Financial Econometrics | Lecture 5

Multiple Regression: Omitted Variable Bias and Control Variables

In Class Group Activity

· SJTU, Summer School 2026 · Fatih Kansoy

Lesson of the day. A single-regressor slope can be precise and significant and still be wrong, because a confounder you left out is hiding in the error term. The cure is to put the confounder into the regression and read each slope as "other things equal".

Phase 1 – Concept checks

ANSWERS

ConcepTest 1. Answer. Option **(b)**. By the sign rule the bias is corr(size, perf)×(effect of perf on pay) = (negative)×(positive) = negative, so the simple slope sits *below* the true effect. Option (a) assumes both effects are positive; (c) confuses measurement error with omitted variable bias; (d) denies that the sign rule can predict direction.

You can sign the bias before running the bigger regression: that is what the sign rule is for.

ConcepTest 2. Answer. Option **(a)**. A partial slope is the effect holding the other regressors fixed: same age, same sex, compare degree against no degree, +$8.30 an hour. Option (b) reads a slope as a correlation; (c) confuses the partial slope with the simple one (the simple premium is 7.82, not 8.30); (d) reverses the regression.

ConcepTest 3. Answer. Option **(c)**. Ordinary R^2 can never fall when you add a regressor, so a rise is not evidence the variable belongs; check whether *adjusted* R^2 rose. Option (a) is the misconception; (b) invents a threshold; (d) confuses R^2 with adjusted R^2 (only the latter carries the penalty).

Phase 2 – The desk problem

WORKED SOLUTIONS

(a) In Model 2 the coefficient on `risk` is 3.5. Reading it ceteris paribus: among traders on desks of the *same* profit, a trader trusted with one extra £million of daily VaR limit earns about £3,500 more bonus a year. It is a partial effect: the simple slope was 6.0, the partial slope is 3.5.

(b) Bias = simple − multiple = 6.0 − 3.5 = +2.5. The sign rule: corr(risk, desk_pnl) > 0 (risk-trusted traders sit on profitable desks) and the effect of `desk_pnl` on bonus > 0, so the bias is positive, matching +2.5. Part of what looked like a reward for running risk was really the reward for sitting on a profitable desk.

(c) $\bar{R}^2 = 1 - (1 - R^2)\dfrac{n-1}{n-k-1}$, $n = 400$.

$$\text{Model 1 } (k=1): \; 1 - (1-0.250)\tfrac{399}{398} = 1 - 0.750(1.00251) = \mathbf{0.248}.$$

$$\text{Model 2 } (k=2): \; 1 - (1-0.360)\tfrac{399}{397} = 1 - 0.640(1.00504) = \mathbf{0.357}.$$

Adjusted R^2 rose from 0.248 to 0.357, a clear gain even after the penalty, so `desk_pnl` earned its place. Ordinary R^2 (0.250 → 0.360) would have risen even if the control were useless, so it cannot

make this call alone.

(d) (1) `seniority_index`$= 2 \times$ `years_at_bank` + const is an exact linear function of an included regressor: **perfect multicollinearity**. OLS cannot separate the two effects, so the software errors or silently drops one. This is the dummy-variable trap, and the fix is to remove the redundant regressor. (2) Adding `var_limit` (corr() $= 0.93$ with `risk`) is **imperfect multicollinearity**: the slopes stay *unbiased*, but the standard errors on the two near-identical regressors inflate (here the VIF is about $1/(1-0.93^2) \approx 7.3$). Nothing biases the slope; the remedy is a wider interval or more data, not dropping a variable that belongs.

In the CPS wage model the VIFs are 1.02/1.00/1.02*: collinearity is a property of these particular regressors, not a disease of multiple regression.*

Phase 3 – The case

REVEAL AND ANSWERS

---

1. The fund's size and momentum exposures end up bundled into the alpha and the error term, because they are omitted determinants of the fund's return.

2. Yes. Size and momentum determine the fund's return (condition one) and are correlated with its market-driven return (condition two), so both conditions for omitted variable bias hold.

3. By the sign rule, the omitted factors had positive average returns and positive loadings, so the one-factor model credits their premia to alpha. Adding them as controls should pull alpha *down* towards zero.

4. **Instructor reveal.** Add the size (SMB) and momentum (MOM) factors. The four-factor result: $\widehat{\alpha} = 0.05\%$ per month ($t = 0.3$, not significant), $\widehat{\beta}_{\text{mkt}} = 0.98$, $\widehat{\beta}_{\text{smb}} = 0.55$ ($t = 4.1$), $\widehat{\beta}_{\text{mom}} = 0.40$ ($t = 3.5$), adjusted $R^2 = 0.91$. The alpha collapses from 0.42% ($t = 2.6$) to 0.05% ($t = 0.3$); the omitted variable bias was about +0.37% per month. The "skill" was almost entirely compensation for bearing small-cap and momentum risk, which an investor can buy cheaply through an index. Holding the factor exposures fixed, Meridian shows no skill.

Same mechanism as the gender gap, opposite direction: there controls made the gap wider ($-2.50 \to -3.43$) because the bias hid part of the effect; here controls made alpha smaller ($0.42 \to 0.05$) because the bias manufactured part of it. Direction comes from the sign rule, not from a habit of "controls shrink everything".

Reserve alternate (confounded anomaly). A low-price (penny-stock) signal predicts high returns in a simple regression (+0.30%/month, $t = 2.2$), but the effect vanishes once size and book-to-market enter (+0.04%, $t = 0.3$): cheap-priced stocks are mostly small and distressed.

End of worked solutions.