Distributional effects of universal child care:

Methods and evidence

Magne Mogstad, UCL

Universal and targeted child care policy

Broadly speaking, two distinct models of child care:

1. Universal programs: widely available, publicly subsidized child care
   - as offered e.g. in Scandinavia
2. Targeted programs: public investment in child care focused on low income families
   - as offered e.g. in the U.S.

Universal vs. targeted policy: Arguments

"The simple economics of intervention therefore suggests that society should focus its investment where it's likely to have very high returns.

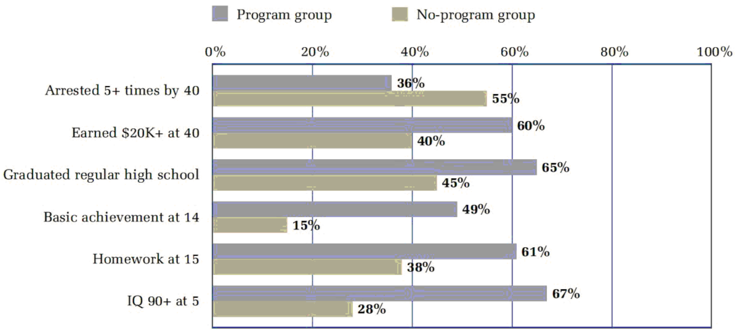
Right now, that is the disadvantaged population....Functioning middle-class homes are producing healthy, productive kids.. It is foolish to try to substitute for what the middle-class and upper-class parents are already doing."

Heckman (2005)

The High Scope Perry Pre-School program

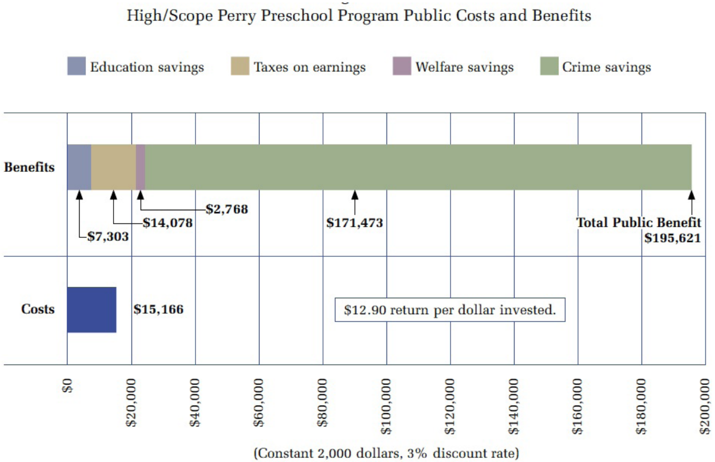
- RCT carried out in Michigan, US
- 58 of 123 high risk children aged 3 and 4 were assigned to a high quality preschool program in the early 1960s
- These children were followed into adulthood.

Major Findings: High/Scope Perry Preschool Study at 40



Schweinhart et al (2005) The High/Scope Perry Preschool Study through age 40.

Perry Preschool: Costs and Benefits



Schweinhart et al (2005) The High/Scope Perry Preschool Study through age 40.

Universal vs. targeted policy: Evidence base

Counterargument: Even if the returns are greater for the poor

- publicly subsidized child care may still have benefits for middle or upper-class children that exceed its costs

To assess this argument, we need credible evidence on the net benefits of subsidized child care

- for middle and upper-class children
- as compared to children from low-income families

Current evidence base for universal child care is insufficient:

- small, nonexperimental, and offers mixed results
- focused on mean impacts

(Literature reviews: Baker, 2011; Almond and Currie, 2010; Ruhm and Waldfogel; 2011)

This talk:

WHAT: Investigate the effects of universal child care on child development

- in a way that allows the effects to vary systematically over the outcome distribution

HOW: Using nonlinear difference-in-differences (DiD) methods, we

- examine how the introduction of large-scale, publicly subsidized child care in Norway
- affected the earnings distribution of exposed children as adults

WHY: The estimated quantile treatment effects (QTE) allows us to assess:

- the impact of subsidized child care in the lower, middle and upper part of the outcome distribution
- what (overall and subgroup) mean impacts miss, because they are averaging together effects of different magnitude and sign

Outline

Methods:

- Potential outcomes framework
- Standard and nonlinear DiD methods
- Inclusion of covariates in QTE estimation

Empirical analysis:

- The child care reform
- Main results
- Specification checks
- Interpretation: Theoretical framework

DiD: Potential and observed outcomes

For each child i, we have three *potential outcomes:*

$$\begin{aligned}
\text{Potential outcome in period 1} &= Y_{i1}^0 \\
\text{Potential outcome in period 2} &= \begin{cases} Y_{i2}^1 & \text{if } T_{i2} = 1 \\ Y_{i2}^0 & \text{if } T_{i2} = 0 \end{cases}
\end{aligned}$$

The *observed outcome* in period t:

$$Y_{it} = T_{it} Y_{it}^1 + (1 - T_{it}) Y_{it}^0 = T_{it}(Y_{it}^1 - Y_{it}^0) + Y_{it}^0$$

DID: Mean impacts

Standard DiD identifies ATET:

$$E(Y^1_{i2} - Y^0_{i2} | T_{i2} = 1) \quad = \quad E(Y_{i2} - Y_{i1} | T_{i2} = 1) - E(Y_{i2} - Y_{i1} | T_{i2} = 0)$$

under the folllowing assumption:

Common trend in the absence of intervention

$$E(Y^0_{i2} - Y^0_{i1} | T_{i2} = 1) = E(Y^0_{i2} - Y^0_{i1} | T_{i2} = 0)$$

$\Rightarrow$no selection on the change in non-treatment outcome level

DID: Mean impacts

Standard DiD identifies ATET:

$$E(Y_{i2}^1 - Y_{i2}^0 | T_{i2} = 1) \quad = \quad E(Y_{i2} - Y_{i1} | T_{i2} = 1) - E(Y_{i2} - Y_{i1} | T_{i2} = 0)$$

under the folllowing assumption:

Common trend in the absence of intervention

$$E(Y_{i2}^0 - Y_{i1}^0 | T_{i2} = 1) = E(Y_{i2}^0 - Y_{i1}^0 | T_{i2} = 0)$$

$\Rightarrow$no selection on the change in non-treatment outcome level

Common trend assumption allows for:

- selection on non-treatment levels:

$$E(Y_{it}^0 | T_{i2} = 1) \neq E(Y_{it}^0 | T_{i2} = 0), \; t = 1, 2$$

- selection on gains:

$$E(Y_{i2}^1 - Y_{i2}^0 | T_{i2} = 1) \neq E(Y_{i2}^1 - Y_{i2}^0 | T_{i2} = 0)$$

DiD: Quantile treatment effects (QTE)

Nonlinear DiD methods to estimate QTE:

- Quantile DiD (QDID)
- Changes in changes (CiC): proposed by Athey and Imbens (2006, Econometrica)
- RIF-DiD: extension of Firpo et al. (2010, Econometrica)

Allow estimation of counterfactual outcome distribution

- in the absence of intervention

Differ in:

- identifying assumption
- how to handle covariates
- invariance wrt. monotone transformation of dependent variable

What is QTE?

Remember that the τ-quantile is defined as:

$$Y_\tau = F^{-1}(\tau)$$

For instance, the median is $Y_{0.5} = F^{-1}(0.5)$

QTE at quantile τ is defined as:

$$QTE_\tau = Y_\tau^1 - Y_\tau^0$$

That is, the differences in the τ-quantile in the distributions of potential outcomes

- with treatment (Y_τ^1) and without treatment (Y_τ^0)

Benchmark: Estimating QTE with a RCT

With randomization, the QTE can be identified from the

- *observed* outcome distributions of the treatment and the control group:

$$QTE_\tau = F^{-1}(\tau | T_i = 1) - F^{-1}(\tau | T_i = 0)$$

Because the treatment and control group will asymptotically have identical distributions of

- (pre-assignment) unobservables and observables, and thus potential outcomes

Benchmark: Estimating QTE with a RCT

With randomization, the QTE can be identified from the

- *observed* outcome distributions of the treatment and the control group:

$$QTE_\tau = F^{-1}(\tau | T_i = 1) - F^{-1}(\tau | T_i = 0)$$

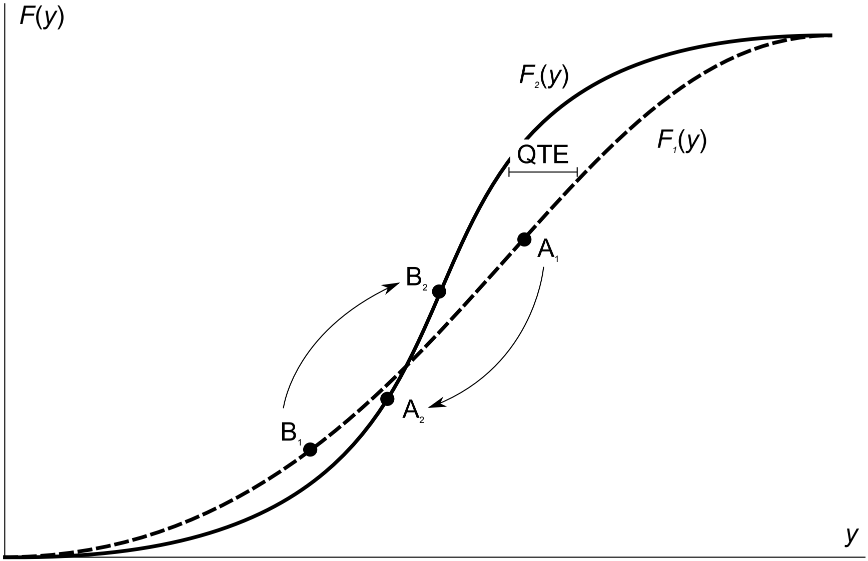
Because the treatment and control group will asymptotically have identical distributions of

- (pre-assignment) unobservables and observables, and thus potential outcomes

But note that QTE does not identify the distribution of effects

- if treatment causes rank reversals in the outcome distribution

Quantile treatment effects and rank reversals



Nonlinear DiD estimation of QTE

Recall that standard DiD identifies

$$E(Y_{i2}^1 - Y_{i2}^0 | T_{i2} = 1)$$

by estimating the counterfactual mean outcome:

$$E(Y_{i2}^0 | T_{i2} = 1) = E(Y_{i1}^0 | T_{i2} = 1) + E(Y_{i2}^0 - Y_{i1}^0 | T_{i2} = 0)$$

Nonlinear DiD estimation of QTE

Recall that standard DiD identifies

$$E(Y_{i2}^1 - Y_{i2}^0 | T_{i2} = 1)$$

by estimating the counterfactual mean outcome:

$$E(Y_{i2}^0 | T_{i2} = 1) = E(Y_{i1}^0 | T_{i2} = 1) + E(Y_{i2}^0 - Y_{i1}^0 | T_{i2} = 0)$$

Nonlinear DiD methods rely on a similar idea for quantiles

Below, for $t = 1, 2$ let

- $F_t(Y)$ be the distribution of Y in the treatment group
- $G_t(Y)$ be the distribution of Y in the control group

Quantile DiD (QDiD)

Three steps:

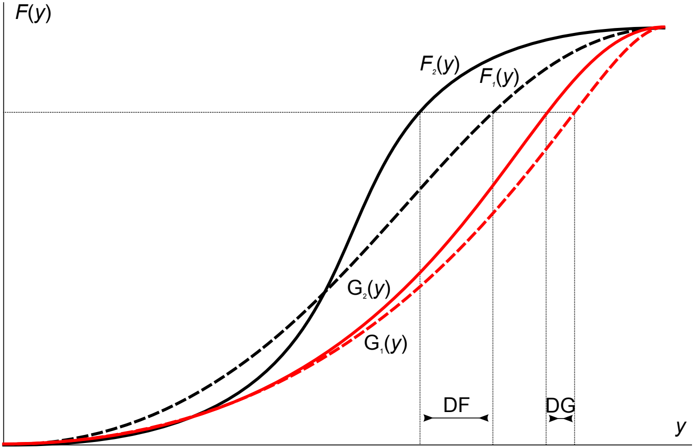
1. Fix the quantile of Y in the pre-reform outcome distribution of the treatment group, $F_1(Y) = \tau$
2. Counterfactual post-reform outcome at that quantile in the treatment group:

$$\begin{aligned} k^{QDID}(\tau) &= F_1^{-1}(\tau) + \Delta^{QDID} \\ &= F_1^{-1}(\tau) + \left(G_2^{-1}(\tau) - G_1^{-1}(\tau)\right) \end{aligned}$$

3. QTE estimate at quantile τ is then

$$F_2^{-1}(\tau) - k^{QDID}(\tau)$$

QDID: Graphical representation



Identifying assumption: common trend in levels at the quantile

RIF-DID

Three steps:

1. Fix the level of Y in the pre-reform outcome distribution of the treatment group, y
2. Counterfactual post-reform outcome at outcome level y in the treatment group:

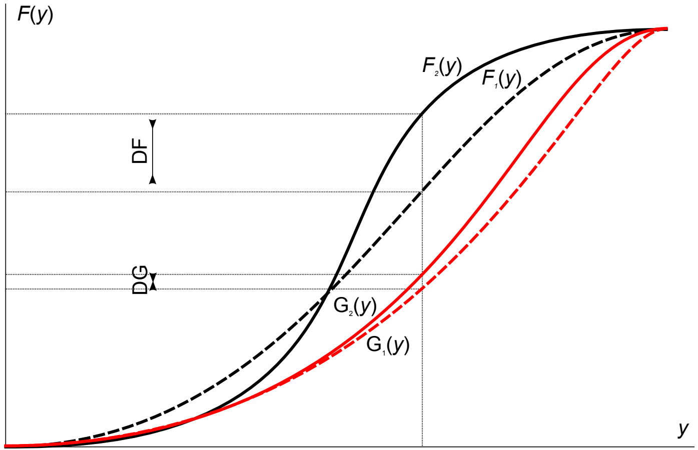
$$k^{UQDID}(y) = F_1(y) + \Delta^{UQDID} = F_1(y) + (G_2(y) - G_1(y))$$

3. The nonlinear DiD-estimate in pop. shares at level y is then

$$F_2(y) - k^{UQDID}(y)$$

- which can then be inverted to get associated QTE

RIF-DID: Graphical representation



Identifying assumption: common trend in the population share at the outcome level

CIC

Three steps:

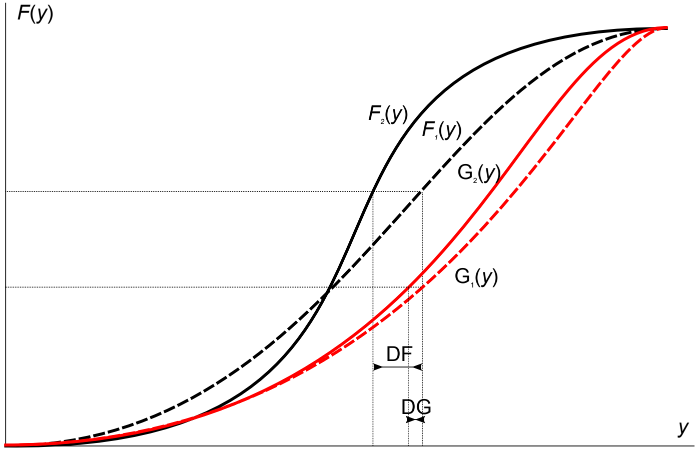
1. Fix the outcome level y, giving the quantiles in the two groups pre-reform, $F_1(y)$ and $G_1(y)$
2. Counterfactual post-reform outcome at y in the treatment group:

$$\begin{aligned} k^{CIC}(y) &= F_1^{-1}\left[F_1(y)\right] + \Delta^{CIC} \\ &= y + \left(G_2^{-1}\left[G_1(y)\right] - G_1^{-1}\left[G_1(y)\right]\right) \\ &= G_2^{-1}\left[G_1(y)\right] \end{aligned}$$

3. The CIC-estimate at level y is then

$$F_2^{-1}(y) - k^{CIC}(y)$$

CIC: Graphical representation



Common trend in outcome at the same quantile value in the treatment and comparison group ⇒ invariance wrt mon. trans

Including covariates: The problem

Conditional quantile regressions finds the effect of a variable for a given value of all other variables. That is, it finds effects on the *conditional* quantile

Unfortunately, the effects on conditional quantiles do *not* average up to the effect on the unconditional quantile

- thus, difficult to interpret estimates from conditional quantile regressions (see Firpo, 2007, Econometrica)

For linear regressions, including covariates is not a problem since

$$\begin{aligned} E[Y|X] &= X'\beta \\ E[Y] &= \int X'\beta dF_x(X) = E\left[X'\right] \cdot \beta \end{aligned}$$

and the effect on the conditional mean is also the effect on the unconditional mean

⇒But this property is *not* shared by the quantile

Nonlinear DiD with covariates

Two ways to estimate the effects on the unconditional quantiles with covariates:

1. Propensity score weighting:
   - use the propensity score to balance out the covariates before estimating quantile regression (Firpo, 2007, Econometrica)
     - used in QDiD
2. RIF-regression:
   - transform the problem by considering effects on population shares rather than quantiles (Firpo et al, 2010, Econometrica)
     - used in RIF-DiD

Not clear how to include covariates in CIC

RIF-DID with covariates

RIF-DID with covariates:

1. Generate a set of binary variables $I^y = 1\{Y_i > y\}$, where $F_1^{-1}(\tau) = y$
2. Estimate a prob model with DID-structure

$$I^y = g\left(\gamma_1^y D_i + \gamma_2^y G_i + \gamma_3^y D_i G_i + x_i' \beta^y + \epsilon_i^y\right)$$

3. Invert estimated average marginal effect using the empirical distribution,

$$\hat{\gamma}_3^y / \hat{f}_1(y)$$

where $\hat{f}_1(y)$ is a kernel estimate of the pdf in the treatment group pre-reform

Universal child care: The Norwegian case

Norway was among the first to introduce subsidized child care on a large scale

- Unique source to information about its long-run consequences

Exceptionally rich panel data set

- Covering the entire population from 1967 and onwards
  - Possible to link all parents to their children

Desirable institutional features for identification

- Homogenous population
- Unitary school system
- Similar availability, quality and spending level on local public services

Child care reform

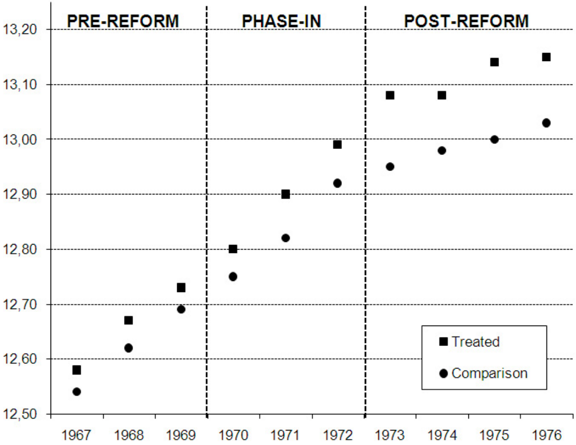
A major reform led to a large positive shock to the supply of subsidized care:

- From 1976 to 1979 coverage rates for 3 to 6 year olds grew by 18 percentage points on average, from 10% to 28%
- Largest supply shocks in municipalities where subsidized child care was mostly rationed before the reform
  - received higher federal subsidies

Havnes and Mogstad (2011a, AEJ: Policy; 2011b, JPubEc) use the staged expansion of subsidized child care induced by the reform:

- to estimate its mean impacts on (a) child outcomes and (b) maternal labor supply
- controlling for unobserved differences between children born in different years and children born in different municipalities

DiD: Graphical representation (years of schooling)

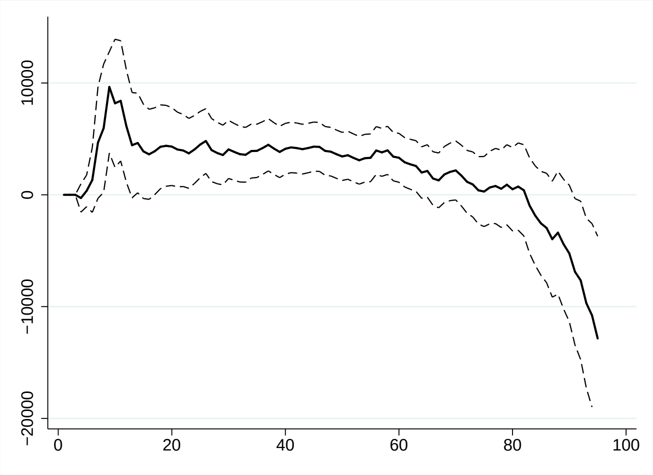


- Absolute terms: the extra 17,500 child care places $\Longrightarrow$ 6,200 years of schooling
- ATET estimate: .4 years of schooling per child care place
- ITT estimate: .07 years of schooling per child in treatment area

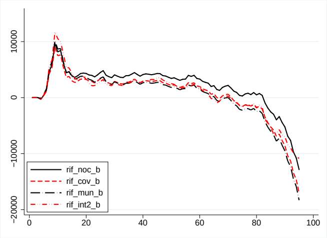
Descriptive statistics: Earnings (aged 30-36)

	Pre-reform	Pre-reform	Phase-in	Post-reform
5th percentile	0	0	0	0
10th percentile	31,685	-13	3,211	8,081
25th percentile	215,559	2,735	3,424	9,352
50th percentile	328,825	3,601	4,083	6,346
75th percentile	431,591	7,65	8,713	7,668
90th percentile	588,319	20,891	18,489	14,401
95th percentile	718,938	30,293	23,727	19,812
Mean (SD)	343,361 (270,402)			

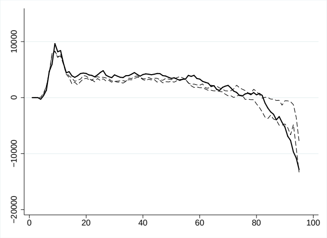
RIF-DID: QTE on earnings distribution (aged 30–36)



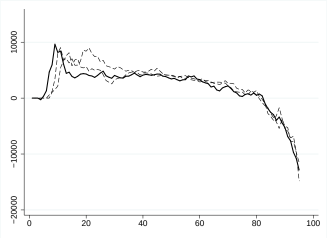
Specification checks



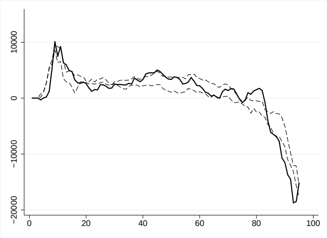
(a) Controls



(b) Alternative DiD models

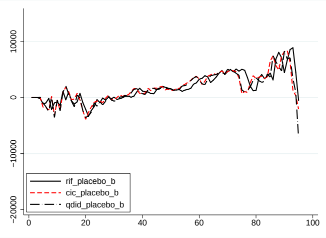


(c) Alternative earnings measures

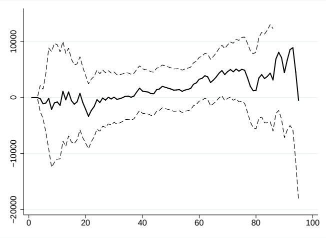


(d) Alternative treatment def.

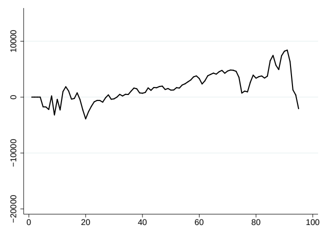
Robustness: Placebo reform



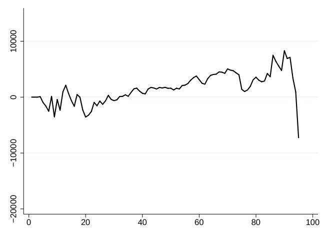
(e) Comparison



(f) RIF-DID



(g) CIC

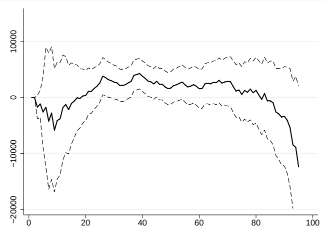


(h) QDID

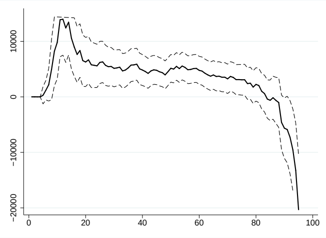
DiD: Mean impacts on earnings (aged 30–36)

	Estimate	SE	Mean	N
Overall	333.5	1,596.0	361,860	498,947
Child gender				
Boys	-977.7	2,503.9	440,020	253,677
Girls	631.3	1,616.3	281,020	245,270
Family income				
High	-2,047.5	3,005.7	393,094	195,081
Low	3,989.6***	1,855.7	341,807	303,866
Parental education				
Father's high	-3,172.1	3,003.3	397,914	186,365
— low	2,973.1*	1,812.0	340,363	312,582
Mother's high	-3,723.7	4,434.2	415,234	101,834
— low	2,374.7*	1,671.6	348,172	397,113

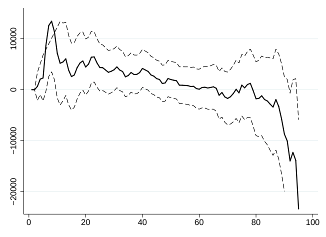
RIF-DID: Subsamples



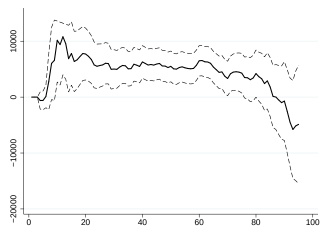
(i) Boys



(j) Girls

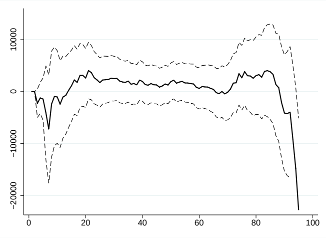


(k) Family income – high

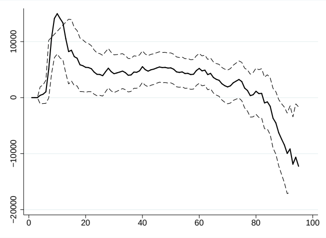


(l) Family income – low

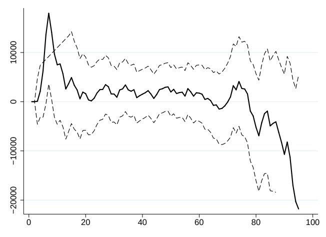
RIF-DID: Subsamples



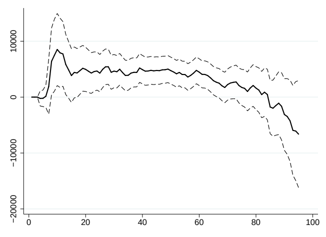
(m) Father's education – high



(n) Father's education – low

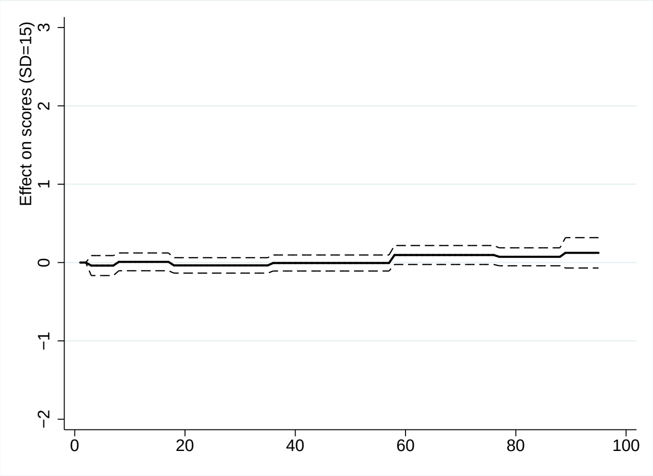


(o) Mother's education – high



(p) Mother's education – low

Intermediate outcomes: Cognitive vs non-cognitive abilitiy



Notes: Test scores are normalized to mean 100 and standard deviation 15. The sample consists of males only).

Interpretation: Theoretical framework

Say parents can affect their child's quality Q by investing in

1. quality of care q
2. market goods k

Specifically, assume a CES production function

$$Q = a\left(\omega q^{\lambda} + (1-\omega)\, k^{\lambda}\right)^{\tau/\lambda}, \quad \lambda < 1,\ \tau \leq 1,\ \omega \in (0,1) \qquad (1)$$

- $\tau = 1$: constant returns to scale. $\tau < 1$: decreasing returns to scale.
- $\lambda \uparrow$, q and k become closer substitutes. $\lambda \to -\infty$ approaches Leontief.

Theoretical framework

Assume

- quality of care q and child goods k can be purchased in the market at prices p_q and p_k, respectively
- parents have a total time endowment L and potential wage w
- c (numeraire) consumption/leisure that does not affect child quality

Theoretical framework

Assume

- quality of care q and child goods k can be purchased in the market at prices p_q and p_k, respectively
- parents have a total time endowment L and potential wage w
- c (numeraire) consumption/leisure that does not affect child quality

Family budget is then

$$wL = c + p_q q + p_k k \tag{2}$$

Theoretical framework

Assume

- quality of care q and child goods k can be purchased in the market at prices p_q and p_k, respectively
- parents have a total time endowment L and potential wage w
- c (numeraire) consumption/leisure that does not affect child quality

Family budget is then

$$wL = c + p_q q + p_k k \qquad (2)$$

Finally, let the family objective function be CRRA,

$$u(c, Q) = (\alpha c^{\rho} + (1-\alpha) Q^{\rho})^{1/\rho}, \quad \rho < 1, \alpha \in (0,1) \qquad (3)$$

- $\rho \uparrow$ implies consumption c and child quality Q are closer substitutes for the family.

Theoretical framework: No subsidized child care

Parents equate marginal costs, such that

$$k = \left(\frac{1-\omega}{\omega}\right)^{\frac{1}{1-\lambda}} q, \qquad q = b\left(\omega, \lambda\right)\left(\frac{Q}{a}\right)^{1/\tau}$$

Theoretical framework: No subsidized child care

Parents equate marginal costs, such that

$$k = \left(\frac{1-\omega}{\omega}\right)^{\frac{1}{1-\lambda}} q, \qquad q = b(\omega, \lambda)\left(\frac{Q}{a}\right)^{1/\tau}$$

Using the budget, this gives

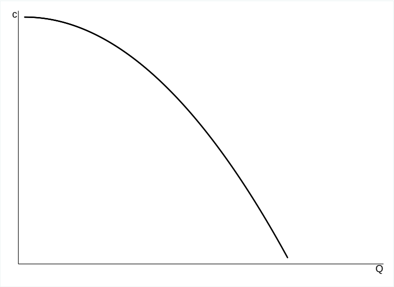
$$wL = c + p_Q(Q)$$

where

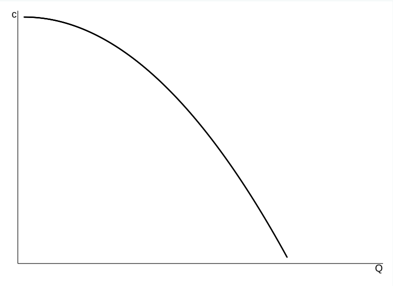
$$p_Q(Q) = b(\omega, \lambda)\left(\frac{Q}{a}\right)^{1/\tau}\left(p_q + p_k\left(\frac{1-\omega}{\omega}\right)^{\frac{1}{1-\lambda}}\right)$$

- $p_Q(Q)$ is the total cost in consumption units of producing Q
- which is increasing and convex

Theoretical framework: No subsidized child care



Theoretical framework: No subsidized child care

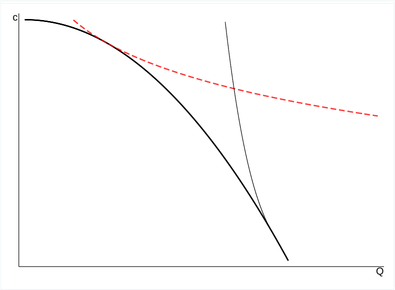


Families equate

- marginal benefit of child quality
- marginal cost of provision in terms of foregone consumption

$$\frac{1-\alpha}{\alpha}\left(\frac{c}{Q}\right)^{1-\rho} = p_Q^{'}(Q)$$

Theoretical framework: No subsidized child care



Families equate

- marginal benefit of child quality
- marginal cost of provision in terms of foregone consumption

$$\frac{1-\alpha}{\alpha}\left(\frac{c}{Q}\right)^{1-\rho} = p_Q^{'}(Q)$$

Theoretical framework: With subsidized child care

1. Provides child care *of a particular quality*, say q_f, at a lower price $p_q^f < p_q$
2. Families may choose *not to use* subsidized child care

Using subsidized child care locks in the care quality, *higher or lower* than before

- family may partially offset this with child goods k

Theoretical framework: With subsidized child care

1. Provides child care *of a particular quality*, say q_f, at a lower price $p_q^f < p_q$
2. Families may choose *not to use* subsidized child care

Using subsidized child care locks in the care quality, *higher or lower* than before

- family may partially offset this with child goods k

The total cost of child quality is now

$$p^f(Q) = p_q^f q_f + p_k \left[\left(\frac{Q}{a} \right)^{\lambda/\tau} \frac{1}{1-\omega} - \omega q_f^{\lambda} \right]^{1/\lambda}$$

Theoretical framework: With subsidized child care

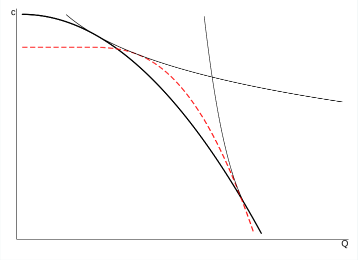
Note that subsidized child care is not free

- $p^f(0) > p(0) = 0$ implies
  - $\max c > \max c^f$
- $p_q^f < p_q$ implies
  - there are c on the frontier such that $c < c^f$
  - e.g. if $q^* = q^f$, then clearly $c^* < c^f$

Theoretical framework: With subsidized child care

Note that subsidized child care is not free

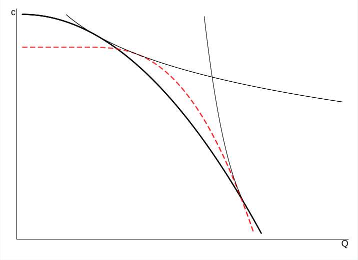
- $p^f(0) > p(0) = 0$ implies
  - $\max c > \max c^f$
- $p^f_q < p_q$ implies
  - there are c on the frontier such that $c < c^f$
  - e.g. if $q^* = q^f$, then clearly $c^* < c^f$



Theoretical framework: With subsidized child care

Note that subsidized child care is not free

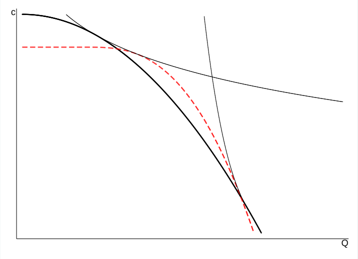
- $p^f(0) > p(0) = 0$ implies
  - $\max c > \max c^f$
- $p^f_q < p_q$ implies
  - there are c on the frontier such that $c < c^f$
  - e.g. if $q^* = q^f$, then clearly $c^* < c^f$
- $\max Q^f$ depends on
  1. quality of and price of subsidized care, q^f and p_q
  2. substitutability with k

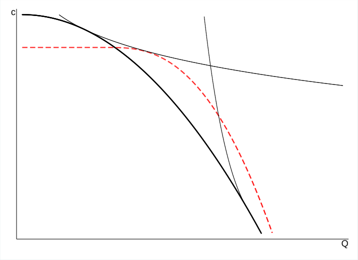


Theoretical framework: With subsidized child care

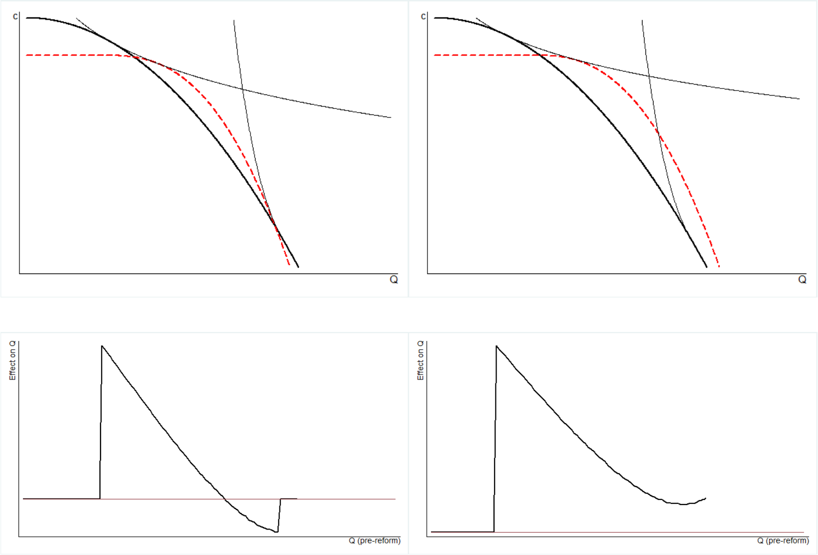
Note that subsidized child care is not free

- $p^f(0) > p(0) = 0$ implies
  - $\max c > \max c^f$
- $p^f_q < p_q$ implies
  - there are c on the frontier such that $c < c^f$
  - e.g. if $q^* = q^f$, then clearly $c^* < c^f$
- $\max Q^f$ depends on
  1. quality of and price of subsidized care, q^f and p_q
  2. substitutability with k





Theoretical framework: Predicted effects



Concluding remarks

Nonlinear DiD methods can be used to estimate

- the counterfactual outcome distribution in the absence of the policy intervention
- and compare it to the actual outcome distribution when subject to the policy intervention

Estimating the counterfactual distribution is useful to assess:

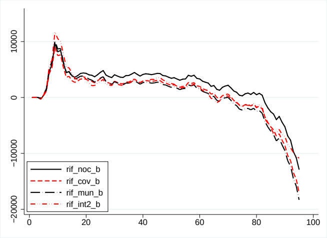
- responses when theory makes heterogenous predictions
  - e.g. Bitler et al. (2005, AER)
- distributional effects of public policy, e.g.
  - a policy that reduces inequality may be socially desirable, even if there is zero or even negative mean impact

Concluding remarks

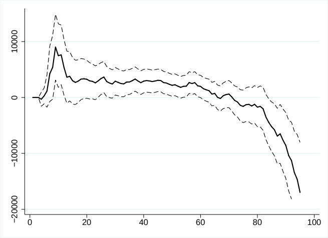
Our study suggests that:

1. subsidized child care has:
    - positive effects in the lower part of the distribution
    - negative effects in the uppermost part
        - mean impacts miss a lot
        - targeted policies may be preferable
2. the policymaker has to be quite inequality averse to
    - conclude that the introduction of universal child care reform improved child outcomes

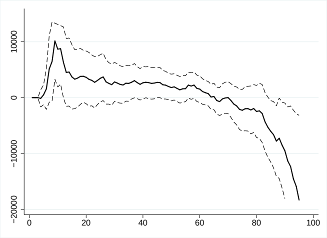
Robustness: Covariates



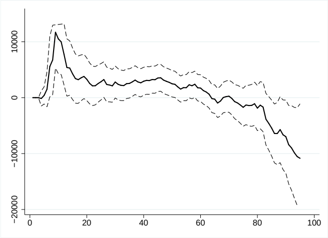
(q) Comparison



(r) Covariates

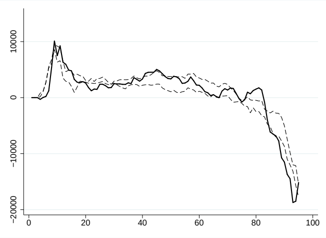


(s) Municipality FE

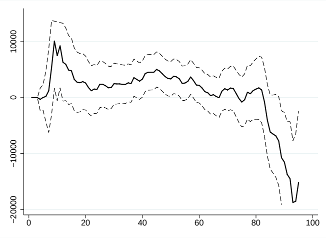


(t) Flexible trends

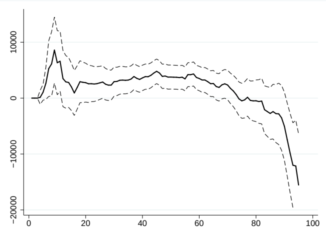
Robustness: Treatment definition



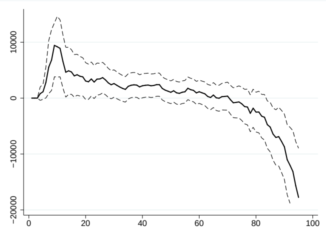
(u) Comparison



(v) 1976–1979, 33rd vs 67th



(w) 1977–1979, 50th vs 50th



(x) 1976–1978, 50th vs 50th