# Lecture 3 The Multiple Regression Model: Inference

LECTURED BY WENYA WANG

# Outline

1. Sampling distributions of the OLS estimators
2. Inference on a single population estimator
   a) t test: one-sided vs two-sided
   b) p-value for t statistic
   c) Confidence intervals
3. Inference on a linear combination of parameters
   a) Unrestricted vs restricted models
   b) F statistic
4. Convention in reporting regression results

# Aside: the matrix representation

Population regression model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \mu$$

In a random sample with $\{y_i\}_{i=1}^{n}, \{x_{i1}, x_{i2}\}_{i=1}^{n}$

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ \vdots & \vdots \\ x_{n1} & x_{n2} \end{pmatrix} \times \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix}$$

Or using the matrix representation $Y = X\vec{\beta} + \vec{\mu}$

# Aside: the matrix representation

The matrix representation

- Easier to show the sampling distributions of OLS estimator, and SSR later.

For simplicity, use $\beta$ to refer to $\vec{\beta}$, similarly for $\mu$

The OLS estimator is now

$$\begin{aligned}\hat{\beta} &= (X'X)^{-1}X'Y \\ &= (X'X)^{-1}X'(X\beta + \mu) \\ &= \beta + (X'X)^{-1}X'\mu \\ &= \beta + \sum_{i=1}^{n} f_i(X)\mu_i\end{aligned}$$

# Aside: the matrix representation

The sum of squared residuals is now

$$SSR = (Y - X\hat{\beta})'(Y - X\hat{\beta})$$
$$= (\mu - X(X'X)^{-1}X'\mu)'(\mu - X(X'X)^{-1}X'\mu)$$
$$= \mu'(I - X(X'X)^{-1}X')\mu$$
$$= \sum_{i=1}^{n} g_i(X)\mu_i^2$$

Note the unbiased estimator of $\sigma^2$ is $\hat{\sigma}^2 = \sum_{i=1}^{n} \hat{\mu}_i{}^2/(n-k-1) = SSR/(n-k-1)$

Now we're ready to study inference with the t- and F-statistics

# Sampling distributions of the OLS estimators

Last class: expected values and variances of the OLS estimators

- the distribution of the OLS estimators unknown
- hinder further analysis in statistical inference, especially for a small sample

Thus, we impose the 6th assumption: normality assumption

**Assumption MLR.6** Normality

The population error $u$ is *independent* of the explanatory variables $x_1, x_2, \ldots, x_k$ and is normally distributed with zero mean and variance $\sigma^2$: $u \sim \text{Normal}(0,\sigma^2)$.

Or succinctly, we can write

$$y|\mathbf{x} \sim \text{Normal}(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k, \sigma^2),$$

# Sampling distributions of the OLS estimators

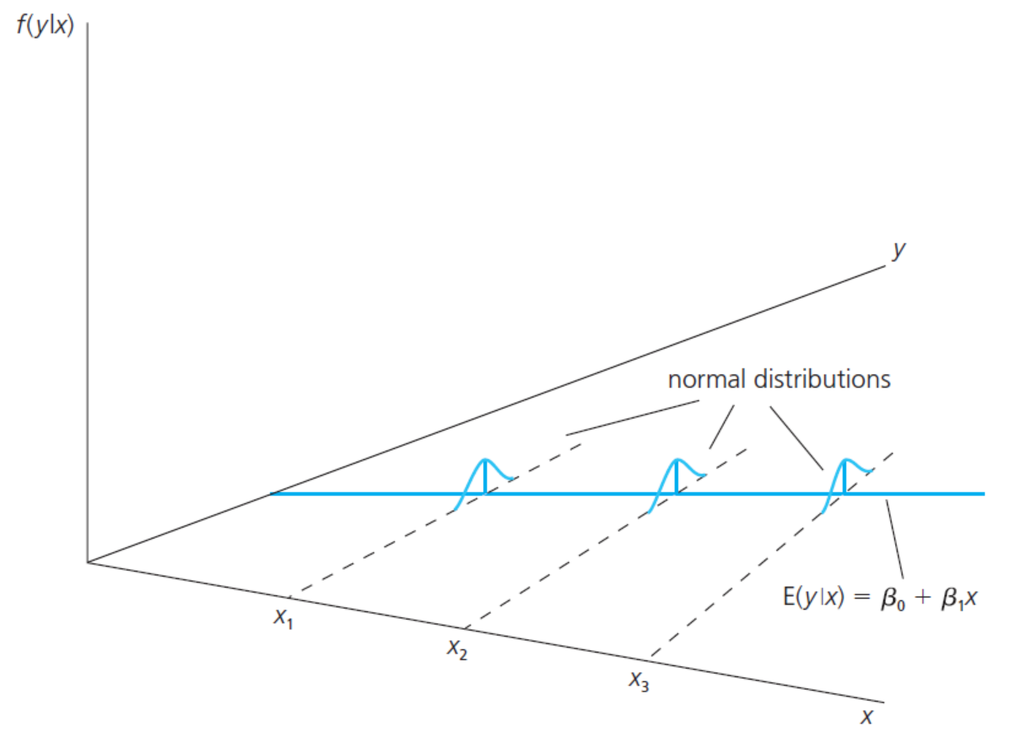
Note assumption MLR.6 implies assumptions MLR. 4 (zero conditional mean) and MRL. 5 (homoscedasticity)

Assumptions MLR.1 – 6: classical linear model (CLM) assumptions

A model that satisfies MLR.1 – 6 : classical linear model

# Normality condition (graphically)

**FIGURE 4.1** The homoskedastic normal distribution with a single explanatory variable.



# Checking normality in data

How could we check normality in data?

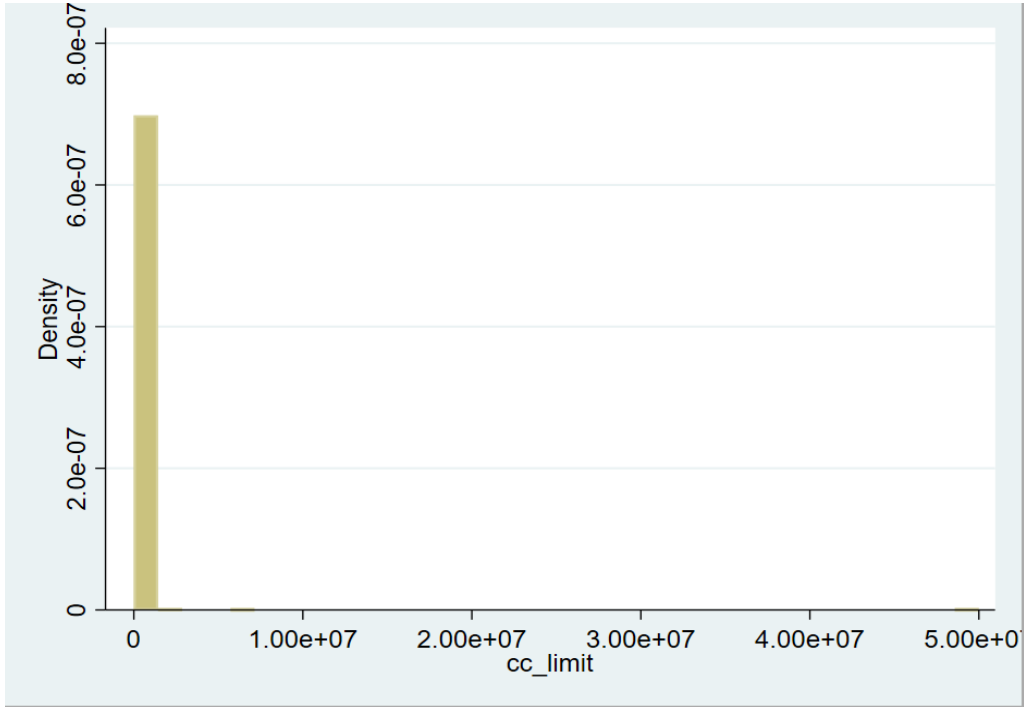
- Note we cannot directly see $\mu$, but we see y, which also follows a normal distribution given X
- Good habit to check normality before running regression

Using the SCF example,

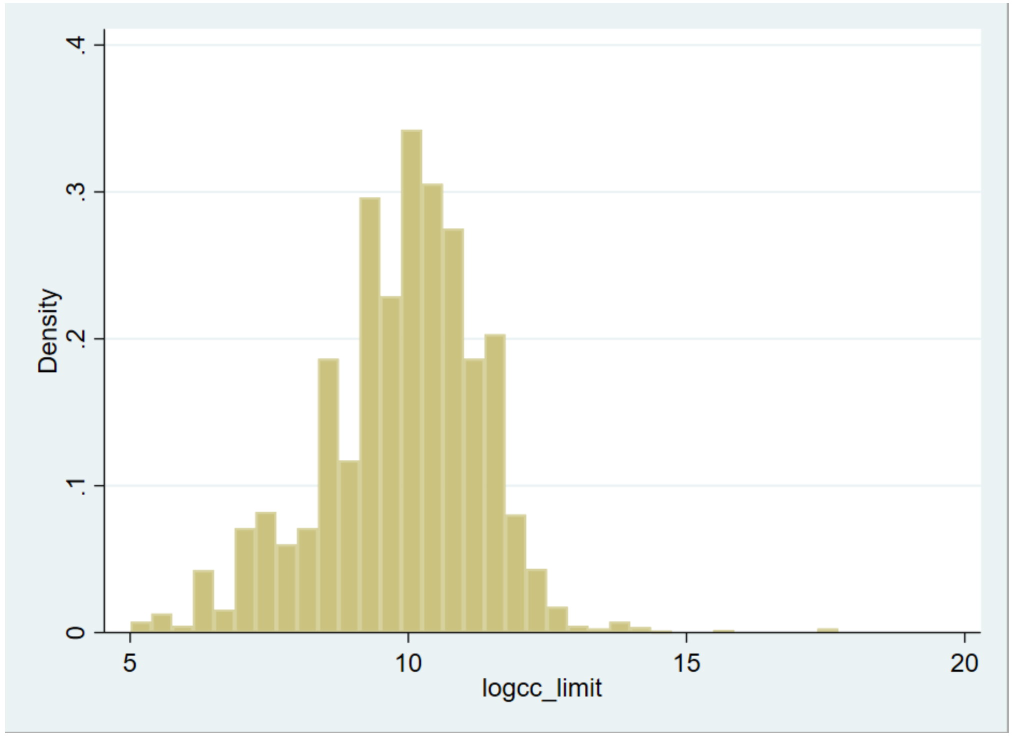
let’s check the histogram of credit card limit

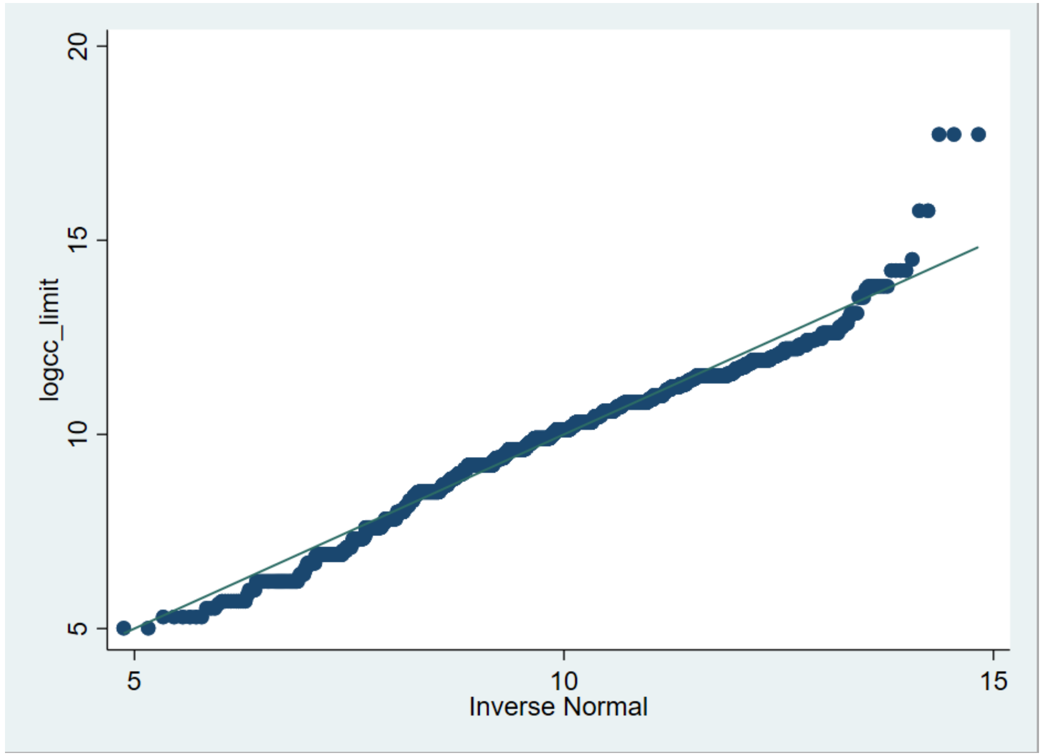
=> hardly normal, but the picture changes

dramatically when we take the log of cc_limit



# QQ-plot for testing normality




A more precise test of normality is to check the qq-plot

# Sampling distributions of OLS estimators

**THEOREM 4.1**

**NORMAL SAMPLING DISTRIBUTIONS**

Under the CLM assumptions MLR.1 through MLR.6, conditional on the sample values of the independent variables,

$$\hat{\beta}_j \sim \text{Normal}[\beta_j, \text{Var}(\hat{\beta}_j)], \quad [4.1]$$

where $\text{Var}(\hat{\beta}_j)$ was given in Chapter 3 [equation (3.51)]. Therefore,

$$(\hat{\beta}_j - \beta_j)/\text{sd}(\hat{\beta}_j) \sim \text{Normal}(0,1).$$

1. expected value and variance of $\hat{\beta}_j$ derived from the last lecture

2. normality holds since $\hat{\beta}_j = \beta_j + \sum_i f_i(X)\mu_i$ (i.e., a linear combination of normally distributed $\mu_i$ follows a normal distribution)

# t statistic: a single $\beta_j$

Recall from the last class, the population parameter $\sigma^2$ not observable in

$$\mathrm{Var}(\hat{\beta}_j) = \frac{\sigma^2}{\mathrm{SST}_j(1 - R_j^2)},$$

Has to replace $sd(\hat{\beta}_j)$ by $se(\hat{\beta}_j)$, and the t statistic is

$$t_{\hat{\beta}_j} \equiv \hat{\beta}_j / \mathrm{se}(\hat{\beta}_j).$$

Why we call it t statistic?

# t distribution for the standardized estimators

**THEOREM 4.2**

**$t$ DISTRIBUTION FOR THE STANDARDIZED ESTIMATORS**

Under the CLM assumptions MLR.1 through MLR.6,

$$(\hat{\beta}_j - \beta_j)/\mathrm{se}(\hat{\beta}_j) \sim t_{n-k-1} = t_{df}, \qquad [4.3]$$

where $k + 1$ is the number of unknown parameters in the population model $y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + u$ ($k$ slope parameters and the intercept $\beta_0$) and $n - k - 1$ is the degrees of freedom ($df$).

# t distribution for the standardized estimators

$$\frac{(\hat{\beta}_j - \beta_j) \,/\, sd(\hat{\beta}_j)}{((n-k-1)\hat{\sigma}^2/\sigma^2)^{0.5}}$$

$(\hat{\beta}_j - \beta_j) \,/\, sd(\hat{\beta}_j)$ $\Rightarrow$ Standard normal distribution N(0,1)

$((n-k-1)\hat{\sigma}^2/\sigma^2)^{0.5}$ $\Rightarrow$ Denominator squared:

Chi-squared distribution $\chi^2_{n-k-1}$

t distribution

# Hypothesis testing

Null hypothesis concerns whether there is a statistical relation b/w $x_j$ and y, ceteris paribus, i.e.,

$$H_0: \beta_j = 0,$$

This forms the null hypothesis

Naiive thinking: check if $\hat{\beta}_j = 0$

- Even the population parameter $\beta_j = 0$ , we could get $\hat{\beta}_j \neq 0$ given a random sample
- A metric that measures how $\hat{\beta}_j$ is away from 0 in *the probabilistic sense*
- which means we could make mistake: nonzero probability that $H_0$ is true but we reject (Type I error)

# Hypothesis testing: type I & II errors

| | Reject | Not reject |
|---|---|---|
| $H_0$ true | Type I error | |
| $H_0$ false | | Type II error |

Significance level: the probability of type I error

- 10%, 5%, 1% most commonly used

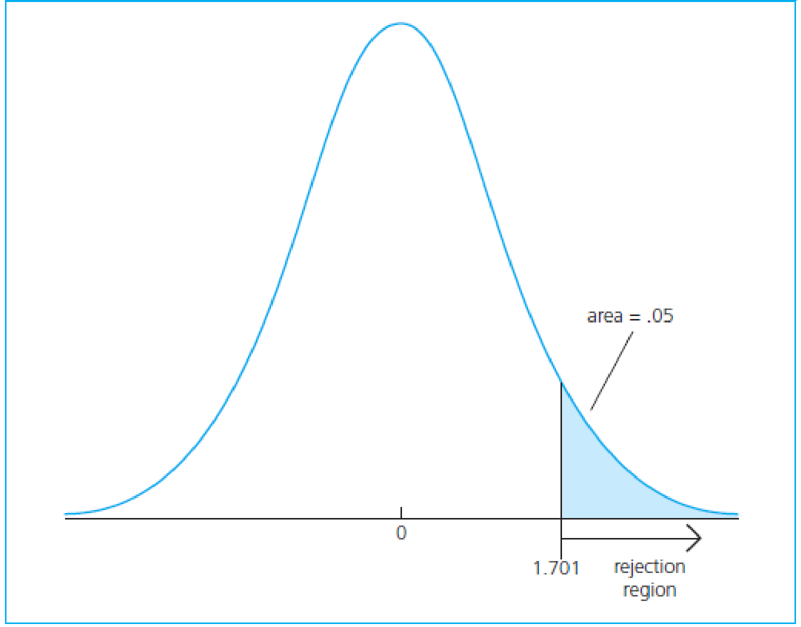
# Testing against one-sided alternatives

One-sided alterative

$$H_0\colon \beta_j \leq 0$$

$$H_1\colon \beta_j > 0$$

FIGURE 4.2 5% rejection rule for the alternative $H_1\colon \beta_j > 0$ with 28 *df*.




Suppose the null hypothesis is true, our t statistic

$$t_{\hat{\beta}_j} \equiv \hat{\beta}_j / \mathrm{se}(\hat{\beta}_j).$$

should not be too large

- i.e., the probability that we get a random sample with t statistic greater than a critical value
- equals to the designated significance level

Rejection rule $t_{\hat{\beta}_j} > c$

# Example: one-sided testing

Type the following command in Stata

```
ssc install bcuse

bcuse wage1
```

This gives us the wage1 data illustrated on the textbook

Using the data in WAGE1.RAW gives the estimated equation

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

$$n = 526,\ R^2 = .316,$$

where standard errors appear in parentheses below the estimated coefficients.

# Example: one-sided testing

Suppose we are interested in testing

$$H_0: \beta_{exper} = 0 \text{ versus } H_1: \beta_{exper} > 0$$

t statistic=0.0041/0.0017 = 2.41

Degree of freedom for the t distribution=526-3-1=522

- Critical value is 1.645 for 5% significance level,
- and 2.326 for 1% significance level

Conclusion: reject null. $\hat{\beta}_{exper}$ is statistically greater than 0 at the 1% significance level.

**TABLE G.2 Critical Values of the *t* Distribution**

| | Significance Level | | | | |
|---|---|---|---|---|---|
| 1-Tailed: | .10 | .05 | .025 | .01 | .005 |
| 2-Tailed: | .20 | .10 | .05 | .02 | .01 |
| 1 | 3.078 | 6.314 | 12.706 | 31.821 | 63.657 |
| 2 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 |
| 3 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 |
| 4 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 |
| 5 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 |
| 6 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 |
| 7 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 |
| 8 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 |
| 9 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 |
| 10 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 |
| 11 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 |
| 12 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 |
| 13 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 |
| 14 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 |
| 15 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 |
| 16 | 1.337 | 1.746 | 2.120 | 2.583 | 2.921 |
| 17 | 1.333 | 1.740 | 2.110 | 2.567 | 2.898 |
| 18 | 1.330 | 1.734 | 2.101 | 2.552 | 2.878 |
| 19 | 1.328 | 1.729 | 2.093 | 2.539 | 2.861 |
| 20 | 1.325 | 1.725 | 2.086 | 2.528 | 2.845 |
| 21 | 1.323 | 1.721 | 2.080 | 2.518 | 2.831 |
| 22 | 1.321 | 1.717 | 2.074 | 2.508 | 2.819 |
| 23 | 1.319 | 1.714 | 2.069 | 2.500 | 2.807 |
| 24 | 1.318 | 1.711 | 2.064 | 2.492 | 2.797 |
| 25 | 1.316 | 1.708 | 2.060 | 2.485 | 2.787 |
| 26 | 1.315 | 1.706 | 2.056 | 2.479 | 2.779 |
| 27 | 1.314 | 1.703 | 2.052 | 2.473 | 2.771 |
| 28 | 1.313 | 1.701 | 2.048 | 2.467 | 2.763 |
| 29 | 1.311 | 1.699 | 2.045 | 2.462 | 2.756 |
| 30 | 1.310 | 1.697 | 2.042 | 2.457 | 2.750 |
| 40 | 1.303 | 1.684 | 2.021 | 2.423 | 2.704 |
| 60 | 1.296 | 1.671 | 2.000 | 2.390 | 2.660 |
| 90 | 1.291 | 1.662 | 1.987 | 2.368 | 2.632 |
| 120 | 1.289 | 1.658 | 1.980 | 2.358 | 2.617 |
| ∞ | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 |

Degrees of Freedom

# Example: one-sided testing

Economic significance:

- Change in $x_j$ is associated with how many units change in y
- In this example, 1 more year of experience is associated with a 0.41% higher wage
  - Not that large in the economic sense

Difference from statistical significance

- Always discuss statistical significance first, economic significance second

# One-sided test: alternative

Alterative one-sided alterative

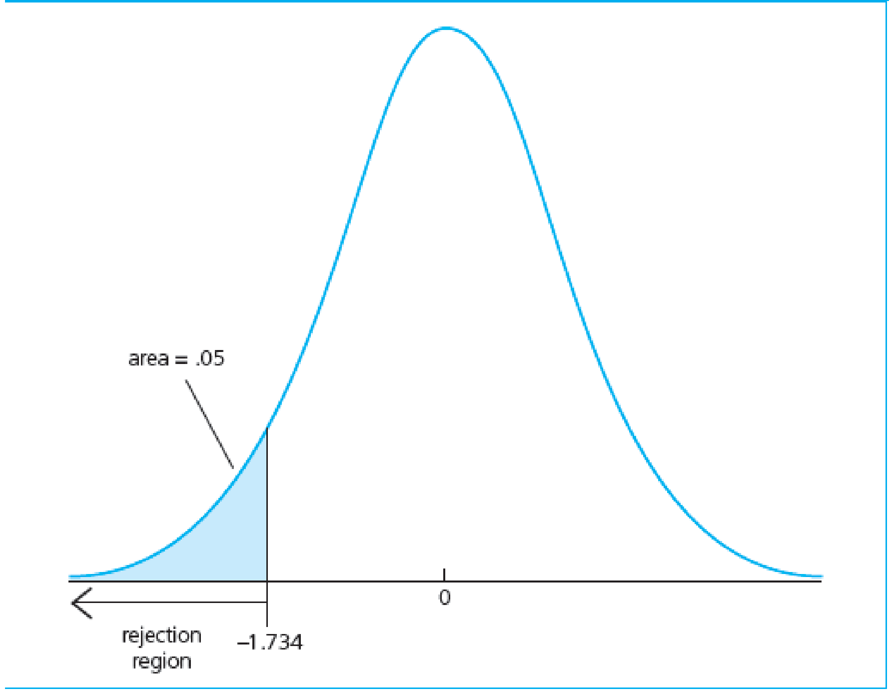
$$H_0: \beta_j \geq 0$$

$$H_1: \beta_j < 0$$

Rejection rule $t_{\hat{\beta}_j} < -c$

Suppose we have the t statistic -1.5, degree of freedom 18. Reject or not reject at the 5% significance level?

FIGURE 4.3 5% rejection rule for the alternative $H_1: \beta_j < 0$ with 18 $df$.



# Testing against two-sided alternatives

Two-sided alternatives

$$\mathrm{H}_0\colon \beta_j = 0$$

$$\mathrm{H}_1\colon \beta_j \neq 0$$

more common in empirical test

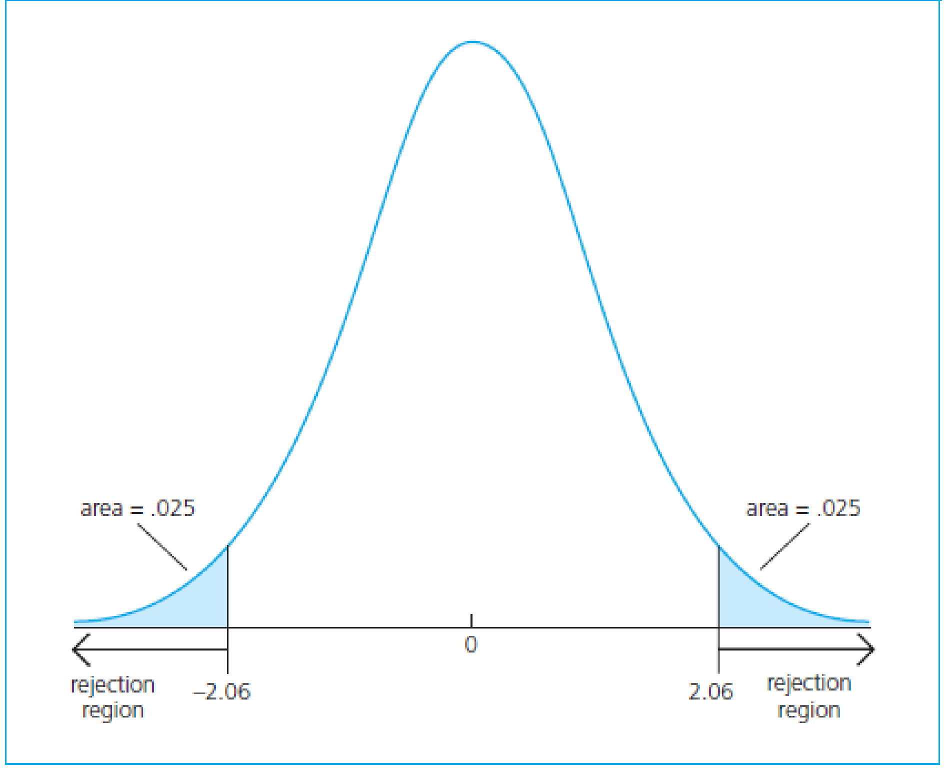
- No prior on the sign of $\beta_j$, for example, because either sign is supported by one economic theory

Rejection rule $|t_{\hat{\beta}_j}| > c$

- Suppose we have the t statistic -1.5, degree of freedom 18. Reject or not reject at the 5% significance level?

# Two-sided testing (graphically)

FIGURE 4.4 5% rejection rule for the alternative $H_1$: $\beta_j \neq 0$ with 25 *df*.




Degree of freedom 25, 5% significance level

# Variant of hypothesis

Sometimes, we are interested in testing

$$\mathrm{H}_0 : \beta_j = a_j$$

$$\mathrm{H}_1 :\ \beta_j \neq 0$$

For example, in international trade

- Interested in whether 1% devaluation of local currency increases the price of imported goods by 1% (pass-through)
- i.e., $a_j$=1

# Variant of hypothesis

Under the null hypothesis

$$t = (\hat{\beta}_j - a_j)/\mathrm{se}(\hat{\beta}_j).$$

follows a t distribution

Depending our alternative hypothesis:

- Rejection rule for this new t statistic could be $t_{\hat{\beta}_j} > c, t_{\hat{\beta}_j} < -c, or\ |t_{\hat{\beta}_j}| > c$
- In the exchange rate pass through example, theory rules out $\beta_j > 1,$ so $H_1\colon \beta_j < 1$

Previous case: a special case $a_j = 0$

# P-values for t tests

Note we assign a 5% significance level for testing

- ➢ then compare the t statistic to the corresponding critical value

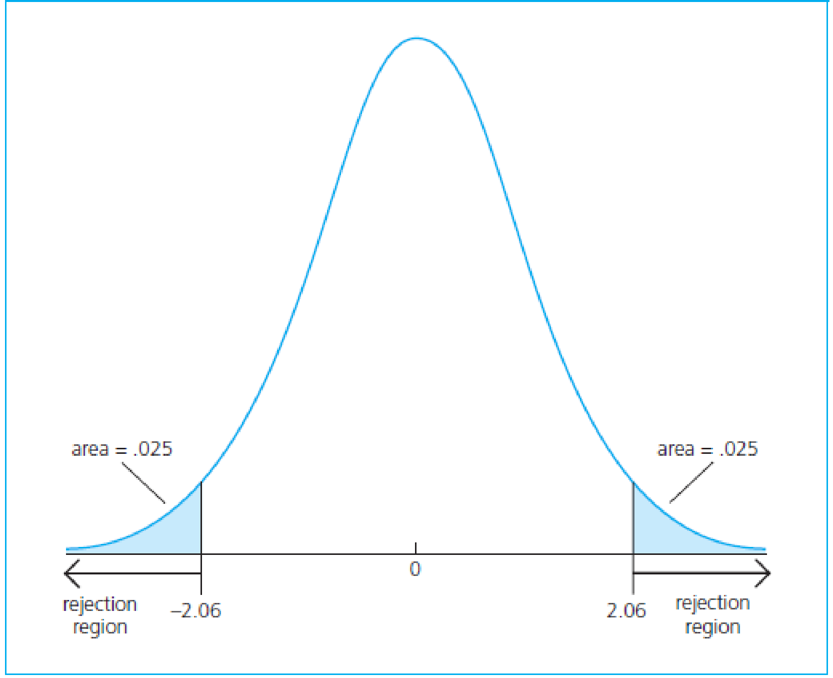
Another way is to find the p-value for the t statistic, i.e., find

$$P(|T| > |t|)$$

where $T$ denote a t distributed random variable with n-k-1 degree of freedom

- ➢ for two-sided test
- ➢ Then compare this p-value to the designated significance level

FIGURE 4.4 5% rejection rule for the alternative $H_1: \beta_j \neq 0$ with 25 $df$.




Suppose df=25, t statistic is 1.93, find the p-value? Reject or not reject?

# Confidence intervals

When we use critical value or p-value for testing,

- assume if null hypothesis $\mathrm{H}_0: \beta_j = 0$ is true

A different perspective:

- don't know $\beta_j$
- but with a certain confidence level that the $\beta_j$ fall into a certain interval

Suppose the confidence level is 95%

$$Prob\left(\frac{|\hat{\beta}_j - \beta_j|}{se\left(\hat{\beta}_j\right)} < c\right) = 95\%$$

$\Rightarrow$ if I sample N times ($N \to \infty$), the number of sample that $\frac{|\hat{\beta}_j - \beta_j|}{se\left(\hat{\beta}_j\right)} < c$ holds is 95%*N

$\Rightarrow$ Of course, the $\hat{\beta}_j$ and $se\left(\hat{\beta}_j\right)$ may change from sample to sample

# Confidence intervals

Hence we derive the 95% confidence interval (CI) as

$$[\hat{\beta}_j - c * se(\hat{\beta}_j), \hat{\beta}_j + c * se(\hat{\beta}_j)]$$

Suppose df=25, $\hat{\beta}_j$ is 0.67, $se(\hat{\beta}_j)$=0.2, construct the 95% CI, 90% CI.

# Hypothesis testing on multiple population parameters

So far, we test on a single parameter.

Scenarios that we are interested in multiple population parameters

- linear combination $\mathrm{H}_0: a_1\beta_1 + a_2\beta_2 + \cdots + a_k\beta_k = 0$
  - e.g., estimating a production function and testing constant return to scale
- multiple restrictions $\mathrm{H}_0: \beta_{k-q-1} = 0, \ldots, \beta_k = 0$
  - e.g., testing redundant variables in a regression

# Linear combination

Example of $\mathrm{H}_0: a_1\beta_1 + a_2\beta_2 + \cdots + a_k\beta_k = 0$ from the textbook: wage equation

$$\log(wage) = \beta_0 + \beta_1 jc + \beta_2 univ + \beta_3 exper + u,$$

$jc$ = number of years attending a two-year college.
$univ$ = number of years at a four-year college.
$exper$ = months in the workforce.

Interested in testing whether the increases in wage by a additional year in collage and university equal

$$\mathrm{H}_0: \beta_1 = \beta_2$$

$$\mathrm{H}_1: \beta_1 < \beta_2$$

# Linear combination

$(\hat{\beta}_1 - \beta_1)/sd(\hat{\beta}_1)$ and $(\hat{\beta}_2 - \beta_2)/sd(\hat{\beta}_2)$ follow normal distributions

=> $(\hat{\beta}_1 - \hat{\beta}_2 - \beta_1 + \beta_2)/\ sd(\hat{\beta}_1 - \hat{\beta}_2)$ normal

=> $(\hat{\beta}_1 - \hat{\beta}_2 - \beta_1 + \beta_2)/\ se(\hat{\beta}_1 - \hat{\beta}_2)$ *(numerator $\hat{\beta}_1 - \hat{\beta}_2$ if null is true)* t distribution

From the estimation results, we have $\hat{\beta}_1$, $\hat{\beta}_2$, but no $se(\hat{\beta}_1 - \hat{\beta}_2)$

$$\widehat{\log(wage)} = \underset{(.021)}{1.472} + \underset{(.0068)}{.0667}\ jc + \underset{(.0023)}{.0769}\ univ + \underset{(.0002)}{.0049}\ exper$$

$$n = 6{,}763,\ R^2 = .222.$$

# Linear combination

An alterative way is to let $\theta_1 = \beta_1 - \beta_2$ and test

$$H_0: \theta_1 = 0 \ \ against \ \ H_1: \theta_1 < 0$$

The new regression equation is

$$\begin{aligned} \log(wage) &= \beta_0 + (\theta_1 + \beta_2) jc + \beta_2 univ + \beta_3 exper + u \\ &= \beta_0 + \theta_1 jc + \beta_2 (jc + univ) + \beta_3 exper + u. \end{aligned}$$

with estimated results

$$\begin{aligned} \widehat{\log(wage)} = \underset{(.021)}{1.472} - \underset{(.0069)}{.0102}\, jc + \underset{(.0023)}{.0769}\, totcoll + \underset{(.0002)}{.0049}\, exper \\ n = 6{,}763,\ R^2 = .222. \end{aligned}$$

Reject or not at the 5% significance level?

# Multiple linear restrictions

Joint (multiple) hypothesis test

$$H_0: \beta_{k-q-1} = 0, \dots, \beta_k = 0$$

Extreme case $H_0: \beta_1 = 0, \dots, \beta_k = 0$

- Whether the chosen set of explanatory variables are statistically significant

An example using MLB1.raw by typing `bcuse mlb1`

Interested in testing $H_0: \beta_3 = 0, \beta_4 = 0, \beta_5 = 0.$ in the rhs equation

$$\widehat{\log(salary)} = \underset{(0.29)}{11.19} + \underset{(.0121)}{.0689}\, years + \underset{(.0026)}{.0126}\, gamesyr + \underset{(.00110)}{.00098}\, bavg + \underset{(.0161)}{.0144}\, hrunsyr + \underset{(.0072)}{.0108}\, rbisyr$$

$$n = 353, \text{SSR} = 183.186, R^2 = .6278,$$

# Multiple linear restrictions

Cannot check the significance of $\beta_3, \beta_4, and\ \beta_5$ one-by-one

- Not a joint test

Rather, if the null hypothesis is true, we get the following restricted model,

$$\log(salary) = \beta_0 + \beta_1 years + \beta_2 gamesyr + u.$$

- Intuitively, if the null is true, the unexplained sum of squared residuals (SSR) in the restricted model should not be too smaller than that in the unrestricted model

# Multiple linear restrictions

Given that

$$\frac{SSR}{\sigma^2} = \frac{(n-k-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-k-1}$$

=> $\frac{SSR_{ur}}{\sigma^2} = \frac{(n-k-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-k-1}$; $\frac{SSR_r}{\sigma^2} = \frac{(n-k-1+q)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-k-1+q}$

=> $\frac{SSR_r - SSR_{ur}}{\sigma^2} \sim \chi^2_q$

=> F statistic $F \equiv \frac{(\mathrm{SSR}_r - \mathrm{SSR}_{ur})/q}{\mathrm{SSR}_{ur}/(n-k-1)}$, follows F(q,n-k-1) distribution

# Multiple linear restrictions

In this example, estimation results for the restricted model

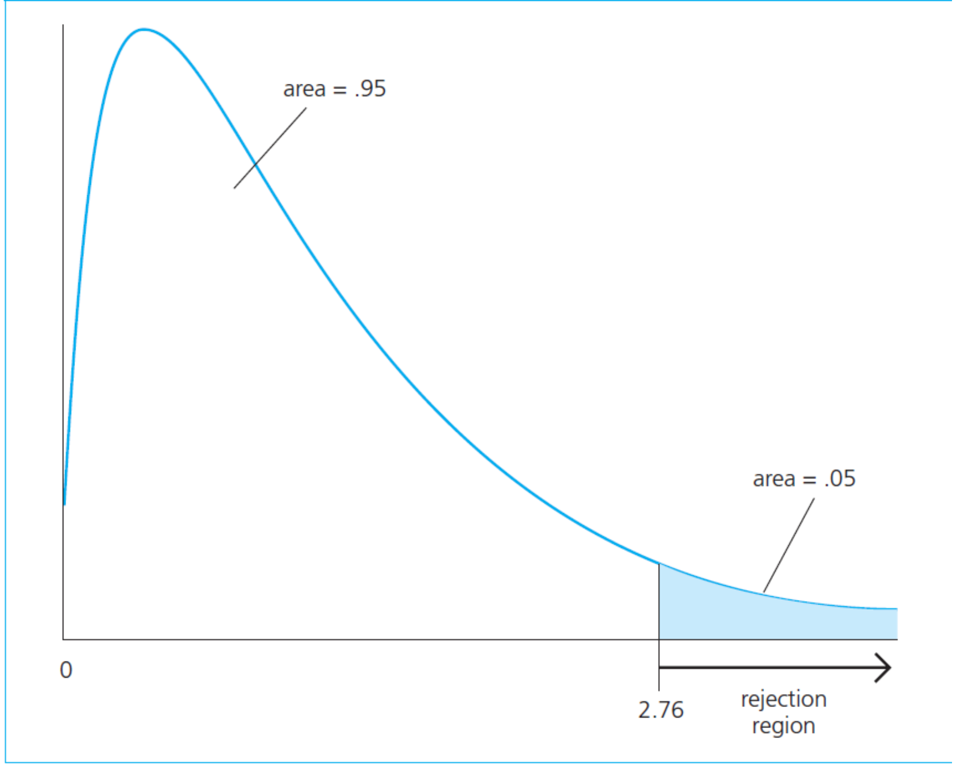
$$\widehat{\log(salary)} = \underset{(.11)}{11.22} + \underset{(.0125)}{.0713}\, years + \underset{(.0013)}{.0202}\, gamesyr$$
$$n = 353,\ \mathrm{SSR} = 198.311,\ R^2 = .5971.$$

F statistic=9.55

Reject if F statistic greater than the critical value

- $x_{k-q+1}, \ldots, x_k$ $are$ jointly statistically significant

FIGURE 4.7 The 5% critical value and rejection region in an $F_{3,60}$ distribution.



# R-squared form of the F statistic

$$F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)}$$

$$= \frac{\frac{SST - SSR_r - SST - SSR_{ur}}{SST * q}}{\frac{SST - SSR_{ur}}{SST * (n-k-1)}}$$

$$= \frac{(R_{ur}^2 - R_r^2)/q}{(1 - R_{ur}^2)/(n-k-1)}$$

# Computing p-values

Similar to a single parameter testing, we can define

$$p\text{-value} = \mathrm{P}(\mathcal{F} > F),$$

Find the p-value for the above F statistic.

# Relation b.w. a single parameter testing vs joint testing

1. $x_{k-q+1}, \ldots, x_k$ *are* jointly statistically significant => each x statistically significant?

2. Each x statistically not significant => $x_{k-q+1}, \ldots, x_k$ *are* jointly statistically insignificant?
   - No in our MLB1 data. This is because a high correlation b.w. hrunsyr and rbisyr
   - hrunsyr statistically significant if we drop rbisyr

| | hrunsyr | rbisyr |
|---|---|---|
| hrunsyr | 1.0000 | |
| rbisyr | 0.8907 | 1.0000 |

| logsalary | Coef. | Std. Err. | t | P>\|t\| |
|---|---|---|---|---|
| years | .0677325 | .0121128 | 5.59 | 0.000 |
| gamesyr | .0157595 | .0015636 | 10.08 | 0.000 |
| bavg | .0014185 | .0010658 | 1.33 | 0.184 |
| hrunsyr | .0359434 | .0072408 | 4.96 | 0.000 |
| _cons | 11.02091 | .2657191 | 41.48 | 0.000 |

# Relation b.w. a single parameter testing vs joint testing

3. $x_{k-q+1}, \ldots, x_k$ $are$ jointly statistically insignificant => each x statistically insignificant?

4. Each x statistically significant => $x_{k-q+1}, \ldots, x_k$ $are$ jointly statistically significant?

# Relation b.w. F and t statistics

Special case of the joint test

$$\mathrm{H}_0\colon \beta_j = 0$$

The restricted model could be

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_{j-1} x_{j-1} + \beta_{j+1} x_{j+1} + \cdots + \beta_k x_k + \mu$$

Can construct the F statistic that follows (1, n—k,1) distribution

- Same result when using the t statistic
- In fact, $t_{n-k-1}^2 = F(1, n-k-1)$

# Overall significance of regression

Another special case for the joint test

$$\mathrm{H}_0: \beta_1 = 0, \dots, \beta_k = 0$$

Restricted model in this case

$$y = \beta_0 + \mu$$

$$F = \frac{(R_{ur}^2 - R_r^2)/q}{(1 - R_{ur}^2)/(n-k-1)} = \frac{(R_{ur}^2 - 0)/k}{(1 - R_{ur}^2)/(n-k-1)} = \frac{R^2 /k}{(1 - R^2)/(n-k-1)}$$

Find the F statistic in the following results from Stata, verify the p-value and if the above F equation holds.

# Overall significance of regression

| Source | SS | df | MS |
|---|---|---|---|
| Model | 308.9892 | 5 | 61.79784 |
| Residual | 183.186335 | 347 | .52791451 |
| Total | 492.175535 | 352 | 1.39822595 |

| | | |
|---|---|---|
| Number of obs | = | 353 |
| F(5, 347) | = | 117.06 |
| Prob > F | = | 0.0000 |
| R-squared | = | 0.6278 |
| Adj R-squared | = | 0.6224 |
| Root MSE | = | .72658 |

| logsalary | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| years | .0688626 | .0121145 | 5.68 | 0.000 | .0450355 | .0926898 |
| gamesyr | .0125521 | .0026468 | 4.74 | 0.000 | .0073464 | .0177578 |
| bavg | .0009786 | .0011035 | 0.89 | 0.376 | -.0011918 | .003149 |
| hrunsyr | .0144295 | .016057 | 0.90 | 0.369 | -.0171518 | .0460107 |
| rbisyr | .0107657 | .007175 | 1.50 | 0.134 | -.0033462 | .0248776 |
| _cons | 11.19242 | .2888229 | 38.75 | 0.000 | 10.62435 | 11.76048 |

# Variant of multiple restrictions

$H_0: \beta_{k-q-1} = a \quad , \beta_{k-q} = 0, \dots, \beta_k = 0$ ?

Use the following regression as the unrestricted model

$$y - a \quad x_{k-q-1} = \beta_0 + \beta_1 x_1 + \cdots + \beta_{k-q-2} x_{k-q-2} + \theta_{k-q-1} x_{k-q-1} + \beta_{k-q} x_{k-q} + \cdots + \beta_k x_k + \mu$$

Now the new null hypothesis $H_0: \theta_{k-q-1} = a \quad , \beta_{k-q} = 0, \dots, \beta_k = 0$

Can construct the F statistic as the earlier case

# Reporting regression results

Rules of thumb

1. Estimated OLS coefficients should always be reported, regardless of their significance
2. Report standard errors, or t statistics
3. Report R squared
4. We may estimate several equations with many different sets of independent variables. Report them column-by-column.

# Example

Table 3: Effects of Provincial Financial Development on Misallocation for Industries with Different Financial Vulnerabilities, Industry-Province Clustered

| | $FinDevp = LoanMkt$ | | | | | | $FinDevp = FinMkt$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CV(MRPM) | | | | | CV(MRPK) | CV(MRPM) | | | | | CV(MRPK) |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| $FinDevp_{pt}$ | -0.0120** | -0.0117** | -0.0131** | -0.0012 | -0.0103 | -0.0276 | -0.0130* | -0.0130* | -0.0134* | 0.0018 | -0.0117 | -0.0133 |
| | (0.0040) | (0.0048) | (0.0029) | (0.6357) | (0.0504) | (0.1511) | (0.0113) | (0.0117) | (0.0118) | (0.5882) | (0.0657) | (0.6553) |
| $FinDevp_{pt} \times AssetTang_s$ | 0.0176** | 0.0172** | 0.0124 | -0.0005 | 0.0154* | 0.0520 | 0.0182** | 0.0180** | 0.0165* | -0.0067 | 0.0161 | 0.0385 |
| | (0.0015) | (0.0019) | (0.0513) | (0.8783) | (0.0234) | (0.0660) | (0.0087) | (0.0091) | (0.0362) | (0.1305) | (0.0546) | (0.3735) |
| $FinDevp_{pt} \times ExtDep_s$ | -0.0024 | -0.0022 | -0.0026 | -0.0034* | -0.0023 | -0.0192* | -0.0032 | -0.0031 | -0.0033 | -0.0043* | -0.0032 | -0.0219* |
| | (0.1152) | (0.1397) | (0.0839) | (0.0369) | (0.1201) | (0.0103) | (0.1206) | (0.1313) | (0.1115) | (0.0351) | (0.1231) | (0.0380) |
| $FinDevp_{pt} \times CCC_s$ | 0.0111*** | 0.0110*** | 0.0097*** | | 0.0108*** | 0.0172 | 0.0149*** | 0.0148*** | 0.0144*** | | 0.0146*** | 0.0146 |
| | (0.0000) | (0.0000) | (0.0000) | | (0.0000) | (0.1267) | (0.0000) | (0.0000) | (0.0000) | | (0.0000) | (0.3965) |
| $CV(MRPK_{spt})$ | | 0.0079* | | | | | | 0.0043 | | | | |
| | | (0.0167) | | | | | | (0.1511) | | | | |
| $FinDevp_{pt} \times Upstream_s$ | | | 0.0089 | | | | | | 0.0029 | | | |
| | | | (0.1612) | | | | | | (0.6947) | | | |
| $FinDevp_{pt} \times CCC\_SSBF_s$ | | | | 0.0122*** | | | | | | 0.0156*** | | |
| | | | | (0.0003) | | | | | | (0.0002) | | |
| $FinDevp_{pt} \times CCC\_LST_s$ | | | | | -0.0007 | | | | | | -0.0007 | |
| | | | | | (0.5770) | | | | | | (0.6793) | |
| $SOEShare_{spt}$ | 0.0459** | 0.0439** | 0.0467** | 0.0483** | 0.0454** | 0.2508** | 0.0299* | 0.0288* | 0.0301* | 0.0323* | 0.0301* | 0.2550** |
| | (0.0017) | (0.0025) | (0.0014) | (0.0011) | (0.0019) | (0.0029) | (0.0391) | (0.0455) | (0.0386) | (0.0264) | (0.0383) | (0.0046) |
| $ExporterShare_{spt}$ | 0.0324* | 0.0305* | 0.0342** | 0.0342** | 0.0307* | 0.2443*** | 0.0230* | 0.0219 | 0.0234* | 0.0236* | 0.0213 | 0.2479*** |
| | (0.0105) | (0.0153) | (0.0073) | (0.0073) | (0.0157) | (0.0002) | (0.0487) | (0.0592) | (0.0460) | (0.0436) | (0.0673) | (0.0004) |
| Constant | 0.3382*** | 0.3249*** | 0.3428*** | 0.3374*** | 0.3385*** | 1.6834*** | 0.2758*** | 0.2687*** | 0.2772*** | 0.2744*** | 0.2776*** | 1.6369*** |
| | (0.0000) | (0.0000) | (0.0000) | (0.0000) | (0.0000) | (0.0000) | (0.0000) | (0.0000) | (0.0000) | (0.0000) | (0.0000) | (0.0000) |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| Province FE | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES |
| N | 6669 | 6669 | 6669 | 6669 | 6640 | 6669 | 5999 | 5999 | 5999 | 5999 | 5975 | 5999 |
| Adj. R-sq | 0.446 | 0.447 | 0.447 | 0.446 | 0.447 | 0.255 | 0.449 | 0.449 | 0.449 | 0.447 | 0.450 | 0.257 |

*Note:* $FinMkt_{pt}$ index starts from 1999 and hence the number of observations drops when the $FinDevp_{pt}$ is proxied by $FinMkt_{pt}$. Industry-province cells with fewer than 20 firms are dropped. P-values in parentheses are 0.05 for *, 0.01 for **, and 0.001 for ***.

# Summary

1. Sampling distributions of the OLS estimators
2. Inference on a single population estimator
   a) t test: one-sided vs two-sided
   b) p-value for t statistic
   c) Confidence intervals
3. Inference on a linear combination of parameters
   a) Unrestricted vs restricted models
   b) F statistic
4. Convention in reporting regression results