

# Parenthood Timing and Gender Inequality

Julius Ilciukas

University of Amsterdam

# Motivation

Gender inequality in Western labor markets emerges when individuals become parents (Goldin, 2014; Blau & Kahn, 2017; Bertrand, 2020; Cortés & Pan, 2023)

Quantifying the causal effects of parenthood is central to understanding gender inequity and designing policies to address it

This is challenging for two reasons:

1. Parenthood (timing) may be selective: human capital, wealth, health, career prospects
2. Effects may depend on parenthood timing: age of children, career stage when becoming a parent

Existing methods only address 1. or 2. but yield conflicting results

# This Paper

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

- ▶ How would labor market outcomes of parents change if they did not have children
- 1. New methodological approach robust to selection and dynamic effects
  - ▶ **Application using quasi-experimental variation in the success of assisted conception procedures**
- 2. Empirical evidence using administrative Dutch data
  - ▶ Focus on couples undergoing **artificial insemination**
- 3. Unified framework to disentangle and quantify the extent of selective timing and timing-dependent effects
  - ▶ Quantify bias in leading methods that disregard either factor

## Preview of Main Results

- ▶ Parenthood persistently reduces women's work hours and income
  - ▶ Yearly reductions between 9 and 24 percent
- ▶ Parenthood causes a large share, but far from all, of post-child gender inequality
  - ▶ Between 36 and 54 percent in work hours and up to 46 percent in income
- ▶ Both selection and dynamic effects are substantial
  - ▶ Bias as extreme as attributing all or none of gender inequality to parenthood

# Literature and Contribution

1. Existing methods for quantifying parenthood's impacts rely on restrictive assumptions about dynamic effects or selection

- ▶ Address selection: Hotz et al. (2005); Agüero & Marks (2008); Cristia (2008); Miller (2011); [Lundborg et al. \(2017\)](#); Bensnes et al. (2023); Gallen et al. (2023); Lundborg et al. (2024)
- ▶ Address dynamic effects: Fitzenberger et al. (2013); Angelov et al. (2016); Adda et al. (2017); Bütikofer et al. (2018); [Kleven et al. \(2019\)](#); Melentyeva & Riedel (2023); Kleven et al. (2024)

**I provide first estimates robust to both selective fertility and dynamic effects**

# Literature and Contribution

1. Existing methods for quantifying parenthood's impacts rely on restrictive assumptions about dynamic effects or selection

- ▶ Address selection: Hotz et al. (2005); Agüero & Marks (2008); Cristia (2008); Miller (2011); [Lundborg et al. \(2017\)](#); Bensnes et al. (2023); Gallen et al. (2023); Lundborg et al. (2024)
- ▶ Address dynamic effects: Fitzenberger et al. (2013); Angelov et al. (2016); Adda et al. (2017); Bütikofer et al. (2018); [Kleven et al. \(2019\)](#); Melentyeva & Riedel (2023); Kleven et al. (2024)

**I provide first estimates robust to both selective fertility and dynamic effects**

2. Leading methods offer conflicting evidence yet are hard to compare Comparison

- ▶ Lundborg et al. (2017): parenthood has little impact on gender inequality
- ▶ Kleven et al. (2019): parenthood causes nearly all gender inequality

**I provide a unified comparison framework and demonstrate that accounting for selection and dynamic effects is sufficient to reconcile the conflicting results**

# Literature and Contribution

1. Existing methods for quantifying parenthood's impacts rely on restrictive assumptions about dynamic effects or selection

- ▶ Address selection: Hotz et al. (2005); Agüero & Marks (2008); Cristia (2008); Miller (2011); [Lundborg et al. \(2017\)](#); Bensnes et al. (2023); Gallen et al. (2023); Lundborg et al. (2024)
- ▶ Address dynamic effects: Fitzenberger et al. (2013); Angelov et al. (2016); Adda et al. (2017); Bütikofer et al. (2018); [Kleven et al. \(2019\)](#); Melentyeva & Riedel (2023); Kleven et al. (2024)

**I provide first estimates robust to both selective fertility and dynamic effects**

2. Leading methods offer conflicting evidence yet are hard to compare Comparison

- ▶ Lundborg et al. (2017): parenthood has little impact on gender inequality
- ▶ Kleven et al. (2019): parenthood causes nearly all gender inequality

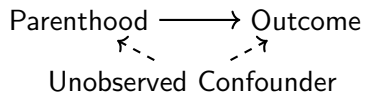
**I provide a unified comparison framework and demonstrate that accounting for selection and dynamic effects is sufficient to reconcile the conflicting results**

3. Addressing selection and dynamic effects is a common challenge

- ▶ Education programs with multiple admission cycles, assignment to judges, promotion tournaments

**Method applicable to many other settings with sequential quasi-experiments**

## Identification Challenge



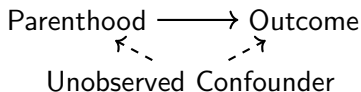


## Identification Challenge



- ▶ Addressing selection requires a natural experiment
- ▶ Consider couples undergoing assisted conception procedures (ACP) for first child
- ▶ Sperm is inserted into uterus—assume conception is as-good-as-random
  - ▶ Uncorrelated with past labor market outcomes conditional on age
- ▶ 75% become mothers after their first ACP fails

# Identification Challenge



- ▶ Addressing selection requires a natural experiment
- ▶ Consider couples undergoing assisted conception procedures (ACP) for first child
- ▶ Sperm is inserted into uterus—assume conception is as-good-as-random
  - ▶ Uncorrelated with past labor market outcomes conditional on age
- ▶ 75% become mothers after their first ACP fails



## Outline of the Approach

Objective is a clean comparison uncontaminated by timing differences:

- ▶ Women conceiving at first ACP vs **similar** childless women with failed first ACP

Challenge: women remaining childless after first ACP fails may be a selective group:

1. Decision to undergo more ACPs may be selective—complicates using later ACPs
2. Conception via non-ACP means might be selective

Step 1: selection via subsequent ACPs can be addressed using the realized ACP count

- ▶ Women who tried more were less likely to never succeed ex-ante: weight them more

Step 2: non-ACP conceptions can be addressed using a worst-case bounding approach

- ▶ Assume most extreme selection into non-ACP births consistent with the data

Crucial assumption: ACP outcomes are as good as random (cond. on age)

## Model: Outcomes

Moment  $t \in \{1, \dots, T\}$  since woman's first ACP

- ▶ Outcome when motherhood begins at first ACP:

$$Y_t(1)$$

- ▶ Childless outcome:

$$Y_t(0)$$

- ▶ Outcome when motherhood begins in period  $k$ :

$$Y_t(k)$$

These scenarios involve women trying to conceive through ACPs

- ▶ I will first focus on quantifying impacts in these scenarios

To simplify exposition:

- ▶ We are at  $t = T$
- ▶  $Y_T(k) = Y_T(\text{later})$  for all  $k > 1$

## Model: Latent Variables and Treatment Effect

Women differ in two unobserved characteristics:

- ▶ “Willingness” to undergo ACPs,  $W \in \{1, \dots, \bar{w}\}$ 
  - ▶ Would undergo  $W$  ACPs for the first child if all ACPs failed
- ▶ “Reliance” on ACPs,  $R \in \{0, 1\}$ 
  - ▶ No child if all ACPs fail,  $R = 1$
  - ▶ “Reliers”  $\supseteq$  “compliers” (no child if first ACP fails)

Average treatment effect for reliers:

$$\tau_{ATR} = \mathbb{E}[Y(1) - Y(0) | R = 1]$$

- ▶ Focusing on reliers rather than compliers is not only appealing because they are a more general group but it is also essential for informative results
- ▶ As compliers, reliers may change over time, which I will address in extensions

## Model: Observables

- ▶ ACP  $j$  success indicator,  $Z_j$ 
  - ▶  $Z_j = 0$  if failed or did not happen
  - ▶ Only ACPs before the first child
- ▶ Number of realized ACPs:

$$A = \min(\{j : Z_j = 1\} \cup \{W\})$$

- ▶ Parenthood indicator:

$$D = Z_A + (1 - Z_A)(1 - R)$$

- ▶ One-sided non-compliance
- ▶ Realized outcome:

$$Y = Y(0)(1 - D) + Y(1)DZ_1 + Y(\text{later})D(1 - Z_1)$$

# Sequential Unconfoundedness

## Assumption (Sequential Unconfoundedness)

$$(Y(1), Y(0), R, W) \perp\!\!\!\perp Z_j | A \geq j.$$

In words: once sperm/embryo at ACP  $j$  are implanted, whether this results in a conception is as-good-as-random

- ▶ **The decision to undergo the procedure can be endogenous**
- ▶  $Y(1), Y(0), R$  and  $W$  can be related
- ▶ Main method relaxes to covariate-conditional version
  - ▶ E.g. age at procedure, technology

## Simple World: Max 1 ACP, All Reliers

$W = 1$   
(willing to try once)



## Simple World: Max 1 ACP, All Reliers

$W = 1$   
(willing to try once)

$$Z_1 = 1$$

$$Z_1 = 0$$

## Simple World: Max 1 ACP, All Reliers

$W = 1$   
(willing to try once)

$$Z_1 = 1$$

$$\mathbb{E}[Y(1)]$$

$$Z_1 = 0$$

$$\mathbb{E}[Y(0)]$$

## Simple World: Max 2 ACPs, All Reliers

$$W = 1$$

(willing to try once)

$$Z_1 = 1$$

$$Z_1 = 0$$

## Simple World: Max 2 ACPs, All Reliers

$$W = 1$$

(willing to try once)

$$Z_1 = 1$$

$$Z_1 = 0$$

$$W = 2$$

(willing to try twice)

$$Z_1 = 1$$

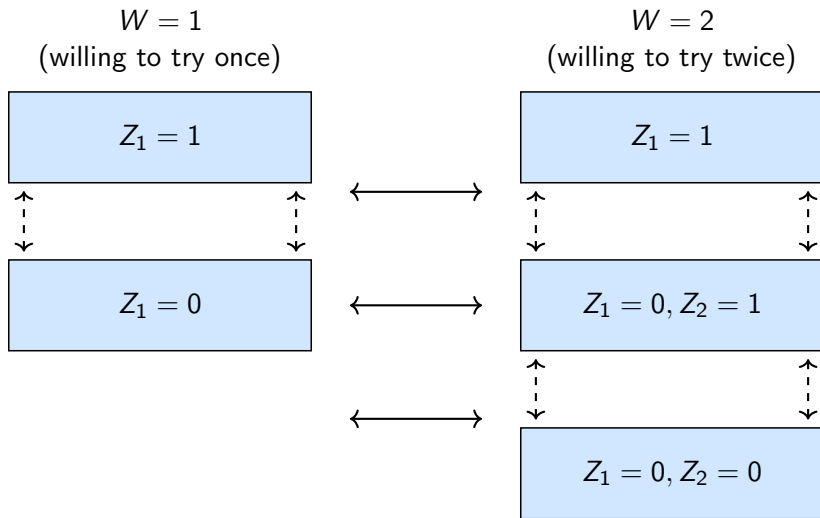
$$Z_1 = 0, Z_2 = 1$$

$$Z_1 = 0, Z_2 = 0$$

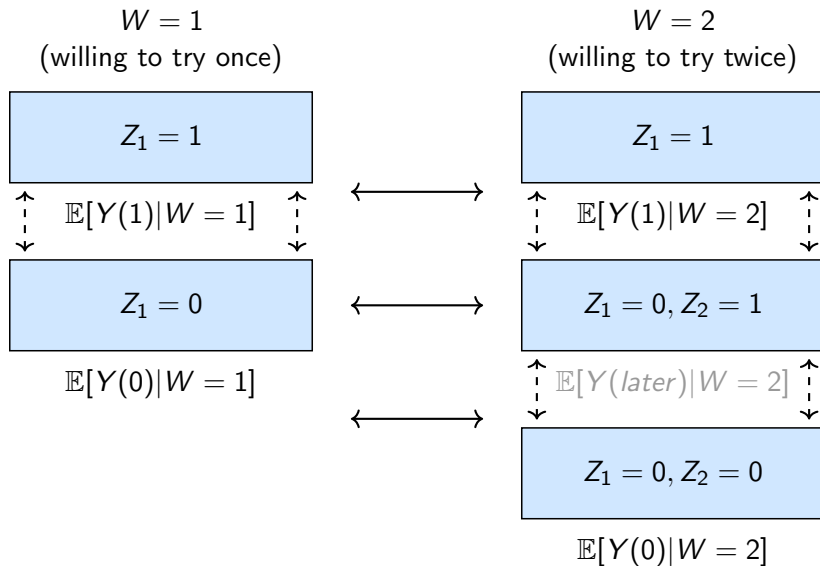
## Simple World: Max 2 ACPs, All Reliers



## Simple World: Max 2 ACPs, All Reliers



## Simple World: Max 2 ACPs, All Reliers



## Simple World (Observed): Max 2 ACPs, All Reliers

$W = 1$   
(willing to try once)

$W = 2$   
(willing to try twice)

$$Z_1 = 1$$

$$Z_1 = 0$$

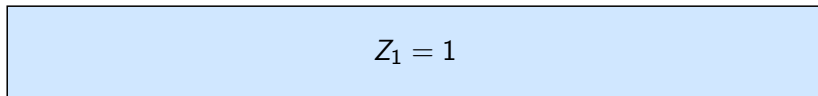
$$Z_1 = 0, Z_2 = 1$$

$$Z_1 = 0, Z_2 = 0$$

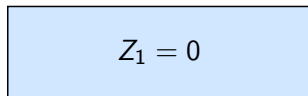


## Simple World (Observed): Max 2 ACPs, All Reliers

$W = 1$   
(willing to try once)

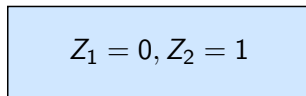


$\mathbb{E}[Y(1)]$

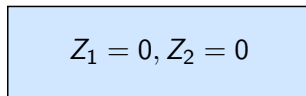


$\mathbb{E}[Y(0)|W = 1]$

$W = 2$   
(willing to try twice)

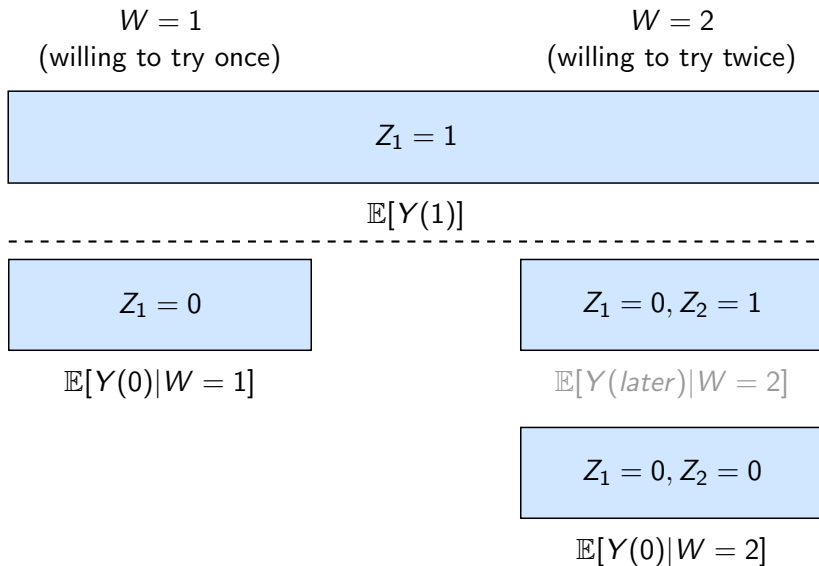


$\mathbb{E}[Y(\textit{later})|W = 2]$



$\mathbb{E}[Y(0)|W = 2]$

## Simple World (Observed): Max 2 ACPs, All Reliers



## Simple World (Observed): Max 2 ACPs, All Reliers

$W = 1$   
(willing to try once)

$W = 2$   
(willing to try twice)



$\mathbb{E}[Y(1)]$



$\mathbb{E}[Y(0)|W = 1]$



$\mathbb{E}[Y(\textit{later})|W = 2]$

$Pr(W = 1) =$



$\mathbb{E}[Y(0)|W = 2]$

## Simple World: Max 1 ACP with Non-reliers

$R = 1$   
(no child if fail)

$R = 0$   
(child if fail)

## Simple World: Max 1 ACP with Non-reliers

$R = 1$   
(no child if fail)

$R = 0$   
(child if fail)


$$Z_1 = 1$$

## Simple World: Max 1 ACP with Non-reliers

$R = 1$   
(no child if fail)

$R = 0$   
(child if fail)

$$Z_1 = 1$$

$$Z_1 = 0, D = 0$$

$$Z_1 = 0, D = 1$$

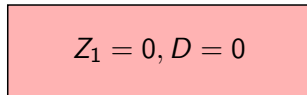
## Simple World: Max 1 ACP with Non-reliers

$R = 1$   
(no child if fail)

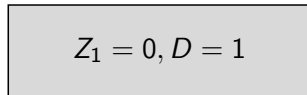
$R = 0$   
(child if fail)



Distribution of  $Y(1)$



$\mathbb{E}[Y(0)|R = 1]$



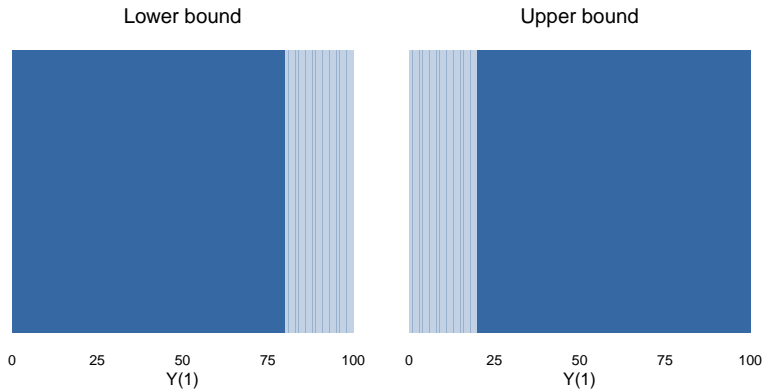
$\mathbb{E}[Y(\text{later})|R = 0]$

$Pr(R = 1) =$



## Intuition: Motherhood Outcome $Y(1)$

1. Treated group is a representative sample but their types are unobserved
2. Identify  $\Pr(R = 1) = 0.8$  on control group
3. Assume most extreme distributions of types in treated group
4. Bound  $\mathbb{E}[Y(1)|R = 1]$





# Technical Details

Formal identification →

- ▶ Covariate-conditional sequential unconfoundedness
- ▶ Combine the two steps in a semi-parametric moment equation

Using covariates to narrow the bounds →

- ▶ The bounds are sharp

Inference complicated by trimming of the outcome distribution →

- ▶ Build on a double/debiased machine learning approach by Semenova (2023)

# Background and Data

## Assisted conception procedures →

- ▶ **Intrauterine insemination:** direct sperm injection, minimally invasive, free
- ▶ In-vitro fertilization: invasive medical procedure, first 3 free

## Dutch family policies similar to OECD average →

- ▶ 16 weeks maternity + pregnancy leave, 1 week paternity leave
- ▶ Average net childcare cost 10% median household income
- ▶ Part-time work more common, average work hours similar to OECD average

## Data →

- ▶ Hospital records on treatment dates and types: success imputed as birth within 10 months without additional ACPs
- ▶ Work hours and income include leave; results for hours corrected for uncertainty
- ▶ 15,523 cohabiting opposite-sex couples undergoing intrauterine insemination for their first child between 2013 and 2016
- ▶ Balance: ACP success at each attempt uncorr. with past outcomes cond. on age

Details

Balance in 1st ACP first

Balance in later ACPs

Success and willingness

Rep. samp.

## Results: Bounds

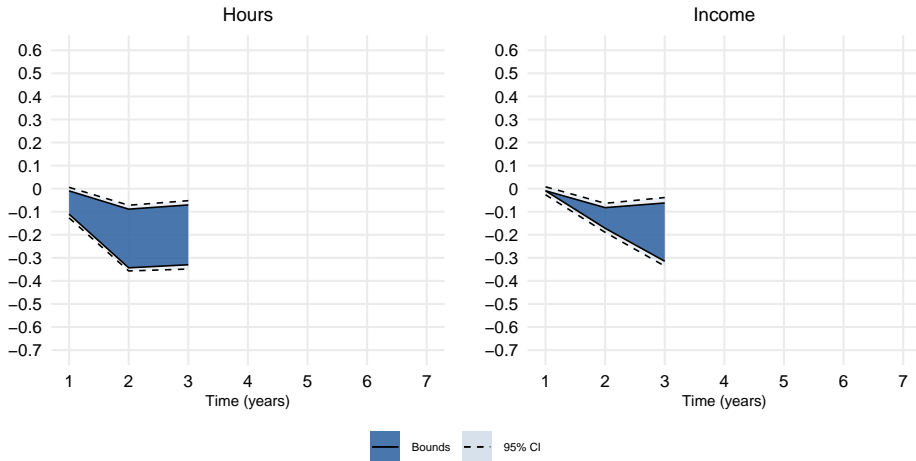


Figure 1: Bounds for Women (% Change Relative to Childless)

## Results: Bounds

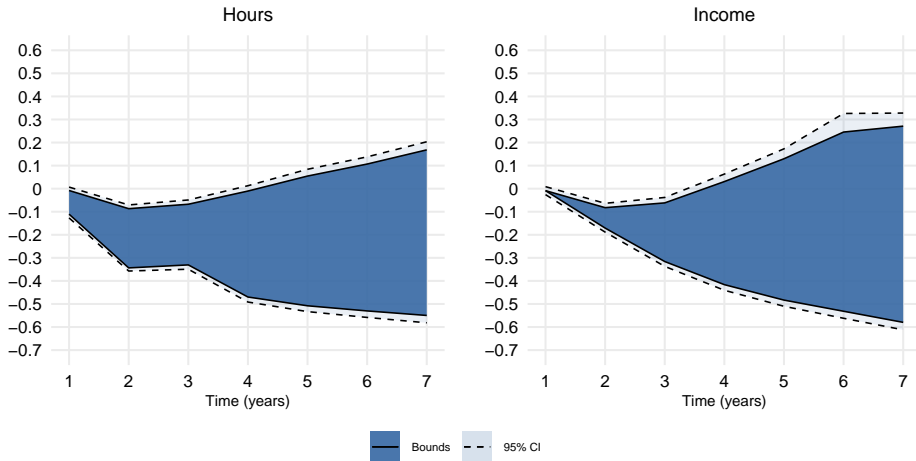


Figure 1: Bounds for Women (% Change Relative to Childless)

## Narrowing the Bounds Further

Use additional information which on mother are reliers:

- ▶ Some women have non-ACP children after ACP succeeds
- ▶ May be reasonable to assume they are not reliant on ACPs
- ▶ Consistent with being determined to have at least one child
- ▶ Reduces uncertainty around which women are reliers



# Bounds with Monotonicity



Figure 2: Bounds for Women Under Monotonicity

# Bounds for Men



Figure 3: Bounds fo Men

# Gender Inequality

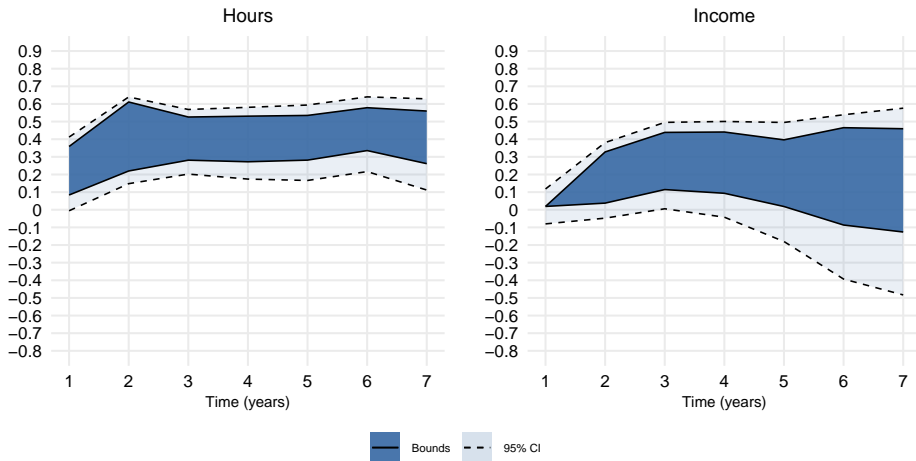


Figure 4: Share of Gender Inequality Caused by Parenthood



# Extensions

- ▶ **Bias in leading methods** Naive comp. Less naive comp. Formal procedure Bias in IV Bias in ES
- ▶ **Mental health side effects** Discussion Bounds for non-depresses
- ▶ Relation to methodological literature Theoretical comparison Results
- ▶ Confidence interval comparison Confidence intervals
- ▶ Inequality correcting for age De-aging partners
- ▶ Stable complier group Childless final period
- ▶ Estimator without DML Estimates
- ▶ Monotonicity Discussion Direction Partnered only Partnership and depression Test
- ▶ Testing Bensnes et al. (2023); Gallen et al. (2023) Estimates
- ▶ Heterogeneity Willingness to try
- ▶ Population imputation\* ES pop. Mother. imp. Childless imp. Effect imp.

# Conclusion

Method for evaluating the career cost of parenthood:

- ▶ Robust to selective fertility and dynamic effects
- ▶ Applicable to various settings with sequential treatment assignment and selection:
  - ▶ Assignment to job training and educational programs, legal settings with assignment to varying leniency “judges”, promotion tournaments, clinical trials in extension phase

Application to Dutch data:

- ▶ Motherhood reduces work hours and income by 9% to 24%
- ▶ Parenthood causes up to 50% of post-child gender inequality

Reconciling conflicting findings in literature:

- ▶ Naive: IV might understate the role of parenthood in gender ineq., ES overstates it
- ▶ This also holds after accounting for differences in population and effect definition

Policy:

- ▶ Large share of gender inequality may not be due to parenthood per se

# Instrumental Variable vs Event Study: Percent Reduction in Earnings



Source: (Lundborg et al., 2024)

► “Naive” comparison with differing sub-populations and treatment definitions

[Back \(literature\)](#)

[Back \(extensions\)](#)

# Theoretical Bias



Figure 5: Motherhood among unsuccessfully treated

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

$$\tau_{IV} = \tau_{Parenthood} - 3\tau_{Delay}$$

## Bounding $\tau_{ATR}$

Construct the moment:

$$m^L(G, \eta^0) = Y 1_{\{Y < q(r(X_1), X_1)\}} \frac{Z_1}{e_1(X_1)} - Y(1 - D) \prod_{j=1}^A \frac{(1 - Z_j)}{(1 - e_j(X_j))}$$

- ▶  $G$  is the observed data vector
- ▶  $\eta^0$  contains the following:
  - ▶  $e_j(X_j) = \Pr(Z_j = 1 | X_j)$
  - ▶  $q(r(X_1), X_1)$  is the  $r(X_1)$ -th quantile of  $Y$  given  $X_1$  and  $Z_1 = 1$
  - ▶  $r(X_1)$  identifies the covariate-conditional relier share

### Assumption (Conditional Sequential Unconfoundedness)

$(Y(k), R, W) \perp\!\!\!\perp Z_j \mid X_j$  for all  $j, k$ , and  $X_j, A \geq j$ .

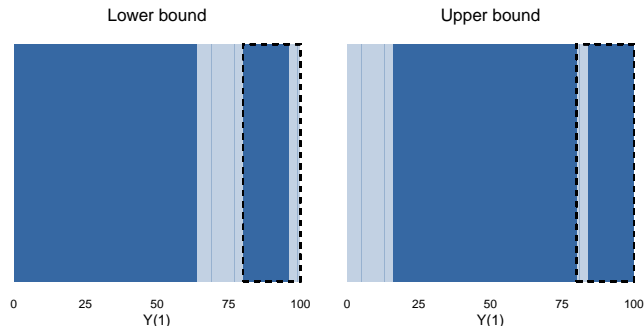
### Theorem (Lower Bound)

*Under conditional sequential unconfoundedness and regularity, the sharp lower bound on  $\tau_{ATR}$  is  $\mathbb{E}[m^L(G, \eta^0)] / \mathbb{E}[r(X_1)]$ .*

## Intuition: Motherhood Outcome $Y(1)$ —Covariates

Pre-ACP covariates can help narrow the bounds:

- ▶ Can identify relier share at each covariate value
- ▶ Baseline bounds assume extreme scenarios where reliers have highest or lowest treated outcomes
- ▶ These distributions of treated outcomes might be inconsistent with conditional relier shares



# Estimating the Bounds

Distribution of  $m^L(G, \eta^0)$  is complicated by  $q(r(X_1), X_1)$

- ▶ Semenova (2023) addresses a closely related inference challenge
- ▶ Double/debiased machine learning approach
  1. Adjust  $m^L(G, \eta^0)$  to make it insensitive to small error in  $q(r(X_1), X_1)$
  2. Sample splitting
- ▶ Asymptotic inference as if  $q(r(X_1), X_1)$  was known

New moment:

$$\psi^L(G, \xi^0) = m^L(G, \eta^0) + \text{corr}(G, \xi^0)$$

Identifies same parameter:

$$\mathbb{E}[\psi^L(G, \xi^0)] = \mathbb{E}[m^L(G, \eta^0)]$$

Insensitive to estimation error in  $q(r(X_1), X_1)$ :

$$\partial_{q(\cdot)} \mathbb{E}[\psi^{L+}(G, \xi_r) | X_1] |_{\xi_r = \xi_r^0} = 0 \text{ a.s.}$$

# Assisted Conception Procedures

- ▶ IUI (main procedure): sperm injected into uterus
  - ▶ Minimally invasive, primary ACP in most countries
  - ▶ “Free” in NL
- ▶ IVF (secondary procedure): embryo inserted into uterus
  - ▶ Invasive treatment, performed under sedation/anesthesia
  - ▶ Eggs retrieved through the vaginal wall using a specialized needle
  - ▶ In NL, first 3 free; each subsequent costs between 1000 and 4000 EUR

[Back](#)



# Institutions

- ▶ Dutch family friendly policies similar to OECD average
  - ▶ 16 weeks of fully paid pregnancy+maternity leave
  - ▶ 1 week of paternity leave
  - ▶ Average time in child care similar to OECD average
  - ▶ Net child care cost 10% median household income
- ▶ Dutch employment intensity similar to OECD average
  - ▶ Employment among parents and non-parents relatively high
  - ▶ Part time work much more common
  - ▶ Approximately 15% two-parent families have both partners working part-time

[Back](#)

# Data

## Administrative data from Statistics Netherlands

- ▶ Comprehensive hospital records cover fertility treatments from 2012 to 2017: procedure date and type
  - ▶ Success imputed as having child born within 10 months
- ▶ Tax records cover work hours and income from 2011 to 2023
  - ▶ Include maternity leave and pay
  - ▶ Main bounds account for uncertainty around actual work hours
- ▶ Birth dates, legal family connections, cohabitation
- ▶ Dispensed medication registry

Main sample: cohabiting opposite-sex couples undergoing IUI for their first child between 2013 and 2016: 15,523

[Back](#)

# Overview of Descriptives

- ▶ First and subsequent ACP success uncorrelated with past labor market outcomes conditional on age [Table first](#) [Table later](#)
  - ▶ Support for independence of  $Z_j$  and  $(Y(1), Y(0))$
- ▶ Success probability stable across ACPs conditional on age [Figure](#)
  - ▶ Support for independence of  $Z_j$  and  $W$
- ▶ Representative sample worked less and had lower income before parenthood, but differences relatively small [Table](#)
  - ▶ ACP sample older before parenthood

[Back \(summary\)](#)

# Balance in 1<sup>st</sup> ACP

Table 1: First ACP Outcomes and Descriptives

	Success (1)	Fail (2)	Difference (1)-(2)	Dif. cond. age & educ. (1)-(2) cond.
Work (W)	0.882 [0.323]	0.863 [0.344]	0.019 (0.009)	0.008 (0.009)
Work (P)	0.884 [0.320]	0.865 [0.342]	0.019 (0.009)	0.013 (0.009)
Hours (W)	1240.315 [604.666]	1207.860 [635.194]	32.455 (16.183)	18.702 (16.560)
Hours (P)	1474.530 [658.231]	1438.590 [695.692]	35.940 (17.713)	18.579 (17.870)
Income 1000s € (W)	28.065 [19.559]	27.418 [20.219]	0.647 (0.516)	0.745 (0.546)
Income 1000s € (P)	37.205 [26.482]	36.952 [29.452]	0.252 (0.746)	0.364 (0.730)
Bachelor deg. (W)	0.480 [0.500]	0.451 [0.498]	0.029 (0.013)	
Bachelor deg. (P)	0.394 [0.489]	0.381 [0.486]	0.013 (0.012)	
Age (W)	31.638 [4.015]	32.388 [4.383]	-0.750 (0.111)	
Age (P)	34.675 [5.513]	35.461 [5.996]	-0.786 (0.152)	
Observations	1,714	13,809		
Joint $p$ -val.			0.000	0.928

Note: Labor market outcomes measured year before first ACP. (W) - woman, (P) - partner. Last column uses inverse probability weights for the first ACP that follow the main specification. Standard deviations in brackets. Standard errors in parentheses.

# Balance in Subsequent ACPs

Table 2: Balance in Later ACPs

	$Z_2$	$Z_3$	$Z_4$	$Z_5$	$Z_6$	$Z_7$	$Z_8$	$Z_9$	$Z_{10}$
Work (W)	0.009 (0.010)	-0.004 (0.011)	0.022 (0.011)	0.014 (0.012)	0.039 (0.012)	-0.003 (0.017)	-0.011 (0.018)	0.022 (0.019)	0.030 (0.024)
Work (P)	0.006 (0.010)	0.016 (0.010)	0.012 (0.012)	0.020 (0.012)	-0.004 (0.015)	-0.004 (0.015)	-0.019 (0.019)	0.017 (0.020)	0.030 (0.027)
Hours (W)	32.885 (18.721)	-4.482 (20.032)	52.999 (21.045)	41.332 (22.686)	81.957 (25.131)	11.894 (31.187)	-18.836 (32.937)	72.659 (38.210)	24.819 (48.490)
Hours (P)	21.655 (21.018)	24.730 (21.089)	23.756 (23.574)	38.965 (25.255)	9.666 (30.585)	-6.580 (31.513)	-28.458 (37.976)	30.525 (44.856)	43.722 (52.821)
Income 1000s € (W)	1.481 (0.615)	-0.015 (0.624)	1.685 (0.767)	1.802 (0.830)	2.086 (0.913)	0.150 (1.000)	-0.043 (1.092)	0.866 (1.234)	-0.444 (1.629)
Income 1000s € (P)	-0.749 (0.835)	1.002 (0.912)	2.040 (1.066)	0.800 (1.115)	0.774 (1.424)	0.025 (1.424)	0.259 (1.563)	-0.324 (1.737)	0.149 (2.203)
Observations	12,974	10,774	8,726	6,977	5,411	3,944	2,723	1,850	1,174
Joint $p$ -val.	0.175	0.976	0.234	0.303	0.140	1.000	0.956	0.704	0.917

Note: Each column describes the difference in average characteristics between women for whom the respective ACP succeeds and those for whom it fails, among those who undergo the procedure, using inverse probability weights for each ACP following the main specification. Labor market outcomes and age measured year before first treatment. (W) - woman, (P) - partner. Standard errors in parentheses.

# Estimated Success Probabilities

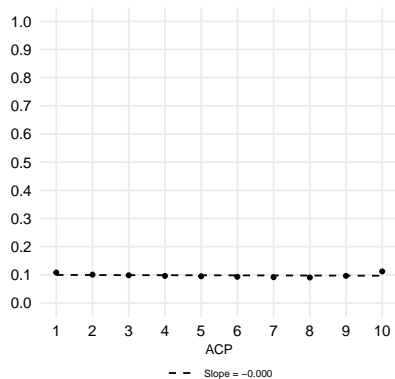


Figure 6: Estimated Success Probabilities

# Comparison to Representative Sample

Table 3: Full Sample, Reliers, and Representative Sample

	Success (1)	Fail (2)	Reliers (3)	Rep. (4)	Success vs rep. (1)-(4)	Rel. vs rep. (3)-(4)
Work (W)	0.882 [0.323]	0.863 [0.344]	0.820 [0.333]	0.801 [0.399]	0.080 (0.010)	0.019 (0.005)
Work (P)	0.884 [0.320]	0.865 [0.342]	0.849 [0.344]	0.783 [0.412]	0.101 (0.010)	0.066 (0.005)
Hours (W)	1240.315 [604.666]	1207.860 [635.194]	1117.711 [582.334]	1076.204 [696.245]	164.111 (16.856)	41.508 (8.412)
Hours (P)	1474.530 [658.231]	1438.590 [695.692]	1390.699 [662.920]	1250.948 [793.536]	223.582 (19.211)	139.752 (9.576)
Income 1000s € (W)	28.065 [19.559]	27.418 [20.219]	24.976 [15.359]	21.362 [18.330]	6.703 (0.444)	3.615 (0.222)
Income 1000s € (P)	37.205 [26.482]	36.952 [29.452]	35.299 [24.304]	28.107 [29.076]	9.098 (0.704)	7.193 (0.351)
Bachelor deg. (W)	0.480 [0.500]	0.451 [0.498]	0.398 [0.411]	0.411 [0.492]	0.069 (0.012)	-0.012 (0.006)
Bachelor deg. (P)	0.394 [0.489]	0.381 [0.486]	0.329 [0.397]	0.345 [0.475]	0.049 (0.012)	-0.015 (0.006)
Age (W)	31.638 [4.015]	32.388 [4.383]	33.480 [3.897]	28.713 [4.658]	2.926 (0.113)	4.767 (0.056)
Age (P)	34.675 [5.513]	35.461 [5.996]	36.580 [3.928]	28.713 [4.665]	5.962 (0.113)	7.868 (0.057)
Observations	1,714	13,809	4,882	376,152		

Note: Labor market outcomes measured year before first ACP for main sample and year and 9 months before birth of first child for the representative sample. Representative sample is selected to match the main sample by year of conception. Average relier outcomes are based on sample of women who remain childless 7 years after their first ACP with weights described under implementation. (W) - woman, (P) - partner. Standard deviations in brackets. Standard errors in parentheses.

# ACP Histories



Figure 7: ACP Histories and Reliance



# What are the counterfactuals?

Broadly:

- ▶ Do not want/plan children
- ▶ **Want/plan children**

Motherhood outcome:

- ▶ Get immediately
- ▶ Get naturally after few attempts
- ▶ **Get with medical assistance**

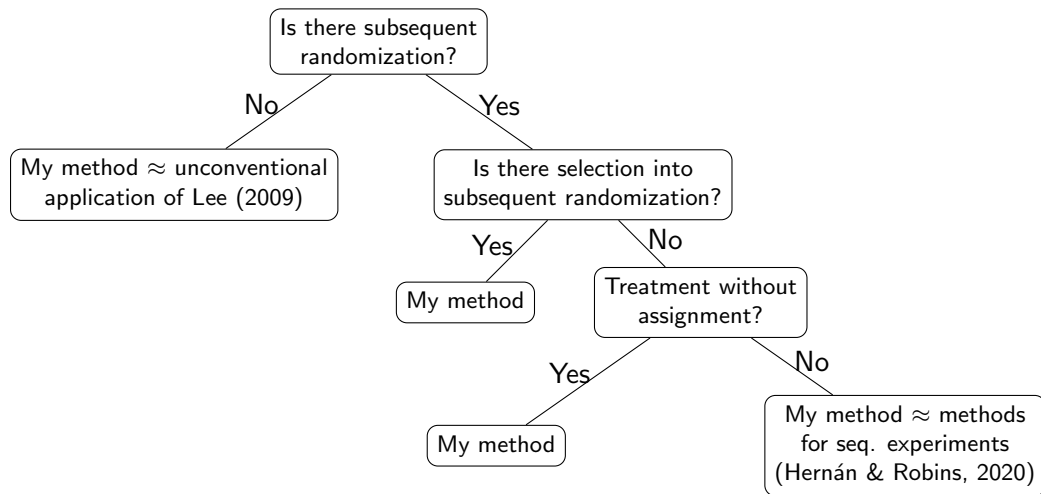
Childless outcome:

- ▶ Do not try
- ▶ Try and fail naturally
- ▶ **Try and fail with medical assistance (+ naturally?)**

Extrapolation requires carefully addressing mental health consequences of failure (and medical procedures)

[Back \(model\)](#) [Back \(extensions\)](#) [Depr. effect](#) [Bounds non-depr.](#) [Arguments](#)

## Relation to Methodological Literature



# Comparison with Lee (2009)

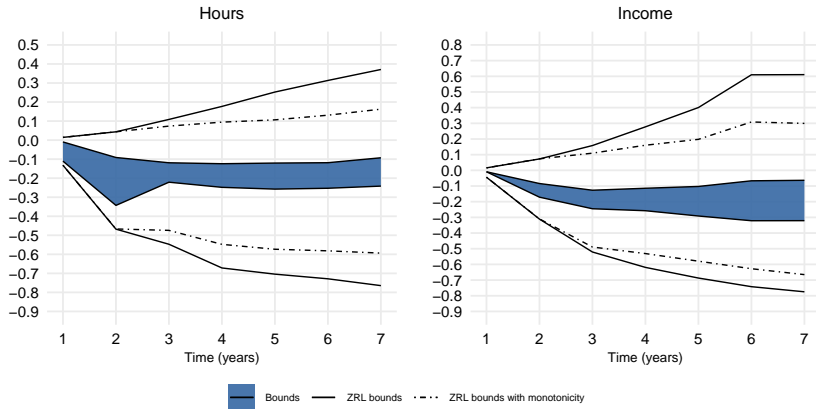


Figure 8: Comparison with Lee (2009) Bounds for Effects on Women

# Less Naive Comparison to Existing Methods

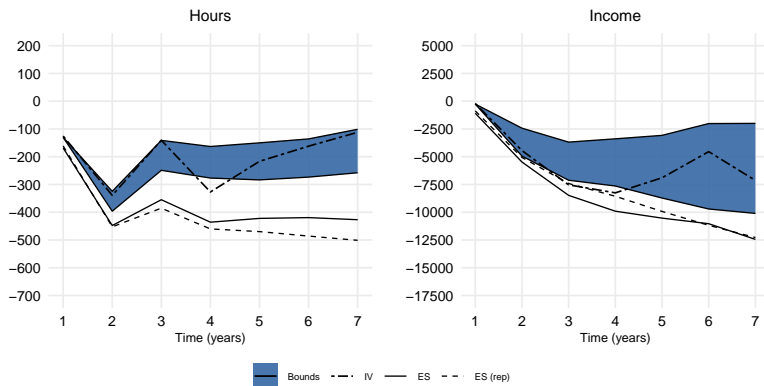


Figure 9: Estimates Using on Different Methods (Absolute Effects)

- ▶ Using women whose ACP succeeds for ES makes treatment definition consistent
- ▶ The three methods still target different sub-populations

# Quantifying Bias in Existing Methods

Instrumental Variable (Lundborg et al., 2017):

- ▶ I have bounded  $\tau_{ATR}$  allowing for dynamic effects
- ▶ I can point-identify  $\tau_{ATR}$  assuming static effects à la instrumental variable
- ▶ Subtracting it from bounds on  $\tau_{ATR}$  gives bounds on bias

Event Study (Kleven et al., 2019):

- ▶ ES imputes childless career trajectories from pre-parenthood outcomes of later mothers
- ▶ I proxy when women chose to have children with ACP timing
- ▶ Relier childless trajectories from my method vs imputed à la event study from older relies
- ▶ Comparing the two identifies the bias

[Back \(extensions\)](#)

[Conclusion](#)

# Effect of Delaying Motherhood

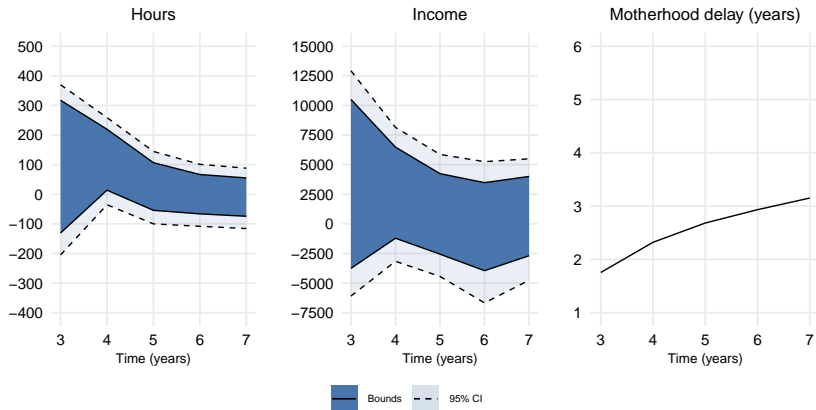


Figure 10: Effect of Delaying Relative to Motherhood at First Attempt (Absolute)

$$\tau_{IV} = \tau_{Parenthood} - 3\tau_{Delay}$$

# Placebo Event

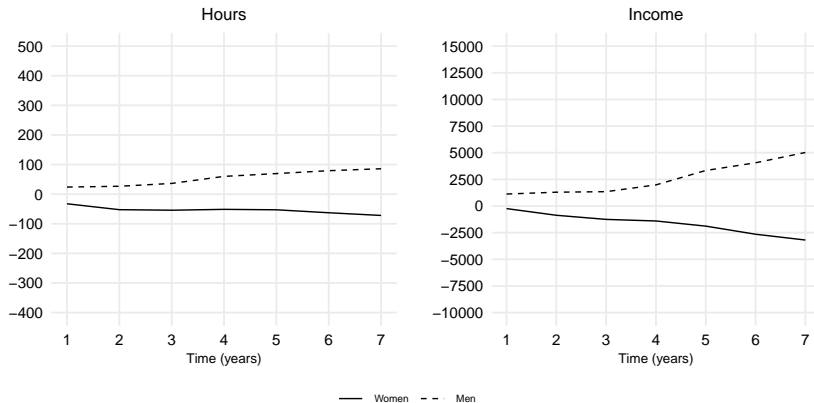


Figure 11: Placebo Event Study (Absolute)

- Negative selection of early mothers and positive selection of fathers

# Gender Inequality: Parenthood and Selection



Figure 12: Share of Gender Inequality Explained by Section and Parenthood

- Consistent with ES estimates attributing almost all gender inequality to parenthood



# Mental Health and ACPs

Mental health consequences associated with failure to conceive are a part of the story:

- ▶ Unmet fertility goals may negatively impact mental health, and in turn, labor market outcomes

There are, however, additional concerns:

- ▶ Mental health issues caused specifically by failed conception or ACPs (external)
  - ▶ Focusing on artificial insemination helps mitigate this
- ▶ Large impacts unique to ACP families (external)
- ▶ Worsened mental health by threatening monotonicity (internal)

In practice, these impacts are likely small (Lundborg et al., 2024)

[Antidepressant uptake](#)

[Back \(extensions\)](#)

[Conclusion](#)

# Monotone Bounds for Non-depressed Childless Women



**Figure 13:** Monotone Bounds for Women Who Would Not Uptake Antidepressants if They Were to Remain Childless

# Confidence Intervals

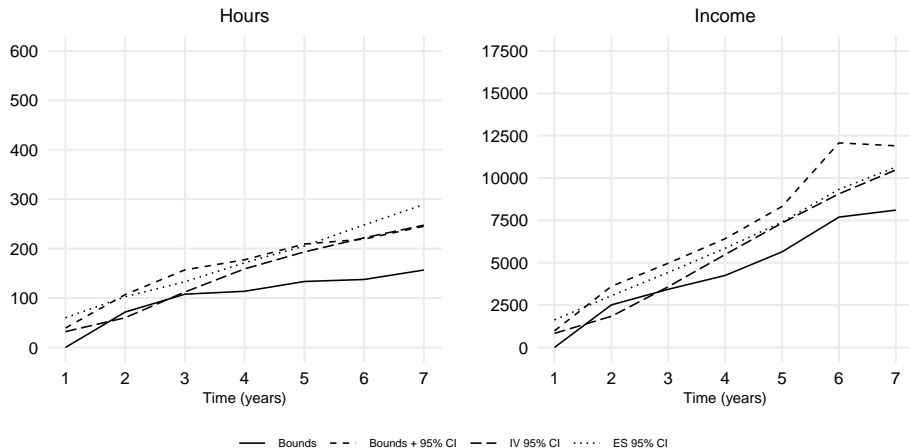


Figure 14: 95% CI for Different Methods (Absolute)

# Monotonicity

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 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

[Back](#)

# Monotone Bounds: Women who Remain Childless

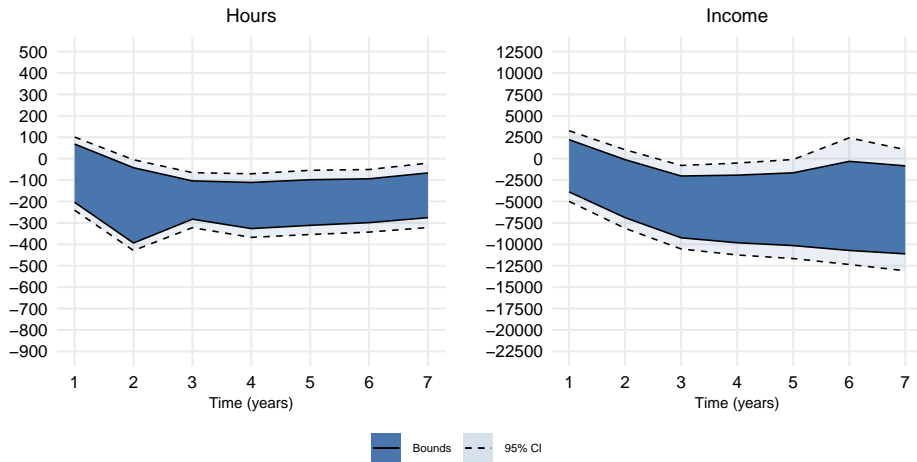


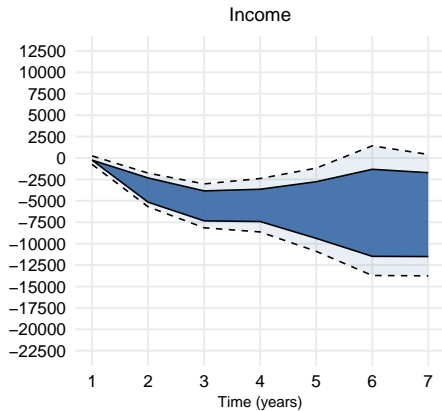
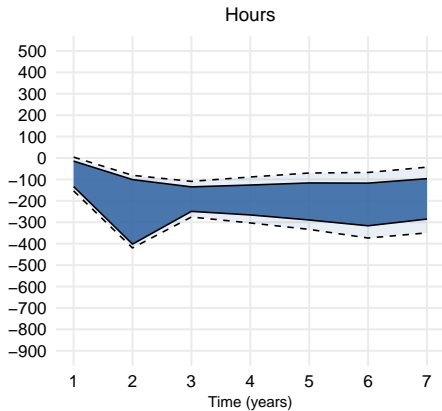
Figure 15: Monotone Bounds Using Completed Fertility (Absolute)

# Simple estimator



- ▶  $\mathbb{E}[Y(1)|R = 1] = \mathbb{E}[g(X_1) + \varepsilon|R = 1]$
- ▶  $\mathbb{E}[g(X_1)|R = 1]$  identified on chillness reliers using baseline method
- ▶ Only need to bound  $\mathbb{E}[\varepsilon|R = 1]$

# Relaxing Monotonicity Direction



— Bounds — 95% CI

[Back](#)

[ToC](#)

# Relaxing Monotonicity to Partnered Women

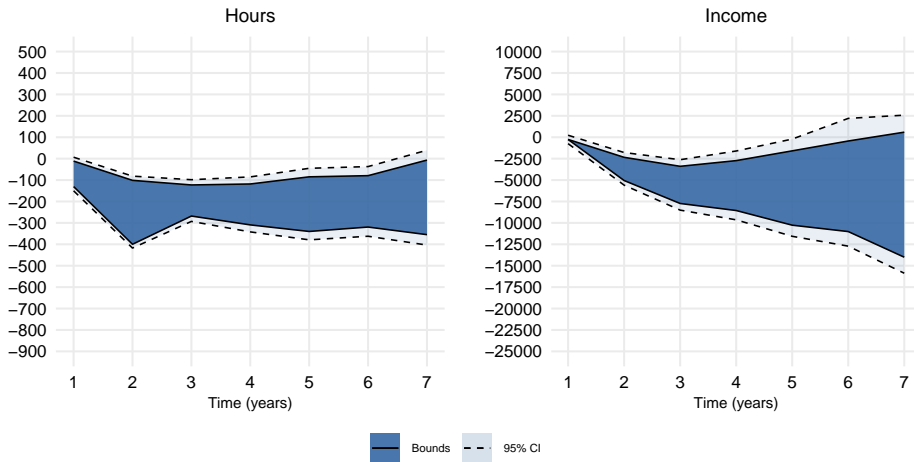


Figure 16: Monotone Bounds Using Women Who Stay Partnered



# Relaxing Monotonicity for Depression and Partnership

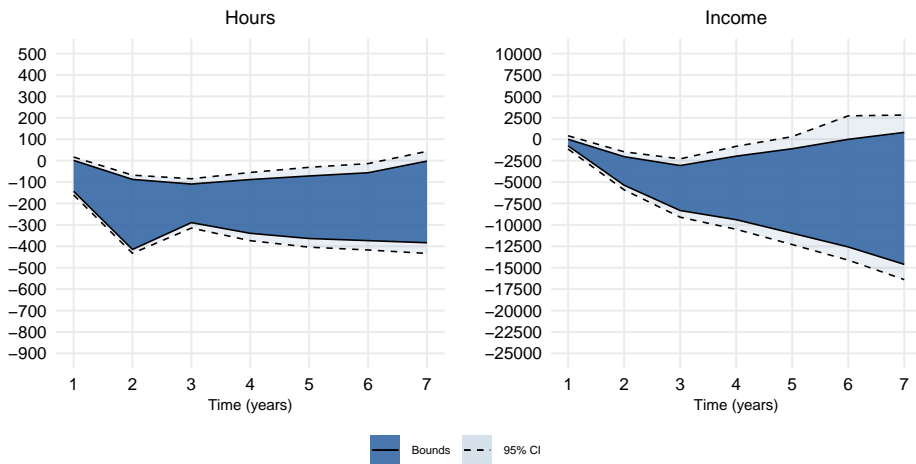
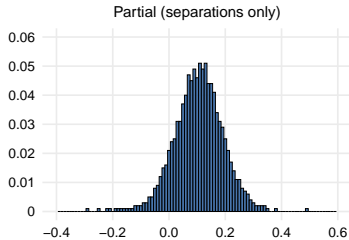
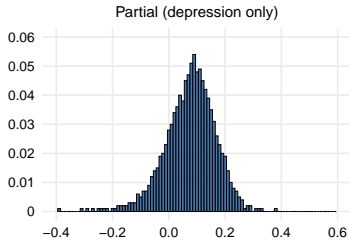
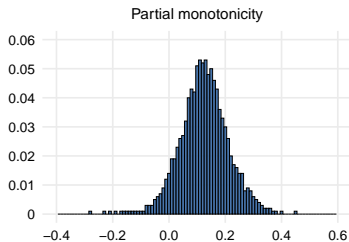
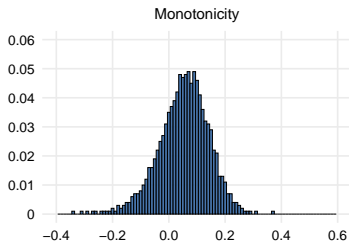


Figure 17: Monotone Bounds For Women Who Stay Partnered and Do Not Uptake Antidepress.

# Testing Monotonicity



# Heterogeneity by Willingness to Undergo Procedures



Figure 18: Cumulative Outcomes 6 Years After,  $G$  Above or Below 6

# Monotone Bounds: Excluding Depression



Figure 19: Effect on Antidepressant Take-Up

# Monotone Bounds: Correcting for Partner's age



Figure 20: Monotone Bounds Using Male Income at Same Age as Female

# Testing the Plug-in Approach

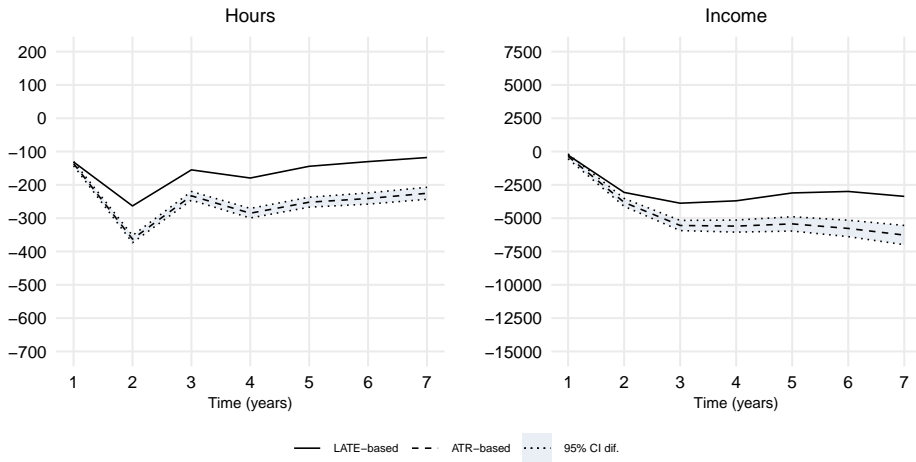


Figure 21: Plug-In Estimators Exploiting Different Numbers of Treatments

# Event Study: Population vs IUI Sample

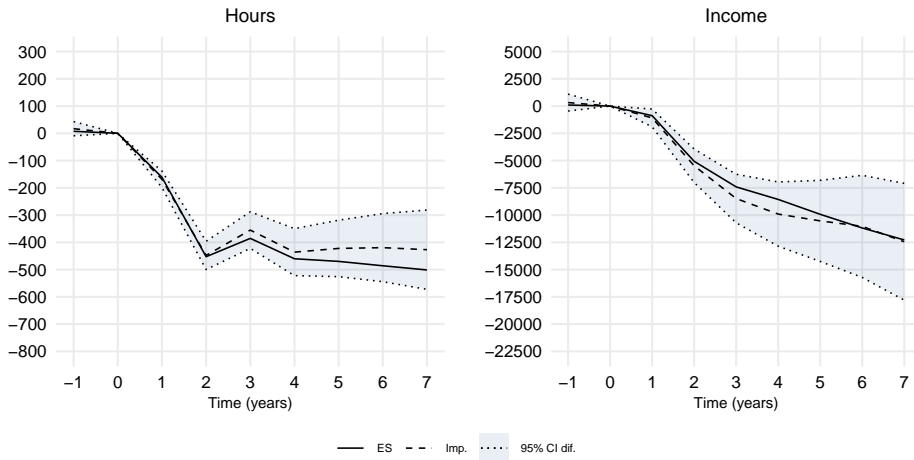


Figure 22: ES for Population and Women with First IUI Success

# Imputing Population Motherhood Outcomes Using IUI Sample



Figure 23: Population Outcomes vs IUI-Imputation (Age & Education)



# Imputing Population Childless Outcomes Using IUI Sample

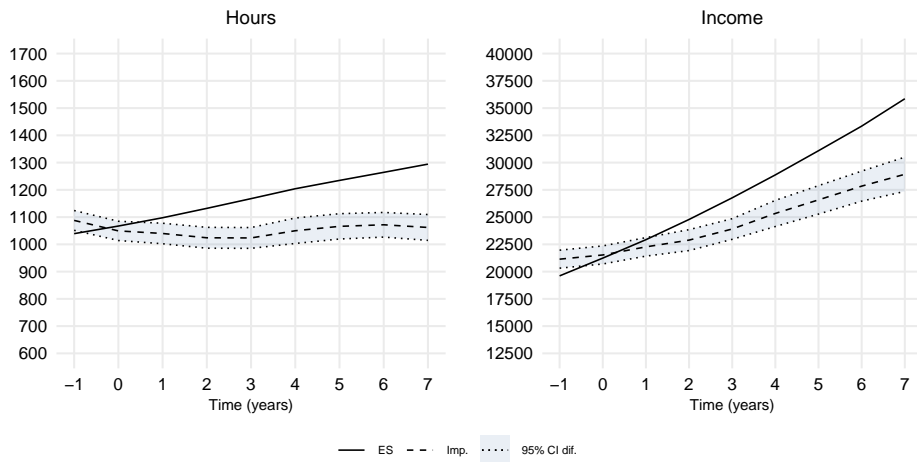


Figure 24: Population Outcomes vs IUI-Imputation (Age & Education)

# Event Study vs IUI-imputation for Population



Figure 25: Event Study vs IUI-Imputation for Population (Age & Education)

# References I

- Adda, J., Dustmann, C., & Stevens, K. (2017). The career costs of children. *Journal of Political Economy*, 125(2), 293–337.
- 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.
- Bensnes, S., Huitfeldt, I., & Leuven, E. (2023). Reconciling estimates of the long-term earnings effect of fertility. *IZA Discussion Paper*.
- Bertrand, M. (2020). Gender in the twenty-first century. *AEA Papers and proceedings*, 110, 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.
- Bütikofer, A., Jensen, S., & Salvanes, K. G. (2018). The role of parenthood on the gender gap among top earners. *European Economic Review*, 109, 103–123.
- Cortés, P., & Pan, J. (2023). Children and the remaining gender gaps in the labor market. *Journal of Economic Literature*, 61(4), 1359–1409.
- 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.
- Fitzenberger, B., Sommerfeld, K., & Steffes, S. (2013). Causal effects on employment after first birth—a dynamic treatment approach. *Labour Economics*, 25, 49–62.
- 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.
- Hernán, M. A., & Robins, J. M. (2020). *Causal inference: What if*. Boca Raton: Chapman & Hall/CRC.
- 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.
- Kleven, H., Landais, C., & Leite-Mariante, G. (2024). The child penalty atlas. *Review of Economic Studies*, rdae104.
- 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. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, 76(3), 1071–1102.

## References II

- 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.
- Lundborg, P., Plug, E., & Rasmussen, A. W. (2024). Is there really a child penalty in the long run? new evidence from ivf treatments. *IZA Discussion Paper*.
- Melentyeva, V., & Riedel, L. (2023). Child penalty estimation and mothers' age at first birth. *ECONtribute Discussion Paper*.
- Miller, A. R. (2011). The effects of motherhood timing on career path. *Journal of population economics*, 24, 1071–1100.
- Semenova, V. (2023). Generalized lee bounds. *arXiv preprint arXiv:2008.12720v3*.