# Financial Econometrics | Lecture 5

Multiple Regression: Omitted Variable Bias and Control Variables

**In Class Group Activity**

· SJTU, Summer School 2026 · Fatih Kansoy

**How this block runs.** You work in your home group of three or four, textbook open. Roles rotate: **Scribe**, **Reporter**, **Sceptic**, **Timekeeper**. All money today is in US dollars ($) and pounds sterling (£).

- **Phase 1, Concept checks (~10 min).** Vote on your own, talk for 90 seconds, then vote again.
- **Phase 2, Desk problem (~20 min).** Add a control to a pay regression and read what happens.
- **Phase 3, Case (~25 min).** One real story, ending hidden. You take a position, then we reveal what happened.

A short word list. **Omitted variable bias** arises when a left-out variable both affects $Y$ and is correlated with an included $X$; the **sign rule** says the bias has the sign of $\text{corr}(X, Z) \times$ (effect of $Z$ on $Y$). A **partial** (ceteris-paribus) slope is the effect of one regressor holding the others fixed. **Adjusted** $R^2$ penalises extra regressors, because ordinary $R^2$ never falls when you add one. **Perfect multicollinearity** (the dummy-variable trap) makes a model impossible to estimate; **imperfect multicollinearity** only inflates the standard errors, measured by the **variance inflation factor**. A **control** need not be causal; it only has to make the variable of interest as good as randomly assigned.

## Phase 1 – Concept checks ~10 MIN

**ConcepTest 1.** You regress firms' annual CEO pay on a single regressor, firm size (log assets), and get a positive, significant slope. A colleague says you have omitted firm performance (return on assets). In your sample, larger firms perform slightly *worse* (size and performance are negatively correlated), and better performance raises pay. Compared with the true ceteris-paribus effect of size, the simple slope on size is:

(a) Too high, because both omitted-variable effects are positive.

(b) Too low. By the sign rule the bias is $\text{corr}(\text{size}, \text{perf}) \times (\text{effect of perf}) = (-) \times (+) = \text{negative}$, so the simple slope sits below the truth.

(c) Unbiased, because firm size is measured without error.

(d) Impossible to sign without running the bigger regression first.

**ConcepTest 2.** From the controlled wage regression on the CPS data, $\widehat{\text{ahe}} = 0.96 + 8.30\,\text{bachelor} + 0.46\,\text{age} - 3.43\,\text{female}$ (US dollars per hour). Which statement reads the coefficient on `bachelor` correctly?

(a) Graduates earn $8.30 an hour more than non-graduates, holding age and sex fixed.

(b) The correlation between holding a degree and pay is 8.30.

(c) Graduates earn \$8.30 more, the same number you would get regressing pay on `bachelor` alone.

(d) A one-dollar rise in pay raises the chance of a degree by 8.30.

**ConcepTest 3.** You add one extra regressor (a noisy momentum signal) to a fund-return model. Ordinary $R^2$ rises from 0.300 to 0.305. A teammate says "$R^2$ went up, so momentum belongs." The safest conclusion is:

(a) Correct: any rise in $R^2$ shows the variable adds explanatory power.

(b) Correct only if $R^2$ rose by more than 0.05.

(c) Not established: ordinary $R^2$ can never fall when you add a regressor, so a rise is not evidence. Check whether *adjusted* $R^2$ rose.

(d) Wrong: adding any regressor must lower $R^2$ because of the penalty.

## Phase 2 – The desk problem

~20 MIN

### A London trading desk's pay

**The setting.** A bank studies the pay of $n = 400$ traders on its London desks. Annual bonus is in thousands of pounds (£). The variable of interest is `risk` (the trader's daily value-at-risk limit, in £millions, how much risk they are trusted to run). The candidate confounder is `desk_pnl` (their desk's profit last year, in £millions); risk-trusted traders tend to sit on more profitable desks.

| | | |
|---|---|---|
| **Model 1 (simple)** | $\widehat{\text{bonus}} = 40.0 + 6.0\,\text{risk},$ | $R^2 = 0.250$ |
| **Model 2 (add control)** | $\widehat{\text{bonus}} = 22.0 + 3.5\,\text{risk} + 1.8\,\text{desk\_pnl},$ | $R^2 = 0.360$ |

**(a) Read the partial slope.** Interpret the coefficient on `risk` in Model 2, ceteris paribus, with units. How does it differ from the simple slope of 6.0?

**(b) Compute and sign the bias.** The omitted variable bias is (simple slope − multiple slope). Compute it, then use the sign rule to explain its direction from the correlation of `desk_pnl` with `risk` and the effect of `desk_pnl` on bonus.

**(c) Adjusted $R^2$: did the control earn its place?** Using $\bar{R}^2 = 1 - (1 - R^2)\dfrac{n-1}{n-k-1}$ with $n = 400$, compute $\bar{R}^2$ for Model 1 ($k = 1$) and Model 2 ($k = 2$). What does ordinary $R^2$ fail to tell you here?

**(d) The multicollinearity twist.**

(1) The bank adds `seniority_index` $= 2 \times$ `years_at_bank` $+$ const while `years_at_bank` is already in the model. What goes wrong, and what is this called?

(2) In a separate trial it adds `var_limit`, which correlates 0.93 with `risk`. What happens to the standard errors? What does *not* happen to the slopes?

## Phase 3 – The case

~25 MIN

## Meridian Growth: skill, or just exposure?

**How to run this.** Read the story with your group. The ending is hidden on purpose. Commit a written position before the reveal. Do not search the internet.

**The story.** Meridian Growth is a US equity fund that has beaten the S&P 500 for six straight years. A wealth manager runs a one-factor CAPM regression of the fund's monthly excess return on the market's excess return, over 72 months:

$$R_{\text{fund}} - R_f = \alpha + \beta_{\text{mkt}}(R_{\text{mkt}} - R_f) + u.$$

The result: $\widehat{\alpha} = 0.42\%$ per month ($t = 2.6$), $\widehat{\beta}_{\text{mkt}} = 1.15$, $R^2 = 0.78$. A monthly alpha of 0.42% is about 5% a year of "skill", and it is significant. The manager is about to recommend the fund as a genuine stock-picker.

An analyst objects: Meridian is a *growth* fund holding mostly *small, high-momentum* stocks. Over this period small-cap and momentum exposures themselves earned premia. None of that is in the one-factor model, so it all sits in the fund's alpha and in the error.

*The story stops here.*

**Four discussion questions** (answer all four; take a position):

1. In the one-factor regression, where do the fund's size and momentum exposures end up?
2. Are the two conditions for omitted variable bias met for the size and momentum factors?
3. Use the sign rule to predict what happens to the estimated alpha once size and momentum are added as controls. Up or down?
4. Write your group's one-sentence recommendation to clients, before the reveal: is Meridian skilled, or just exposed?

End of activity handout. Solutions are distributed after the block.