Principle: Occam's Razor:
All other things being equal,
simpler explanations are better
models

When selecting which model is "best", you need to balance goodness-of-fit vs complexity
how well does model fit points vs how complex

Use more complex model if complexity is warranted.

For Data in ModernDive
Fig 6.8
- simpler parallel slopes model has k=4 parameters
- Complex interaction model has k=6 parameters.

Using the eyeball test, Bert prefers the simpler model

Numerical measures for model

selection:

<u>Recall</u> $R^2 = [0, 1]$

0: if points do not fit regression line at all

1: if regression line fits points perfectly

① $R^2_{adj} = 1 - \frac{(n-1)}{n-(k+1)}(1-R^2)$

where n = # of points
k = " " parameters
$R^2 = R^2$

balances fit as represented by R^2 vs complexity as measured by k.

<u>Ex</u>: As $R^2 \uparrow$
$(1-R^2) \downarrow$
$- \frac{\quad}{\quad}(1-R^2) \uparrow$
as fit gets better,
$R^2_{adj} \uparrow$

<u>Ex</u>: As $k \uparrow$
denominator $n-(k+1) \downarrow$
term $\frac{\quad}{\quad}(1-R^2) \uparrow$
$1 - \frac{\quad}{\quad}(1-R^2) \downarrow$

[illegible] $\downarrow$

a) complexity of model goes up, R^2_{adj} $\downarrow$

<u>Moral</u>: Models with larger R^2_{adj} are to be preferred.

For data in ModernDive Fig 6.8 b/c interaction model has $\underline{k=6}$ vs parallel slopes model has $k=4$.

More complex, it has a lower R^2_{adj} of 0.694 (vs 0.696)

② Akaike Information Criteria AIC is another numerical measure for model selection where models with smaller AIC are to be preferred.

$$AIC = 2k - 2\,\text{log-likelihood}$$

[illegible]

balance between
fit → log-likelihood
& complexity → K

a) All other things being equal,
as models get more complex
i.e. have larger K.
AIC ↑

b) All other things being equal,
as models fit the points
better, the log-likelihood ↑,
then $-2 \times$ log-likelihood ↓,
then AIC ↓

Looking @ code.
b/c parallel slopes model
has lower AIC (3280 vs
3283), you are led to select
this model over the more
complex interaction model