STUDENTS' CHEATING AS A SOCIAL INTERACTION

EVIDENCE FROM A RANDOMIZED EXPERIMENT IN A NATIONAL EVALUATION PROGRAM

Claudio Lucifora
(Catholic University Milan & IZA)

Marco Tonello
(Bank of Italy & Defap)

February 8, 2013
INVALSI Workshop, Rome

Introduction and motivation

The growing evidence of cheating in test scores

- A variety of recent studies show that students often violate standards of academic integrity and that the problem has worsened over the last decade hand in hand with the introduction of high-stake testing systems (Davies et al. 2009; Rimer 2003):
  - Colleges: 21% of undergraduates admitted to have cheated on exams at least once a year (McCabe, 2005).
  - High schools: 59.3% of the U.S. students interviewed cheated at least once during a test (Josephson Institute of Ethics, 2011).
- Large-scale cheating has been uncovered over the last year at some of the US most competitive schools:
  - Stuyvesant High School, Air Force Academy and Harvard.
- Students' cheating has detrimental consequences:
  - it contaminates information used in educational decisions (e.g. awarding a diploma, tracking);
  - it detracts from the signaling validity of education titles on the labor market.

Students' cheating as a social interaction

The social mechanism

D. K., 17, student at Stuyvesant High School, Manhattan:

-It's seen as helping your friend out. If you ask people, they'd say it's not cheating. I have your back, you have mine-

Student at Harvard Graduate School of Education:

-We want to be famous and successful, we think our colleagues are cutting corners, we'll be damned if we'll lose out to them, and some day, when we've made it, we'll be role models. But until then, give us a pass-

(The New York Times, September 25th, 2012)

Students' cheating as a social interaction

The social multiplier

- We define students' cheating when a student breaks an ethical code of behavior exchanging information, cooperating with other students or using any prohibited materials during an exam (Cizek, 2003).
- Even an isolate cheating behavior may propagate and become larger through social interactions and originate the students cheating social multiplier as other students who might otherwise have behaved honestly - end up being influenced thus reacting to such behavior (Carrel et al. 2009).
  - the aggregate outcome (class average test score) depends on a direct effect (a reaction via private incentives to cheat) and an indirect effect (a reaction to the cheating behavior of others);
  - the social multiplier (γ) is the ratio between the equilibrium aggregate response and the sum of the reactions of individuals to cheating (Glaeser et al. 2003).

Literature and contributions

Educational psychology literature

- Stanard and Bowers (1970); Jordan (2001); Grimes and Rezek (2005)

Economic literature

- Unintended consequences of testing programs in education:
  - teachers cheating (Jacob and Levitt, 2003; Jacob, 2005); plagiarism (Dee and Jacob, 2012); Bertoni et al. 2012.
- Endogenous peer effects and social multipliers:
  - identification of endogenous interactions (De Giorgi et al. 2010; Bramoulle et al. 2009; Calvo-Armengol et al., 2009)
  - Carrel et al. 2009; Glaeser et al. 1996, 2003; Drago and Galbiati, 2012; Maurin and Moschion, 2009.
- Methodological approach:
  - structural estimations of peer effects parameters (Cooley, 2010)
  - Excess-Variance approach (Graham, 2008; Galbiati and Zanella, 2012)

Outline

- Theoretical framework
  - Students' effort maximization problem
  - Derivation of the cheating social multiplier
- Institutional setting
  - the natural experiment in SNV survey
- Data and descriptive statistics
- Empirical strategy
  - the E-V approach and identification assumption
- Results
  - baseline results
  - heterogeneous effects: social ties
- Robustess checks
- Conclusions and policy implications

Theoretical framework

A model of student's behavior

- Student's maximization problem (BrockDurlauf,2001; Cooley,2010):

$$max_{e_i} \;\; U_i = \beta_y y_i - \frac{1}{2}\beta_e e_i^2 + \tilde{\beta}_e e_i e_j - \bar{c}(\mu) \tag{1}$$

$$s.t. \;\; y_i = x_i^{'}\pi_x + x_j^{'}\tilde{\pi}_x + e_i + \tilde{\pi}_e e_j + \mu$$

- FOC: effort best response

$$e_i^{BR} = \frac{\beta_y}{\beta_e} + \frac{\tilde{\beta}_e}{\beta_e} e_j \tag{2}$$

Theoretical framework

Achievement best response

- Assuming achievement monotonically increasing in cooperative effort behavior (e_i), the effort BR can be mapped into an achievement BR which is observable to the researcher:

$$y_i^{BR} = \delta_0 + x_i^{'}\delta_x + x_j^{'}\tilde{\delta}_x + Jy_j + \mu^{'}\tilde{\delta}_\mu \quad (3)$$

- Given the linear-in-parameters form of the achievement BR, it can be shown that a unique Nash equilibrium exists $(y_i^*; y_j^*)$ so that eq. (3) can be rewritten as:

$$y_i^* = \delta_0 + x_i^{'}\delta_x + x_j^{'}\tilde{\delta}_x + Jy_j^* + \mu^{'}\tilde{\delta}_\mu \quad (4)$$

- The parameter J corresponds to the *unobserved endogenous social effects* (Manski, 1993) and it is a measure of the endogenously determined effect of individual behavior on the reference group average behavior:

$$J = \frac{\tilde{\beta}_e + \beta_e\tilde{\pi}_e}{\beta_e + \tilde{\beta}_e\tilde{\pi}_e} \quad (5)$$

Theoretical framework

The cheating social multiplier (CSM)

- Rearranging eq. (5) w.l.g. using class averages:

$$y^*_{ic} = \delta_0 + x_i^{'}\delta_x + \bar{x}_c^{'}\tilde{\delta}_x + J\bar{y}^*_c + \mu^{'}\tilde{\delta}_\mu \quad (6)$$

- Averaging within the reference group (i.e. the classroom) and solving for $\bar{y}$ yields:

$$\bar{y}_c = \gamma\delta_0 + \bar{x}_c^{'}\gamma(\delta_x + \tilde{\delta}_x) + \mu^{'}\gamma\tilde{\delta}_\mu \quad (7)$$

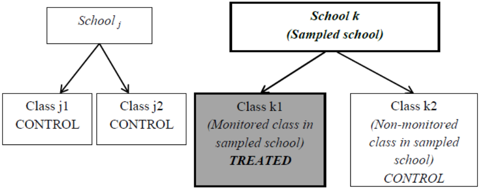
$$y^*_i = \gamma\delta_0 + x_i^{'}\delta_x + \gamma\bar{x}_c^{'}\tilde{\delta}_x + (\gamma - 1)\bar{x}_c^{'}\delta_x + \gamma\mu^{'}\tilde{\delta}_\mu \quad (8)$$

- where $\gamma = \frac{1}{1-J}$ represents the CSM in students' cooperative efforts when, during the exam, cheating can occur.

Institutional setting

The natural experiment in Invalsi SNV data

- The Protocol for the SNV survey entails the use of external inspectors for the administration of the tests, in a representative and random sample of classrooms.
- Inspectors' tasks:
  - invigilate students during the tests
  - provide specific information on the test administration
  - compute scores and send results and documentation to Invalsi within a couple of days.
- It introduces a random treatment with respect to the possibility of students to interact exchanging information or cooperating during the test.



Data and descriptive statistics

SNV survey s.y. 2009-10

- 6^{th} grade students (junior high schools), census;
- Invalsi excludes the possibility of cheating in monitored classes also computing a statistical index of cheating, which, in turn, highlights severe cheating in non-monitored classes.

% Sampled schools	22.47
% Monitored classes	7.78
% Monitored students	8.01
% Non-monitored class in sampled school	13.07
% Absent students	0.71
Average school size	131.75
Average class size	20.58
Total No. Schools	5824
Total No. Classes	26707
Total No. Students	522655

Data and descriptive statistics

SNV survey s.y. 2009-10

	Monitored	Non-monitored	Delta
Female	48.3	48.34	0.04
Regular Path	91.5	91.56	0.06
Retained	7.31	7.09	-0.22
Retained and Italian	3.67	3.63	-0.04
Immigrant	10.26	9.94	-0.32**
Kindergarten Attendance	96.83	96.82	-0.01
Speak Dialect at Home	16.93	17.08	0.15
N	41,550	477,395	

	LANGUAGE		
	All Pop.	Monitored	Non-monitored
Mean	61.50	61.42	61.51
Median	62.07	62.07	63.79
Total Var.	231.44	235.35	231.10
Var. Between Classrooms	47.67	42.46	48.13
Var. Within Classrooms	183.77	192.90	182.97
	MATH		
	All Pop.	Monitored	Non-monitored
Mean	52.00	51.46	52.05
Median	50.00	47.62	50.00
Total Var.	328.42	329.35	328.94
Var. Between Classrooms	76.80	68.89	77.41
Var. Within Classrooms	252.24	260.45	251.53

Empirical strategy

The E-V approach

The E-V approach (Graham, 2008) allows a direct estimation of γ in eq. (7) by relying only on the cross-group variation in conditional variances originating from endogenous social effects. ▸ Calculations

- We observe N classrooms composed by N_c students; y_i, the outcome variable (test score); Z_c and Ψ_c, vectors containing group-level information. Individual-level (ϵ_i) and classroom-level heterogeneity (μ_c) are unobserved latent variables;
- γ^2 can be identified if there exists a "treatment variable" (Z_c) s.t. it generates an exogenous variation that affects the between-classroom variance in students' achievement only via the effect that cheating interactions have on the within-classroom variance. Then:

$$\gamma^2 = \frac{E(G_c^b|Z_c = 1) - E(G_c^b|Z_c = 0)}{E(G_c^w|Z_c = 1) - E(G_c^w|Z_c = 0)} \tag{9}$$

Empirical strategy

Identifying assumption

- Identifying assumption: $Z_c = 1$ for classrooms with external inspector (natural experiment in SNV, perfect randomization):
  - Z_c is uncorrelated with classroom characteristics so that no sorting or matching occurs;
  - Z_c determines two equal sub-groups in observable and unobservable characteristics (e.g. students' quality).
- Advantages:
  - It overcomes Manski's (1993) reflection problem: direct estimation of structural parameter
  - Robust to individual level heterogeneity: it limits confounding effects in contextual and correlated effects
  - Limited data requirement: it limits OVB in contextual effects

Results

Baseline estimations

The structural parameters identified are γ and J, eq.(6) and (7):

$$y_{ic}^{*} = \delta_0 + x_i^{'}\delta_x + \bar{x}_c^{'}\tilde{\delta}_x + J\bar{y}_c^{*} + \mu^{'}\tilde{\delta}_\mu$$

$$\bar{y}_c^{*} = \gamma[\delta_0 + \bar{x}_c^{'}(\delta_x + \tilde{\delta}_x) + \mu^{'}\tilde{\delta}_\mu]$$

- γ captures the equilibrium social effect on individual achievement (i.e. test score) due students' cheating cooperative behavior during the exam.
- E-V approach retrieves γ^2:
  - $\gamma > 0$ comes from positive complementarities assumed in theoretical model and we test the null of no social interactions: H_0: $\gamma^2 = 1$
- Additional controls:
  - Ψ_c^{NMC}: dummy for non-monitored class in sampled school;
  - Ψ_c^{Imm}: dummy for high immigrant share ($P > 90$) ▸ Controls

Results

Math and Language (SNV Invalsi, 6^{th} grade)

	MATH			
γ^2	5.135	5.136	5.390	5.401
	(0.211)	(0.211)	(0.244)	(0.240)
P-value (H_0: $\gamma^2 = 1$)	0.00	0.00	0.00	0.00
Model Parameters				
γ	2.266	2.266	2.322	2.324
	(0.047)	(0.047)	(0.053)	(0.052)
J	0.559	0.559	0.569	0.570
	(0.009)	(0.009)	(0.010)	(0.010)
1st stage F-Stat	10772.44	10772.02	6161.20	6838.82
	LANGUAGE			
γ^2	4.189	4.182	4.370	4.383
	(0.169)	(0.169)	(0.198)	(0.195)
P-value (H_0: $\gamma^2 = 1$)	0.00	0.00	0.00	0.00
Model Parameters				
γ	2.047	2.045	2.090	2.094
	(0.041)	(0.041)	(0.047)	(0.047)
J	0.511	0.511	0.522	0.522
	(0.010)	(0.010)	(0.011)	(0.011)
1st stage F-Stat	8290.73	8290.42	5172.91	5641.34
No. Classrooms	25959	25959	25959	25959
Ψ_c^{NMC}		yes		yes
Ψ_c^{Imm}			yes	yes

Results

Heterogeneous effects: Language

	LANGUAGE			
	High	Low	High	Low
	Books at home		*Outside-school activities*	
γ^2	4.0339	4.7029	3.8383	4.9494
	(0.2606)	(0.2817)	(0.2410)	(0.3025)
1st stage F-Stat	2679.83	2875.71	2881.15	2682.26
P-val H_0: $\gamma^2_H = \gamma^2_L$	0.08		0.00	
No. Classrooms	12031	12107	12040	12098
	Play with friends		*Sport practice*	
γ^2	4.4492	4.2456	3.8128	4.6742
	(0.3320)	(0.2303)	(0.2577)	(0.2737)
1st stage F-Stat	2133.98	3435.55	2244.61	3547.14
P-val H_0: $\gamma^2_H = \gamma^2_L$	0.61		0.02	
No. Classrooms	11897	12241	11903	12235
Ψ_c^{NMC}	yes	yes	yes	yes
Ψ_c^{Imm}	yes	yes	yes	yes

Results

Heterogeneous effects: Math

	MATH			
	High	Low	High	Low
	Books at home		*Outside-school activities*	
γ^2	4.9199	5.8522	4.7388	6.1082
	(0.3203)	(0.3781)	(0.2932)	(0.4092)
1st stage F-Stat	2876.05	3635.15	3270.83	3187.27
P-val H_0: $\gamma_H^2 = \gamma_L^2$	0.06		0.01	
No. Classrooms	12031	12107	12040	12098
	Play with friends		*Sport practice*	
γ^2	5.4594	5.2770	5.5002	5.0890
	(0.4135)	(0.3044)	(0.3907)	(0.3151)
1st stage F-Stat	2633.23	3811.77	2517.54	4066.95
P-val H_0: $\gamma_H^2 = \gamma_L^2$	0.72		0.41	
No. Classrooms	11897	12241	12209	12538
Ψ_c^{NMC}	yes	yes	yes	yes
Ψ_c^{Imm}	yes	yes	yes	yes

Results

Summary

- Our results imply a strong amplifying role played by social interactions within students in the classroom.
- CSM is increasing in the strength of social ties: absence of monitoring amplify cheating behaviors more the more homogeneous is the class in terms of social links (i.e. complementary inputs).
- We estimate a value for the γ ranging between 2 and 3 in all the specifications.
  - Cooperative behaviors, when a strict external monitoring is missing, may generate a change in the equilibrium of students' achievements that is twice as big as the class average achievement without behavioral interactions.
  - In terms of individual test score, our estimates imply that the marginal contribution due to cheating increases individual test score by almost a half of the class average test result.

Robustness checks

Alternative social mechanisms

- Teachers' cheating:
  - Comparison across grades based on previous studies on cheating in SNV data (Bertoni et al. 2012; Ferrer-Esteban, 2012)
    See Table
- Students anxiety while taking the test:
  - Motivational questions in SNV Students' Questionnaire (Bertoni et al. 2012)
    See Figures
- Achievement peer effects:
  - Alternative instrument, class-size (Sacerdote, 2001; Lavy et al. 2012; Imberman et al. 2012)
    See Table

Conclusion and policy implications

Tolerating cheating behavior - as it is often done in schools - is a dangerous practice, since the social multiplier amplifies the negative effects on both students' performance and on the signaling role of education in the labor market.

- ENFORCE ETHICAL CODES and COMMIT TO STRICTER POLICIES: commitment to academic integrity and sanctions to violations; where (and when) the pressure is higher more resources should be devoted to monitoring activities.
- SPREAD INSPECTORS: to benefit from indirect effects of monitoring in non-monitored classrooms of sampled schools (Bertoni et al. 2012; Ferrer-Esteban, 2012)
- BREAKING SOCIAL LINKS: cheating interactions are complementary inputs to higher social ties. Thus, reshuffling students and teachers randomly in different classrooms the same day of the exam would break social links and make cheating more difficult...

Conclusion and policy implications

Breaking the social links: expected effects...

Appendix: robustness

Teachers' cheating

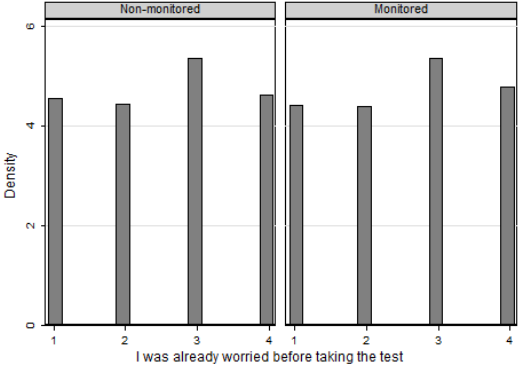
Panel A: 5^{th} **grade**				
	MATH			
γ^2	7.482	7.471	7.562	7.590
	(0.365)	(0.364)	(0.408)	(0.399)
P-value (H_0: $\gamma^2 = 1$)	0.00	0.00	0.00	0.00
F-Stat	7113.20	7112.93	4840.38	5393.57
	LANGUAGE			
γ^2	5.245	5.227	5.267	5.280
	(0.323)	(0.326)	(0.368)	(0.361)
P-value (H_0: $\gamma^2 = 1$)	0.00	0.00	0.00	0.00
F-Stat	6504.80	6504.56	4440.85	4911.43
No. Classrooms	26942	26942	26942	26942
Panel B: 2^{nd} **grade**				
	MATH			
γ^2	7.419	7.364	7.228	7.297
	(0.379)	(0.375)	(0.418)	(0.411)
P-value (H_0: $\gamma^2 = 1$)	0.00	0.00	0.00	0.00
F-Stat	7867.09	7866.80	5115.35	5648.61
	LANGUAGE			
γ^2	4.437	4.385	4.273	4.296
	(0.201)	(0.198)	(0.219)	(0.215)
P-value (H_0: $\gamma^2 = 1$)	0.00	0.00	0.00	0.00
F-Stat	8809.94	8809.61	5608.41	6218.09
No. Classrooms	26850	26850	26850	26850
Ψ_c^{NMC}		yes		yes
Ψ_c^{Imm}			yes	yes

▸ Back

Appendix: robustness

Stress induced by external inspectors

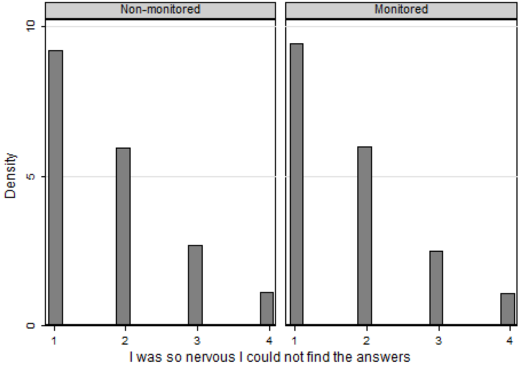
- "I was already worried before taking the test"



Appendix: robustness

Stress induced by external inspectors

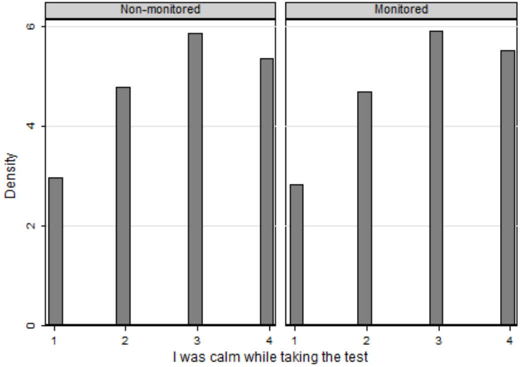
- "I was so nervous I could not find the answers"



Appendix: robustness

Stress induced by external inspectors

- "I was calm while sitting the test"



▸ Back

Appendix: robustness

Achievement peer effects: instrument class-size

	Non-monitored classrooms			
	Math		Language	
γ^2	5.9350	3.2160	5.0901	2.0785
	(0.1059)	(0.4398)	(0.0976)	(0.3375)
P-value (H_0: $\gamma_2 = 1$)	0.00	0.00	0.00	0.00
1st stage F-Stat	49602.35	3333.62	48037.98	3914.55
No. Classrooms	23901	23901	23901	23901
	Monitored classrooms			
	Math		Language	
γ^2	5.3363	1.7149	4.2599	1.5793
	(0.2956)	(1.2544)	(0.2333)	(0.6559)
P-value (H_0: $\gamma_2 = 1$)	0.01	0.57	0.00	0.38
1st stage F-Stat	5457.98	257.31	4767.00	309.41
No. Classrooms	2058	2058	2058	2058
Ψ_c^{NMC} and Ψ_c^{Imm}	yes	yes	yes	yes
School-district FE		yes		yes

▸ Back

Empirical strategy

The behavioral equations

Eq. (7) and (8) can be rewritten as behavioral equations:

$$y_{ci} = \epsilon_i + (\gamma - 1)\epsilon_c + \gamma\mu_c \tag{10}$$

$$\bar{y}_c = \gamma\epsilon_c + \gamma\mu_c \tag{11}$$

where: $\mu_c = \tilde{\delta}_x \bar{x}_c + \delta_\mu \mu + \delta_0$ is class-level heterogeneity; $\epsilon_i = \delta_x x_i$ is individual-level heterogeneity; $\epsilon_c = \delta_x \bar{x}_c$ is the class level average of individual heterogeneity.

- γ captures the equilibrium social effect on individual achievement (i.e. test score) due students' cheating cooperative behavior during the exam.
- Eq. (11) shows that γ is related to both ϵ_c and μ_c such that an exogenous shocks to contextual factors can also contribute (feeding-back through individual behaviors) to amplify the effects social externalities.

Back

Empirical strategy

The E-V approach (Graham, 2008)

Define the following simplified notation:

- $\sigma^2_\epsilon(Z_c, \Psi_c) = \sigma^2_\epsilon$ cond. var. of individual heterogeneity
- $\sigma_{\epsilon\epsilon}(Z_c, \Psi_c) = \sigma_{\epsilon\epsilon}$ cond. covar. across individual heterogeneity
- $\sigma^2_\mu(Z_c, \Psi_c) = \sigma^2_\mu$ cond. var. of group-level heterogeneity
- $\sigma_{\mu\epsilon}(Z_c, \Psi_c) = \sigma_{\mu\epsilon}$ con. covar. of group-level heterogeneity with individual heterogeneity
- $V^w_c(Z_c, \Psi_c) = V^w_c$ and $V^b_c(Z_c, \Psi_c) = V^b_c$ are within-group and between-groups cond. var.

Assuming that σ^2_μ, $\sigma_{\epsilon\epsilon}$ and $\sigma_{\mu\epsilon}$ are independent of Z_c, Graham (2008) shows that V^w_c and V^b_c can be written as:

$$V^w_c = E[\frac{\sigma^2_\epsilon(Z_c, \Psi_c) - \sigma_{\epsilon\epsilon}(\Psi_c)}{N_c}|Z_c, \Psi_c] \tag{12}$$

$$V^b_c = \gamma^2[\sigma^2_\mu(\Psi_c) + 2\sigma_{\mu\epsilon}(\Psi_c) + \sigma_{\epsilon\epsilon}(\Psi_c) + V^w_c] \tag{13}$$

Empirical strategy

The E-V approach (Graham, 2008)

Assuming:

$$\gamma^2[\sigma^2_\mu(\Psi_c) + 2\sigma_{\mu\epsilon}(\Psi_c) + \sigma_{\epsilon\epsilon}(\Psi_c)] = \theta\Psi_c \quad (14)$$

and re-writing the con. var. as conditional expectations of the correspondent statistics G^w_c and G^b_c:

$$V^w_c = E(G^w_c|Z_c, \Psi_c) \quad (15)$$

$$V^b_c = E(G^b_c|Z_c, \Psi_c) \quad (16)$$

Using eq. (14), eq. (13) becomes:

$$E(G^b_c|Z_c, \Psi_c) = \theta\Psi_c + \gamma^2[E(G^w_c|Z_c, \Psi_c)] \quad (17)$$

which determines the following conditional and unconditional moment restrictions:

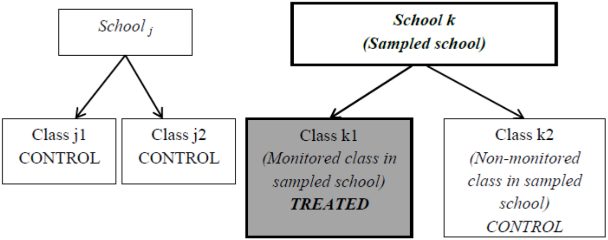
$$E(G^b_c - \theta\Psi_c - \gamma_2 G^w_c|Z_c, \Psi_c) = 0 \quad (18)$$

$$E[\begin{pmatrix} Z_c \\ \Psi_c \end{pmatrix} (G^b_c - \theta\Psi_c - \gamma_2 G^w_c)\Bigg] = 0 \quad (19)$$

Back

Additional controls

Dummy for non-monitored classes in sampled schools (Ψ_c^{NMC})



Back

Additional controls

Dummy for high immigrant share (Ψ_c^{Imm})

	Monitored	Non-monitored	Delta
Female	48.3	48.34	0.04
Regular Path	91.5	91.56	0.06
Retained	7.31	7.09	-0.22
Retained and Italian	3.67	3.63	-0.04
Immigrant	10.26	9.94	-0.32**
Kindergarten Attendance	96.83	96.82	-0.01
Speak Dialect at Home	16.93	17.08	0.15
N	41,550	477,395	

Back