

Desenvolvimento Econômico

Aula 3: Qualidade de educação

Ricardo Dahis

Hoje

- ▶ Qualidade de educação: Duflo et al. (2011)
 - ▶ *Tracking*
 - ▶ *Peer effects*
- ▶ Poder estatístico na prática
- ▶ Outros temas em qualidade de educação
 - ▶ *Vouchers*, redução de tamanho de turmas.

Qualidade escolar em países em desenvolvimento

- ▶ Qualidade da educação vem sendo uma área de pesquisa ativa em desenvolvimento, com RCTs sobre vários aspectos da questão.
 - ▶ Incentivos a estudantes, pais, professores.
 - ▶ Inputs escolares: livros-texto, quadros, etc
 - ▶ Informação e envolvimento de pais (*report cards*, comitês escolares, etc)
 - ▶ Pedagogia
 - ▶ Escolas públicas vs. privadas (*vouchers*)
 - ▶ “Para-professores” vs. regulares

O programa de professores extras no Kenya

- ▶ Depois da educação primária virar grátis no Kenya, turmas ficaram com tamanho enorme.
- ▶ Banco Mundial estava disposto a pagar por um(a) professor(a) extra em algumas escolas e financiar a avaliação.
- ▶ Professor(a) extra: contratado(a) por pais, menos experiência, em um contrato mais curto.
- ▶ O que aprenderíamos de um experimento onde aleatoriamente atribuímos o(a) professor(a) para algumas escolas e não outras?

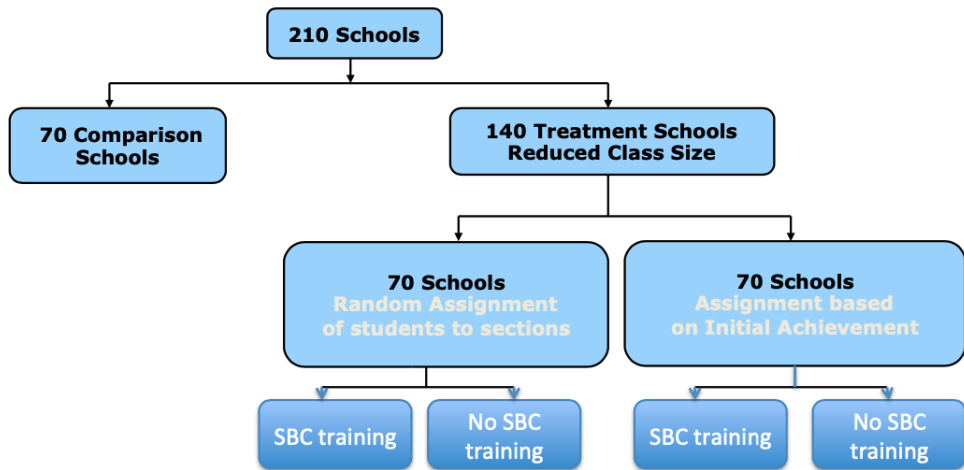
Desenhando um estudo mais ambicioso

- ▶ O que se sabia na época:
 - ▶ Efeito pequeno de tamanho de turma em países em desenvolvimento (Banerjee et al., 2007)
 - ▶ Impacto alto de educação remedial (Banerjee et al., 2007)
 - ▶ Livros-texto só efetivos para bons alunos (Glewwe et al., 2009)
 - ▶ Evidência inconclusiva sobre pais se envolvendo em projetos de escola e saúde
- ▶ Esse projeto tentou desenhar um experimento para falar com várias dessas perguntas abertas em qualidade de educação.

Perguntas de pesquisa

- ▶ Quais são os efeitos de *tracking*?
 - ▶ Como esse efeito varia com performance de estudantes ex ante?
- ▶ Quais são os efeitos diretos (aluno-aluno) e indiretos (via resposta comportamental de professores) de pares em escolas?
- ▶ (Qual é o efeito de diminuição do tamanho de turmas?)
- ▶ (Qual é o efeito de professores por contrato vs regular?)
- ▶ (O envolvimento de pais faz diferença?)
- ▶ O que tudo isso nos diz sobre o setor de educação?

Desenho experimental



Como esse desenho nos ajuda com as perguntas de pesquisa?

Controle	Seções aleatórias	Seções por performance
Professor regular	Professor Regular	Professor regular
	Professor contratado	Professor contratado

- ▶ Quais grupos deveríamos comparar para responder as várias perguntas:
 1. Qual é o efeito de redução de tamanho de turma?
 2. Qual é o efeito de professores contratados vs. regulares?
 3. Qual é o efeito de *tracking*? Como difere por performance no baseline?
 4. Quais são os efeitos diretos e indiretos de pares?

Debate sobre políticas de *tracking*

▶ Prós

- ▶ Se o ensino for mirado no meio da distribuição, aqueles no topo se beneficiariam com material mais adequado.
- ▶ Na média, distância entre todos alunos para o nível de ensino diminui.
- ▶ Metade superior se beneficia de pares melhores.
 - ▶ Se não há aprendizado de pares muito melhores que você, metade de baixo pode se beneficiar também.

▶ Contras

- ▶ Meio da distribuição tem material menos adequado à sua posição.
 - ▶ Metade inferior sofre com pares piores.
- ▶ Um grupo que perderia com certeza, por esses argumentos, é o de melhores alunos da metade inferior.

Modelo no artigo: viés de elite no Kenya

- ▶ Alunos aprendem melhor de pares. Quanto melhor o aluno, mais benefícios de pares.
- ▶ O aprendizado de alunos depende da distância entre seu nível e o alvo do professor. Se muito longe, não aprende nada.
- ▶ Professores escolhem nível de esforço e quem visar.
- ▶ Logo, o alvo dependerá da distribuição de habilidades de alunos.
- ▶ Também dependerá da função de recompensas/custos do professor.
- ▶ Com uma função convexa (prêmio maior para alunos do topo performando melhor), professores terão o topo da distribuição como alvo.

Previsões do modelo de *tracking*

- ▶ Quem ganha ou perde em termos de efeito direto de pares?
 - ▶ A metade superior ganha, a inferior perde.
- ▶ Quem ganha ou perde em termos do que professores ensinam?
 - ▶ Quase todos se beneficiam, exceto os da metade superior que estão abaixo do nível original de ensino.
- ▶ Quem ganha ou perde em termos do esforço de professores?
 - ▶ Esforço sobe na metade superior.

Resultados

TABLE 2—OVERALL EFFECT OF TRACKING

	Total score				Math score		Literacy score	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Panel A. Short-run effects (after 18 months in program)</i>								
(1) Tracking school	0.139 (0.078)*	0.176 (0.077)**	0.192 (0.093)**	0.182 (0.093)*	0.139 (0.073)*	0.156 (0.083)*	0.198 (0.108)*	0.166 (0.098)*
(2) In bottom half of initial distribution × tracking school			−0.036 (0.07)		0.04 (0.07)		−0.091 (0.08)	
(3) In bottom quarter × tracking school				−0.045 (0.08)		0.012 (0.09)		−0.083 (0.08)
(4) In second-to-bottom quarter × tracking school				−0.013 (0.07)		0.026 (0.08)		−0.042 (0.07)
(5) In top quarter × tracking school				0.027 (0.08)		−0.026 (0.07)		0.065 (0.08)
(6) Assigned to contract teacher		0.181 (0.038)***	0.18 (0.038)***	0.18 (0.038)***	0.16 (0.038)***	0.161 (0.037)***	0.16 (0.038)***	0.16 (0.038)***
Individual controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,795	5,279	5,279	5,279	5,280	5,280	5,280	5,280
<i>Total effects on bottom half and bottom quarter</i>								
Coeff (Row 1) + Coeff (Row 2)			0.156		0.179		0.107	
Coeff (Row 1) + Coeff (Row 3)				0.137		0.168		0.083
F-test: total effect = 0			4.40	2.843	5.97	3.949	2.37	1.411
p-value (total effect for bottom = 0)			0.038	0.095	0.016	0.049	0.127	0.237
p-value (effect for top quarter = effect for bottom quarter)				0.507		0.701		0.209

Resultados

Panel B. Longer-run effects (a year after program ended)

(1) Tracking school	0.163 (0.069)**	0.178 (0.073)**	0.216 (0.079)***	0.235 (0.088)***	0.143 (0.064)**	0.168 (0.075)**	0.231 (0.089)**	0.241 (0.096)**
(2) In bottom half of initial distribution \times tracking school			-0.081 (0.06)		-0.027 (0.06)		-0.106 (0.06)	
(3) In bottom quarter \times tracking school				-0.117 (0.09)		-0.042 (0.10)		-0.152 (0.085)*
(4) In second-to-bottom quarter \times tracking school				-0.096 (0.07)		-0.073 (0.07)		-0.091 (0.07)
(5) In top quarter \times tracking school				-0.028 (0.07)		-0.04 (0.06)		-0.011 (0.08)
(6) Assigned to contract teacher		0.094 (0.032)***	0.094 (0.032)***	0.094 (0.032)***	0.061 (0.031)**	0.061 (0.031)**	0.102 (0.031)***	0.103 (0.031)***
Individual controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,490	5,001	5,001	5,001	5,001	5,001	5,007	5,007
<i>Total effects on bottom half and bottom quarter</i>								
Coeff (Row 1) + Coeff (Row 2)			0.135		0.116		0.125	
Coeff (Row 1) + Coeff (Row 3)				0.118		0.126		0.089
p-value (total effect for bottom = 0)			0.091	0.229	0.122	0.216	0.117	0.319
p-value (effect for top quarter = effect for bottom quarter)				0.365		0.985		0.141

Impacto de *tracking* para o aluno mediano

- ▶ O que o grupo onde alunos medianos caem pode nos dizer sobre como professores reagem?
- ▶ Como podemos identificar esse efeito? Com um RDD.

Básicos de RDD

- ▶ Queremos estimar $Y_i = \alpha + \beta X_i + \varepsilon_i$ mas X_i é endógeno.
- ▶ Usamos outra variável Z_i como *running/forcing* variable onde X_i varia descontinuamente.
 - ▶ Sharp RD: $X_i = 1 \iff Z_i \geq \bar{Z}$.
 - ▶ Estimamos $\beta = \lim_{\varepsilon \downarrow 0} E(Y_i | Z_i = \bar{Z} + \varepsilon) = \lim_{\varepsilon \uparrow 0} E(Y_i | Z_i = \bar{Z} + \varepsilon)$
 - ▶ Fuzzy RD: $Z_i \geq \bar{Z} \implies$ mudança discreta em X_i .
 - ▶ Estimação com em IV usando $1[Z_i \geq \bar{Z}]$ como instrumento para X_i .
 - ▶ IV de fato só reescala o efeito para algo discreto ($0 \rightarrow 1$).
- ▶ Testes-padrão de validade
 - ▶ Gráfico de RD para “teste visual” (bins, polynomials, bandwidth, etc)
 - ▶ Instrumento não afetar outros *outcomes*
 - ▶ Continuidade de covariadas no *baseline*
 - ▶ Não-manipulação da *running variable* (McCrary, 2008)

Descontinuidade na nota de pares para alunos medianos

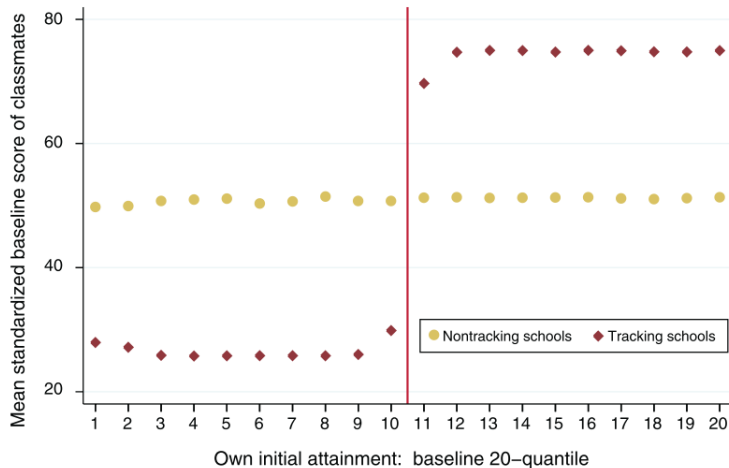


FIGURE 2. EXPERIMENTAL VARIATION IN PEER COMPETITION

Descontinuidade no baseline

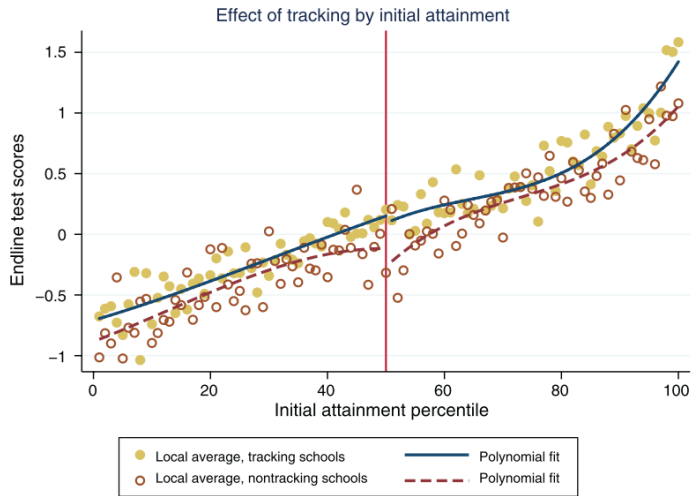


FIGURE 3. LOCAL POLYNOMIAL FITS OF ENDLINE SCORE BY INITIAL ATTAINMENT

Resultados do RDD

TABLE 5—PEER QUALITY: REGRESSION DISCONTINUITY APPROACH (TRACKING SCHOOLS ONLY)

	Total						
	Specification 1: With third order polynomial in baseline attainment		Specification 2: With second order polynomial in baseline attainment estimated separately on either side		Specification 3: With local linear regressions	Specification 4: Pair around the median	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Panel A. Reduced form</i>							
Estimated effect of bottom section at 50th percentile	0.010 (0.093)	0.001 (0.079)	−0.045 (0.106)	−0.051 (0.089)	−0.016 (0.204)	0.034 (0.136)	0.027 (0.145)
Observations (students)	2,959	2,959	2,959	2,959	2,959	149	149
School fixed effects	No	Yes	No	Yes	No	No	Yes
<i>Panel B. IV</i>							
Mean total score of peers	−0.012 (0.117)	−0.002 (0.106)				−0.068 (0.205)	−0.004 (0.277)
Observations (students)	2,959	2,959				149	149
School fixed effects	No	Yes				No	Yes
<i>Panel C. First stage for IV</i>							
In bottom half of initial distribution	−0.731 (0.047)***	−0.743 (0.021)***				−0.612 (0.090)***	−0.607 (0.058)***
Observations (students)	2,959	2,959				149	149
R^2	0.42	0.78				0.25	0.57
School fixed effects	No	Yes				No	Yes

Interpretação da não-discontinuidade

- ▶ Qualquer efeito positivo de pares mais fortes é compensado por nível de instrução ajustado para melhores alunos da metade inferior com *payoffs* convexos para professores.
- ▶ Poderia ser o caso também que professores tem *payoffs* côncavos e põem mais esforço na metade inferior. Como o artigo desconsidera essa possibilidade?

Estimando efeitos de pares

- ▶ Literatura grande sobre *peer effects* e efeitos de redes em educação (Sacerdote, 2011).
 - ▶ Eu aprendo de meus pares? (e.g. adoção de tecnologias)
 - ▶ Eu imito meus pares porque eles criam uma norma social? (e.g. uso de anticoncepcionais)
- ▶ Pares na escola tem um efeito em aprendizado?

Problemas com regressões de efeitos de pares

- ▶ Regressão padrão

$$Y_i = \alpha \bar{Y} + \beta Y_i^0 + \varepsilon_i$$

onde \bar{Y} é a média “leave-one-out” no *baseline*.

- ▶ Se usarmos dados contemporâneos ao invés de *baseline*, \bar{Y} é endógeno e temos o problema da reflexão (Manski, 1993).

- ▶ Comportamento de grupo afeta um indivíduo ou é simplesmente uma agregação de comportamentos individuais (correlacionados)?

- ▶ Possibilidade de quaisquer outras coisas estarem carregando os efeitos.

- ▶ Soluções

- ▶ Aleatorizar alunos a grupos de pares (e.g. colegas de turma, quartos em dormitórios)
 - ▶ Aleatorizar mudanças em alguns dos pares (e.g. *deworming* em Miguel and Kremer (2004) ou *bednets* em Cohen and Dupas (2010))

Resultados de pares

TABLE 4—PEER QUALITY: EXOGENOUS VARIATION IN PEER QUALITY (NONTRACKING SCHOOLS ONLY)

	All			25th–75th percentiles only	Bottom 25th percentiles	Top 25th percentiles only
	Total score (1)	Math score (2)	Lit score (3)	Total score (4)	Total score (5)	Total score (6)
<i>Panel A. Reduced form</i>						
Average baseline score of classmates ^a	0.346 (0.150)**	0.323 (0.160)**	0.293 (0.131)**	−0.052 (0.227)	0.505 (0.199)**	0.893 (0.330)***
Observations	2,188	2,188	2,188	2,188	2,188	2,188
School fixed effects	x	x	x	x	x	x
<i>Panel B. IV</i>						
Average endline score of classmates (predicted)	0.445 (0.117)***	0.470 (0.124)***	0.423 (0.120)***	−0.063 (0.306)	0.855 (0.278)***	1.052 (0.368)***
Observations	2,188	2,188	2,189	1,091	524	573
School fixed effects	x	x	x	x	x	x
<i>Panel C. First-Stage for IV: average endline score of classmates</i>						
	Average total score	Average math score	Average lit score	Average total score	Average total Score	Average total score
Average (standardized) baseline score of classmates [‡]	0.768 (0.033)***	0.680 (0.033)***	0.691 (0.030)***	0.795 (0.056)***	0.757 (0.066)***	0.794 (0.070)***

Resultados de pares: interpretação

- ▶ Há efeito positivo de pares quando controlando por habilidade individual.
- ▶ Em outra parte dos resultados, sensibilidade para controlar por notas individuais (como esperado)
- ▶ Efeitos diferentes dos de RDD. Explicações?

Testes estatísticos: tamanho e poder

- ▶ Sabemos que $E[Y_i(0) | T_i = 1] = E[Y_i(0) | T_i = 0]$ dado aleatorização.
- ▶ Mas em amostras finitas a diferença entre médias amostrais de Y_i para $T_i = 1$ e $T_i = 0$ pode não ser zero.
- ▶ O *tamanho* (ou *nível*) do teste (e.g. testando que $H_0 : ATE = 0$) é a probabilidade de erro tipo I (rejeitar H_0 quando H_0 é verdadeira).
- ▶ Padrão ao desenhar estudos é escolher tamanho $\alpha = 5\%$.
- ▶ Poder de um teste: 1 - probabilidade de erro tipo II (aceitar H_0 quando H_0 é falsa)

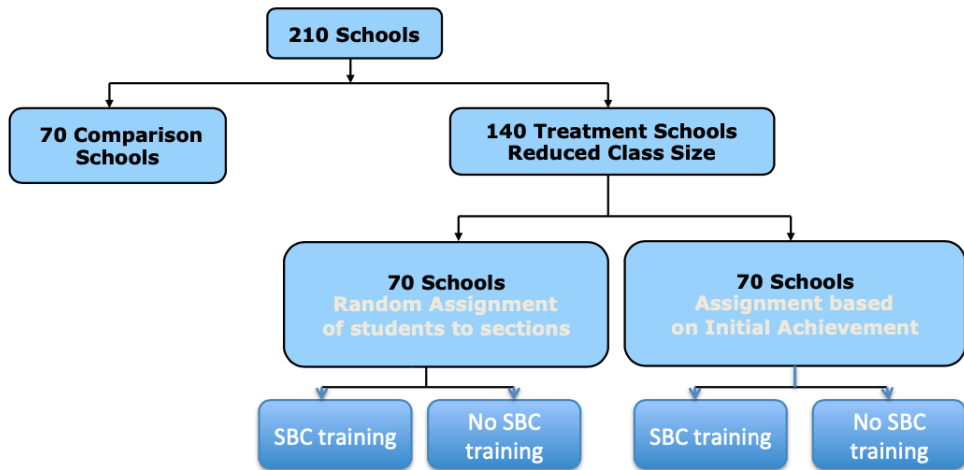
Poder

- ▶ Poder de um teste: 1 - probabilidade de erro tipo II (aceitar H_0 quando H_0 é falsa)
 - ▶ Padrão é escolher poder $\kappa = 80\%$.
- ▶ Com aleatorização simples, o efeito mínimo detectável (MDE) é:

$$MDE = (t_{1-\kappa} + t_{\alpha/2}) \sqrt{\frac{1}{P(1-P)}} \sqrt{\frac{\sigma^2}{N}}$$

- ▶ Nível de aleatorização
 - ▶ Com aleatorização *clustered*, poder depende do tamanho do *cluster* e da correlação esperadas de *outcomes* intra-*cluster*.
 - ▶ Seja ρ a correlação do *outcome* entre unidades intra-*cluster* e m o tamanho do *cluster*. O tamanho do menor efeito detectável dado o poder cresce com $D = \sqrt{1 + (m-1)\rho}$ (*design effect*).
- ▶ Leitura recomendada: Duflo et al. (2008), seção 4.

Desenho experimental



Na prática: casos simples

▶ No Stata, rodar *clustersampsi, samplesize mu1(.5) mu2(.67) m(80) rho(0.13)*.

▶ Parâmetros

- ▶ Média 1: 0.5
- ▶ Média 2: 0.67
- ▶ Desvio-padrão 1: 1.0
- ▶ Desvio-padrão 2: 1.0
- ▶ Significância: 0.05
- ▶ Poder: 0.8
- ▶ Ajuste por medidas de *baseline* (correlação): 0
- ▶ Tamanho médio de *cluster*: 80
- ▶ Correlação entra-*cluster* (ICC): 0.13

⇒ **Número de *clusters* por braço: 78**

Na prática: casos simples – controlando por covariadas no *baseline*

- ▶ No Stata, rodar *clustersampsi*, *samplesize* *mu1*(.5) *mu2*(.67) *m*(80) *rho*(0.13) *base_correl*(0.35).
- ▶ Parâmetros
 - ▶ Média 1: 0.5
 - ▶ Média 2: 0.67
 - ▶ Desvio-padrão 1: 1.0
 - ▶ Desvio-padrão 2: 1.0
 - ▶ Significância: 0.05
 - ▶ Poder: 0.8
 - ▶ Ajuste por medidas de *baseline* (correlação): 0.35
 - ▶ Tamanho médio de *cluster*: 80
 - ▶ Correlação entra-*cluster* (ICC): 0.13

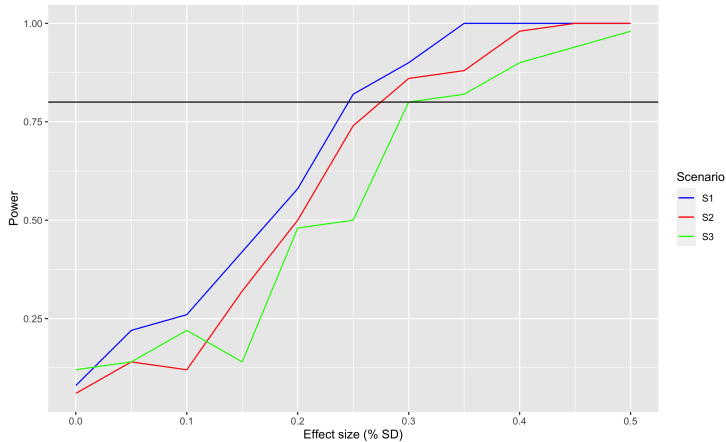
⇒ **Número de *clusters* por braço: 69**

Na prática: casos complexos

- ▶ Para casos onde solução analítica não é direta, é mais fácil produzir uma simulação.
- ▶ Passos:
 1. Estrutura do *setting*
 2. Aleatorização (simples, com *clusters*, estratificação, etc)
 3. Simulação dos dados (cenários de comparação, ICC, etc)
 4. Estimar regressões e calcular cada p-valor abaixo de 0.05.
- ▶ Recursos: **EGAP**, **DeclareDesign**

Na prática: casos complexos

- Exemplo meu, para um RCT sobre resiliência climática nas Filipinas:



Outros tópicos em qualidade de educação

- ▶ *Vouchers* para escolas privadas
 - ▶ Angrist et al. (2002) e Angrist et al. (2006)
 - ▶ Muralidharan and Sundararaman (2015)
 - ▶ Neilson (2021)
- ▶ Envolvimento parental
- ▶ Educação remedial
- ▶ Redução de tamanho de turma
 - ▶ Angrist and Lavy (1999)
 - ▶ Urquiola and Verhoogen (2009)

Vouchers para escolas privadas

- ▶ **Angrist et al. (2002)** e **Angrist et al. (2006)**: programa de vouchers da Colômbia teve impactos de curto e longo prazo no avanço de série, notas em testes e conclusão do ensino médio.
- ▶ **Muralidharan and Sundararaman (2015)**: Experimento de dois estágios na Índia que atribuiu vouchers aleatoriamente para escolas particulares em vilas tratadas.
 - ▶ Podem medir os impactos sobre os vencedores dos vouchers, à la Angrist et al.
 - ▶ Podem medir os efeitos a nível de mercado usando vilas de controle (onde indivíduos aplicaram para o programa); potenciais *spillovers* incluem mudança na composição de pares, mais recursos por aluno em escolas do governo.
- ▶ **Neilson (2021)**: estima a resposta de escolas privadas em preço e qualidade ao grande programa de vouchers no Chile.
 - ▶ O programa teve grande impacto em notas de alunos.
 - ▶ O canal principal é o aumento da qualidade da escola à medida que as escolas competem para atrair alunos.

Precaução: redução de tamanho de turma

- ▶ **Urquiola and Verhoogen (2009)** mostram que tamanho de turma não necessariamente é exógeno em desenhos de RDD.
- ▶ Em contextos onde escolas podem escolher tamanho de turma e preço, e famílias preferem turmas menores e tem liberdade de escolha, tamanho de turma passa a ser endógeno.
 - ▶ Violação da hipótese de não-manipulação de RDD (**McCrary, 2008**).
- ▶ Contexto: Chile.
 - ▶ 45% do mercado era privado em 2003.
 - ▶ Escolas pequenas: 95% das urbanas tinham <135 alunos na quarta série.

Padrão de “dentes”

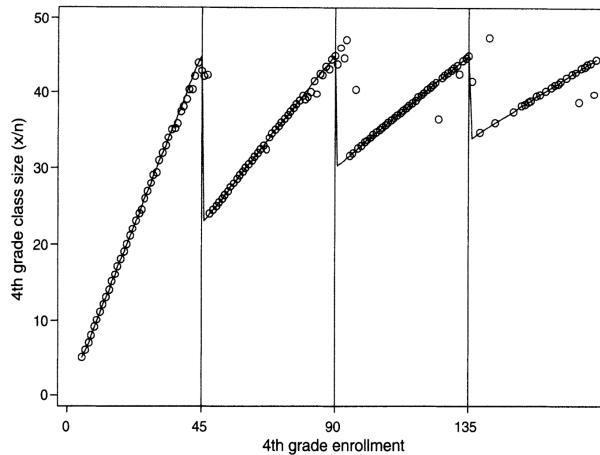
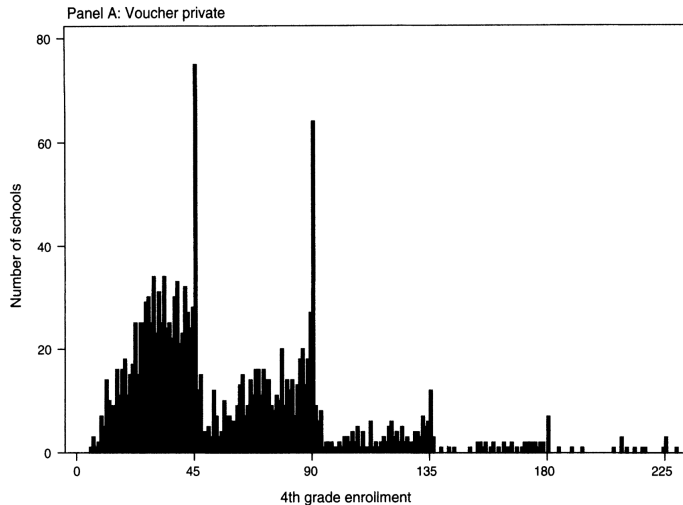


FIGURE 5. FOURTH GRADE ENROLLMENT AND CLASS SIZE IN URBAN PRIVATE VOUCHER SCHOOLS, 2002

Escolas ajustam para evitar adicionar turmas



Covariadas descontínuas nos *cutoffs*



Conclusão

- ▶ Literatura rica e com muitas perguntas relevantes.
- ▶ Muitos dados e variação: fértil para pesquisa (especialmente no Brasil, Chile, etc).

Referências I

- Angrist, Joshua D. and Victor Lavy**, “Using Maimonides Rule to Estimate the Effect of Class Size on Scholastic Achievement,” *Quarterly Journal of Economics*, 1999, 114 (2), 533–575.
- Angrist, Joshua, Eric Bettinger, and Michael Kremer**, “Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia,” *American Economic Review*, 2006, 96 (3), 847–862.
- , — , **Erik Bloom, Elizabeth King, and Michael Kremer**, “Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment,” *American Economic Review*, 2002, 92 (5), 1535–1558.
- Banerjee, Abhijit V., Shawn Cole, Esther Duflo, and Leigh Linden**, “Remedying Education: Evidence from Two Randomized Experiments in India,” *Quarterly Journal of Economics*, 2007, 122 (3), 1235–1264.
- Cohen, Jessica and Pascaline Dupas**, “Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment,” *Quarterly Journal of Economics*, 2010, 125 (1), 1–45.

Referências II

- Duflo, Esther, Pascaline Dupas, and Michael Kremer**, “Peer Effects, Teacher Incentives, and the Impact of Early Tracking: Evidence from a Randomized Evaluation in Kenya,” *American Economic Review*, 2011, 101 (5), 1739–1774.
- , **Rachel Glennerster, and Michael Kremer**, “Chapter 61 Using Randomization in Development Economics Research: A Toolkit,” in “Handbook of Development Economics,” Vol. 4 2008, pp. 3895–3962.
- Glewwe, Paul, Michael Kremer, and Sylvie Moulin**, “Many Children Left Behind? Textbooks and Test Scores in Kenya,” *American Economic Journal: Applied Economics*, 2009, 1 (1), 112–135.
- Manski, Charles F.**, “Identification of Endogenous Social Effects: The Reflection Problem,” *Review of Economic Studies*, 1993, 60 (3), 531–542.
- McCrary, Justin**, “Manipulation of the running variable in the regression discontinuity design: A density test,” *Journal of Econometrics*, 2008, 142 (2), 698–714.
- Miguel, Edward and Michael Kremer**, “Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities,” *Econometrica*, 2004, 72 (1), 159–217.

Referências III

- Muralidharan, Karthik and Venkatesh Sundararaman**, “The Aggregate Effect of School Choice: Evidence From a Two-Stage Experiment in India,” *Quarterly Journal of Economics*, 2015, 130 (August), 1011–1066.
- Neilson, Christopher**, “Targeted Vouchers, Competition Among Schools, and the Academic Achievement of Poor Students,” 2021.
- Sacerdote, Bruce**, *Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?*, 1 ed., Vol. 3, Elsevier B.V., 2011.
- Urquiola, Miguel and Eric Verhoogen**, “Class-Size Caps, Sorting, and the Regression-Discontinuity Design,” *American Economic Review*, 2009, 99 (1), 179–215.