HW11: Explanation Practice

The Principal Components - Ed Brown, Daphne Lin, Linh Tran, Lisa Wu

Part 1 - Explanation Practice

Imagine you are working at a company that is considering whether to purchase a cybersecurity training for its employees. A colleague has gathered data about a set of Fortune 500 companies. They present the following regression estimate:

Outcome variable: Security Breaches	
Cybersecurity Training Hours	0.052** (0.009)
Emails Encrypted	-1.23*** (0.01)
Constant	10.790** (5.078)
Note: *p<0.5; **p<0.01; ***p<0.001	

Here, Emails Encrypted is a covariate that measures the fraction of employees that reported using encrypted email either "all of the time" or "most of the time" in a questionnaire. Based on this regression, your colleague suggests that cybersecurity training leads to more security breaches, so your company should not invest in it.

You are convinced that the statistical assumptions underlying the estimate are sufficiently justified. However, you have concerns with your colleague's causal interpretation. Provide a response as follows:

1. An omitted variable is whether the company has high value assets that might be attractive targets for criminals. Argue whether the omitted variable bias is towards zero or away from zero (5 sentences max).

- We expect that companies with high value assets (omitted variable) would be more attractive targets for criminals and may experience more security breaches (outcome variable), which means that the omitted variable has a positive effect on the outcome variable. We also expect high value assets to have a positive effect on cybersecurity training hours (measured variable), as companies with more valuable assets would seek to protect them through more cybersecurity training. Since the omitted variable has a positive effect on both the measured variable and the outcome variable, we believe the omitted variable bias is away from zero, which is worse than those bias towards zero. We may not be able to trust the coefficient of cybersecurity training hours and the hypothesis test results.

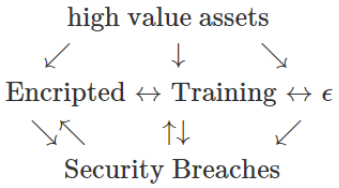
2. Explain why there is a possibility of reverse causality. Argue whether the direction of bias is towards zero or away from zero (5 sentences max).

- It is possible that reverse causality exists, as security breaches may lead to increased cybersecurity training hours. The more security breaches occur, the more a company may require their employees to take cybersecurity training. In this case, there is a positive effect of security breaches on cybersecurity training hours, and the bias is away from zero. Because reverse causality is a key violation for this one-equation structural model, we may not be able to trust the cybersecurity coefficient and the hypothesis test results.

3. Explain why there is an outcome variable on the right hand side. Argue whether removing it would make the coefficient on Cybersecurity Training Hours move up or down (5 sentences max).

- Emails encrypted is an outcome variable on the right hand side, as the fraction of employees that reported using encrypted email either all or most of the time may increase as security breaches increase and the company requires employees to encrypt all or key emails. Email encryption could also increase, as a result of increased cybersecurity training. Removing the Email Encrypted variable could make the Cybersecurity Training Hours coefficient move down, and may even move to the negative sign (and we expect more cybersecurity training would reduce security breaches), invalidating our company's conclusion of 'not investing in cybersecurity".

4. Provide a one-sentence conclusion addressing the idea that your company should not invest in cybersecurity training.



- As discussed in the first three questions and shown in the above graph, there are three violations (omitted variable, reverse causality and an outcome variable on the right hand side) for this one-equation structure model, so we can't trust the coefficient of Cybersecurity Training Hours and the hypothesis testing conclusion that the company should not invest in cybersecurity training.