Model Fit

Applied Regression in R

Aleš Vomáčka

08. 03. 2026

Faculty of Arts, Charles University

How to Check Model Fit?

How to Check Model Fit?

(Almost) any model can be fitted to our data. No all models will fit equally well.

There are three ways to evaluate model fit:

- Assumption checks using diagnostic plots (later)
- Fit indeces summarizing fit.
- Formal tests of fit (e.g. ANOVA)

Coefficient of Determination

Coefficient of Determination

The most common fit index for linear regression is **coefficient of determination** (R^2).

The proportion of variance of the outcome variable, which can be predicted using our predictors.

e.g. $R^2 = \mathbf{0.32}$ means our model sucessfuly predicts 32% of outcome's variance. (This is not causal!)

Coefficient of Determination

Coefficient of determination works with sums of squares.

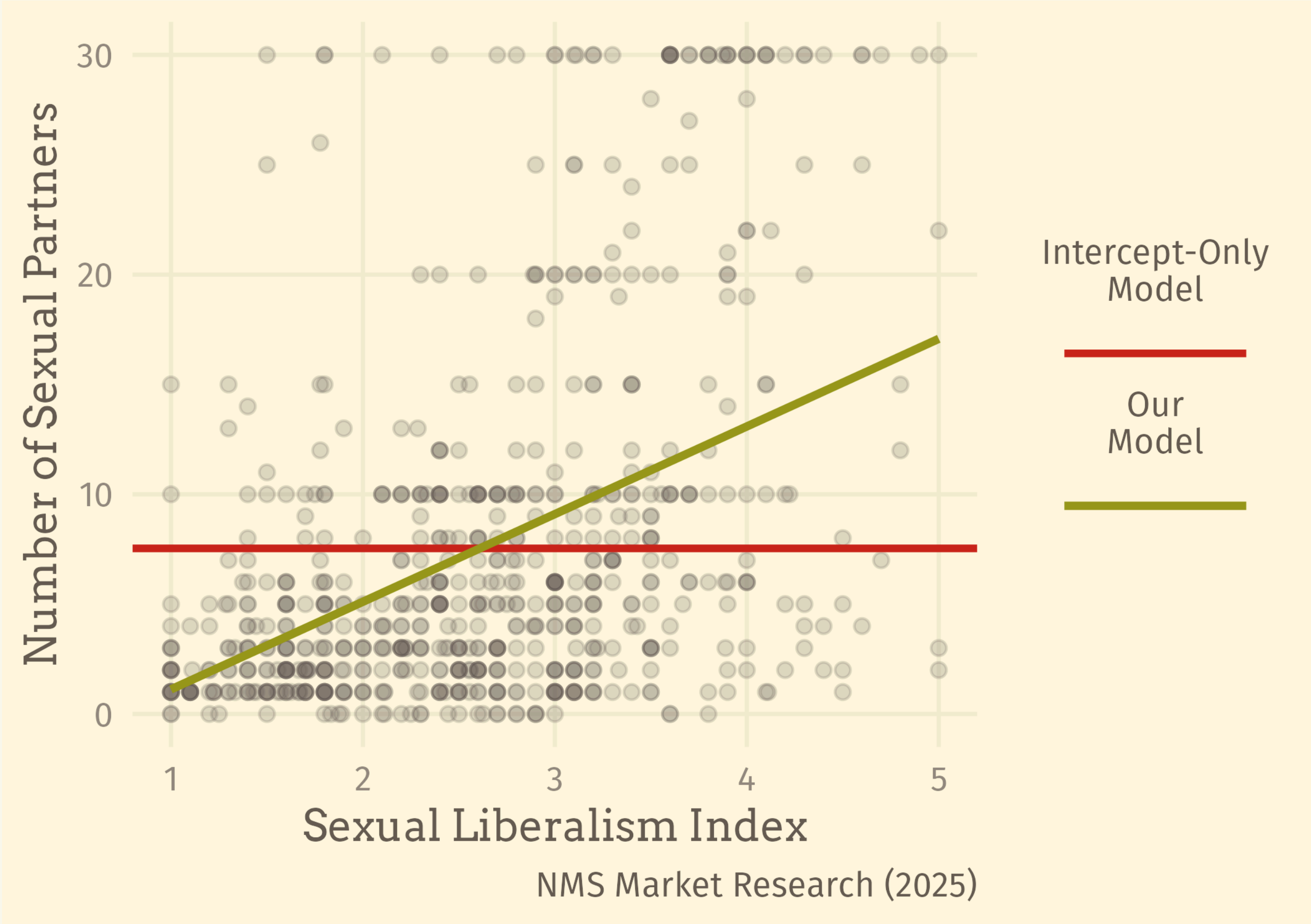
$$R^2 = 1 - \frac{\text{Sum of squares}_{\text{residual}}}{\text{Sum of squares}_{\text{total}}}$$

Unexplained „variance“ (numerator)

Total „variance“ (denominator)

Coefficient of Determination

Visually, R^2 can be viewed as **comparing our model to intercept--only model** (the „worst“ possible one).



Coefficient of Determination

R^2 tells us how much we've **reduced prediction error** by adding predictors.

$R^2 = 0$ means our model is as „good" as if we had no predictors.

$R^2 = 1$ means we predict the outcome perfectly.

Coefficient of Determination

There is **no universal cut-off** for good and bad R^2

In laboratory calibrations, $R^2 < 0.99$ is bad and a sign of equipment failure.

In day stock trading, $R^2 > 0.02$ is good and used for trading.

Questions?

Formal Test of Fit

Formal Test of Fit

We can formally test a hypothesis that two models predict the same amount of outcome variance.

This includes comparing our model with intercept-only model.

In linear regression, this is usually done using ANOVA.

Formal Test of Fit

ANOVA compares, if the amount of predicted variance is the same for both models.

For example, does our model predicts the same amount of variance as a model with no predictors.

Formal Test of Fit

ANOVA test for model `lm(sex_partners ~ sexlib_index)`:

Parameter	Sum Squares	df	Mean Square	F	p
sexlib index	8816	1	8816	165	<.001
Residuals	35974	675	53	NA	NA

Questions?

InteRmezzo!

Limitations

Limitations

ANOVA has all the classic limitations with null hypothesis testing.

- It's etremely unlikely two models two models will explain the exact same amount of outcome variance → the null hypothesis is almost certainly false.
- Magnitude of the difference matters more than simple its existence.
- Power matters - lack of significance doesn't mean lack of difference.

Limitations

R^2 is a measure of predictive strength.

It doesn’t behave intuitively when the goal isn’t prediction.

We’ll have a look at four „gotchas“.

R^2 depends on the number of predictors

Imagine we have outcome y and 15 predictors x_i $(x_1, x_2, ...x_{15})$.

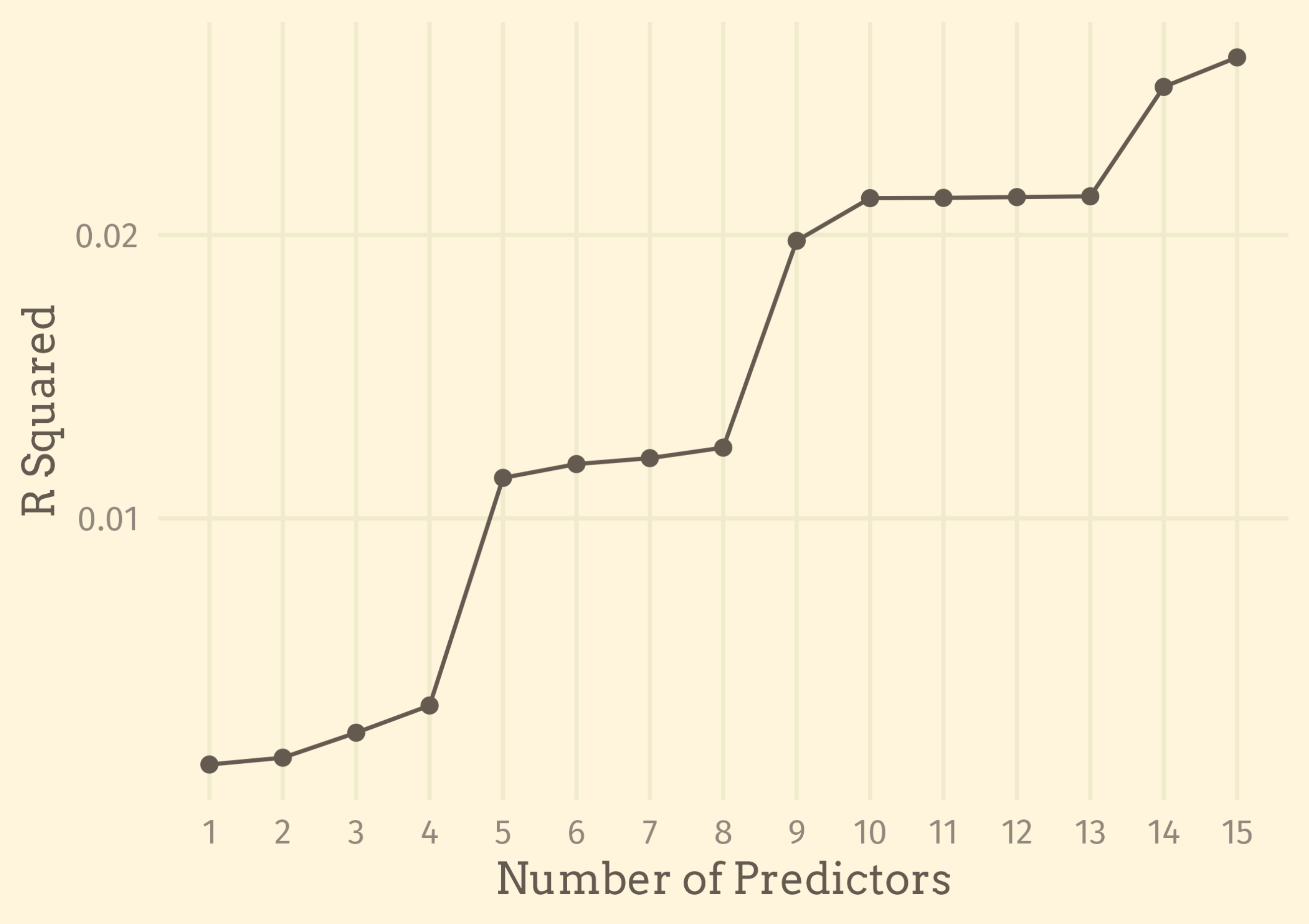
The outcome and all predictors are independent.

What happens to R^2 when we start adding the predictors into our model?

R^2 depends on the number of predictors

Every time we add a predictor, R^2 increases.

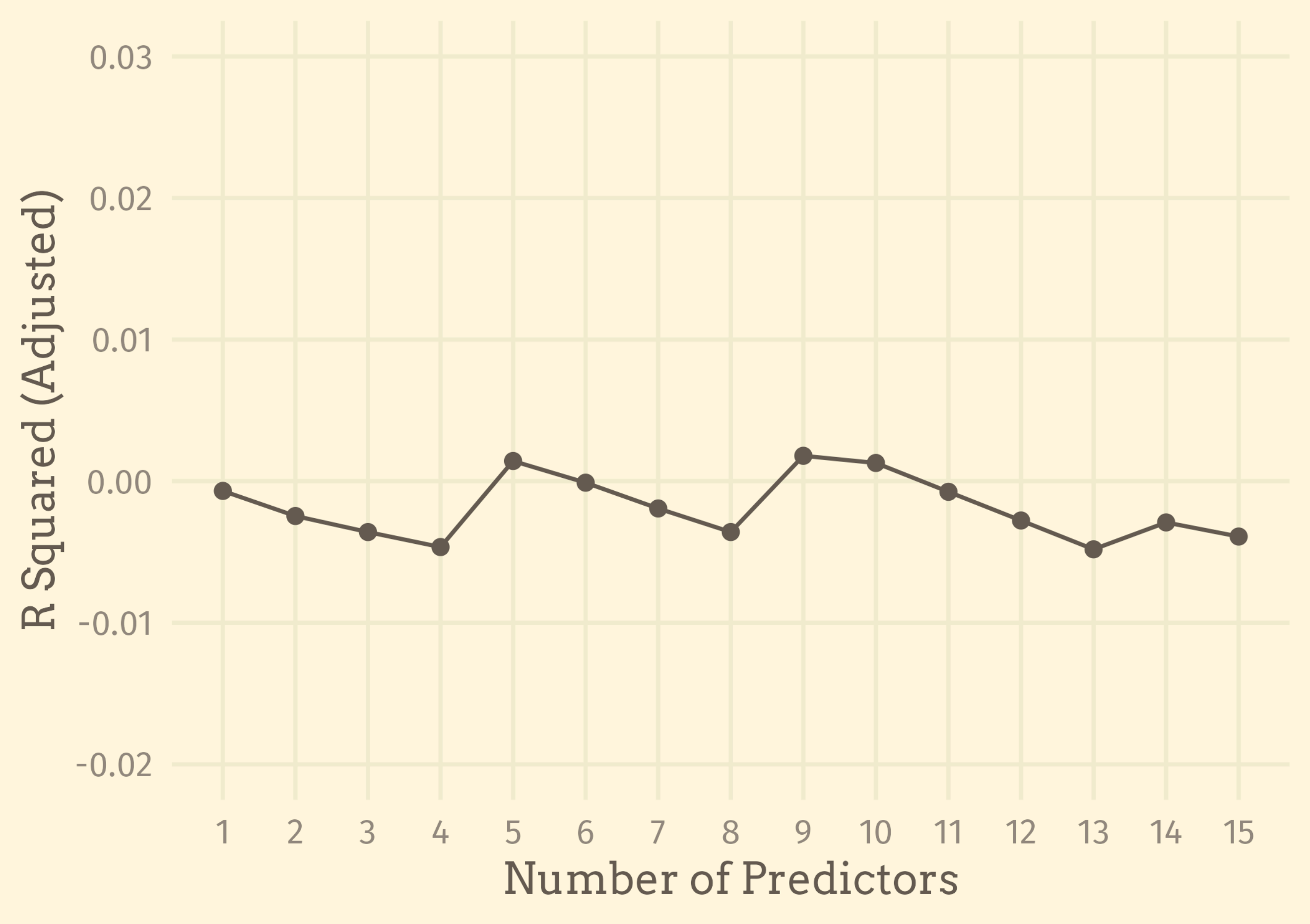
Even if the predictors are unrelated, there will be **some correlation due to random noise**.



R^2 depends on the number of predictors

Use adjusted R^2 when comparing models with different number of predictors.

It subtracks amount of explained variance, we would expect due to random noise.



Questions?

Higher R^2 Doesn't Neccesarily Mean Better Model

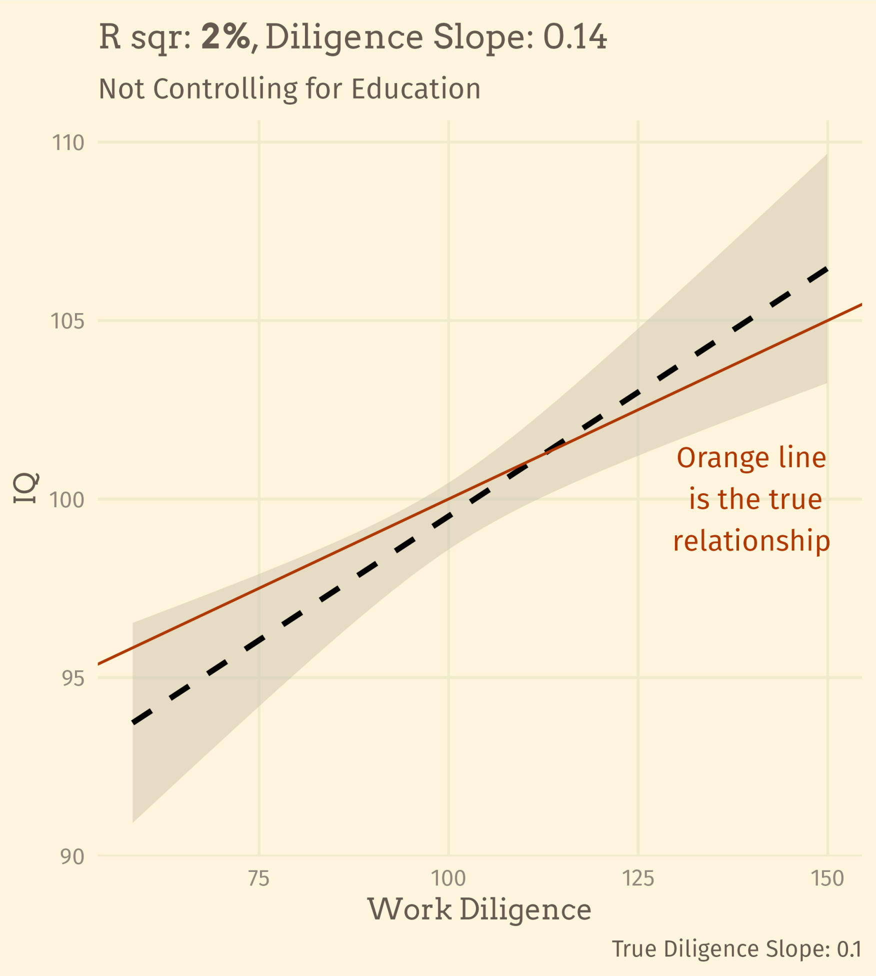
Our psychologist friend wants to study relationship between intelligence (IQ) and work dilligince.

They also know whether their respondents are university graduates or not.

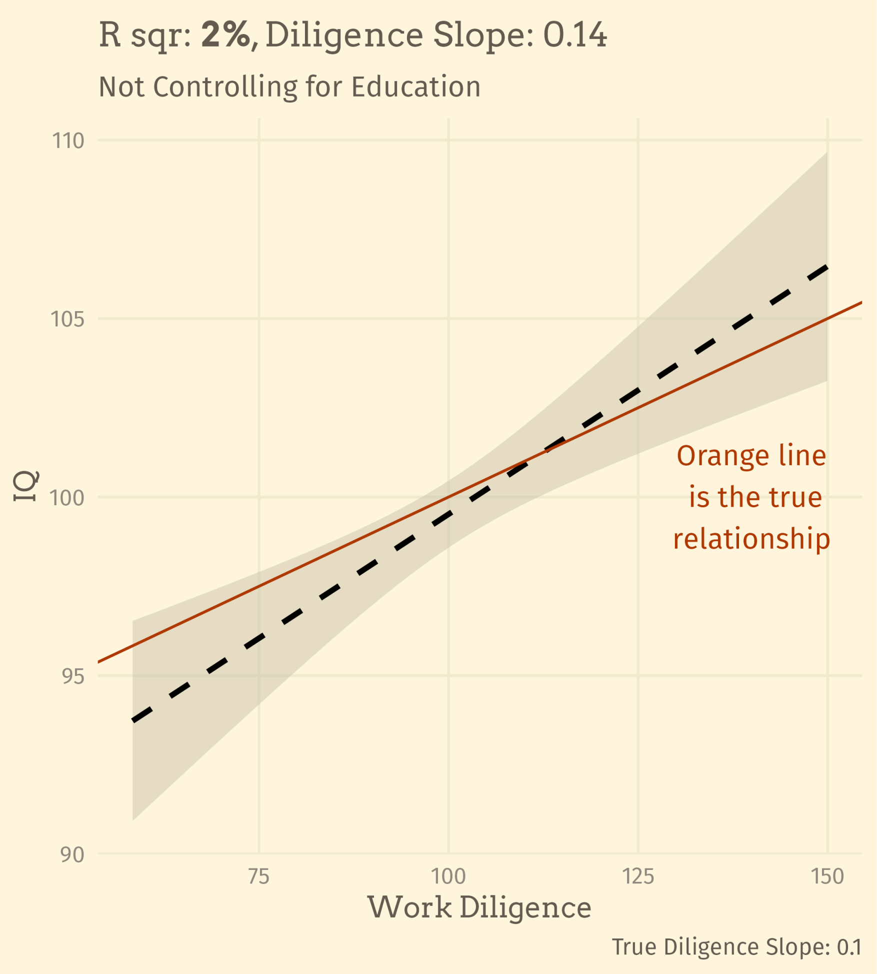
Should we control for university education or not?

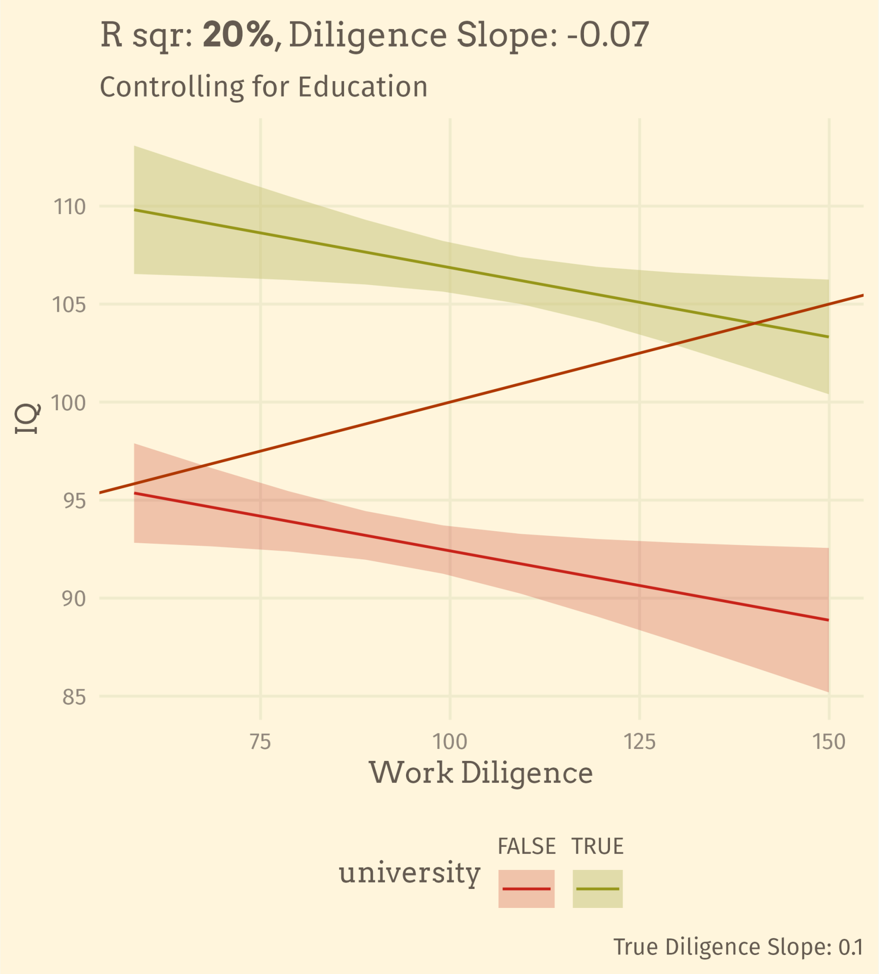
(The true relationship is $\mathbf{IQ} = 0.1 \cdot \mathbf{diligence}$. Let's pretend we don't know.)

Higher R^2 Doesn't Neccesarily Mean Better Model



Higher R^2 Doesn't Neccesarily Mean Better Model

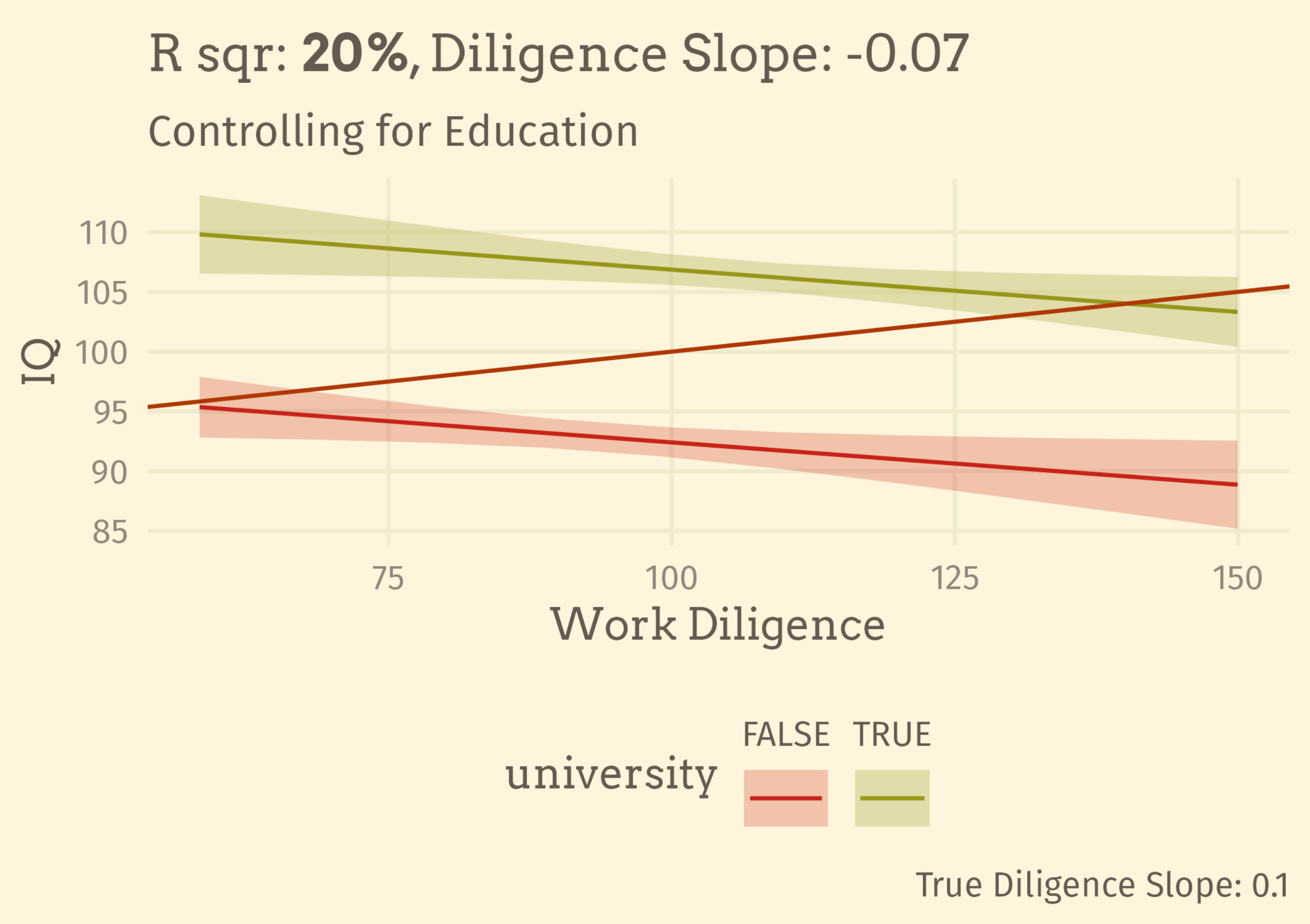




Higher R^2 Doesn't Neccesarily Mean Better Model

Controlling for university degree leads to **higher R^2**, but completly **incorrect slope**!

If the goal is estimate relationships, R^2 can't tell you how good your model is.



Questions?

R^2 Depends on The Outcome Variance

We have predictor x and three outcomes: y_1, y_2 and y_3.

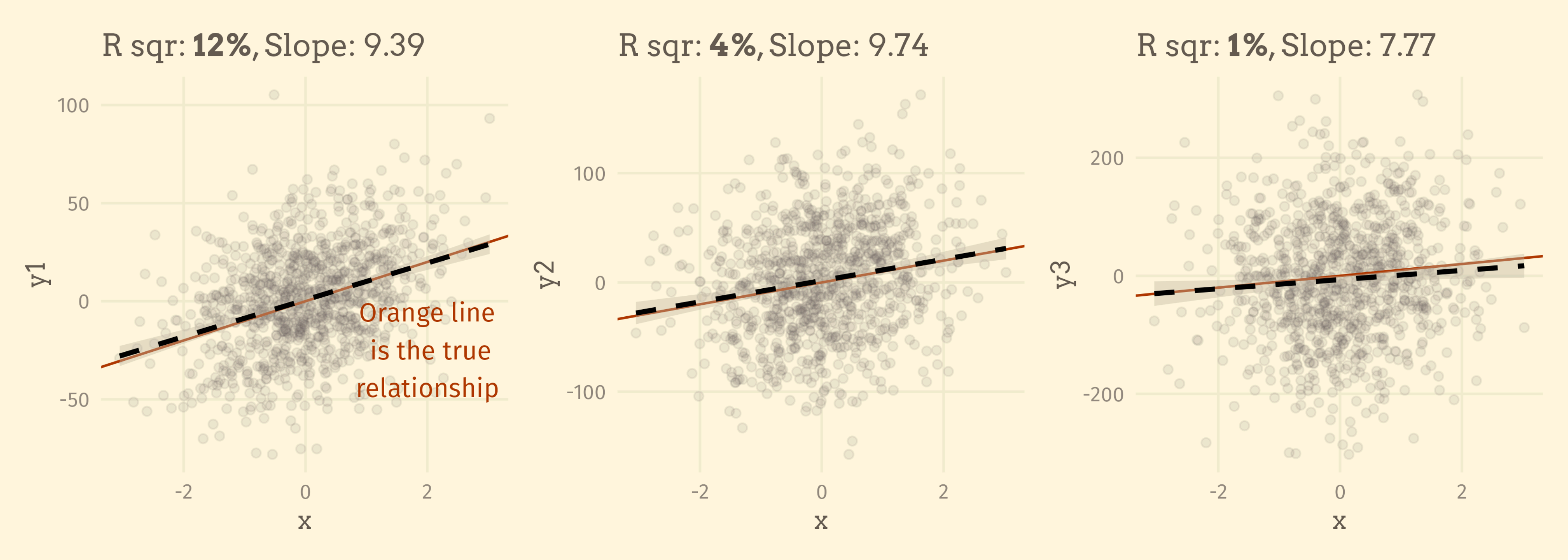
The relationship between the outcome and all predictors is the same:

$$y_i = 0 + 10 \cdot x$$

Each y_i has different standard deviation.

$$\mathrm{sd}_1 = 50, \mathrm{sd}_2 = 100, \mathrm{sd}_3 = 150$$

R^2 Depends on The Outcome Variance



The higher the residual variance, the lower the R^2, even though the relationship is estimated correctly (true slope = 10).

R^2 Depends on The Outcome Variance

Even if the model is perfectly specified (i.e. all relevant variables present in correct form), R^2 can be **low due to absence of variables irrelevant to our study**.

When the goal is estimating relationship, R^2 isn't a good indicator of model quality.

Questions?

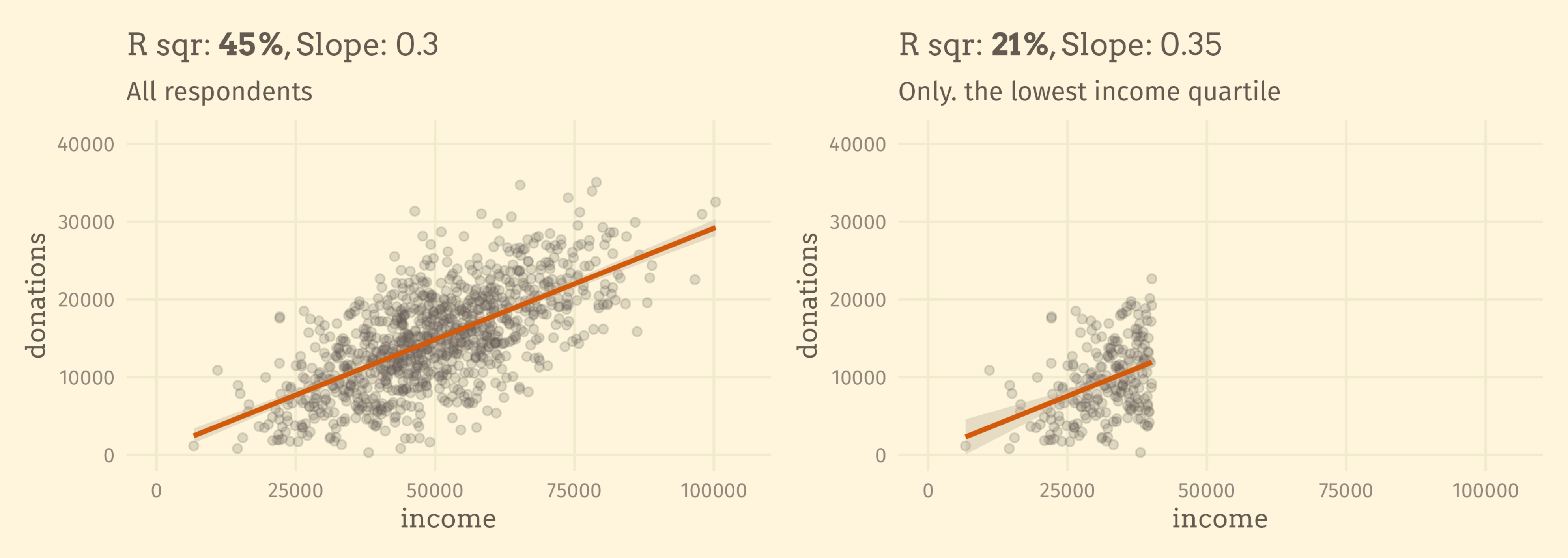
R^2 Depends on The Range of Independent Variables

We are intersted in relationship between income and charity donations.

The true relationships between them is $y = 0 + 0.3 \cdot x$.

What happens if we restrict income to people in the lowest income quartile?

R^2 Depends on The Range of Independent Variables



Restricting predictors lower R^2, but the model is still correct (True slope is 0.3).

Limitations

The **smaller the range** of independent variables/predictors, the **lower the R^2**.

Keep in mind when comparing different populations!

Questions?

Wrapping Up

Wrapping Up

R^2 is **measure of predictive strength**. It can't be used when the goal is estimating relationships.

Low R^2 doesn't neccesarily mean the model is bad.

High R^2 doesn't neccecarily mean the model is good.

The best way to use R^2 is to compare your model to industry standard.

Questions?