Topic 3 Multivariate Regression Model
MG205 Econometrics

José-Ignacio González-Rojas
j.i.gonzalez-rojas@lse.ac.uk
London School of Economics

October 28, 2024

Recap

Almost 1/4 done!
Topics 1 and 2 are foundations.
You need to know these topics by heart.
I promise it will pay off later on the course.
These topics were concisely exposed and it's difficult to cut any material.

Plan for today

1. Key takeaways of Topic 3.
2. Solve Topic 3 Class Exercises.
3. Questions, comments, and concerns.

Key takeaways for Topic 3.

World is complex.
All models are wrong, some are useful.
Focus on a specific relationship between variables.
How does the covariance between variables included and excluded from the model threaten the unbiasedness property of the estimates?

Example True / population model

Probability of buying a Samsung = const + α · price + β_1 · weight + β_2 · battery capacity + e_i

What is the meaning of multivariate regression? (3C.)

How does an increase of 1 pound affect the probability that I buy a Samsung, holding the battery capacity and weight constant / fixed / the same?

Let's say that we only care about the effect of price on demand. Can we get away with just including the price?

It depends on how the independent variables correlate among themselves.

Let's estimate the following model:

$$lP_i = \tilde{\gamma} + \tilde{\alpha} \cdot \varphi_i + \tilde{\beta}_1 \cdot w_i + v_i$$

$$lP = \gamma + \alpha \cdot \varphi_i + \beta_1 \cdot w_i + \beta_2 \cdot bcte_i$$

How do these models relate?

$\gamma \hookrightarrow \tilde{\gamma}$ $\alpha \hookrightarrow \tilde{\alpha}$ $\beta_1 \hookrightarrow \tilde{\beta}_1$ $\beta_2 \hookrightarrow 0$ $\beta_2 \cdot bc_i + e_i \hookrightarrow v_i$

What do we need for an unbiased estimator?

$\mathbb{E}[v_i | p_i, w_i, bc_i] = 0 \Rightarrow$

$\mathbb{E}[v_i \cdot p_i] = 0$

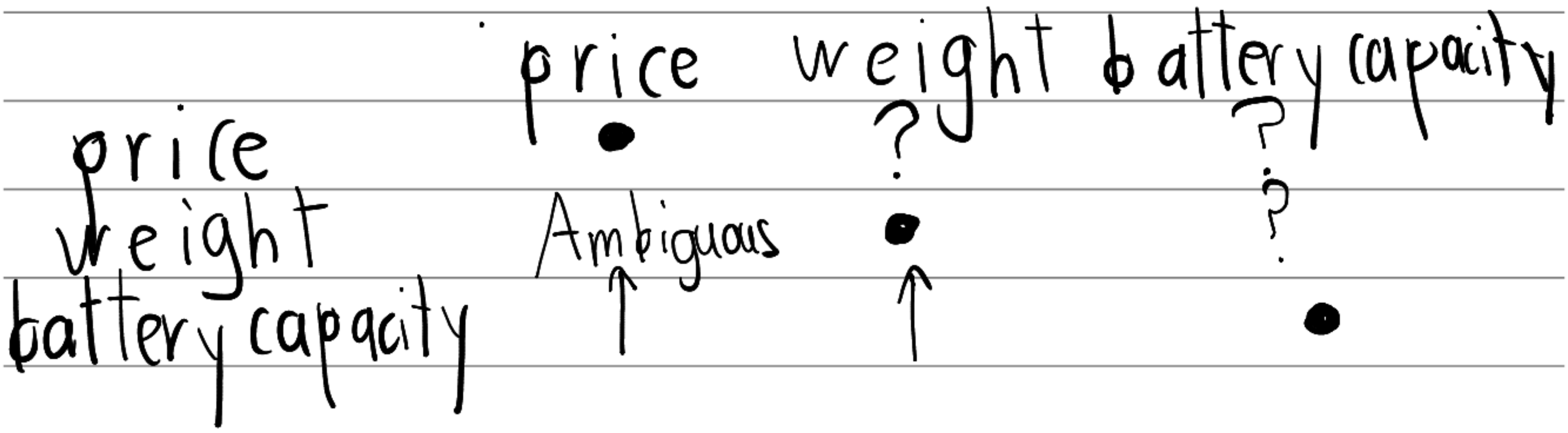
$\mathbb{E}[v_i \cdot w_i] = 0$

$\mathbb{E}[v_i \cdot bc_i] = 0$

$$\mathbb{E}[(\beta_2 \cdot bc_i + e_i)\, p_i] = \beta_2 \underbrace{\mathbb{E}[bc_i \cdot p_i]}_{\neq 0} + \underbrace{\mathbb{E}[e_i p_i]}_{0}$$

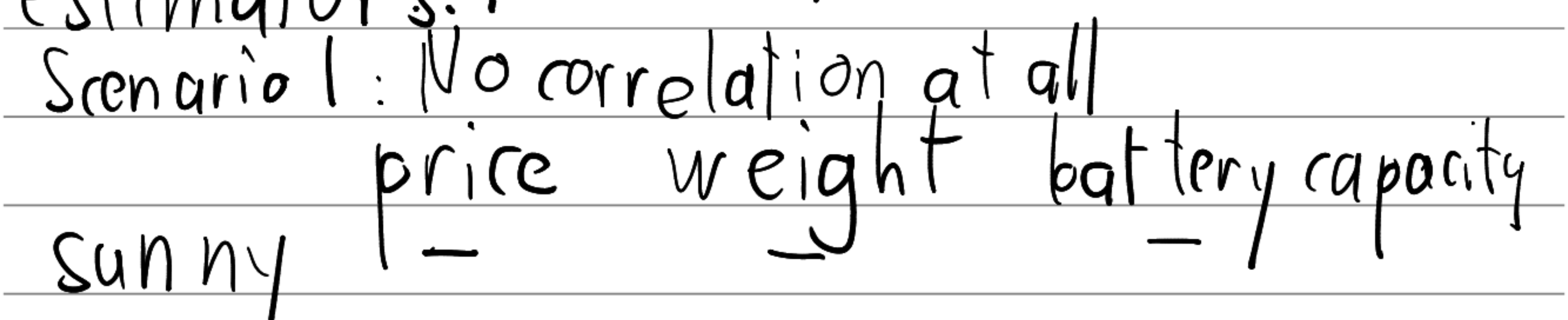
(3C.S)

Hence, we do not have unbiased estimators. You can do the same for the other cases and convince yourselves.

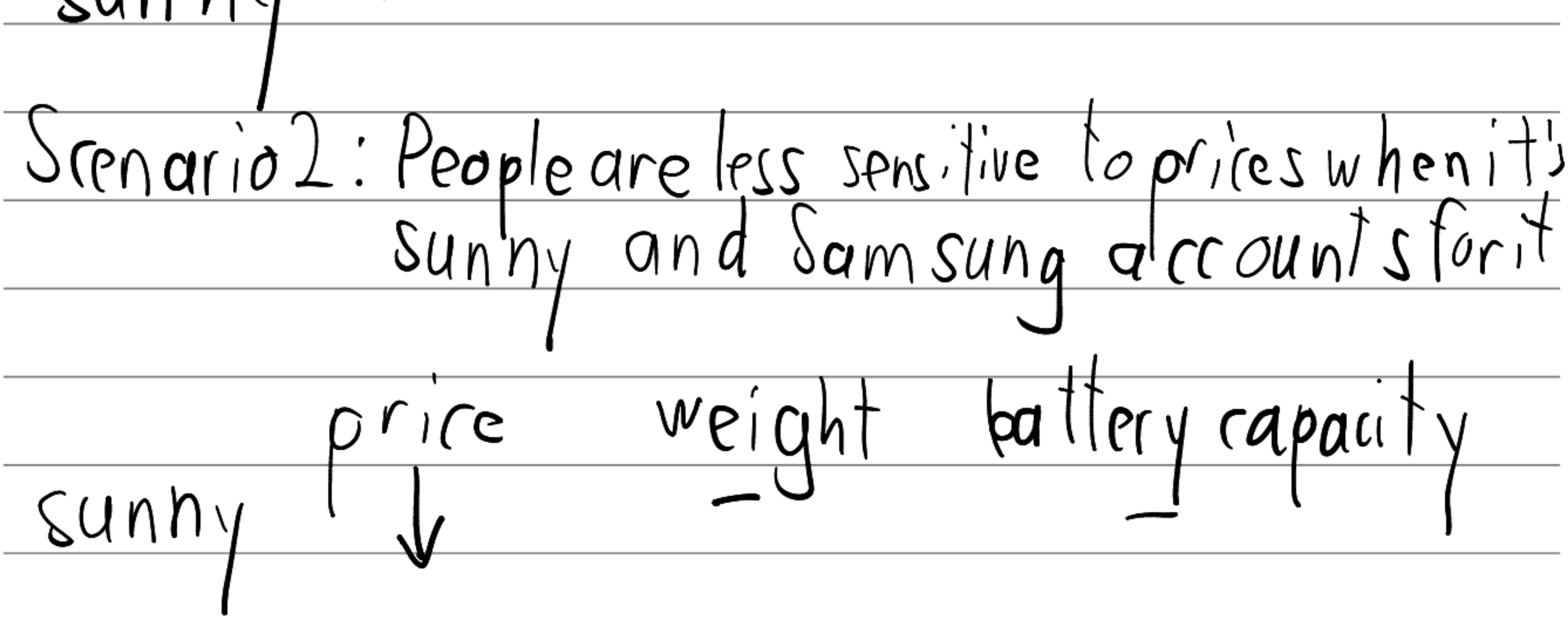


What if buying a Samsung depends on whether today is sunny? Can I get unbiased estimators?

Scenario 1: No correlation at all



Scenario 2: People are less sensitive to prices when it's sunny and Samsung accounts for it



$$P_i = \gamma + \alpha p_i + \beta_1 \cdot w_i + \beta_2 \cdot bc_i + \beta_3 \cdot sunny + e_i$$

And I estimate (3C.4).

$$P_i = \dot{\gamma} + \dot{\alpha} p_i + \dot{\beta}_1 \cdot w_i + \dot{\beta}_2 bc_i + u_i$$

$$E[u_i p_i] = \underbrace{\beta_3 E[sunny \cdot p_i]}_{?} + \underbrace{E[e_i]}_{0}$$

Scenario 1	Scenario 2
0	$\neq 0$
unbiased estimator	biased estimator

But then everything can be correlated with everything! We cannot say anything about the World!

Narrow down your model
Make sensible assumptions
Justify them.

Wait! But I can control for everything I can collect. (3C.1) a, b, c.

The more independent variables you add, the larger the variance of your estimater.

Significance? Goodness of fit? Causality?

1. Significance: wider confidence intervals
   => harder to reject $H_0: \beta = 0$.
   3C.6-7.
2. Goodness of fit: The more controls you add the R^2 <u>weakly</u> increase.
   (3C.4)
3. Causality: You may not observe everything that impads the dependent variable.
   Hence you don't get rid of Omitted Variable Bias.

Solutions: (Qaasi) Experiments, Randomised Controlled Trials.

Research Designs: DiD
IV
RDD.

Solve exercises 3C.1 - 3C.7.

My summary shows the mechanics of this problem set.

Does anyone have any doubts that we should go through over the whiteboard?

Q & A

Summary

1. Need for multivariate regression.
2. Controlling.
3. Unbiased estimates.
4. Goodness of fit
5. Modelling: focus on key mechanisms and defend assumptions.