Regressions: *Empirical Examples*

Brunel MSc Economics and Finance

Prof. Menelaos Karanasos

Brunel University

1. **Soybean Yield and Fertilizer**

Suppose the soybean yield is determined by the model

$$\text{yield}_i = a + b\text{fertilizer}_i + e_i, \;\; i = 1, \ldots, N.$$

The agricultural researcher is interest in the effect of fertilizer on yield, holding other factors fixed. This effect is given by b.

The error term e_i contain factors such as land quality, rainfall and so on.

2. **A simple wage equation**

A model relating a person's wage to observed education and other unobserved factors is

$$\text{Wage}_i = a + b\text{Educ}_i + e_i, \;\; i = 1, \ldots, N$$

If wage is measured in dollars per hour and educ is years of education, then b measures the change in hourly wage given another year of education, holding other factors fixed.
Some other factors include labor force experience, innate ability, tenure with current employer, work ethic, and innumerable other things.

Using data from Wooldridge's book (Chapter 2), where $N = 526$ individuals from a population of people in the workforce in 1976, we obtain the following OLS regression line:

$$\widehat{\text{Wage}}_i = -0.90 + 0.54\text{Educ}_i, \;\; i = 1, \ldots, N$$

We must interpret the equation with caution. The intercept literally means that a person with no education has a predicted hourly wage of -90 cents an hour. This, of course, is silly.
It turns out that only 18 people in the sample have less than eight hour of education. Consequently, it is not surprising that the regression line does poorly at very low education.

For a person with eight hours of education, the predicted wage is $-0.90 + 0.54(8) = 3.42$ dollars per hour (in 1976 dollars).
The slope estimate implies that one more year of education increases hourly wage by 54 cents an hour. Therefore, four more years of education increase the predicted wage by $0.54(4) = 2.16$ dollars per hour. These are fairly large effects.
Because of the linear nature of the above equation, another year of education increases the wage by the same amount, regardless of the initial level of education.
Later on, we discuss some methods that allow for nonconstant marginal effects of our explanatory variables.

3. **CEO Salary and Return on Equity**

For the population of chief executive officers, let y be annual salary in thousands of dollars.
Let x be the average return on equity (roe) for the CEO's firm for the previous three years (return on equity is defined in terms of net income as a percentage of common equity).
To study the relationship between this measure of firm performance and CEO compensation, we postulate the simple model

$$\text{salary}_i = a + b\text{roe}_i + e_i, \;\; i = 1, \ldots, N.$$

The data set in Wooldridge's book contains information on 209 CEOs for the year 1990.

using the data, the OLS regression line relating salary to roe is

$$\widehat{\text{salary}}_i = 963.191 + 18.501\text{roe}_i, \;\; i = 1, \ldots, N.$$

The predicted change in salary as function of the change in roe: $\Delta\, \widehat{\text{salary}} = 18.501(\Delta\text{roe})$.

This means that if the return on equity increases by one percentage point, $\Delta\text{roe} = 1$, then salary is predicted to change by about 18.5 or 18,500 dollars.

We can easily use the above equation to compare predicted salaries at different values of roe.

Suppose $\text{roe} = 30$. Then $\widehat{\text{salary}} = 963.191 + 18.501(30) = 1518,221$, which is just over 1.5 million dollars.

4. **Voting Outcomes and Campaign Expenditures**

Another example is with data on election outcomes and campaign expenditures for 173 two-party races for the U.S. House Representatives in 1988.

There are two candidates in each race. Let voteA be the percentage of the vote received by candidate A and shareA be the percentage of total campaign expenditures accounted for by candidateA.

Many factors other than shareA affect the election outcome (including the quality of the candidates).

Nevertheless, we can estimate a simple regression model to find out whether spending more relative to one's challenge implies higher percentage of the vote.

The estimate model using the 173 observation of the vote is

$$\widehat{\text{voteA}}_i = 26.81 + 0.464\text{shareA}_i, \;\; i = 1, \ldots, N.$$

This means that if the share of Candidate A's spending increases by one percentage point, Candidate A receives almost one-half a percentage point (0.464) more of the total vote.
Whether or not this is a causal effect is unclear, but it is not unbelievable.
In some cases, regression analysis is not used to determine causality but to simply look at whether two variables are positively or negatively correlated like a standard correlation analysis.

2A. **A log wage equation**

Using the same data as in example 2, but using log(wage) as the dependent variable, we obtain the following relationship

$$\widehat{\log(\text{Wage})}_i = 0.584 + 0.083\text{Educ}_i, \ \ i = 1, \dots, N$$

The coefficient on educ has a percentage interpretation: wage increases by 8.3% for every additional year of education.

It is important to remember that the main reason for using the log of wage in the above equation is to impose a constant percentage effect of education on wage.

3A.**CEO Salary and Return on Equity in Logs**

We can estimate a constant elasticity model relating CEO salary to firm sales.

The data set is the same one used in Example 3, except we now relate salary to sales.

Let sales be annual firm sales, measured in millions of dollars.

A constant elasticity model is

$$\widehat{\log(\text{salary})}_i = 4.822 + 0.257 \log(\text{sales})_i, \;\; i = 1, \ldots, N.$$

The coefficient of log(sales) is the estimated elasticity of salary with respect to sales.

It implies that a 1 percent increase in firm sales increases CEO salary by about 0.257 percent.

5. **Test Scores and the Student-Teacher ratio**

When OLS is used to estimate a line relating the student-teacher ratio (STR) to test scores using 420 observations (districts), we have:

$$\widehat{\text{TestScore}}_i = 698.9 - \underset{(0.52)}{2.28}\,\text{STR}_i,\ \ i = 1, \ldots, N.$$

To test the null hypothesis that $b = 0$, against the alternative hypothesis that $b \neq 0$ we construct the $t-$statistic:

$$t = \frac{\beta - b_0}{se(\beta)} \frac{(-2.28 - 0)}{0.52} = -4.38.$$

This $t-$statistic exceeds the two-sided critical value of 1.96, so the null hypothesis is rejected

5A. **Binary Regression**

Suppose that you have a variables D_i that equals either 0 or 1, depending on whether the STR is less than 20:

$$D_i = \begin{cases} 1 & \text{if} \quad \text{STR in } i\text{th district } < 20 \\ 0 & \text{if} \quad \text{STR in } i\text{th district } \geq 20 \end{cases}.$$

Such a regression estimated by OLS yields:

$$\widehat{\text{TestScore}}_i = \underset{(1.3)}{650} + \underset{(1.8)}{7.4}\, D_i, \;\; i = 1, \ldots, N,$$

where the standard errors of the OLS estimates are given in parentheses below the OLS estimates.

Thus the average test score for the subsample with STR grater than or equal to 20 is 650, and the average score for the subsample with STR less than 20 is $650 + 7.4 = 657.4$

Is the difference in the population mean scores in the two groups statistically significant different from zero at the 5% level?
To find out, construct the $t-$statistic on β:

$$t = \frac{\beta - b_0}{se(\beta)} \frac{(7.4 - 0)}{1.8} = 4.04 > 1.96$$

so we reject the null hypothesis that $b = 0$.

5B. **Test Scores and the Class Size**

Suppose that a researcher, using data on class size (CS) and average test scores from 100 third-grade classes, estimates the OLS regression

$$\widehat{\text{Test Score}} = \underset{(20.4)}{520.4} + \underset{(2.21)}{5.82}\,\text{CS}_i, \;\; i = 1, \ldots, 100.$$

a. A classroom has 22 students. What is the regression's prediction for that classroom's average test score?

b. Last year a classroom had 19 students, and this year it has 23 students. What is the regression's prediction for the change in the classroom average test score?

c. Do you reject the null hypothesis that $b = 0$, at the 5% level?

The $t-$test for the null hypothesis, $H_0 : b = 0$ against the alternative $H_1 : b \neq 0$ is: $t = \frac{\beta}{se(\beta)} = \frac{5.82}{2.21} = 2.63$.

Thus we reject the null hypothesis.

6. **Wage and Male/Female Workers**

Suppose that a researcher, using wage data on 250 randomly selected male workers and 280 female workers, estimates the OLS regression,

$$\widehat{\text{Wage}} = \underset{(0.18)}{12.68} + \underset{(0.84)}{2.79}\,\text{Male}_i, \;\; i = 1, \ldots, 530,$$

where wage is measured in \$/hour and male is a binary variable that is equal to one if the person is a male and zero if the person is female. Define the wage gender cap as the difference in the mean earnings between men and women.

a. What is the estimated gender gap?

b. Is the estimated gender gap significantly different from zero?

The $t-$test for the null hypothesis, $H_0 : b = 0$ against the alternative $H_1 : b \neq 0$ is: $t = \frac{\beta}{se(\beta)} = \frac{2.79}{0.84} = 3.32$.

Thus we reject the null hypothesis.

7. **Weight and Height**

Suppose that a random sample of 200 20-year old men is selected from a population and their height and weight is recorded. A regression of weight on height yields:

$$\widehat{\text{Weight}} = \underset{(2.15)}{-99.41} + \underset{(0.31)}{3.94}\,\text{Height}_i, \;\; i = 1, \ldots, 200,$$

where weight is measured in pounds and height is measured in inches.

a. What is the regression's weight prediction for someone who is 70 inches tall? 65 inches tall? 74 inches tall?

b. Test the null hypothesis that $b = 5$:

The $t-$test for the null hypothesis, $H_0 : b = 5$ against the alternative $H_1 : b \neq 5$ is: $|t| = \left|\frac{\beta - 5}{se(\beta)}\right| = \left|\frac{3.94-5}{0.31}\right| = 3.42.$

Thus we reject the null hypothesis.

8. Okun's Law

Okun (1962) developed an empirical relationship, using quarterly data from 1947:2 to 1960:4, between changes in the state of the economy (captured by changes in GNP) and changes in the unemployment rate, known as Okun's law as follows:

$$\Delta \text{UNEMPL}_t = a + b\Delta \text{GNP}_t + e_t.$$

Applying OLS the sample regression equation that Okun obtained was:

$$\Delta\widehat{\text{UNEMPL}}_t = \underset{(0.09)}{0.3} + \underset{(0.07)}{-0.3}\Delta\text{GNP}_t + e_t.$$

From the obtained results we conclude that when the economy does not grow the unemployment rate rises by 3 per cent.

The negative b coefficient suggests that when the state of the economy improves, the unemployment rate falls. The relationship, though is less than one to one.

The $t-$test for the null hypothesis, $H_0 : b = 0$ against the alternative $H_1 : b \neq 0$ is: $|t| = \left|\frac{\beta}{se(\beta)}\right| = \frac{0.3}{0.07} = 4.28$.

Thus we reject the null hypothesis.

9. **Keynesian Consumption Function**

Another basic relationship in economic theory is the Keynesian consumption function that simply states that consumption (C_t) is a positive linear function of disposable (after tax) income (Y_t^d). The relationship is as follows

$$C_t = a + bY_t^d,$$

where a is the autonomous consumption (consumption even when disposable income is zero) and b is the marginal propensity to consume. In this function we expect $a > 0$, $0 > b > 1$. A $\beta = 0.7$ means that the marginal propensity to consume is 0.7.

A Keynesian consumption function is estimated below:

$$\widehat{C}_t = \underset{(6.565)}{15.116} + \underset{(0.038)}{0.160}\, Y_t^d.$$

From this we can (a) read estimated effects of changes in the explanatory variable on the dependent variable
(b) predict values of the dependent variable for given values of the explanatory variable
(c) perform hypothesis testing for the estimated coefficients
(d) construct confidence intervals for the estimated coefficients
For example, the $t-$test for the null hypothesis, $H_0 : b = 0$ against the alternative $H_1 : b \neq 0$ is: $t = \frac{\beta}{se(\beta)} = \frac{0.160}{0.038} = 4.21.$
Thus we reject the null hypothesis.

10. **CAPM Theory**

The following equation has been estimated by OLS:

$$\widehat{R}_t = \underset{(0.33)}{0.567} + \underset{(0.066)}{1.04}\, R_{mt}, \quad n = 250,$$

where R_t and R_{mt} denote the excess return of a stock and the excess return of the market index for the London Stock Exchange.

(a) Are these coefficients statistically significant? Explain what is the meaning of your findings regarding the CAPM theory.

(b) Test the hypothesis $H_0 : b = 1$ and $H_1 : b < 1$ at the 5% level of significance. If you reject H_0 what does this indicate about this stock?

The $t-$test for the null hypothesis, $H_0 : b = 0$ against the alternative $H_1 : b \neq 0$ is: $t = \frac{\beta}{se(\beta)} = \frac{1.04}{0.066} = 15.75$.
Thus we reject the null hypothesis.
The $t-$test for the null hypothesis, $H_0 : b = 1$ against the alternative $H_1 : b < 1$ is: $t = \frac{\beta - 1}{se(\beta)} = \frac{0.4}{0.066} = 6.06$.
Thus we reject the null hypothesis.

Multivariate regressions

2* **Wage equation**

The first example is a simple variation of the wage equation introduced in example 2 (with 526 workers). We will assume that wage is determined by three explanatory variables, education (years of education), experience (years of labour market experience) and tenure (years with current employer):

$$\widehat{\log(\text{wage})} = \underset{(.104)}{.284} + \underset{(.007)}{.092}\,\text{educ} + \underset{(.0017)}{.0041}\,\text{exper} + \underset{(.003)}{.022}\,\text{tenure},$$

where standard errors appear in parentheses below the estimated coefficients.

The coefficient .092 means that, holding exper and tenure fixed, another year of education is predicted to increase log(wage) by .092, which translates into an approximate 9.2 percent $[100(.092)]$ increase in wage.

In the above equation we can obtain the estimated effect on wage when an individual stays at the same firm for another year: exper and tenure both increase by one year. the total effect (holding educ) fixed is $.0041 + .022 = .0261$ or about 2.6 percent.

$$\widehat{\log(\text{wage})} = \underset{(.104)}{.284} + \underset{(.007)}{.092}\,\text{educ} + \underset{(.0017)}{.0041}\text{exper} + \underset{(.003)}{.022}\,\text{tenure},$$

This equation can be used to test whether the return to exper, controlling for educ and tenure, is zero in the population, against the alternative that it is positive.

Write this as $H_0 : b_{\exp er} = 0$ against the alternative $H_1 : b_{\exp er} > 0$.

Since we have 522 degrees of freedom, we can use the standard normal critical values. The 5% critical value is 1.645.

The t statistic is $t = \frac{\beta_{\exp er}}{se(\beta)} = \frac{0.041}{0.0017} = 2.41$ and so $\beta_{\exp er}$ is statistically significant.

5* **Test Scores and Student-Teacher Ratio**

Can we reject the hypothesis that a change in the student-teacher ratio has no effect on test scores, once we control for the percentage of English learners in the district?

$$\widehat{\text{TestScore}}_i = \underset{(8.7)}{686.0} - \underset{(0.43)}{1.10}\,\text{STR}_i - \underset{(0.031)}{0.650}\,\text{PctEL}, \;\; n = 420.$$

The $|t|$-statistic for β_{STR} is $\frac{1.10}{0.43} = 2.54$ Thus the null hypothesis can be rejected at the 5% level.

Next we want to see what is the effect on test scores of reducing the student teacher ratio, holding expenditures per pupil (and the percentage of English learners) constant? The OLS regression line is

$$\widehat{\text{TestScore}}_i = \underset{(15.5)}{649.6} - \underset{(0.48)}{0.29}\,\text{STR}_i + \underset{(1.59)}{3.87}\,\text{Expn} - \underset{(0.032)}{0.656}\,\text{PctEL}, \;\; n = 420,$$

where expn is total annual expenditures per pupil in the district in thousands of dollar.
The result is striking. Holding expenditures per pupil and the percentage of English learners constant, changing the student-teacher ratio is estimated to have a very small effect on test scores, β_{STR} is only -0.29. Moreover, the $|t|$-statistic for testing that the true value of the coefficient is zero is now $\frac{0.29}{0.48} = 0.60$.

11. **Determinants of College GPA**

We obtain the following OLS regression line to predict college GPA (grade point average) from high school GPA and achievement test score (ACT)

$$\widehat{colGPA} = 1.29 + .453GPA + .0094ACT.$$

:

Holding ACT fixed, another point on GPA is associated with .453 of a point on the college GPA, or almost half a point.

In other words if we choose two students, A and B, and these students have the same ACT score, but the high school GPA of student A is one point higher than the high school GPA of student B, then we predict student A to have a college GPA .453 higher than that of student B.

11* **Determinants of College GPA**

We estimate a model explaining college GPA with the average number of lectures missed per week (skipped) as an additional explanatory variable. the estimated model is

$$\widehat{colGPA} = \underset{(0.33)}{1.39} + \underset{(.094)}{.412}\,GPA + \underset{(.011)}{.015}\,ACT - \underset{(.026)}{.083}\,skipped, \quad n = 141.$$

We can easily compute t statistics to see which variables are statistically significant. The 5% critical value is about 1.96, since the degrees of freedom $141 - 4 = 137$ is large enough to use the standard normal approximation.

The t statistic on GPA is 4.38 which is significant. The t statistic on ACT is 1.36, which is not statistically significant. The coefficient on ACT is also practically small.

The coefficient on skipped has a t statistic of $\frac{-0.83}{.026} = -3.19$, so skipped is statistically significant $(3.19 > 1.96)$.

Holding GPA and ACT fixed, the predicted difference in colGPA between a student who misses no lectures per week and a student who misses five lectures per week is about .42.

12. **Housing Prices**

For a sample of 506 communities in the Boston area, we estimate a model relating median housing price (price) in the community to various community characteristics:

nox is the amount of nitrogen oxide in the air, in parts per million

dist is a weighted distance of the community from five employment centers, in miles

rooms is the average number of rooms in houses in the community

stratio is the average student-teacher ratio of schools in the community.

The population model is

$$\log(\text{price}) = \underset{(0.32)}{11.08} - \underset{(.117)}{.954}\log(\text{nox}) - \underset{(.043)}{.134}\log(\text{dist}) + \underset{(.019)}{.255}(\text{rooms}) - \underset{(.006)}{.052}(\text{stratio}),\ n = 506.$$

The slope estimates all have the anticipated signs. Each coefficient is statistically different from zero, including the coefficient on log(nox). But we do not want to test that $\beta_{nox} = 0$. the null hypothesis of interest is $H_0 : \beta_{nox} = -1$, with corresponding t-statistic $\frac{(-.954+1)}{.117} = .393$. There is no need to look in the t table for a critical value when the t statistic is this small: the estimated elasticity is not statistically different from -1.

13. **Participation rates in 401(k) Plans**

We use data on 401(k) plans to estimate a model describing participation rates in terms of the firm's match rate (mrate), the age of the plan (age) and a measure of the firm size, the total number of firm employees(totemp). The estimated equation is

$$\widehat{\text{prate}} = \underset{(0.78)}{80.29} + \underset{(0.52)}{5.44}\,\text{mrate} + \underset{(.045)}{.269}\,\text{age} - \underset{(.00004)}{.0013}\,\text{totemp}, \quad n = 1,534.$$

The smallest t statistic in absolute value is that on the variable totemp: $t = \frac{-.00013}{.00004} = -3.25$, and thesis statistically significant. Thus, all of the variables are statistically significant.

Holding mrate and age fixed, if a firm grows by $10,000$ employees, the participation rate falls by $10,000(.00013) = 1.3$ percentage points.

This is a huge increase in number of employees with only a modest effect on the participation rate.

Thus, while firm size does affect the participation rate, the effect is not practically very large.

14. **Effect of Job Training Grants on firm Scrap Rates**

The scrap rate for a manufacturing firm is the number of defective items out of every 100 items produced that must be discarded. Thus, a decrease in the scrap rate reflects higher productivity. We can use the scrap rate to measure the effect of worker training on productivity.

For a sample ($n = 30$) of Michigan manufacturing firms in 1987, the following equation is estimated:

$$\log(\widehat{\text{scrap}}) = \underset{(4.91)}{13.72} - \underset{(.019)}{.028}\,\text{hrsemp} - \underset{(0.41)}{1.21}\,\log(\text{sales}) + \underset{(0.43)}{1.48}\,\log(\text{employ}),$$

The variable hrsemp is annual hours of training per employee, sales is annual firm sales (in dollars), and employ is number of firm employees.

The main variable of interest is hrsemp. One more hour of training per employee lowers log(scrap) by .028, which means the scrap rate is about 2.8% lower.
What about the statistical significance of the training variable? The t statistic on hrsemp is $\frac{-.028}{.019} = -1.47$, and now you probably recognize this as not being a large enough in magnitude to conclude that hrsemp is statistically significant at the 5% level.
In fact with $30 - 4 = 26$ degrees of freedom for the one sided alternative, $H_1 : \beta_{hrsemp} < 0$, the 5% critical value is about 1.71. Thus, using a strict 5% level test, we must conclude that hrsemp is not statistically significant. Because of the sample size is pretty small, we might be more liberal with the significance level. The 10% critical level is 1.32, and so hrsemp is significant against the one sided alternative at the 10% level.

15. **Hedonic price Model for Houses**

A model that explains the price of a good in terms of the good's characteristics is called an hedonic price model.

The following equation is an hedonic price model for housing prices; the characteristics are: square footage (sqrft), number of bedrooms (bdrms), and number of bathrooms (bthrms).

Often, prices appears in logarithmic form, as do some of the explanatory variables.

Using $n = 19$ observations on houses that were sold in Waltham, Massachusetts, in 1990, the estimated equation is

$$\widehat{\log(\text{price})} = \underset{(1.15)}{7.46} + \underset{(.184)}{.634}\log(\text{sqrft}) - \underset{(.059)}{.066}\,\text{bdrms} + \underset{(.075)}{.158}\,\text{bthrms}, \quad n = 19.$$

Since price and sqrft both appear in logarithmic form, the price elasticity with respect to square footage is .634, so that, holding number of bedrooms and bathrooms fixed, a 1% increase in square footage increases the predicted housing price by about .634%. We also reject $H_0 : \beta_{sqrft} = 0$, against the two sided alternative at the 5% level.

The coefficient on bdrms is negative, which seems counterintuitive. In any case, bdrms does not have a statistically significant effect on housing price. Given size and number of bedrooms, one more bathroom is predicted to increase housing price by about 15.8%. However, we fail to reject $H_0 : \beta_{bthrms} = 0$, against the two sided alternative at the 5% level.

16. **Testing hypothesis about a single linear combination of the parameters**

We consider a simple model to compare returns to education at junior colleges and four-year colleges (we refer to the latter as universities).
The estimated model is

$$\log(\widehat{\text{wage}}) = \underset{(.021)}{1.47} + \underset{(.0068)}{.0667}\text{jc} + \underset{(.0023)}{.0769}\text{univ} + \underset{(.0002)}{.0049}\text{exper}, \quad n = 6,763$$

where jc is number of years attending a two-year college and univ is number of years at a four-year college.
It is clear that jc and univ have both economically and statistically significant effects on wage.

This certainly of interest, but we more concerned about testing whether the estimated difference in the coefficients is statistically significant.

The difference is estimated as $\beta_{jc} - \beta_{univ} = -.0102$, so the return to a year at a junior college is about one percentage point less than a year at a university.

The $t-$statistic for the $H_0 : \beta_{jc} - \beta_{univ} = 0$, is given by $\frac{\beta_{jc} - \beta_{univ}}{se(\beta_{jc} - \beta_{univ})}$. To find $se(\beta_{jc} - \beta_{univ})$ we obtain first the variance:

$var(\beta_{jc} - \beta_{univ}) = var(\beta_{jc}) + var(\beta_{univ}) - 2cov(\beta_{jc}, \beta_{univ})$.

Unfortunately we are not given the $cov(\beta_{jc}, \beta_{univ})$.

We suggest another route that is much simpler to compute, and readily applied to a variety of problems.
Define a new parameter as the difference between $\theta = \beta_{jc} - \beta_{univ}$. Then we want to test $H_0 : \widehat{\theta} = 0$ against $H_1 : \widehat{\theta} < 0$.
We can do this by rewriting the model so that θ appears directly on one of the independent variables. Since $\beta_{jc} = \theta + \beta_{univ}$ we have

$$\log(\widehat{\text{wage}}) = c + (\theta + \beta_{univ})\text{jc} + \beta_{univ}\text{univ} + \beta_{\exp er}\text{exper},$$

or

$$\begin{aligned} \log(\widehat{\text{wage}}) &= c + \theta\text{jc} + \beta_{univ}(\text{jc+univ}) + \beta_{\exp er}\text{exper, or} \\ \log(\widehat{\text{wage}}) &= c + \theta\text{jc} + \beta_{univ}\text{totcoll} + \beta_{\exp er}\text{exper.} \end{aligned}$$

The estimated model is

$$\log(\widehat{\text{wage}}) = \underset{(.021)}{1.47} - \underset{(.0069)}{.0102}\text{jc} + \underset{(.0023)}{.0769}\text{totcoll} + \underset{(.0002)}{.0049}\text{exper}, \quad n = 6,763$$

The $|t|$ statistic for testing $H_0 : \widehat{\theta} = 0$ is $\frac{.0102}{.0069} = 1.48$, so there is no strong evidence against $\beta_{jc} = \beta_{univ}$.
The coefficient on the new variable totcoll, is the same as the coefficient on univ, and the standard error is also the same.

16. **Imports**

Variable	Coefficient	Std. Error	t-statistic
Constant	0.21	0.36	0.60
LOG(GDP)	1.97	0.16	12.56
LOG(CPI)	1.02	0.32	3.17
LOG(PPI)	−0.77	0.30	−2.52