Hypothesis Testing in Multiple Regression

Same Logic, More Coefficients, New Questions

Jake Anderson

March 21, 2026

Outline

1. From SLR to MR: What Changes?
2. Individual t-Tests in MR
3. Marginal Effects with Polynomial Terms
4. Testing Linear Combinations of Coefficients
5. The General Linear Combination Formula
6. Individual $\neq$ Joint Significance
7. Summary

Outline

1. From SLR to MR: What Changes?
2. Individual t-Tests in MR
3. Marginal Effects with Polynomial Terms
4. Testing Linear Combinations of Coefficients
5. The General Linear Combination Formula
6. Individual $\neq$ Joint Significance
7. Summary

Where We Left Off

In simple regression, you learned to test hypotheses about β_k using:

$$t = \frac{b_k - c}{\text{se}(b_k)} \sim t_{(N-2)}$$

Where We Left Off

In simple regression, you learned to test hypotheses about β_k using:

$$t = \frac{b_k - c}{\text{se}(b_k)} \sim t_{(N-2)}$$

In multiple regression, the mechanics are **identical**. Two things change:

Where We Left Off

In simple regression, you learned to test hypotheses about β_k using:

$$t = \frac{b_k - c}{\text{se}(b_k)} \sim t_{(N-2)}$$

In multiple regression, the mechanics are **identical**. Two things change:

1. **Degrees of freedom:** $N - 2 \longrightarrow N - K$

   where K = total number of parameters (intercept + all slopes)

Where We Left Off

In simple regression, you learned to test hypotheses about β_k using:

$$t = \frac{b_k - c}{\text{se}(b_k)} \sim t_{(N-2)}$$

In multiple regression, the mechanics are **identical**. Two things change:

1. **Degrees of freedom:** $N - 2 \longrightarrow N - K$

   where K = total number of parameters (intercept + all slopes)

2. **New questions become possible:**

   With multiple regressors, you can test linear combinations of coefficients, not just individual ones

Where We Left Off

In simple regression, you learned to test hypotheses about β_k using:

$$t = \frac{b_k - c}{\text{se}(b_k)} \sim t_{(N-2)}$$

In multiple regression, the mechanics are **identical**. Two things change:

1. **Degrees of freedom:** $N - 2 \longrightarrow N - K$

   where K = total number of parameters (intercept + all slopes)

2. **New questions become possible:**

   With multiple regressors, you can test linear combinations of coefficients, not just individual ones

$\Longrightarrow$ If you can do t-tests in SLR, you can do them in MR. Today is about the new questions MR opens up.

The Wage Model

We want to understand how education and experience affect wages. Start with:

$$\text{WAGE}_i = \beta_1 + \beta_2\,\text{EDUC}_i + \beta_3\,\text{EXPER}_i + e_i$$

The Wage Model

We want to understand how education and experience affect wages. Start with:

$$\text{WAGE}_i = \beta_1 + \beta_2\,\text{EDUC}_i + \beta_3\,\text{EXPER}_i + e_i$$

But this assumes each additional year of experience adds the same amount to wages, regardless of how much experience you already have.

The Wage Model

We want to understand how education and experience affect wages. Start with:

$$\text{WAGE}_i = \beta_1 + \beta_2\,\text{EDUC}_i + \beta_3\,\text{EXPER}_i + e_i$$

But this assumes each additional year of experience adds the same amount to wages, regardless of how much experience you already have.

A 25-year-old with 3 years of experience gets the same bump from year 4 as a 55-year-old gets from year 33?

The Wage Model

We want to understand how education and experience affect wages. Start with:

$$\text{WAGE}_i = \beta_1 + \beta_2\,\text{EDUC}_i + \beta_3\,\text{EXPER}_i + e_i$$

But this assumes each additional year of experience adds the same amount to wages, regardless of how much experience you already have.

A 25-year-old with 3 years of experience gets the same bump from year 4 as a 55-year-old gets from year 33?

That seems unlikely. We probably need **diminishing returns to experience**.

The Wage Model

We want to understand how education and experience affect wages. Start with:

$$\text{WAGE}_i = \beta_1 + \beta_2\,\text{EDUC}_i + \beta_3\,\text{EXPER}_i + e_i$$

But this assumes each additional year of experience adds the same amount to wages, regardless of how much experience you already have.

A 25-year-old with 3 years of experience gets the same bump from year 4 as a 55-year-old gets from year 33?

That seems unlikely. We probably need **diminishing returns to experience**.

$\Longrightarrow$ Add EXPER^2 to allow the marginal effect of experience to change:

$$\text{WAGE}_i = \beta_1 + \beta_2\,\text{EDUC}_i + \beta_3\,\text{EXPER}_i + \beta_4\,\text{EXPER}_i^2 + e_i$$

The Estimated Model

Using the CPS data ($N = 1000$, $K = 4$, $df = 996$):

$$\widehat{\text{WAGE}} = \underset{(2.87)}{-10.40} + \underset{(0.16)}{2.44}\ \text{EDUC} + \underset{(0.11)}{0.55}\ \text{EXPER} \underset{(0.0020)}{-0.0063}\ \text{EXPER}^2$$

(standard errors in parentheses)

The Estimated Model

Using the CPS data ($N = 1000$, $K = 4$, $df = 996$):

$$\widehat{WAGE} = \underset{(2.87)}{-10.40} + \underset{(0.16)}{2.44}\ EDUC + \underset{(0.11)}{0.55}\ EXPER \underset{(0.0020)}{-0.0063}\ EXPER^2$$

(standard errors in parentheses)

	b_1	b_2	b_3	b_4
Estimate	-10.40	2.44	0.55	-0.0063
Std. Error	2.87	0.16	0.11	0.0020

The Estimated Model

Using the CPS data ($N = 1000$, $K = 4$, $df = 996$):

$$\widehat{WAGE} = \underset{(2.87)}{-10.40} + \underset{(0.16)}{2.44}\ EDUC + \underset{(0.11)}{0.55}\ EXPER \underset{(0.0020)}{-0.0063}\ EXPER^2$$

(standard errors in parentheses)

	b_1	b_2	b_3	b_4
Estimate	-10.40	2.44	0.55	-0.0063
Std. Error	2.87	0.16	0.11	0.0020

Today's questions:

1. Are individual coefficients statistically significant?
2. What is the **marginal return to experience** at a specific level, and is it significant?
3. Does "individually significant" mean "jointly significant"?

Outline

1. From SLR to MR: What Changes?
2. Individual t-Tests in MR
3. Marginal Effects with Polynomial Terms
4. Testing Linear Combinations of Coefficients
5. The General Linear Combination Formula
6. Individual $\neq$ Joint Significance
7. Summary

Testing a Single Coefficient: The Formula

Exactly as in SLR, but with $N - K$ degrees of freedom:

Testing a Single Coefficient: The Formula

Exactly as in SLR, but with $N - K$ degrees of freedom:

$$t = \frac{b_k - c}{\text{se}(b_k)} \sim t_{(N-K)} \quad \text{under } H_0: \beta_k = c$$

Testing a Single Coefficient: The Formula

Exactly as in SLR, but with $N - K$ degrees of freedom:

$$t = \frac{b_k - c}{\text{se}(b_k)} \sim t_{(N-K)} \quad \text{under } H_0 : \beta_k = c$$

Significance test (most common): $H_0 : \beta_k = 0$ vs. $H_1 : \beta_k \neq 0$

$$t = \frac{b_k}{\text{se}(b_k)}$$

Testing a Single Coefficient: The Formula

Exactly as in SLR, but with $N - K$ degrees of freedom:

$$t = \frac{b_k - c}{\text{se}(b_k)} \sim t_{(N-K)} \quad \text{under } H_0 : \beta_k = c$$

Significance test (most common): $H_0 : \beta_k = 0$ vs. $H_1 : \beta_k \neq 0$

$$t = \frac{b_k}{\text{se}(b_k)}$$

Rejection rules are the same as before:

- Two-tail: reject if $|t| \geq t_{(1-\alpha/2,\ N-K)}$
- Right-tail ($H_1 : \beta_k > c$): reject if $t \geq t_{(1-\alpha,\ N-K)}$
- Left-tail ($H_1 : \beta_k < c$): reject if $t \leq -t_{(1-\alpha,\ N-K)}$

Example: Is Education Significant?

$$H_0 : \beta_2 = 0 \quad \text{vs.} \quad H_1 : \beta_2 \neq 0 \qquad (\alpha = 0.05)$$

Example: Is Education Significant?

$H_0 : \beta_2 = 0$ vs. $H_1 : \beta_2 \neq 0$ $(\alpha = 0.05)$

$$t = \frac{b_2}{\text{se}(b_2)} = \frac{2.44}{0.16} = 15.25$$

Example: Is Education Significant?

$H_0 : \beta_2 = 0 \quad \text{vs.} \quad H_1 : \beta_2 \neq 0 \qquad (\alpha = 0.05)$

$$t = \frac{b_2}{\text{se}(b_2)} = \frac{2.44}{0.16} = 15.25$$

Critical value: $t_{(0.975,\ 996)} \approx 1.962$

Example: Is Education Significant?

$H_0 : \beta_2 = 0$ vs. $H_1 : \beta_2 \neq 0$ $(\alpha = 0.05)$

$$t = \frac{b_2}{\text{se}(b_2)} = \frac{2.44}{0.16} = 15.25$$

Critical value: $t_{(0.975,\ 996)} \approx 1.962$

$|15.25| \geq 1.962?$ Yes. $p \approx 0.000$.

Example: Is Education Significant?

$H_0: \beta_2 = 0$ vs. $H_1: \beta_2 \neq 0$ $(\alpha = 0.05)$

$$t = \frac{b_2}{\text{se}(b_2)} = \frac{2.44}{0.16} = 15.25$$

Critical value: $t_{(0.975,\ 996)} \approx 1.962$

$|15.25| \geq 1.962$? Yes. $p \approx 0.000$.

Reject H_0. Strong evidence that education affects wages, holding experience constant.

Example: Is Education Significant?

$H_0 : \beta_2 = 0$ vs. $H_1 : \beta_2 \neq 0$ $(\alpha = 0.05)$

$$t = \frac{b_2}{\text{se}(b_2)} = \frac{2.44}{0.16} = 15.25$$

Critical value: $t_{(0.975,\ 996)} \approx 1.962$

$|15.25| \geq 1.962?$ Yes. $p \approx 0.000$.

Reject H_0. Strong evidence that education affects wages, holding experience constant.

$\implies$ This is no different from what you did in SLR. The only change is the degrees of freedom (996 instead of 998).

Example: Is EXPER^2 Significant?

$$H_0: \beta_4 = 0 \quad \text{vs.} \quad H_1: \beta_4 \neq 0 \qquad (\alpha = 0.05)$$

Example: Is EXPER^2 Significant?

$H_0 : \beta_4 = 0 \quad \text{vs.} \quad H_1 : \beta_4 \neq 0 \qquad (\alpha = 0.05)$

$$t = \frac{b_4}{\text{se}(b_4)} = \frac{-0.0063}{0.0020} = -3.15$$

Example: Is EXPER^2 Significant?

$H_0 : \beta_4 = 0 \quad \text{vs.} \quad H_1 : \beta_4 \neq 0 \qquad (\alpha = 0.05)$

$$t = \frac{b_4}{\text{se}(b_4)} = \frac{-0.0063}{0.0020} = -3.15$$

$|-3.15| \geq 1.962? \quad$ Yes. $\quad p \approx 0.002.$

Example: Is EXPER^2 Significant?

$H_0: \beta_4 = 0 \quad$ vs. $\quad H_1: \beta_4 \neq 0 \qquad (\alpha = 0.05)$

$$t = \frac{b_4}{\mathrm{se}(b_4)} = \frac{-0.0063}{0.0020} = -3.15$$

$|-3.15| \geq 1.962$? Yes. $p \approx 0.002$.

Reject H_0. The quadratic term is significant.

Example: Is EXPER^2 Significant?

$H_0 : \beta_4 = 0 \quad \text{vs.} \quad H_1 : \beta_4 \neq 0 \qquad (\alpha = 0.05)$

$$t = \frac{b_4}{\text{se}(b_4)} = \frac{-0.0063}{0.0020} = -3.15$$

$|-3.15| \geq 1.962?$ Yes. $p \approx 0.002$.

Reject H_0. The quadratic term is significant.

$\Longrightarrow$ The data support diminishing returns to experience. The linear model (EXPER only) was too restrictive.

Example: Is EXPER^2 Significant?

$H_0: \beta_4 = 0 \quad$ vs. $\quad H_1: \beta_4 \neq 0 \qquad (\alpha = 0.05)$

$$t = \frac{b_4}{\text{se}(b_4)} = \frac{-0.0063}{0.0020} = -3.15$$

$|-3.15| \geq 1.962$? Yes. $p \approx 0.002$.

Reject H_0. The quadratic term is significant.

$\Longrightarrow$ The data support diminishing returns to experience. The linear model (EXPER only) was too restrictive.

But notice: $b_4 = -0.0063$ looks tiny. Is it economically meaningful? We need to think about what the **marginal effect** of experience actually is.

Outline

1. From SLR to MR: What Changes?
2. Individual t-Tests in MR
3. Marginal Effects with Polynomial Terms
4. Testing Linear Combinations of Coefficients
5. The General Linear Combination Formula
6. Individual $\neq$ Joint Significance
7. Summary

The Marginal Effect of Experience

With a quadratic term, the coefficient β_3 is **not** the marginal effect of experience. The marginal effect is the derivative:

The Marginal Effect of Experience

With a quadratic term, the coefficient β_3 is **not** the marginal effect of experience. The marginal effect is the derivative:

$$\frac{\partial\, \text{E(WAGE)}}{\partial\, \text{EXPER}} = \beta_3 + 2\beta_4\, \text{EXPER}$$

The Marginal Effect of Experience

With a quadratic term, the coefficient β_3 is **not** the marginal effect of experience. The marginal effect is the derivative:

$$\frac{\partial \text{ E(WAGE)}}{\partial \text{ EXPER}} = \beta_3 + 2\beta_4 \text{ EXPER}$$

$\Longrightarrow$ The effect of one more year of experience **depends on how much experience you already have**.

The Marginal Effect of Experience

With a quadratic term, the coefficient β_3 is **not** the marginal effect of experience. The marginal effect is the derivative:

$$\frac{\partial\, \text{E(WAGE)}}{\partial\, \text{EXPER}} = \beta_3 + 2\beta_4\, \text{EXPER}$$

$\implies$ The effect of one more year of experience **depends on how much experience you already have**.

Plugging in our estimates:

$$\widehat{\text{Marginal effect}} = 0.55 + 2(-0.0063) \times \text{EXPER} = 0.55 - 0.0126 \times \text{EXPER}$$

The Marginal Effect of Experience

With a quadratic term, the coefficient β_3 is **not** the marginal effect of experience. The marginal effect is the derivative:

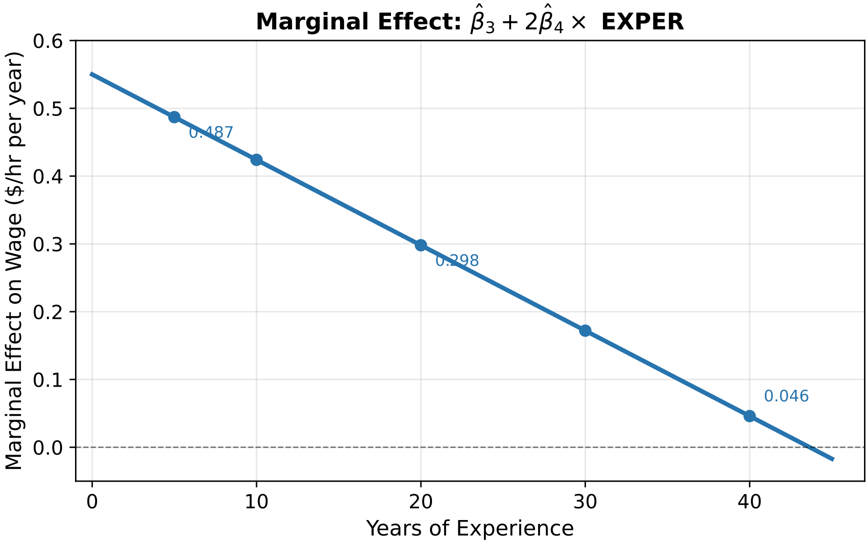
$$\frac{\partial \, E(WAGE)}{\partial \, EXPER} = \beta_3 + 2\beta_4 \, EXPER$$

$\implies$ The effect of one more year of experience **depends on how much experience you already have.**

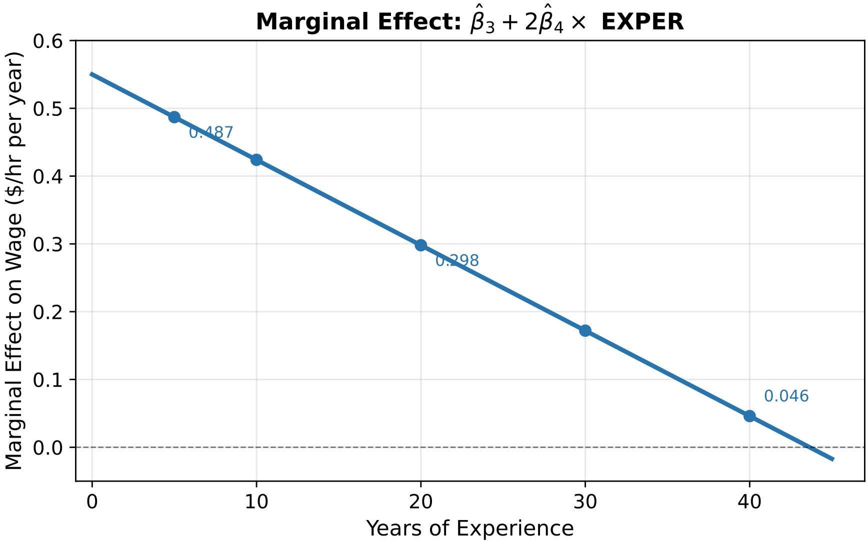
Plugging in our estimates:

$$\widehat{\text{Marginal effect}} = 0.55 + 2(-0.0063) \times EXPER = 0.55 - 0.0126 \times EXPER$$

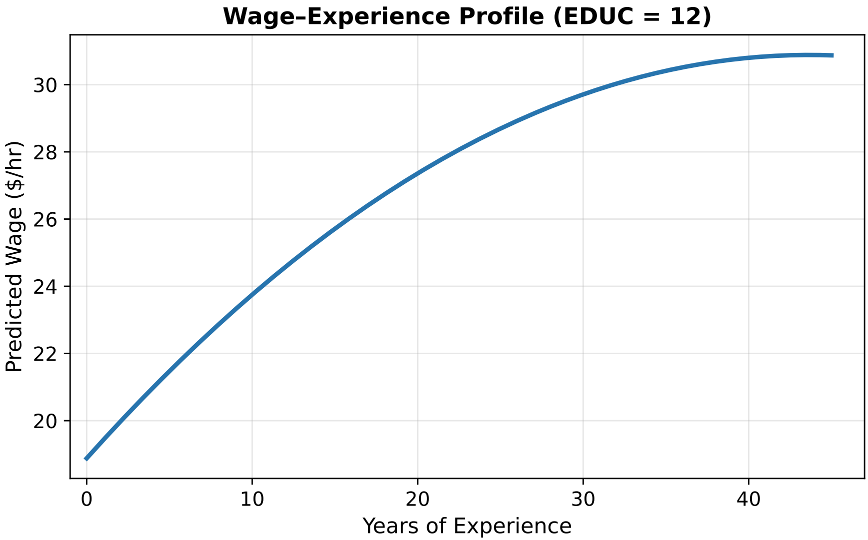
Experience level	**Marginal effect ($/hr)**
5 years	$0.55 - 0.063 = 0.487$
10 years	$0.55 - 0.126 = 0.424$
20 years	$0.55 - 0.252 = 0.298$
30 years	$0.55 - 0.378 = 0.172$


Marginal Effect: $\hat{\beta}_3 + 2\hat{\beta}_4 \times$ EXPER
Marginal Effect on Wage ($/hr per year)
Years of Experience
0.487
0.298
0.046
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
10
20
30
40

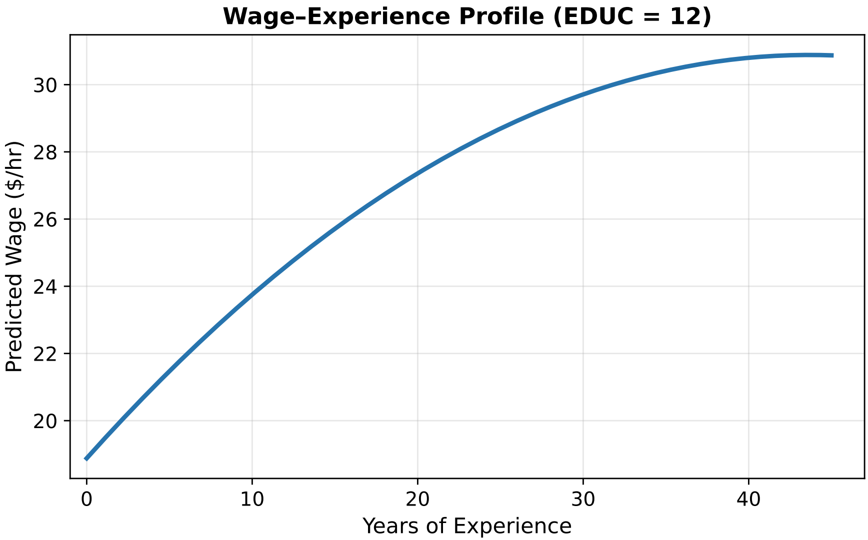
Visualizing Diminishing Returns



Visualizing the Wage–Experience Profile



Visualizing the Wage–Experience Profile



Outline

1. From SLR to MR: What Changes?
2. Individual t-Tests in MR
3. Marginal Effects with Polynomial Terms
4. Testing Linear Combinations of Coefficients
5. The General Linear Combination Formula
6. Individual $\neq$ Joint Significance
7. Summary

Why Individual t-Tests Are Not Enough

Suppose you want to know: “Is the marginal effect of experience significantly different from zero for a worker with 10 years?”

Why Individual t-Tests Are Not Enough

Suppose you want to know: "Is the marginal effect of experience significantly different from zero for a worker with 10 years?"

You might think: just look at the t-statistic for β_3. But the marginal effect at 10 years is $\beta_3 + 20\beta_4$, not β_3 alone.

Why Individual t-Tests Are Not Enough

Suppose you want to know: "Is the marginal effect of experience significantly different from zero for a worker with 10 years?"

You might think: just look at the t-statistic for β_3. But the marginal effect at 10 years is $\beta_3 + 20\beta_4$, not β_3 alone.

The individual t-test for β_3 asks a different question: "Is $\beta_3 = 0$?" That is the marginal effect at EXPER $= 0$, which is not what we want.

Why Individual t-Tests Are Not Enough

Suppose you want to know: "Is the marginal effect of experience significantly different from zero for a worker with 10 years?"

You might think: just look at the t-statistic for β_3. But the marginal effect at 10 years is $\beta_3 + 20\beta_4$, not β_3 alone.

The individual t-test for β_3 asks a different question: "Is $\beta_3 = 0$?" That is the marginal effect at EXPER $= 0$, which is not what we want.

$\implies$ To test the marginal effect at a specific experience level, we need a test for a **linear combination** of coefficients.

A New Type of Question

Individual t-tests ask: “Is this coefficient zero?”

A New Type of Question

Individual t-tests ask: "Is this coefficient zero?"

But sometimes the question involves **multiple coefficients together**:

> *"What is the return to one more year of experience for a worker who already has 10 years?"*

A New Type of Question

Individual t-tests ask: “Is this coefficient zero?”

But sometimes the question involves **multiple coefficients together**:

> *“What is the return to one more year of experience for a worker who already has 10 years?”*

The marginal effect at EXPER $= 10$ is:

$$\theta = \beta_3 + 2\beta_4 \times 10 = \beta_3 + 20\beta_4$$

A New Type of Question

Individual t-tests ask: "Is this coefficient zero?"

But sometimes the question involves **multiple coefficients together**:

> *"What is the return to one more year of experience for a worker who already has 10 years?"*

The marginal effect at EXPER $= 10$ is:

$$\theta = \beta_3 + 2\beta_4 \times 10 = \beta_3 + 20\beta_4$$

This is a **linear combination** of β_3 and β_4:

$$\theta = c_3\beta_3 + c_4\beta_4 \qquad \text{with } c_3 = 1,\ c_4 = 20$$

A New Type of Question

Individual t-tests ask: "Is this coefficient zero?"

But sometimes the question involves **multiple coefficients together**:

> *"What is the return to one more year of experience for a worker who already has 10 years?"*

The marginal effect at EXPER $= 10$ is:

$$\theta = \beta_3 + 2\beta_4 \times 10 = \beta_3 + 20\beta_4$$

This is a **linear combination** of β_3 and β_4:

$$\theta = c_3\beta_3 + c_4\beta_4 \qquad \text{with } c_3 = 1,\ c_4 = 20$$

$\Longrightarrow$ We can test $H_0 : \theta = 0$ using the same t-test framework. We just need the standard error of $\hat{\theta}$.

Standard Error of a Linear Combination

For $\hat{\theta} = c_3 b_3 + c_4 b_4$:

Standard Error of a Linear Combination

For $\hat{\theta} = c_3 b_3 + c_4 b_4$:

$$\widehat{\text{Var}}(\hat{\theta}) = c_3^2\,\widehat{\text{Var}}(b_3) + c_4^2\,\widehat{\text{Var}}(b_4) + 2c_3 c_4\,\widehat{\text{Cov}}(b_3, b_4)$$

Standard Error of a Linear Combination

For $\hat{\theta} = c_3 b_3 + c_4 b_4$:

$$\widehat{\text{Var}}(\hat{\theta}) = c_3^2 \widehat{\text{Var}}(b_3) + c_4^2 \widehat{\text{Var}}(b_4) + 2c_3 c_4 \widehat{\text{Cov}}(b_3, b_4)$$

Do not forget the covariance term. This is where students make mistakes.

Standard Error of a Linear Combination

For $\hat{\theta} = c_3 b_3 + c_4 b_4$:

$$\widehat{\text{Var}}(\hat{\theta}) = c_3^2\,\widehat{\text{Var}}(b_3) + c_4^2\,\widehat{\text{Var}}(b_4) + 2c_3 c_4\,\widehat{\text{Cov}}(b_3, b_4)$$

Do not forget the covariance term. This is where students make mistakes.

If you only use the variances, you are implicitly assuming $\text{Cov}(b_3, b_4) = 0$, which is almost never true in practice.

Standard Error of a Linear Combination

For $\hat{\theta} = c_3 b_3 + c_4 b_4$:

$$\widehat{\text{Var}}(\hat{\theta}) = c_3^2\, \widehat{\text{Var}}(b_3) + c_4^2\, \widehat{\text{Var}}(b_4) + 2 c_3 c_4\, \widehat{\text{Cov}}(b_3, b_4)$$

Do not forget the covariance term. This is where students make mistakes.

If you only use the variances, you are implicitly assuming $\text{Cov}(b_3, b_4) = 0$, which is almost never true in practice.

Where does $\widehat{\text{Cov}}(b_3, b_4)$ come from?

Standard Error of a Linear Combination

For $\hat{\theta} = c_3 b_3 + c_4 b_4$:

$$\widehat{\text{Var}}(\hat{\theta}) = c_3^2\, \widehat{\text{Var}}(b_3) + c_4^2\, \widehat{\text{Var}}(b_4) + 2c_3 c_4\, \widehat{\text{Cov}}(b_3, b_4)$$

Do not forget the covariance term. This is where students make mistakes.

If you only use the variances, you are implicitly assuming $\text{Cov}(b_3, b_4) = 0$, which is almost never true in practice.

Where does $\widehat{\text{Cov}}(b_3, b_4)$ come from?

$\implies$ The **estimated covariance matrix** of the coefficients. Software computes this for every regression. You just need to read off the right entry.

The Covariance Matrix

Software reports a matrix with variances on the diagonal and covariances off the diagonal:

The Covariance Matrix

Software reports a matrix with variances on the diagonal and covariances off the diagonal:

$$\widehat{\text{Cov}}(b_1, b_2, b_3, b_4) = \begin{bmatrix} 8.2369 & -0.3984 & -0.0639 & 0.00040 \\ -0.3984 & 0.0256 & 0.0001 & -0.00001 \\ -0.0639 & 0.0001 & 0.0121 & -0.00022 \\ 0.00040 & -0.00001 & -0.00022 & 0.0000040 \end{bmatrix}$$

The Covariance Matrix

Software reports a matrix with variances on the diagonal and covariances off the diagonal:

$$\widehat{\text{Cov}}(b_1, b_2, b_3, b_4) = \begin{bmatrix} 8.2369 & -0.3984 & -0.0639 & 0.00040 \\ -0.3984 & 0.0256 & 0.0001 & -0.00001 \\ -0.0639 & 0.0001 & 0.0121 & -0.00022 \\ 0.00040 & -0.00001 & -0.00022 & 0.0000040 \end{bmatrix}$$

Reading the matrix:

- $\widehat{\text{Var}}(b_3) = 0.0121 \implies \text{se}(b_3) = \sqrt{0.0121} = 0.11$
- $\widehat{\text{Var}}(b_4) = 0.0000040 \implies \text{se}(b_4) = 0.0020$
- $\widehat{\text{Cov}}(b_3, b_4) = -0.00022$

Testing the Marginal Effect at 10 Years

$H_0 : \beta_3 + 20\beta_4 = 0$ vs. $H_1 : \beta_3 + 20\beta_4 \neq 0$ $(\alpha = 0.05)$

Testing the Marginal Effect at 10 Years

$H_0 : \beta_3 + 20\beta_4 = 0$ vs. $H_1 : \beta_3 + 20\beta_4 \neq 0$ $(\alpha = 0.05)$

Step 1: Point estimate.

$$\hat{\theta} = b_3 + 20b_4 = 0.55 + 20(-0.0063) = 0.424$$

Testing the Marginal Effect at 10 Years

$H_0 : \beta_3 + 20\beta_4 = 0$ vs. $H_1 : \beta_3 + 20\beta_4 \neq 0$ $(\alpha = 0.05)$

Step 1: Point estimate.

$$\hat{\theta} = b_3 + 20b_4 = 0.55 + 20(-0.0063) = 0.424$$

Step 2: Standard error.

$$\begin{aligned}\widehat{\text{Var}}(\hat{\theta}) &= (1)^2(0.0121) + (20)^2(0.0000040) + 2(1)(20)(-0.00022) \\ &= 0.0121 + 0.0016 - 0.0088 \\ &= 0.0049\end{aligned}$$

$$\text{se}(\hat{\theta}) = \sqrt{0.0049} = 0.070$$

Testing the Marginal Effect at 10 Years

$H_0: \beta_3 + 20\beta_4 = 0 \quad$ vs. $\quad H_1: \beta_3 + 20\beta_4 \neq 0 \qquad (\alpha = 0.05)$

Step 1: Point estimate.

$$\hat{\theta} = b_3 + 20b_4 = 0.55 + 20(-0.0063) = 0.424$$

Step 2: Standard error.

$$\begin{aligned}\widehat{\text{Var}}(\hat{\theta}) &= (1)^2(0.0121) + (20)^2(0.0000040) + 2(1)(20)(-0.00022) \\ &= 0.0121 + 0.0016 - 0.0088 \\ &= 0.0049\end{aligned}$$

$$\text{se}(\hat{\theta}) = \sqrt{0.0049} = 0.070$$

Step 3: t-statistic.

$$t = \frac{0.424 - 0}{0.070} = 6.06$$

Testing the Marginal Effect at 10 Years

$H_0: \beta_3 + 20\beta_4 = 0 \quad \text{vs.} \quad H_1: \beta_3 + 20\beta_4 \neq 0 \qquad (\alpha = 0.05)$

Step 1: Point estimate.

$$\hat{\theta} = b_3 + 20b_4 = 0.55 + 20(-0.0063) = 0.424$$

Step 2: Standard error.

$$\begin{aligned}\widehat{\text{Var}}(\hat{\theta}) &= (1)^2(0.0121) + (20)^2(0.0000040) + 2(1)(20)(-0.00022) \\ &= 0.0121 + 0.0016 - 0.0088 \\ &= 0.0049\end{aligned}$$

$$\text{se}(\hat{\theta}) = \sqrt{0.0049} = 0.070$$

Step 3: t-statistic.

$$t = \frac{0.424 - 0}{0.070} = 6.06$$

$|6.06| \geq 1.962?$ Yes. $p \approx 0.000.$

Testing the Marginal Effect at 10 Years

$H_0: \beta_3 + 20\beta_4 = 0 \quad$ vs. $\quad H_1: \beta_3 + 20\beta_4 \neq 0 \qquad (\alpha = 0.05)$

Step 1: Point estimate.

$$\hat{\theta} = b_3 + 20b_4 = 0.55 + 20(-0.0063) = 0.424$$

Step 2: Standard error.

$$\begin{aligned}\widehat{\text{Var}}(\hat{\theta}) &= (1)^2(0.0121) + (20)^2(0.0000040) + 2(1)(20)(-0.00022) \\ &= 0.0121 + 0.0016 - 0.0088 \\ &= 0.0049\end{aligned}$$

$$\text{se}(\hat{\theta}) = \sqrt{0.0049} = 0.070$$

Step 3: t-statistic.

$$t = \frac{0.424 - 0}{0.070} = 6.06$$

$|6.06| \geq 1.962? \quad$ Yes. $\quad p \approx 0.000.$

Reject H_0. At 10 years of experience, the marginal return is significantly different from zero.

What About at 30 Years?

$H_0 : \beta_3 + 60\beta_4 = 0 \quad \text{vs.} \quad H_1 : \beta_3 + 60\beta_4 \neq 0 \qquad (\alpha = 0.05)$

What About at 30 Years?

$H_0 : \beta_3 + 60\beta_4 = 0$ vs. $H_1 : \beta_3 + 60\beta_4 \neq 0$ $(\alpha = 0.05)$

Point estimate:

$$\hat{\theta} = 0.55 + 60(-0.0063) = 0.172$$

What About at 30 Years?

$H_0: \beta_3 + 60\beta_4 = 0$ vs. $H_1: \beta_3 + 60\beta_4 \neq 0$ $(\alpha = 0.05)$

Point estimate:

$$\hat{\theta} = 0.55 + 60(-0.0063) = 0.172$$

Standard error:

$$\begin{aligned}\widehat{\text{Var}}(\hat{\theta}) &= (1)^2(0.0121) + (60)^2(0.0000040) + 2(1)(60)(-0.00022) \\ &= 0.0121 + 0.0144 - 0.0264 \\ &= 0.0001\end{aligned}$$

$$\text{se}(\hat{\theta}) = \sqrt{0.0001} = 0.010$$

What About at 30 Years?

$H_0: \beta_3 + 60\beta_4 = 0$ vs. $H_1: \beta_3 + 60\beta_4 \neq 0$ $(\alpha = 0.05)$

Point estimate:

$$\hat{\theta} = 0.55 + 60(-0.0063) = 0.172$$

Standard error:

$$\begin{aligned}\widehat{\text{Var}}(\hat{\theta}) &= (1)^2(0.0121) + (60)^2(0.0000040) + 2(1)(60)(-0.00022) \\ &= 0.0121 + 0.0144 - 0.0264 \\ &= 0.0001\end{aligned}$$

$$\text{se}(\hat{\theta}) = \sqrt{0.0001} = 0.010$$

t-statistic:

$$t = \frac{0.172}{0.010} = 17.2$$

What About at 30 Years?

$H_0 : \beta_3 + 60\beta_4 = 0 \quad \text{vs.} \quad H_1 : \beta_3 + 60\beta_4 \neq 0 \qquad (\alpha = 0.05)$

Point estimate:

$$\hat{\theta} = 0.55 + 60(-0.0063) = 0.172$$

Standard error:

$$\begin{aligned} \widehat{\text{Var}}(\hat{\theta}) &= (1)^2(0.0121) + (60)^2(0.0000040) + 2(1)(60)(-0.00022) \\ &= 0.0121 + 0.0144 - 0.0264 \\ &= 0.0001 \end{aligned}$$

$$\text{se}(\hat{\theta}) = \sqrt{0.0001} = 0.010$$

t-statistic:

$$t = \frac{0.172}{0.010} = 17.2$$

Still highly significant ($p \approx 0.000$). Even at 30 years, each additional year of experience still raises wages.

Comparison: How the Test Changes with Experience

EXPER	$\hat{\theta}$	se($\hat{\theta}$)	t	**Significant?**
5	0.487	0.090	5.41	Yes
10	0.424	0.070	6.06	Yes
20	0.298	0.030	9.93	Yes
30	0.172	0.010	17.2	Yes
40	0.046	0.050	0.92	No

Comparison: How the Test Changes with Experience

EXPER	$\hat{\theta}$	se($\hat{\theta}$)	t	**Significant?**
5	0.487	0.090	5.41	Yes
10	0.424	0.070	6.06	Yes
20	0.298	0.030	9.93	Yes
30	0.172	0.010	17.2	Yes
40	0.046	0.050	0.92	No

$\Longrightarrow$ The marginal return to experience shrinks as experience grows. At 40 years, the estimated return is \$0.05/hr and is no longer distinguishable from zero.

Comparison: How the Test Changes with Experience

EXPER	$\hat{\theta}$	se($\hat{\theta}$)	t	**Significant?**
5	0.487	0.090	5.41	Yes
10	0.424	0.070	6.06	Yes
20	0.298	0.030	9.93	Yes
30	0.172	0.010	17.2	Yes
40	0.046	0.050	0.92	No

$\implies$ The marginal return to experience shrinks as experience grows. At 40 years, the estimated return is $0.05/hr and is no longer distinguishable from zero.

Notice the standard error is **not constant** across experience levels. It depends on the covariance matrix and the weights c_3, c_4. The standard error is smallest near the center of the data and rises at extreme values.

Outline

1. From SLR to MR: What Changes?
2. Individual t-Tests in MR
3. Marginal Effects with Polynomial Terms
4. Testing Linear Combinations of Coefficients
5. The General Linear Combination Formula
6. Individual $\neq$ Joint Significance
7. Summary

General Setup

Suppose the quantity of interest is a linear combination of any number of coefficients:

$$\lambda = c_1\beta_1 + c_2\beta_2 + \cdots + c_K\beta_K$$

General Setup

Suppose the quantity of interest is a linear combination of any number of coefficients:

$$\lambda = c_1\beta_1 + c_2\beta_2 + \cdots + c_K\beta_K$$

Estimated by: $\hat{\lambda} = c_1 b_1 + c_2 b_2 + \cdots + c_K b_K$

General Setup

Suppose the quantity of interest is a linear combination of any number of coefficients:

$$\lambda = c_1\beta_1 + c_2\beta_2 + \cdots + c_K\beta_K$$

Estimated by: $\hat{\lambda} = c_1 b_1 + c_2 b_2 + \cdots + c_K b_K$

Variance:

$$\widehat{\text{Var}}(\hat{\lambda}) = \sum_{k=1}^{K} c_k^2\,\widehat{\text{Var}}(b_k) + 2\sum_{j<k} c_j\,c_k\,\widehat{\text{Cov}}(b_j, b_k)$$

General Setup

Suppose the quantity of interest is a linear combination of any number of coefficients:

$$\lambda = c_1\beta_1 + c_2\beta_2 + \cdots + c_K\beta_K$$

Estimated by: $\hat{\lambda} = c_1 b_1 + c_2 b_2 + \cdots + c_K b_K$

Variance:

$$\widehat{\text{Var}}(\hat{\lambda}) = \sum_{k=1}^{K} c_k^2 \, \widehat{\text{Var}}(b_k) + 2\sum_{j<k} c_j \, c_k \, \widehat{\text{Cov}}(b_j, b_k)$$

Test statistic:

$$t = \frac{\hat{\lambda} - \lambda_0}{\text{se}(\hat{\lambda})} \sim t_{(N-K)} \quad \text{under } H_0 : \lambda = \lambda_0$$

General Setup

Suppose the quantity of interest is a linear combination of any number of coefficients:

$$\lambda = c_1\beta_1 + c_2\beta_2 + \cdots + c_K\beta_K$$

Estimated by: $\hat{\lambda} = c_1 b_1 + c_2 b_2 + \cdots + c_K b_K$

Variance:

$$\widehat{\text{Var}}(\hat{\lambda}) = \sum_{k=1}^{K} c_k^2\,\widehat{\text{Var}}(b_k) + 2\sum_{j<k} c_j\,c_k\,\widehat{\text{Cov}}(b_j, b_k)$$

Test statistic:

$$t = \frac{\hat{\lambda} - \lambda_0}{\text{se}(\hat{\lambda})} \sim t_{(N-K)} \quad \text{under } H_0 : \lambda = \lambda_0$$

$\implies$ This is the same t-test structure. The only new ingredient is computing $\text{se}(\hat{\lambda})$ from the covariance matrix.

Example: Predicted Wage Difference

How much more does a worker with 16 years of education earn than one with 12 years, holding experience constant?

Example: Predicted Wage Difference

How much more does a worker with 16 years of education earn than one with 12 years, holding experience constant?

$$\Delta \text{WAGE} = \beta_2 \times (16 - 12) = 4\beta_2$$

Example: Predicted Wage Difference

How much more does a worker with 16 years of education earn than one with 12 years, holding experience constant?

$$\Delta \text{WAGE} = \beta_2 \times (16 - 12) = 4\beta_2$$

This is a linear combination with $c_2 = 4$ and all other $c_k = 0$.

Example: Predicted Wage Difference

How much more does a worker with 16 years of education earn than one with 12 years, holding experience constant?

$$\Delta \text{WAGE} = \beta_2 \times (16 - 12) = 4\beta_2$$

This is a linear combination with $c_2 = 4$ and all other $c_k = 0$.

Point estimate: $4 \times 2.44 = 9.76$ ($9.76/hr)

Example: Predicted Wage Difference

How much more does a worker with 16 years of education earn than one with 12 years, holding experience constant?

$$\Delta \text{WAGE} = \beta_2 \times (16 - 12) = 4\beta_2$$

This is a linear combination with $c_2 = 4$ and all other $c_k = 0$.

Point estimate: $4 \times 2.44 = 9.76$ (\$9.76/hr)

Standard error: $4 \times \text{se}(b_2) = 4 \times 0.16 = 0.64$

Example: Predicted Wage Difference

How much more does a worker with 16 years of education earn than one with 12 years, holding experience constant?

$$\Delta \text{WAGE} = \beta_2 \times (16 - 12) = 4\beta_2$$

This is a linear combination with $c_2 = 4$ and all other $c_k = 0$.

Point estimate: $4 \times 2.44 = 9.76$ (\$9.76/hr)

Standard error: $4 \times \text{se}(b_2) = 4 \times 0.16 = 0.64$

95% CI: $9.76 \pm 1.962 \times 0.64 = [8.50,\ 11.02]$

Example: Predicted Wage Difference

How much more does a worker with 16 years of education earn than one with 12 years, holding experience constant?

$$\Delta\text{WAGE} = \beta_2 \times (16 - 12) = 4\beta_2$$

This is a linear combination with $c_2 = 4$ and all other $c_k = 0$.

Point estimate: $4 \times 2.44 = 9.76$ (\$9.76/hr)

Standard error: $4 \times \text{se}(b_2) = 4 \times 0.16 = 0.64$

95% CI: $9.76 \pm 1.962 \times 0.64 = [8.50,\ 11.02]$

$\Longrightarrow$ Four extra years of college are associated with \$8.50 to \$11.02 more per hour. No covariance terms needed here because only one coefficient is involved.

Outline

1. From SLR to MR: What Changes?
2. Individual t-Tests in MR
3. Marginal Effects with Polynomial Terms
4. Testing Linear Combinations of Coefficients
5. The General Linear Combination Formula
6. Individual $\neq$ Joint Significance
7. Summary

A Surprising Possibility

Suppose you run a regression with EXPER and $EXPER^2$, and the results are:

A Surprising Possibility

Suppose you run a regression with EXPER and EXPER^2, and the results are:

Variable	Coefficient	Std. Error	t-stat
EXPER	0.12	0.08	1.50
EXPER^2	-0.001	0.001	-1.00

A Surprising Possibility

Suppose you run a regression with EXPER and EXPER^2, and the results are:

Variable	Coefficient	Std. Error	t-stat
EXPER	0.12	0.08	1.50
EXPER^2	-0.001	0.001	-1.00

Neither is individually significant at $\alpha = 0.05$.

A Surprising Possibility

Suppose you run a regression with EXPER and EXPER^2, and the results are:

Variable	Coefficient	Std. Error	t-stat
EXPER	0.12	0.08	1.50
EXPER^2	-0.001	0.001	-1.00

Neither is individually significant at $\alpha = 0.05$.

Question: Does this mean experience has no effect on wages?

A Surprising Possibility

Suppose you run a regression with EXPER and EXPER^2, and the results are:

Variable	Coefficient	Std. Error	t-stat
EXPER	0.12	0.08	1.50
EXPER^2	-0.001	0.001	-1.00

Neither is individually significant at $\alpha = 0.05$.

Question: Does this mean experience has no effect on wages?

No. EXPER and EXPER^2 are highly correlated ($r > 0.95$ in most datasets), which inflates both standard errors.

A Surprising Possibility

Suppose you run a regression with EXPER and EXPER^2, and the results are:

Variable	Coefficient	Std. Error	t-stat
EXPER	0.12	0.08	1.50
EXPER^2	-0.001	0.001	-1.00

Neither is individually significant at $\alpha = 0.05$.

Question: Does this mean experience has no effect on wages?

No. EXPER and EXPER^2 are highly correlated ($r > 0.95$ in most datasets), which inflates both standard errors.

$\Longrightarrow$ Each variable's t-test asks: "Does this variable contribute, *given that the other is already in the model*?"

The Problem with One-at-a-Time Testing

When regressors are correlated, individual t-tests do not answer: "Do these variables **jointly** affect y?"

The Problem with One-at-a-Time Testing

When regressors are correlated, individual t-tests do not answer: “Do these variables **jointly** affect y?”

Consider: EXPER and EXPER^2 are **jointly** capturing the curvature of the wage–experience profile. Dropping one makes the other look different.

The Problem with One-at-a-Time Testing

When regressors are correlated, individual t-tests do not answer: "Do these variables **jointly** affect y?"

Consider: EXPER and EXPER^2 are **jointly** capturing the curvature of the wage–experience profile. Dropping one makes the other look different.

What we need is a test for:

$$H_0 : \beta_3 = 0 \text{ AND } \beta_4 = 0$$

$$H_1 : \beta_3 \neq 0 \text{ OR } \beta_4 \neq 0 \text{ (or both)}$$

The Problem with One-at-a-Time Testing

When regressors are correlated, individual t-tests do not answer: "Do these variables **jointly** affect y?"

Consider: EXPER and EXPER^2 are **jointly** capturing the curvature of the wage–experience profile. Dropping one makes the other look different.

What we need is a test for:

$$H_0 : \beta_3 = 0 \text{ AND } \beta_4 = 0$$

$$H_1 : \beta_3 \neq 0 \text{ OR } \beta_4 \neq 0 \text{ (or both)}$$

This is a **joint hypothesis**, and the t-test cannot handle it.

The Problem with One-at-a-Time Testing

When regressors are correlated, individual t-tests do not answer: “Do these variables **jointly** affect y?”

Consider: EXPER and EXPER^2 are **jointly** capturing the curvature of the wage–experience profile. Dropping one makes the other look different.

What we need is a test for:

$$H_0 : \beta_3 = 0 \text{ AND } \beta_4 = 0$$

$$H_1 : \beta_3 \neq 0 \text{ OR } \beta_4 \neq 0 \text{ (or both)}$$

This is a **joint hypothesis**, and the t-test cannot handle it.

$\Longrightarrow$ We need the ***F*-test**, which is the topic for next lecture. For now, remember: individually insignificant does **not** mean jointly insignificant.

Preview: Why the F-Test Exists

Question	**Test**	**Statistic**
Is one coefficient zero?	t-test	$t = \frac{b_k}{\text{se}(b_k)}$
Are *several* coefficients zero?	F-test	Next lecture

Preview: Why the F-Test Exists

Question	**Test**	**Statistic**
Is one coefficient zero?	t-test	$t = \dfrac{b_k}{\text{se}(b_k)}$
Are *several* coefficients zero?	F-test	Next lecture

Today's toolkit:

- Individual t-tests (same as SLR, with $N - K$ df)
- Linear combination tests (new: uses covariance matrix)

Preview: Why the F-Test Exists

Question	Test	Statistic
Is one coefficient zero?	t-test	$t = \frac{b_k}{se(b_k)}$
Are *several* coefficients zero?	F-test	Next lecture

Today's toolkit:

- Individual t-tests (same as SLR, with $N - K$ df)
- Linear combination tests (new: uses covariance matrix)

Next lecture's addition:

- F-test for joint significance (tests multiple restrictions at once)

Preview: Why the F-Test Exists

Question	Test	Statistic
Is one coefficient zero?	t-test	$t = \dfrac{b_k}{\text{se}(b_k)}$
Are *several* coefficients zero?	F-test	Next lecture

Today's toolkit:

- Individual t-tests (same as SLR, with $N - K$ df)
- Linear combination tests (new: uses covariance matrix)

Next lecture's addition:

- F-test for joint significance (tests multiple restrictions at once)

$\Longrightarrow$ The t-test handles one constraint. The F-test handles many.

Outline

1. From SLR to MR: What Changes?
2. Individual t-Tests in MR
3. Marginal Effects with Polynomial Terms
4. Testing Linear Combinations of Coefficients
5. The General Linear Combination Formula
6. Individual $\neq$ Joint Significance
7. Summary

What Changed from SLR to MR

	SLR	**MR**
t-statistic formula	$\dfrac{b_k - c}{\text{se}(b_k)}$	Same
Degrees of freedom	$N - 2$	$N - K$
Test individual coefficients?	Yes	Yes
Test linear combinations?	Not needed	Yes (uses cov matrix)
Joint tests?	N/A	Need F-test

What Changed from SLR to MR

	SLR	MR
t-statistic formula	$\frac{b_k - c}{\text{se}(b_k)}$	Same
Degrees of freedom	$N - 2$	$N - K$
Test individual coefficients?	Yes	Yes
Test linear combinations?	Not needed	Yes (uses cov matrix)
Joint tests?	N/A	Need F-test

Polynomial terms (like EXPER^2):

- The marginal effect is a **derivative**, not a single coefficient
- Testing the marginal effect at a specific point requires a linear combination test
- Individual t-tests on correlated terms can both be insignificant even when the variables are jointly significant

Thank you!

jakeanderson@g.ucla.edu