R Notebook

```r
set.seed(123)

n <- 10000

beta_xy <- 2
beta_ux <- 3
beta_uy <- 4
beta_zx <- 1.5

sd_u <- 1
sd_x <- 1
sd_y <- 1
sd_z <- 1

U <- rnorm(n, mean = 0, sd = sd_u)
Z <- rnorm(n, mean = 0, sd = sd_z)

X <- beta_zx * Z + beta_ux * U + rnorm(n, mean = 0, sd = sd_x)
Y <- beta_xy * X + beta_uy * U + rnorm(n, mean = 0, sd = sd_y)

naive_model <- lm(Y ~ X)
summary(naive_model)
```

```
##
## Call:
## lm(formula = Y ~ X)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.3015 -1.5390  0.0182  1.5393  8.8255
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.012248   0.022795   0.537    0.591
## X           2.967904   0.006474 458.445   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.279 on 9998 degrees of freedom
## Multiple R-squared:  0.9546, Adjusted R-squared:  0.9546
## F-statistic: 2.102e+05 on 1 and 9998 DF,  p-value: < 2.2e-16
```

```r
first_stage <- lm(X ~ Z)
X_hat <- predict(first_stage)
```

```r
iv_model <- lm(Y ~ X_hat)
summary(iv_model)
```

```
##
## Call:
## lm(formula = Y ~ X_hat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -40.765  -6.919  -0.080   6.926  39.098
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.01442    0.10254  -0.141    0.888
## X_hat        2.01130    0.06757  29.766   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.25 on 9998 degrees of freedom
## Multiple R-squared:  0.08141,    Adjusted R-squared:  0.08132
## F-statistic:   886 on 1 and 9998 DF,  p-value: < 2.2e-16
```

```r
cat("\nTrue coefficient for X on Y:", beta_xy, "\n")
```

```
##
## True coefficient for X on Y: 2
```

```r
cat("Naive estimate (without instrument):", coef(naive_model)["X"], "\n")
```

```
## Naive estimate (without instrument): 2.967904
```

```r
cat("IV estimate (with Z as instrument):", coef(iv_model)["X_hat"], "\n")
```

```
## IV estimate (with Z as instrument): 2.011305
```