Studying the Impact of Program Participation in Multi-Site Trials Using Instrumental Variables

Stephen W. Raudenbush

Early Childhood Interventions Working Group
Inaugural Conference
University of Chicago, April, 2012

The work reported here was supported by the William T. Grant Foundation under the grant “Building Capacity for Evaluating Group-Level Interventions” and by the Institute of Education Sciences under the grant “Using Instrumental Variables Analysis Coupled with Rigorous Multi-Site Impact Studies to Study the Causal Paths by which Educational Interventions Affect Student Outcomes” (grant R305D090009.

This joint work with **Sean F Reardon** and **Takako Nomi**.

Aims

1. Define, identify, estimate
   - the average effect of program participation
   - variance of this effect across sites

2. Expand to include multiple mediators

Conception: Random Coefficient Model
- Person-specific
- Site-specific

Illustration: “Double Dose Algebra” in 60 Chicago schools, 2003

Multi-Site Trials

Pervasive in Education now

* Most IES RCTs are multi site (Spybrook and Raudenbush, 2009)
* National Head Start Experiment
* Tennessee Class Size
* Reading Recovery
* KIPP

"Planned meta-analysis"

Head Start

5000 children in 380 program sites

Program participation varies

Counter factual child care varies

Need to discovery and account for heterogeneity

Aims in Multi-Site Trials

Estimate

- the average ITT effect
- the variance of the ITT effect
- site-specific ITT effects

Similarly for Effect of Progam Participation under partial compliance

Identify Multiple Mediators

Outline of Talk

Random Coefficient Model in a Single Site

Random Coefficients Across a Sample of Sites

Expand to Multiple Mediators

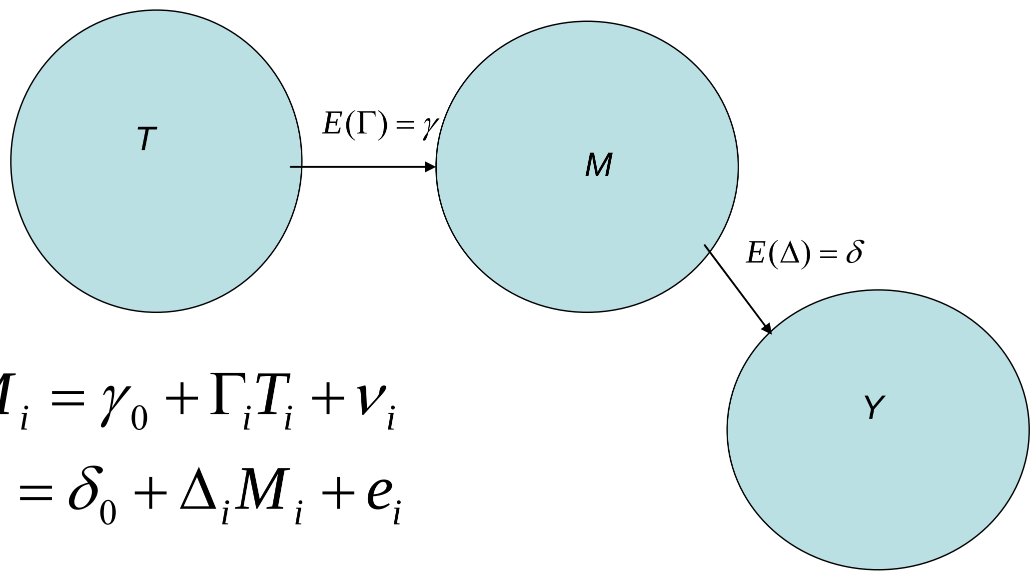
Random Coefficient Model in One Site

e.g., Heckman, J. J., & Vytlacil, E. (1998)

Generate Key Assumptions

Single site, heterogeneous treatment effect

Figure 1



$$M_i = \gamma_0 + \Gamma_i T_i + \nu_i$$
$$Y_i = \delta_0 + \Delta_i M_i + e_i$$

$$Y(T=1) - Y(T=0) \equiv B_i = \Gamma_i \Delta_i$$

Covariance between compliance and impact

1. Assume no covariance, hence δ=ATE

$$E(B) = E(\Gamma\Delta) \equiv \beta = \gamma\delta + Cov(\Gamma, \Delta)$$
$$= \gamma\delta$$

2. Assume $\Pr(\Gamma \geq 0) = 1$, $\delta = \delta_{late}$ (Angrist, Imbens, Rubin, 1996)

$$E(\mathrm{B}) = \beta = E(B \mid \Gamma = 1)\Pr(\Gamma = 1)$$
$$\equiv \delta_{late}\gamma.$$

Interpretation Problems with LATE

Unobserved Population

(unless controls cannot participate)

Instrument-Dependent Effect

Interpretation with continuous mediator

Summary of assumptions, single site case

(i) SUTVA

(ii) Exclusion restriction

(iii) No compliance-effect covariance – or – monotonicity

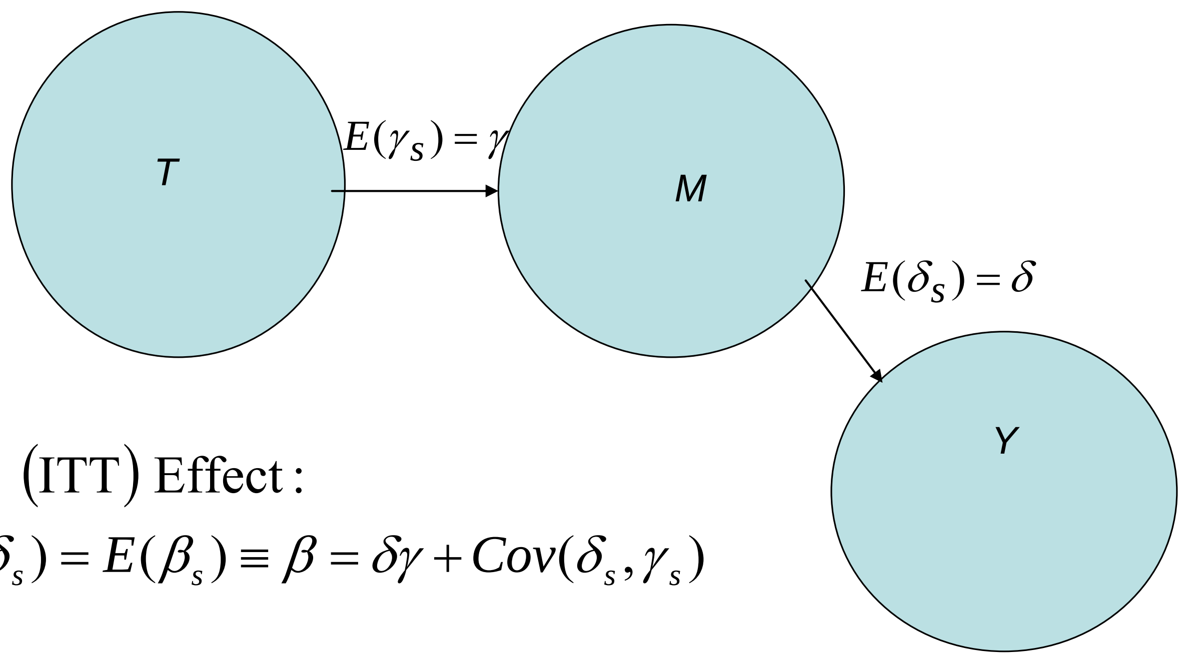
(iv) ignorable assignment of *T*; and

(v) effectiveness of the instrument

Hence $\beta / \gamma = \delta$

Multiple sites, heterogeneous treatment effect

Figure 3



Total (ITT) Effect:

$$E(\gamma_s \delta_s) = E(\beta_s) \equiv \beta = \delta\gamma + Cov(\delta_s, \gamma_s)$$

or

$$E(\gamma_s \delta_s) = E(\beta_s \mid \Gamma = 1) P(\Gamma = 1) = \delta\gamma$$

What to do?

Assume $Cov(\gamma_S, \delta_S) = 0$

(Sites with large compliance have no larger or smaller benefits than average)

Or

Monotonicity (if mediator is binary)

Variance of the site-average ITT effect

In general

$$Var(\beta_s) = (\tau_\gamma^2 + \gamma^2)\tau_\delta^2 + \delta^2\tau_\gamma^2$$

In the case of binary M

$$Var(\beta_s) = \tau_\delta^2\gamma + \delta^2\gamma(1-\gamma)$$

Summary of assumptions for the multi-site case

(i) SUTVA within each site

(ii) Exclusion restriction in each site

(iii) No compliance-effect covariance – or – monotonicity – within each site

(iv) ignorable assignment of *T* within each site,

(v) effectiveness of the instrument within each site

(vi) independence of the site-average compliance and the site-average effect of program participation – Or monotonicity.

(vii a) effectiveness of the instrument on average (Model 1)

Or

(vii a) effectiveness of the instrument somewhere (model 2)

Model 1

Random coefficient model, within each site:

$$Y_{is} - \overline{Y}_s = \beta_s (T_{is} - \overline{T})_s + \varepsilon_{is}$$

Site-specific ITT estimates vary across sites

$$\beta_s = \beta + U_s, \qquad U_s \sim (0, \tau_\beta^2)$$

$$\beta = \gamma\delta$$

$$\tau_\beta^2 = \gamma\tau_\delta^2 + \delta^2\gamma(1-\gamma)$$

Model 2

Random coefficient model, within each site:

$$Y_{is} - \overline{Y}_s = \beta_s (T_{is} - \overline{T})_s + \varepsilon_{is}$$

Site-specific ITT estimates vary across sites

$$\beta_s = \gamma_s \delta_s = \gamma_s \delta + \gamma_s (\delta_s - \delta)$$

$$Assume \ \ \delta_s \perp \gamma_s$$

$$E(\beta_s \mid \gamma_s) = \gamma_s \delta + \gamma_s E[(\delta_s - \delta) \mid \gamma_s]$$

$$= \gamma_s \delta$$

$$Var(\beta_s \mid \gamma_s) = \gamma_s^2 \tau_\delta^2$$

Fixed Effects Estimators

- Same as above, but set

$$\tau_{\delta}^{2} = \tau_{\beta}^{2} = 0$$

- Note
  - Option 1=2SLS with single instrument and site fixed effects
  - Option 2=2SLS with site-specific instruments and site fixed effects

Illustrative Example

1997 Algebra for All

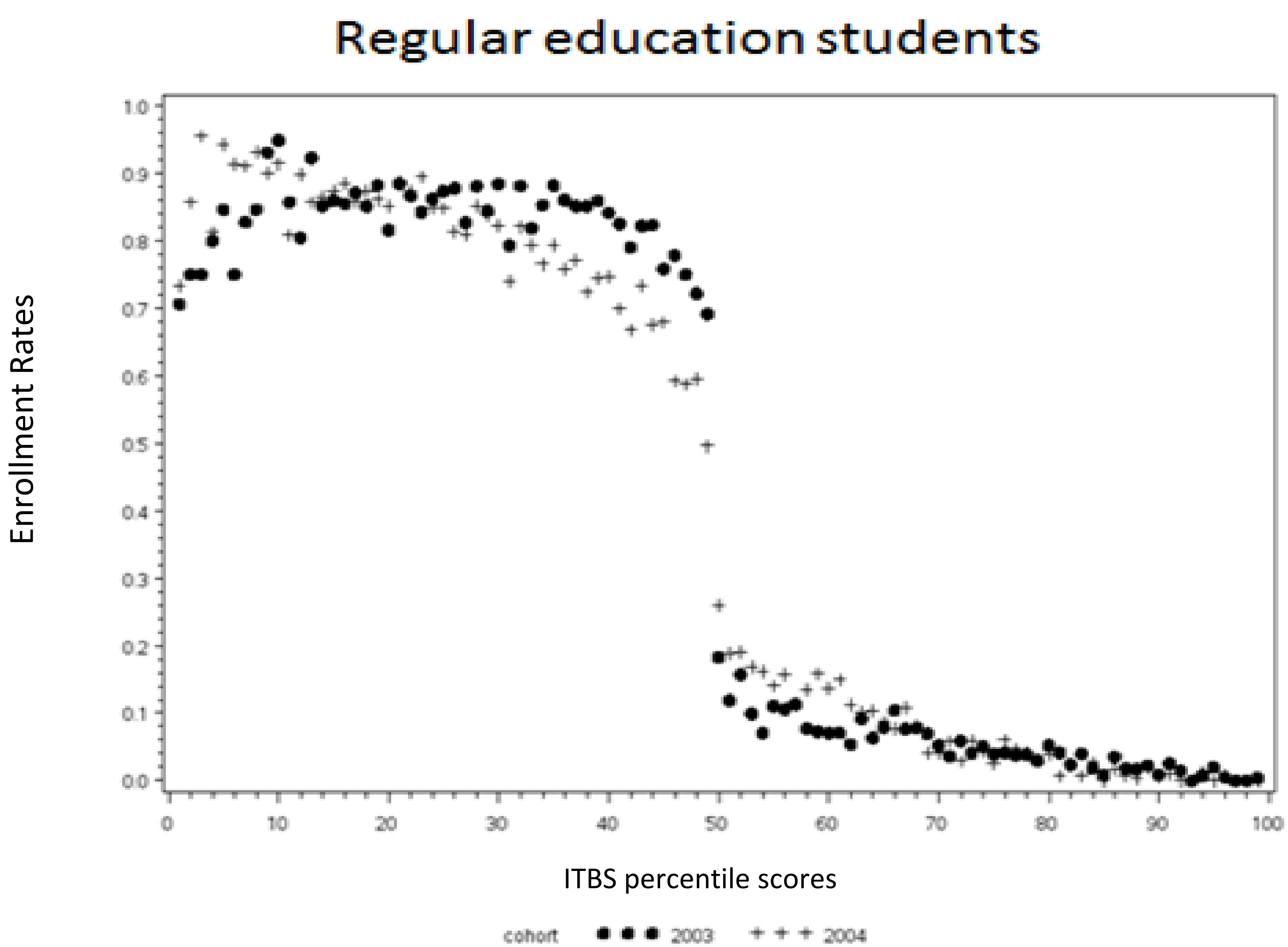
  Disappointing results

2003 Double-Dose Algebra

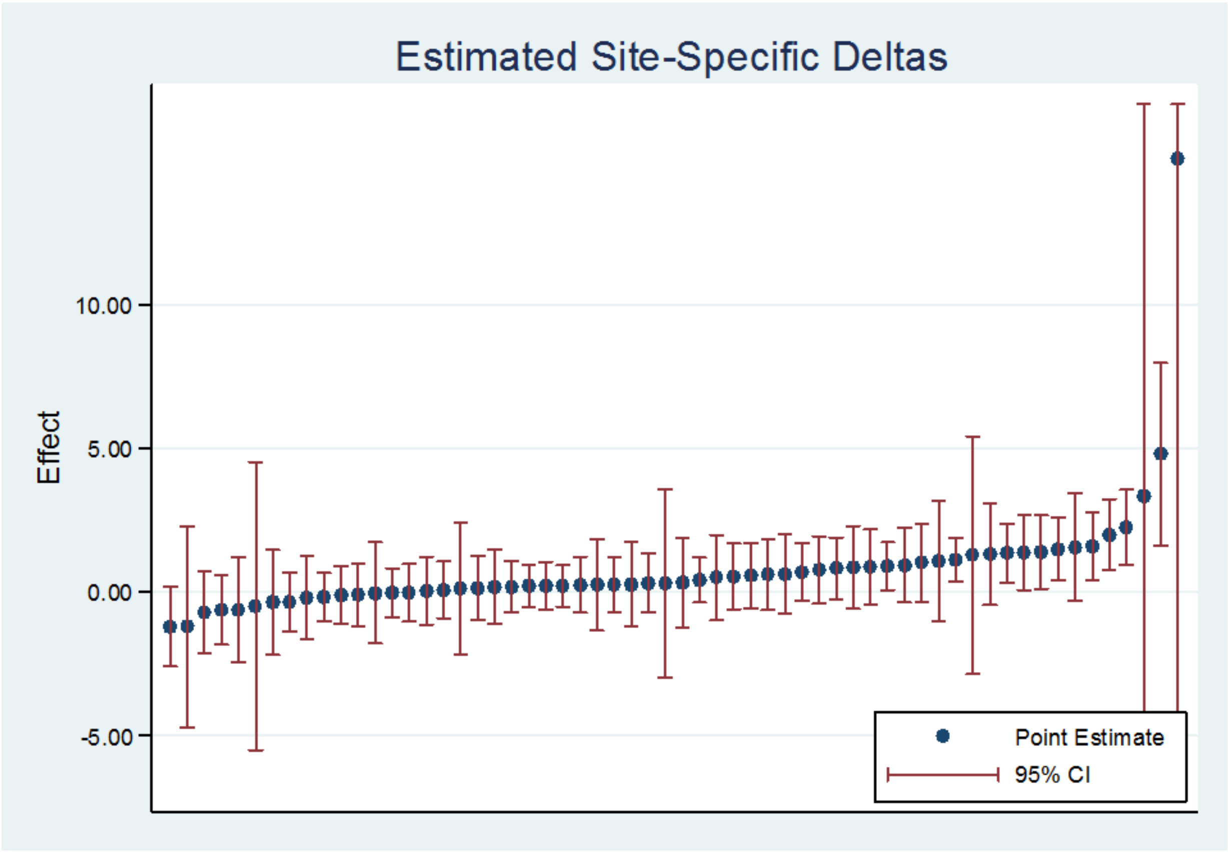
12,000 ninth grades in 60 schools

RDD

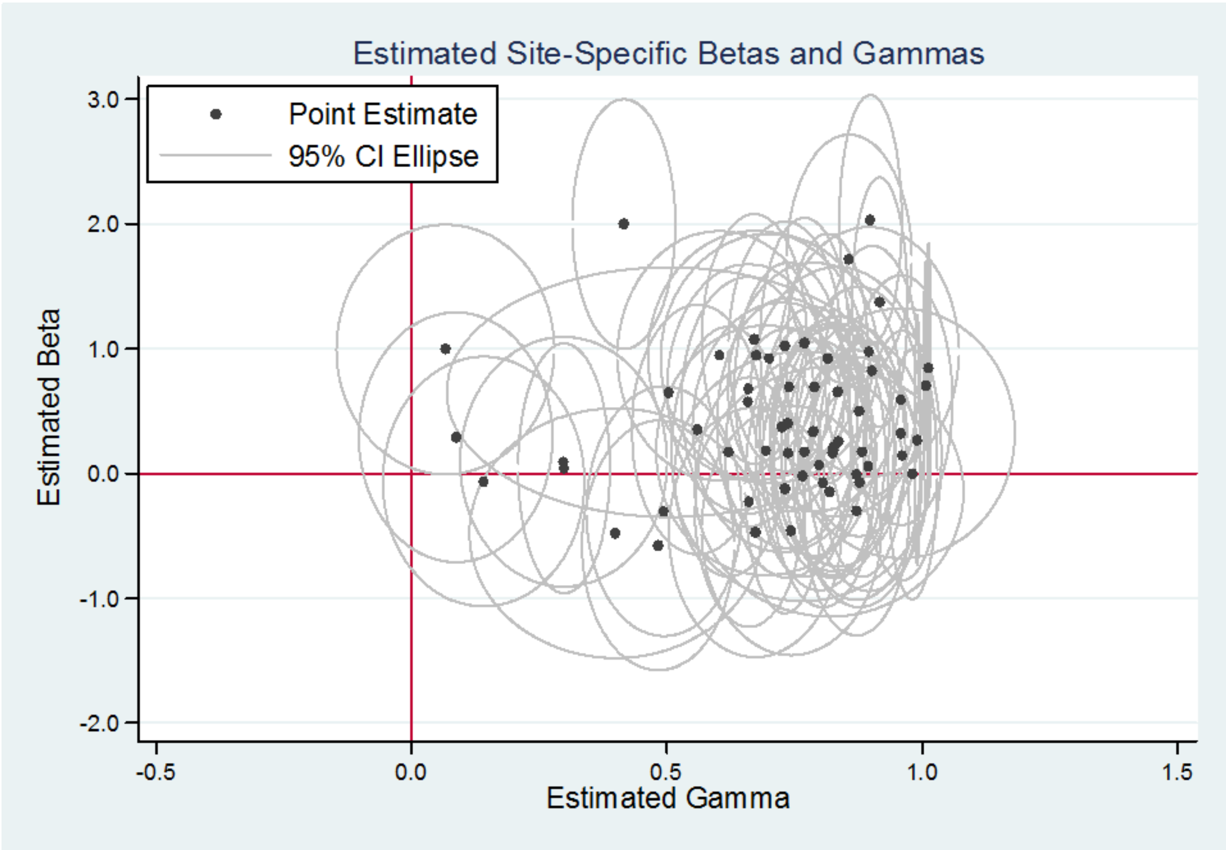
Double-dose Algebra enrollment rate by math percentile scores



Site-by-site estimates



Plot of ITT on Y by ITT on M



Summary of results

	Random Coefficient Model	Fixed Coefficient Model
Model 1	$\hat{\delta} = 0.49$ $S_{\hat{\delta}} = 0.10$ $\hat{\tau}_{\delta}^{2} = 0.26$	$\hat{\delta} = 0.54$ $S_{\hat{\delta}} = 0.09$
Model 2	$\hat{\delta} = 0.47$ $S_{\hat{\delta}} = 0.09$ $\hat{\tau}_{\delta}^{2} = 0.28$	$\hat{\delta} = 0.53$ $S_{\hat{\delta}} = 0.08$

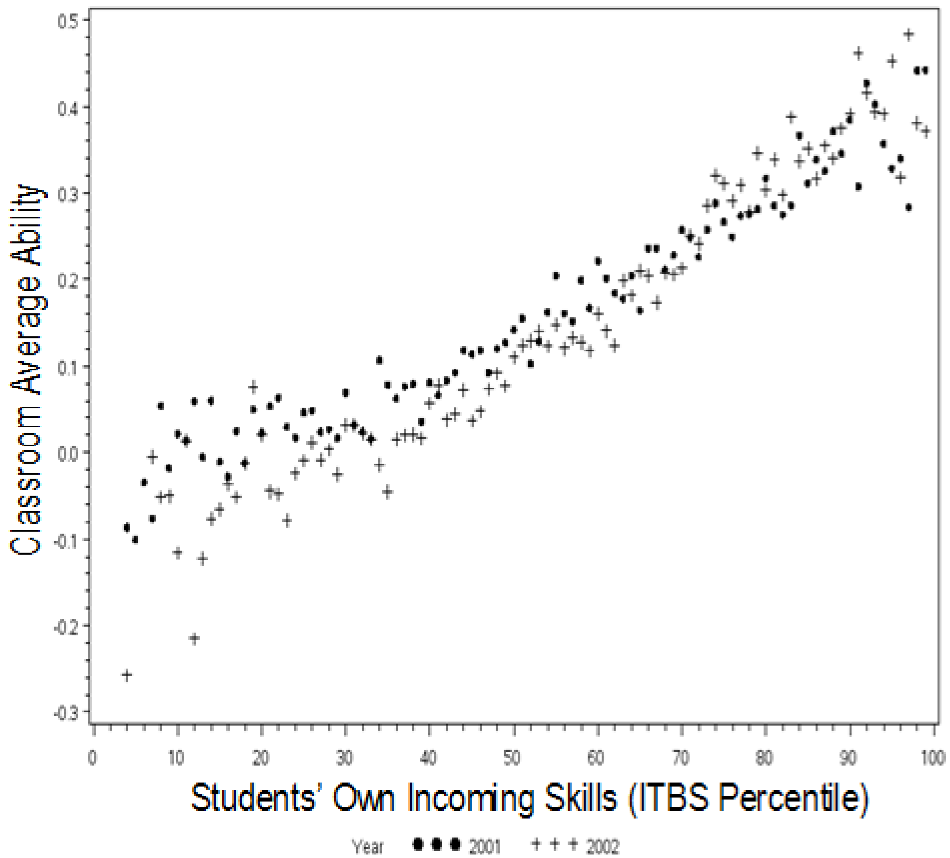
However,….

This curricular reform increased classroom segregation!

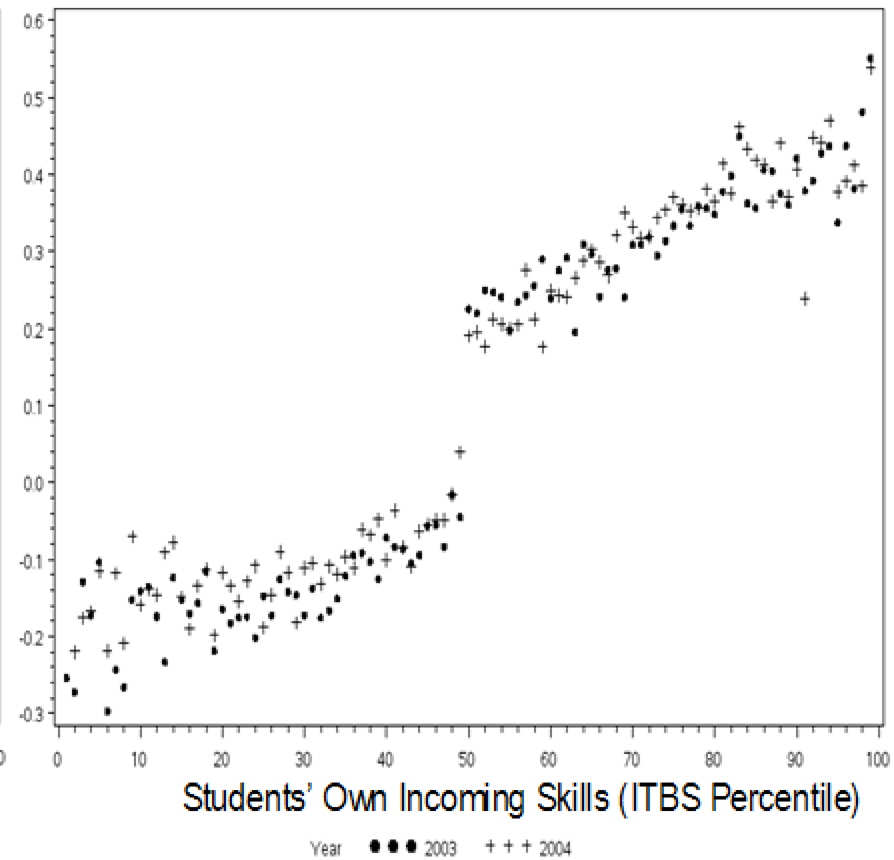
SUTVA fails

Classroom average skill levels by math percentile scores

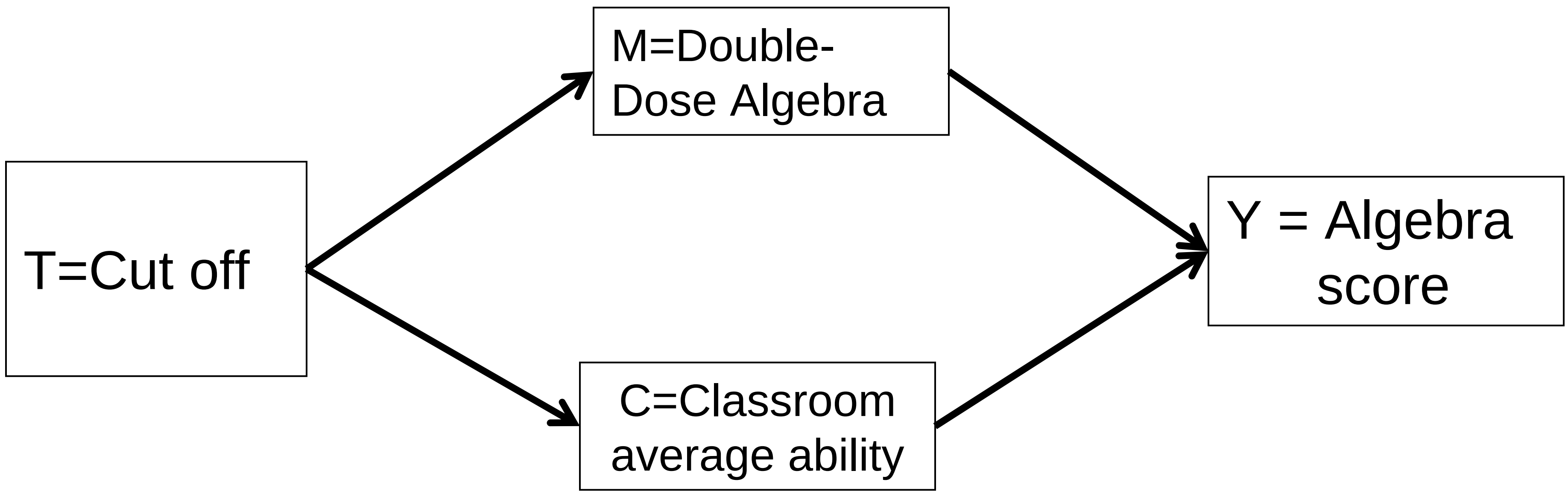
Pre-policy
(2001-02 and 2002-03 cohorts)



Post-policy
(2003-04 and 2004-05 cohorts)



Elaborated Causal Model
(T,M,C,Y model)



The cutoff scores affects student outcome only through talking double-dose algebra course and changes in peer composition

Instrumental Variable Approach

- Problem of 1 instrument, 2 mediators
- Use school-by-cut off dummies as instruments
  - Liebman, Katz, Kling, 2007
  - What are the assumptions?

Model 2 extends (Model 1 doesn't)

Random coefficient model, within each site:

$$Y_{is} - \overline{Y}_s - f(X_{is}) = \beta_s (T_{is} - \overline{T})_s + \varepsilon_{is}$$

Site-specific ITT estimates vary across sites

$$\beta_s = \gamma_{ms}\delta_m + \gamma_{cs}\delta_c + \gamma_{ms}(\delta_{ms} - \delta_m) + \gamma_{cs}(\delta_{cs} - \delta_c)$$

$$E(\beta_s \mid \gamma_{ms}, \gamma_{cs}) = \gamma_{ms}\delta_m + \gamma_{cs}\delta_c + \gamma_{ms}E[(\delta_{ms} - \delta_m) \mid \gamma_{ms}, \gamma_{cs}] + \gamma_{cs}E[(\delta_{cs} - \delta_c) \mid \gamma_{mcs}, \gamma_{cs}]$$
$$= \gamma_{ms}\delta_m + \gamma_{cs}\delta_c$$
$$Var(\beta_s \mid \gamma_{ms}, \gamma_{cs}) = \gamma_{ms}^2 \tau_{\delta_m}^2 + \gamma_{cs}^2 \tau_{\delta_c}^2 + 2\gamma_{ms}\gamma_{cs} Cov(\delta_{ms}, \delta_{cs})$$

Summary of Assumptions

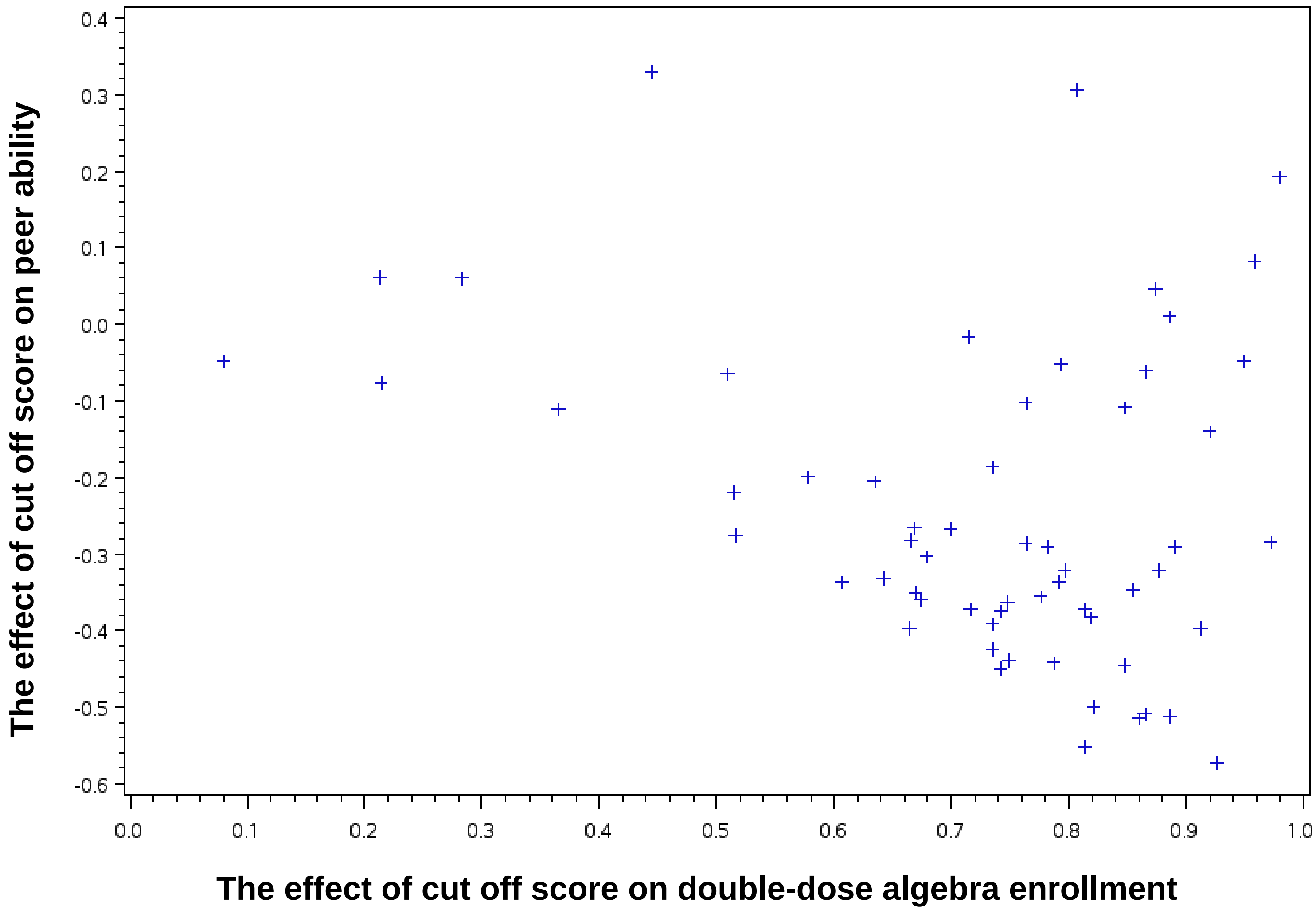
A1: Relaxed SUTVA
A2: Functional form for interference
A3: Exclusion restriction
A4: Linearity of Y in C
A5: No within-site compliance-effect covariance
A6: Between-site independence between compliance and effect
A7: Full rank design matrix
A8: Mediators operate in parallel
A9: Ignorable assignment of T given X

No-covariance assumptions most robust when

1. Compliances vary a lot
2. Compliances not too correlated
3. Compliances far from zero on average

Context specific effects:

The effects of cutoff score on double-dose algebra enrollment and peer ability



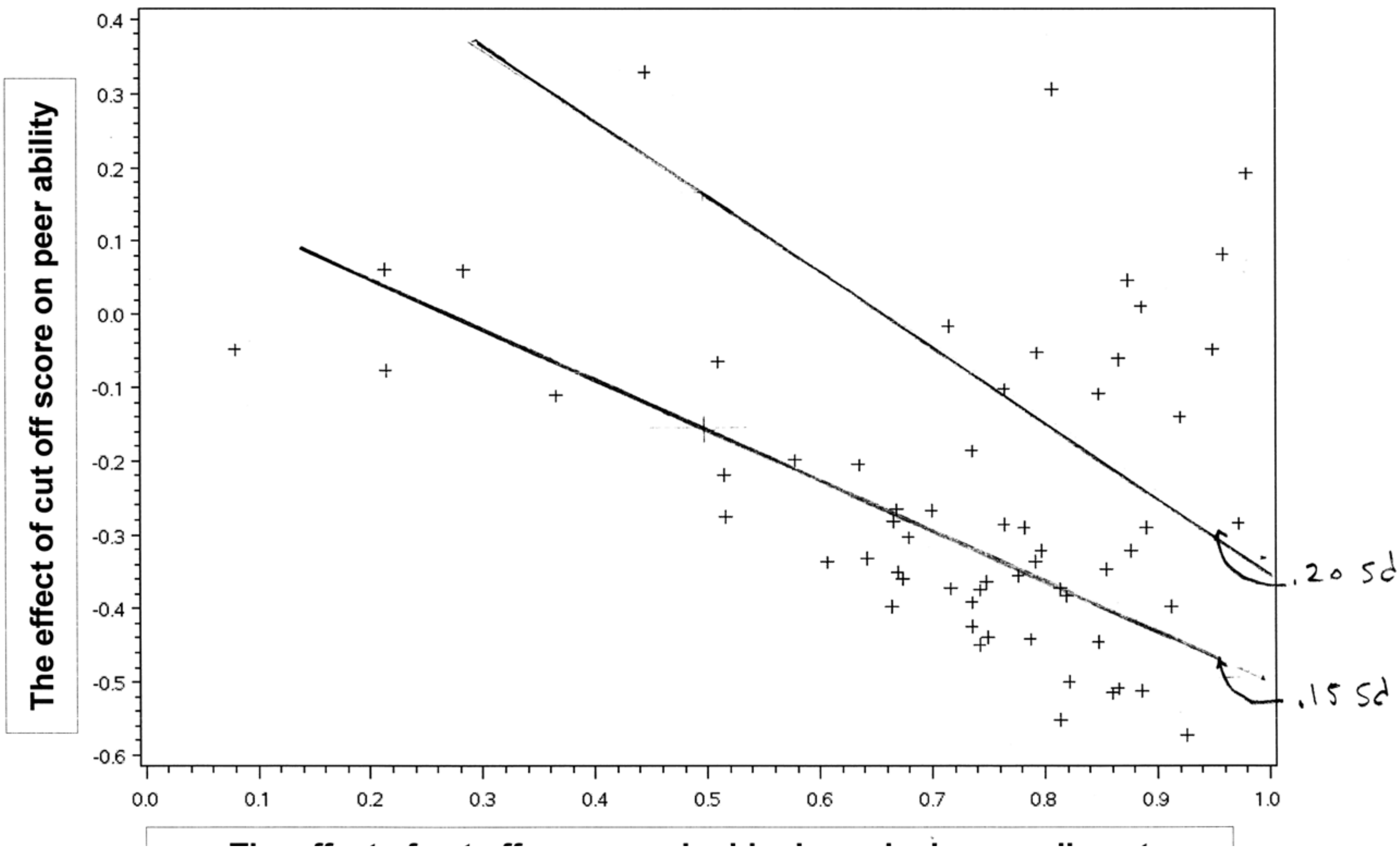
Stage 2 results:
The effect of M and C on Y

The average effect of taking double-dose algebra (M) and peer ability (C) on Algebra test scores

	Double-dose algebra enrollment	**Classroom Peer composition**
Coeff	0.69*** (d=.30)	0.81*** (d=.40)
SE	0.14	0.29

Context specific effects:

The effects of cutoff score on double-dose algebra enrollment and peer ability



But are effects heterogeneous?

- Yes
  - Both effects larger in 30 all African American Schools
  - Weaker in predominately Latino schools

Summary

The reform enhanced math instruction for low-skill students, and that helped a lot

The reform also intensified tracking and that hurt

On balance the effect was positive, but much more so in schools that implemented double dose with minimal tracking

Next Steps

Can we fruitfully apply to Head Start Experiment? (Many sites, small n per site)

Need rich modeling of mediators

Smoothing of estimates of γ_s

Correction for Bias caused by covariance between compliance and effect

Or extensions of LATE