CA1a

Ties van der Veen

31-8-2019

III. Potential Outcomes

(a) Y(0,i) = health status without health insurance Y(1,i) = health status with health insurance

(b)
1. Better aftercare in case of an accident
2. Lower threshold to get something checked out, preventing possible injuries from developing.

IV. Causal diagram

(a) Over-estimate, since income positively affects both variables. Not adjusting for this effect will thus attribute some of that effect to the causal relationship.
(b) Yes, this would also have the same result of overestimation.
(c) Depends on whether it is assumed that this risky behavior is randomly distributed over the population. If it is, then there is no issue.
(d) Same thing as c, depends on the randomness of this variable.

V. Descriptive Statistics

Table 1

hi	mean_hlth
0	3.70
1	4.01

Table 2 (probably wrong, got error ‘’object hi not found”)

hi	mean_inc
0	44234.
1	98921.

VI. Regression analysis

Naive regression output

Call: lm(formula = hlth ~ hi, data = nhis, subset = fml == 0, weights = perweight)

Weighted Residuals:

Min	1Q	Median	3Q	Max
-347.42	-41.47	-0.47	49.69	175.89

Coefficients:

	Estimate	Std. Error	t value	Pr(>t)
(Intercept)	3.69565	0.02628	140.65	<2e-16 ***
hi	0.31325	0.02827	11.08	<2e-16 ***

Signif. codes: 0 ‘’ ***0.001*** ‘’ *0.01* ‘’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 56.68 on 9393 degrees of freedom Multiple R-squared: 0.0129, Adjusted R-squared: 0.01279 F-statistic: 122.7 on 1 and 9393 DF, p-value: < 2.2e-16

The estimation results suggest a somewhat higher health status for those with health insurance compared to those without health insurance.

Regression output with income level

Call: lm(formula = hlth ~ hi + incmp, data = nhis, subset = fml == 0, weights = perweight)

Weighted Residuals: Min 1Q Median 3Q Max -304.749 -36.118 2.906 44.129 194.547

Coefficients: Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.495e+00 2.734e-02 127.831 <2e-16 ***hi 6.502e-02 2.997e-02 2.169 0.0301***
incmp 4.539e-06 2.131e-07 21.303 <2e-16 * —

Signif. codes: 0 ‘’ ***0.001*** ‘’ *0.01* ‘’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 55.36 on 9392 degrees of freedom Multiple R-squared: 0.0584, Adjusted R-squared: 0.0582 F-statistic: 291.2 on 2 and 9392 DF, p-value: < 2.2e-16

The estimation results suggest no higher health status for those with health insurance compared to those without health insurance.