

PS1 - Controls and Bias

Part A

If we regress $earnings = \beta_0 + \beta_1 college$ we find that $\beta_1 \approx 13.33$. By comparison, under $earnings = \alpha_0 + \alpha_1 college + \alpha_2 ability$ we find $\alpha_1 \approx 1.97$. This reveals an upwards bias of $\beta_1 - \alpha_1 \approx 11.36$. These numbers are fairly consistent with the way earnings was defined.

Part B

1. Under $earnings = \beta_0 + \beta_1 college + \beta_2 parental_occupation$ we find $\beta_1 \approx 11.51$. This means we have an upwards bias of $11.51 - 1.97 = 9.54$, where 1.97 is the estimate achieved when regressing earnings on college and true ability.
2. Pushing the final r-squared up by 30% finds a positive bias $\approx 8.71 - 1.97 = 6.74$.
3. Pushing the final r-squared to 1 finds a negative bias $\approx 1.64 - 1.97 = -0.33$.

Part C

1. With a tighter relationship between parental and child ability, we find $\beta_1 = 10.79$ gives us a positive bias, $\approx 10.79 - 1.97 = 8.82$.
2. Pushing the r-squared up 30% leaves a positive bias $\approx 7.68 - 1.97 = 5.71$.
3. Pushing the r-squared to 1 finds a negative bias $\approx 1.19 - 1.97 = -0.78$.
4. Under the standard regression and Oster's 30% rule of thumb, assuming a tighter relationship between child and parent ability produces less biased estimates, which makes sense. In both cases, pushing the r-squared to 1 flips the sign of the bias, consistent with the conclusion that this is may be an overly conservative procedure.

Part D

1.
 1. For $white_collar = \beta_0 + \beta_1 college$ we find $\beta_1 = 0.129$ with an r-squared of 0.0169. The low coefficient and small r-squared indicates that college degree receipt is not a significant predictor of having a white-collar occupation. This makes sense considering we only took college (and thereby ability) in account for approximately half of the population, "low-ability" types, and left the remainder as a random function of $N(0, 1) > 0$ (no dependence on college).
 2. We find $\beta_{white-collar} \approx 11.29$ with positive bias $\approx 11.29 - 1.97 = 9.32$. For non-white-collar workers, we find $\beta_{non-white-collar} \approx 16.86$ and a relatively larger positive bias of $\approx 16.86 - 1.97 = 14.89$.
- 2.

1. Find $\beta_1 \approx 0.37$ with an r-squared of 0.1561. This is greater and more significant effect than found under the first variable specification. This makes sense because higher-ability types are more likely to go to college, so redefining occupation based on the intersection of high-ability and college would have a greater effect.
2. We find $\beta_{white-collar} \approx 14.78$ with positive bias $\approx 14.78 - 1.97 = 12.81$. For non-white-collar workers, we find $\beta_{non-white-collar} \approx 3.23$ for a relatively *smaller* positive bias $\approx 3.23 - 1.97 = 1.26$.
3. Under Part I (b), college helps low-ability types get white-collar jobs, so within the white-collar sample college degree recipients are lower-ability than average, creating a negative selection bias compared to the undivided regression (the selection bias is still positive overall however, since even within the low-ability group, higher ability types are still more likely to attend college). In comparison, the non-white-collar sample includes **no** low-ability college graduates, leading to a more positive bias compared to the undivided regression.

Under Part II (b), college helps high-ability types get white-collar jobs, so within the white-collar sample college degree holders are higher-ability than average, creating a greater positive bias compared to the undivided regression. In comparison, the non-white-collar sample includes **no** high-ability college graduates, meaning we are only seeing the effect of college on earnings low-ability types (compared to earnings for all abilities with no degree). The means the selection bias is significantly positive less compared to the undivided regression.

	Part I	Part II
White Collar	College helps low-ability types get WC jobs → College=1 lower ability than average	College helps high-ability types get WC jobs → College=1 higher ability than average
Not	Includes no low-ability college graduates → College=1 much higher ability than average	Includes no high-ability college graduates → Significantly less upwards bias

3. Without controlling for occupation, we have standard upwards selection bias caused by ability. Since occupation is impacted by both college and ability, if we hold occupation fixed then within each occupation category, a person's probability of having attended college is tied to their underlying ability through a second channel, adding another source of bias from ability differences. This could show up as an estimate either closer or further from the true effect (as we see above) but any reduction in bias is not because the estimates are cleaner but rather because there is another source of bias pulling in the opposite direction.