Mediation Analysis in Difference-in-Differences Designs

Timo Schenk

Aarhus University ($\rightarrow$ Erasmus University Rotterdam)

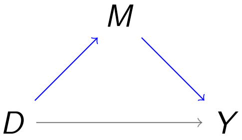
Microeconometrics Class of 2024 Conference

Motivation

- Standard DiD: total effect of D on Y
- But *why* did D affect Y?
- This paper: indirect effect D on Y mediated by M

Motivation

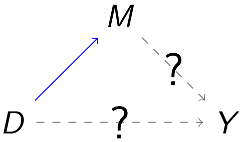
- Standard DiD: total effect of D on Y
- But *why* did D affect Y?
- This paper: indirect effect D on Y mediated by M



Empirical practice

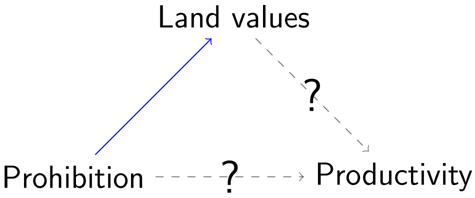
Most empirical DiD papers study mechanisms informally:

- Estimate the effect of D on M,
- If significant, conclude that M might be an important mechanism.

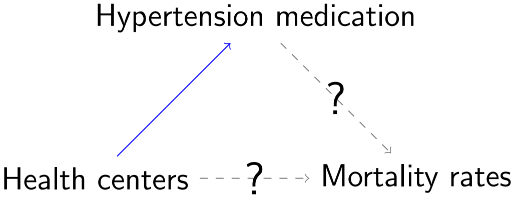


To quantify the mechanism we need the effect of M on Y!

Empirical examples



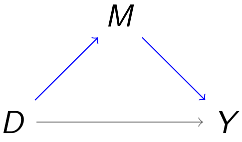
Howard and Ornaghi (2021)



Bailey and Goodman-Bacon (2015)

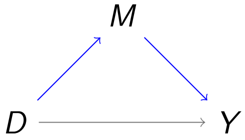
Identifying indirect effects is challenging

- Need additional assumptions even if D is randomized
  - Sequential conditional independence (Imai et al., 2010; Heckman and Pinto, 2013; Huber et al., 2017),
  - Instruments for D and M (Frölich and Huber, 2017)
  - Parallel trends in Y for different levels of M (Deuchert et al., 2019)
- In DiD designs, we typically face selection into D and M.



This paper

1. Identification strategies for the indirect effect in DiD designs
   - Sequential DiD approach: sequential parallel trends & restricted effect heterogeneity
   - Two-sample approach: external validity
2. Robust inference methods,
3. Two types of applications (one for each approach)



More related literature

- DiD when covariates are affected by the treatment (Caetano et al., 2022; Brown et al. 2023)
- Dynamic effects of staggered natural experiments (de Chaisemartin and D'Haultfoeuille, 2024)
- LATE without exclusion restriction (Flores and Flores-Lagunes, 2013; Huber and Mellace, 2015; Kitagawa, 2015; Kwon and Roth, 2024)

Simple setup

- Two periods: $t = 0$ (pre-treatment) and $t = 1$ (post-treatment)
- Binary treatment D_{it}, binary mediator M_{it}
- Sharp timing: $D_{i0} = M_{i0} = 0$ (relaxed in paper)
- Potential outcome $Y_{it}(d, M_{it}(d))$, potential mediator $M_{it}(d)$
- Observed outcome and mediator:

$$Y_{it} = \begin{cases} Y_{it}(1, M_{it}(1)), & D_{it} = 1 \\ Y_{it}(0, M_{it}(0)), & D_{it} = 0 \end{cases}, \quad M_{it} = \begin{cases} M_{it}(1), & D_{it} = 1 \\ M_{it}(0), & D_{it} = 0 \end{cases}$$

The total effect is well-identified

Assumption (Parallel trends across treatment groups)
$\mathrm{E}[Y_1(0, M(0)) - Y_0(0, M(0))|D]$ *does not depend on* D*.*

The DiD estimand identifies the total effect

$$\begin{aligned}\tau_{att} &:= \mathrm{E}\left[Y_1(1, M(1)) - Y_1(0, M(0))|D = 1\right] \\ &= \mathrm{E}[\Delta Y|D = 1] - \mathrm{E}[\Delta Y|D = 0]\end{aligned}$$

with $\Delta Y = Y_1 - Y_0$

Effect of interest: the indirect effect

Definition (Indirect effect)

$$\tau_{ind} := \mathrm{E}[Y_1(0, M(1)) - Y_1(0, M(0)) | D = 1]$$

$=$ the effect of D on Y mediated by M

Decomposition of the total effect:

$$\tau_{att} = \tau_{ind} + \tau_{dir}$$

The indirect effect is driven by switchers

Table: Unobserved mediator groups

Group	G	$M(0)$	$M(1)$
Always mediators	a	1	1
Never mediators	n	0	0
Compliers	c	0	1
Defiers	d	1	0

The indirect effect is

$$\tau_{ind} = \pi_c \beta_c - \pi_d \beta_d$$

with local average mediator effects

$$\beta_g = \mathrm{E}[Y_1(0,1) - Y_1(0,0) | D = 1, G = g]$$

and share of compliers (defiers)

$$\pi_g = \Pr[G = g | D = 1]$$

for $g = c, d$

Sequential DiD: implementation

1. Total (reduced form) effect τ_{att}

$$\Delta Y_i = \delta + \tau_{att} D_i + \varepsilon_i$$

2. Treatment effect on the mediator π

$$M_i = \alpha + \pi^* D_i + \nu_i$$

3. Mediator effect β

$$\Delta Y_i = \tilde{\delta} + \beta^* M_i + \gamma_0 D_i + \gamma_1 D_i M_i + \tilde{\varepsilon}_i$$

4. Indirect effect

$$\tau^*_{ind} = \tau_{att} - (\gamma_0 + \pi^* \gamma_1) = \pi^* \beta^*$$

Sequential DiD: identification

Sequential DiD identifies the product

$$\tau_{ind}^{*} = \pi^{*}\beta^{*}$$

- Ingredient 1: average effect of D on M

$$\pi^{*} = \mathrm{E}[M|D=1] - \mathrm{E}[M|D=0]$$

- Ingredient 2: average mediator effect

$$\beta^{*} = \mathrm{E}[\Delta Y|D=0, M=1] - \mathrm{E}[\Delta Y|D=0, M=0]$$

Main decomposition: the three sources of bias

Theorem

$$\tau^*_{ind} = \tau_{ind} + SM + EH + NW$$

with

1. *Selection into mediation:* $SM = \pi^*(\delta_1 - \delta_0)$ *with* $\delta_m = \mathrm{E}[\Delta Y(0,0)|D, M(0) = m]$,
2. *Mediator effect heterogeneity:* $EH = \pi_c(\beta_a - \beta_c)$, *and*
3. *Negative weights bias:* $NW = \pi_d \tilde{\pi}(\beta_d - \beta_a)$, $\tilde{\pi} = \frac{\pi_c - \pi_a}{\pi_d + \pi_a}$.

Regularity conditions

Identification based on sequential parallel trends

$$\tau_{ind}^{*} = \tau_{ind} + SM + EH + NW$$

So $\tau_{ind}^{*} = \tau_{ind}$ under

1. Parallel trends across mediator groups:
   $\mathrm{E}[\Delta Y(0,0)|D, M(0)]$ does not depend on $M(0) \Rightarrow\; SM = 0$
2. Mediator effect homogeneity: $\beta_a = \beta_c \Rightarrow\; EH = 0$
3. Monotonicity in treatment effects on the mediator (no defiers)
   $M(1) \geq M(0)\ a.s. \;\Rightarrow\; NW = 0$

▸ Visualization

Two-sample approach

- Problem: Main approach not feasible in some common settings
- Example: $\Pr(M = 1|D = 0) = 0$ (mediator "part of" treatment)
- Consequence: β^* not identified in the main sample

Solution

- Mediator effects ($M \to Y$) identified in a separate sample (e.g. RCT)

$$\tau^{\dagger}_{ind} = \pi^* \beta^{\dagger}$$

- Parallel trends across mediator groups no longer required

Inference challenges

1. Conventional tests for $\tau_{ind} = 0$ are conservative. ▸ Details

   Solution: improved test for mediation van Garderen and van Giesbergen (2022)

2. Conventional inference on $\frac{\tau_{ind}}{\tau_{att}}$ not uniformly valid.

   Solution: robust CIs based on the Fieller method (aka Anderson-Rubin)

Robust inference on the proportion mediated

- Important object: the proportion mediated $r = \frac{\tau_{ind}}{\tau_{att}}$
- Conventional confidence intervals (CIs) based on the delta method

$$CI_{conv} = \{\hat{r} \pm 1.96 SE[\hat{r}]\}$$

- Robust CIs based on the Fieller method

$$CI_{rob} = \{r : AR(r) \leq q_{0.95}\}$$

i.e. collect all r for which $H_0(r)$: $\tau_{ind} - r\tau_{att} = 0$ cannot be rejected

Asymptotics Simulations

Two types of applications

1. (Economic history) How Prohibition increased productivity through farm values. Howard and Ornaghi (2021)

   ▸ Details

2. (Public health) How Community Health Centers decreased mortality rates through increased use of anti-hypertensive medication. Bailey and Goodman-Bacon (2015)

   ▸ Details

How did Community Health Centers decrease mortality rates?

- Treatment D: County establishes CHC ▸ Map
- Outcome Y: Mortality rate among aged 50+ years, 5-9 years after treatment
- Total effect: $\hat{\tau}_{att} = 274$ avoided deaths per 100k (SE 51.5)
- Proposed mediator M: prescription of anti-hypertensive medication
- Problem: no (observed) variation of M in untreated counties

How important is the hypertension channel?

Table: Calculation of the indirect effect in Bailey and Goodman-Bacon (2015)

Effect	Source	Est	Std. err.
Effect of D on Y ($\hat{\tau}_{att}$)		274	(51.5)
Share of CHC beneficiaries	$P[takeup \mid D = 1]$	0.16	-
Hypertension prevalance (α_{hyp})	NHE Survey[1]	0.262	-
Effect of D on M (π^*)	$P[takeup \mid D = 1] \times \alpha_{hyp}$	0.042	-
Local effect of M on Y ($\hat{\beta}_{hyp}$)	HDFP RCT[2]	2,168	(760)
Indirect effect ($\hat{\tau}_{ind}$)	$\hat{\beta}_{hyp} \times \pi^*$	91	(31.9)
Relative indirect effect ($\hat{r}$)	$\hat{\tau}_{ind} / \hat{\tau}_{att}$	33.2%	(13.2%)

[1] National Health Examination Survey

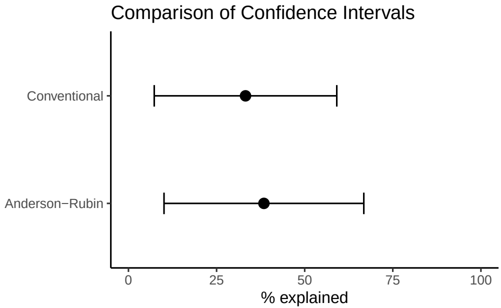
[2] Hypertension Detection and Follow-Up Program

How credible is this ratio?

Identification

- Selection into M: no problem as we took β from RCT
- Monotonicity: did CHCs prevent some people from M?
- Mediator effect heterogeneity: how different is the population of the RCT from the CHC beneficiaries?

Inference



Back

Summary

1. Identification strategies for the indirect effect in DiD designs
   - Main approach: sequential parallel trends & restricted effect heterogeneity
   - Two sample approach: external validity
2. Robust inference methods,
3. Two types of applications (one for each approach)

Robust parallel trends

Parallel trends across treatment groups do not hold just by coincidence.

Assumption

$\mathrm{E}[\Delta Y_i(0,0)|D_i, M_i(0) = m]$ *does not depend on* D_i *for both* $m = 0, 1$.

Assumption

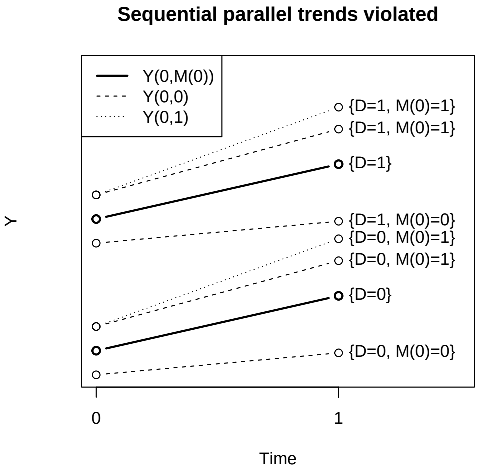
$\mathrm{E}[M_i(0)|D_i]$ *does not depend on* D_i.

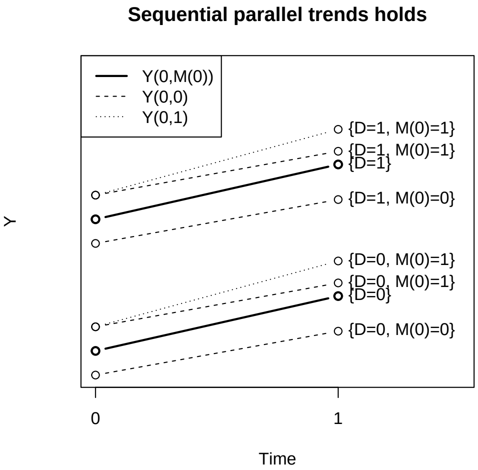
Assumption

The mediator effects are the same between treatment groups: $\beta_{1,1} = \beta_{0,1}$.

Back

Visualizing sequential parallel trends





Back

Inference on the indirect effect

Consider testing for the presence on an indirect effect

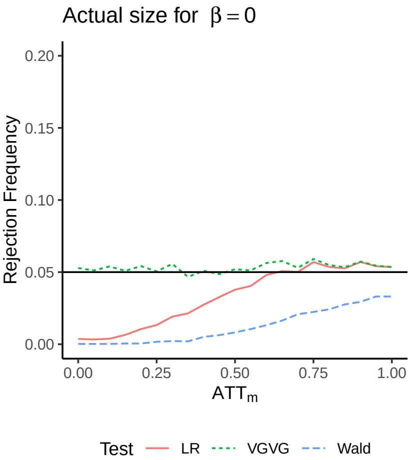
$$H_0 \colon \pi\beta = 0 \quad \Leftrightarrow \quad H_{0,a} \colon \pi = 0 \quad \vee \quad H_{0,b} \colon \beta = 0$$

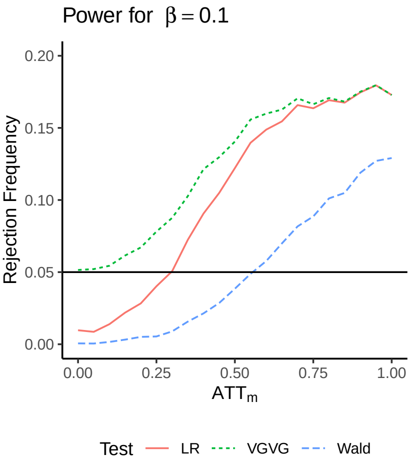
Conventional tests:

- Wald / Sobel test based on $\frac{\hat{\tau}_{ind}}{\mathrm{se}[\tau_{ind}]}$
- LR: test $H_{0,a}$ and $H_{0,b}$ individually based on corresponding t-statistics t_π, t_β
  Reject H_0 iff both $H_{0,a}$ and $H_{0,b}$ reject

Problem: power below nominal size when π and β are both small.

Conventional tests are conservative





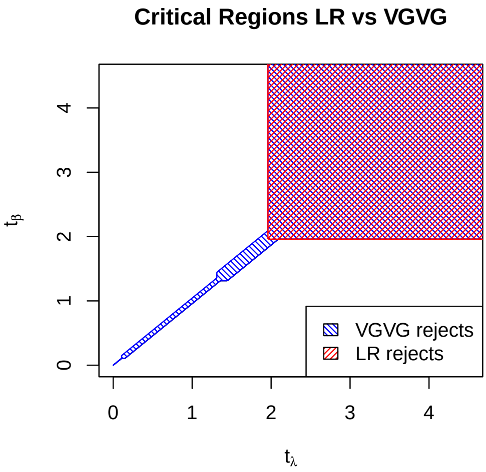
DGP: $\Delta Y = \beta \Delta M + \varepsilon$ and $\Delta M = ATT_m D + \nu$ over a grid $ATT_m \in \{0, 0.05, \ldots, 1\}$. Sample size $N = 100$.

Solution: improved test for mediation

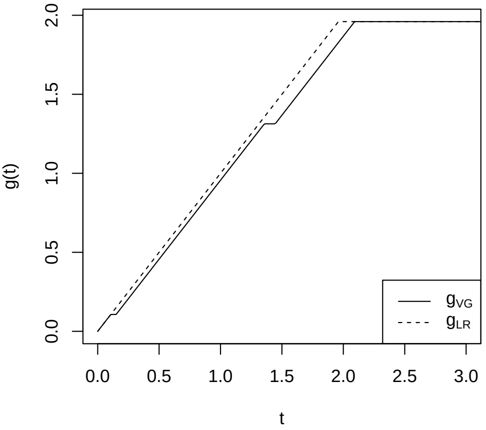
van Garderen and van Giesbergen (2022) [VGVG]

- Order the t-statistics
  $t_{(1)} = \min\{|t_\beta|, |t_\pi|\}$,
  $t_{(2)} = \max\{|t_\beta|, |t_\pi|\}$
- Reject if $g(t_{(2)}) < t_{(1)}$ with $g(.)$ a transformation given in VGVG

Back



Details on the VGVG test



Asymptotic distribution of the effect ratio

Let $\boldsymbol{\theta}$ collect the parameters of the two linear regressions

$$\begin{aligned}\Delta Y_i &= \delta + \beta M_i + \gamma_0 D_i(1 - M_i) + \gamma_1 D_i M_i + \varepsilon_i \\ M_i &= \pi_0 + \pi_1 D_i + \nu_i\end{aligned}$$

Theorem

Let $\{Y_{i0}, Y_{i1}, M_i, D_i\}_{i=1}^N$ be an independent sample. Let $\hat{\boldsymbol{\theta}}_\tau = \boldsymbol{g}(\hat{\boldsymbol{\theta}}) = (\hat{\tau}_{ind}, \hat{\tau}_{dir})'$. Then $\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^*) \xrightarrow{d} \mathcal{N}[\mathbf{0}, \boldsymbol{\Sigma}]$ as $N \to \infty$ with $\boldsymbol{\Sigma}$ the full-rank variance matrix. Consequently,

$$\sqrt{N}(\hat{\boldsymbol{\theta}}_\tau - \boldsymbol{\theta}_\tau^*) \xrightarrow{d} \mathcal{N}[\mathbf{0}, \boldsymbol{\Omega}]$$

where $\boldsymbol{\Omega} = \boldsymbol{G}_0' \boldsymbol{\Sigma} \boldsymbol{G}_0$ and $\boldsymbol{G}_0 = [\nabla g_1, \nabla g_2]_{\theta=\theta^*}$ the jacobian of $\boldsymbol{g} : \Theta \to \mathbb{R}^2$.

Asymptotic distribution of the effect ratio

Consider now the ratio $\hat{r} = \frac{\hat{\tau}_{ind}}{\hat{\tau}_{att}}$.

Corollary

If $\tau_{att} \neq 0$, then

$$\sqrt{N}(\hat{r} - r_0) \xrightarrow{d} \mathcal{N}[0, \sigma^2_{\hat{r}}]$$

with $\sigma^2_{\hat{r}} = \nabla r_0 \boldsymbol{\Omega} \nabla r_0$ and $\nabla r_0 = \frac{1}{\tau_{att}}(\tau_{dir}, -\tau_{ind})'$.

Corollary

Let $AR(r) = \frac{(\hat{\tau}_{ind} - r\hat{\tau}_{att})^2}{\hat{\sigma}^2_{ar}}$ and suppose $\boldsymbol{\Omega}$ has full rank. Then $AR(r_0) \xrightarrow{d} \chi^2(1)$ for all τ_{att}.

▸ Back

Simulation: strong ATT

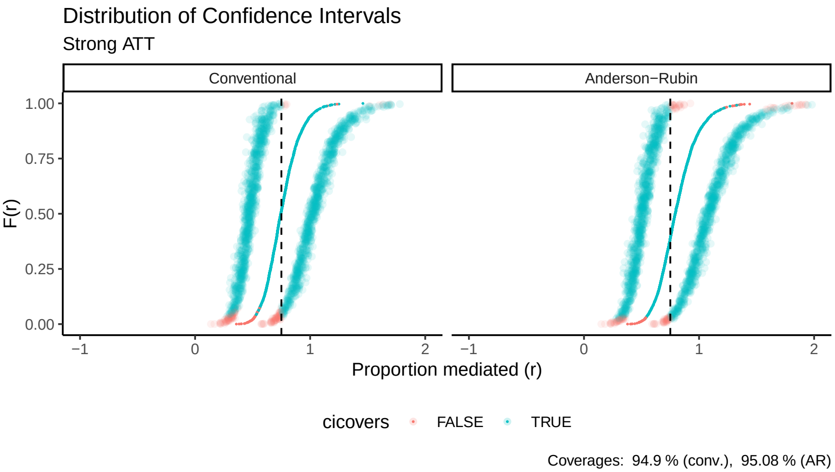


Figure: $\Delta Y_i | (M_i, D_i) \sim \mathcal{N}[3M_i + \gamma_0 D_i, 16],\ M_i | D_i \sim Bern(0.25 + \pi^* D_i),\ (\pi^*, \gamma_0) = (0.5, 0.5)$

Simulation: weak ATT

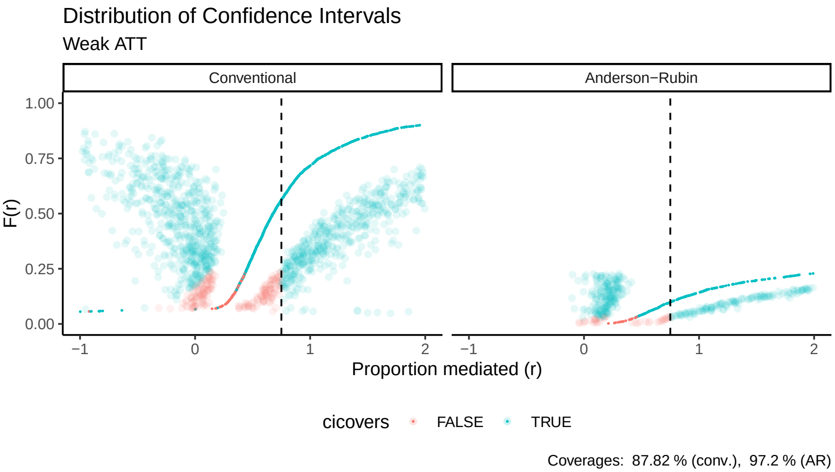


Figure: $\Delta Y_i|(M_i, D_i) \sim \mathcal{N}[3M_i + \gamma_0 D_i, 16],\ M_i|D_i \sim Bern(0.25 + \pi^* D_i),\ (\pi^*, \gamma_0) = (0.1, 0.1)$

Back

The indirect effects of Prohibition

Setting

$t = 1$	1910
$t = 0$	1900
D_i	early adopting county
$Y_{i,t}$	log productivity
$M_{i,t}$	log farm value
$X_{i,1900}$	baseline demographics

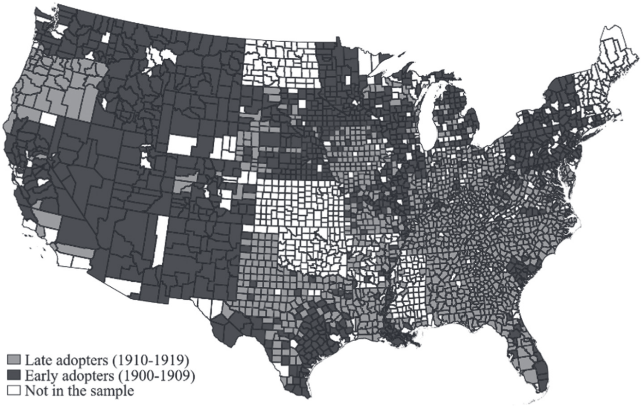


FIGURE 2
TREATMENT STATUS MAP

Figure: Howard and Ornaghi (2021, Fig. 2).

Results

Table: Decomposed Effect of Prohibition on Productivity

	Estimate	Std. Err.	95% CI
Total effect	0.072*	0.043	[-0.012, 0.156]
Effect D on M	0.081***	0.024	[0.034, 0.128]
Effect M on Y	0.567***	0.104	[0.363, 0.771]
Indirect effect	$0.046^{\dagger}$	0.014	[0.019, 0.074]
% mediated (conv.)	63.89%	36.79%	[-7.6%, 136.6%]
% mediated (AR)	-	-	$(-\infty, \infty)$
Direct effect	0.026	0.042	[-0.06, 0.108]

Empirical strategy

Back to applications

Empirical strategy

Reduced form

$$\Delta Y_i = \tau_{att} D_i + X'_{i,1900}\vartheta + \alpha_{s(i)} + \varepsilon_i$$

First stage

$$\Delta M_i = \pi D_i + X'_{i,1900}\vartheta + \alpha_{s(i)} + \varepsilon_i$$

Second stage

$$\Delta Y_i = \gamma D_i + \beta \Delta M_i + X'_{i,1900}\vartheta + \alpha_{s(i)} + \varepsilon_i$$

Back

Establishment of CHCs

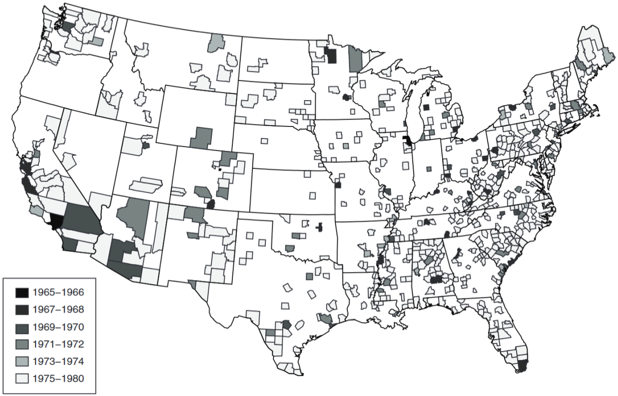


FIGURE 3. ESTABLISHMENT OF COMMUNITY HEALTH CENTERS BY COUNTY OF SERVICE DELIVERY, 1965–1980

Note: Dates are the first year that a CHC was established in the county.
Source: Information on CHCs drawn from NACAP and PHS reports.

Back

Survey of Empirical Studies

Table: Proposed mechanisms in DiD papers surveyed in Roth (2022)

Paper	Outcome	Treatment	Mechanism
Bailey and Goodman-Bacon (2015)	mortality rate	Community Health Centers	anti-hypertensive medication
Deryugina (2017)	government transfers	hurricane	demographics, earnings, employment
Fitzpatrick and Lovenheim (2014)	test scores	teacher retirements	shift in resources, teacher assignment
Gallagher (2014)	flood insurance take up	large regional floods	flood costs, migration, protective measures