



Looking Beyond Test Scores: The Effect of State Accountability Programs on Educational & Labor Market Outcomes

Kathleen Wong
University of California, Irvine

Research Question

- How did state accountability programs affect educational and labor market outcomes?
- Outline
 - Institutional background
 - Literature review
 - Econometric models
 - Data
 - Results
 - Policy Implications/Conclusion

Accountability Programs

- State-level programs
- First introduced in 1993 by 3 states
- Federal mandate in 2001: No Child Left Behind (NCLB)
 - Signed in 2002
- All 50 states & D.C. adopted by 2003



Accountability Programs II

- Goal: place responsibility on schools to ensure all students proficiently educated
- State standardized testing to meet Academic Yearly Progress (AYP) goal
- Consequences
 - Rewards: additional school funding
 - Punishment: loss of school funding or school autonomy

Literature

- State or district level analysis: no consensus
 - Improvement in test scores
 - No effect/negative impact on students
- Test scores contain measurement error:
 - Gaming the system, cheating
 - Figlio & Getzler (2002)
 - Cullen & Reback (2002)
 - Mix of signal and noise (Kane & Staiger, 2002)
 - Mean reversion (Chay et al, 2005)

Literature II

- School outcomes, not individual
- Do accountability programs lead to improvements in long-run outcomes?
- Educational:
 - Highest Grade Completed (HGC)
 - High School Graduation
 - College Enrollment
- Labor Market:
 - Employment Status
 - Log Earnings
 - Log Hourly Wage

Summary of Results

- Initially, results indicate TE led to large gains in education outcomes, log earnings
- After controlling for trends and fixed effects, more modest results:
 - White males: ↑ schooling by 0.10 years
 - White females: ↑ 0.10 years of schooling, 0.8 pp ↑ in HS graduation and 1.2 pp ↑ in college enrollment
- Variation in state program rules leads to heterogeneous treatment effects

Econometric Specification: Average Treatment Effect

$$outcome_{ist} = \alpha_0 + \alpha_1 Account_i + X'_{ist} \alpha_2 + SQ'_{ist} \alpha_3 + \alpha_4 Adopt_{s'} + \gamma_s + \gamma_t + \gamma_{st} + \varepsilon_{ist}$$

- Basic regression, by race and by sex
- Build up model by adding additional explanatory variables, fixed effects
 - Account: treatment variable (state of birth)
 - X: age, age², married, family size, children
 - SQ: per-pupil spending, ethnic diversity
 - Adopt: year that state adopted accountability

Econometric Specification: Heterogeneous Treatment Effects

- Length of accountability exposure
 - Ranges from 1 to 12 years
- Cross-state (Fletcher & Raymond, 2002)
 - Number of grade levels assessed between 1st grade and 12th grade
- Intensity of testing
 - Minimum: < 5 times → Low intensity
 - Better: Btw 5 and 8 times → Med intensity
 - Best: 9+ times → High intensity

Data

- Scope of study: 1990 - 2006
 - 1990, 2000: 5% Census sample
 - 2001-2006: American Community Survey
- Individual-level data: state of birth, current state, demographics, educational attainment and labor-market outcomes
- US-born individuals between 18-30
 - White: 351,163 men, 405,921 women
 - Black: 82,076 men, 111,540 women

Results: Basic Regression

Table 5: Basic Regression

	Males					
	Highest Grade Completed	High School Diploma	College Enrollment	Employment Status	Log Earnings	Log Hourly Wages
Treat Effect for Whites	0.125** (0.026)	0.038** (0.005)	0.043** (0.007)	0.000 (0.003)	0.187** (0.023)	0.000 (0.000)
Treat Effect for Blacks	0.173** (0.044)	0.046** (0.007)	0.026** (0.008)	-0.008 (0.007)	0.131** (0.031)	0.002 (0.001)
Observations (Whites)	351,163	351,163	351,163	174,107	164,360	164,360
Observations (Blacks)	82,076	82,076	82,076	40,013	35,627	35,627
	Females					
	Highest Grade Completed	High School Graduation	College Enrollment	Employment Status	Log Earnings	Log Hourly Wages
Treat Effect for Whites	0.247** (0.029)	0.046** (0.005)	0.068** (0.008)	0.004 (0.006)	0.228** (0.020)	0.000 (0.000)
Treat Effect for Blacks	0.214** (0.026)	0.062** (0.006)	0.058** (0.006)	0.022** (0.009)	0.133** (0.029)	0.001 (0.001)
Observations (Whites)	405,921	405,921	405,921	196,090	166,111	166,111
Observations (Blacks)	111,540	111,540	111,540	52,494	44,073	44,073
School Quality Measures	No					
Year of Adoption	No					
State and Year Dummies	No					
State by Year Dummies	No					

State by Year Fixed Effects

Table 9: State By Year Fixed Effects

	Males					
	Highest Grade Completed	High School Diploma	College Enrollment	Employment Status	Log Earnings	Log Hourly Wages
Treat Effect for Whites	0.071** (0.029)	0.004 (0.004)	0.009 (0.006)	0.002 (0.003)	0.001 (0.016)	-0.001 (0.000)
Treat Effect for Blacks	0.013 (0.039)	0.001 (0.007)	-0.008 (0.009)	0.001 (0.009)	-0.009 (0.030)	-0.001 (0.002)
Observations (Whites)	351,163	351,163	351,163	174,107	164,360	164,360
Observations (Blacks)	82,076	82,076	82,076	40,013	35,627	35,627
	Females					
	Highest Grade Completed	High School Graduation	College Enrollment	Employment Status	Log Earnings	Log Hourly Wages
Treat Effect for Whites	0.065** (0.023)	0.008** (0.003)	0.012** (0.005)	0.001 (0.005)	0.004 (0.017)	-0.001* (0.000)
Treat Effect for Blacks	0.028 (0.032)	0.003 (0.005)	0.011 (0.009)	0.005 (0.011)	-0.018 (0.023)	0.000 (0.001)
Observations (Whites)	405,921	405,921	405,921	196,090	166,111	166,111
Observations (Blacks)	111,540	111,540	111,540	52,494	44,073	44,073
School Quality Measures	Yes					
Year of Adoption	Yes					
State and Year Dummies	Yes					
State by Year Dummies	Yes					

Heterogeneous Treatment Effects

- Longer lengths of treatment → larger TE
- Intensity of testing:

Table 11: Cross-State Variation in Testing Intensity

		Males					
		Highest Grade Completed	High School Diploma	College Enrollment	Employment Status	Log Earnings	Log Hourly Wages
Low Intensity Testing							
	Treat Effect for Whites	0.057 (0.047)	0.001 (0.005)	0.004 (0.008)	0.000 (0.003)	0.026 (0.019)	-0.001* (0.000)
	Treat Effect for Blacks	0.051 (0.045)	0.007 (0.012)	0.001 (0.011)	0.000 (0.018)	-0.044 (0.045)	0.002 (0.007)
Medium Intensity Testing							
	Treat Effect for Whites	0.058 (0.041)	0.002 (0.004)	0.008 (0.007)	-0.000 (0.004)	-0.013 (0.021)	-0.000 (0.000)
	Treat Effect for Blacks	-0.009 (0.044)	-0.003 (0.007)	-0.007 (0.009)	-0.005 (0.011)	0.024 (0.034)	-0.002 (0.002)
High Intensity Testing							
	Treat Effect for Whites	0.160** (0.065)	0.016** (0.006)	0.022** (0.010)	0.021** (0.007)	-0.010 (0.034)	-0.001 (0.001)
	Treat Effect for Blacks	0.046 (0.064)	0.005 (0.008)	-0.017 (0.025)	0.014 (0.011)	-0.069 (0.042)	0.001 (0.001)
		Females					
		Highest Grade Completed	High School Graduation	College Enrollment	Employment Status	Log Earnings	Log Hourly Wages
Low Intensity Testing							
	Treat Effect for Whites	0.028 (0.047)	0.000 (0.004)	0.006 (0.006)	0.011* (0.007)	-0.009 (0.023)	-0.001 (0.001)
	Treat Effect for Blacks	0.042 (0.038)	-0.001 (0.008)	0.008 (0.013)	0.019 (0.014)	0.090 (0.062)	-0.003 (0.002)
Medium Intensity Testing							
	Treat Effect for Whites	0.059* (0.033)	0.009** (0.004)	0.011* (0.006)	-0.004 (0.005)	0.004 (0.019)	-0.001 (0.001)
	Treat Effect for Blacks	-0.001 (0.032)	0.004 (0.004)	0.011 (0.009)	0.001 (0.015)	-0.031 (0.021)	0.001 (0.001)
High Intensity Testing							
	Treat Effect for Whites	0.178** (0.045)	0.019** (0.005)	0.028** (0.007)	-0.007 (0.009)	0.048 (0.045)	-0.001** (0.001)
	Treat Effect for Blacks	0.096* (0.052)	0.003 (0.009)	0.014 (0.017)	0.002 (0.019)	-0.084* (0.043)	-0.000 (0.001)

Conclusions

- NCLB is up for renewal → successful?
 - Small average treatment effects
 - Improvements only for whites, further widening the black-white achievement gap
- Room for improvement? Maybe
 - Correlation between the strength of state program rules and the magnitude of the treatment effect