Tests & Model Validation

Dr. Patrick Toche

Textbook:

James H. Stock and Mark W. Watson, *Introduction to Econometrics*, 4th Edition, Pearson.

Other references:

Jeffrey M. Wooldridge, *Introductory Econometrics: A Modern Approach*, 7th Edition, Cengage Learning.
Joshua D. Angrist and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, 1st Edition, Princeton University Press.

The textbook comes with online resources and study guides. Other references will be given from time to time.

Contents

In this lesson you will learn ...

- Tests of hypothesis for a single coefficient in a multivariate regression.
- Tests of joint hypothesis in multivariate regressions.
- The F-statistic with q restrictions.
- The overall-regression F-statistic.
- Tests of single restrictions involving multiple coefficients.
- Confidence intervals for a single coefficient.
- Confidence sets for multiple coefficients.

In this lesson you will learn …

- Tests of hypothesis for a single coefficient in a multivariate regression.
- Tests of joint hypothesis in multivariate regressions.
- The F-statistic with q restrictions.
- The overall-regression F-statistic.
- Tests of single restrictions involving multiple coefficients.
- Confidence intervals for a single coefficient.
- Confidence sets for multiple coefficients.

In this lesson you will learn ...

- Tests of hypothesis for a single coefficient in a multivariate regression.
- Tests of joint hypothesis in multivariate regressions.
- The F-statistic with q restrictions.
- The overall-regression F-statistic.
- Tests of single restrictions involving multiple coefficients.
- Confidence intervals for a single coefficient.
- Confidence sets for multiple coefficients.

In this lesson you will learn ...

- Tests of hypothesis for a single coefficient in a multivariate regression.
- Tests of joint hypothesis in multivariate regressions.
- The F-statistic with q restrictions.
- The overall-regression F-statistic.
- Tests of single restrictions involving multiple coefficients.
- Confidence intervals for a single coefficient.
- Confidence sets for multiple coefficients.

In this lesson you will learn ...

- Tests of hypothesis for a single coefficient in a multivariate regression.
- Tests of joint hypothesis in multivariate regressions.
- The F-statistic with q restrictions.
- The overall-regression F-statistic.
- Tests of single restrictions involving multiple coefficients.
- Confidence intervals for a single coefficient.
- Confidence sets for multiple coefficients.

In this lesson you will learn ...

- Tests of hypothesis for a single coefficient in a multivariate regression.
- Tests of joint hypothesis in multivariate regressions.
- The F-statistic with q restrictions.
- The overall-regression F-statistic.
- Tests of single restrictions involving multiple coefficients.
- Confidence intervals for a single coefficient.
- Confidence sets for multiple coefficients.

In this lesson you will learn …

- Tests of hypothesis for a single coefficient in a multivariate regression.
- Tests of joint hypothesis in multivariate regressions.
- The F-statistic with q restrictions.
- The overall-regression F-statistic.
- Tests of single restrictions involving multiple coefficients.
- Confidence intervals for a single coefficient.
- Confidence sets for multiple coefficients.

Single Hypothesis Test

Hypothesis Tests

Test the Hypothesis $\beta_j = \beta_{j,0}$

1. State the Null Hypothesis and the Alternate:

$$H_0: \beta_j = \beta_{j,0}; \qquad H_1: \beta_j \neq \beta_{j,0}$$

2. Compute the t-statistic: $\quad t^{\text{act}} = \frac{\hat{\beta}_j - \beta_{j,0}}{\text{SE}(\hat{\beta}_j)}$

3. Compute the p-value: $\quad p = 2\Phi(-|t^{\text{act}}|)$

4. Conclude: $\quad p < \alpha \implies$ reject H_0 $\quad$ (fail to reject otherwise).

Hypothesis Tests

Test the Hypothesis $\beta_j = \beta_{j,0}$

1. State the Null Hypothesis and the Alternate:

$$H_0: \beta_j = \beta_{j,0}; \qquad H_1: \beta_j \neq \beta_{j,0}$$

2. Compute the t-statistic: $t^{\text{act}} = \frac{\beta_j - \beta_{j,0}}{\text{SE}(\hat{\beta}_j)}$

3. Compute the p-value: $p = 2\Phi(-|t^{\text{act}}|)$

4. Conclude: $p < \alpha \implies$ reject H_0 (fail to reject otherwise).

Hypothesis Tests

Test the Hypothesis $\beta_j = \beta_{j,0}$

1. State the Null Hypothesis and the Alternate:

$$H_0 : \beta_j = \beta_{j,0}; \qquad H_1 : \beta_j \neq \beta_{j,0}$$

2. Compute the t-statistic: $\quad t^{\text{act}} = \frac{\beta_j - \beta_{j,0}}{\text{SE}(\hat{\beta}_j)}$

3. Compute the p-value: $\quad p = 2\Phi(-|t^{\text{act}}|)$

4. Conclude: $\quad p < \alpha \implies$ reject H_0 (fail to reject otherwise).

Hypothesis Tests

Test the Hypothesis $\beta_j = \beta_{j,0}$

1. State the Null Hypothesis and the Alternate:

$$H_0 : \beta_j = \beta_{j,0}; \qquad H_1 : \beta_j \neq \beta_{j,0}$$

2. Compute the t-statistic: $t^{\text{act}} = \frac{\beta_j - \beta_{j,0}}{\text{SE}(\hat{\beta}_j)}$
3. Compute the p-value: $p = 2\Phi(-|t^{\text{act}}|)$
4. Conclude: $p < \alpha \implies$ reject H_0 (fail to reject otherwise).

Effect of Student-Teacher Ratio on Test Scores

Test the Hypothesis $\boldsymbol{STR} = 0$

$$\widehat{TestScore} = \underset{(8.7)}{686.0} - \underset{(0.43)}{1.10}\ STR - \underset{(0.031)}{0.650}\ PctEL$$

1. State the hypothesis: $H_0: \ STR = 0; \quad H_1: \ STR \neq 0.$
2. Compute the t-statistic: $t^{\text{act}} = \frac{-1.10-0}{0.43} \approx -2.54.$
3. Compute the p-value: $p = 2\Phi(-2.54) \approx 1.1\%.$
4. Reject H_0 at the 5% level.
5. 95% Confidence interval: $-1.10 \pm 1.96 \cdot 0.43 \approx (-1.95, -0.26).$

Effect of Student-Teacher Ratio on Test Scores

Test the Hypothesis $STR = 0$

$$\widehat{TestScore} = \underset{(8.7)}{686.0} - \underset{(0.43)}{1.10}\ STR - \underset{(0.031)}{0.650}\ PctEL$$

1. State the hypothesis: $H_0: \ STR = 0; \quad H_1: \ STR \neq 0.$
2. Compute the t-statistic: $t^{\text{act}} = \frac{-1.10-0}{0.43} \approx -2.54.$
3. Compute the p-value: $p = 2\Phi(-2.54) \approx 1.1\%.$
4. Reject H_0 at the 5% level.
5. 95% Confidence interval: $-1.10 \pm 1.96 \cdot 0.43 \approx (-1.95, -0.26).$

Effect of Student-Teacher Ratio on Test Scores

Test the Hypothesis $STR = 0$

$$\widehat{TestScore} = \underset{(8.7)}{686.0} - \underset{(0.43)}{1.10}\ STR - \underset{(0.031)}{0.650}\ PctEL$$

1. State the hypothesis: $H_0 : \ STR = 0; \quad H_1 : \ STR \neq 0.$
2. Compute the t-statistic: $t^{\text{act}} = \frac{-1.10-0}{0.43} \approx -2.54.$
3. Compute the p-value: $p = 2\Phi(-2.54) \approx 1.1\%.$
4. Reject H_0 at the 5% level.
5. 95% Confidence interval: $-1.10 \pm 1.96 \cdot 0.43 \approx (-1.95, -0.26).$

Effect of Student-Teacher Ratio on Test Scores

Test the Hypothesis $STR = 0$

$$\widehat{TestScore} = \underset{(8.7)}{686.0} - \underset{(0.43)}{1.10}\ STR - \underset{(0.031)}{0.650}\ PctEL$$

1. State the hypothesis: $H_0: \ STR = 0; \quad H_1: \ STR \neq 0.$
2. Compute the t-statistic: $t^{\text{act}} = \frac{-1.10-0}{0.43} \approx -2.54.$
3. Compute the p-value: $p = 2\Phi(-2.54) \approx 1.1\%.$
4. Reject H_0 at the 5% level.
5. 95% Confidence interval: $-1.10 \pm 1.96 \cdot 0.43 \approx (-1.95, -0.26).$

Effect of Student-Teacher Ratio on Test Scores

Test the Hypothesis $STR = 0$

$$\widehat{TestScore} = \underset{(8.7)}{686.0} - \underset{(0.43)}{1.10}\ STR - \underset{(0.031)}{0.650}\ PctEL$$

1. State the hypothesis: $H_0: \ STR = 0; \quad H_1: \ STR \neq 0.$
2. Compute the t-statistic: $t^{\text{act}} = \frac{-1.10-0}{0.43} \approx -2.54.$
3. Compute the p-value: $p = 2\Phi(-2.54) \approx 1.1\%.$
4. Reject H_0 at the 5% level.
5. 95% Confidence interval: $-1.10 \pm 1.96 \cdot 0.43 \approx (-1.95, -0.26).$

Controlling for Expenditure Per Pupil

Test the Hypothesis $STR = 0$

$$\widehat{TestScore} = \underset{(15.5)}{649.6} - \underset{(0.48)}{0.29}\ STR - \underset{(0.032)}{0.656}\ PctEL + \underset{(1.59)}{3.87}\ Expn$$

where $Expn$ is total annual expenditures per pupil in thousands of dollars.

- Adding $Expn$ as a regressor, the estimated coefficient on STR changes from -1.10 to -0.29.
- The t-statistic associated with $H_0: \ STR = 0$ is $t^{\text{act}} = \frac{-0.29-0}{0.48} \approx -0.60$.
- H_0 cannot be rejected at the 10% significance level.
- Holding expenditures per pupil and the percentage of English learners constant, the student-teacher ratio has no significant effect on test scores.
- Interpretation: school administrators allocate their budgets efficiently.
- In other words: Reallocating spending to reduce class sizes — without increasing spending — will not help raise test scores. There is no “free lunch”.

Controlling for Expenditure Per Pupil

Test the Hypothesis $STR = 0$

$$\widehat{TestScore} = \underset{(15.5)}{649.6} - \underset{(0.48)}{0.29}\ STR - \underset{(0.032)}{0.656}\ PctEL + \underset{(1.59)}{3.87}\ Expn$$

where $Expn$ is total annual expenditures per pupil in thousands of dollars.

- Adding $Expn$ as a regressor, the estimated coefficient on STR changes from -1.10 to -0.29.
- The t-statistic associated with $H_0:\ STR = 0$ is $t^{\text{act}} = \frac{-0.29-0}{0.48} \approx -0.60$.
- H_0 cannot be rejected at the 10% significance level.
- Holding expenditures per pupil and the percentage of English learners constant, the student-teacher ratio has no significant effect on test scores.
- Interpretation: school administrators allocate their budgets efficiently.
- In other words: Reallocating spending to reduce class sizes — without increasing spending — will not help raise test scores. There is no “free lunch”.

Controlling for Expenditure Per Pupil

Test the Hypothesis $STR = 0$

$$\widehat{TestScore} = \underset{(15.5)}{649.6} - \underset{(0.48)}{0.29}\ STR - \underset{(0.032)}{0.656}\ PctEL + \underset{(1.59)}{3.87}\ Expn$$

where $Expn$ is total annual expenditures per pupil in thousands of dollars.

- Adding $Expn$ as a regressor, the estimated coefficient on STR changes from -1.10 to -0.29.
- The t-statistic associated with $H_0: \ STR = 0$ is $t^{\text{act}} = \frac{-0.29-0}{0.48} \approx -0.60$.
- H_0 cannot be rejected at the 10% significance level.
- Holding expenditures per pupil and the percentage of English learners constant, the student-teacher ratio has no significant effect on test scores.
- Interpretation: school administrators allocate their budgets efficiently.
- In other words: Reallocating spending to reduce class sizes — without increasing spending — will not help raise test scores. There is no “free lunch”.

Controlling for Expenditure Per Pupil

Test the Hypothesis $STR = 0$

$$\widehat{TestScore} = \underset{(15.5)}{649.6} - \underset{(0.48)}{0.29}\ STR - \underset{(0.032)}{0.656}\ PctEL + \underset{(1.59)}{3.87}\ Expn$$

where $Expn$ is total annual expenditures per pupil in thousands of dollars.

- Adding $Expn$ as a regressor, the estimated coefficient on STR changes from -1.10 to -0.29.
- The t-statistic associated with $H_0 : \ STR = 0$ is $t^{\text{act}} = \frac{-0.29-0}{0.48} \approx -0.60$.
- H_0 cannot be rejected at the 10% significance level.
- Holding expenditures per pupil and the percentage of English learners constant, the student-teacher ratio has no significant effect on test scores.
- Interpretation: school administrators allocate their budgets efficiently.
- In other words: Reallocating spending to reduce class sizes — without increasing spending — will not help raise test scores. There is no “free lunch”.

Controlling for Expenditure Per Pupil

Test the Hypothesis $STR = 0$

$$\widehat{TestScore} = \underset{(15.5)}{649.6} - \underset{(0.48)}{0.29}\ STR - \underset{(0.032)}{0.656}\ PctEL + \underset{(1.59)}{3.87}\ Expn$$

where $Expn$ is total annual expenditures per pupil in thousands of dollars.

- Adding $Expn$ as a regressor, the estimated coefficient on STR changes from -1.10 to -0.29.
- The t-statistic associated with $H_0: \ STR = 0$ is $t^{\text{act}} = \frac{-0.29-0}{0.48} \approx -0.60$.
- H_0 cannot be rejected at the 10% significance level.
- Holding expenditures per pupil and the percentage of English learners constant, the student-teacher ratio has no significant effect on test scores.
- Interpretation: school administrators allocate their budgets efficiently.
- In other words: Reallocating spending to reduce class sizes — without increasing spending — will not help raise test scores. There is no “free lunch”.

Controlling for Expenditure Per Pupil

Test the Hypothesis $STR = 0$

$$\widehat{TestScore} = \underset{(15.5)}{649.6} - \underset{(0.48)}{0.29}\ STR - \underset{(0.032)}{0.656}\ PctEL + \underset{(1.59)}{3.87}\ Expn$$

where $Expn$ is total annual expenditures per pupil in thousands of dollars.

- Adding $Expn$ as a regressor, the estimated coefficient on STR changes from -1.10 to -0.29.
- The t-statistic associated with $H_0: \; STR = 0$ is $t^{\text{act}} = \frac{-0.29-0}{0.48} \approx -0.60$.
- H_0 cannot be rejected at the 10% significance level.
- Holding expenditures per pupil and the percentage of English learners constant, the student-teacher ratio has no significant effect on test scores.
- Interpretation: school administrators allocate their budgets efficiently.
- In other words: Reallocating spending to reduce class sizes — without increasing spending — will not help raise test scores. There is no “free lunch”.

Joint Hypothesis Test

Joint Hypothesis Tests

F-Test for Two Restrictions:

- The F-statistic is used to test a joint hypothesis about regression coefficients.
- Joint Hypothesis: $H_0: \beta_1 = 0 \quad \text{and} \quad \beta_2 = 0$
- F-statistic with $q = 2$ restrictions:

$$F = \frac{1}{2}\left(\frac{t_1^2 + t_2^2 - 2\hat{\rho}_{t_1,t_2} t_1 t_2}{1 - \hat{\rho}^2_{t_1,t_2}}\right)$$

  where $\hat{\rho}^2_{t_1,t_2}$ is an estimator of the correlation between the two t-statistics.
- In large samples, under the null hypothesis, the F-statistic has sampling distribution $F_{q,\infty}$.

Joint Hypothesis Tests

F-Test for Two Restrictions:

- The F-statistic is used to test a joint hypothesis about regression coefficients.
- Joint Hypothesis: $H_0: \beta_1 = 0 \quad \text{and} \quad \beta_2 = 0$
- F-statistic with $q = 2$ restrictions:

$$F = \frac{1}{2}\left(\frac{t_1^2 + t_2^2 - 2\hat{\rho}_{t_1,t_2} t_1 t_2}{1 - \hat{\rho}_{t_1,t_2}^2}\right)$$

  where $\hat{\rho}_{t_1,t_2}^2$ is an estimator of the correlation between the two t-statistics.
- In large samples, under the null hypothesis, the F-statistic has sampling distribution $F_{q,\infty}$.

Joint Hypothesis Tests

F-Test for Two Restrictions:

- The F-statistic is used to test a joint hypothesis about regression coefficients.
- Joint Hypothesis: $H_0: \ \beta_1 = 0 \quad \text{and} \quad \beta_2 = 0$
- F-statistic with $q = 2$ restrictions:

$$F = \frac{1}{2}\left(\frac{t_1^2 + t_2^2 - 2\hat{\rho}_{t_1,t_2} t_1 t_2}{1 - \hat{\rho}_{t_1,t_2}^2}\right)$$

  where $\hat{\rho}_{t_1,t_2}^2$ is an estimator of the correlation between the two t-statistics.
- In large samples, under the null hypothesis, the F-statistic has sampling distribution $F_{q,\infty}$.

Joint Hypothesis Tests

F-Test for Two Restrictions:

- The F-statistic is used to test a joint hypothesis about regression coefficients.
- Joint Hypothesis: $H_0 : \beta_1 = 0 \quad \text{and} \quad \beta_2 = 0$
- F-statistic with $q = 2$ restrictions:

$$F = \frac{1}{2}\left(\frac{t_1^2 + t_2^2 - 2\hat{\rho}_{t_1,t_2} t_1 t_2}{1 - \hat{\rho}^2_{t_1,t_2}}\right)$$

  where $\hat{\rho}^2_{t_1,t_2}$ is an estimator of the correlation between the two t-statistics.
- In large samples, under the null hypothesis, the F-statistic has sampling distribution $F_{q,\infty}$.

Joint Hypothesis Tests

F-Test for Two Restrictions:

$$\widehat{TestScore} = \underset{(15.5)}{649.6} - \underset{(0.48)}{0.29}\ STR - \underset{(0.032)}{0.656}\ PctEL + \underset{(1.59)}{3.87}\ Expn$$

1. $H_0: \ STR = 0$ and $Expn = 0$
2. Compute heteroskedasticity-robust F-statistic: $F^{\text{act}} = 5.43$.
   In large samples, $F \sim F(2, \infty)$.
3. Compute critical $F_{\alpha/2}(2, \infty)$ values:
   $$F_{0.05/2}(2, \infty) \approx 3.00$$
   $$F_{0.01/2}(2, \infty) \approx 4.61$$
4. Conclude: $F^{\text{act}} > F_{\alpha/2}(2, \infty) \implies$ reject H_0.
   At least one of STR or $Expn$ has a statistically significant effect on $TestScore$.

Joint Hypothesis Tests

F-Test for Two Restrictions:

$$\widehat{TestScore} = \underset{(15.5)}{649.6} - \underset{(0.48)}{0.29}\ STR - \underset{(0.032)}{0.656}\ PctEL + \underset{(1.59)}{3.87}\ Expn$$

1. $H_0: \ STR = 0$ and $Expn = 0$
2. Compute heteroskedasticity-robust F-statistic: $F^{\text{act}} = 5.43$.
   In large samples, $F \sim F(2, \infty)$.
3. Compute critical $F_{\alpha/2}(2, \infty)$ values:

$$F_{0.05/2}(2, \infty) \approx 3.00$$

$$F_{0.01/2}(2, \infty) \approx 4.61$$

4. Conclude: $F^{\text{act}} > F_{\alpha/2}(2, \infty) \implies$ reject H_0.
   At least one of STR or $Expn$ has a statistically significant effect on $TestScore$.

Joint Hypothesis Tests

F-Test for Two Restrictions:

$$\widehat{TestScore} = \underset{(15.5)}{649.6} - \underset{(0.48)}{0.29}\ STR - \underset{(0.032)}{0.656}\ PctEL + \underset{(1.59)}{3.87}\ Expn$$

1. $H_0: \ STR = 0$ and $Expn = 0$
2. Compute heteroskedasticity-robust F-statistic: $F^{\text{act}} = 5.43.$

   In large samples, $F \sim F(2, \infty)$.
3. Compute critical $F_{\alpha/2}(2, \infty)$ values:

$$F_{0.05/2}(2, \infty) \approx 3.00$$
$$F_{0.01/2}(2, \infty) \approx 4.61$$

4. Conclude: $F^{\text{act}} > F_{\alpha/2}(2, \infty) \implies$ reject H_0.

   At least one of STR or $Expn$ has a statistically significant effect on $TestScore$.

Joint Hypothesis Tests

F-Test for Two Restrictions:

$$\widehat{TestScore} = \underset{(15.5)}{649.6} - \underset{(0.48)}{0.29}\ STR - \underset{(0.032)}{0.656}\ PctEL + \underset{(1.59)}{3.87}\ Expn$$

1. $H_0:\ STR = 0$ and $Expn = 0$
2. Compute heteroskedasticity-robust F-statistic: $F^{\text{act}} = 5.43$.
   In large samples, $F \sim F(2, \infty)$.
3. Compute critical $F_{\alpha/2}(2, \infty)$ values:
$$F_{0.05/2}(2, \infty) \approx 3.00$$
$$F_{0.01/2}(2, \infty) \approx 4.61$$
4. Conclude: $F^{\text{act}} > F_{\alpha/2}(2, \infty) \implies$ reject H_0.
   At least one of STR or $Expn$ has a statistically significant effect on $TestScore$.

Confidence Sets

Confidence Intervals for a Single Coefficient

- A two-sided confidence interval for coefficient β_j, for given significance level α, is an interval that contains the true value of β_j with probability $(1-\alpha)\%$.

$$\hat{\beta}_j \pm t_{\alpha/2,dof}\,\text{SE}(\hat{\beta}_j)$$

where $t_{\alpha/2,dof}$ is the critical t-value for significance α and degrees of freedom dof.

- For large dof and $\alpha = 0.10$, $t_{\alpha/2,dof} \to 1.64$.
- For large dof and $\alpha = 0.05$, $t_{\alpha/2,dof} \to 1.96$.
- For large dof and $\alpha = 0.01$, $t_{\alpha/2,dof} \to 2.58$.

Confidence Intervals for a Single Coefficient

- A two-sided confidence interval for coefficient β_j, for given significance level α, is an interval that contains the true value of β_j with probability $(1-\alpha)\%$.

$$\hat{\beta}_j \pm t_{\alpha/2,dof}\,\text{SE}(\hat{\beta}_j)$$

  where $t_{\alpha/2,dof}$ is the critical t-value for significance α and degrees of freedom dof.
- For large dof and $\alpha = 0.10$, $t_{\alpha/2,dof} \rightarrow 1.64$.
- For large dof and $\alpha = 0.05$, $t_{\alpha/2,dof} \rightarrow 1.96$.
- For large dof and $\alpha = 0.01$, $t_{\alpha/2,dof} \rightarrow 2.58$.

Confidence Intervals for a Single Coefficient

- A two-sided confidence interval for coefficient β_j, for given significance level α, is an interval that contains the true value of β_j with probability $(1-\alpha)\%$.

$$\hat{\beta}_j \pm t_{\alpha/2,dof}\,\mathrm{SE}(\hat{\beta}_j)$$

  where $t_{\alpha/2,dof}$ is the critical t-value for significance α and degrees of freedom dof.
- For large dof and $\alpha = 0.10$, $t_{\alpha/2,dof} \to 1.64$.
- For large dof and $\alpha = 0.05$, $t_{\alpha/2,dof} \to 1.96$.
- For large dof and $\alpha = 0.01$, $t_{\alpha/2,dof} \to 2.58$.

Confidence Intervals for a Single Coefficient

- A two-sided confidence interval for coefficient β_j, for given significance level α, is an interval that contains the true value of β_j with probability $(1-\alpha)\%$.

$$\hat{\beta}_j \pm t_{\alpha/2,dof}\,\mathsf{SE}(\hat{\beta}_j)$$

  where $t_{\alpha/2,dof}$ is the critical t-value for significance α and degrees of freedom dof.
- For large dof and $\alpha = 0.10$, $t_{\alpha/2,dof} \to 1.64$.
- For large dof and $\alpha = 0.05$, $t_{\alpha/2,dof} \to 1.96$.
- For large dof and $\alpha = 0.01$, $t_{\alpha/2,dof} \to 2.58$.

Confidence Sets for Multiple Coefficients

Joint Hypotheses in Matrix Notation

- A confidence set is the generalization to two or more coefficients of a confidence interval for a single coefficient.

$$\mathbf{R}\beta = \mathbf{r}$$

  where $\mathbf{R}$ is a $q \times (k+1)$ non-random matrix with full row rank and $\mathbf{r}$ is a non-random $q \times 1$ vector. The number of rows q of $\mathbf{R}$ is the number of restrictions under the null hypothesis.

- Example: Let $k = 2$ and $H_0 : \beta_1 + \beta_2 = 0$. This can be represented in matrix form with $\beta = [\beta_0\ \beta_1\ \beta_0]'$, $\mathbf{R} = [0\ 1\ 1]$, where $r = 0$ and $q = 1$.

- **Asymptotic Distribution of the F-Statistic:**

  The heteroskedasticity-robust F-statistic testing the joint hypothesis

$$F = (\mathbf{R}\hat{\beta} - \mathbf{r})'[(\mathbf{R}\hat{\Sigma}\mathbf{R}']^{-1} - \mathbf{r})'](\mathbf{R}\hat{\beta} - \mathbf{r})/q$$

  Under the null hypothesis $F \xrightarrow{d} F_{q,\infty}$.

Confidence Sets for Multiple Coefficients

Joint Hypotheses in Matrix Notation

- A confidence set is the generalization to two or more coefficients of a confidence interval for a single coefficient.

$$\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$$

  where $\mathbf{R}$ is a $q \times (k+1)$ non-random matrix with full row rank and $\mathbf{r}$ is a non-random $q \times 1$ vector. The number of rows q of $\mathbf{R}$ is the number of restrictions under the null hypothesis.
- Example: Let $k = 2$ and $H_0 : \beta_1 + \beta_2 = 0$. This can be represented in matrix form with $\boldsymbol{\beta} = [\beta_0\ \beta_1\ \beta_0]'$, $\mathbf{R} = [0\ 1\ 1]$, where $r = 0$ and $q = 1$.
- Asymptotic Distribution of the F-Statistic:

  The heteroskedasticity-robust F-statistic testing the joint hypothesis

$$F = (\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})'[(\mathbf{R}\hat{\Sigma}\mathbf{R}']^{-1} - \mathbf{r})'](\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})/q$$

  Under the null hypothesis $F \xrightarrow{d} F_{q,\infty}$.

Confidence Sets for Multiple Coefficients

Joint Hypotheses in Matrix Notation

- A confidence set is the generalization to two or more coefficients of a confidence interval for a single coefficient.

$$\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$$

  where $\mathbf{R}$ is a $q \times (k+1)$ non-random matrix with full row rank and $\mathbf{r}$ is a non-random $q \times 1$ vector. The number of rows q of $\mathbf{R}$ is the number of restrictions under the null hypothesis.

- Example: Let $k = 2$ and $H_0 : \beta_1 + \beta_2 = 0$. This can be represented in matrix form with $\boldsymbol{\beta} = [\beta_0\ \beta_1\ \beta_0]'$, $\mathbf{R} = [0\ 1\ 1]$, where $r = 0$ and $q = 1$.

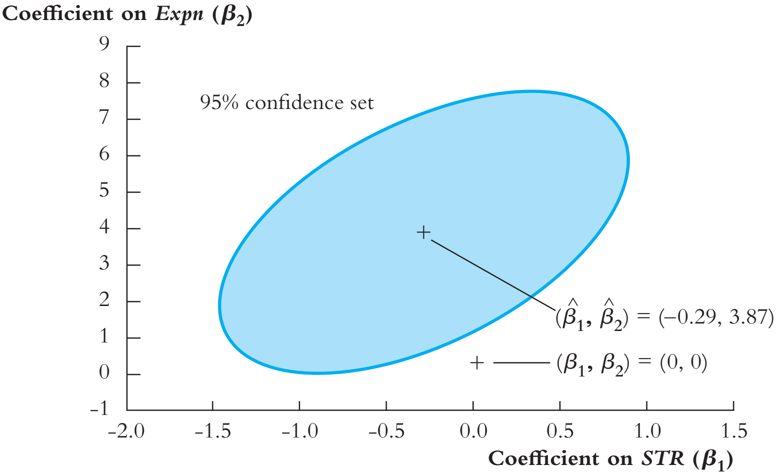
- **Asymptotic Distribution of the F-Statistic:**

  The heteroskedasticity-robust F-statistic testing the joint hypothesis

$$F = (\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})'[(\mathbf{R}\hat{\boldsymbol{\Sigma}}\mathbf{R}']^{-1} - \mathbf{r})'](\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{r})/q$$

  Under the null hypothesis $F \xrightarrow{d} F_{q,\infty}$.

Confidence Sets for Multiple Coefficients



95% **Confidence Set for Coefficients on** $STR\,(\beta_1)$ **and** $Expn\,(\beta_2)$**.** The ellipse contains the pairs of values of β_1 and β_2 that cannot be rejected using the F-statistic at the 5% significance level. The point $(0, 0)$ is not contained in the confidence set, so the null hypothesis $H_0:\ \beta_1 = 0$ and $\beta_2 = 0$ is rejected at the 5% significance level.

Model Specification

Model Specification

Choose a regression specification

- **Starting point:**

  Think through the possible sources of omitted variable bias.

- **Base specification:**

  Include the variables of primary interest and the control variables suggested by economic theory.

- **Alternative specifications:**

  If the estimates of the coefficients of interest are numerically similar across the alternative specifications, then this provides evidence that the estimates of the base specification are reliable.

Model Specification

Choose a regression specification

- **Starting point:**
  Think through the possible sources of omitted variable bias.
- **Base specification:**
  Include the variables of primary interest and the control variables suggested by economic theory.
- **Alternative specifications:**
  If the estimates of the coefficients of interest are numerically similar across the alternative specifications, then this provides evidence that the estimates of the base specification are reliable.

Model Specification

Choose a regression specification

- **Starting point:**

  Think through the possible sources of omitted variable bias.

- **Base specification:**

  Include the variables of primary interest and the control variables suggested by economic theory.

- **Alternative specifications:**

  If the estimates of the coefficients of interest are numerically similar across the alternative specifications, then this provides evidence that the estimates of the base specification are reliable.

Interpreting R-Squared

Interpreting R^2 and Adjusted-R^2

- An increase in the R^2 or Adjusted-R^2 does not necessarily mean that an added variable is statistically significant. Perform a hypothesis test using the t-statistic.
- A high R^2 or Adjusted-R^2 does not mean that the regressors are a true cause of the dependent variable.
- A high R^2 or Adjusted-R^2 does not mean that there is no omitted variable bias.
- A high R^2 or Adjusted-R^2 does not mean that you have the most appropriate set of regressors — A low R^2 or Adjusted-R^2 does not mean that you have an inappropriate set of regressors.

Interpreting R-Squared

Interpreting R^2 and Adjusted-R^2

- An increase in the R^2 or Adjusted-R^2 does not necessarily mean that an added variable is statistically significant. Perform a hypothesis test using the t-statistic.
- A high R^2 or Adjusted-R^2 does not mean that the regressors are a true cause of the dependent variable.
- A high R^2 or Adjusted-R^2 does not mean that there is no omitted variable bias.
- A high R^2 or Adjusted-R^2 does not mean that you have the most appropriate set of regressors – A low R^2 or Adjusted-R^2 does not mean that you have an inappropriate set of regressors.

Interpreting R-Squared

Interpreting R^2 and Adjusted-R^2

- An increase in the R^2 or Adjusted-R^2 does not necessarily mean that an added variable is statistically significant. Perform a hypothesis test using the t-statistic.
- A high R^2 or Adjusted-R^2 does not mean that the regressors are a true cause of the dependent variable.
- A high R^2 or Adjusted-R^2 does not mean that there is no omitted variable bias.
- A high R^2 or Adjusted-R^2 does not mean that you have the most appropriate set of regressors — A low R^2 or Adjusted-R^2 does not mean that you have an inappropriate set of regressors.

Interpreting R-Squared

Interpreting R^2 and Adjusted-R^2

- An increase in the R^2 or Adjusted-R^2 does not necessarily mean that an added variable is statistically significant. Perform a hypothesis test using the t-statistic.
- A high R^2 or Adjusted-R^2 does not mean that the regressors are a true cause of the dependent variable.
- A high R^2 or Adjusted-R^2 does not mean that there is no omitted variable bias.
- A high R^2 or Adjusted-R^2 does not mean that you have the most appropriate set of regressors — A low R^2 or Adjusted-R^2 does not mean that you have an inappropriate set of regressors.

California School Districts

California School Districts

Selecting control variables

- Consider three variables that control for background characteristics of the students that could affect test scores: the fraction of students who are still learning English, the percentage of students who are eligible to receive a subsidized or free lunch at school, and the percentage of students in the district whose families qualify for a California income assistance program.
- The last two variables are different measures of the fraction of economically disadvantaged children in the district. Eligibility for the income assistance program requires a stricter threshold than the subsidized lunch program.
- Theory and expert judgment do not tell us which of these two variables to use to control for determinants of test scores related to economic background.

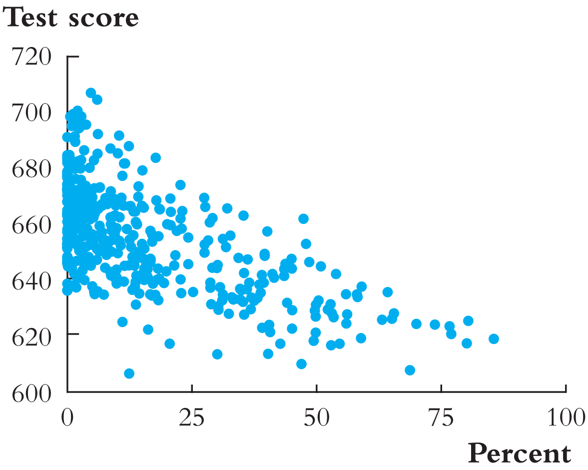
Selecting control variables

- Consider three variables that control for background characteristics of the students that could affect test scores: the fraction of students who are still learning English, the percentage of students who are eligible to receive a subsidized or free lunch at school, and the percentage of students in the district whose families qualify for a California income assistance program.
- The last two variables are different measures of the fraction of economically disadvantaged children in the district. Eligibility for the income assistance program requires a stricter threshold than the subsidized lunch program.
- Theory and expert judgment do not tell us which of these two variables to use to control for determinants of test scores related to economic background.

Selecting control variables

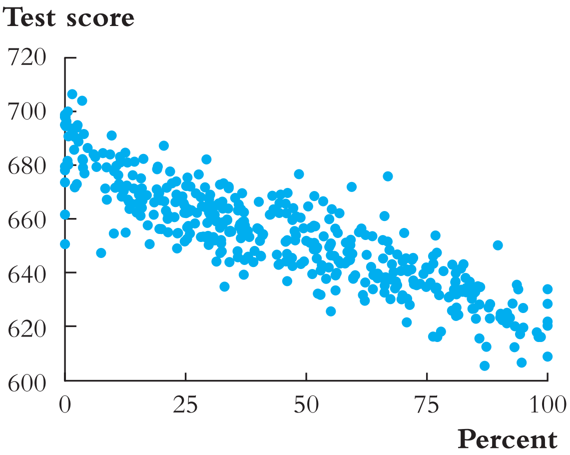
- Consider three variables that control for background characteristics of the students that could affect test scores: the fraction of students who are still learning English, the percentage of students who are eligible to receive a subsidized or free lunch at school, and the percentage of students in the district whose families qualify for a California income assistance program.
- The last two variables are different measures of the fraction of economically disadvantaged children in the district. Eligibility for the income assistance program requires a stricter threshold than the subsidized lunch program.
- Theory and expert judgment do not tell us which of these two variables to use to control for determinants of test scores related to economic background.

Test Scores vs Student Characteristic



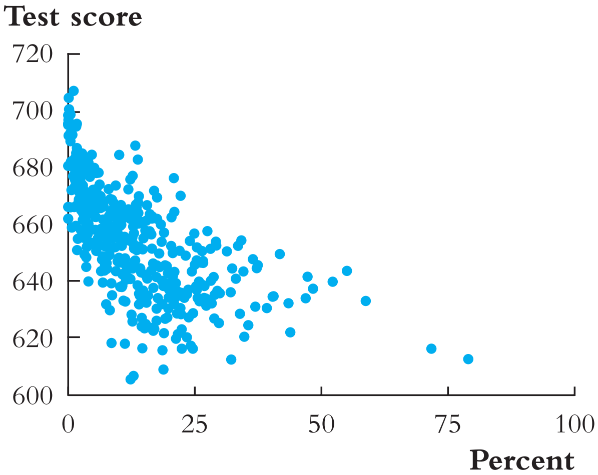
(a) Percentage of English learners

Test Scores vs Student Characteristic



(b) Percentage eligible for subsidized lunch

Test Scores vs Student Characteristic



(c) Percentage qualifying for income assistance

Regression of Test Scores on Student-Teacher Ratio: Selecting Controls

Dependent variable: average test score in the district.

Regressor	(1)	(2)	(3)	(4)	(5)
Student–teacher ratio (X_1)	-2.28 (0.52) $[-3.30, -1.26]$	-1.10 (0.43) $[-1.95, -0.25]$	-1.00 (0.27) $[-1.53, -0.47]$	-1.31 (0.34) $[-1.97, -0.64]$	-1.01 (0.27) $[-1.54, -0.49]$
Control variables					
Percentage English learners (X_2)		-0.650 (0.031)	-0.122 (0.033)	-0.488 (0.030)	-0.130 (0.036)
Percentage eligible for subsidized lunch (X_3)			-0.547 (0.024)		-0.529 (0.038)
Percentage qualifying for income assistance (X_4)				-0.790 (0.068)	0.048 (0.059)
Intercept	698.9 (10.4)	686.0 (8.7)	700.2 (5.6)	698.0 (6.9)	700.4 (5.5)
Summary Statistics					
SER	18.58	14.46	9.08	11.65	9.08
$\overline{R}^2$	0.049	0.424	0.773	0.626	0.773
n	420	420	420	420	420

Selecting control variables

- To mitigate a potential omitted variable bias, we augment the regression by including variables that control for various student characteristics.
- These controls cut the estimated effect of the student-teacher ratio on test scores approximately in half. This estimated effect is not very sensitive to which specific control variables are included in the regression. In all cases, the hypothesis that the coefficient on the student-teacher ratio is 0 can be rejected at the 5% level.
- The student characteristic variables are potent predictors of test scores. The student-teacher ratio alone explains only a small fraction of the variation in test scores: The $\bar{R}^2$ in column (1) is 0.049. The $\bar{R}^2$ jumps, however, when the student characteristic variables are added. Districts with many English learners and districts with many poor children have lower test scores.
- The percentage qualifying for income assistance appears to be redundant. As reported in regression (5), adding it to regression (3) has a negligible effect on the estimated coefficient on the student-teacher ratio or its standard error.

Selecting control variables

- To mitigate a potential omitted variable bias, we augment the regression by including variables that control for various student characteristics.
- These controls cut the estimated effect of the student-teacher ratio on test scores approximately in half. This estimated effect is not very sensitive to which specific control variables are included in the regression. In all cases, the hypothesis that the coefficient on the student-teacher ratio is 0 can be rejected at the 5% level.
- The student characteristic variables are potent predictors of test scores. The student-teacher ratio alone explains only a small fraction of the variation in test scores: The $\bar{R}^2$ in column (1) is 0.049. The $\bar{R}^2$ jumps, however, when the student characteristic variables are added. Districts with many English learners and districts with many poor children have lower test scores.
- The percentage qualifying for income assistance appears to be redundant. As reported in regression (5), adding it to regression (3) has a negligible effect on the estimated coefficient on the student-teacher ratio or its standard error.

California School Districts

Selecting control variables

- To mitigate a potential omitted variable bias, we augment the regression by including variables that control for various student characteristics.
- These controls cut the estimated effect of the student-teacher ratio on test scores approximately in half. This estimated effect is not very sensitive to which specific control variables are included in the regression. In all cases, the hypothesis that the coefficient on the student-teacher ratio is 0 can be rejected at the 5% level.
- The student characteristic variables are potent predictors of test scores. The student-teacher ratio alone explains only a small fraction of the variation in test scores: The $\bar{R}^2$ in column (1) is 0.049. The $\bar{R}^2$ jumps, however, when the student characteristic variables are added. Districts with many English learners and districts with many poor children have lower test scores.
- The percentage qualifying for income assistance appears to be redundant. As reported in regression (5), adding it to regression (3) has a negligible effect on the estimated coefficient on the student-teacher ratio or its standard error.

California School Districts

Selecting control variables

- To mitigate a potential omitted variable bias, we augment the regression by including variables that control for various student characteristics.
- These controls cut the estimated effect of the student-teacher ratio on test scores approximately in half. This estimated effect is not very sensitive to which specific control variables are included in the regression. In all cases, the hypothesis that the coefficient on the student-teacher ratio is 0 can be rejected at the 5% level.
- The student characteristic variables are potent predictors of test scores. The student-teacher ratio alone explains only a small fraction of the variation in test scores: The $\bar{R}^2$ in column (1) is 0.049. The $\bar{R}^2$ jumps, however, when the student characteristic variables are added. Districts with many English learners and districts with many poor children have lower test scores.
- The percentage qualifying for income assistance appears to be redundant. As reported in regression (5), adding it to regression (3) has a negligible effect on the estimated coefficient on the student-teacher ratio or its standard error.

Summary

Summary

- Hypotheses involving more than one restriction on the coefficients are called joint hypotheses. Joint hypotheses can be tested using an F-statistic.
- Regression specification proceeds by first determining a base specification chosen to address concern about omitted variable bias. The base specification can be modified by including additional regressors that control for other potential sources of omitted variable bias.
- Choosing the specification with the highest R^2 can lead to regression models that do not estimate the causal effect of interest.
- A study is internally valid if the statistical inferences about causal effects are valid for the population being studied. A study is externally valid if its inferences and conclusions can be generalized from the population and setting studied to other populations and settings.
- Regressors and error terms may be correlated when there are omitted variables, an incorrect functional form is used, one or more of the regressors are measured with error, the sample is not randomly selected, or there is simultaneous causality between the regressors and dependent variables.
- When regression models are used for prediction, it is not necessary for regression coefficients to be unbiased estimates of causal effects, but the model must be externally valid.

Summary

- Hypotheses involving more than one restriction on the coefficients are called joint hypotheses. Joint hypotheses can be tested using an F-statistic.
- Regression specification proceeds by first determining a base specification chosen to address concern about omitted variable bias. The base specification can be modified by including additional regressors that control for other potential sources of omitted variable bias.
- Choosing the specification with the highest R^2 can lead to regression models that do not estimate the causal effect of interest.
- A study is internally valid if the statistical inferences about causal effects are valid for the population being studied. A study is externally valid if its inferences and conclusions can be generalized from the population and setting studied to other populations and settings.
- Regressors and error terms may be correlated when there are omitted variables, an incorrect functional form is used, one or more of the regressors are measured with error, the sample is not randomly selected, or there is simultaneous causality between the regressors and dependent variables.
- When regression models are used for prediction, it is not necessary for regression coefficients to be unbiased estimates of causal effects, but the model must be externally valid.

Summary

- Hypotheses involving more than one restriction on the coefficients are called joint hypotheses. Joint hypotheses can be tested using an F-statistic.
- Regression specification proceeds by first determining a base specification chosen to address concern about omitted variable bias. The base specification can be modified by including additional regressors that control for other potential sources of omitted variable bias.
- Choosing the specification with the highest R^2 can lead to regression models that do not estimate the causal effect of interest.
- A study is internally valid if the statistical inferences about causal effects are valid for the population being studied. A study is externally valid if its inferences and conclusions can be generalized from the population and setting studied to other populations and settings.
- Regressors and error terms may be correlated when there are omitted variables, an incorrect functional form is used, one or more of the regressors are measured with error, the sample is not randomly selected, or there is simultaneous causality between the regressors and dependent variables.
- When regression models are used for prediction, it is not necessary for regression coefficients to be unbiased estimates of causal effects, but the model must be externally valid.

Summary

- Hypotheses involving more than one restriction on the coefficients are called joint hypotheses. Joint hypotheses can be tested using an F-statistic.
- Regression specification proceeds by first determining a base specification chosen to address concern about omitted variable bias. The base specification can be modified by including additional regressors that control for other potential sources of omitted variable bias.
- Choosing the specification with the highest R^2 can lead to regression models that do not estimate the causal effect of interest.
- A study is internally valid if the statistical inferences about causal effects are valid for the population being studied. A study is externally valid if its inferences and conclusions can be generalized from the population and setting studied to other populations and settings.
- Regressors and error terms may be correlated when there are omitted variables, an incorrect functional form is used, one or more of the regressors are measured with error, the sample is not randomly selected, or there is simultaneous causality between the regressors and dependent variables.
- When regression models are used for prediction, it is not necessary for regression coefficients to be unbiased estimates of causal effects, but the model must be externally valid.

Summary

- Hypotheses involving more than one restriction on the coefficients are called joint hypotheses. Joint hypotheses can be tested using an F-statistic.
- Regression specification proceeds by first determining a base specification chosen to address concern about omitted variable bias. The base specification can be modified by including additional regressors that control for other potential sources of omitted variable bias.
- Choosing the specification with the highest R^2 can lead to regression models that do not estimate the causal effect of interest.
- A study is internally valid if the statistical inferences about causal effects are valid for the population being studied. A study is externally valid if its inferences and conclusions can be generalized from the population and setting studied to other populations and settings.
- Regressors and error terms may be correlated when there are omitted variables, an incorrect functional form is used, one or more of the regressors are measured with error, the sample is not randomly selected, or there is simultaneous causality between the regressors and dependent variables.
- When regression models are used for prediction, it is not necessary for regression coefficients to be unbiased estimates of causal effects, but the model must be externally valid.

Summary

- Hypotheses involving more than one restriction on the coefficients are called joint hypotheses. Joint hypotheses can be tested using an F-statistic.
- Regression specification proceeds by first determining a base specification chosen to address concern about omitted variable bias. The base specification can be modified by including additional regressors that control for other potential sources of omitted variable bias.
- Choosing the specification with the highest R^2 can lead to regression models that do not estimate the causal effect of interest.
- A study is internally valid if the statistical inferences about causal effects are valid for the population being studied. A study is externally valid if its inferences and conclusions can be generalized from the population and setting studied to other populations and settings.
- Regressors and error terms may be correlated when there are omitted variables, an incorrect functional form is used, one or more of the regressors are measured with error, the sample is not randomly selected, or there is simultaneous causality between the regressors and dependent variables.
- When regression models are used for prediction, it is not necessary for regression coefficients to be unbiased estimates of causal effects, but the model must be externally valid.

Problems & Applications

Dependent variable: average hourly earnings (*AHE*).

Regressor	(1)	(2)	(3)
College (X_1)	10.47 (0.29)	10.44 (0.29)	10.42 (0.29)
Female (X_2)	−4.69 (0.29)	−4.56 (0.29)	−4.57 (0.29)
Age (X_3)		0.61 (0.05)	0.61 (0.05)
Northeast (X_4)			0.74 (0.47)
Midwest (X_5)			−1.54 (0.40)
South (X_6)			−0.44 (0.37)
Intercept	18.15 (0.19)	0.11 (1.46)	0.33 (1.47)
Summary Statistics and Joint Tests			
F-statistic testing regional effects $= 0$			9.32
SER	12.15	12.03	12.01
R^2	0.165	0.182	0.185
n	7178	7178	7178

Problems and Applications

Stock & Watson, Introduction (4th), Chapter 7, Exercise 3.

Using the regression results in column (2):

1. Is age an important determinant of earnings? Use an appropriate statistical test and/or confidence interval to explain your answer.
2. Sally is a 29-year-old female college graduate. Betsy is a 34-year-old female college graduate. Construct a 95% confidence interval for the expected difference between their earnings.

Stock & Watson, Introduction (4th), Chapter 7, Exercise 3.

Using the regression results in column (2):

1. Is age an important determinant of earnings? Use an appropriate statistical test and/or confidence interval to explain your answer.
2. Sally is a 29-year-old female college graduate. Betsy is a 34-year-old female college graduate. Construct a 95% confidence interval for the expected difference between their earnings.

Problems and Applications

Stock & Watson, Introduction (4th), Chapter 7, Exercise 5.

The regression shown in column (2) was estimated again, this time using data from 1992 (4000 observations selected at random from the March 1993 Current Population Survey, converted into 2015 dollars using the Consumer Price Index). The results are

$$\widehat{AHE} = \underset{(1.65)}{-1.3} + \underset{(0.34)}{8.94}\, College - \underset{(0.30)}{4.38}\, Female + \underset{(0.05)}{0.67}\, Age,$$

$$\bar{R}^2 = 0.21, \quad SER = 9.88$$

Comparing this regression to the regression for 2015 shown in column (2), was there a statistically significant change in the coefficient on College?