Quantitative Economics Problem Set 2

1 No.

The recommendation is premised on the proposition that consuming chocolate up to five times a week has a negative causal effect on weight. This proposition is not supported by the result of the survey. The survey found that very frequent chocolate consumption compared to no chocolate consumption was, on average, associated with a 2.5kg increase in weight. This supports the conclusion that such an increase in chocolate consumption has a negative ~~effect on~~ causal effect on weight only if the other causal determinants of weight are uncorrelated with chocolate consumption. This condition does not necessarily hold. In fact, the survey ~~suggests~~ results suggest that this does not hold. The survey found that chocolate consumption is positively related to calorie intake, and calorie intake is a causal determinant of obesity. The ~~observation that~~ association between ~~high~~ chocolate assumption and obesity could be explained by the ~~following~~ following. Individuals who consume no chocolate disproportionately consist of diabetic persons, who are disproportionately obese.

2 Other determinants of nchildren include the ~~mother's~~ ~~date of birth (hence~~ country of origin, woman's country of residence, ~~state/county/city~~ area (urban/rural) of residence, year of birth, wealth, income, and religious belief.

Country of residence, ~~state/county/city~~ area of residence wealth, and income are likely to be correlated with educ. In general, educ varies significantly between countries, ~~states,~~ and between urban and rural areas. The total years ~~of~~ in education of a woman in a more developed country ~~of~~ is on average higher than that of a woman in a less developed country. Likewise for ~~[illegible]~~ urban and rural areas. A woman with higher total years in education is also likely to have higher wealth and income.

b No.

~~This~~ The unobserved causal determinants of ~~A child~~ nchildren collected in u are correlated with educ. Then, the causal model nchildren = $\beta_0 + \beta_1 educ + u$ does not coincide with the population linear regression model nchildren = $\rho_0 + \rho_1 educ + e$, where $Ee = 0$ and $E\, educ\, e = 0$ (hence $cov(educ, e) = 0$). The sample linear regression of nchildren and educ yields ~~paramt~~ parameters $\hat{\beta}_0$ and $\hat{\beta}_1$ that are consistent for ρ_0 and ρ_1 but not for β_0 and β_1 because ρ_0 and ρ_1 do not coincide with β_0 and β_1.

3a ~~Predicted birthweight when cigs = 0 is 3395. When cigs = 20, this is 3395 − 15(20) = 3095, the difference is 300g. On average, an increase in the number of cigarettes smoked per day during pregnancy is associated from 0 to 20 is associated with a 300g decrease in infant birthweight.~~

$\hat{E}(bwght \mid cigs = 0) = 3395$

$\hat{E}(bwght \mid cigs = 20) = 3395 - 15(20) = 3095$

Difference = ~~30~~ 300

On average, women who ~~smoked~~ consumed 20 cigarettes a day during pregnancy had infants with birthweight 300g less than those of women who consumed 0 cigarettes a day during pregnancy.

b No.

The regression in (1) yields a reliable estimate of the causal effect of smoking on birthweight only if the causal effect is linear and cigs is uncorrelated with the unobserved determinants of bwght collected in u.

There is no a priori reason to think that the causal effect ~~on~~ of cigs on bwght is linear.

Unobserved determinants of bwght collected in u include eating habits, drinking habits, access to healthcare. ~~and~~ These determinants are unlikely to be uncorrelated with cigs. For example, plausibly, a person who indulges in the vice of smoking is also more likely to indulge in the vice of drinking. Then orthogonality fails in the causal model, the population regression model does not coincide with the causal model, ~~the~~ and the sample regression coefficients are consistent for the coefficients of the population regression model but not those of the causal model.

c $\hat{E}(bwght \mid cigs = c) = 3500 \Rightarrow c = -7$

The linear regression model never predicts a birthweight greater than 3395. This is the average birthweight for a mother who ~~smoked~~ consumed zero cigarettes a day during pregnancy and on average, ~~the~~ a higher number of cigarettes ~~smoked~~ consumed a day during pregnancy is associated with a lower infant birthweight. Then, given only the number of cigarettes consumed a day during pregnancy, the optimal prediction is no greater than 3395.

4a The OLS regression problem is
$\min_{b_0, b_1} \sum_{i=1}^n (Y_i - b_0 - b_1 X_i)^2$.

FOC b_0: $\partial/\partial b_0 \sum_{i=1}^n (Y_i - b_0 - b_1 X_i)^2 = \sum_{i=1}^n \partial/\partial b_0 (Y_i - b_0 - b_1 X_i)^2$
$= \sum_{i=1}^n 2(Y_i - b_0 - b_1 X_i)(-1)$
$= 0$

FOC b_1: $\partial/\partial b_1 \sum_{i=1}^n (Y_i - b_0 - b_1 X_i)^2 = \sum_{i=1}^n \partial/\partial b_1 (Y_i - b_0 - b_1 X_i)^2$
$= \sum_{i=1}^n 2(Y_i - b_0 - b_1 X_i)(X_i)$
$= 0$

FOC $b_0 \Rightarrow \hat{E}(Y - b_0 - b_1 X) = 0$
$\Rightarrow \hat{E}(Y) - b_0 - b_1 \hat{E}(X) = 0$
$\Rightarrow b_0 = \hat{E}Y - b_1 \hat{E}X$

FOC $b_1 \Rightarrow \hat{E}[X(Y - b_0 - b_1 X)] = 0$
$\Rightarrow \hat{E}[X(Y - (\hat{E}Y - b_1 \hat{E}X) - b_1 X)] = 0$
$\Rightarrow \hat{E}[X(Y - \hat{E}Y - b_1(X - \hat{E}X))] = 0$
~~$\Rightarrow \hat{E}[X(Y - \hat{E}Y)] - b_1 \hat{E}[X - \hat{E}X$~~
$\Rightarrow \hat{E}[X(Y - \hat{E}Y)] - b_1 \hat{E}[X(X - \hat{E}X)] = 0$
$\Rightarrow \hat{E}[XY - X\hat{E}Y] - b_1 \hat{E}[X^2 - X\hat{E}X] = 0$
$\Rightarrow (\hat{E}XY - \hat{E}X\hat{E}Y) - b_1(\hat{E}X^2 - \hat{E}X\hat{E}X) = 0$
$\Rightarrow b_1 = \widehat{cov}(X, Y) / \widehat{var}(X)$

Supposing that the FOC are sufficient for a minimum,
$\hat{\beta}_0, \hat{\beta}_1 = \operatorname{argmin}_{b_0, b_1} \sum_{i=1}^n (Y_i - b_0 - b_1 X_i)^2$

b $\hat{u}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i$

From the above, FOC b_0, FOC b_1 are satisfied at $\hat{\beta}_0, \hat{\beta}_1$, then evaluate at $\hat{\beta}_0, \hat{\beta}_1$ and substitute $\hat{u}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i$

$\sum_{i=1}^n (-2)\hat{u}_i = 0$, $\sum_{i=1}^n (-2) X_i \hat{u}_i = 0$
$\Rightarrow 1/n \sum_{i=1}^n \hat{u}_i = 0$, $1/n \sum_{i=1}^n X_i \hat{u}_i = 0$

c Note that $\widehat{var}(X) \geq 0$

For $\widehat{var}(X) = 0$, $\hat{\beta}_1$ is undefined.
For $\widehat{var}(X) \neq 0$, $\hat{\beta}_1$ is directly proportionate to the sample ~~correlation between~~ ~~cor~~ covariance between X and Y. Hence, $\hat{\beta}_1$ is ~~less than directly~~ ~~more~~ more than proportionately varies in the same direction as the sample correlation.

5 $cov(Y_i, X_i) = cov(\beta_0 + \beta_1 X_i + u_i, X_i)$
$= \beta_1 cov(X_i, X_i) + cov(u_i, X_i)$
$= \cancel{cov(X_i, X_i)}\ \beta_1 var(X_i)$

$\beta_1 = cov(Y_i, X_i) / var(X_i)$

This will be recognised as the solution to the population linear regression problem (of Y_i on X_i). Then such population linear regression recovers β_1.

6 $\beta_1 = cov(Y_i, X_i) / var(X_i)$
Denote the parameters of the population regression of X_i on Y_i as β_0', β_1'.
$\beta_1' = cov(X_i, Y_i) / var(Y_i)$
$1/\beta_1 = var(X_i) / cov(Y_i, X_i)$
$\beta_1' = 1/\beta_1$ iff ~~$var(X_i) = var(Y_i)$~~
$\Leftrightarrow cov(X_i, Y_i) / var(Y_i) = var(X_i) / cov(Y_i, X_i)$
$\Leftrightarrow [cov(X_i, Y_i)]^2 = var(X_i) var(Y_i)$

$X_i = -\beta_0/\beta_1 + 1/\beta_1 Y_i - 1/\beta_1 u_i$
coincides with the population linear regression of X_i on Y_i iff $E(u_i) = E(Y_i u_i) = 0$. Given that $Y_i = \beta_0 + \beta_1 X_i + u_i$ is the population linear regression of Y_i on X_i, we have $E(X_i u_i) = 0$. Then $cov(X_i, u_i) = cov(Y_i, u_i) = 0$. $cov(Y_i, u_i) = cov(\beta_0 + \beta_1 X_i + u_i, u_i) = 0$
$\Rightarrow var(u_i) = 0 \Rightarrow u_i$ is a constant, 0.

$1/\beta_1 = \beta_1'$ iff $u_i = 0$, which is iff Y_i is simply a linear function of X_i.

7 $var(Y) =$ ~~$var(E(Y|X) + Y - E(Y|X))$~~
$var(E(Y|X) + \varepsilon)$
$= var(E(Y|X)) + var\,\varepsilon + 2cov(E(Y|X), \varepsilon)$
~~$= var(E(Y|X)) + var\,\varepsilon + 2cov(Y - E(Y|X), E(Y|X))$~~
~~$= var(E(Y|X)) + 2cov(Y, E(Y|X)) - 2var(E(Y|X))$~~
~~$+ var\,\varepsilon$~~
~~$= -var(E(Y|X))$~~

$cov(E(Y|X), \varepsilon) = E(E(Y|X) - E(E(Y|X)))(\varepsilon - E(\varepsilon))$
$= E(E(Y|X) - E(Y))(\varepsilon - E(\varepsilon))$
$= [E(E(Y|X)) - E(E(Y))](\varepsilon - E(\varepsilon))$
$= [EY - EY](\varepsilon - E\varepsilon)$
$= 0$

$var(Y) = var(E(Y|X)) + var\,\varepsilon$

$E\varepsilon = E(Y - E(Y|X)) = EY - EE(Y|X) = EY - EY = 0$
$var\,\varepsilon = E(\varepsilon - E\varepsilon)^2 = E\varepsilon^2$

$var(Y) = var(E(Y|X)) + E\varepsilon^2$

$E(var[Y|X]) = E[E(Y - E(Y|X))^2 | X]$
$= E(Y - E(Y|X))^2$
$= E\varepsilon^2$

$var(Y) = var(E(Y|X)) + E(var(Y|X))$