The Effects of Parents' Receiving Unemployment Benefits on Children's Educational Attainment

Tiansui Tu (Sean)

University of Oxford

01/03/2023

Clarifying the question

Interested in

$$D \longrightarrow Y$$

$Y, D \in \{0, 1\}$

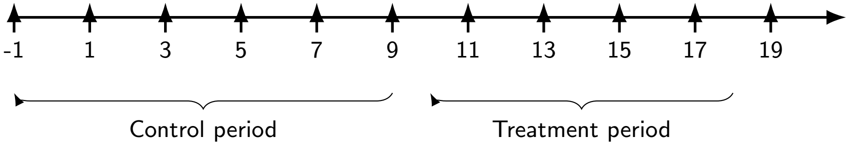
$D = 1$: mother (NOT father) receives a positive amount of UI benefit when the child is 10 to 18

$Y = 1$: child graduates from high school at 19

Focusing on

1. US: Unemployment Insurance (UI) provides cash benefits to workers who lose their jobs **through no fault of their own**, and they meet work/wage requirements.
2. Mothers unemployed for ≥ 1 week when the child is 10 to 18
3. Mothers average lagged wage ≥ 0 when the child is 10 to 18

Timeline



Preview of results

OLS: close to 0, NOT statistically significantly different from 0

Linear IV: negative, statistically significantly different from 0

Non-parametric IV: insignificantly negative

Falsification tests: evidence unclear, inclined to not favor the validity of IV, probably my IV is rubbish

For falsification tests, not sure how to cluster as they all assume i.i.d.

Simple Model

Mother is employed with probability s and unemployed with probability $1 - s$,

$$\max_{i,s} \log(e) - s^2$$

subject to

$$i = sw + (1 - s)b$$

$$e = i$$

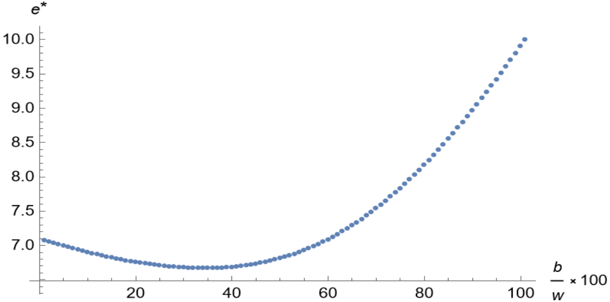
$$i \geq 0, 0 \leq s \leq 1$$

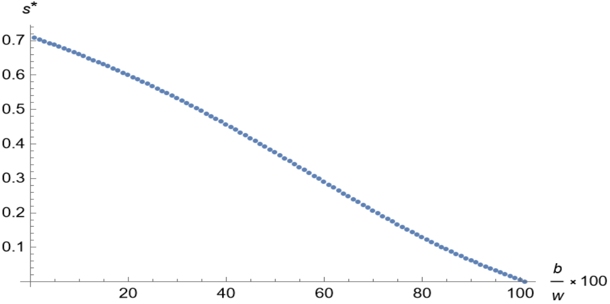
- s: search effort
- i: investment in child's education
- e: education output from the child
- w: wage
- b: UI benefit

What is $\partial e^* / \partial b$?

Comparative Statics

Set $w = 10$,





Possible causal channels

Among unemployed parents, UI affects child's education through:

Pecuniary channels:

$+$: Improving home material conditions

$-$: Prolonging parents' unemployment spell, losing potential income

Non-pecuniary channels:

$+$: Reducing parental stress, better family functioning

$-$: Stigma/ Exhibits negative role model effect on child

Ambiguous overall effect

(Ku and Plotnick, 2003) (Heflin and Acevedo, 2011)

Why OLS may not be sufficient?

E.g.

$$Y = \beta_0 + \beta_1 D + \beta_2 X + U$$

$X \in [0,1]$: fraction of years (child is -1 to 10) mother has high school degree

At least 2 selection biases: (Acevedo and Heflin, 2014)

1. Mothers eligibile for UI differ from those ineligible (conditioning on sample selection) in some characteristics

2. Among eligible mothers, mothers who take up UI differ from those who do not in some characteristics

These characteristics may be unobservable and related to both D and Y

OLS may overestimate the effect if mothers getting UI are better in some sense?

IV motivation

Dummy Z: whether mother unemployed due to business closure for more than once when child is 10 to 18

Relevance: eligible for UI if unemployed not because of the fault of your own

Exogeneity:

i. Cylus, Jonathan (2015) used this as an IV for receiving UI on health, balance table looks alright for PSID

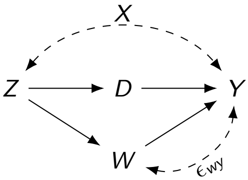
ii. Maybe some randomness/unexpectedness

iii. Even if ability $\rightarrow$ firm size $\rightarrow$ business closure, it is arguably not correlated with welfare preference, so better than nothing (maybe)

iv. No clear reason for direct effect, maybe exists indirect effect on Y

Causal diagram for IV

For unemployed mothers,



$Y, D, Z \in \{0, 1\}$

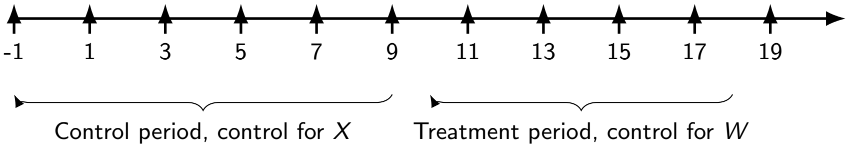
Control for X, W

Typically, chronological order is X, Z, W, Y so

Mostly, X from ages -1 to 9; W from ages 10 to 18

Possibly doomed if there exists **unobservable** ϵ_{wy} or $Z \rightarrow Y$ directly

Timeline



X may include "D" and "Z" from age -1 to 9

If ϵ_{wy} observed, may also control for it

Data

Longitudinal surveys: (`https://www.nlsinfo.org/content/cohorts`)

- NLSY79 cohort: 12,686 American youths ages 14-22 when first interviewed, available annually 1979-1994 and biennially 1996-2018
- NLSY79 Child and Young Adult cohort: currently included 11,545 children born to interviewed NLSY79 mothers

Transform it to cross-sectional:

- Y by definition fixed across time
- Take central tendency measures of control variables
- Unit is mother-child pair, independence across mother clusters
- Cannot use child weights

Sample selection:

- Mothers unemployed/have positive average wage
- Filter out the observations of child being 17 (or 18) and graduating at 17 (or 18), very few obs

Y, D, Z distribution

```
. tab Y
```

Y	Freq.	Percent	Cum.
0	917	44.19	44.19
1	1,158	55.81	100.00
Total	2,075	100.00	

```
. tab D
```

D	Freq.	Percent	Cum.
0	1,246	60.05	60.05
1	829	39.95	100.00
Total	2,075	100.00	

```
. tab Z
```

Z	Freq.	Percent	Cum.
0	1,820	87.71	87.71
1	255	12.29	100.00
Total	2,075	100.00	

Linear IV with homogeneous effects

Stable Unit Treatment Value Assumption **violated**:
(i) effect of mothers' receiving UI once different from those receiving UI twice;
(ii) receiving UI at child age 15 less impactful than 18
Not sure how this affects the results

Relevance: $Cov(D, Z) \neq 0$?

Exogeneity: $Cov(U, Z) = 0$?

$Cov(D, Z) \neq 0$?

```
.          reg D Z, r cluster(mid)

Linear regression                               Number of obs     =      2,075
                                                F(1, 1312)        =      26.10
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0199
                                                Root MSE          =     .48512

                               (Std. err. adjusted for 1,313 clusters in mid)
```

D	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
Z	.2106874	.0412409	5.11	0.000	.129782	.2915927
_cons	.3736264	.0145581	25.66	0.000	.3450667	.4021861

Likely holds, also robust to adding controls (not shown)

$Cov(U, Z) = 0?$

Subsample of mothers unemployed because of either laid off or closure, where $Z \not\to D$, so now check if $Z \to Y$ (reg D on Z gives -.016 with pval 0.724)

```
.          reg Y Z if layfclos == 1, r cluster(mid)

Linear regression                               Number of obs     =        871
                                                F(1, 591)         =       9.49
                                                Prob > F          =     0.0022
                                                R-squared         =     0.0126
                                                Root MSE          =     .49566

                                   (Std. err. adjusted for 592 clusters in mid)
```

Y	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
Z	-.1230201	.0399394	-3.08	0.002	-.2014605	-.0445797
_cons	.5779221	.0220729	26.18	0.000	.5345713	.6212729

Violated, so need to condition on X or W

$Cov(U, Z) = 0$?

```
.         reg Y Z ///
>                 hs uni afqt lwage_ lnfaminc_ cfemale if layfclos == 1, r cluster(mid)

Linear regression                               Number of obs     =        832
                                                F(7, 567)         =      14.15
                                                Prob > F          =     0.0000
                                                R-squared         =     0.0987
                                                Root MSE          =     .47528

                                 (Std. err. adjusted for 568 clusters in mid)
```

Y	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
Z	-.0729455	.0404357	-1.80	0.072	-.1523675	.0064765
hs_	.1323256	.0485867	2.72	0.007	.0368938	.2277574
uni_	.0887129	.0538617	1.65	0.100	-.0170798	.1945057
afqt	1.03e-06	9.33e-07	1.10	0.270	-8.03e-07	2.86e-06
lwage_	3.33e-06	1.93e-06	1.73	0.085	-4.55e-07	7.12e-06
lnfaminc_	1.07e-06	6.74e-07	1.59	0.113	-2.53e-07	2.40e-06
cfemale	.1278452	.033785	3.78	0.000	.0614862	.1942041
_cons	.305127	.0435296	7.01	0.000	.2196281	.390626

Z gets smaller coefficient and less significance, can be just because of small sample size

Welfare Preference with D and Z

Using "D" and "Z" from child age -1 to 9, see how it related to whether mother wants to go on welfare if needed (answered in 1979)

```
.         logistic D_ wc_onwelf_79

Logistic regression                                     Number of obs =  2,026
                                                        LR chi2(1)    =   8.50
                                                        Prob > chi2   = 0.0036
Log likelihood = -1236.0621                             Pseudo R2     = 0.0034
```

D_	Odds ratio	Std. err.	z	P>|z|	[95% conf.	interval]
wc_onwelf_79	.7406709	.0769227	-2.89	0.004	.6042594	.9078773
_cons	.4775785	.0281265	-12.55	0.000	.4255143	.536013

Note: _cons estimates baseline odds.

```
.         logistic Z_ wc_onwelf_79

Logistic regression                                     Number of obs =  1,703
                                                        LR chi2(1)    =   0.23
                                                        Prob > chi2   = 0.6289
Log likelihood = -436.44611                             Pseudo R2     = 0.0003
```

Z_	Odds ratio	Std. err.	z	P>|z|	[95% conf.	interval]
wc_onwelf_79	1.100476	.2171009	0.49	0.627	.7475776	1.619962
_cons	.0740038	.0086838	-22.19	0.000	.0587993	.09314

Note: _cons estimates baseline odds.

Control for pecuniary channel?

Not control for wages/net family income at child age 10-18, as will eliminate the pecuniary channel, assume this is ok

lwage: wage when 10-18 (treatment period)

lwage_: wage when -1 to 9 (control period)

```
.         reg lwage Z lwage_ if layfclos == 1, r cluster(mid)

Linear regression                               Number of obs     =        871
                                                F(2, 591)         =      65.89
                                                Prob > F          =     0.0000
                                                R-squared         =     0.6640
                                                Root MSE          =     9640.5

                                  (Std. err. adjusted for 592 clusters in mid)
```

lwage	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
Z	-149.3183	812.0841	-0.18	0.854	-1744.24	1445.604
lwage_	1.151619	.1083839	10.63	0.000	.9387542	1.364483
_cons	5654.998	813.7682	6.95	0.000	4056.769	7253.228

OLS v.s. IV without controls

```
.         reg Y D, r cluster(mid)

Linear regression                               Number of obs     =      2,075
                                                F(1, 1312)        =       1.92
                                                Prob > F          =     0.1658
                                                R-squared         =     0.0012
                                                Root MSE          =     .49656

                              (Std. err. adjusted for 1,313 clusters in mid)
```

Y	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
D	.0348696	.0251482	1.39	0.166	-.0144655	.0842047
_cons	.5441413	.0158713	34.28	0.000	.5130054	.5752771

```
.         ivregress 2sls Y (D = Z), r cluster(mid)

Instrumental variables 2SLS regression            Number of obs   =      2,075
                                                  Wald chi2(1)    =       6.40
                                                  Prob > chi2     =     0.0114
                                                  R-squared       =          .
                                                  Root MSE        =      .5751

                              (Std. err. adjusted for 1,313 clusters in mid)
```

Y	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
D	-.5582941	.2206005	-2.53	0.011	-.9906632	-.1259251
_cons	.7811209	.0890902	8.77	0.000	.6065073	.9557345

```
Instrumented: D
 Instruments: Z
```

Kitchen-sink IV 1

Instrumental variables 2SLS regression

Number of obs	=	1,551
Wald chi2(65)	=	1293.63
Prob > chi2	=	0.0000
R-squared	=	.
Root MSE	=	.51059

(Std. err. adjusted for 1,030 clusters in mid)

Y	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
D	-.4893457	.2564322	-1.91	0.056	-.9919437	.0132522
D_	.0356348	.0452778	0.79	0.431	-.0531079	.1243776
Z_	.0521681	.0600173	0.87	0.385	-.0654637	.1698
cfemale	.1030439	.0276488	3.73	0.000	.0488532	.1572346
chispanic	.0714761	.0467947	1.53	0.127	-.0202399	.163192
cblack	.0543934	.0464901	1.17	0.242	-.0367255	.1455123
lwage_	-3.63e-06	3.07e-06	-1.18	0.237	-9.65e-06	2.38e-06
lnfaminc_	9.73e-07	6.04e-07	1.61	0.107	-2.11e-07	2.16e-06
wkunemp	.0084127	.007102	1.18	0.236	-.0055071	.0223324
hrwrked	.005513	.0023374	2.36	0.018	.0009318	.0100942
numjob	-.0422308	.0343592	-1.23	0.219	-.1095736	.0251121
wkunemp_	-.0028777	.0036468	-0.79	0.430	-.0100253	.00427
hrwrked_cdf_	.0005897	.0003167	1.86	0.063	-.000031	.0012103
hrwrked_cdfsq_	-3.61e-07	2.54e-07	-1.42	0.156	-8.59e-07	1.38e-07
firmsz	7.67e-07	6.96e-06	0.11	0.912	-.0000129	.0000144
numjob_	.0004487	.0435222	0.01	0.992	-.0848532	.0857507

Kitchen-sink IV 2

industry_						
agric	.1946883	.1684314	1.16	0.248	-.1354312	.5248078
const	.1005961	.1247875	0.81	0.420	-.1439829	.3451752
finan	.1103448	.1304852	0.85	0.398	-.1454016	.3660911
missing	.0803939	.1188532	0.68	0.499	-.1525541	.313342
perso	.0885184	.135851	0.65	0.515	-.1777447	.3547815
profe	.0359227	.124032	0.29	0.772	-.2071756	.2790209
publi	.146733	.1673786	0.88	0.381	-.1813231	.4747891
recre	.0998959	.1932579	0.52	0.605	-.2788826	.4786745
retai	.0524533	.120571	0.44	0.664	-.1838615	.2887681
trans	.3904062	.2201318	1.77	0.076	-.0410443	.8218566
hs_	.1915829	.0483595	3.96	0.000	.0968	.2863658
uni_	.0572455	.0443942	1.29	0.197	-.0297656	.1442566
afqt	1.60e-09	8.99e-07	0.00	0.999	-1.76e-06	1.76e-06
married_	.1010687	.0463787	2.18	0.029	.0101681	.1919693
livewmo_	.0902171	.1091119	0.83	0.408	-.1236383	.3040725
famsize_	-.0320578	.017904	-1.79	0.073	-.067149	.0030334
cmom_agbrth	-.0081644	.0089799	-0.91	0.363	-.0257646	.0094358
cbirth_order	-.0701825	.0261611	-2.68	0.007	-.1214573	-.0189077
numchild_	.0393047	.031738	1.24	0.216	-.0229005	.10151
religionr_79	.0145515	.0092064	1.58	0.114	-.0034928	.0325958
region_						
2	-.0432694	.0545702	-0.79	0.428	-.150225	.0636863
3	-.1022311	.0570762	-1.79	0.073	-.2140984	.0096362
4	-.0353054	.053567	-0.66	0.510	-.1402948	.069684
urban_	.0113271	.0610875	0.19	0.853	-.1084021	.1310563
smsa_	.0459327	.0559995	0.82	0.412	-.0638244	.1556897

Kitchen-sink IV 3

cbirth_y						
1972	0	(empty)				
1973	0	(empty)				
1974	0	(empty)				
1975	0	(empty)				
1976	-.3603743	.2546512	-1.42	0.157	-.8594814	.1387328
1977	-.3864126	.240964	-1.60	0.109	-.8586933	.0858681
1978	-.4682505	.2341702	-2.00	0.046	-.9272156	-.0092854
1979	-.4540273	.2074307	-2.19	0.029	-.8605839	-.0474707
1980	-.4826437	.2237538	-2.16	0.031	-.921193	-.0440944
1981	-.4201759	.2059493	-2.04	0.041	-.8238292	-.0165227
1982	-.4949668	.2106725	-2.35	0.019	-.9078774	-.0820562
1983	-.547367	.2119718	-2.58	0.010	-.9628241	-.1319098
1984	-.6487694	.2042371	-3.18	0.001	-1.049067	-.2484721
1985	-.5328462	.186536	-2.86	0.004	-.89845	-.1672424
1986	-.670939	.1985861	-3.38	0.001	-1.060161	-.2817175
1987	-.5389211	.1889141	-2.85	0.004	-.909186	-.1686563
1988	-.6713117	.2021405	-3.32	0.001	-1.0675	-.2751236
1989	-.4927662	.1692618	-2.91	0.004	-.8245132	-.1610192
1990	-.6453117	.1949007	-3.31	0.001	-1.02731	-.2633134
1991	-.5224169	.1807786	-2.89	0.004	-.8767365	-.1680974
1992	-.5080718	.1893704	-2.68	0.007	-.8792309	-.1369127
1993	-.0886561	.1386673	-0.64	0.523	-.360439	.1831269
1994	-.1480053	.1787939	-0.83	0.408	-.4984348	.2024243
1995	-.0943346	.1366725	-0.69	0.490	-.3622077	.1735386
1996	-.3653381	.1906135	-1.92	0.055	-.7389336	.0082574
1997	-.2756392	.1593375	-1.73	0.084	-.5879349	.0366566
1998	-.1997993	.1599481	-1.25	0.212	-.5132918	.1136932
1999	-.0178646	.1238507	-0.14	0.885	-.2606075	.2248783
2000	0	(omitted)				
_cons	.883801	.3757263	2.35	0.019	.1473909	1.620211

Partialing-out lasso IV

```
Estimating lasso for Y using cv
Estimating lasso for D using cv
Estimating lasso for pred(D) using cv

Partialing-out IV linear model     Number of obs                  =      1,551
                                   Number of controls             =         71
                                   Number of instruments          =          1
                                   Number of selected controls    =         61
                                   Number of selected instruments =          1
                                   Wald chi2(1)                   =       3.52
                                   Prob > chi2                    =     0.0607

                              (Std. err. adjusted for 1,030 clusters in mid)
```

Y	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
D	-.4936906	.2631932	-1.88	0.061	-1.00954	.0221586

Endogenous: **D**

Non-parametric IV

Assume heterogeneous treatment effect across compliance pop

Linear IV generally has negative weightings in computation of LATE

Need a fully saturated model or at least rich covariates, so use non-parametric IV

Frolich (2007): need $D, Z \in \{0, 1\}$

Allow me to use Prof Romu Meango's slides: page 51/52 of Chapter 5: The LATE Model for Adv ERM

Frolich (2007)

LATE: Extension with covariates

Frölich (2007) suggests three estimators in the case with covariates, where D and Z are binary. Define:

$\widehat{m}_d(x) = \widehat{E}(Y|X = x, Z = d)$: a nonparametric estimator of the conditional moment,
$\widehat{\mu}_d(x) = \widehat{E}(D|X = x, Z = d)$: a nonparametric estimator of the propensity score.

The ATE of the population of compliers is nonparametrically estimated by

1. A ratio between to *matching estimator*:

$$\widehat{E}(Y_1 - Y_0|T = c) = \frac{\sum_{i:Z_i=1}\left(Y_i - \widehat{m}_0(X_i)\right) - \sum_{i:Z_i=0}\left(Y_i - \widehat{m}_1(X_i)\right)}{\sum_{i:Z_i=1}\left(D_i - \widehat{\mu}_0(X_i)\right) - \sum_{i:Z_i=0}\left(D_i - \widehat{\mu}_1(X_i)\right)}$$

LATE: Extension with covariates

Define:

$\widehat{\pi}(x) = \hat{E}(D|X = x)$: a nonparametric estimator of the propensity score.
$\widehat{m}_{\pi d}(\rho) = \hat{E}(D|\pi(X) = \rho, Z = d)$: an estimator of the conditional moment at a given propensity score.

2. An inverse propensity score weighting estimator:

$$\widehat{E}(Y_1 - Y_0|T = c) = \frac{\sum_i \left(Y_i Z_i / \widehat{\pi}(X_i) - Y_i(1 - Z_i)/(1 - \widehat{\pi}(X_i))\right)}{\sum_i \left(D_i Z_i / \widehat{\pi}(X_i) - D_i(1 - Z_i)/(1 - \widehat{\pi}(X_i))\right)}$$

3. A propensity score matching estimator (R package: causalweight):

$$\widehat{E}(Y_1 - Y_0|T = c) = \frac{\sum_i \left(\hat{m}_{\pi 1}(\hat{\pi}(X_i)) - \hat{m}_{\pi 0}(\hat{\pi}(X_i))\right)}{\sum_{i:Z_i=1} \left(\hat{\mu}_{\pi 1}(\hat{\pi}(X_i)) - \hat{\mu}_{\pi 0}(\hat{\pi}(X_i))\right)}$$

Stata nplate package

```
Bootstrap results                                   Number of obs = 1,551
                                                    Replications  =   100

                              (Replications based on 1,030 clusters in mid)
```

	Observed coefficient	Bootstrap std. err.	z	P>|z|	Normal-based [95% conf.	interval]
late	**-.2757542**	**.7461053**	**-0.37**	**0.712**	**-1.738094**	**1.186585**

Fitting issue with R package (occurredWarning: glm.fit: fitted probabilities numerically 0 or 1), need to investigate into this, maybe too many variables

LATE assumptions

Notations: potential outcomes $Y(z,d)$, $D(z)$ and omit the conditioning on X:

Stronger:

- S1 Unconfoundedness: $Y(1,1), Y(1,0), Y(0,1), Y(0,0), D(1), D(0) \perp\!\!\!\perp Z$
- S2 Monotonicity: $D(1) \geq D(0)$
- S3 Exclusion Restriction: $Y(1,d) = Y(0,d) := Y(d)$
- S4 Existence of Compliers: $P(D=1|Z=1) > P(D=1|Z=0)$

Weaker:

- W1 Unconfounded Type: For all compliance types $t \in \{a,c,n,d\}$,
  $P(T=t) = P(T=t|Z=0) = P(T=t|Z=1)$
- W2 No Defiers: $P(T=d) = 0$
- W3 Mean Exclusion: $E[Y(z,d)|Z=z, T=t] = E[Y(d)|T=t] \; \forall (z,t)$
- W4 Existence of Compliers. $P(T=c) > 0$

S3 $\implies$ W3 but W3 $\not\Longrightarrow$ S3
W4 equivalent to S4 under W1 and W2

Falsification test MW (Mourifie and Wan, 2017)

MW tests for implications (necessary not sufficient for LATE) of S1-S4: *forall* $A \in \mathcal{B}_{\mathcal{Y}}$

$$P(Y \in A, D = 1|Z = 0) = P(Y(z, 1) \in A, D(0) = 1|Z = 0)$$
$$= P(Y(1) \in A, D(0) = 1)$$
$$P(Y \in A, D = 1|Z = 1) = P(Y(z, 1) \in A, D(0) = 1|Z = 1)$$
$$= P(Y(1) \in A, D(1) = 1)$$

$$P(Y(1) \in A, D(0) = 1) \leq P(Y(1) \in A, D(1) = 1)$$

$$\implies P(Y \in A, D = 1|Z = 0) \leq P(Y \in A, D = 1|Z = 1)$$

Similarly

$$\implies P(Y \in A, D = 0|Z = 1) \leq P(Y \in A, D = 0|Z = 0)$$

Intuition: Given treatment status, density of compliers must be nonnegative at any point in the distribution of the outcome variable.

Falsification test MW (Mourifie and Wan, 2017)

MW tests $\forall y \in \{0,1\}$

$$\theta(y,1) \equiv E[P(Z=1)D(1-Z) - P(Z=0)DZ \,|Y=y] \leq 0$$

$$\theta(y,0) \equiv E[P(Z=0)(1-D)Z - P(Z=1)(1-D)(1-Z) \,|Y=y] \leq 0$$

Let $\mathcal{V} = \mathcal{Y} \times \{0,1\}$

$$H_0 : sup_{\{\theta(v)\in\mathcal{V}\}}\theta(v) \leq 0, \;\; H_1 : sup_{\{\theta(v)\in\mathcal{V}\}}\theta(v) > 0$$

But MW assumes

$$\{(D_i, Y_i, Z_i)\}_{i=1}^{n}$$

for consistency and validity of the proposed testing procedure, and I am not sure how clustering on mothers affects the results

Falsification test MW

Full Sample **without** conditioning on covariates: H_0 not rejected at 1% level but rejected at 5%, 10%
To condition on covariates, zoom in sub-samples: e.g. the child is female: not rejected; the child is male: rejected only at 10%, not at 1% and not 5%

So conditioning on child gender seems to be useful

But the sample size is limited, there are many variables to split the sample. We also need to discretize continuous variables.

Falsification test FGK: Farbmacher, Guber and Klaassen (2022)

Assume unconfoundednees S1, (jointly) tests for S2-S4

Intuition:

If condition on observables, even if the test is passed for the FULL sample, LATE assumptions may be violated in some subpopulation defined by observables. If the subpopulation is small, then such violation is diluted in FULL sample and NOT detected by the test

Need a data-driven way of splitting the sample to detect as many violations as possible

Falsification test FGK

In MW, rearranging,

$$P(Y \in A, D = 1 | Z = 0, X = x) - P(Y \in A, D = 1 | Z = 1, X = x) \leq 0$$

$$P(Y \in A, D = 0 | Z = 1, X = x) - P(Y \in A, D = 0 | Z = 0, X = x) \leq 0$$

Similar to saying, conditioning of $X = x$, how does changing Z (treatment) change the inequalities (outcome)?

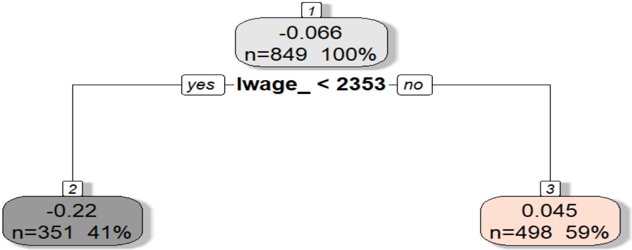
Reformulate as a conditional average treatment effect (CATE) question, and use causal forest to estimate CATE.

Use trees to split the sample in covariate space that delivers the largest heterogeneity between the newly formed subgroup

Finally, if the CATE of that subgroup has positive sign, this means violations of the null hypothesis implied by LATE asssumptions

FGK tree

I guess the sample size is not large enough for causal forest...



What to do next?

maybe try Huber and Mellace (2015) who assume weaker LATE assumptions not full independence as in MW and FGK

estimate compiler population and characteristics

Sensitivity analysis, Baysian IV?

Lewbel's (2012) generated IV based on higher moments, do not need an actual IV

Manski's Monotone IV to bound LATE

Stein-like 2SLS (interesting but maybe not useful)

other suggestions?

Don’t think I have time for this ...

Structural Model (Simplified)

$$\max_{c_t, i_t, s_t} \sum_{t=1}^{T} \beta^{t-1} \left[\log(c_t) - \gamma s_t^2 \right] + \alpha \log\left(e_T\right) \text{ subject to}$$

$$c_t + i_t + a_{t+1} = (1 - p_t) b_t + (1 - \tau_t) \left[p_t w_t + (1 + r_t) a_t \right]$$

$$e_T = \text{sigmoid}\left(\sum_{t=1}^{T} i_t \right)$$

$$s_t = p_t, c_t, i_t \geq 0,\ 0 \leq s_t \leq 1, T = 18$$

α includes discounting factor and weight on child's education; τ_t is marginal tax rate from NBER TAXSIM; p_t is fraction of time employed.

Parameters/estimand: $\beta > 0, \alpha > 0, \gamma > 0$

Inputs: w_t, b_t, r, a_1, a_T; r is retrieved from data on $(1 + r)a_t$

Simulated variables: c_t, i_t, s_t and for $1 < t < T$, a_t

For non-pecuniary channel, add b, s, and p to the production function of education, but the parameters are probably unidentified

Matching moments

Some first moments, average across i, for each t

$\bar{c}_t, \bar{i}_t$

$\bar{s}_t$ with $\bar{p}_t$

$\overline{\text{sigmoid}\left(\sum_{t=1}^{T} i_t\right)}$ with $\bar{e}_T \in \{0, 1\}$ or generally some production function

Maybe consider second moments

Matching regression coefficients (e.g. of search effort on benefit level)

Currently not identified