Bounding the Child Penalty

Julius Ilciukas

University of Amsterdam

Motivation

Children are the primary suspects behind gender gaps in the labor market.

Motivation

Children are the primary suspects behind gender gaps in the labor market.

- Public discourse:
  - The New York Times: "The Gender Pay Gap Is Largely Because of Motherhood" (C. C. Miller, 2017)
  - Business Insider: "What's the major reason a woman might get paid less than men in the same field, and with the same education? Kids." (Kaplan, 2023)

Motivation

Children are the primary suspects behind gender gaps in the labor market.

- Public discourse:
  - The New York Times: "The Gender Pay Gap Is Largely Because of Motherhood" (C. C. Miller, 2017)
  - Business Insider: "What's the major reason a woman might get paid less than men in the same field, and with the same education? Kids." (Kaplan, 2023)
- Academic discourse:
  - "Not surprisingly, children are the main contributors to women's labor supply changes." (Goldin, 2014)
  - "The effects of parenthood (...) account for most of the observed gender inequality in labor market outcomes" (Kleven et al., 2023)

Identification Challenge

Identification Challenge

Fertility is endogenous

- Human capital, wealth, health, career prospects, the cost of parenthood

Identification Challenge

Fertility is endogenous

- Human capital, wealth, health, career prospects, the cost of parenthood

Effects of parenthood are dynamic

- Time spent in parenthood, number and age of children, career stage at and during parenthood.

Identification Challenge

Fertility is endogenous

- Human capital, wealth, health, career prospects, the cost of parenthood

Effects of parenthood are dynamic

- Time spent in parenthood, number and age of children, career stage at and during parenthood.

Existing literature only addresses each separately

Existing Approaches

"Some of the most compelling evidence of the crucial role children (. . .) has been produced over the past few years" (Bertrand, 2020)

Existing Approaches

"Some of the most compelling evidence of the crucial role children (. . .) has been produced over the past few years" (Bertrand, 2020)

- Event study (Kleven, Landais, & Søgaard, 2019)
  - Compare mothers to yet-to-be mothers controlling for age
  - If fertility is random $\rightarrow$ unbiased estimates

Existing Approaches

"Some of the most compelling evidence of the crucial role children (. . .) has been produced over the past few years" (Bertrand, 2020)

- Event study (Kleven, Landais, & Søgaard, 2019)
  - Compare mothers to yet-to-be mothers controlling for age
  - If fertility is random $\rightarrow$ unbiased estimates
- IV using in vitro fertilization (Lundborg et al., 2017)
  - Treatment success as a natural experiment
  - If effect is static $\rightarrow$ unbiased estimates

Existing Approaches

"Some of the most compelling evidence of the crucial role children (...) has been produced over the past few years" (Bertrand, 2020)

- Event study (Kleven, Landais, & Søgaard, 2019)
  - Compare mothers to yet-to-be mothers controlling for age
  - If fertility is random $\rightarrow$ unbiased estimates
- IV using in vitro fertilization (Lundborg et al., 2017)
  - Treatment success as a natural experiment
  - If effect is static $\rightarrow$ unbiased estimates

Literature short Literature long

IV-IVF

IV-IVF

- Couples enter assisted fertility treatment to conceive a child

IV-IVF

- Couples enter assisted fertility treatment to conceive a child
- Embryos transferred into uterus

IV-IVF

- Couples enter assisted fertility treatment to conceive a child
- Embryos transferred into uterus
- Success as-good-as-random (cond. on important determinants)

IV-IVF

- Couples enter assisted fertility treatment to conceive a child
- Embryos transferred into uterus
- Success as-good-as-random (cond. on important determinants)
- Random variation in parenthood

IV-IVF

- Couples enter assisted fertility treatment to conceive a child
- Embryos transferred into uterus
- Success as-good-as-random (cond. on important determinants)
- Random variation in parenthood
- Compare women whose first treatment succeeds to women whose first treatment fails.

What does IV-IVF identify?

What does IV-IVF identify?

Linear combination of two effects:

What does IV-IVF identify?

Linear combination of two effects:

- Effect of parenthood (women who remain childless after failure)

What does IV-IVF identify?

Linear combination of two effects:

- Effect of parenthood (women who remain childless after failure)
- Effect of expediting parenthood (women who conceive after failure)

What does IV-IVF identify?

Linear combination of two effects:

- Effect of parenthood (women who remain childless after failure)
- Effect of expediting parenthood (women who conceive after failure)
- Example: 200 women enter treatment

What does IV-IVF identify?

Linear combination of two effects:

- Effect of parenthood (women who remain childless after failure)
- Effect of expediting parenthood (women who conceive after failure)
- Example: 200 women enter treatment
  - 100 first attempt succeeds
    - 5 years after: all mothers for 5 years

What does IV-IVF identify?

Linear combination of two effects:

- Effect of parenthood (women who remain childless after failure)
- Effect of expediting parenthood (women who conceive after failure)
- Example: 200 women enter treatment
  - 100 first attempt succeeds
    - 5 years after: all mothers for 5 years
  - 100 first attempt fails
    - 5 years after: some childless, some mothers for <5 years

Motherhood Among Unsuccessfully Treated

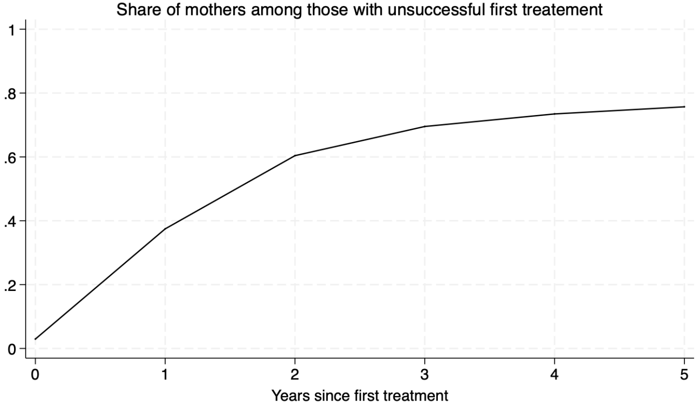


Figure 1: Motherhood among unsuccessfully treated

Motherhood Among Unsuccessfully Treated

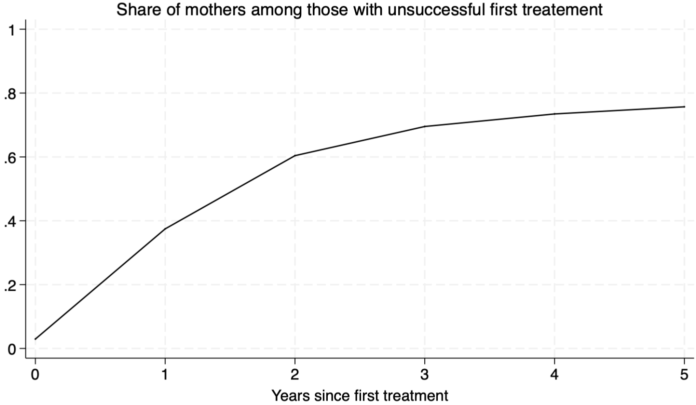


Figure 1: Motherhood among unsuccessfully treated

$$\tau_{RF} = 0.25\tau_{Parenthood} - 0.75\tau_{Delay}$$

Motherhood Among Unsuccessfully Treated

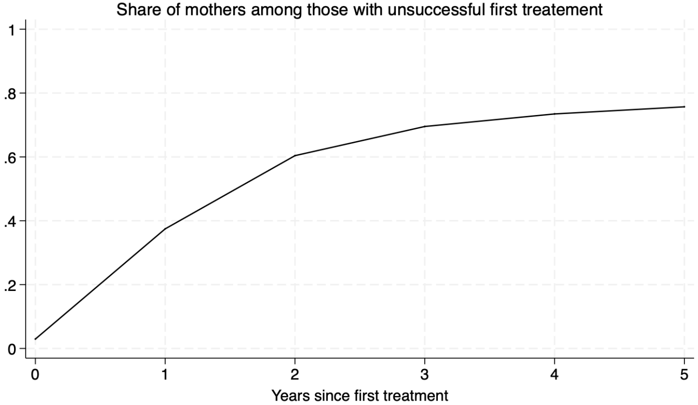


Figure 1: Motherhood among unsuccessfully treated

$$\tau_{RF} = 0.25\tau_{Parenthood} - 0.75\tau_{Delay}$$
$$\tau_{IV} = \tau_{Parenthood} - 3\tau_{Delay}$$

How much can we say about the causal effect of parenthood?

This Paper

- Minimal assumptions $\rightarrow$ see what is possible

This Paper

- Minimal assumptions $\rightarrow$ see what is possible
- Only crucial assumption: randomness in treatment success

This Paper

- Minimal assumptions $\rightarrow$ see what is possible
- Only crucial assumption: randomness in treatment success
- No restrictions on:

This Paper

- Minimal assumptions $\rightarrow$ see what is possible
- Only crucial assumption: randomness in treatment success
- No restrictions on:
  - Selection into treatment attempts
  - Dynamic structure/heterogeneity in effects
  - Non-treatment (endogenous) fertility

This Paper

- Minimal assumptions $\rightarrow$ see what is possible
- Only crucial assumption: randomness in treatment success
- No restrictions on:
  - Selection into treatment attempts
  - Dynamic structure/heterogeneity in effects
  - Non-treatment (endogenous) fertility
- Identification approach:
  - Exploit first **and later** treatments to handle treatment births
  - Bounds to handle non-treatment births

This Paper

- Minimal assumptions $\rightarrow$ see what is possible
- Only crucial assumption: randomness in treatment success
- No restrictions on:
  - Selection into treatment attempts
  - Dynamic structure/heterogeneity in effects
  - Non-treatment (endogenous) fertility
- Identification approach:
  - Exploit first **and later** treatments to handle treatment births
  - Bounds to handle non-treatment births
- Narrowing bounds

This Paper

- Minimal assumptions $\rightarrow$ see what is possible
- Only crucial assumption: randomness in treatment success
- No restrictions on:
  - Selection into treatment attempts
  - Dynamic structure/heterogeneity in effects
  - Non-treatment (endogenous) fertility
- Identification approach:
  - Exploit first **and later** treatments to handle treatment births
  - Bounds to handle non-treatment births
- Narrowing bounds
- Application to Dutch data

This Paper

- Minimal assumptions $\rightarrow$ see what is possible
- Only crucial assumption: randomness in treatment success
- No restrictions on:
  - Selection into treatment attempts
  - Dynamic structure/heterogeneity in effects
  - Non-treatment (endogenous) fertility
- Identification approach:
  - Exploit first **and later** treatments to handle treatment births
  - Bounds to handle non-treatment births
- Narrowing bounds
- Application to Dutch data
  - Sizable effect on women's outcomes in short-medium run

This Paper

- Minimal assumptions $\rightarrow$ see what is possible
- Only crucial assumption: randomness in treatment success
- No restrictions on:
  - Selection into treatment attempts
  - Dynamic structure/heterogeneity in effects
  - Non-treatment (endogenous) fertility
- Identification approach:
  - Exploit first **and later** treatments to handle treatment births
  - Bounds to handle non-treatment births
- Narrowing bounds
- Application to Dutch data
  - Sizable effect on women's outcomes in short-medium run
  - Explains up to half of post-child gender inequality
  - **Both** event study and IV might **overstate** the negative effect
  - Delaying motherhood might have a positive effect

Model

Model

- Time starts at first attempt: $t = 1, \ldots, T$
- Motherhood history: vector of length T where element C_t take value 1 if woman is a mother in period t and 0 otherwise

Model

- Time starts at first attempt: $t = 1, \ldots, T$
- Motherhood history: vector of length T where element C_t take value 1 if woman is a mother in period t and 0 otherwise
- Each motherhood history associated with different potential outcome

$$Y_t = \left(\begin{pmatrix} C_1 & C_2 & C_3 & \ldots & C_T \end{pmatrix}\right)$$

Model

- Time starts at first attempt: $t = 1, \ldots, T$
- Motherhood history: vector of length T where element C_t take value 1 if woman is a mother in period t and 0 otherwise
- Each motherhood history associated with different potential outcome

$$Y_t = \left(\begin{pmatrix} C_1 & C_2 & C_3 & \ldots & C_T \end{pmatrix}\right)$$

$$\begin{gathered}
Y_t\left(\begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \end{pmatrix}\right) \\
Y_t\left(\begin{pmatrix} 0 & 1 & 1 & \ldots & 1 \end{pmatrix}\right) \\
\vdots \\
Y_t\left(\begin{pmatrix} 0 & 0 & 0 & \ldots & 0 \end{pmatrix}\right)
\end{gathered}$$

For simplicity we are at $t = T$.

Model (cont.)

Model (cont.)

Women differ in:

- How many attempts would undergo if all failed ($G \in \{1, \ldots, \overline{g}\}$)?

Model (cont.)

Women differ in:

- How many attempts would undergo if all failed ($G \in \{1, \ldots, \overline{g}\}$)?
- Would enter motherhood if all attempts failed ($N \in \{0, 1\}$)?

Model (cont.)

Women differ in:

- How many attempts would undergo if all failed ($G \in \{1, \ldots, \overline{g}\}$)?
- Would enter motherhood if all attempts failed ($N \in \{0, 1\}$)?
- **Both are endogenous but independent of success:**

Model (cont.)

Women differ in:

- How many attempts would undergo if all failed ($G \in \{1, \dots, \overline{g}\}$)?
- Would enter motherhood if all attempts failed ($N \in \{0, 1\}$)?
- **Both are endogenous but independent of success:**

$$(Y(\mathbf{1}), Y(\mathbf{0}), N, G) \perp W_j | P \geq j$$

- P - number of realized attempts (observed)
- W_j - success of realized attempt j (observed)

Model (cont.)

Women differ in:

- How many attempts would undergo if all failed ($G \in \{1, \dots, \overline{g}\}$)?
- Would enter motherhood if all attempts failed ($N \in \{0, 1\}$)?
- **Both are endogenous but independent of success:**

$$(Y(\mathbf{1}), Y(\mathbf{0}), N, G) \perp W_j | P \geq j$$

- P - number of realized attempts (observed)
- W_j - success of realized attempt j (observed)

Object of interest:

$$\mathrm{E}[Y(\mathbf{1}) - Y(\mathbf{0}) | N = 0]$$

Effect of motherhood for women reliant on treatments to conceive.

Intuition

1. Use randomness at each treatment attempt to construct on average representative counterfactual where all attempts fail

Intuition

1. Use randomness at each treatment attempt to construct on average representative counterfactual where all attempts fail
    1.1 For women who depend on treatments: expected childless outcome

Intuition

1. Use randomness at each treatment attempt to construct on average representative counterfactual where all attempts fail
   1.1 For women who depend on treatments: expected childless outcome
   1.2 Share of women who depend on treatments

Intuition

1. Use randomness at each treatment attempt to construct on average representative counterfactual where all attempts fail
   1.1 For women who depend on treatments: expected childless outcome
   1.2 Share of women who depend on treatments
2. Use randomness at first attempt to "identify" the distribution of counterfactual outcomes where all first attempts succeed

Intuition

1. Use randomness at each treatment attempt to construct on average representative counterfactual where all attempts fail
   1.1 For women who depend on treatments: expected childless outcome
   1.2 Share of women who depend on treatments
2. Use randomness at first attempt to "identify" the distribution of counterfactual outcomes where all first attempts succeed
   2.1 Assume all women who depend on treatments are in the tails

Intuition

1. Use randomness at each treatment attempt to construct on average representative counterfactual where all attempts fail
   1.1 For women who depend on treatments: expected childless outcome
   1.2 Share of women who depend on treatments
2. Use randomness at first attempt to “identify” the distribution of counterfactual outcomes where all first attempts succeed
   2.1 Assume all women who depend on treatments are in the tails
   2.2 Bound their the expected motherhood outcome

Intuition

1. Use randomness at each treatment attempt to construct on average representative counterfactual where all attempts fail
   1.1 For women who depend on treatments: expected childless outcome
   1.2 Share of women who depend on treatments
2. Use randomness at first attempt to "identify" the distribution of counterfactual outcomes where all first attempts succeed
   2.1 Assume all women who depend on treatments are in the tails
   2.2 Bound their the expected motherhood outcome
3. Bounds on the effect combining 1.1 and 2.2

Intuition

1. Use randomness at each treatment attempt to construct on average representative counterfactual where all attempts fail
   1.1 For women who depend on treatments: expected childless outcome
   1.2 Share of women who depend on treatments
2. Use randomness at first attempt to “identify” the distribution of counterfactual outcomes where all first attempts succeed
   2.1 Assume all women who depend on treatments are in the tails
   2.2 Bound their the expected motherhood outcome
3. Bounds on the effect combining 1.1 and 2.2

(4.) Exploit pre-treatment covariates to narrow the bounds

Graph intuition | Math intuition | Detailed intuition | Detailed math | Trimming intuition

Assisted Fertility Treatments

Assisted Fertility Treatments

In vitro fertilization (IVF)

- Relatively invasive procedure performed under sedation/anesthesia.
- $\sim$25% success rate.

Assisted Fertility Treatments

In vitro fertilization (IVF)

- Relatively invasive procedure performed under sedation/anesthesia.
- $\sim$25% success rate.

Intrauterine insemination (IUI)

- Sperm injected directly into the uterus.
- $\sim$10% success rate.
- First-line infertility treatment in most countries.

Assisted Fertility Treatments

In vitro fertilization (IVF)

- Relatively invasive procedure performed under sedation/anesthesia.
- $\sim$25% success rate.

Intrauterine insemination (IUI)

- Sperm injected directly into the uterus.
- $\sim$10% success rate.
- First-line infertility treatment in most countries.

I exploit the moment embryos/sperm are transferred into the uterus.

Is success as-good-as-random?

Data

- Administrative data from Statistics Netherlands.
  - Data on fertility treatments from 2013 to 2017.
  - Labor market outcomes from 2011 to 2021.

Data

- Administrative data from Statistics Netherlands.
  - Data on fertility treatments from 2013 to 2017.
  - Labor market outcomes from 2011 to 2021.
- Sample of opposite sex couples cohabiting before first IUI attempt.

Balance Treatment success

Results

Bounds

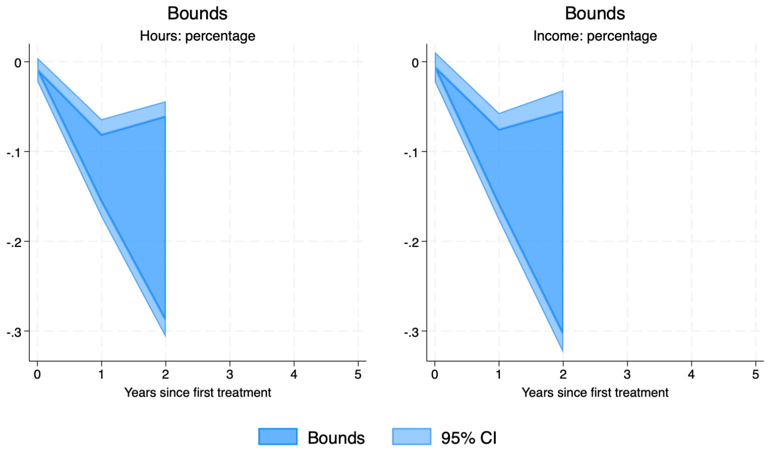


Figure 2: Bounds - short run

Baseline Lee bounds Absolute effects

Bounds

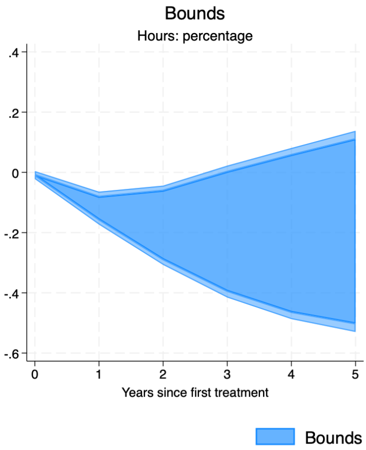


Figure 3: Bounds - medium run

Baseline Lee bounds Absolute effects

Monotonicity

Monotonicity

- Some women whose first treatment attempt succeeds eventually conceive more children without treatments

Monotonicity

- Some women whose first treatment attempt succeeds eventually conceive more children without treatments
- It may be reasonable to assume that they would have eventually conceived at least one child if all treatments had failed

Discussion

Monotone Bounds

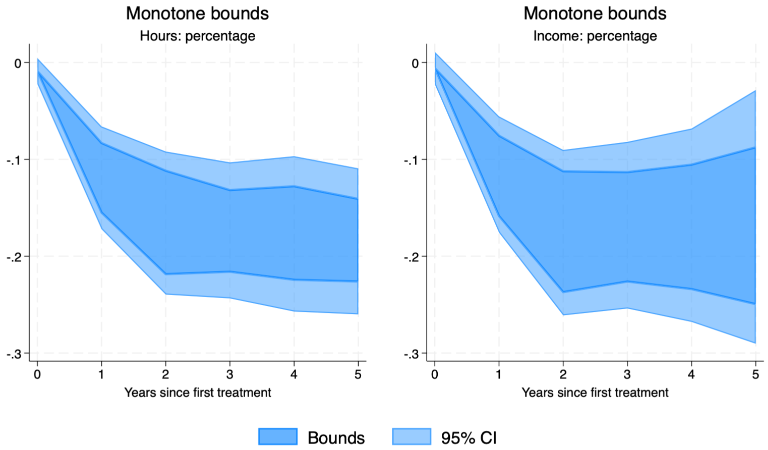


Figure 4: Monotone bounds for percent effects

Comparison to base Absolute How wide?

Extensions

- Fatherhood penalty Absolute Percentage
- Decomposing gender inequality Share explained by children
- Are existing estimates biased? IV-IVF equivalent Placebo event Placebo %
- Effect of delaying motherhood Bounds and delays Delays and effects
- Are estimates less informative that existing? Confidence intervals
- Bias due to depression Bounds for non-depressed
- Stable complier group Childless final period

Conclusion

- Estimator for the career cost of children that simultaneously address selection and dynamic effects

Conclusion

- Estimator for the career cost of children that simultaneously address selection and dynamic effects
  - Can be applied in other settings where individuals can re-apply for "treatment assignment" and/or enter treatment "endogenously" and where these treatments might "differ"

Conclusion

- Estimator for the career cost of children that simultaneously address selection and dynamic effects
  - Can be applied in other settings where individuals can re-apply for "treatment assignment" and/or enter treatment "endogenously" and where these treatments might "differ"
- Application to assisted fertility treatments:
  - Sizable motherhood penalty in the first few years

Conclusion

- Estimator for the career cost of children that simultaneously address selection and dynamic effects
  - Can be applied in other settings where individuals can re-apply for "treatment assignment" and/or enter treatment "endogenously" and where these treatments might "differ"
- Application to assisted fertility treatments:
  - Sizable motherhood penalty in the first few years
  - Under monotonicity: stable penalty after 5 years

Conclusion

- Estimator for the career cost of children that simultaneously address selection and dynamic effects
  - Can be applied in other settings where individuals can re-apply for "treatment assignment" and/or enter treatment "endogenously" and where these treatments might "differ"
- Application to assisted fertility treatments:
  - Sizable motherhood penalty in the first few years
  - Under monotonicity: stable penalty after 5 years
  - Responsible for 40%-50% of gender inequality in post-child work hours and 10%-40% of inequality in post-child earnings

Conclusion

- Estimator for the career cost of children that simultaneously address selection and dynamic effects
  - Can be applied in other settings where individuals can re-apply for "treatment assignment" and/or enter treatment "endogenously" and where these treatments might "differ"
- Application to assisted fertility treatments:
  - Sizable motherhood penalty in the first few years
  - Under monotonicity: stable penalty after 5 years
  - Responsible for 40%-50% of gender inequality in post-child work hours and 10%-40% of inequality in post-child earnings
- Comparison to existing approaches

Conclusion

- Estimator for the career cost of children that simultaneously address selection and dynamic effects
  - Can be applied in other settings where individuals can re-apply for "treatment assignment" and/or enter treatment "endogenously" and where these treatments might "differ"
- Application to assisted fertility treatments:
  - Sizable motherhood penalty in the first few years
  - Under monotonicity: stable penalty after 5 years
  - Responsible for 40%-50% of gender inequality in post-child work hours and 10%-40% of inequality in post-child earnings
- Comparison to existing approaches
  - IV-IVF might **overstate** penalty in the medium run

Conclusion

- Estimator for the career cost of children that simultaneously address selection and dynamic effects
  - Can be applied in other settings where individuals can re-apply for "treatment assignment" and/or enter treatment "endogenously" and where these treatments might "differ"
- Application to assisted fertility treatments:
  - Sizable motherhood penalty in the first few years
  - Under monotonicity: stable penalty after 5 years
  - Responsible for 40%-50% of gender inequality in post-child work hours and 10%-40% of inequality in post-child earnings
- Comparison to existing approaches
  - IV-IVF might **overstate** penalty in the medium run
  - Event study might overstate penalty in both short and medium run

Appendix

Related Literature

Gender inequality in labor market outcomes.

- Bertrand (2011); Blau & Kahn (2017) for review.
- Role of parenthood with selection and (partially) dynamic effects:
  - Bensnes et al. (2023); Gallen et al. (2023) Discussion
  - Static by time since birth and homogeneous across individuals.

No paper to date addresses endogenous timing and dynamic (and heterogeneous) effects.

Main methodological ideas closely related to:

- Van den Berg & Vikström (2022): sequential treatment assignment.
- Lee (2005); Zhang & Rubin (2003): bounds with missing data.

Back Literature

Bensnes et al. (2023); Gallen et al. (2023)

Main idea:

1. Estimate effect in first period after treatment (while there are no later-mothers)
2. For individuals who are treated in second period, plug in estimate from the first
3. Repeat for all periods . . .

Required (intuitive) assumptions:

1. Effect must be similar between women who do and who do not enter motherhood later
   - Hard to defend because these women differ a lot in observables
   - Might differ even more on unobservables
2. Effect cannot vary over the life-cycle
   - E.g. being a mother for 2 years at 25 should have the same effect as as being a mother for 2 years at 35
   - Hard to defend because these are different career stages

Back

ToC

Literature (cont.)

Parenthood, labor market outcomes, and gender equality:

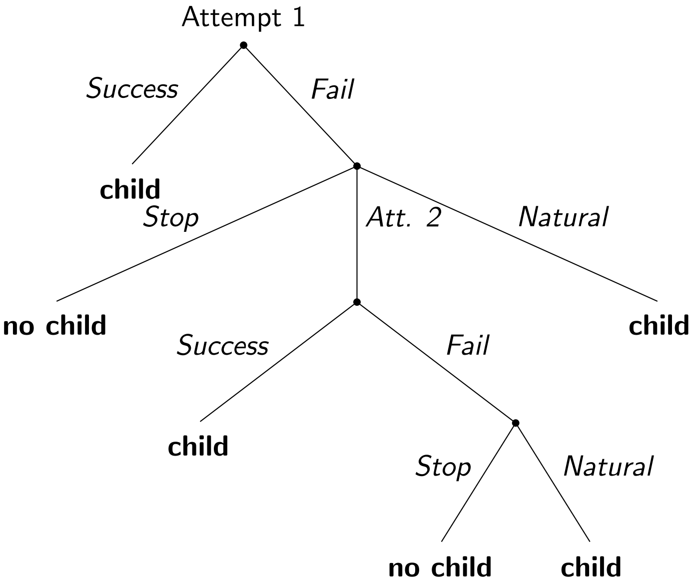
- Event studies: Angelov et al. (2016); Chung et al. (2017); Kleven, Landais, Posch, et al. (2019); Eichmeyer & Kent (2022),
- Abortion access: A. R. Miller (2011); Brooks & Zohar (2021)
- Infertility: Agüero & Marks (2008); Cristia (2008)
- Sibling sex mix: Angrist & Evans (1996); Iacovou (2001); Cruces & Galiani (2007); Maurin & Moschion (2009); Hirvonen (2009)
- Twins: Rosenzweig & Wolpin (1980); Bronars & Grogger (1994); Jacobsen et al. (1999); Vere (2011)
- Miscarriages: Hotz et al. (2005)

Methodological paper on dynamic compliance and treatment effects:

- Heckman et al. (2016); Han (2021); Van den Berg & Vikström (2022)

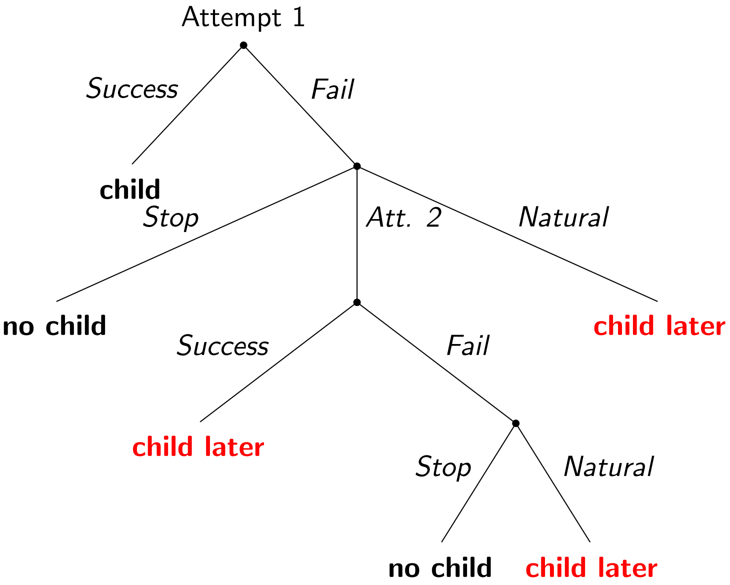
Back

Model Intuition



ToC

Model Intuition: Dynamic Effects



ToC

Wald Estimator

- Y - observed outcome
- C - motherhood dummy
- W - treatment success dummy

$$\frac{E[Y|W=1]-E[Y|W=0]}{E[C|W=1]-E[C|W=0]}$$

$$= E[Y(\text{mother}) - Y(\text{childless})|\text{complier}]$$

$$+\frac{1-\Pr(\text{complier})}{\Pr(\text{complier})}E[Y(\text{mother}) - Y(\text{mother})|\text{always-taker}]$$

Wald Estimator

- Y - observed outcome
- C - dummy for having children
- W - treatment success dummy

$$\frac{\mathrm{E}[Y|W=1]-\mathrm{E}[Y|W=0]}{\mathrm{E}[C|W=1]-\mathrm{E}[C|W=0]}$$

$$= \mathrm{E}[Y(\text{mother}) - Y(\text{childless})|\text{complier}]$$

$$+\frac{1-\Pr(\text{complier})}{\Pr(\text{complier})}\,\mathrm{E}[\underbrace{Y(\text{mother}) - Y(\text{mother})}_{=0}|\text{always-taker}]$$

$$= \tau_{ATC}$$

Wald Estimator: Dynamic Effect

Always-takers among unsuccessfully treated become mothers later:

$$\frac{E[Y|W=1]-E[Y|W=0]}{E[C|W=1]-E[C|W=0]}$$

$$= \tau_{ATC}$$

$$+\frac{1-\Pr(\text{complier})}{\Pr(\text{complier})}\, E[\underbrace{Y(\text{mother})-Y(\text{mother later})}_{\neq 0}|\text{always-taker}]$$

- Differ in career stage at childbirth
- Differ in age of children
- Differ in number of children

Model: Latent Variables

- Time: starts at first attempt, $t = 0, \ldots, T$.
- Types: pair (G, N):
  - G - how many treatment attempts woman would undergo if all previous attempts failed.
  - N_t - 1 if woman would conceive naturally by period t if all attempts were to fail and 0 otherwise.
- Treatment success: W_j indicator for potential success of treatment j.
  - "Potential" because attempt might never happen.
- Potential outcomes: vary by time since first treatment and the period in which woman becomes a mother:
  - $Y_t(k)$ - outcome in period t if woman becomes a mother in period k.
  - $Y_t(\mathbf{0})$ - outcome in period t if woman remains childless.

ToC

Model: Observed Variables

- Number of realized attempts - P.
- Success of first P attempts - $W_1, \ldots, W_P$.
- Indicator for having children - C_t.
- Realized outcome - Y_t.
  - Assuming all later treatment attempts and births happen in period t.
  - Assuming no anticipation: $Y_t(j) = Y_t(\mathbf{0})$ if $j > t$.

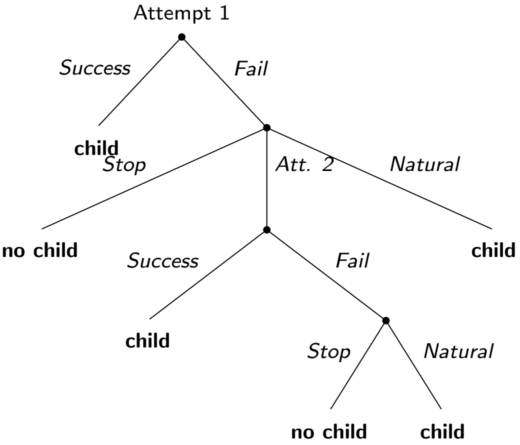
$$Y_t = \begin{cases} Y_t(\mathbf{1}), & \text{if } W_1 = 1 \\ Y_t(\mathbf{0}), & \text{if } \prod_{g=1}^{G}(1 - W_g) = 1 \text{ and } N_t = 0 \\ Y_t(t), & \text{otherwise.} \end{cases}$$

Model Implications

- Women who would conceive naturally cannot be observed childless.
  - If they are systematically different from the rest we cannot learn anything about their potential childless outcome.
- Effect of interest without additional assumptions:

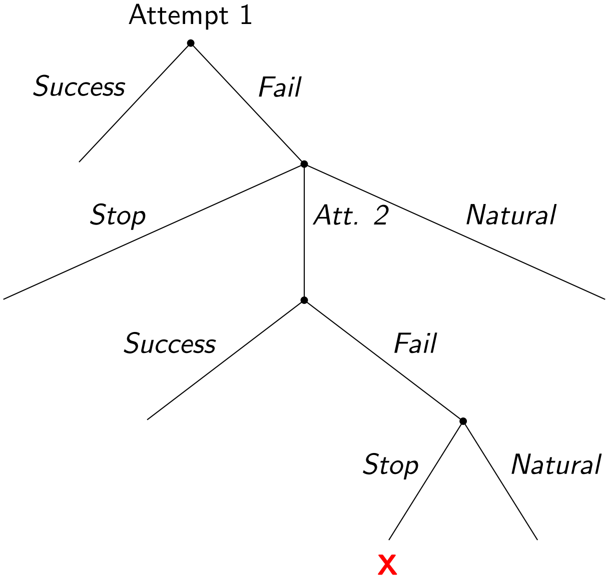
$$\mathrm{E}[Y_t(\mathbf{1}) - Y_t(\mathbf{0})|N_t = 0] \text{ for } t = 0, \ldots, T$$

Intuition: Childless Outcome $Y_t(\mathbf{0})$



1. Each childless node contains exclusively a specific type
2. If success is random, it is a random sub-sample of that type
3. Can identify expected childless outcome for that type

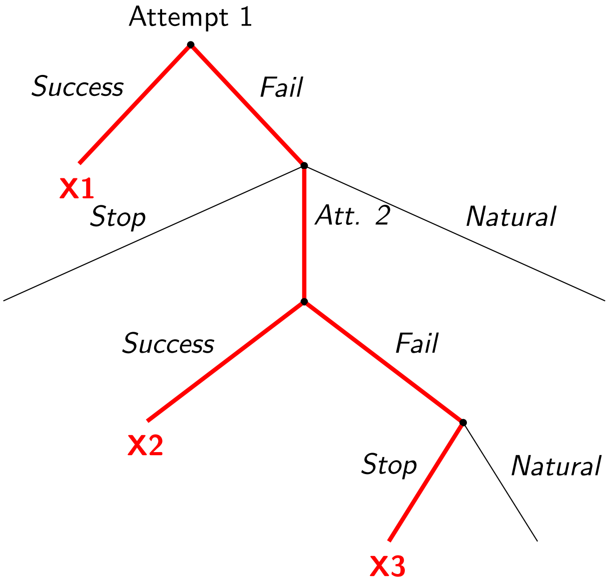
Intuition: Childless Outcome $Y_t(\mathbf{0})$ - Step 1



ToC X$\Rightarrow G = 2, N_t = 0$

Intuition: Childless Outcome $Y_t(\mathbf{0})$ - Step 2

Where can $G = 2, N_t = 0$ end up?

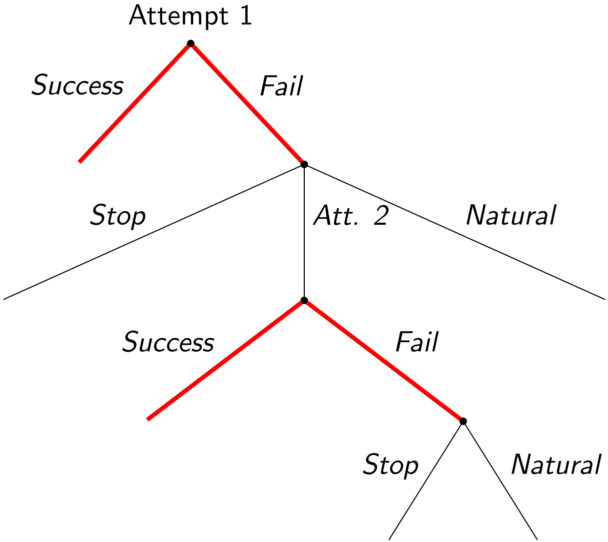


Identification Assumption

Identification assumption - (potential) treatment success is independent of potential outcomes, types, and potential succes of other attempts:

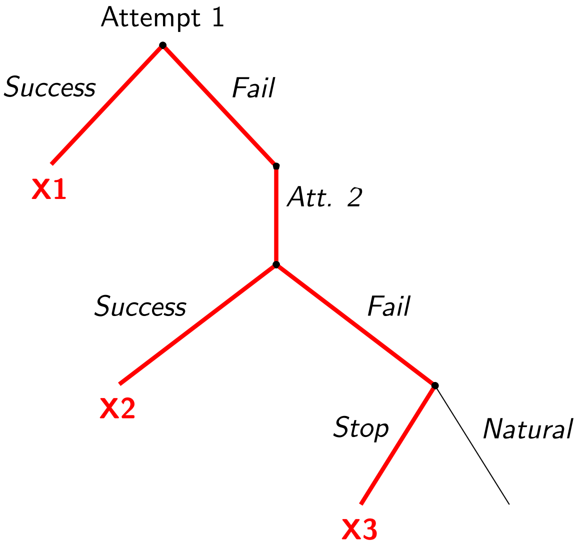
$$Y_t(k), W_l, G, N_t \perp W_j \forall j, k, l.$$

Intuition: Identification Assumption



Intuition: Childless Outcome $Y_t(\mathbf{0})$ - Step 2 (cont.)

Where can $G = 2, N_t = 0$ end up?



ToC

X3: random subsample of $G = 2, N_t = 0$ type, **observed childless**.

Intuition: Motherhood Outcome $Y_t(\mathbf{1})$

- Women successfully treated at first attempt are a random sample.
- Do not known which of them would have eventually conceived naturally.
- Do know the share that would have conceived naturally.
- Assume the most extreme scenarios consistent with the data.

Intuition: Motherhood Outcome $Y_t(\mathbf{1})$

1. Identify distribution of motherhood outcomes using women with successful first treatment.

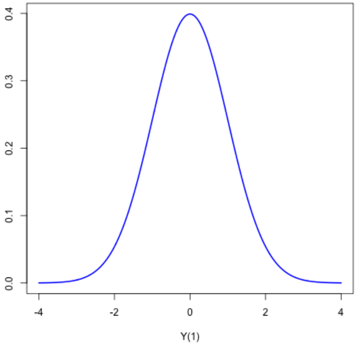


Figure 5: Distribution of potential motherhood outcomes

ToC

Intuition: Motherhood Outcome $Y_t(\mathbf{1})$ (cont.)

2. Estimate $\Pr(N_t = 0) = 0.9$.
3. Assume most extreme distributions of types.

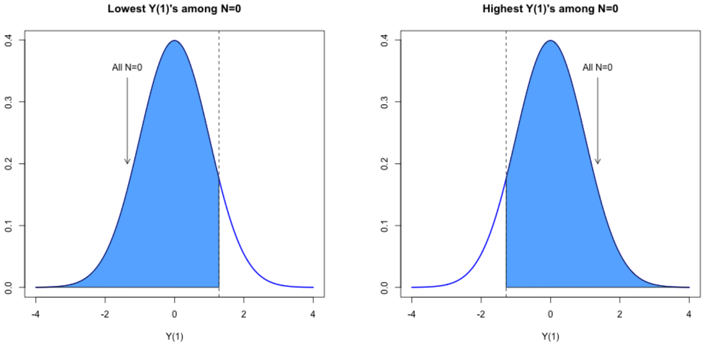


Figure 6: Distributions of potential motherhood outcomes

Intuition: Motherhood Outcome $Y_t(\mathbf{1})$ (cont.)

4. The means of the two trimmed distributions give bounds:

$$L \leq \mathrm{E}[Y_t(0)|N_t = 0] \leq U$$

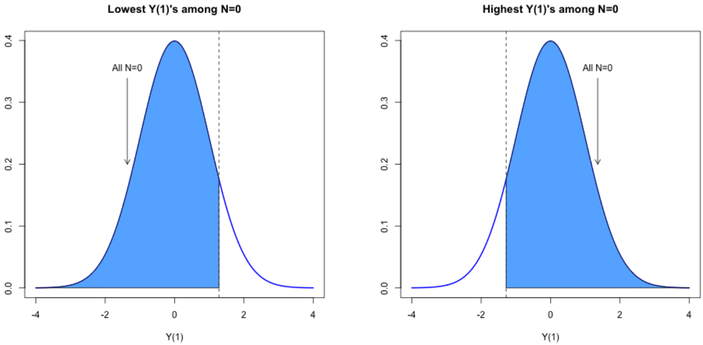


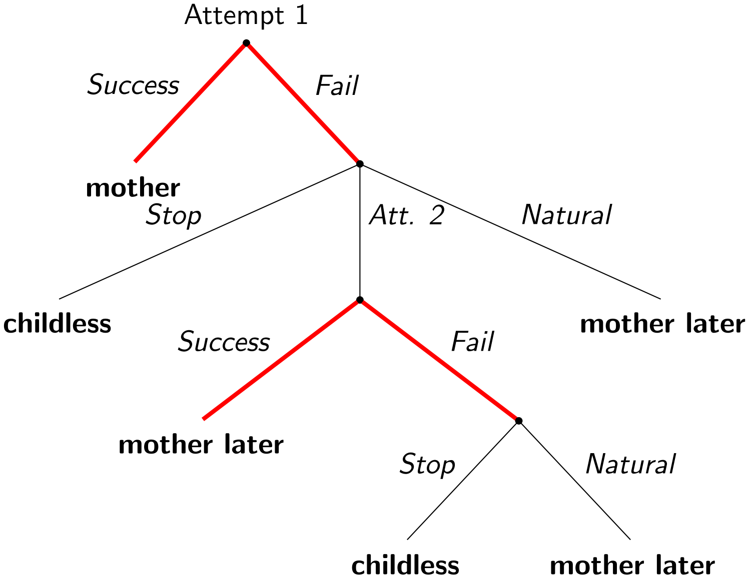
Figure 7: Distributions of potential outcomes in treatment group among $N_t = 0$'s

Identification: Effect of Motherhood

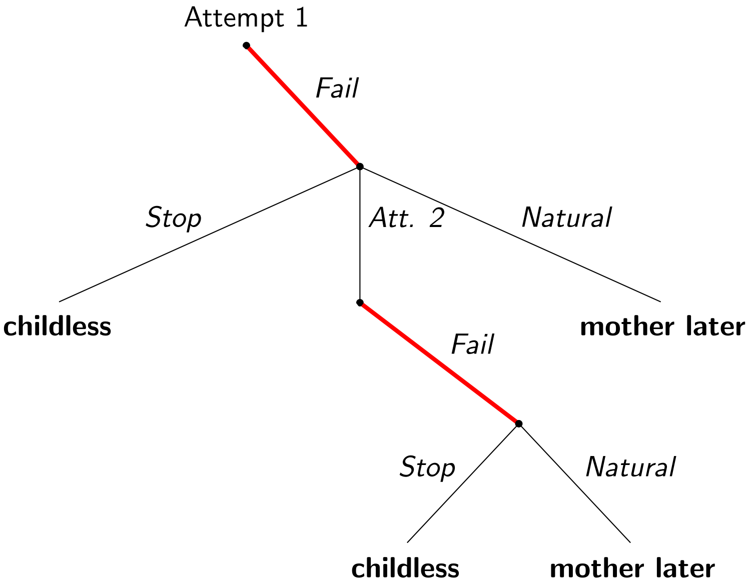
1. Point identified $\mathrm{E}[Y_t(\mathbf{0})|N_t = 0]$
2. Bounds on $\mathrm{E}[Y_t(\mathbf{1})|N_t = 0]$
3. Bounds on $\mathrm{E}[Y_t(\mathbf{1}) - Y_t(\mathbf{0})|N_t = 0]$

Back

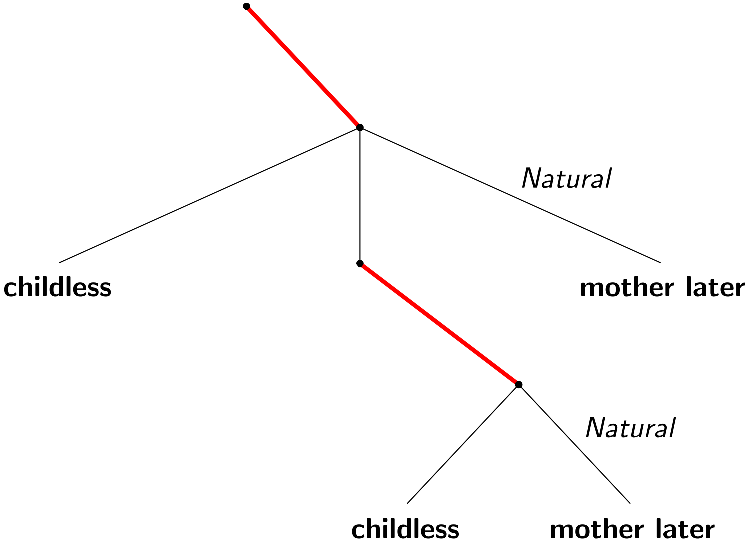
Identification Intuition: Graph



Identification Intuition: Graph



Identification Intuition: Graph



ToC Back

Identification Intuition (cont.)

Step 0:

$$
\begin{aligned}
\mathrm{E}[Y|C=0,P=j] &= \mathrm{E}[Y(\mathbf{0})|j \text{ fails}, N=0, G=j] \\
&= \mathrm{E}[Y(\mathbf{0})|N=0, G=j]
\end{aligned}
$$

Step 1:

$$
\mathrm{E}[w(P)Y|C=0] = \mathrm{E}[Y(\mathbf{0})|N=0],
$$

with:

$$
w(P) \propto \frac{1}{\prod_{g=1}^{P} \Pr(W_g=0|P \geq g)} \overset{\text{if Pr const.}}{=} \frac{1}{\alpha^P}.
$$

Step 2:

$$
\mathrm{E}[N] = \mathrm{E}\left[w(P)C|\text{no success}\right],
$$

Step 3:

1. Y's among $W_1 = 1$ "reveal" the distribution of $Y(\mathbf{1})'s$
2. Assume women with $N=0$ are in the left/right tail
3. Bounds on $\mathrm{E}[Y(\mathbf{1})|N=0]$ Graph int. Coins Det. int Trimming int. Back

ToC

Estimator Intuition: Math with Coins

- Each individuals flips a coin once
- Some may chose to flip again if heads come up
- Number of flips (P) observed
- Y only revealed for those who never flip heads

$$\mathrm{E}[Y] = \mathrm{E}\left[\frac{1}{(1/2)^P} Y \, \mathbf{1}\{\text{no heads}\}\right]$$

Back

Formal Identification

$$
\begin{aligned}
\Delta_L &= \mu_L - \mu_C \\
\Delta_U &= \mu_U - \mu_C \\
\mu_C &= \mathrm{E}\left[\frac{Y}{\prod_j^P (1 - p_j(X_j))}\middle| \mathbf{1}_{Child} = 0\right] \mathrm{E}\left[\prod_j^P (1 - p_j(X_j)) \middle| \mathbf{1}_{Child} = 0\right] \\
\mu_L &= \mathrm{E}\left[\frac{Y}{p_1(X_1)}\middle| W_1 = 1, Y < y(1-s)\right] \mathrm{E}\left[p_1(X_1) | W_1 = 1, Y < y(1-s)\right] \\
\mu_U &= \mathrm{E}\left[\frac{Y}{p_1(X_1)}\middle| W_1 = 1, Y > y(s)\right] \mathrm{E}\left[p_1(X_1) | W_1 = 1, Y > y(s)\right] \\
y(q) &= G^{-1}(q) \\
G(q) &= \mathrm{E}\left[\frac{\mathbf{1}(Y \leq q)}{p_1(X_1)}\middle| W_1 = 1\right] \mathrm{E}[p_1(X_1) | W_1 = 1] \\
s &= \mathrm{E}\left[\frac{\mathbf{1}_{Child}}{\prod_j^P (1 - p_j(X_j))}\middle| W = 0\right] \mathrm{E}\left[\prod_j^P (1 - p_j(X_j)) \middle| W = 0\right],
\end{aligned}
$$

where $W = 1 - \prod_{j=1}^P (1 - W_j)$.

Additional Details and Extensions (cont.)

- Conditional independence
  - Weights inversely proportional conditional probability of treatment failure.
  - Propensities estimated sequentially for each treatment.
  - I assume a logistic form, easy to do non-parametrically in my application.
- Using covariates to narrow bounds
  - Split sample based on discrete pre-treatment covariates and estimate bounds on each subsample (Lee)
    - I split the sample into 5 binds based on income.
  - "Regress out" covariates and estimate bounds on residuals and a point-identified fitted component (new?)
    - I "regress out" the outcome in the year before first treatment in each specification.
  - Could use modern approaches to Lee bounds that fully exploit all information contained in covariates (Semenova, 2020; Heiler, 2022).
- Asymptotic properties and inference
  - Proof following Lee (2005), I bootstrap standard errors to facilitate comparisons between estimators.

Conditional Independence

- All of the results extends to a setting with conditional independence:

$$Y_t(k), W_l, G, N_t, X_m \perp W_j | X_j \forall j, k, l, m \neq j,$$

- In estimation of the expected childless outcome $\Pr(W_j = 1)$ needs to be replaced by $\Pr(W_j = 1 | X_j)$
- New weight function:

$$w(P, X_1, \ldots, X_P) = \frac{c}{\prod_{g=1}^{P} \Pr(W_g = 0 | X_g)}.$$

- In estimation of the expected motherhood outcome the sample first needs to be reweighted using weights $\frac{1}{\Pr(W_1 = 1 | X_1)}$

Back

ToC

Narrowing the Bounds with Covariates

Take some covariates unaffected by treatment X

1. Discrete covariates in Lee (2005)
   - Split the sample based on some discrete covariate X.
   - Estimate the bounds in every group and aggregate across groups.
   - Intuition: Suppose 100 women from before are split into two groups based on some covariate. Suppose it turns out that the 20 women who would conceive naturally are equally distributed across the two groups. New bounds require picking 40 women with lowest/highest outcomes in each group of 50. This restricts the possible ways to pick the 80 women, which will generally narrow the bounds.

The two approaches can be combined. Back

Narrowing the Bounds with Covariates (cont.)

Take some covariates unaffected by treatment X

2. Extension of Lee bounds (to my knowledge)
    - Take some function $g(X)$
    - $\mathrm{E}[g(X)|N_t = 0]$ can be identified on women who remain childless using similar steps as before.
    - Take $\mathrm{E}[Y_t(\mathbf{1})|N_t = 0] = \mathrm{E}[g(X) + \varepsilon|N_t = 0]$
    - Only need to bound $\mathrm{E}[\varepsilon|N_t = 0]$
    - $g(X)$ can be directly chosen to minimize the spread in residuals, e.g. OLS of Y_t on X for women with $W_1 = 1$.
    - Since residuals typically have a narrower distribution, this gives narrower bounds.
    - Also, likely to improve asymptotic efficiency: only need to compare residuals in both groups, $\mathrm{E}[g(X)|N_t = 0]$ is the same for treated and control.

Back

ToC

Narrowest Bounds with Covariates

- The narrowest bounds that can be achieved require estimating $\Pr(N_t = 0|X = x)$ and $Q_{Y_t(\mathbf{1})}(x)$ where $Q(x)$ is the $\Pr(N_t = 0|X = x)$ conditional quantile of $Y_t(\mathbf{1})$ given x.
- Ideally, this should be done non-parametrically.
- Semenova (2020)
- Asymptotic results are readily available if $G = 1$.
- In my setting the bounds do not change much.

Back

ToC

Randomness in IUI and IVF

Treatment success is not completely random.

- Arguably the most important factor for success also related to labor market outcomes is the age of the woman and their partner.
- I allow the probability of success at each attempt to depend on the age of the woman and their partner at the time of the attempt interacted with treatment type.

Other (medical) factors relevant for IVF success appear to be largely unrelated to labor market outcomes (Lundborg et al., 2017).

Back

Treatment Success

- I impute treatment success as having a child born within 10 months after treatment without subsequent treatments in between.
  - Validated to reflect medical records of treatment success (Lundborg et al., 2017).

Back

Balance

Table 1: First treatment outcomes and descriptives

	Success (1)	Fail (2)	Difference (1)-(2)	Dif. cond. age FE (1)-(2) cond. age
Work (W)	0.881	0.863	0.018	0.008
	[0.324]	[0.344]	(0.009)	(0.009)
Work (P)	0.884	0.865	0.019	0.014
	[0.320]	[0.341]	(0.009)	(0.009)
Hours (W)	1239.696	1208.255	31.441	17.578
	[605.070]	[634.840]	(16.168)	(15.812)
Hours (P)	1473.383	1438.880	34.502	22.690
	[658.917]	[695.345]	(17.699)	(17.587)
Income 1000s € (W)	28.049	27.434	0.615	0.942
	[19.559]	[20.232]	(0.516)	(0.496)
Income 1000s € (P)	37.173	36.959	0.214	0.896
	[26.484]	[29.443]	(0.746)	(0.732)
Bachelor deg. (W)	0.608	0.605	0.002	0.018
	[0.488]	[0.489]	(0.013)	(0.012)
Bachelor deg. (P)	0.593	0.598	-0.004	0.008
	[0.491]	[0.490]	(0.013)	(0.012)
Age (W)	31.643	32.384	-0.741	
	[4.016]	[4.383]	(0.111)	
Age (P)	34.672	35.459	-0.787	
	[5.527]	[5.993]	(0.152)	
Observations	1,716	13,788		
Joint *p*-val.			0.000	0.536

Note: Labor market outcomes measured year before first treatment. (W) - woman, (P) - partner. Standard deviations in brackets. Standard errors in parentheses.

Balance in Later Treatments

Table 2: Balance in later treatments

	D2	D3	D4	D5	D6	D7	D8	D9	D10
Work (W)	0.013	-0.002	0.023	0.008	0.030	0.007	-0.008	0.016	0.041
	(0.009)	(0.010)	(0.011)	(0.012)	(0.013)	(0.014)	(0.017)	(0.019)	(0.026)
Work (P)	0.011	0.014	0.005	0.014	-0.004	-0.008	0.001	0.016	0.040
	(0.010)	(0.010)	(0.011)	(0.012)	(0.013)	(0.014)	(0.017)	(0.020)	(0.027)
Hours (W)	37.050	-0.615	45.477	39.327	68.596	25.780	-5.734	81.149	29.860
	(17.373)	(18.648)	(20.127)	(21.930)	(24.489)	(26.043)	(31.176)	(36.869)	(49.101)
Hours (P)	29.074	28.347	18.441	35.597	-7.332	-15.344	0.360	47.511	49.279
	(19.336)	(20.807)	(22.614)	(24.685)	(27.215)	(28.618)	(34.381)	(41.158)	(55.440)
Income 1000s € (W)	1.786	0.283	1.123	1.672	1.380	0.489	0.417	1.839	-0.297
	(0.548)	(0.592)	(0.647)	(0.710)	(0.786)	(0.831)	(1.030)	(1.240)	(1.714)
Income 1000s € (P)	0.221	1.277	1.588	1.125	-0.542	-0.370	1.567	1.001	-0.202
	(0.820)	(0.846)	(0.923)	(1.018)	(1.123)	(1.212)	(1.423)	(1.666)	(2.277)
Bachelor deg. (W)	0.002	0.026	-0.020	0.001	-0.003	0.003	0.023	-0.012	0.045
	(0.013)	(0.014)	(0.015)	(0.017)	(0.019)	(0.020)	(0.024)	(0.028)	(0.038)
Bachelor deg. (P)	0.005	0.010	0.011	0.007	-0.003	0.013	0.020	0.012	-0.014
	(0.013)	(0.014)	(0.016)	(0.017)	(0.019)	(0.020)	(0.024)	(0.029)	(0.039)
Age (W)	0.001	-0.007	-0.040	0.024	0.013	-0.001	-0.046	-0.027	-0.017
	(0.011)	(0.015)	(0.019)	(0.023)	(0.026)	(0.028)	(0.036)	(0.043)	(0.059)
Age (P)	0.001	-0.007	-0.040	0.024	0.013	-0.001	-0.046	-0.027	-0.017
	(0.011)	(0.015)	(0.019)	(0.023)	(0.026)	(0.028)	(0.036)	(0.043)	(0.059)
Observations	12,955	10,759	8,714	6,969	5,403	3,938	2,718	1,848	1,173
Joint *p*-val.	0.071	0.737	0.057	0.439	0.420	0.991	0.836	0.508	0.437

Note: Each column describes the differences between those treated successfully and those treated unsuccessfully at the respective treatment attempt conditional on fixed effects for age-at-treatment interacted with fixed effects for treatment type. Labor market outcomes and age measured year before first treatment. (W) - woman, (P) - partner. Standard errors in parentheses.

Representative and Relevant Treatment group

Table 3: Full sample, relevant sample, and representative sample

	Success (1)	Fail (2)	Relevant (3)	Representative (4)	Success vs rep. (1)-(4)	Rel. vs rep. (3)-(4)
Work (W)	0.881	0.863	0.822	0.801	0.080	0.021
	[0.324]	[0.344]	[0.334]	[0.399]	(0.010)	(0.005)
Work (P)	0.884	0.865	0.850	0.783	0.101	0.066
	[0.320]	[0.341]	[0.344]	[0.412]	(0.010)	(0.005)
Hours (W)	1239.696	1208.255	1120.310	1071.721	167.975	48.589
	[605.070]	[634.840]	[583.894]	[697.609]	(16.879)	(8.254)
Hours (P)	1473.383	1438.880	1392.628	1245.385	227.998	147.243
	[658.917]	[695.345]	[663.323]	[793.411]	(19.197)	(9.376)
Income 1000s € (W)	28.049	27.434	24.925	20.903	7.146	4.021
	[19.559]	[20.232]	[15.086]	[17.981]	(0.435)	(0.213)
Income 1000s € (P)	37.173	36.959	35.002	27.544	9.630	7.459
	[26.484]	[29.443]	[23.998]	[28.685]	(0.694)	(0.339)
Bachelor deg. (W)	0.608	0.605	0.591	0.576	0.032	0.015
	[0.488]	[0.489]	[0.414]	[0.494]	(0.012)	(0.006)
Bachelor deg. (P)	0.593	0.598	0.582	0.554	0.040	0.029
	[0.491]	[0.490]	[0.416]	[0.497]	(0.012)	(0.006)
Age (W)	31.643	32.384	33.284	28.384	3.259	4.900
	[4.016]	[4.383]	[3.892]	[4.648]	(0.112)	(0.055)
Age (P)	34.672	35.459	36.327	28.384	6.288	7.943
	[5.527]	[5.993]	[3.924]	[4.655]	(0.113)	(0.055)
Observations	1,716	13,788	5,103	374,812		

Note: Labor market outcomes measured year before first treatment for main sample and year and 9 months before birth of first child for represenstative sample. Representative sample selected to match main sample by year of conception. Relevant sample consists of women in the main sample who remain childless weighted to account for differences in the probability to remain childless. (W) - woman, (P) - partner. Standard deviations in brackets. Standard errors in parentheses.

Back

ToC

Attempts

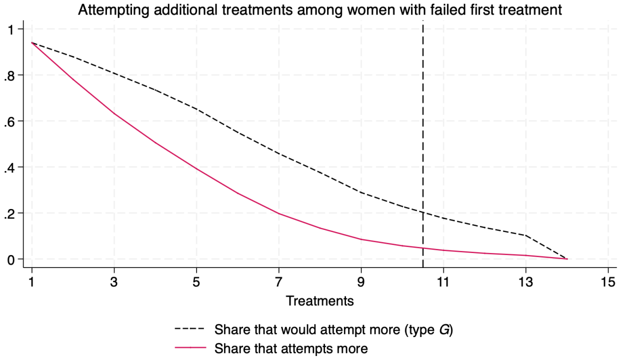


Figure 8: Number of treatments and type

Back

Non-treatment Conception by Type

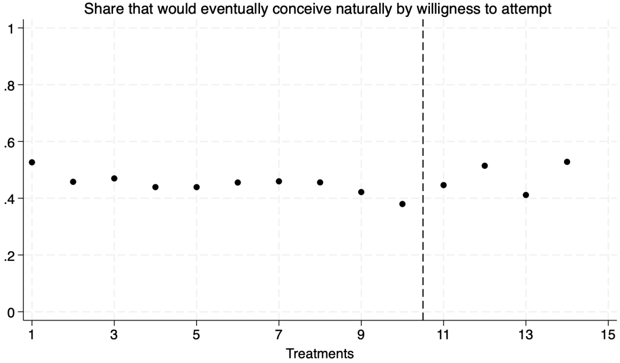


Figure 9: Conceiving naturally and willingness to attempt

Back

ToC

Trimming shares

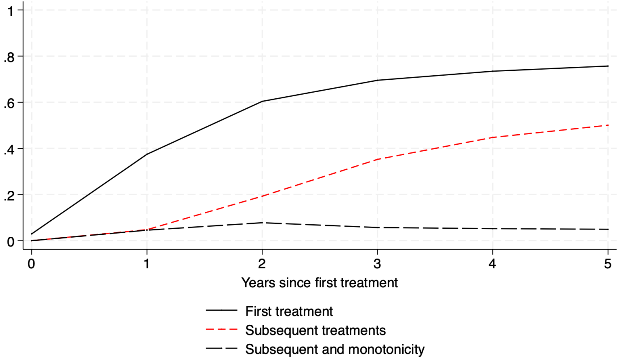


Figure 10: Trimming share under different information

Back

Bounds: Absolute

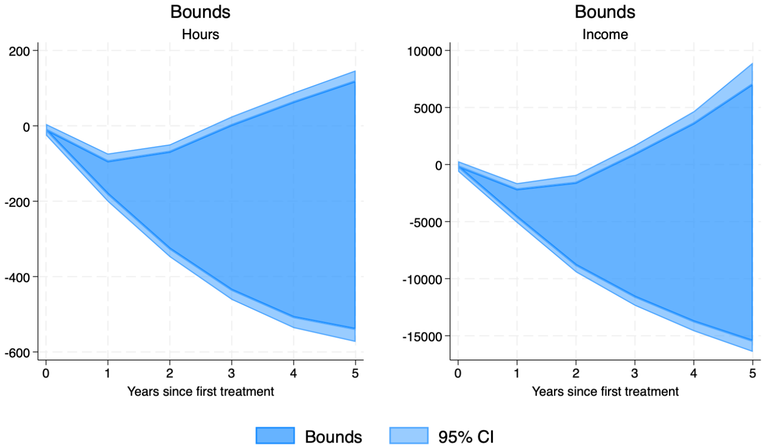


Figure 11: Bounds effects

Bounds: Hours - Comparison to Baseline Lee Bounds

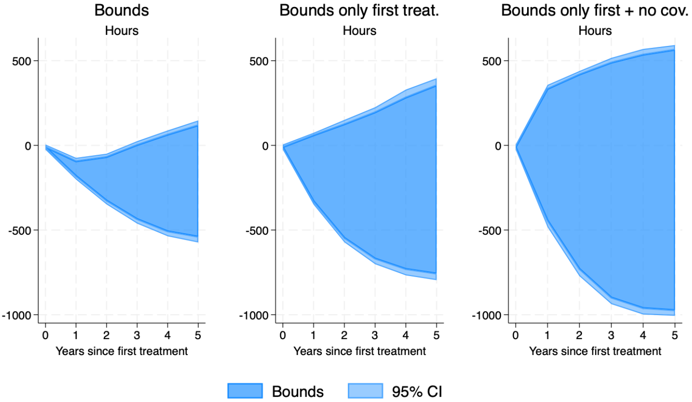


Figure 12: Comparison with baseline Lee: hours

Back

ToC

Bounds: Income - Comparison to Baseline Lee Bounds

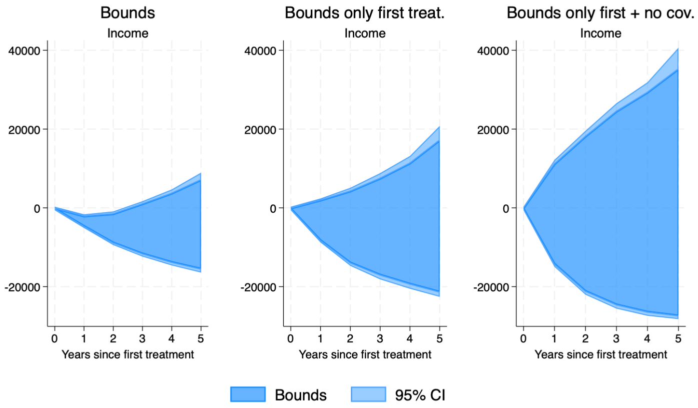


Figure 13: Comparison with baseline Lee: income

Back

Monotonicity (cont.)

Is monotonicity realistic?

- Yes, if families are determined to have at least one child.
  - Decreasing marginal returns to children.
  - Stronger sufficient assumption: success cannot increase total (natural) births.
- No, if first treatment success increases the likelihood of attempting to conceive naturally.
  - Couples may realize they are fertile and try more.
  - First child may "save the relationship" resulting in more attempts to conceive.
  - Couples may dislike having only one child and put more effort into conceiving than if they were to have no children.
- Could assume partial monotonicity.

Back Benefit of monotonicity

ToC

Benefit of Monotonicity

- Women who conceive a second child naturally are not the women who would remain childless.
- This restricts which women among successfully treated could have remained childless had all treatment attempts failed.
- Narrower bounds.

Intuition Back

Monotonicity Intuition (1)

- Suppose there are 100 women whose first treatment attempt succeeded. I know that 20 of them would have eventually conceived naturally.
  - Before: Select 80 women with highest and 80 women with lowest observed outcomes.
    - Trim bottom/top 20%.
  - Now: I additionally observe that 10 of the 100 women do conceive a second child naturally. Select 80 women with highest and 80 women with lowest observed outcomes among the remaining 90.
    - Trim bottom/top 11.1%.

Back

Monotonicity: Intuition (2)

Identify distribution of motherhood outcomes from group whose first treatment succeeds:

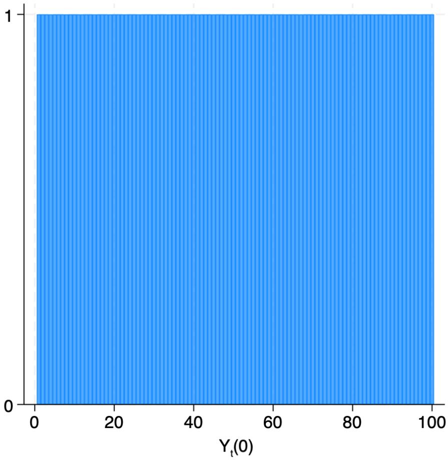


Figure 14: Distribution of potential motherhood outcomes

Monotonicity: Intuition (3)

Without monotonicity: 20 with highest/lowest outcomes:

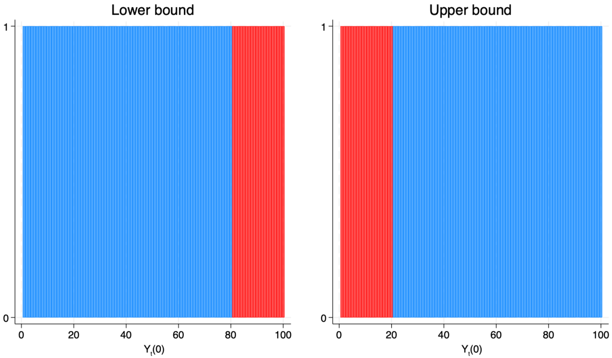


Figure 15: Distribution of potential motherhood outcomes

Back

ToC

Monotonicity: Intuition (4)

With monotonicity: first, drop 10 who conceive again naturally:

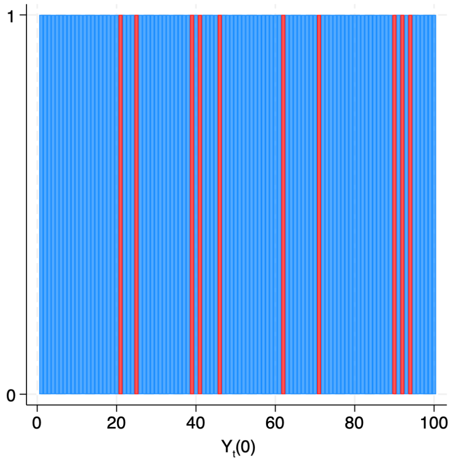


Figure 16: Distribution of potential motherhood outcomes

Monotonicity: Intuition (5)

With monotonicity: second, drop another 10 with highest/lowest outcomes:

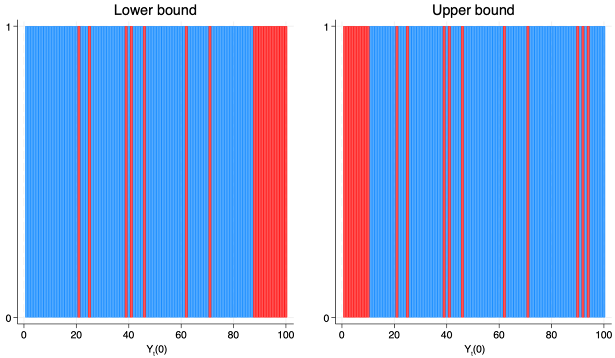


Figure 17: Distribution of potential motherhood outcomes

Back

ToC

Monotonicity: Intuition (6)

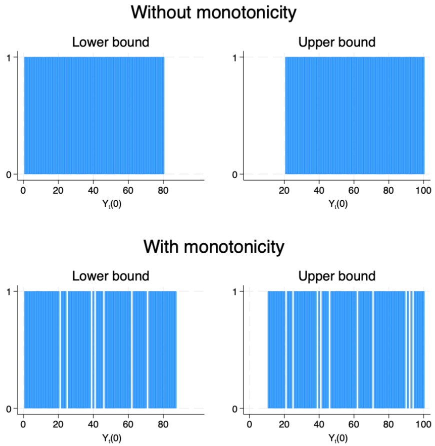


Figure 18: Distribution of potential motherhood outcomes

Monotone Bounds: Absolute

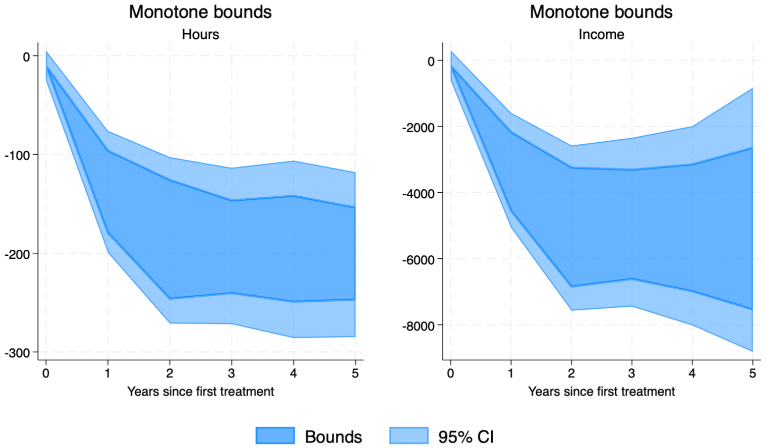


Figure 19: Monotone bounds: absolute terms

Bounds vs Monotone Bounds: Hours

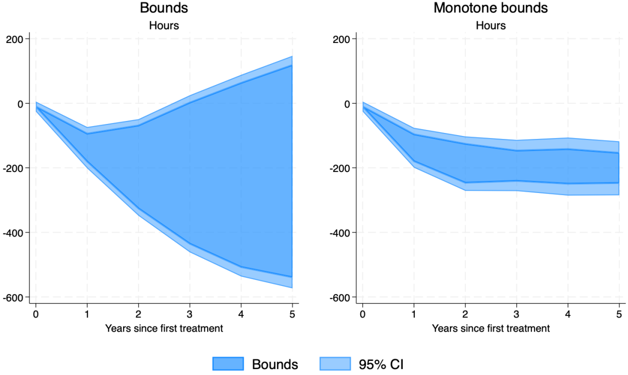


Figure 20: Effect on hours

Back

Bounds vs Monotone Bounds: Income

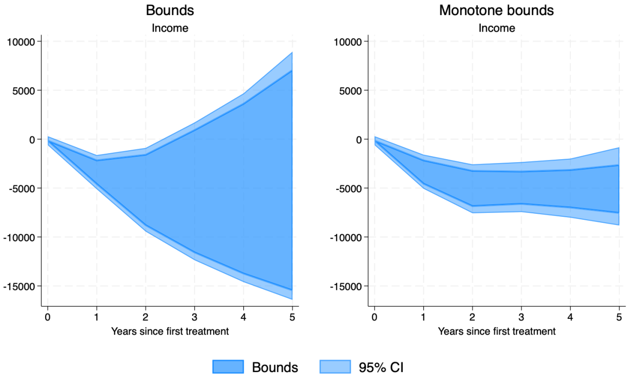


Figure 21: Effect on income

Back

ToC

How Wide are the Bounds?

5 years after first treatment:

- Bounds:
  - 1 SD of pre-treatment hours
  - 1 SD of pre-treatment earnings
- Monotone bounds:
  - 0.15 SD of pre-treatment hours
  - 0.25 SD of pre-treatment earnings

Back

Extensions

- Stable complier group Childless final period
- Bias due to depression Bounds for non-depressed
- Fatherhood penalty Absolute Percentage
- Decomposing gender inequality Share explained by children
- Are estimates less informative that existing? Confidence intervals
- Are existing estimates biased? IV-IVF equivalent Placebo event Placebo %
- Effect of delaying motherhood Bounds and delays

Monotone Bounds: Women who Remain Childless

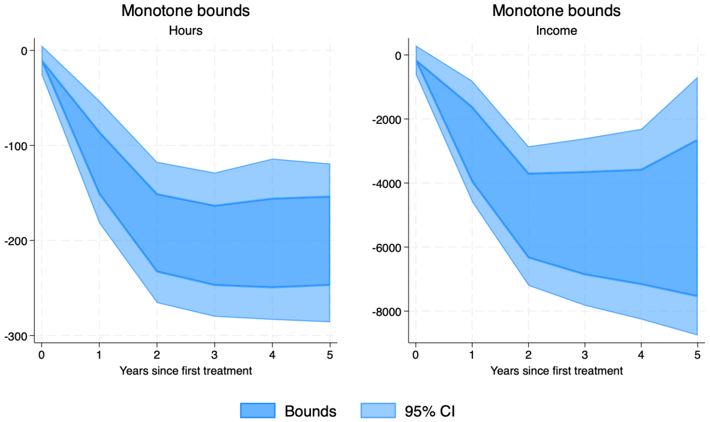


Figure 22: Monotone bounds using final status

Back

ToC

Monotone Bounds: Excluding Depression

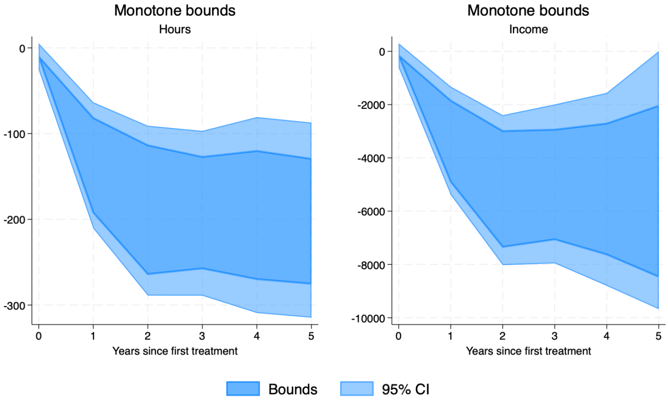


Figure 23: Monotone bounds for women who would not get depressed if they were to remain childless

Back

ToC

Monotone Bounds: Fatherhood Penalty

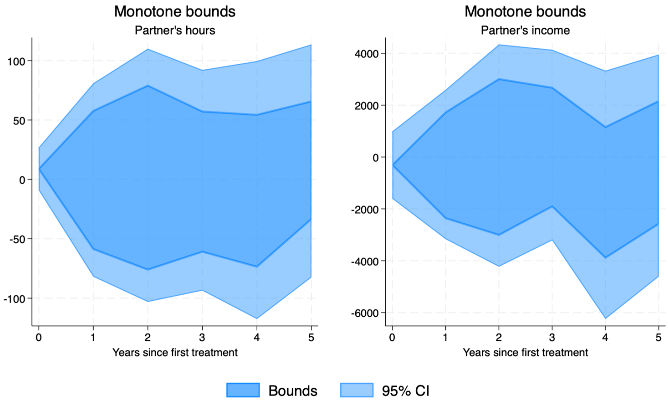


Figure 24: Monotone bounds for partners

Back

ToC

Monotone Bounds: Fatherhood Penalty in Percent

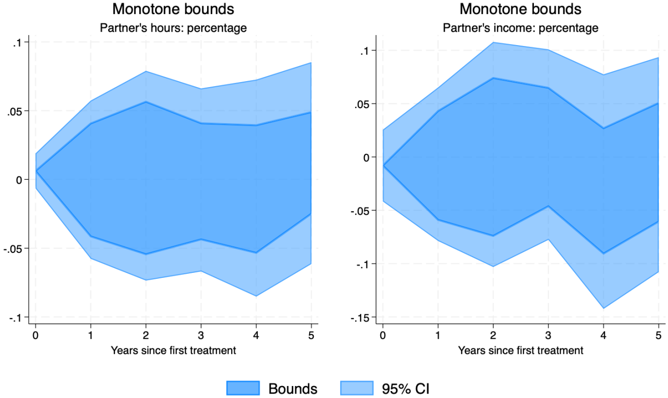


Figure 25: Monotone bounds for partners in percent

Back

ToC

Monotone Bounds: Explaining Gender Inequality

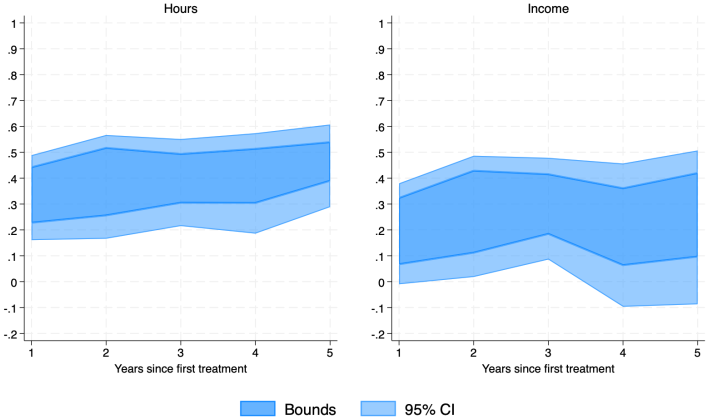


Figure 26: Share of gender inequality explained by parenthood

Back

Are Bounds Less Informative?

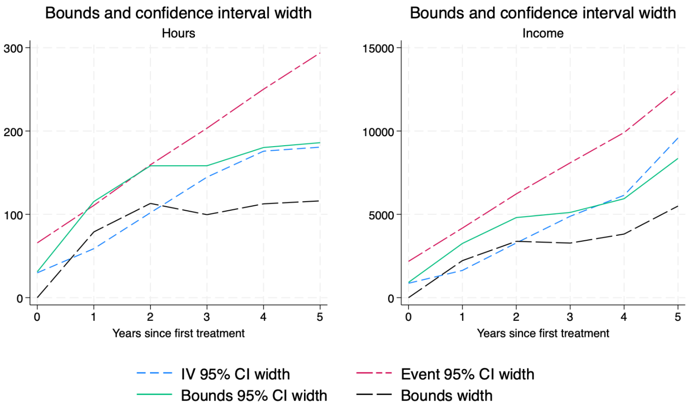


Figure 27: Confidence intervals for different methods

Back

ToC

Monotone Bounds and IV

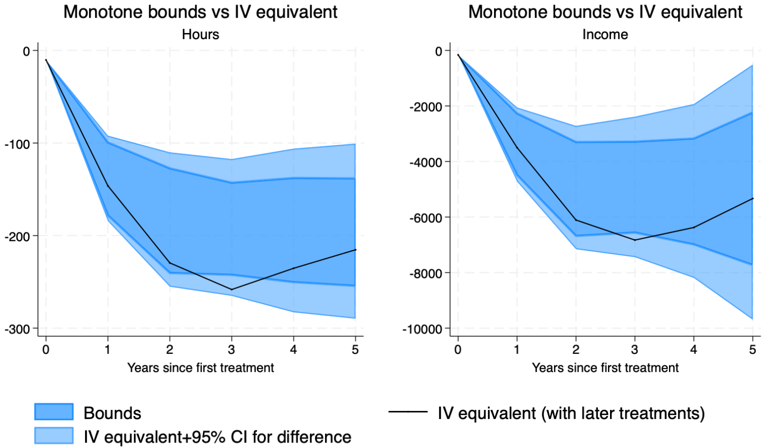


Figure 28: Bounds and IV equivalent for the same population

Placebo Event

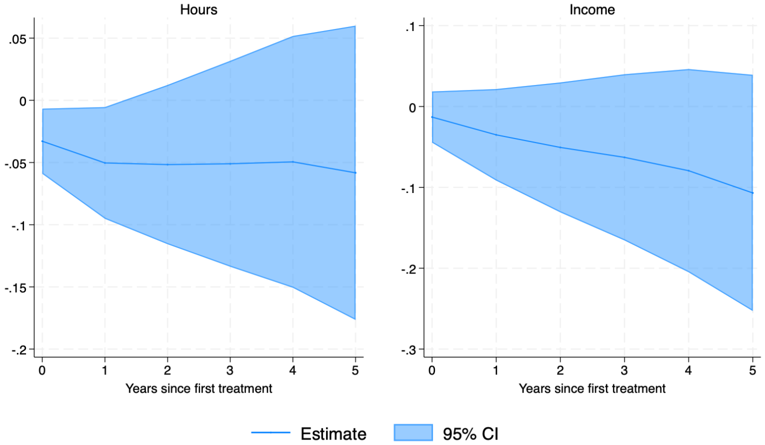


Figure 29: Placebo event study

Back Placebo as share of bounds Placebo as share of regular event

ToC

Placebo Event as Share of Effect

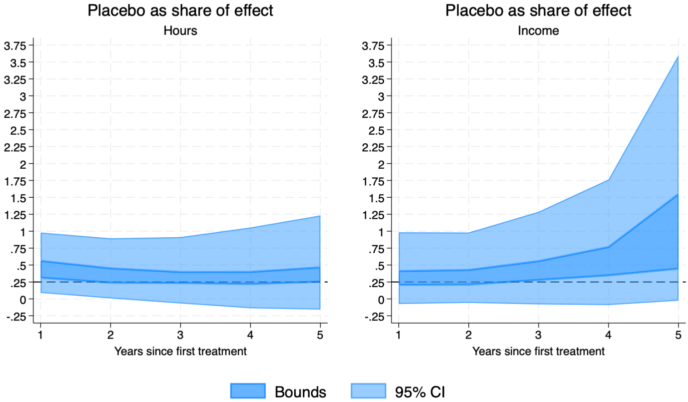


Figure 30: Placebo effects as share of bounds for the same population

Two possible interpretations:

- Event study severely overstates the penalty
- Large share of penalty is due to trying to conceive not parenthood per se

Back Placebo as share of regular event

Placebo Event as Share of Success-at-first Event

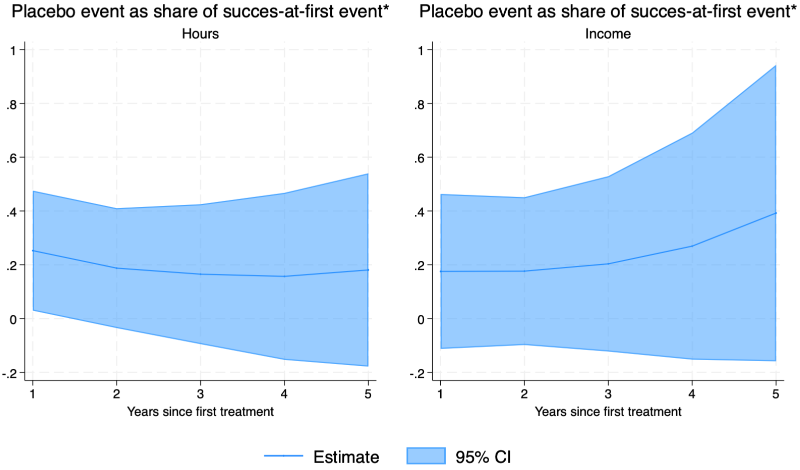


Figure 31: Placebo event estimates as share of success-at-first event estimates

ToC Back

Effect of Delaying Motherhood

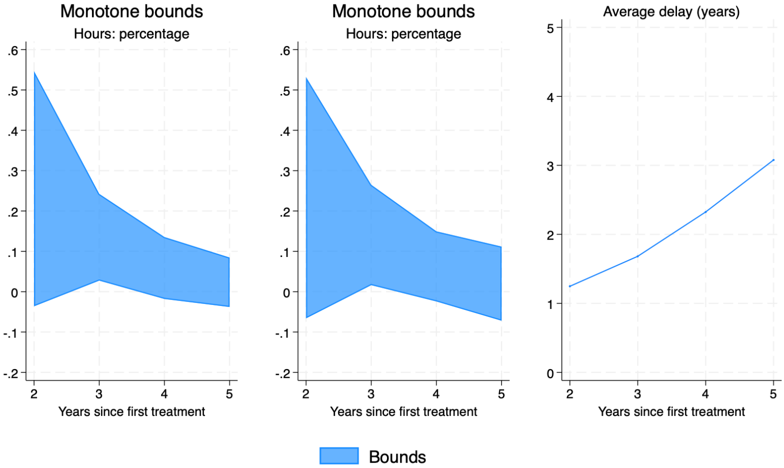


Figure 32: Effect of delaying relative to motherhood at first attempt

Opposite of what is frequently assumed!

Back

ToC

Effect of Motherhood vs Delay

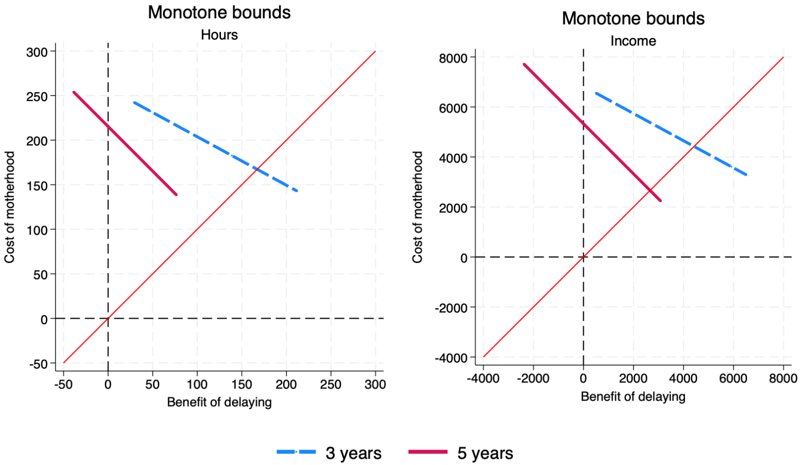


Figure 33: Possible combinations for effects of motherhood and delaying

Back

ToC

Application to Other Settings

Key features:

1. Dynamic treatment effects.
2. Individuals (may choose to) "re-apply" for random treatment assignment.
3. Some may obtain treatment endogenously.

Few examples:

- Education, medical trials, research grants, job training.

Examples

ToC

Application to Other Settings (Examples)

- Education: grade retention, school admission lotteries, special and gifted education programs.
- Medical trials: individuals assigned to control may eventually choose to participate in other medical trials or selectively receive a different treatment.
- Research grants: after unsuccessful application can apply for another or receive funding other ways.
- Job training: those not assigned to training may re-apply, some assignments may be non-random.

Back

Point Identification

Exploiting subsequent treatments enables building a representative sample where none of the women conceive through treatments.

- Remains to address conception by other means.

Two approaches:

1. Restrict selection into natural conception.
2. Restrict the dynamic structure and heterogeneity in the penalty.

- When implementing I also make functional form assumptions.
- Parameter of interest is the ATE.

ToC

First Approach to Point Identification

Restricting selection into natural conceptions.

- *Among women who have no conceived yet, conceiving naturally next year is as-good-as-random after conditioning on age and labor market outcomes in the current year (Z_t).*

$$Y_t(\mathbf{0}), Z_j \perp N_{k+1} | N_k = 0, Z_k, \forall t, j > k.$$

- Does require:
  - Restrictions on selection into natural conception.
- Does not require:
  - Restrictions on heterogeneity and the dynamic structure of the penalty.
- Sequentially estimated probability to conceive naturally in each subsequent period and re-weight the sample using IPW.
- Van den Berg & Vikström (2022) approach on top of the base weights.

Technical details

ToC

Second Approach to Point Identification

Restricting the structure of the penalty for natural conceptions.

- *Penalty t years after conception is similar for women who conceive at first treatment and women who conceive naturally later after conditioning on age and labor market outcomes right before conception (Z).*

$$\mathrm{E}[Y_t(\mathbf{1}) - Y_t(\mathbf{0})|Z] = \mathrm{E}[Y_{t+j}(j) - Y_{t+j}(\mathbf{0})|N_j = 1, N_{j-1} = 0, Z]$$

- Does require:
  - Moment of conception only affects the penalty through time since conception and observable covariates.
  - Penalty is similar for women who conceive at first attempt and those who conceive naturally (conditional on covariates).
- Does not require:
  - Restrictions on selection into natural conception.
- Sequentially estimate the penalty in each period and plug in the estimate for those who conceive naturally.
- Similar to Bensnes et al. (2023); Gallen et al. (2023) approach but on top of the base weights and more flexible.

Technical details

ToC

Point Estimates: Hours

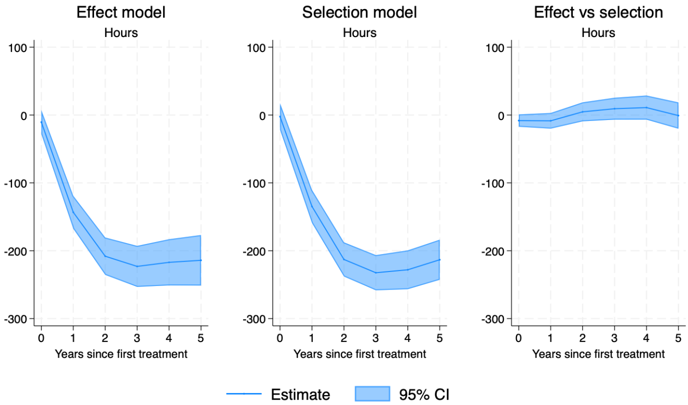


Figure 34: Point estimates for hours

Point Estimates: Income

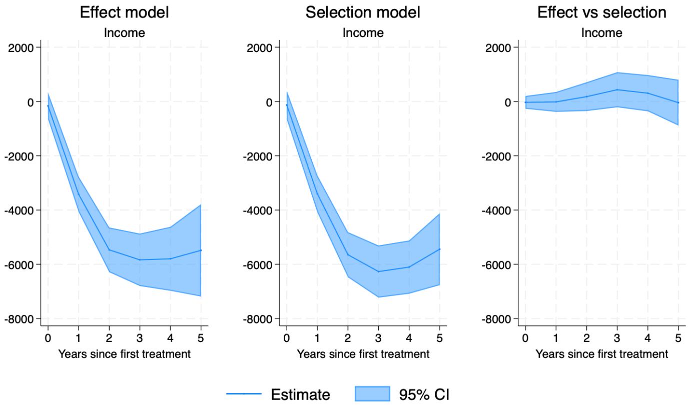


Figure 35: Point estimates for income

ToC

Existing Approaches: Event Study

Event study (Kleven, Landais, & Søgaard, 2019):

$$Y_{ti} = \sum_j \mathbf{1}(age_{ti} = j)\alpha_j + \sum_c \mathbf{1}(year_{ti} = c)\beta_c + \sum_d \mathbf{1}(distanceFB_{ti} = d)\tau_d^{EV} + \varepsilon_{ti},$$

where *distanceFB* is years from the current years to the birth of the first child.

- To make the estimated parameter comparable to mine I implement the event study using women whose first treatment succeeds also applying IPWs to account for differences in the probability of success.

ToC

Existing Approaches: Instrumental Variable

IV (Lundborg et al., 2017):

$$Y_{di} = \hat{C}_{di}\tau_d^{IV} + controls_{di}\beta_d + \nu_{di}$$

$$C_{di} = W_{1i}\mu_d^{IV} + controls_{di}\theta_d + \gamma_{di},$$

where W_1 is the first treatment success indicator.

- I implement the IV estimator in the style of a doubly robust IPW estimator to avoid differences in weights due to covariates.
- Nonetheless, it estimates a difference parameter (LATE).

ToC

Comparison with Event and IV: Hours

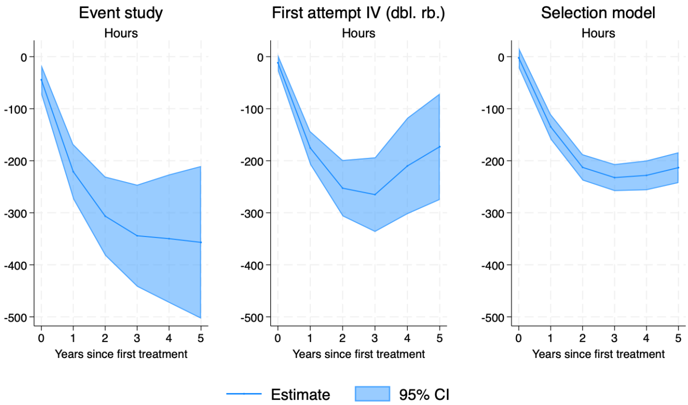


Figure 36: Comparing point estimates for hours

Difference

ToC

Comparison with Event and IV: Income

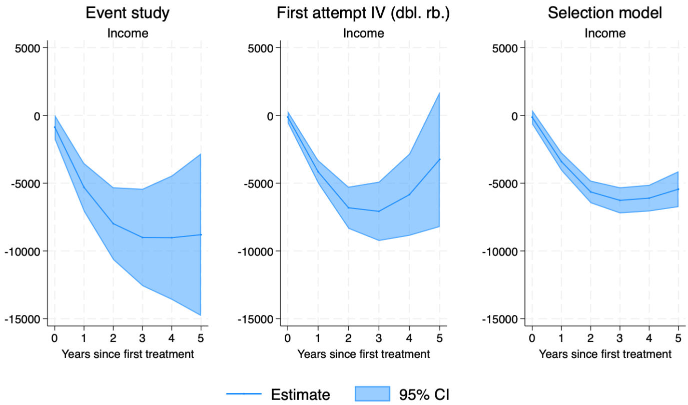


Figure 37: Comparing point estimates for income

Difference

ToC

Reconciling the Differences

- Estimated penalty **in the medium run** is bigger than IV and smaller than event study.
  - IV and event study estimates are must less precise, differences rarely statistically significant.
- IV might understate the penalty:
  - Assumes that conceiving later is same as conceiving earlier.
  - Biased downwards if younger children have a larger impact.
  - This is in contrast to the short run, where IV seems to **overstate**, potentially because in the short run always-takers among successfully treated have more children
- Event study might overstate the penalty:
  - Assumes that women who conceive later have similar career trajectories to those who conceive earlier.
  - Biased upwards if career-oriented women have children later.

Placebo event study

ToC

Sequentially Estimating Selection

1. Construct weights $w(P)$.
2. Using these weights estimate $\Pr(N_1 = 1|Z_1)$ using women for whom all treatment attempts failed.
3. Construct $w_1(P, Z_1) = \frac{w(P)}{\Pr(N_1=0|Z_1)}$
4. Using these weights estimate $\Pr(N_2 = 0|Z_2, N_1 = 0)$ using women for whom all treatment attempts failed and who did not conceive naturally in the first period.
5. Construct $w_2(P, Z_1, Z_2) = \frac{w(P)}{\Pr(N_1=0|Z_1)\Pr(N_2=0|Z_2,N_1=0)}$
6. ... repeat until last period, which gives final weights $w_K(P, X_1, \ldots, X_P, Z_1, \ldots, Z_K)$.
7. Weighted average outcome among childless women together with point identified motherhood outcome gives the ATE

- I take $\Pr(N_k = 0|Z_k, N_1 = 0, \ldots, N_{k-1} = 0)$ as logistic.
- Maybe can be done non-parametrically but more tricky.

Back

Sequentially Estimating the Penalty

1. Estimate $\tau_0(Z) = \mathrm{E}[Y_0(0) - Y_0(\mathbf{0})|Z]$ using Y_0 and W_1 with baseline weights.
2. Construct $\hat{Y}_1 = Y_1(1) - \tau_0(Z)$ if $N_t = 1$ and $\hat{Y}_1 = Y_1$ otherwise.
3. Estimate $\tau_1(Z)$ using $\hat{Y}_1$ and W_1 in the whole sample with baseline weights.
4. Repeat until all $\tau_j(Z)$ are identified and take $\mathrm{E}[\tau_j(Z)]$, which gives the ATE.

- I take $\tau_j(Z)$ as linear.
- Maybe can be done non-parametrically but more tricky.
- Intuition: the penalty 5 years after conception for women who conceive naturally 2 years after the first treatment attempt is similar to the penalty 5 years after conception for women who conceive through their first treatment attempt.

Back

Note on Estimation

- I implement both estimators in the style of a doubly robust IPW estimator to improve efficiency.
- For example, when estimating the treated outcome I add the correction term:

$$\left(1 - \frac{W_1}{\Pr(W_1 = 1|X_1)}\right) Q_t(1, X_1),$$

- where $Q_t(1, X_1)$ is the fitted value from an OLS regression of Y_t on X_1 using women for whom $W_1 = 1$.
- The whole term is zero in expectation but it improves the overall efficiency of the estimators.
- Similar to controlling for covariates in a regression with unconfoundedness.

Back

Difference with Event and IV: Hours

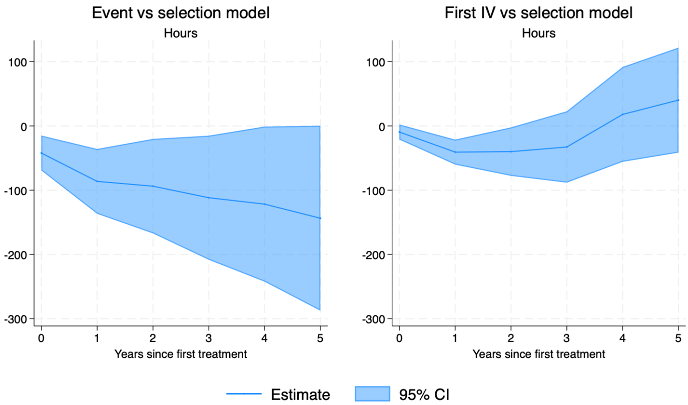


Figure 38: Difference in point estimates for hours

Back

Difference with Event and IV: Income

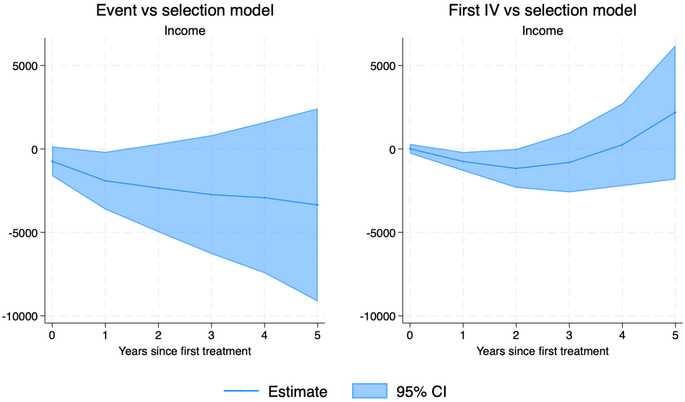


Figure 39: Difference in point estimates for income

Back

Estimated Bias and Placebo Event Study

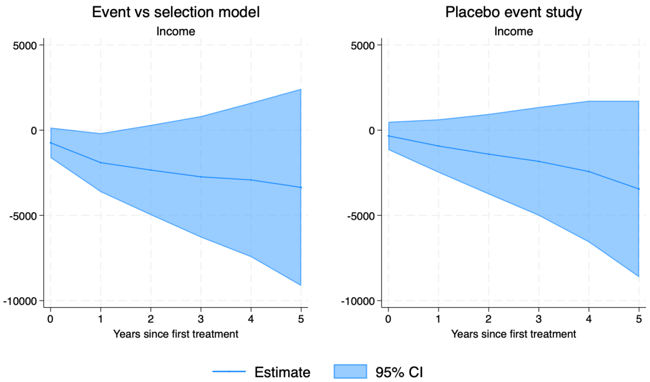


Figure 40: Difference between selection model estimate and event study estimate compared to placebo event study estimate

Back

ToC

References I

Agüero, J. M., & Marks, M. S. (2008). Motherhood and female labor force participation: evidence from infertility shocks. *American Economic Review*, *98*(2), 500–504.

Angelov, N., Johansson, P., & Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of labor economics*, *34*(3), 545–579.

Angrist, J., & Evans, W. N. (1996). *Children and their parents' labor supply: Evidence from exogenous variation in family size.* National bureau of economic research Cambridge, Mass., USA.

Bensnes, S., Huitfeldt, I., & Leuven, E. (2023). *Reconciling estimates of the long-term earnings effect of fertility* (Tech. Rep.). Institute of Labor Economics (IZA).

Bertrand, M. (2011). New perspectives on gender. In *Handbook of labor economics* (Vol. 4, pp. 1543–1590). Elsevier.

Bertrand, M. (2020). Gender in the twenty-first century. In *Aea papers and proceedings* (Vol. 110, pp. 1–24).

Blau, F. D., & Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of economic literature*, *55*(3), 789–865.

Bronars, S. G., & Grogger, J. (1994). The economic consequences of unwed motherhood: Using twin births as a natural experiment. *The American Economic Review*, 1141–1156.

Brooks, N., & Zohar, T. (2021). *Out of labor and into the labor force? the role of abortion access, social stigma, and financial constraints* (Tech. Rep.).

Chung, Y., Downs, B., Sandler, D. H., Sienkiewicz, R., et al. (2017). *The parental gender earnings gap in the united states* (Tech. Rep.).

Cristia, J. P. (2008). The effect of a first child on female labor supply: Evidence from women seeking fertility services. *Journal of Human Resources*, *43*(3), 487–510.

Cruces, G., & Galiani, S. (2007). Fertility and female labor supply in latin america: New causal evidence. *Labour Economics*, *14*(3), 565–573.

Eichmeyer, S., & Kent, C. (2022). *Parenthood in poverty*. Centre for Economic Policy Research.

Gallen, Y., Joensen, J. S., Johansen, E. R., & Veramendi, G. F. (2023). The labor market returns to delaying pregnancy. *Available at SSRN 4554407*.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American economic review*, *104*(4), 1091–1119.

References II

Han, S. (2021). Identification in nonparametric models for dynamic treatment effects. *Journal of Econometrics*, *225*(2), 132–147.

Heckman, J. J., Humphries, J. E., & Veramendi, G. (2016). Dynamic treatment effects. *Journal of econometrics*, *191*(2), 276–292.

Heiler, P. (2022). Estimating heterogeneous bounds for treatment effects under sample selection and non-response. *arXiv preprint arXiv:2209.04329*.

Hirvonen, L. (2009). *The effect of children on earnings using exogenous variation in family size: Swedish evidence.*

Hotz, V. J., McElroy, S. W., & Sanders, S. G. (2005). Teenage childbearing and its life cycle consequences: Exploiting a natural experiment. *Journal of Human Resources*, *40*(3), 683–715.

Iacovou, M. (2001). *Fertility and female labour supply* (Tech. Rep.). ISER Working Paper Series.

Jacobsen, J. P., Pearce III, J. W., & Rosenbloom, J. L. (1999). The effects of childbearing on married women's labor supply and earnings: using twin births as a natural experiment. *Journal of Human Resources*, 449–474.

Kaplan, J. (2023, October 9). A woman's chance for equal pay plummets when she has a kid — and never recovers, says the newest harvard economist to win the nobel prize. *Business Insider*. Retrieved from `https://www.businessinsider.com/women-pay-gap-motherhood-fatherhood-nobel-prize-harvard-economist-2023-10`

Kleven, H., Landais, C., & Leite-Mariante, G. (2023). *The child penalty atlas* (Tech. Rep.). National Bureau of Economic Research.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., & Zweimüller, J. (2019). Child penalties across countries: Evidence and explanations. In *Aea papers and proceedings* (Vol. 109, pp. 122–126).

Kleven, H., Landais, C., & Søgaard, J. E. (2019). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, *11*(4), 181–209.

Lee, D. S. (2005). *Training, wages, and sample selection: Estimating sharp bounds on treatment effects.* National Bureau of Economic Research Cambridge, Mass., USA.

Lundborg, P., Plug, E., & Rasmussen, A. W. (2017). Can women have children and a career? IV evidence from IVF treatments. *American Economic Review*, *107*(6), 1611–37.

References III

Maurin, E., & Moschion, J. (2009). The social multiplier and labor market participation of mothers. *American Economic Journal: Applied Economics*, *1*(1), 251–272.

Miller, A. R. (2011). The effects of motherhood timing on career path. *Journal of population economics*, *24*, 1071–1100.

Miller, C. C. (2017, May 13). The gender pay gap is largely because of motherhood. *The New York Times*. Retrieved from `https://www.nytimes.com/2017/05/13/upshot/the-gender-pay-gap-is-largely-because-of-motherhood.html`

Rosenzweig, M. R., & Wolpin, K. I. (1980). Life-cycle labor supply and fertility: Causal inferences from household models. *Journal of Political economy*, *88*(2), 328–348.

Semenova, V. (2020). Generalized lee bounds. *arXiv preprint arXiv:2008.12720*.

Van den Berg, G. J., & Vikström, J. (2022). Long-run effects of dynamically assigned treatments: A new methodology and an evaluation of training effects on earnings. *Econometrica*, *90*(3), 1337–1354.

Vere, J. P. (2011). Fertility and parents' labour supply: new evidence from us census data: Winner of the oep prize for best paper on women and work. *Oxford Economic Papers*, *63*(2), 211–231.

Zhang, J. L., & Rubin, D. B. (2003). Estimation of causal effects via principal stratification when some outcomes are truncated by "death". *Journal of Educational and Behavioral Statistics*, *28*(4), 353–368.