Parenthood Timing and Gender Inequality

Julius Ilciukas

University of Amsterdam

Motivation

Parenthood can explain most of the gender inequality in Western labor markets (Kleven et al., 2024)

- Women's and men's careers follow similar trajectories up to parenthood
- Substantial gaps emerge at the birth of the first child

Family-friendly policies generally seem to have little effect on gender inequality (Cortés & Pan, 2023)

- Existing policies are inadequate to address the career impacts of parenthood?
- Or parenthood not primary cause of gender disparities?

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

Causal Identification Has Proven Challenging

Parenthood (timing) may be selective

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

Effects of parenthood may depend on parenthood timing

- Time spent in parenthood, career stage and age at the time of becoming a parent, age of the children

Addressing both simultaneously is an econometric challenge

Leading Methods to Quantify the Effects of Parenthood

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

Two methods dominate the debate:

1. Event study (Kleven et al., 2019)
2. IV using in vitro fertilization (Lundborg et al., 2017)

Event study:

Age 35, 1st child at 30	$\longleftrightarrow$ Direct comparison at 35	Age 35, 1st child at 36
All mothers for 5 years		All childless

- Clean comparison robust to dynamic effects
- Section is a concern: estimates reflect the effect of parenthood + difference in career trajectories in the absence of children

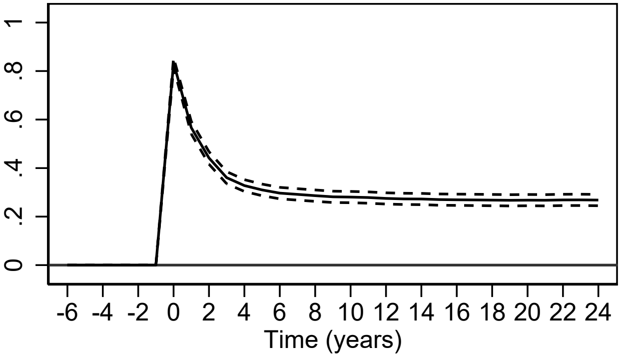
IV-IVF

- IVF is a type of assisted conception procedure
- Consider couples undergoing IVF for their first child
- In last procedure stage an embryo is transferred into uterus
- Assume conception from this transfer is as-good-as-random
  - Uncorrelated with past labor market outcomes cond. on age
- Success of first transfer creates plausibly exogenous variation in parenthood

Age 35, 1^{st} IVF at 30 worked	$\xleftrightarrow{\text{IV for parenthood at 35}}$	Age 35, 1^{st} IVF at 30 failed

- Dynamic effects are a concern: most women eventually conceive after the first procedure fails...

First Stage: 1st IVF Success and Motherhood

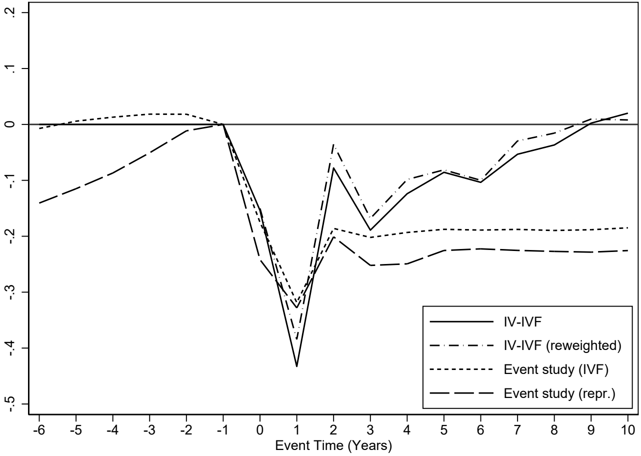


Source: Lundborg et al. (2024)

$$\tau_{RF} = \frac{1}{3}\tau_{Parenthood} + \frac{2}{3}\tau_{Timing}$$

$$\begin{aligned}\tau_{IV} &= \tau_{Parenthood} + 2\tau_{Timing} \\ &= 3\tau_{Parenthood} - 2\tau_{Later\ parenthood}\end{aligned}$$

IV vs ES: Earnings



Source: Lundborg et al. (2024)

Common Challenge in Applied Economics

1. Interested is the effect of program/treatment/intervention
2. Address selection leveraging (quasi-)experimental variation
3. Many (quasi-)experimental "shocks" are temporary
4. Conventional methods cannot separate the effect of treatment from the effect of treatment timing:
   (A) Focus on reduced form: weighted average of treatment effect and timing effect
   (B) Assume no timing effect: instrument treatment with assignment

Sometimes economic theory can suggest the direction of bias; not so much in the context of parenthood

This Paper

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

1. New approach using women's entire assisted conception procedures (ACP) histories
   - Robust to selective timing and dynamic effects
2. Empirical evidence using Dutch data
   - Focus on couples undergoing intrauterine insemination
3. Assess bias in existing methods
   - Quantify the extent of selectivity timing and dynamic effects
4. External validity

Literature

Gender inequality in labor market:

- Bertrand (2011); Blau & Kahn (2017) for overview

Literature on career impacts of children:

- Intensive margin: Rosenzweig & Wolpin (1980); Bronars & Grogger (1994); Angrist & Evans (1996); Jacobsen et al. (1999); Iacovou (2001); Cruces & Galiani (2007); Maurin & Moschion (2009); Hirvonen (2009); Vere (2011)
- Address selection: Hotz et al. (2005); Agüero & Marks (2008); Cristia (2008); Miller (2011); Brooks & Zohar (2021); Gallen et al. (2023)
- Address dynamic effects: Fitzenberger et al. (2013); Angelov et al. (2016); Adda et al. (2017); Chung et al. (2017); Bütikofer et al. (2018); Kleven et al. (2019); Eichmeyer & Kent (2022); Melentyeva & Riedel (2023)
- Selection and dynamic effects: Bensnes et al. (2023); Gallen et al. (2023)
    - Leverage women's first IVF, restrict effect heterogeneity across women and time via parametric assumptions, correct LR est. using SR est.

Methodological literature:

- Sequential experiments in biostatistics: Hernán & Robins (2020)
- Partial identification of treatment effects: Manski (1989, 1990); Zhang & Rubin (2003); Lee (2009)

Contribution

1. Estimates robust to selective fertility and dynamic effects
   - I demonstrate that accounting for these factors can reconcile the conflicting ES and IV estimates
2. External relevance
   - Netherlands vs Scandinavia
   - Intrauterine insemination vs IVF
3. Methodological approach applicable to various settings
   - Assignment to job training and educational programs, legal settings with assignment to varying leniency "judges", clinical trials in extension phase

Model: Outcomes

- Particular moment since woman's first ACP
- Outcome when motherhood begins at first ACP:

$$Y(1)$$

- Childless outcome:

$$Y(0)$$

- Outcome when motherhood begins after first ACP:

$$Y(2)$$

These scenarios involve women trying to conceive through ACPs

- I will first focus on quantifying impacts in these scenarios
- Extrapolation to scenarios where women choose to remain childless afterward (external validity)

Counterfactual

Model: Unobservables

Women differ in two unobserved characteristics:

- "Willingness" to undergo ACPs, $W \in \{1, \ldots, \overline{w}\}$
  - Would try W times total in case all ACPs fail
- "Reliance" on ACPs, $R \in \{0, 1\}$
  - No child if all ACPs fail, $R = 1$
  - "Reliers" $\supseteq$ "compliers" (no child if first ACP fails)

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:

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

- Parenthood indicator:

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

- Realized outcome:

$$Y = Y(0)(1 - D) + Y(1)DZ_1 + Y(2)D(1 - Z_1)$$

Treatment Effect

Average treatment effect for reliers:

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

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

- $Y(1), Y(0), R$ and W can be related
- Main method relaxes to covariate-conditional version

Outline of the Approach

Objective is a clean comparison:

- Women who conceive at first ACP vs similar childless women
- Challenge: selection into parenthood after 1st ACP fails
    1. Conceptions via subsequent ACPs
    2. Conceptions via non-ACP means
- I address 1. by leveraging women's complete ACP histories
- Births from 2. may be selective, I use a bounding approach

To present intuition, I consider 1. and 2. in isolation

Simple World: Max 1 ACP, All Reliers

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

Simple World: Max 2 ACPs, All Reliers

$W = 1$
(willing to try once)

$Z_1 = 1$

$Z_1 = 0$

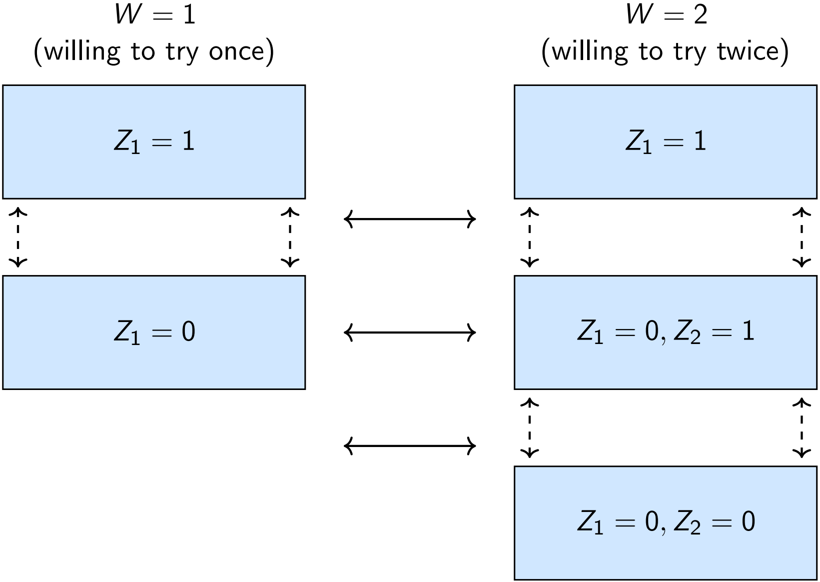
Simple World: Max 2 ACPs, All Reliers

$W = 1$ (willing to try once)	$W = 2$ (willing to try twice)
$Z_1 = 1$	$Z_1 = 1$
$Z_1 = 0$	$Z_1 = 0, Z_2 = 1$
	$Z_1 = 0, Z_2 = 0$

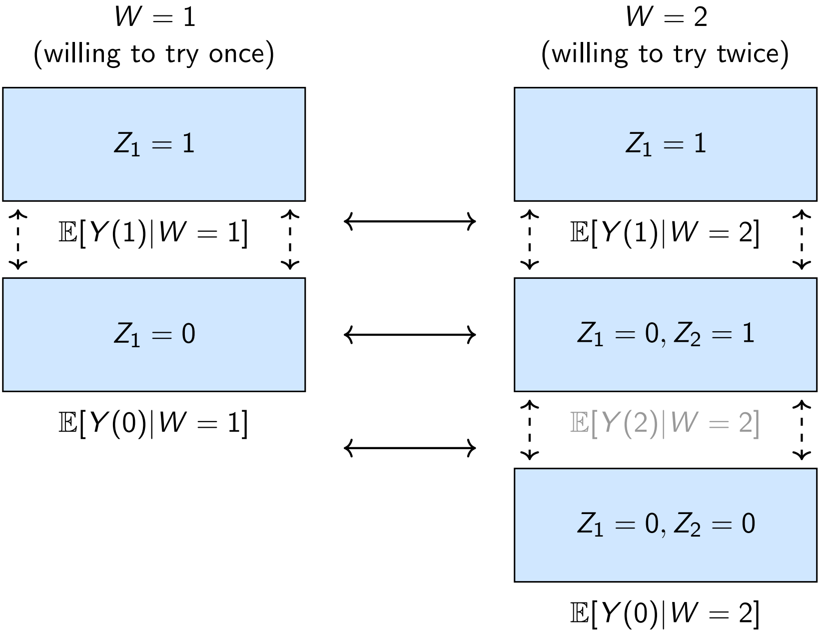
Simple World: Max 2 ACPs, All Reliers

$W = 1$ (willing to try once)		$W = 2$ (willing to try twice)
$Z_1 = 1$		$Z_1 = 1$
	$\longleftrightarrow$	
$Z_1 = 0$	$\longleftrightarrow$	$Z_1 = 0, Z_2 = 1$
	$\longleftrightarrow$	
		$Z_1 = 0, Z_2 = 0$

Simple World: Max 2 ACPs, All Reliers



Simple World: Max 2 ACPs, All Reliers



Simple World (Observed): Max 2 Attempts, All Reliers

$W = 1$ (willing to try once)	$W = 2$ (willing to try twice)
$Z_1 = 1$	$Z_1 = 1$
$Z_1 = 0$	$Z_1 = 0, Z_2 = 1$
	$Z_1 = 0, Z_2 = 0$

Simple World (Observed): Max 2 Attempts, All Reliers

$W = 1$
(willing to try once)

$W = 2$
(willing to try twice)

$Z_1 = 1$

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

$Z_1 = 0$

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

$Z_1 = 0, Z_2 = 1$

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

$Z_1 = 0, Z_2 = 0$

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

Simple World (Observed): Max 2 Attempts, All Reliers

$W = 1$
(willing to try once)

$W = 2$
(willing to try twice)

$Z_1 = 1$

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

$Z_1 = 0$

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

$Z_1 = 0, Z_2 = 1$

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

$Z_1 = 0, Z_2 = 0$

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

Simple World (Observed): Max 2 Attempts, All Reliers

$W = 1$
(willing to try once)

$W = 2$
(willing to try twice)

$Z_1 = 1$

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

$Z_1 = 0$

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

$Z_1 = 0, Z_2 = 1$

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

$Z_1 = 0, Z_2 = 0$

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

$Pr(W = 1) =$

Simple World: Max 1 ACP with Non-reliers

Simple World: Max 1 ACP with Non-reliers

$$W = 1$$

$R = 1$
(no child if fail)

$R = 0$
(child if fail)

Simple World: Max 1 ACP with Non-reliers

$W = 1$

$R = 1$
(no child if fail)

$R = 0$
(child if fail)

$Z_1 = 1$

Simple World: Max 1 ACP with Non-reliers

$W = 1$

$R = 1$
(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

$W = 1$

$R = 1$
(no child if fail)

$R = 0$
(child if fail)

$Z_1 = 1$

$Z_1 = 0, D = 0$

$Z_1 = 0, D = 1$

$Pr(R = 1) =$

Simple World: Max 1 ACP with Non-reliers

$W = 1$

$R = 1$
(no child if fail)

$R = 0$
(child if fail)

$Z_1 = 1$

$F_{Y(1)}$

$Z_1 = 0, D = 0$

$Z_1 = 0, D = 1$

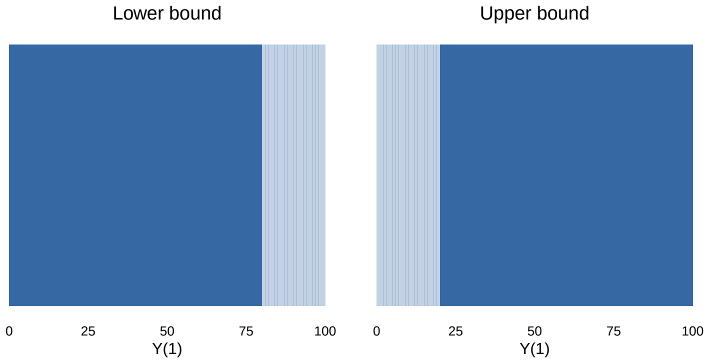
$Pr(R = 1) =$

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

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

Intuition: Motherhood Outcome $Y(1)$

1. Treated group is a rep. 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 →

- Baseline assumes reliers have the highest/lowest outcomes
- This might be inconsistent with covariate-conditional relier shares, making baseline bounds overly conservative

Inference complicated by trimming of the outcome distribution →

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

1. Construct orthogonal moments insensitive to estimation error in the trimming function
2. Estimate the trimming function using sample splitting

- Asymptotic inference as if the trimming function was known

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

Institutions

- Dutch family friendly polices 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

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 couples undergoing ACP for their first child between 2013 and 2016: 15,523

Overview of Descriptives

- First and subsequent ACP success uncorrelated with past labor market outcomes cond. on age Table first Table later
- Success probability stable across ACPs cond. on age Figure
- Representative sample worked less and had lower income before parenthood, but differences relatively small Table
  - ACP sample older before parenthood

Results: Bounds

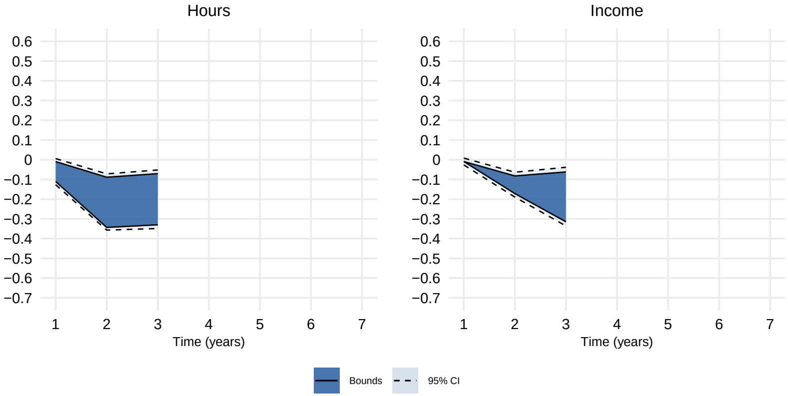


Figure 1: Bounds for women - short run

Results: Bounds

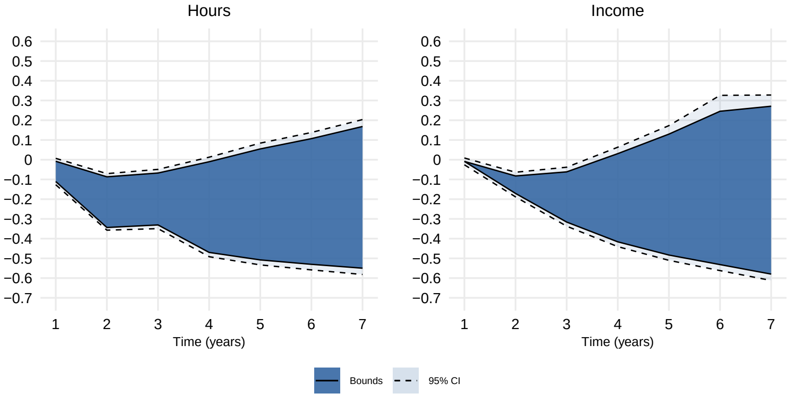
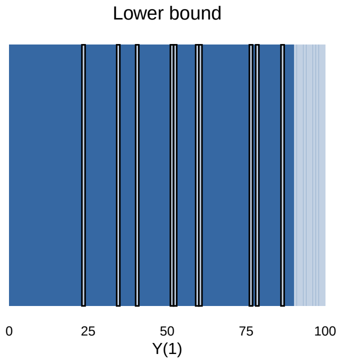


Figure 2: Bounds for women - medium run

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



Plausibility discussion

Bounds with Monotonicity

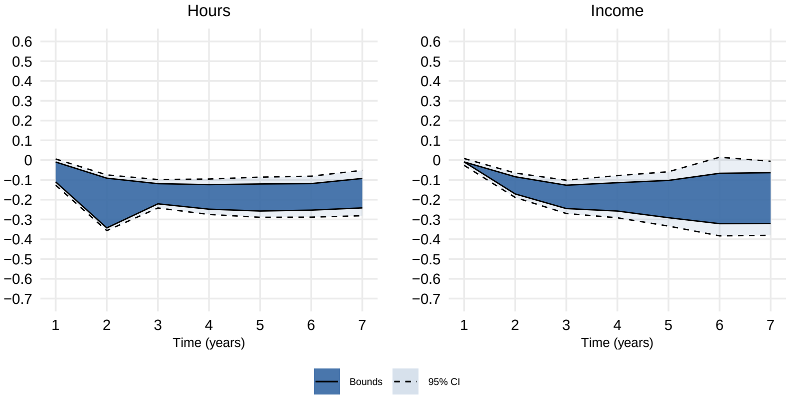


Figure 3: Bounds for women with monotonicity

Bounds for Men

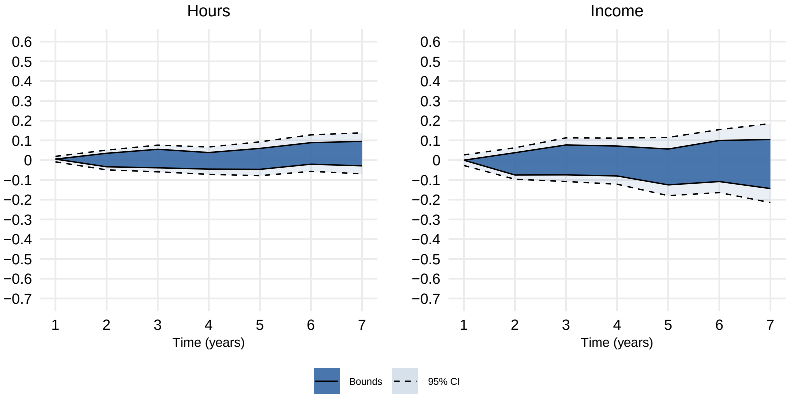


Figure 4: Bounds fo men

Gender Inequality

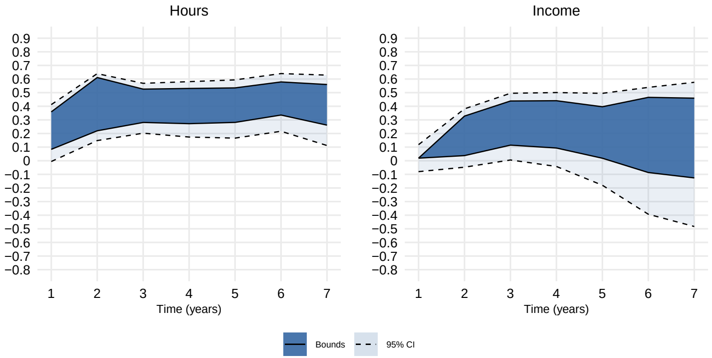


Figure 5: Share of gender inequality caused by parenthood

Next

- Relation to methodological literature →
- Bias in IV and ES →
- Mental health side effects →
- Conclusion →

Relation to Methodological Literature

Conceptions via ACPs: sequential experiments (Hernán & Robins, 2020)

- Key difference 1: treatment without assignment
- Key difference 2: selection into randomization (Van den Berg & Vikström, 2022)
- These methods are no applicable

Non-ACP conceptions: unobserved outcomes in experiments (Zhang & Rubin, 2003; Lee, 2009)

- Can be adapted to bound effects of parenthood for compliers
- Key advantage 1: my bounds cover a more general group
- Key advantage 2: my bounds are mechanically narrower
- Difference: unique monotonicity assumption
- I use Lee (2009) as a benchmark

Next Conclusion

Comparison with Lee (2009)

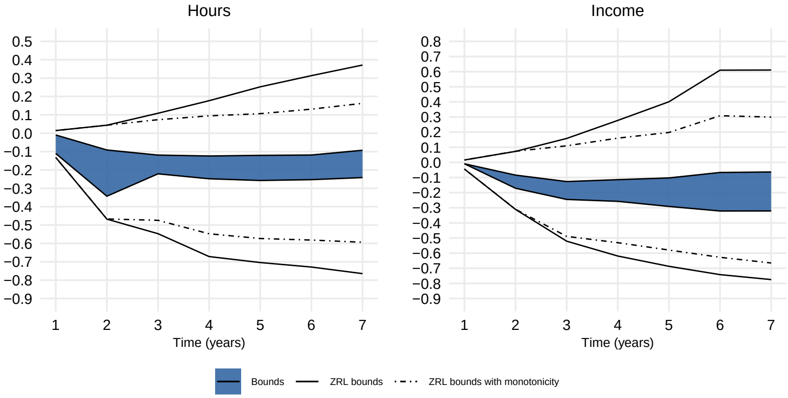


Figure 6: Comparison with Lee (2009) Bounds for Effects on Women

Next Conclusion

Naive Comparison to Existing Methods

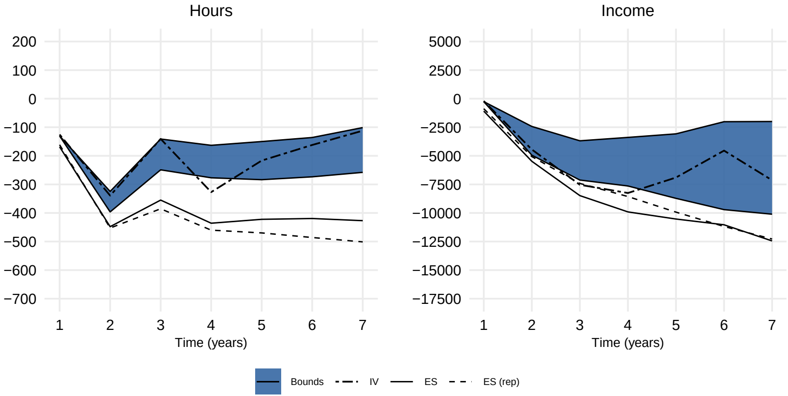


Figure 7: Estimates based on different methods

Next Conclusion

Quantifying Bias in Existing Methods

IV-IVF:

- Linear combination: effect of parenthood and effect of delaying parenthood
- I have bounded τ_{ATR}
- I can point-identify τ_{ATR} assuming statis effects
- Bounds on the effect of delaying parenthood

Event Study:

- Imputes childless career trajectories from pre-parenthood outcomes of older mothers
- I can construct representative group of childless reliers
- Placebo event study
- Quantify difference in career trajectories between women with different fertility timing in the absence of children

Next Conclusion

Effect of Delaying Motherhood

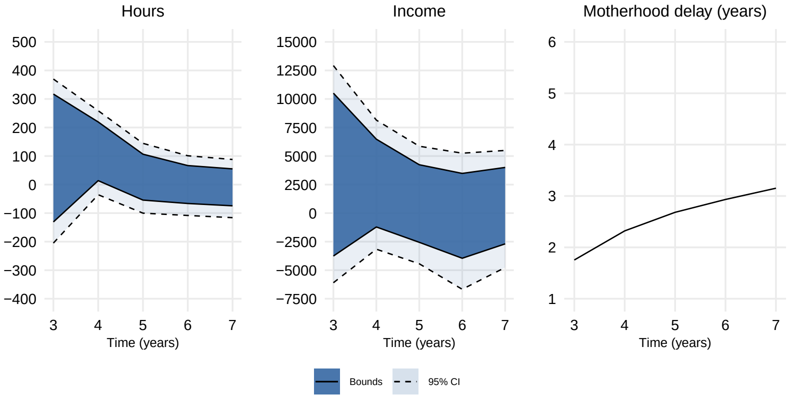


Figure 8: Effect of delaying relative to motherhood at first attempt

$$\tau_{IV} = \tau_{Parenthood} - 2\tau_{Delay}$$

Next Conclusion

Placebo Event

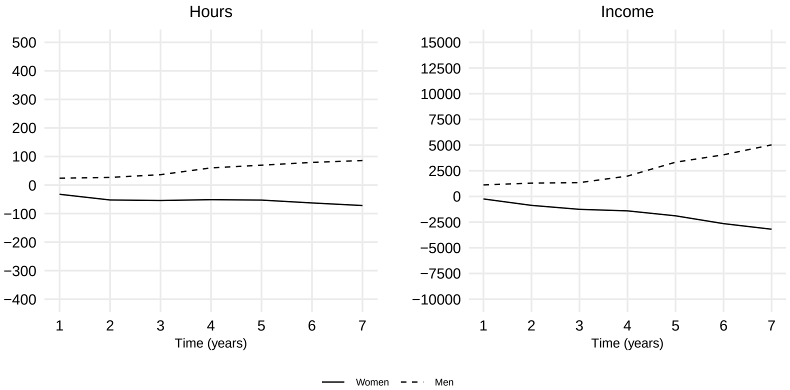


Figure 9: Placebo event study

- Negative selection of early mothers and positive selection of fathers

Next Conclusion

Gender Inquality: Parenthood and Selection

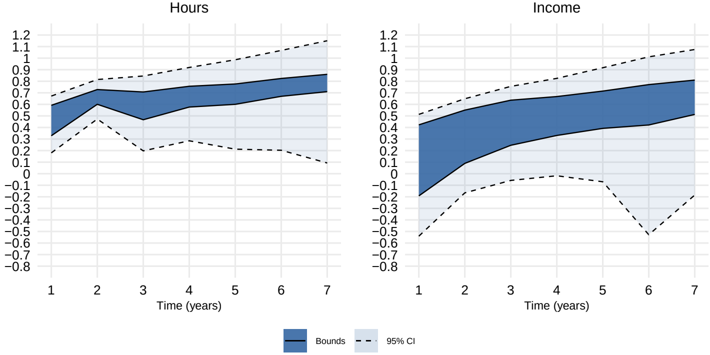


Figure 10: Share of gender inequality explained by section and parenthood

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

Next Conclusion

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 intrauterine insemination helps mitigate this
- Large impacts unique to ACP families (external)
- Worsened mental health by threatening monotonicity (internal)

In practice, these impacts may be small (Lundborg et al., 2024)

Antidepressant uptake Next Conclusion

Monotone Bounds for Non-depressed Childless Women

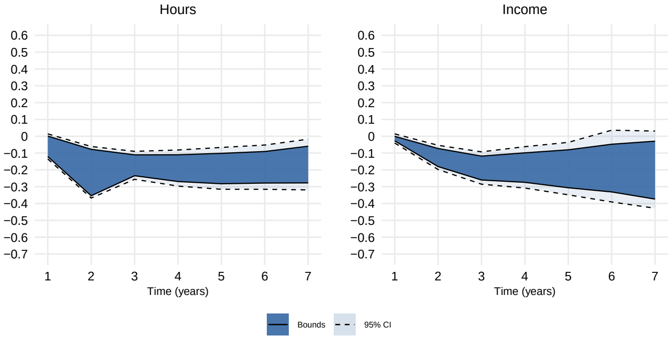


Figure 11: Monotone bounds for women who would not uptake antidepressants if they were to remain childless

Next Conclusion

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

Application to Dutch data:

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

Relative to conventional methods:

- IV-IVF might understate the career cost of motherhood
- ES might overstate the cost of motherhood and benefits of fatherhood
- Accounting for selection and dynamic effects can reconcile the results

Policy:

- Large share of gender ineq. may not be due to parenthood per se
- Family policies may still help by shaping behavior up to parenthood

Extensions

- Confidence inveral 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.

Bounding τ_{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
- η^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 τ_{ATR} *is* $\mathbb{E}[m^L(G, \eta^0)] / \mathbb{E}[r(X_1)]$*.*

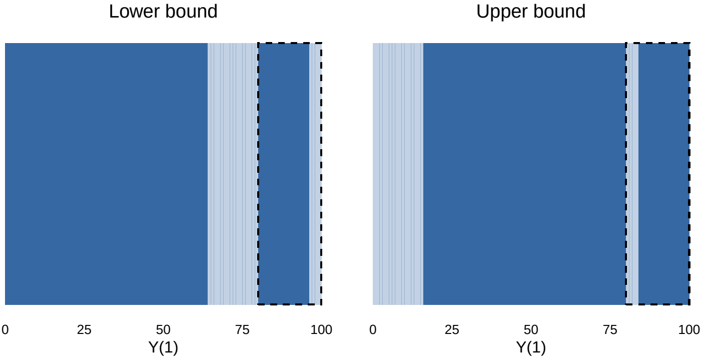
Back

ToC

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:

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

Identifies same parameter:

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

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

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

Back

ToC

Balance in 1^{st} ACP

Table 1: First ACP Outcomes and Descriptives

	Success	Fail	Difference	Dif. cond. age & educ.
	(1)	(2)	(1)-(2)	(1)-(2) cond.
Work (W)	0.882	0.863	0.019	0.008
	[0.323]	[0.344]	(0.009)	(0.009)
Work (P)	0.884	0.865	0.019	0.013
	[0.320]	[0.342]	(0.009)	(0.009)
Hours (W)	1240.315	1207.860	32.455	18.702
	[604.666]	[635.194]	(16.183)	(16.560)
Hours (P)	1474.530	1438.590	35.940	18.579
	[658.231]	[695.692]	(17.713)	(17.870)
Income 1000s € (W)	28.065	27.418	0.647	0.745
	[19.559]	[20.219]	(0.516)	(0.546)
Income 1000s € (P)	37.205	36.952	0.252	0.364
	[26.482]	[29.452]	(0.746)	(0.730)
Bachelor deg. (W)	0.480	0.451	0.029	
	[0.500]	[0.498]	(0.013)	
Bachelor deg. (P)	0.394	0.381	0.013	
	[0.489]	[0.486]	(0.012)	
Age (W)	31.638	32.388	-0.750	
	[4.015]	[4.383]	(0.111)	
Age (P)	34.675	35.461	-0.786	
	[5.513]	[5.996]	(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 prbability weights for the first ACP that follow the main specificaition. 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	-0.004	0.022	0.014	0.039	-0.003	-0.011	0.022	0.030
	(0.010)	(0.011)	(0.011)	(0.012)	(0.012)	(0.017)	(0.018)	(0.019)	(0.024)
Work (P)	0.006	0.016	0.012	0.020	-0.004	-0.004	-0.019	0.017	0.030
	(0.010)	(0.010)	(0.012)	(0.012)	(0.015)	(0.015)	(0.019)	(0.020)	(0.027)
Hours (W)	32.885	-4.482	52.999	41.332	81.957	11.894	-18.836	72.659	24.819
	(18.721)	(20.032)	(21.045)	(22.686)	(25.131)	(31.187)	(32.937)	(38.210)	(48.490)
Hours (P)	21.655	24.730	23.756	38.965	9.666	-6.580	-28.458	30.525	43.722
	(21.018)	(21.089)	(23.574)	(25.255)	(30.585)	(31.513)	(37.976)	(44.856)	(52.821)
Income 1000s € (W)	1.481	-0.015	1.685	1.802	2.086	0.150	-0.043	0.866	-0.444
	(0.615)	(0.624)	(0.767)	(0.830)	(0.913)	(1.000)	(1.092)	(1.234)	(1.629)
Income 1000s € (P)	-0.749	1.002	2.040	0.800	0.774	0.025	0.259	-0.324	0.149
	(0.835)	(0.912)	(1.066)	(1.115)	(1.424)	(1.424)	(1.563)	(1.737)	(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.

ACP histories Back

ToC

Estimated Success Probabilities

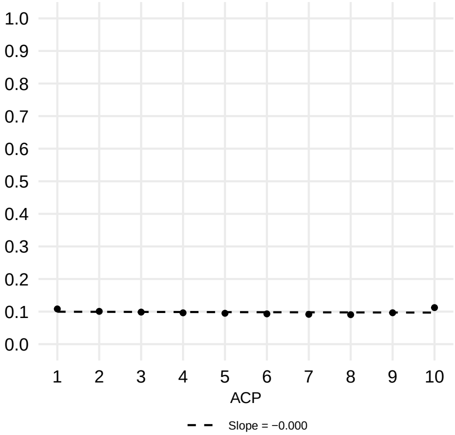


Figure 12: Estimated Success Probabilities

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

ToC Back

ACP Histories

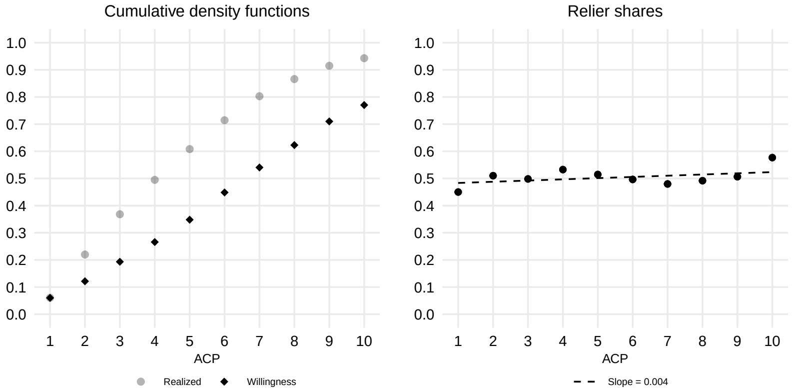


Figure 13: ACP Histories and Reliance

Back

ToC

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

Confidence Intervals

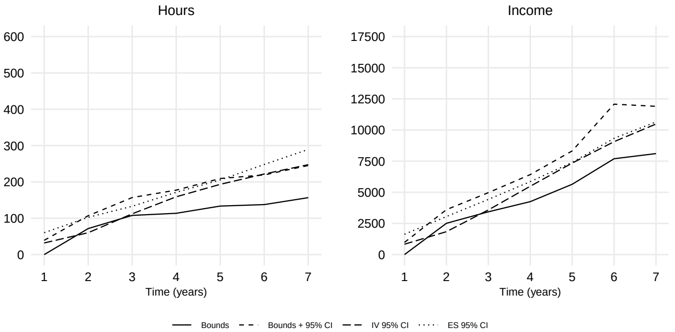


Figure 14: 95% CI for different methods

Back

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

ToC

Monotone Bounds: Women who Remain Childless

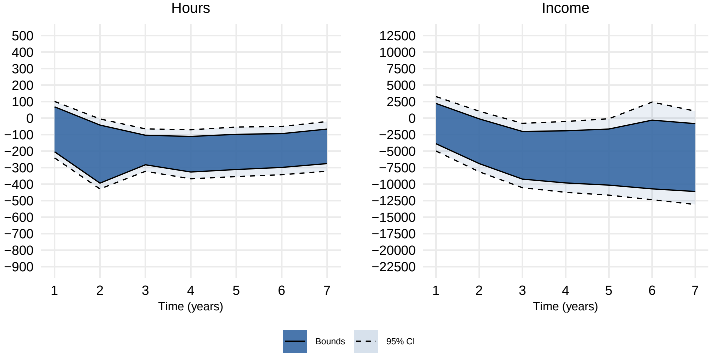
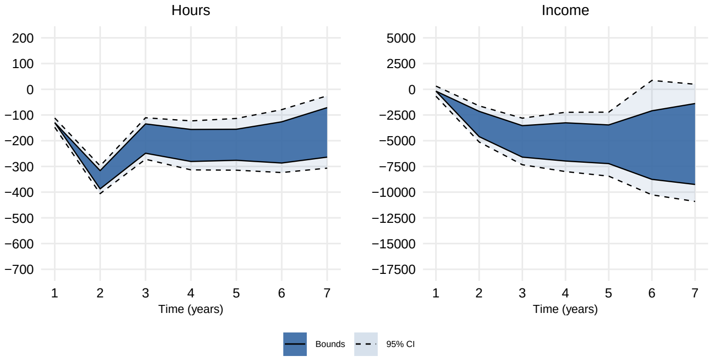


Figure 15: Monotone bounds using final status

Back

ToC

Simple estimator

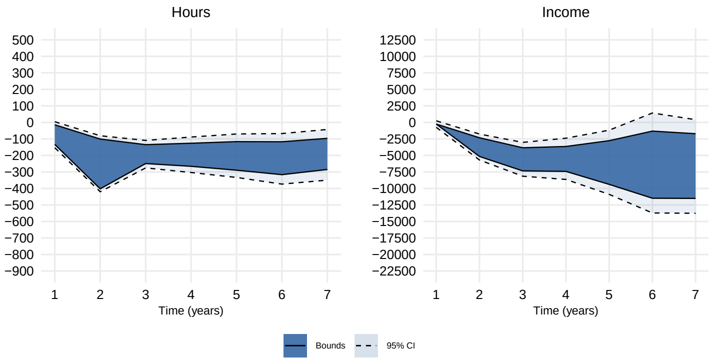


- $\mathbb{E}[Y(1)|R = 1] = \mathbb{E}[g(X_1) + \varepsilon|R = 1]$
- $\mathbb{E}[g(X_1)|R = 1]$ identified on chillness reliers using baseline method
- Only need to bound $\mathbb{E}[\varepsilon|R = 1]$

Back

ToC

Relaxing Monotonicity Direction



Back

Relaxing Monotonicity to Partnered Women

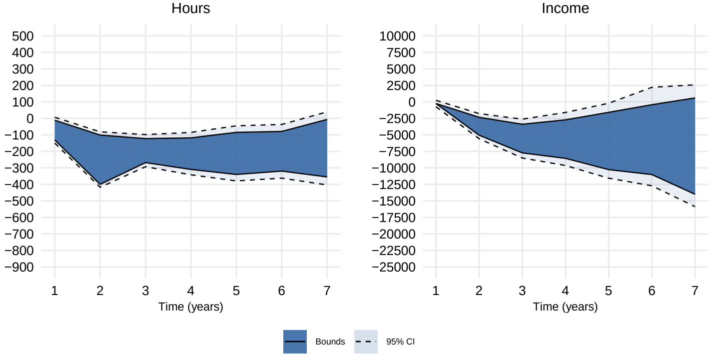


Figure 16: Monotone bounds using women who stay partnered

Back (extensions)

Relaxing Monotonicity for Depression and Partnership

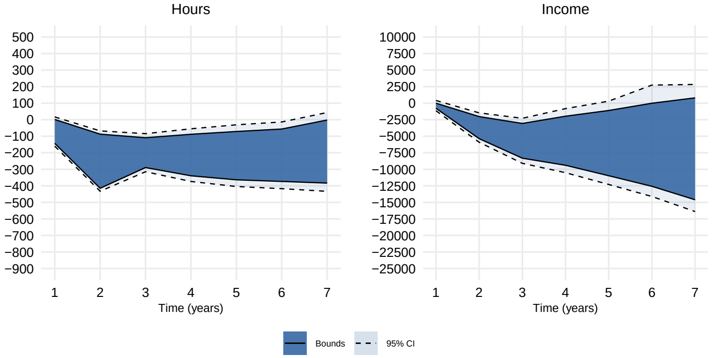
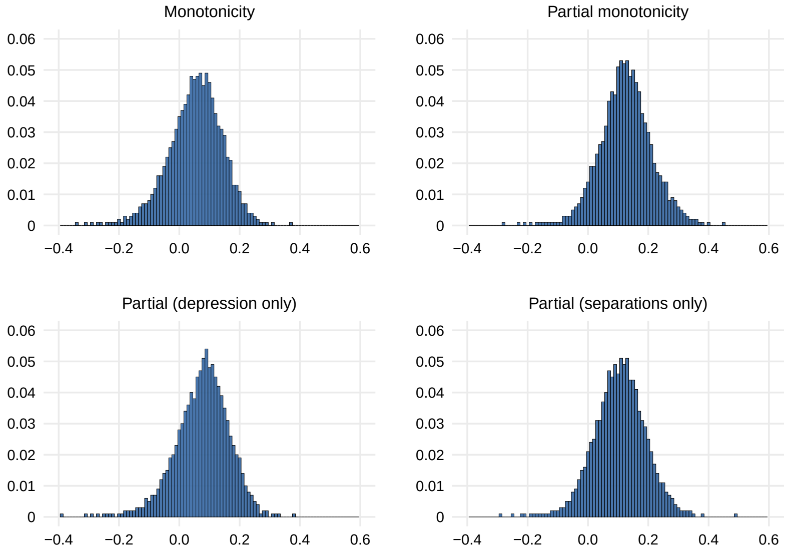


Figure 17: Monotone bounds using women who stay partnered and do not uptake antidepressants

Back (extensions)

ToC

Testing Monotonicity



Heterogeneity by Willingness to Undergo Procedures

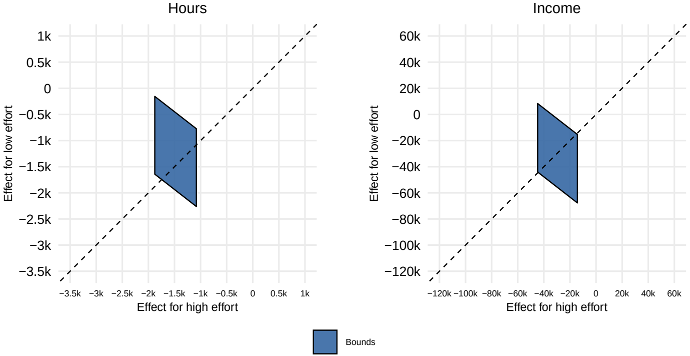


Figure 18: Cumulative outcomes 6 years after, G above or below 6

Back

Monotone Bounds: Excluding Depression

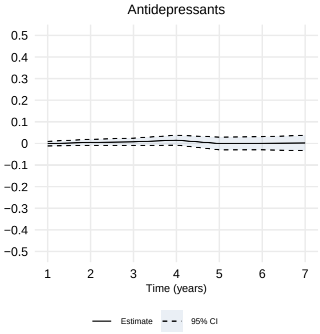


Figure 19: Effect on antidepressant take-up

Back (extensions) Back (mental health)

ToC

Monotone Bounds: Correcting for Partner's age

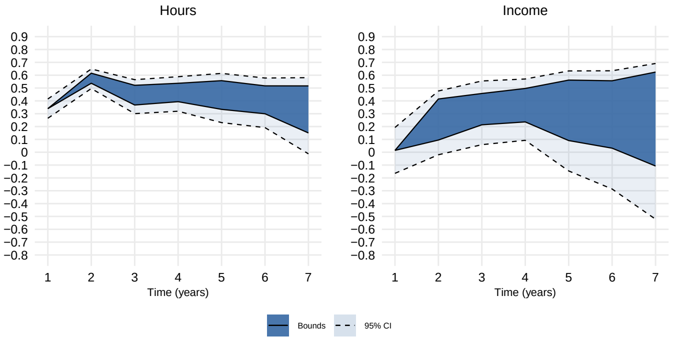


Figure 20: Monotone bounds using male income at same age as female

Back

Monotone Bounds: Women who Remain Childless

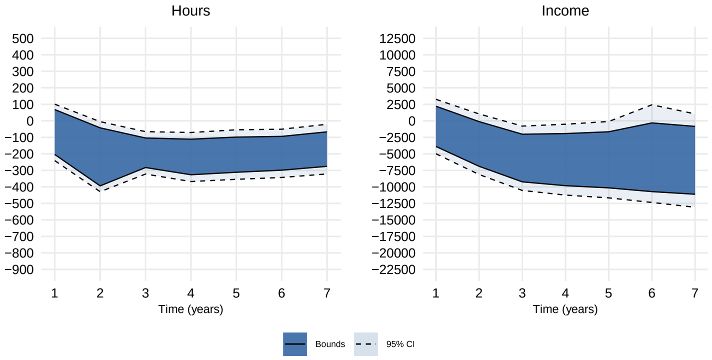


Figure 21: Monotone bounds using final status

Back

Testing the Plug-in Approach

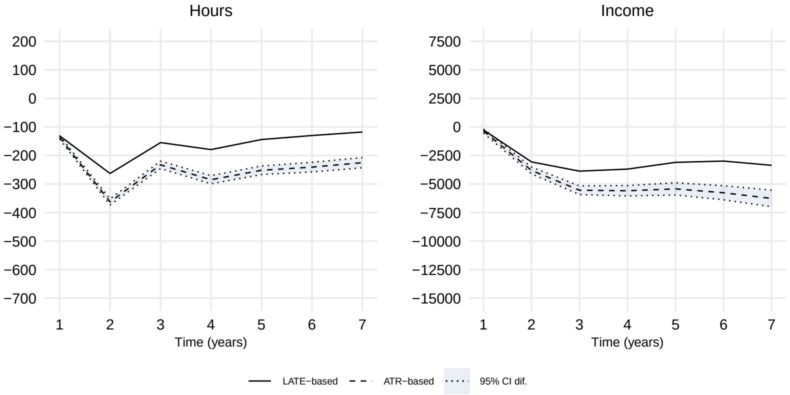


Figure 22: Plug-in estimators exploiting different number of treatments

Back

ToC

Event Study: Population vs IUI Sample

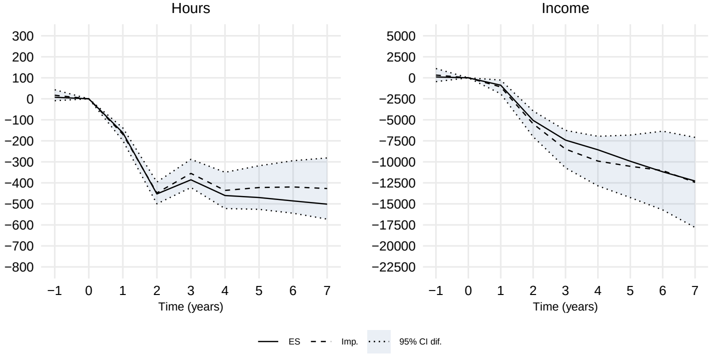


Figure 23: ES for population and women with first IUI success

Back (extensions) Back (intro)

ToC

Imputing Population Motherhood Outcomes Using IUI Sample

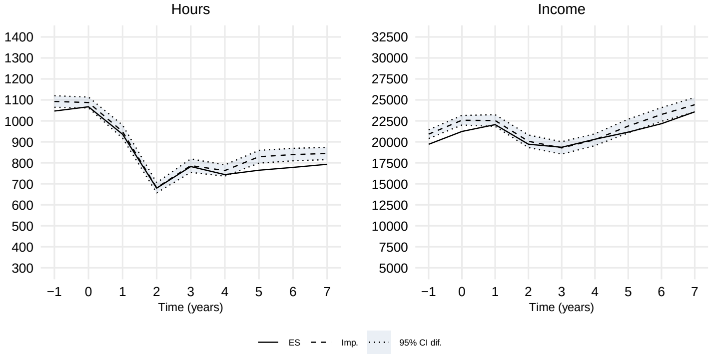


Figure 24: Population Outcomes vs IUI-imputation (age & education)

Imputing Population Childless Outcomes Using IUI Sample

Figure 25: Population Outcomes vs IUI-imputation (age & education)

Back

ToC

Event Study vs IUI-imputation for Population

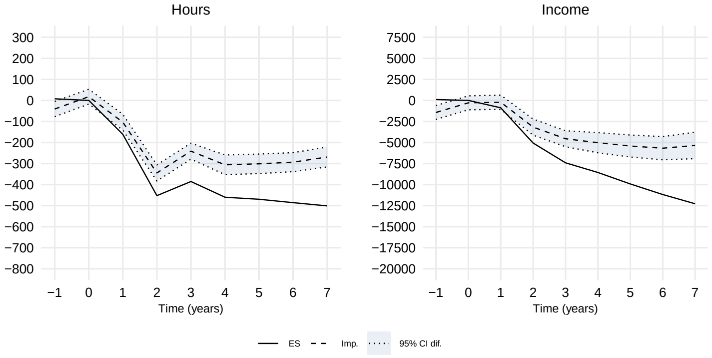


Figure 26: Event study vs IUI-imputation for population (age & education)

Back

ToC

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.

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

Bensnes, S., Huitfeldt, I., & Leuven, E. (2023). Reconciling estimates of the long-term earnings effect of fertility. *IZA Discussion Paper*.

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. *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.

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. *CEMFI Working Paper No. 2111*.

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.

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

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.

References II

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*.

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*.

Hernán, M. A., & Robins, J. M. (2020). *Causal inference: What if*. Boca Raton: Chapman & Hall/CRC.

Hirvonen, L. (2009). The effect of children on earnings using exogenous variation in family size: Swedish evidence. *Swedish Institute for Social Research Working Paper*.

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. *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.

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.

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*.

Manski, C. F. (1989). Anatomy of the selection problem. *Journal of Human resources*, 343–360.

References III

Manski, C. F. (1990). Nonparametric bounds on treatment effects. *The American Economic Review*, *80*(2), 319–323.

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

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.

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. (2023). Generalized lee bounds. *arXiv preprint arXiv:2008.12720v3*.

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.