The Causal Impact of Health on Employment and Earnings: A Partial Identification Approach with Imperfect Instruments

Akbar Ullah
Health Organisation, Policy & Economics
The University of Manchester

August 26, 2025

Overview

Introduction

- A substantial body of empirical literature demonstrates a negative association between health and labour market outcomes (Pintor et al., 2024).

- Current approaches for measuring the impact are:
  1. Matching methods
  2. DID/TWFE
  3. A few studies use instruments

- However, the causal interpretation of this association is unclear:
  1. Omitted variable bias
  2. Reverse causality
  3. Measurement errors

Introduction

- DID/TWFE estimates the ATT and is based on: 1. No anticipation assumption, 2. Parallel trends assumption.
- Event-study DID can rule out anticipation and assess pre-trend balance, but what if past labour outcomes influence current or future health?
- Best practice suggests justifying the parallel trends assumption conceptually, rather than relying solely on statistical tests (Roth et al., 2023; Roth, Jonathan, 2022; Jaeger et al., 2019).

Introduction

- The best instruments for health studies are often genetic markers.
- Genetic markers and childhood events can influence labour outcomes (Germinario et al., 2022; Fletcher, 2008, 2013b; Rosenzweig and Wolpin, 2000).
- Many genetic variants have multiple functions or are correlated with other variants (Von Hinke et al., 2016).

Estimate of interest

- Let $y_i(d_1)$ and $y_i(d_2)$ represent two possible outcomes for individual i associated with health scores d_1 and d_2, where $d_2 > d_1$. We aim to estimate:

$$ATE(d_2, d_1) = E[y(d_2)] - E[y(d_1)]. \tag{1}$$

$$DID{\rightarrow}ATT(d_2, d_1) = E[y(d_2) - y(d_1)|D = 1].$$

- Expressed in the potential outcomes framework:

$$\begin{aligned} E[y(d_2)] &= E[y(d_2)\mathbb{1}\{D = d_2\}] + E[y(d_2)\mathbb{1}\{D{\neq}d_2\}], \\ E[y(d_1)] &= E[y(d_1)\mathbb{1}\{D = d_1\}] + E[y(d_1)\mathbb{1}\{D{\neq}d_1\}]. \end{aligned} \tag{2}$$

- If treatment were externally determined (e.g., randomised), ATE estimation would be straightforward.

Identification

- To address the unobserved counterfactual issue, Manski (and Pepper) proposed bounding (comes at the expense of point estimate):

$$E[y(d_2)\mathbb{1}\{D = d_2\}] + Y_{min}\mathbb{1}\{D \neq d_2\} \leq E[y(d_2)] \leq E[y(d_2)\mathbb{1}\{D = d_2\}] + Y_{max}\mathbb{1}\{D \neq d_2\}.(10) \quad (3)$$

- The ATE can be bounded as follows:

$$LB.E[y(d_2)] - UB.E[y(d_1)] \leq ATE(d_2, d_1) \leq UB.E[y(d_2)] - LB.E[y(d_1)]. \quad (4)$$

Identification

Assumption 1. Same Direction of Correlation (SDC)

$$\text{Cov}(y(d), D)\text{Cov}(y(d), Z) \geq 0 \quad \text{for all } d \in D. \tag{5}$$

Assumption 1.1 Monotone Treatment Selection (MTS)

$$d_2 > d_1 \Rightarrow E[y(d) \mid D = d_2] \geq E[y(d) \mid D = d_1] \quad \text{for all } d \in D. \tag{6}$$

Assumption 1.2 Monotone Instrumental Variable (MIV)

$$z_2 > z_1 \Rightarrow E[y(d) \mid Z = z_2] \geq E[y(d) \mid Z = z_1] \quad \text{for all } d \in D,\ z \in Z. \tag{7}$$

- SDC: Individuals' potential outcomes should be non-decreasing in their own health and parents' qualifications, or vice versa.
- MTS: Individuals with better health have at least the same average latent outcomes as those with poorer health.
- MIV: Individuals' average potential outcomes should be non-decreasing in their parents' qualifications.

Identification

Assumption 2. Monotone treatment response (MTR)

$$y(d) \geq y(d') \quad \text{for all} \quad d > d'. \tag{8}$$

- The MTR assumption states that the potential outcome weakly increases with health scores.
- Individuals with better health have at least the same latent outcomes as those with poorer health.
- Combined with the MTS assumption, it implies that observed average labour outcomes should not decrease with better health.

Identification

- Using SDC, MTR (and LEI), the No assumption bounds in 3 can be replaced with:

$$\hat{\mu}_d^+ = 1 + \alpha\,[\epsilon(1-\beta)[\theta_Z(D-E(D)) - \theta_D(Z-E(Z))] + \beta(D-E(D)) + (1-\beta)(1-\epsilon)(Z-E(Z))]$$

$$\hat{\mu}_d^- = 1 - \alpha\,[\epsilon(1-\beta)[\theta_Z(D-E(D)) - \theta_D(Z-E(Z))] + \beta(D-E(D)) + (1-\beta)(1-\epsilon)(Z-E(Z))]$$

$$Lower^1_{MTR}(d) = (D=d)\hat{\mu}_d^- Y + (D>d)\min\left\{\hat{\mu}^- Y_{min}, \hat{\mu}^- Y\right\} + (D<d)\min\left\{\hat{\mu}^- Y_{max}, \hat{\mu}^- Y\right\}$$

$$Lower^2_{MTR}(d) = (D=d)\hat{\mu}_d^+ Y + (D>d)\min\left\{\hat{\mu}^+ Y_{min}, \hat{\mu}^+ Y\right\} + (D<d)\min\left\{\hat{\mu}^+ Y_{max}, \hat{\mu}^+ Y\right\}$$

$$Upper^1_{MTR}(d) = (D=d)\hat{\mu}_d^+ Y + (D>d)\max\left\{\hat{\mu}^+ Y_{min}, \hat{\mu}^+ Y\right\} + (D<d)\max\left\{\hat{\mu}^+ Y_{max}, \hat{\mu}^+ Y\right\}$$

$$Upper^2_{MTR}(d) = (D=d)\hat{\mu}_d^- Y + (D>d)\max\left\{\hat{\mu}^- Y_{min}, \hat{\mu}^- Y\right\} + (D<d)\max\left\{\hat{\mu}^- Y_{max}, \hat{\mu}^- Y\right\}$$

Data

- Use Waves 2–5 of the United Kingdom Household Longitudinal Study (UKHLS).
- A nationally representative panel launched in 2009, including 100,000 individuals across the UK.
- In Waves 2–3, after the standard annual survey, respondents were invited to a nurse-led health assessment.
- Blood samples were collected and analysed to produce a set of biomarkers.

Health Measures

$\sum_p NCD_p + Disab.$	$\sum_q Test_q$	$SHS(BHS)$
0	0	Excellent Self-reported(nurse-measured) health (4)
1	1	Good Self-reported(nurse-measured) health (3)
2	2	fair Self-reported(nurse-measured) health (2)
3	3	Poor Self-reported(nurse-measured) health (1)
>3	>3	Very poor Self-reported(nurse-measured) health (0)

Notes: NCD, Noncommunicable disease indicators; Disab. Disability indicator; SHS, Self-reported health score; BHS, biomarkers-based health score.

$NCD_p = 1$ if NCD_p is reported for $p \in (1, ..., 17)$, zero otherwise,

$Disab. = 1$ if disability reported, zero otherwise,

$Test_q = 1$ if $test_q$ is not normal for $q \in (1, ..., 14)$, zero otherwise.

Outcomes and Instrument

- Employment (=1 if paid (self-)employed, =0 if unemployed or on sick leave) and monthly gross earnings (in thousands).
- Mother's qualifications when the respondent was aged 14, used as an imperfect instrument.
- We estimate the effect of SHS (BHS) from Wave 3 (Waves 2 and 3) on employment and gross monthly labour earnings in Wave 4 (Waves 4 and 5).
- For SHS, data is available for 19,874 individuals for the employment analysis and 15,539 for the earnings analysis.
- For BHS, complete data is available for 5,422 observations for the employment analysis and 4,275 for the earnings analysis.

MTR+MTS Validation

	Very poor health	Poor health	Fair health	Good health	Excellent health
Full sample					
Wave 4	$SHS = 0$	$SHS = 1$	$SHS = 2$	$SHS = 3$	$SHS = 4$
Monthly wage	1296.78	1597.10	1844.89	1876.32	2005.19
Employment ratio	0.40	0.53	0.64	0.78	0.93
Observations	83	215	1141	6239	15795
Wave 3					
Monthly wage	1268.29	1517.43	1697.13	1853.37	1967.84
Employment ratio	0.34	0.48	0.65	0.76	0.93
Observations	103	250	1284	6286	16187
Waves 2-8					
Monthly wage	1387.54	1694.51	1838.25	1957.71	2096.61
Employment ratio	0.33	0.51	0.66	0.77	0.94
Observations	402	1083	6536	38683	102694
Blood sample					
Wave 4	$BHS = 0$	$BHS = 1$	$BHS = 2$	$BHS = 3$	$BHS = 4$
Monthly wage	1554.27	1582.63	1681.08	1725.60	1729.55
Employment ratio	0.81	0.88	0.93	0.94	0.96
Observations	599	627	894	808	306
Waves 4 & 5					
Monthly wage	1538.03	1587.63	1697.52	1732.30	1748.50
Employment ratio	0.81	0.88	0.92	0.94	0.95
Observations	1149	1239	1729	1577	611

Notes: SHS, Self-reported health scores; BHS, Blood test based health scores.

SDC Validation

Full sample	$SHS = 0$	$SHS = 1$	$SHS = 2$	$SHS = 3$	$SHS = 4$
Age	52.33	50.24	48.78	47.04	40.28
Male	0.46	0.39	0.39	0.42	0.44
Qualification: Degree/Other higher degree	0.17	0.22	0.28	0.34	0.41
Qualification: A-level/GCSE etc.	0.58	0.56	0.53	0.53	0.53
Qualification: No qualification	0.25	0.22	0.19	0.13	0.07
Mother didn't go to school	0.05	0.04	0.04	0.03	0.03
Mother left school without qualifications	0.69	0.55	0.55	0.46	0.35
Mother left school with qualifications	0.18	0.25	0.24	0.29	0.34
Mother gained post school qualifications	0.07	0.12	0.14	0.16	0.19
Mother gained a uni. degree or higher	0.01	0.04	0.03	0.06	0.09
Observations	142	336	1766	9294	21823
Blood sample	$BHS = 0$	$BHS = 1$	$BHS = 2$	$BHS = 3$	$BHS = 4$
Age	50.43	47.34	45.98	42.03	38.53
Male	0.01	0.01	0.00	0.00	0.00
Qualification: Degree/Other higher degree	0.29	0.34	0.42	0.47	0.42
Qualification: A-level/GCSE etc.	0.52	0.55	0.49	0.49	0.55
Qualification: No qualification	0.18	0.11	0.08	0.04	0.03
Mother didn't go to school	0.01	0.01	0.01	0.01	0.01
Mother left school without qualifications	0.57	0.48	0.45	0.32	0.28
Mother left school with qualifications	0.23	0.29	0.31	0.36	0.38
Mother gained post school qualifications	0.16	0.16	0.17	0.21	0.23
Mother gained a uni. degree or higher	0.03	0.06	0.06	0.10	0.10
Observations	1820	1831	2473	2026	730

Notes: SHS, Self-reported health scores; BHS, Blood test based health scores.

Results for Employment

	OLS	*SDC + LEI*		*SDC + LEI + MTR*		*SDC + MTR*	
Full sample		LB	UB	LB	UB	LB	UB
Very poor health (*SHS* = 0)		0.02	0.98	0.02	0.87	0.02	0.88
Poor health (*SHS* = 1)		0.03	0.98	0.87	0.87	0.88	0.88
Fair health (*SHS* = 2)		0.06	0.97	0.87	0.88	0.88	0.88
Good health (*SHS* = 3)		0.24	0.93	0.88	0.90	0.88	0.90
Excellent health (*SHS* = 4)		0.66	0.94	0.90	0.95	0.90	0.95
ATE(1, 4)	-0.39 (0.04)	-0.92	0.32	-0.08	-0.02	-0.08	-0.02
ATE(2, 4)	-0.28 (0.02)	-0.88	0.31	-0.08	-0.02	-0.07	-0.02
ATE(3, 4)	-0.15 (0.01)	-0.70	0.27	-0.07	0.00	-0.07	0.00
Blood sample	OLS	LB	UB	LB	UB	LB	UB
Very poor health (*BHS* = 0)		0.22	0.99	0.22	0.93	0.22	0.93
Poor health (*BHS* = 1)		0.22	0.98	0.93	0.95	0.93	0.95
Fair health (*BHS* = 2)		0.27	0.99	0.95	0.97	0.95	0.97
Good health (*BHS* = 3)		0.28	0.99	0.97	0.99	0.97	0.99
Excellent health (*BHS* = 4)		0.13	1.00	0.99	1.01	0.99	1.01
ATE(1, 4)	-0.07 (0.01)	-0.78	0.85	-0.08	-0.04	-0.08	-0.04
ATE(2, 4)	-0.02 (0.01)	-0.73	0.85	-0.05	-0.02	-0.06	-0.02
ATE(3, 4)	-0.01 (0.01)	-0.72	0.85	-0.04	0.00	-0.04	0.00
ATE(1, 3)	-0.06 (0.01)	-0.77	0.71	-0.06	-0.02	-0.06	-0.02
ATE(2, 3)	-0.01 (0.01)	-0.72	0.71	-0.04	0.00	-0.04	0.00

Notes: OLS, Ordinary least square, SDC, Same direction of correlation; LEI, Less endogenous instrument; MTR, Monotone treatment response; LB, Lower bound (95%); UB, Upper bound (95%); ATE, Average treatment effects.

Results for Earnings

	OLS	*SDC + LEI*		*SDC + LEI + MTR*		*SDC + MTR*	
Full sample		LB	UB	LB	UB	LB	UB
Very poor health (*SHS* = 0)		0.35	2.96	0.35	2.03	0.36	2.04
Poor health (*SHS* = 1)		0.21	4.00	2.03	2.03	2.04	2.04
Fair health (*SHS* = 2)		0.21	7.60	2.03	2.07	2.04	2.08
Good health (*SHS* = 3)		0.62	6.56	2.07	2.27	2.08	2.21
Excellent health (*SHS* = 4)		1.48	3.56	2.27	3.55	2.21	3.48
ATE(1, 4)	-0.40 (0.11)	-3.34	2.52	-1.52	-0.24	-1.44	-0.17
ATE(2, 4)	-0.15 (0.06)	-3.35	6.12	-1.52	-0.20	-1.44	-0.13
ATE(3, 4)	-0.10 (0.03)	-2.94	5.08	-1.49	0.00	-1.40	0.00
Blood sample	OLS	LB	UB	LB	UB	LB	UB
Very poor health (*BHS* = 0)		0.53	5.44	0.56	1.79	0.55	1.82
Poor health (*BHS* = 1)		0.59	6.05	1.79	2.77	1.82	2.70
Fair health (*BHS* = 2)		0.84	5.61	2.77	3.61	2.70	3.39
Good health (*BHS* = 3)		0.76	5.65	3.61	5.10	3.39	4.96
Excellent health (*BHS* = 4)		0.49	5.08	5.10	5.08	4.96	5.15
ATE(1, 4)	-0.12 (0.06)	-4.49	5.55	-3.29	-2.34	-3.33	-2.27
ATE(2, 4)	-0.02 (0.06)	-4.24	5.11	-2.32	-1.50	-2.45	-1.58
ATE(3, 4)	0.01 (0.06)	-4.32	5.16	-1.48	0.00	-1.76	0.00
ATE(1, 3)	-0.14 (0.05)	-5.06	5.29	-3.31	-0.84	-3.14	-0.69
ATE(2, 3)	-0.04 (0.05)	-4.82	4.85	-2.34	0.00	-2.27	0.00

Notes: OLS, Ordinary least square, SDC, Same direction of correlation; LEI, Less endogenous instrument; MTR, Monotone treatment response; LB, Lower bound (95%); UB, Upper bound (95%); ATE, Average treatment effects.

Return to Qualifications

	SDC + LEI		*SDC + LEI + MTR*		*SDC + MTR*	
Employment	LB	UB	LB	UB	LB	UB
No qualification (0)	0.08	0.97	0.08	0.89	0.09	0.89
GCSE (1)	0.27	0.96	0.89	0.92	0.89	0.92
A-level (2)	0.22	0.96	0.92	0.95	0.92	0.95
Degree (3)	0.46	0.97	0.95	0.97	0.95	0.97
$ATE(0,3)$	-0.88	0.51	-0.89	-0.06	-0.88	-0.06
$ATE(1,3)$	-0.70	0.50	-0.08	-0.03	-0.08	-0.03
$ATE(2,3)$	-0.75	0.51	-0.05	0.00	-0.05	0.00
$ATE(0,2)$	-0.88	0.75	-0.86	-0.03	-0.86	-0.03
$ATE(1,2)$	-0.70	0.74	-0.06	0.00	-0.06	0.00
$ATE(0,1)$	-0.87	0.70	-0.84	0.00	-0.83	0.00
Earning	LB	UB	LB	UB	LB	UB
No qualification (0)	0.20	4.85	0.20	2.05	0.20	2.05
GCSE (1)	0.59	6.37	2.05	2.48	2.05	2.47
A-level (2)	0.55	6.51	2.48	3.87	2.47	3.76
Degree (3)	1.28	5.46	3.87	5.46	3.76	5.35
$ATE(0,3)$	-5.26	3.58	-5.26	-1.82	-5.15	-1.71
$ATE(1,3)$	-4.87	5.09	-3.41	-1.39	-3.31	-1.29
$ATE(2,3)$	-4.91	5.24	-2.98	0.00	-2.89	0.00
$ATE(0,2)$	-6.31	4.30	-3.67	-0.43	-3.56	-0.42
$ATE(1,2)$	-5.93	5.82	-1.82	0.00	-1.71	0.00

Notes: OLS, Ordinary least square; SDC, Same direction of correlation; LEI, Less endogenous instrument; MTR, Monotone treatment response; LB, Lower bound (95%); UB, Upper bound (95%); ATE, Average treatment effects.

Impact of Mental Health

	OLS	*SDC + LEI*		*SDC + LEI + MTR*		*SDC + MTR*	
Employment		LB	UB	LB	UB	LB	UB
Severe depression	-0.29 (0.02)	-0.85	0.17	-0.85	-0.03	-0.85	-0.02
Moderate depression	-0.16 (0.01)	-0.82	0.17	-0.04	-0.02	-0.04	-0.01
Mild depression	-0.11 (0.01)	-0.80	0.17	-0.03	0.00	-0.03	0.00
Earnings		LB	UB	LB	UB	LB	UB
Severe depression	-0.20 (0.05)	-2.66	5.05	-2.70	-0.43	-2.57	-0.31
Moderate depression	-0.21 (0.04)	-2.64	5.88	-0.91	-0.19	-0.76	-0.16
Mild depression	-0.12 (0.04)	-2.63	5.25	-0.66	0.00	-0.61	0.00

Notes: SDC, Same direction of correlation; LEI, Less endogenous instrument; MTR, Monotone treatment response; LB, Lower bound (95%); UB, Upper bound (95%); ATE, Average treatment effects.

Conclusion

- This study addresses measurement error, omitted variable bias, and reverse causality in evaluating the impact of health on employment and earnings.
- Our estimates of health impact bounds are somewhat comparable in magnitude to those for differences in employment and earnings between individuals with no qualifications and A-levels, or GCSEs and degrees.
- For example, employment probability differences between degree and GCSE qualifications ranges from 3% to 8%, with earnings gap of £1,390 to £3,410.
- The employment effect of having no versus two health conditions is 2%-8%.
- Provides strong evidence of the significant link between good health and individual economic success.

Thank You!