEACHER EFFECTS

- We Looked only at *Within-School* Variation of Teacher Effects
 - Teacher Effects May also Vary Between Schools
 - Achievement Measures May Not be Closely Aligned with Intention of Instruction
-

LIMITATIONS

- This is a Poor Design for Estimating the Relation Between Teacher *Characteristics* and Teacher Effects
 - Teacher Effects May be Cumulative
-

FOLLOWUP WORK

(BY SPYROS KONSTANTOPOULOS)

- Teacher Effects on Achievement Status
Appear to Persist for up to 3 Years
 - Teacher Effects May Impact Achievement
Gains for Several Years Too
-