Omitted Variable Bias

Economics

The commonest reason a regression coefficient is wrong.

Omitting a variable correlated with both the regressor and the outcome biases the estimate.

Bias direction = sign(effect of Z on Y) × sign(corr of Z and X)

Both positive Overstated.	**Coefficient biased upward**
Opposite signs	**Biased downward**
Z unrelated to X Only loses precision.	**No bias**

THE COMMON MISTAKE

× **Adding every available control variable**

✓ **Choosing controls by causal reasoning**

Controlling for a variable on the causal path removes part of the very effect you are trying to measure.

Draw the causal diagram before choosing controls. Which variables to include is a theory decision, not a statistical one.