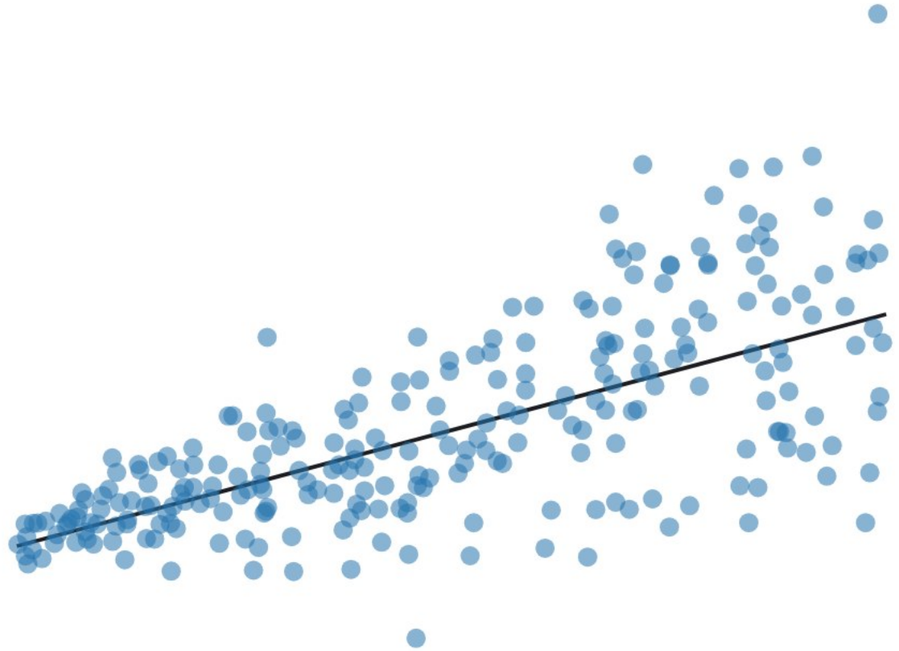
ECON 104 · DISCUSSION · FALL 2026

Heteroskedasticity

Jake Anderson
UCLA



Plan for Today

1

Breusch-Pagan test

Does the error variance change with the regressors?

Regress $\hat{\epsilon}_i^2$ on the regressors, then compare $N \times R^2_{\text{aux}}$ to a χ^2 critical value.

Spring 2026 Midterm (Boswell), Question 17

2

Goldfeld-Quandt test

Is the error variance larger at one end of the data?

Sort, split into two groups, then compare the two error variances with an F test.

Spring 2026 Final (Boswell), Question 29

3

OLS vs WLS

What do similar point estimates and different standard errors tell us?

Both estimators are unbiased. Only the variances differ.

Spring 2026 Midterm (Boswell), Question 19

The Question: Fuel Consumption and Truck Weight

A logistics company analyzes fuel consumption Y_i based on truck weight X_i:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

The Breusch-Pagan auxiliary regression is:

```
Call: lm(formula = I(eh^2) ~ weight)
            Estimate  Std. Error  t value  Pr(>|t|)
(Intercept) 0.77682   0.05336     14.56    < 2e-16   ***
weight      0.22286   0.05422      4.11    4.63e-05  ***
Residual standard error: 1.193 on 498 degrees of freedom
Multiple R-squared: 0.03281, Adjusted R-squared: 0.03086
F-statistic: 16.89 on 1 and 498 DF, p-value: 4.627e-05
```

If $N = 500$ and the 5% critical value is $\chi^2_{(1)} = 3.84$, what is your conclusion?

a Reject H_0; LM $= 16.4$, so heteroskedasticity is present.

b Fail to reject H_0; LM $= 0.035$, not enough evidence to show homoskedasticity is present with respect to weight.

c Reject H_0; LM $= 3.84$, so heteroskedasticity is present.

d Fail to reject H_0; LM $= 16.4$, not enough evidence to show homoskedasticity is present with respect to weight.

e Cannot conclude; need the F-statistic instead.

Source: Spring 2026 Midterm (Boswell), Question 17

Does the Spread Around the Line Increase with Weight?

THE DATA

Fuel consumption Y_i against truck weight X_i, with the OLS line $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$.

SPREAD AROUND THE LINE

Each bar runs two standard deviations above and below the line, so it covers about 95% of the trucks at that weight. The bars get longer as weight rises.

GOAL

Decide whether the spread increases with weight. We can see it in the picture, but we want a formal rule that answers yes or no. The Breusch-Pagan test is that rule.

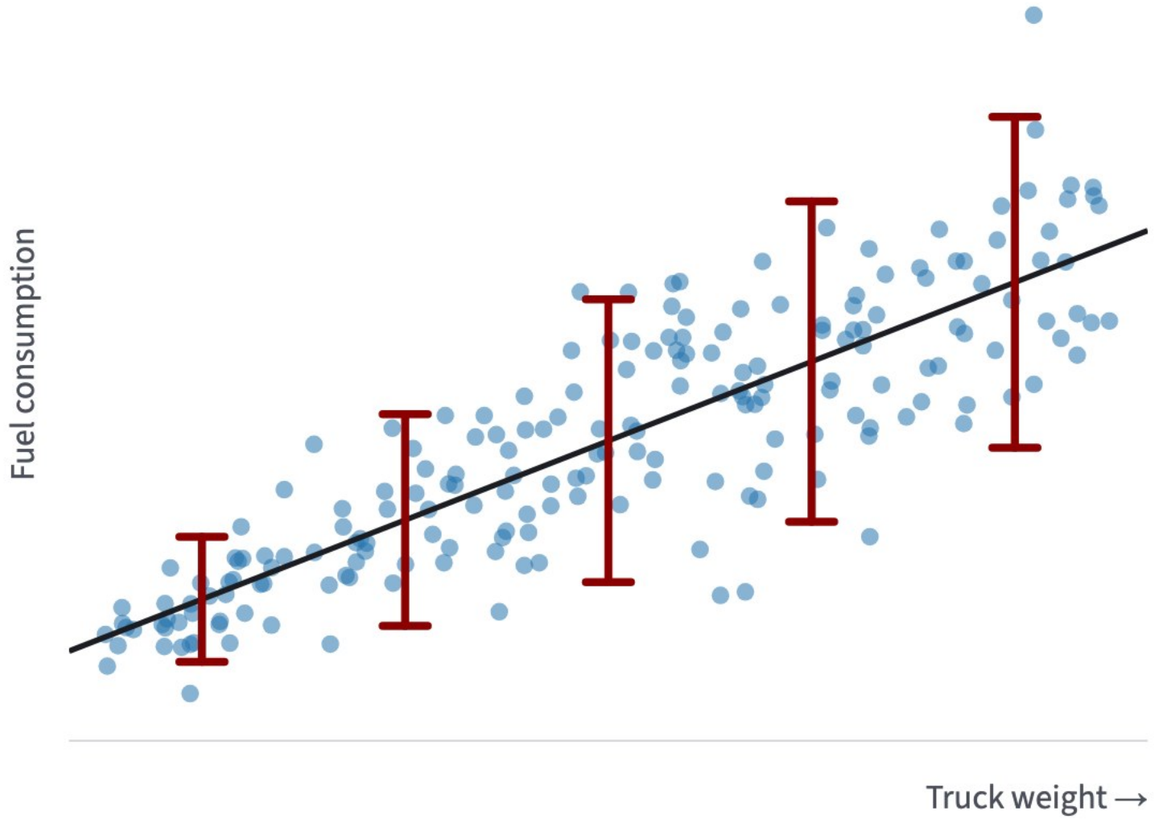


Illustration: 200 trucks whose error variance rises with weight.

Regress the Squared Residuals on Weight

PLOT THE RESIDUALS

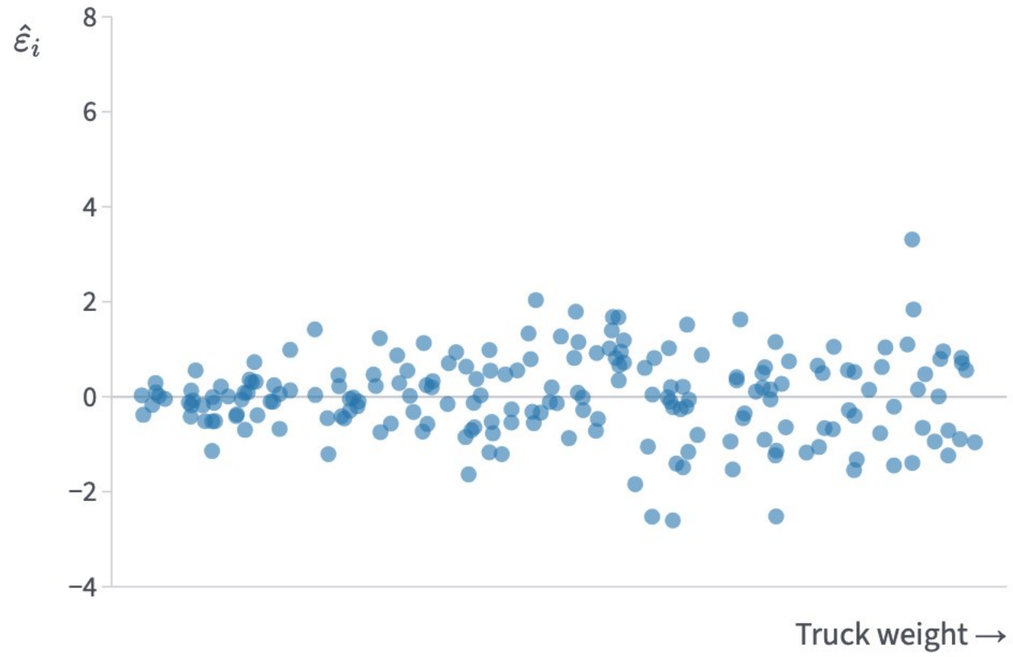
Plot $\hat{\varepsilon}_i$ against weight. The OLS line becomes the zero line; the spread is unchanged.

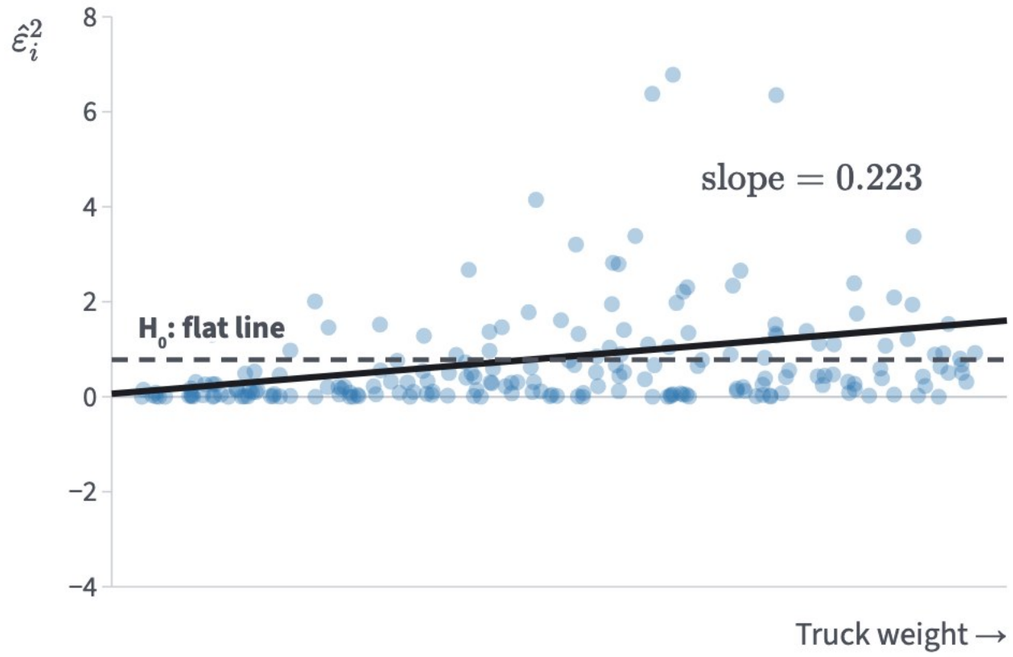
SQUARE THEM

$E(\varepsilon_i^2) = \sigma_i^2$, so $\hat{\varepsilon}_i^2$ tracks the error variance of truck i.

AUXILIARY REGRESSION

Regress $\hat{\varepsilon}_i^2$ on weight: $\hat{\varepsilon}_i^2 = 0.777 + 0.223\,\text{weight}_i$





H_0: $\alpha_2 = 0$ ⟺ flat line ⟺ homoskedastic H_1: $\alpha_2 \neq 0$ ⟺ the variance changes with weight

Compute the Test Statistic

```
Call: lm(formula = I(eh^2) ~ weight)
            Estimate  Std. Error  t value  Pr(>|t|)
(Intercept) 0.77682   0.05336     14.56    < 2e-16   ***
weight      0.22286   0.05422      4.11    4.63e-05  ***
Residual standard error: 1.193 on 498 degrees of freedom
Multiple R-squared: 0.03281, Adjusted R-squared: 0.03086
F-statistic: 16.89 on 1 and 498 DF, p-value: 4.627e-05
```

`lm( )`	linear model: estimate by OLS
`eh`	"e hat": the residuals $\hat{\varepsilon}_i$ saved from the original regression
`I(eh^2)`	square the residuals: the dependent variable is $\hat{\varepsilon}_i^2$
`~ weight`	regressed on weight $\Longrightarrow$ this is the auxiliary regression

> **On the exam:** output like this may appear without any explanation. Recognize `I(eh^2) ~ weight` as the Breusch-Pagan auxiliary regression.

> Use the **Multiple** R^2, not the adjusted R^2.

FORMULA

$$LM = N \times R^2_{\text{aux}} \sim \chi^2_{(S-1)}$$

PLUG IN

$$LM = 500 \times 0.03281 = 16.4$$

DEGREES OF FREEDOM

$S - 1$ counts the regressors in the auxiliary regression. The intercept is not counted. One regressor (weight) $\Longrightarrow$ $\chi^2_{(1)}$.

Compare LM to the $\chi^2_{(1)}$ Distribution

UNDER H_0

LM follows a $\chi^2_{(1)}$ distribution. Most draws are small.

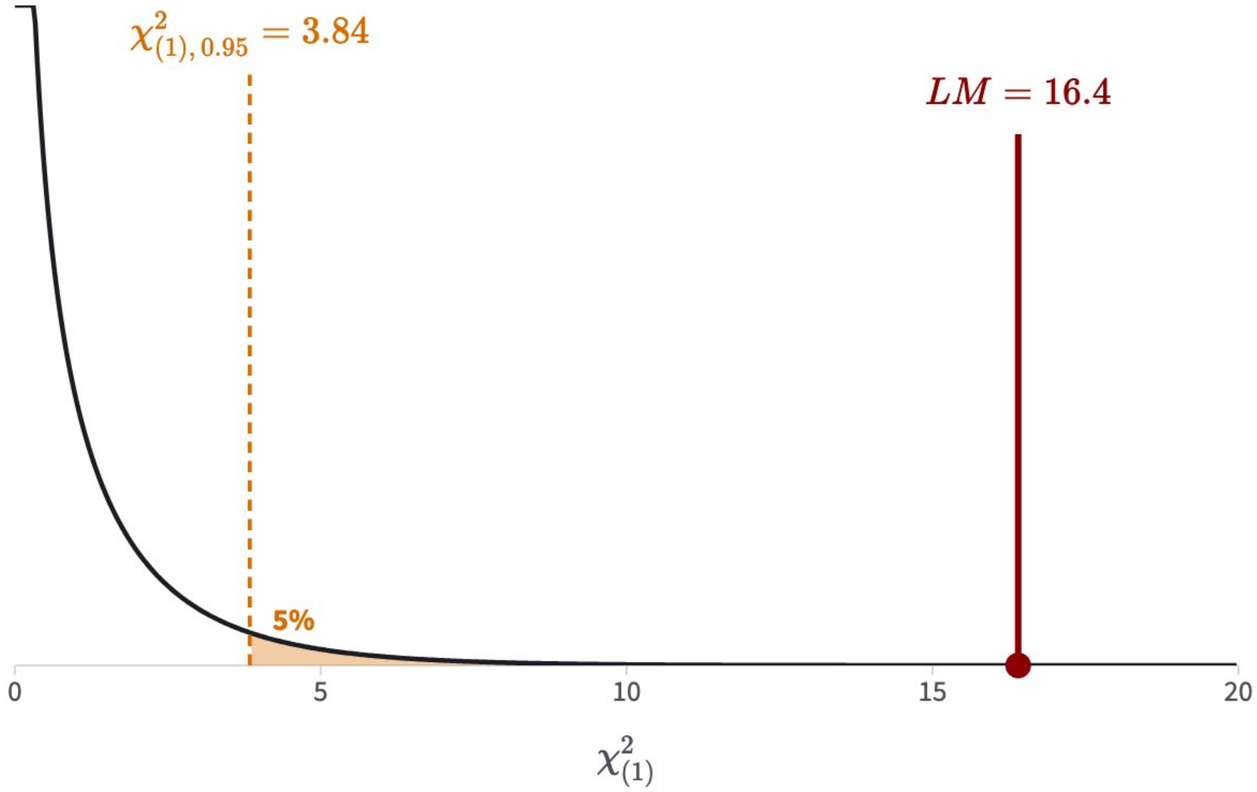
REJECTION REGION

Only 5% of the distribution lies above 3.84.

OUR STATISTIC

$LM = 16.4$ lies far inside the rejection region ($p \approx 0.00005$).

REJECT H_0 $16.4 > 3.84 \implies$ the error variance changes with truck weight.



Checking Each Answer Choice

b Fail to reject H_0; LM $= 0.035$, not enough evidence to show homoskedasticity is present with respect to weight.

0.035 is about R^2 itself. LM multiplies it by $N = 500$.

c Reject H_0; LM $= 3.84$, so heteroskedasticity is present.

3.84 is the critical value, not the test statistic.

d Fail to reject H_0; LM $= 16.4$, not enough evidence to show homoskedasticity is present with respect to weight.

Right statistic, wrong decision: $16.4 > 3.84$.

e Cannot conclude; need the F-statistic instead.

The F-statistic is printed and agrees: $F = t^2 = 4.11^2 \approx 16.9, p = 4.6 \times 10^{-5}$.

a Reject H_0; LM $= 16.4$, so heteroskedasticity is present.

Answer: (a)

WORDING

Failing to reject would mean **no evidence of heteroskedasticity**, not "evidence of homoskedasticity."

INTERPRETATION

The slope on weight is positive, so the error variance **rises** with truck weight. The usual OLS standard errors are invalid $\Longrightarrow$ use robust standard errors, or model the variance with WLS.

The Question: R&D Spending and Firm Size

The original regression: $\text{r\&d}_i = \alpha + \beta\,\text{size}_i + \gamma\,X_i + \varepsilon_i$

A researcher estimates $\widehat{\text{r\&d}}_i = \hat{\alpha} + \hat{\beta}\,\text{size}_i + \hat{\gamma}\,X_i$ to find the determinants of R&D spending, using a cross-section of 40 companies. Observations are sorted by size, the middle 4 observations dropped, and OLS is estimated separately on the large size group ($RSS_1 = 13{,}984.61$) and the small size group ($RSS_2 = 5{,}708.51$).

At the 5% level, $F_{(18,18),\,0.95} = 2.217$. If the GQ test is carried out, the correct conclusion is:

a Reject H_0; the large size group has significantly larger variance.

b Fail to reject H_0; there is insufficient evidence of heteroskedasticity of this form.

c Reject H_0; the model is misspecified because the slope differs across groups.

d Fail to reject H_0; the error term is homoskedastic with certainty.

e Reject H_0; the small size group has significantly larger variance.

Source: Spring 2026 Final (Boswell), Question 29

Is the Spread Around the Line Larger for Large Firms?

THE DATA

R&D spending against firm size for 40 companies, with the OLS line.

SPREAD AT EACH END

Each bar runs two standard deviations above and below the line. The bar for the smallest firms is short; the bar for the largest firms is long.

GOAL

Decide whether the spread is larger among large firms. The Goldfeld-Quandt test answers yes or no by comparing the error variance of the smallest firms with the error variance of the largest firms.

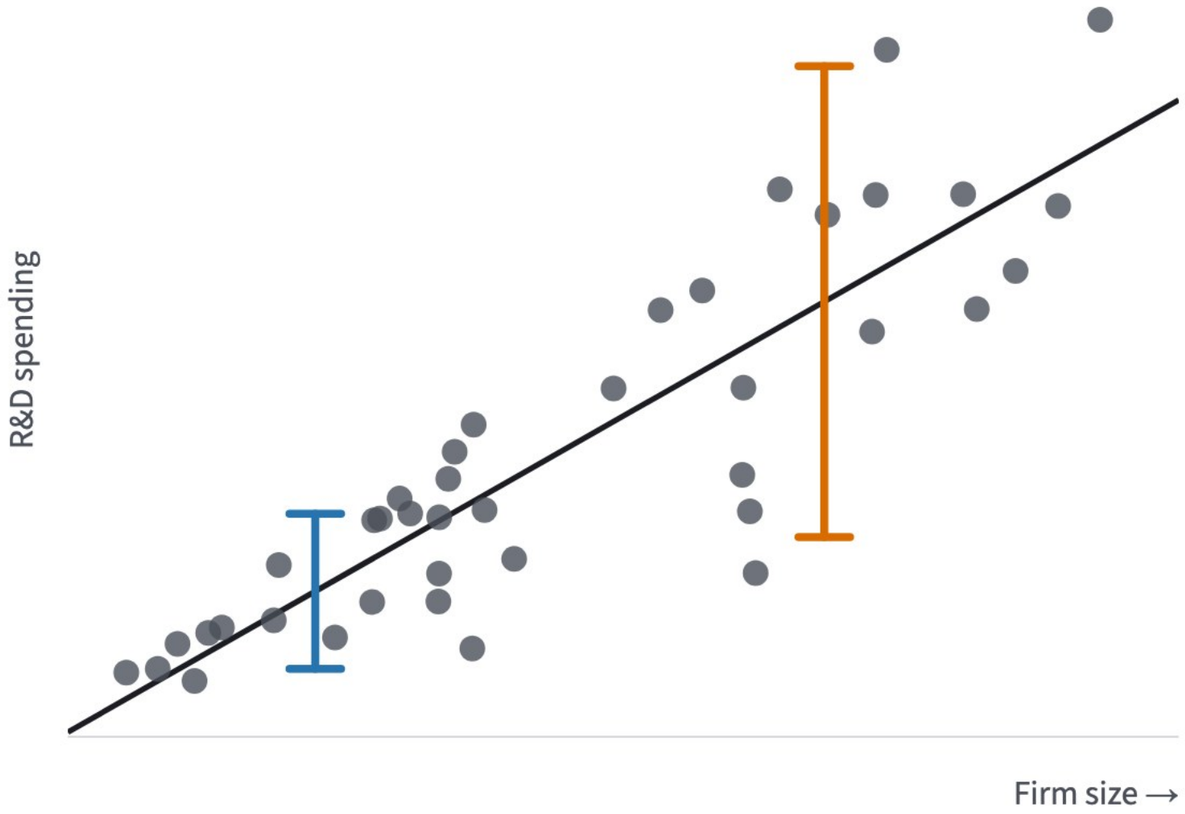


Illustration: 40 firms whose error variance rises with size.

GOLDFELD-QUANDT · STEP 1

Sort by Size, Drop the Middle, Fit Each Group

START

40 firms, in the order they appear in the data.

SORT

Order the firms by size, the variable we suspect drives the variance.

DROP THE MIDDLE

Remove the middle 4 firms. This leaves 18 small and 18 large firms.

FIT EACH GROUP

Run the regression separately on each group.

RESIDUAL SUMS OF SQUARES

Add up the squared residuals in each group: $RSS_2 = 5{,}708.51$ (small), $RSS_1 = 13{,}984.61$ (large).

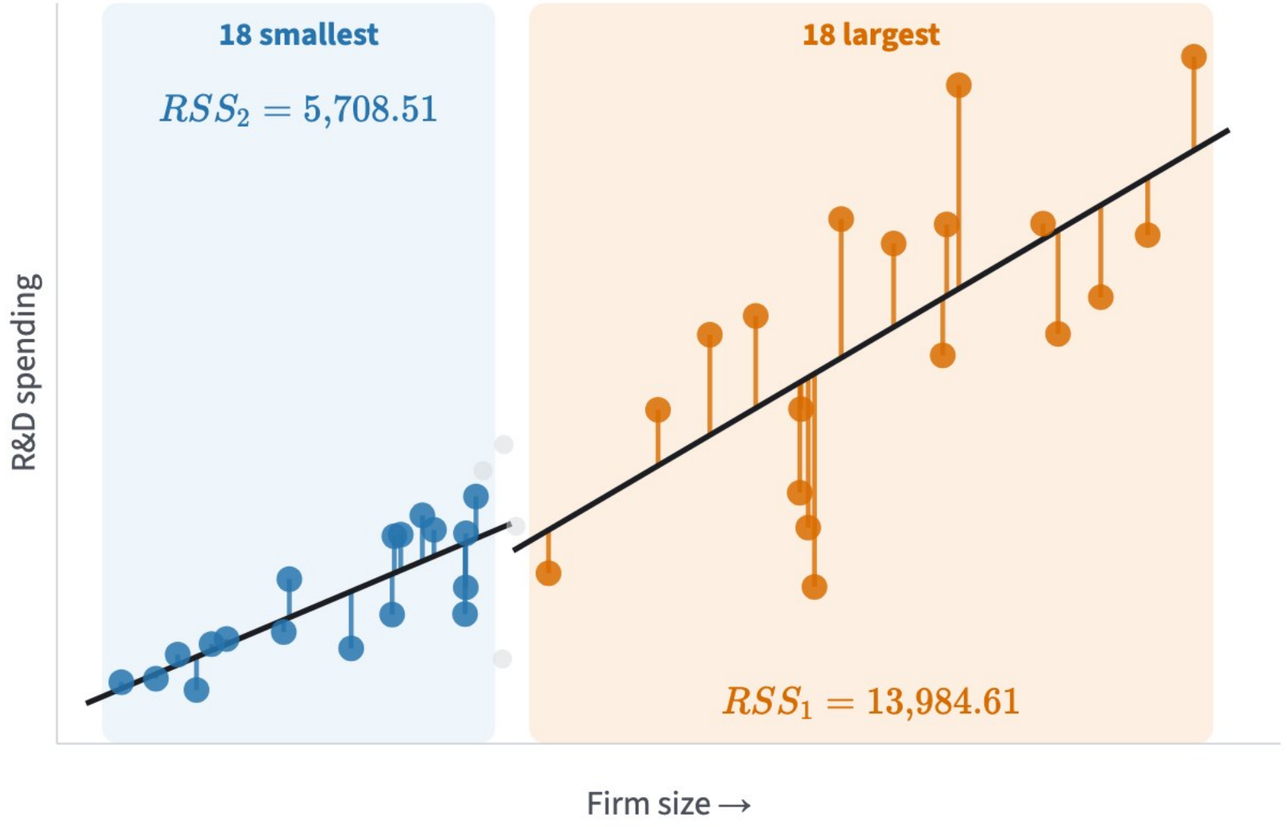


Illustration: 40 firms whose error variance rises with size, plotted against size only.

Compute the Test Statistic

RSS residual sum of squares: $RSS = \sum \hat{\varepsilon}_i^2$, added up within one group

RSS₁ the 18 largest firms: $13{,}984.61$

RSS₂ the 18 smallest firms: $5{,}708.51$

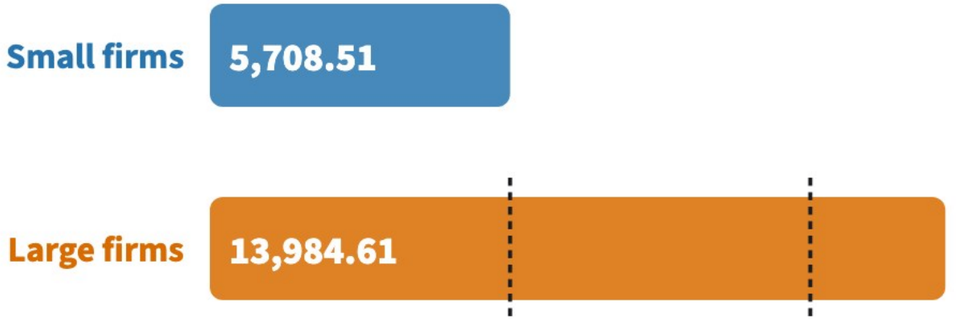
> **On the exam:** the same quantity may be called the sum of squared errors (SSE) or the sum of squared residuals. It is always $\sum \hat{\varepsilon}_i^2$.

HYPOTHESES

$H_0 \colon \sigma^2_{\text{large}} = \sigma^2_{\text{small}}$ vs $H_1 \colon \sigma^2_{\text{large}} > \sigma^2_{\text{small}}$

VARIANCE IN EACH GROUP

Divide by the degrees of freedom: $\hat{\sigma}_j^2 = RSS_j / \text{df}_j$.



STATISTIC

Put the group suspected of the larger variance in the numerator. Both groups have the same degrees of freedom, so they cancel:

$$GQ = \frac{RSS_1/\text{df}}{RSS_2/\text{df}} = \frac{13{,}984.61}{5{,}708.51} = 2.450$$

> Each group has $18 - 3 = 15$ degrees of freedom (18 firms, 3 coefficients) $\Longrightarrow$ $GQ \sim F_{(15,15)}$ under H_0. The exam printed $F_{(18,18),\,0.95} = 2.217$, which uses the wrong degrees of freedom. The decision is the same with either critical value.

Compare GQ to the $F_{(15,15)}$ Distribution

UNDER H_0

The two variances are equal, so the ratio is close to 1.

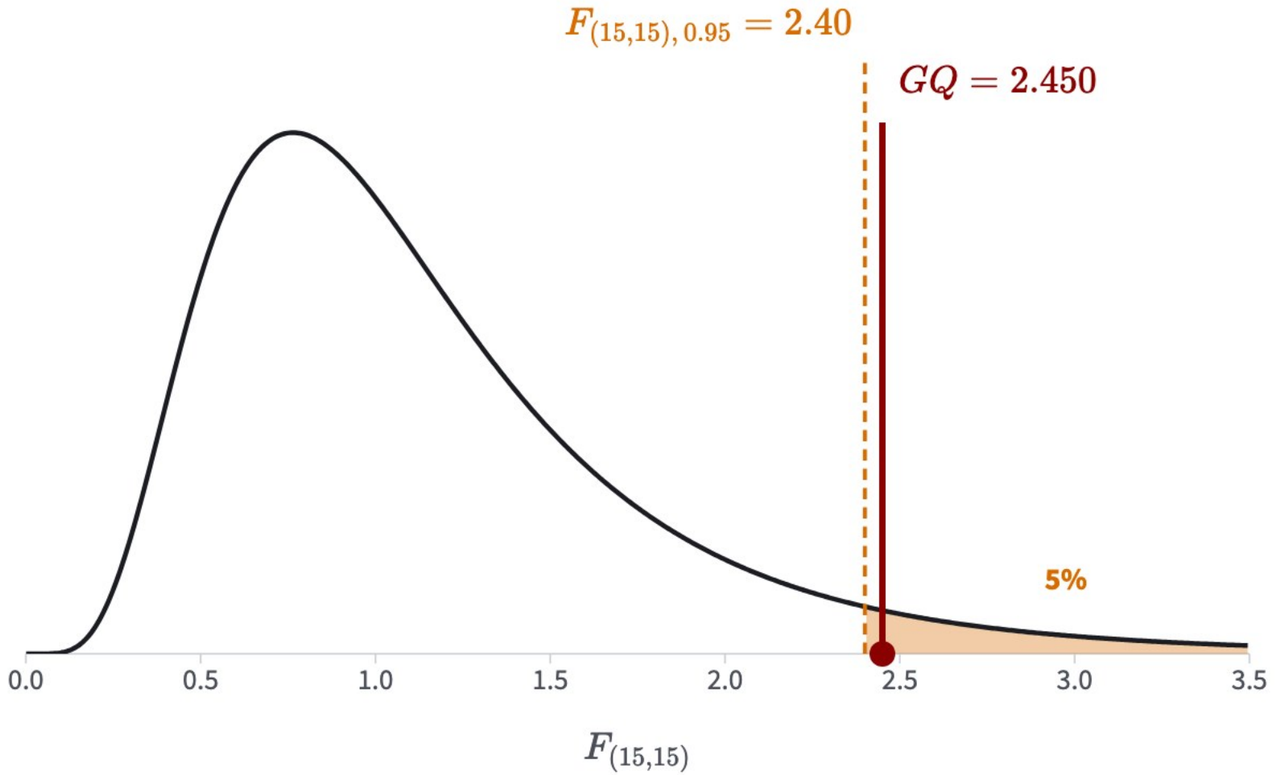
REJECTION REGION

Only 5% of the distribution lies above 2.40.

OUR STATISTIC

$GQ = 2.450$: the large-firm variance is about 2.5 times the small-firm variance.

REJECT H_0 $2.450 > 2.40 \implies$ the error variance rises with firm size.



Checking Each Answer Choice

b Fail to reject H_0; there is insufficient evidence of heteroskedasticity of this form.

The original answer key printed (b), but 2.450 exceeds both 2.40 and 2.217.

c Reject H_0; the model is misspecified because the slope differs across groups.

GQ compares error variances, not slopes.

d Fail to reject H_0; the error term is homoskedastic with certainty.

A test never establishes H_0 "with certainty."

e Reject H_0; the small size group has significantly larger variance.

The large group is in the numerator, and the ratio is above 1.

a Reject H_0; the large size group has significantly larger variance.

Answer: (a)

INTERPRETATION

R&D spending is more variable among large firms than among small firms. The usual OLS standard errors are invalid $\Longrightarrow$ use robust standard errors, or WLS with a variance that rises with size.

The Question: Similar Estimates, Different Standard Errors

The original regression: $\text{wage}_i = \beta_0 + \beta_1 \, \text{educ}_i + \varepsilon_i$

An economist estimates a cross-sectional wage model and suspects heteroskedasticity. Comparing OLS and WLS (weighted by firm size), the slope on education is $\hat{\beta}_{OLS} = 1.42$ with standard error 0.31, while $\hat{\beta}_{WLS} = 1.39$ with standard error 0.180. A colleague argues that because the two point estimates are nearly identical, heteroskedasticity is not a concern and OLS should be preferred for simplicity. Which best identifies the flaw in this reasoning?

a Similarity in point estimates does not imply OLS standard errors are valid; heteroskedasticity leaves OLS coefficients unbiased but renders standard errors inefficient, and the narrower WLS standard error reflects a genuine efficiency gain.

b The colleague is correct; when slope estimates are similar across OLS and WLS, heteroskedasticity is confirmed absent and OLS is BLUE.

c WLS always produces smaller standard errors than OLS by construction, so the difference in standard errors carries no informational content.

d WLS estimates are always biased when the variance function is misspecified, making OLS preferable under any uncertainty about the form of heteroskedasticity.

e OLS is preferred here because WLS is only appropriate when the sample size exceeds 1,000 observations.

Source: Spring 2026 Midterm (Boswell), Question 19

Whose Advice Do You Trust?

THE QUESTION

How many drinks should you have tonight? You ask 8 friends: 4 responsible friends and 4 party animals.

TREAT EVERYONE EQUALLY

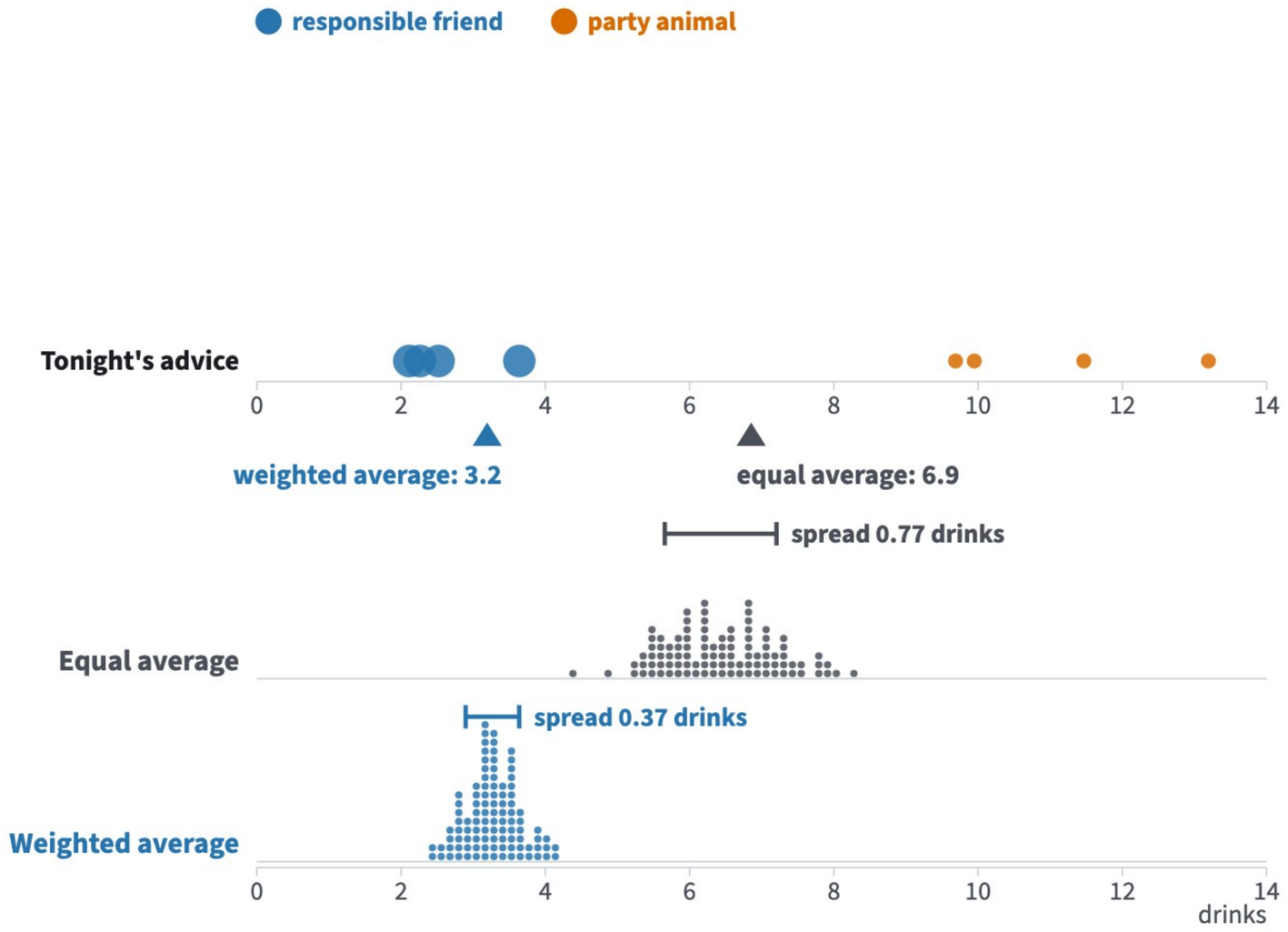
Average all 8 answers: **6.5** drinks. The party animals pull the average up, and you overdo it.

REWEIGHT THE ADVICE

The party animals' answers swing far from sensible, so give them less weight (smaller dots). The weighted average is **3.2** drinks.

ASK 100 FRIEND GROUPS

Every group gives different advice. The equal averages pile up around 6.5 drinks and spread out; the weighted averages pile up near 3 and stay close together.



WLS Gives More Weight to Observations with Small Error Variance

THE DATA

Wage against education, with the OLS line. OLS gives every worker the same weight.

FIRM SIZE

Point size shows firm size. Workers at large firms (big points) stay close to the line; workers at small firms scatter far from it $\Longrightarrow$ heteroskedasticity. WLS gives the large-firm workers more weight.

WLS LINE

The weighted line is close to the OLS line: both estimate the same β_1.

GOAL

The two point estimates barely differ, but the standard errors do. Decide what that comparison says about heteroskedasticity.

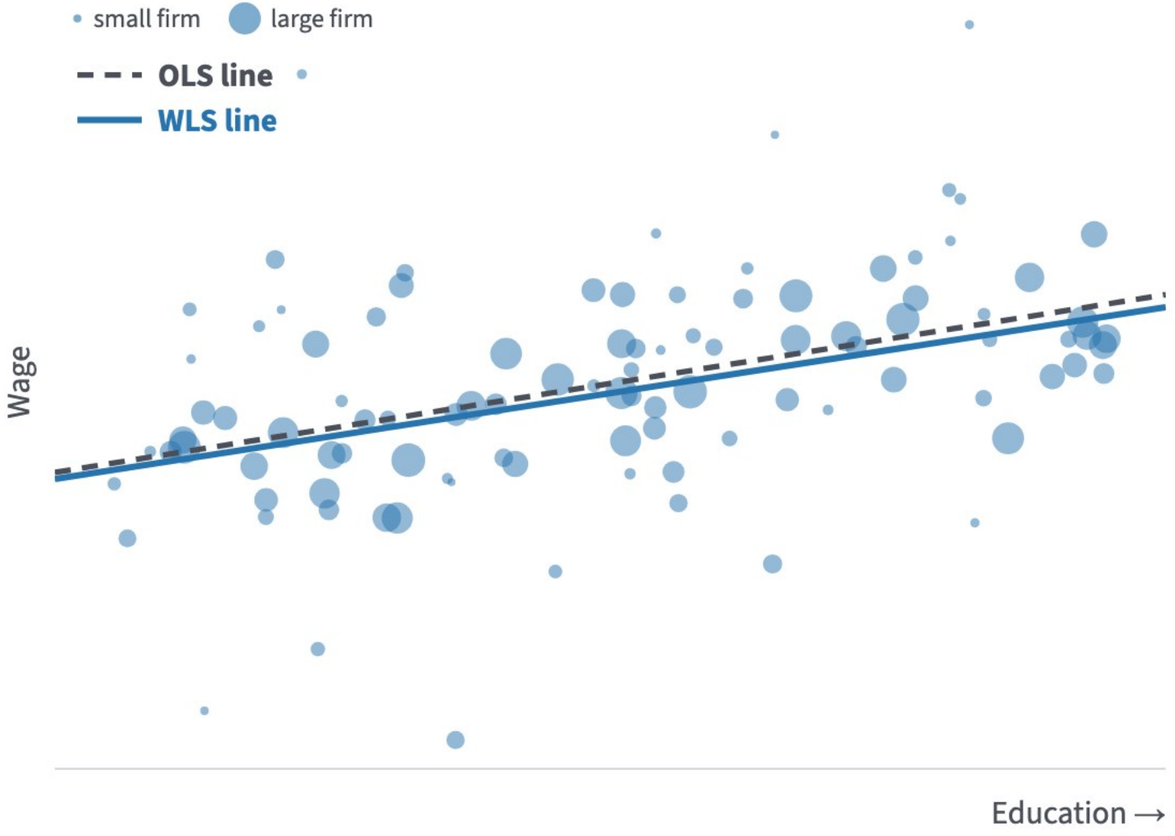
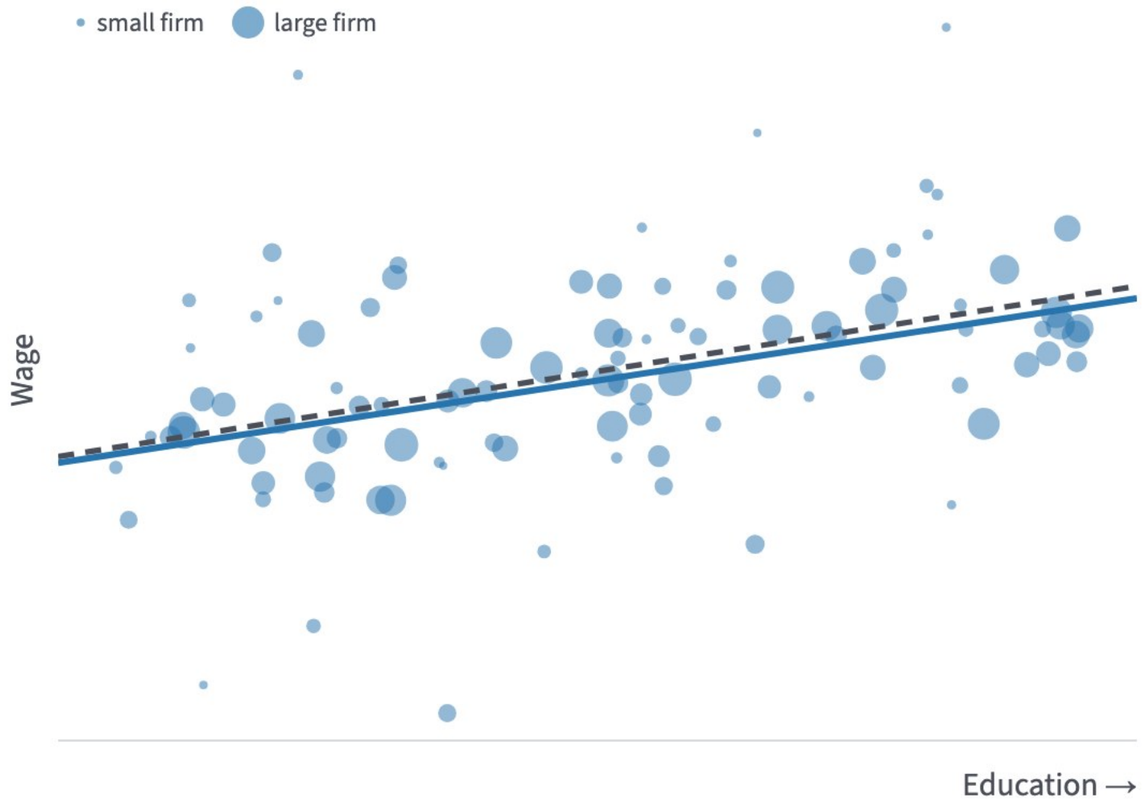


Illustration: 100 workers; the error variance is smaller at larger firms.

Two Questions About Each Estimator



The same 100 workers; point size shows firm size. Dashed gray: OLS line (slope 1.42). Blue: WLS line (slope 1.39).

se standard error: how much the estimate would move from one sample to the next

	Right on average? (unbiased)	**Moves from sample to sample?** (standard error)
OLS	Yes 1.42	se $= 0.31$
WLS	Yes 1.39	**se** $= 0.18$

LEFT COLUMN

The two lines nearly coincide: both estimators are right on average.

THE COLLEAGUE'S MISTAKE

Heteroskedasticity changes only the right column. The colleague looked only at the left column.

Repeat the Study 200 Times

ONE SAMPLE

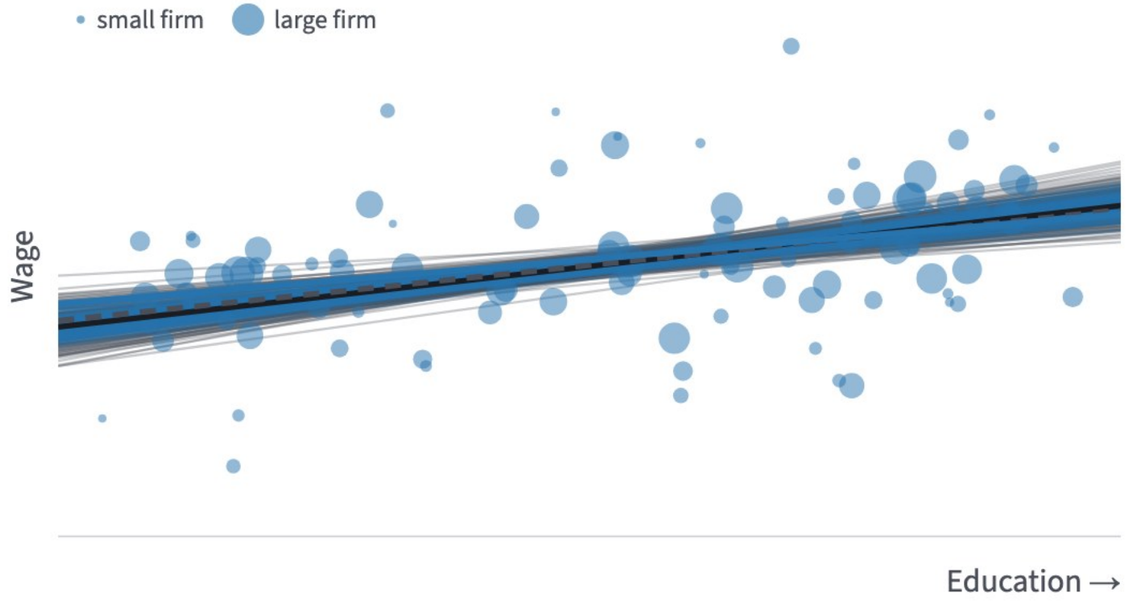
100 workers; point size shows firm size. Big points stay near the line, small points scatter. OLS line: dashed gray. WLS line: blue.

NEW SAMPLES

Each new sample gives new lines. The OLS lines swing more than the WLS lines.

TRUE LINE

Both sets of lines center on the true line (black) $\Longrightarrow$ both estimators are unbiased.

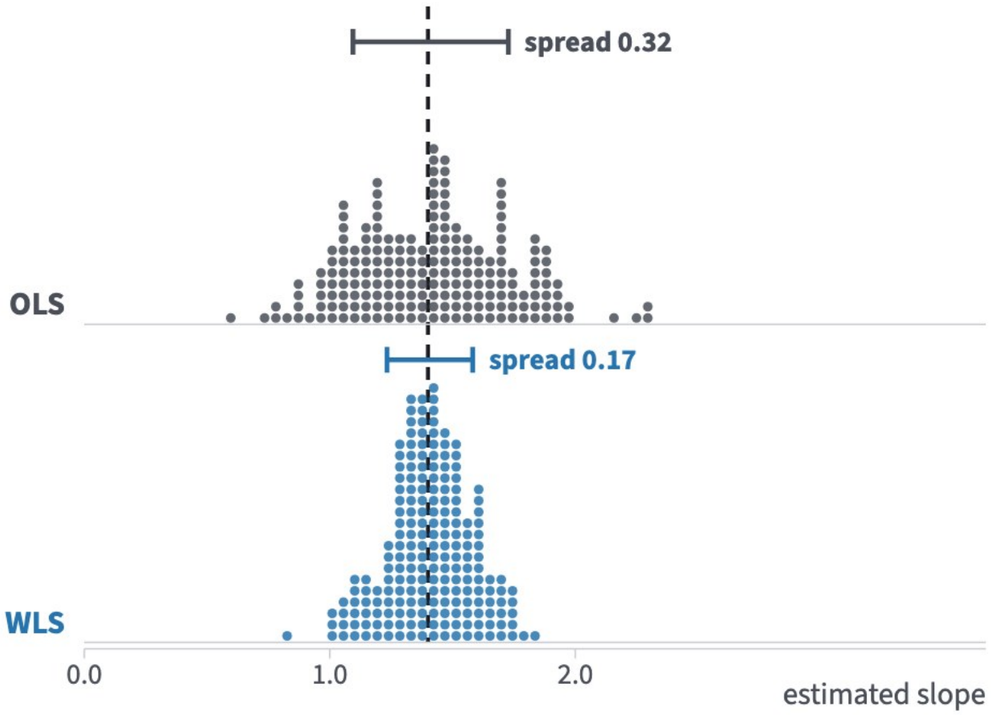


SLOPES

Each line's slope becomes one dot in its row.

SPREAD

The width of each pile is the standard error: 0.32 for OLS, 0.17 for WLS (exam: 0.31 and 0.18).



The exam's 1.42 and 1.39 are one sample: the first one shown here. Both piles land near β_1, so similar estimates are expected.

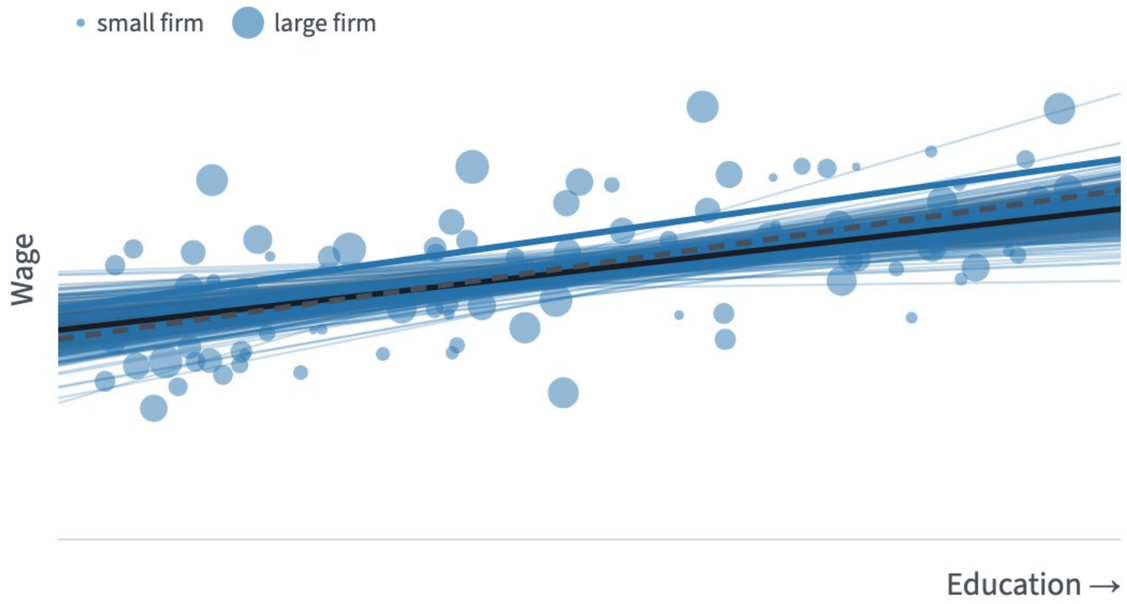
With Constant Error Variance, WLS Cannot Be More Precise Than OLS

ONE SAMPLE

Same firms and weights, but now every worker has the same error variance: big and small points scatter equally.

NEW SAMPLES

Now the WLS lines swing as much as the OLS lines, or more.

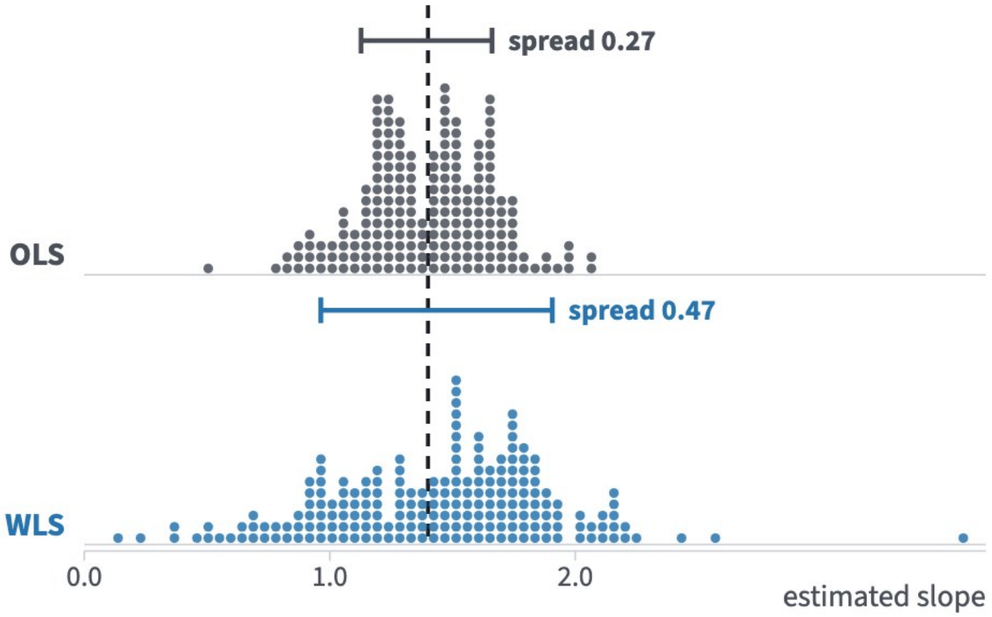


SLOPES

OLS is BLUE here $\Longrightarrow$ the WLS pile cannot be narrower.

SPREAD

These weights do not match the error variance here, so the WLS pile is wider: 0.47 against 0.27 for OLS.



CONCLUSION

- If the error variance were constant, the WLS standard error could not be smaller than the OLS standard error.
- On the exam it is much smaller (0.18 against 0.31) $\Longrightarrow$ evidence of heteroskedasticity.

Checking Each Answer Choice

b The colleague is correct; when slope estimates are similar across OLS and WLS, heteroskedasticity is confirmed absent and OLS is BLUE.

Both estimators are unbiased under heteroskedasticity, so similar slopes are expected either way.

c WLS always produces smaller standard errors than OLS by construction, so the difference in standard errors carries no informational content.

With constant error variance, the WLS pile was not narrower (Step 3).

d WLS estimates are always biased when the variance function is misspecified, making OLS preferable under any uncertainty about the form of heteroskedasticity.

WLS is unbiased with any weights, as long as $E(\varepsilon_i \mid x_i) = 0$. Wrong weights can give a larger variance than OLS.

e OLS is preferred here because WLS is only appropriate when the sample size exceeds 1,000 observations.

No such rule exists. (The original answer key printed (e), a typo later corrected to (a).)

a Similarity in point estimates does not imply OLS standard errors are valid; heteroskedasticity leaves OLS coefficients unbiased but renders standard errors inefficient, and the narrower WLS standard error reflects a genuine efficiency gain.

Answer: (a)

IN PLAIN WORDS

The estimates agree, but WLS is more precise, and that difference is evidence of heteroskedasticity.

CAREFUL WITH THE WORDING OF (A)

Standard errors are not “inefficient.” Under heteroskedasticity the OLS **estimator** is inefficient, and its usual **standard errors** are invalid.

Thank you!

jakeanderson@g.ucla.edu