# Financial Econometrics
# Shanghai Jiao Tong University
# Summer School – 2025
# Empirical Exercise 3A
# **With Solutions**

Fatih Kansoy

*Reference: Stock, J.H. and Watson, M.W., Introduction to Econometrics (4th Edition, Global Edition)*

*Data:* `Earnings_and_Height.csv` *from* `http://www.pearsonglobaleditions.com`

**Empirical Exercise 5.1.** Use the data set `Earnings_and_Height` described in Empirical Exercise 4.2 to carry out the following exercises.

**(a)** Run a regression of *Earnings* on *Height*.

(i) Is the estimated slope statistically significant?

(ii) Construct a 95% confidence interval for the slope coefficient.

**(b)** Repeat (a) for women.

**(c)** Repeat (a) for men.

**(d)** Test the null hypothesis that the effect of height on earnings is the same for men and women. (*Hint:* See Exercise 5.15.)

**(e)** One explanation for the effect of height on earnings is that some professions require strength, which is correlated with height. Does the effect of height on earnings disappear when the sample is restricted to occupations in which strength is unlikely to be important?

**Solution:**

## Key Formulas for Hypothesis Testing

**t-statistic for testing $H_0 : \beta_1 = 0$:**

$$t = \frac{\hat{\beta}_1 - 0}{SE(\hat{\beta}_1)} = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}$$

**Critical values (two-sided test):**

| Significance Level | Critical Value | Reject $H_0$ if |
|---|---|---|
| 10% ($\alpha = 0.10$) | 1.645 | $\lvert t\rvert > 1.645$ |
| 5% ($\alpha = 0.05$) | 1.96 | $\lvert t\rvert > 1.96$ |
| 1% ($\alpha = 0.01$) | 2.576 | $\lvert t\rvert > 2.576$ |

**95% Confidence Interval:**

$$\hat{\beta}_1 \pm 1.96 \times SE(\hat{\beta}_1)$$

## Data Overview

| Sample | n | Mean Height (in) | Mean Earnings ($) |
|---|---|---|---|
| All Workers | 17,870 | 66.96 | 46,875 |
| Women (sex = 0) | 9,974 | 64.5 | 45,400 |
| Men (sex = 1) | 7,896 | 70.1 | 48,900 |

## (a) Regression: All Workers

### Regression Output

| | Coefficient | Std. Error | t-statistic | p-value |
|---|---|---|---|---|
| Height | 707.67 | 50.49 | 14.02 | 0.000 |
| Constant | −512.73 | 3,386.86 | −0.15 | 0.880 |
| $n = 17,870$, $R^2 = 0.0109$, Root MSE = \$26,777 | | | | |

### (i) Is the Slope Statistically Significant?

**Hypothesis Test:**

- $H_0 : \beta_1 = 0$ (height has no effect on earnings)
- $H_1 : \beta_1 \neq 0$ (height has an effect on earnings)

**Test Statistic:**

$$t = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} = \frac{707.67}{50.49} = 14.02$$

**Decision:**

- $|t| = 14.02 > 2.576$ (1% critical value) $\Rightarrow$ **Reject $H_0$ at 1% level**
- $p$-value $\approx 0.0000 < 0.01$

**Answer:** Yes, the slope is **highly statistically significant**. With $t = 14.02$ and $p \approx 0$, we reject $H_0$ at any conventional significance level.

### (ii) 95% Confidence Interval

$$95\% \text{ CI} = \hat{\beta}_1 \pm 1.96 \times SE(\hat{\beta}_1) = 707.67 \pm 1.96 \times 50.49 = 707.67 \pm 98.96$$

**95% Confidence Interval:** (\$608.71, \$806.63)

**Interpretation:** We are 95% confident that each additional inch of height is associated with between \$608.71 and \$806.63 higher annual earnings. Since the interval does not include zero, this confirms the slope is statistically significant.

## (b) Regression: Women Only

### Regression Output

| | Coefficient | Std. Error | t-statistic | p-value |
|---|---|---|---|---|
| Height | 511.22 | 98.90 | 5.17 | 0.000 |
| Constant | 12,650.86 | 6,383.74 | 1.98 | 0.048 |
| $n = 9,974$, $R^2 = 0.0027$, Root MSE = $26,801 | | | | |

### (i) Is the Slope Statistically Significant?

$$t = \frac{511.22}{98.90} = 5.17$$

- $|t| = 5.17 > 2.576 \Rightarrow$ **Reject $H_0$ at 1% level**
- $p$-value $\approx 0.0000$

**Answer:** Yes, the slope is **statistically significant** for women. Each additional inch is associated with $511.22 higher earnings.

### (ii) 95% Confidence Interval

$$95\%\ \text{CI} = 511.22 \pm 1.96 \times 98.90 = 511.22 \pm 193.84$$

**95% Confidence Interval:** ($317.39, $705.06)

## (c) Regression: Men Only

### Regression Output

| | Coefficient | Std. Error | t-statistic | p-value |
|---|---|---|---|---|
| Height | 1,306.86 | 100.77 | 12.97 | 0.000 |
| Constant | $-43,130.34$ | 7,068.48 | $-6.10$ | 0.000 |
| $n = 7,896$, | $R^2 = 0.0209$, | Root MSE = $26,671 | | |

### (i) Is the Slope Statistically Significant?

$$t = \frac{1306.86}{100.77} = 12.97$$

- $|t| = 12.97 > 2.576 \Rightarrow$ **Reject $H_0$ at 1% level**
- $p$-value $\approx 0.0000$

**Answer:** Yes, the slope is **highly statistically significant** for men. Each additional inch is associated with $1,306.86 higher earnings.

### (ii) 95% Confidence Interval

$$95\% \text{ CI} = 1306.86 \pm 1.96 \times 100.77 = 1306.86 \pm 197.51$$

**95% Confidence Interval:** ($1,109.36, $1,504.36)

## Comparison: All Workers vs. Women vs. Men

| | All Workers | Women | Men |
|---|---|---|---|
| Sample Size | 17,870 | 9,974 | 7,896 |
| Slope ($\hat{\beta}_1$) | 707.67 | 511.22 | **1,306.86** |
| Std. Error | 50.49 | 98.90 | 100.77 |
| t-statistic | 14.02 | 5.17 | 12.97 |
| 95% CI Lower | 608.71 | 317.39 | 1,109.36 |
| 95% CI Upper | 806.63 | 705.06 | 1,504.36 |
| Significant? | Yes | Yes | Yes |

**Key Observation:** The height premium for men ($1,307/inch) is more than **2.5 times** the premium for women ($511/inch).

## (d) Testing Whether the Effect Differs by Gender

### Method: Interaction Model

To test whether the effect of height differs between men and women, we estimate an **interaction model**:

$$\text{Earnings}_i = \beta_0 + \beta_1 \cdot \text{Height}_i + \beta_2 \cdot \text{Male}_i + \beta_3 \cdot (\text{Height}_i \times \text{Male}_i) + u_i$$

where $\text{Male}_i = 1$ if the worker is male, and 0 otherwise.

**Interpretation of coefficients:**

- $\beta_1$ = effect of height for women (base group)
- $\beta_1 + \beta_3$ = effect of height for men
- $\beta_3$ = *difference* in the height effect between men and women

**Hypothesis:**

- $H_0 : \beta_3 = 0$ (effect of height is the same for men and women)
- $H_1 : \beta_3 \neq 0$ (effect of height differs by gender)

### Regression Output

| | **Coefficient** | **Std. Error** | **t-statistic** | **p-value** |
|---|---|---|---|---|
| Height | 511.22 | 98.69 | 5.18 | 0.000 |
| Male | $-55,781.20$ | 9,529.61 | $-5.85$ | 0.000 |
| Height × Male | **795.64** | **141.24** | **5.63** | **0.000** |
| Constant | 12,650.86 | 6,370.12 | 1.99 | 0.047 |
| $n = 17,870, \quad R^2 = 0.0135$ | | | | |

### Test Results

**Coefficient on Height × Male ($\hat{\beta}_3$):**

$$\hat{\beta}_3 = 795.64, \quad SE(\hat{\beta}_3) = 141.24, \quad t = \frac{795.64}{141.24} = 5.63$$

**95% Confidence Interval for $\beta_3$:**

$$795.64 \pm 1.96 \times 141.24 = (518.81, 1072.46)$$

**Decision:**

- $|t| = 5.63 > 2.576 \Rightarrow$ **Reject $H_0$ at 1% level**
- $p$-value $\approx 0.0000$
- 95% CI excludes zero

**Verification**

The interaction model confirms our earlier separate regressions:

- Effect for women: $\hat{\beta}_1 = 511.22$ \$/inch
- Effect for men: $\hat{\beta}_1 + \hat{\beta}_3 = 511.22 + 795.64 = 1,306.86$ \$/inch

**Conclusion: Reject $H_0$.** The effect of height on earnings is **significantly different** for men and women. Men gain \$795.64 more per inch of height than women do. This difference is statistically significant at the 1% level.

## (e) Does the Height Effect Disappear in Non-Physical Occupations?

### Motivation

One hypothesis for the height-earnings relationship is that taller workers are stronger, and strength is valuable in certain occupations. If this were the *only* explanation, the height effect should disappear in occupations where physical strength is irrelevant (e.g., office work, professional services).

### Results by Occupation

| Occup. | Slope ($/in) | SE | t-stat | p-value | n |
|---|---|---|---|---|---|
| 1 | 469.46 | 155.20 | 3.02 | 0.003 | 1,906 |
| 2 | 622.76 | 117.27 | 5.31 | 0.000 | 3,158 |
| 3 | 649.72 | 214.63 | 3.03 | 0.003 | 875 |
| 4 | 1,372.38 | 148.80 | 9.22 | 0.000 | 1,957 |
| 5 | 201.22 | 133.74 | 1.50 | 0.133 | 3,124 |
| 6 | −172.89 | 680.40 | −0.25 | 0.800 | 113 |
| 7 | 1,503.04 | 403.08 | 3.73 | 0.000 | 364 |
| 8 | 62.86 | 128.27 | 0.49 | 0.624 | 1,980 |
| 9 | 1,049.20 | 308.82 | 3.40 | 0.001 | 361 |
| 10 | 571.22 | 331.53 | 1.72 | 0.086 | 534 |
| 11 | 967.01 | 306.48 | 3.16 | 0.002 | 616 |
| 12 | 1,080.32 | 286.49 | 3.77 | 0.000 | 439 |
| 13 | 972.91 | 150.64 | 6.46 | 0.000 | 1,268 |
| 14 | 1,138.41 | 268.03 | 4.25 | 0.000 | 684 |
| 15 | 549.12 | 249.24 | 2.20 | 0.028 | 491 |

### Grouped Analysis

| Occupation Group | Slope | t-stat | n | Significant? |
|---|---|---|---|---|
| Professional/Service (Occup. $\geq 6$) | $1,148.43 | 16.45 | 6,850 | Yes |
| Manual/Physical (Occup. $\leq 5$) | $829.78 | 12.64 | 11,020 | Yes |

**Detailed Regression: Non-Physical Occupations**

| | **Coefficient** | **Std. Error** | **t-statistic** | **p-value** |
|---|---|---|---|---|
| Height | 1,148.43 | 69.81 | 16.45 | 0.000 |
| Constant | $-40,740.75$ | 4,724.95 | $-8.62$ | 0.000 |
| $n = 6,850$, $R^2 = 0.0380$ | | | | |

**Answer:** No, the height effect does **NOT** disappear in non-physical occupations.

In fact, the height premium is *larger* in professional/service occupations ($1,148/inch) than in manual/physical occupations ($830/inch). The effect remains highly significant ($t = 16.45$, $p \approx 0$).

This suggests that physical strength is **not** the primary explanation for the height-earnings relationship. Other factors—such as social perceptions, confidence, or correlations with other characteristics—likely play important roles.

## Summary: Key Takeaways

| Sample | Slope ($/inch) | t-stat | 95% CI | Significant? |
|---|---|---|---|---|
| All Workers | 707.67 | 14.02 | (608.71, 806.63) | Yes |
| Women | 511.22 | 5.17 | (317.39, 705.06) | Yes |
| Men | 1,306.86 | 12.97 | (1,109.36, 1,504.36) | Yes |
| Difference (M − W) | 795.64 | 5.63 | (518.81, 1,072.46) | Yes |

## The Big Picture

This exercise illustrates several important concepts:

1. **Statistical Significance:** All slopes are statistically significant at the 1% level, with large t-statistics and p-values near zero.

2. **Heterogeneous Effects:** The height premium differs substantially by gender—men gain $1,307 per inch while women gain only $511. This difference is itself statistically significant.

3. **Testing Differences:** The interaction model provides a formal test for whether coefficients differ across groups. The significant coefficient on Height × Male confirms the gender difference.

4. **Robustness:** The height-earnings relationship persists even in occupations where physical strength is unlikely to matter, suggesting that strength is not the primary mechanism.

5. **Economic vs. Statistical Significance:** While statistically significant, height explains only 1–2% of the variation in earnings ($R^2 = 0.01$ to $0.04$). Many other factors are more important determinants of earnings.

*Calculations performed using Stata. See accompanying `E3A_solution.do` file.*