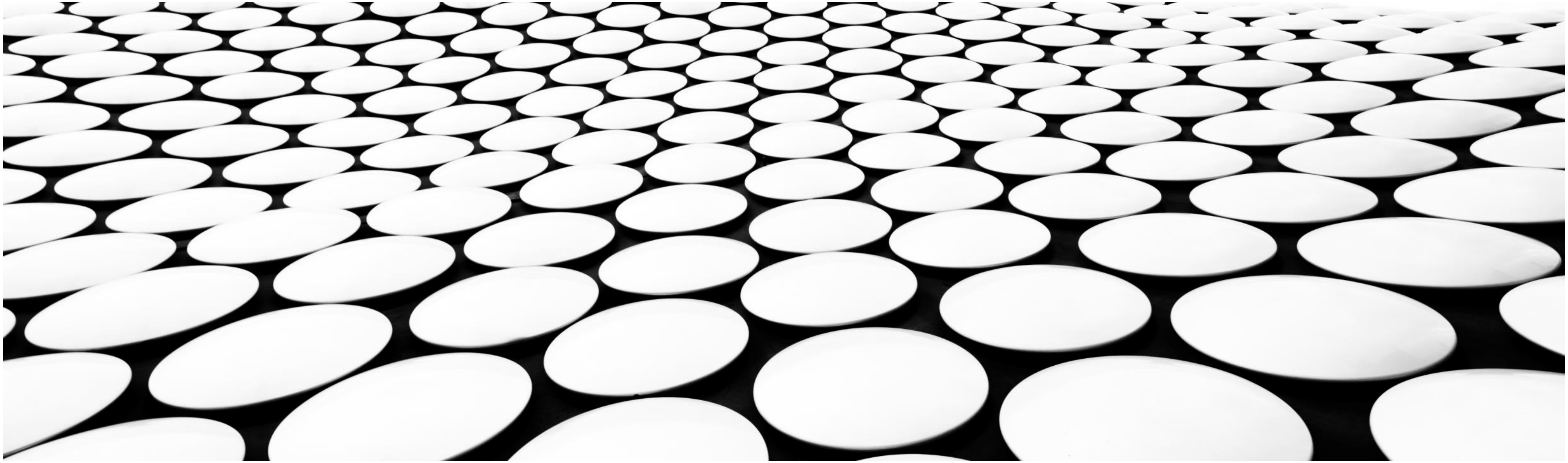
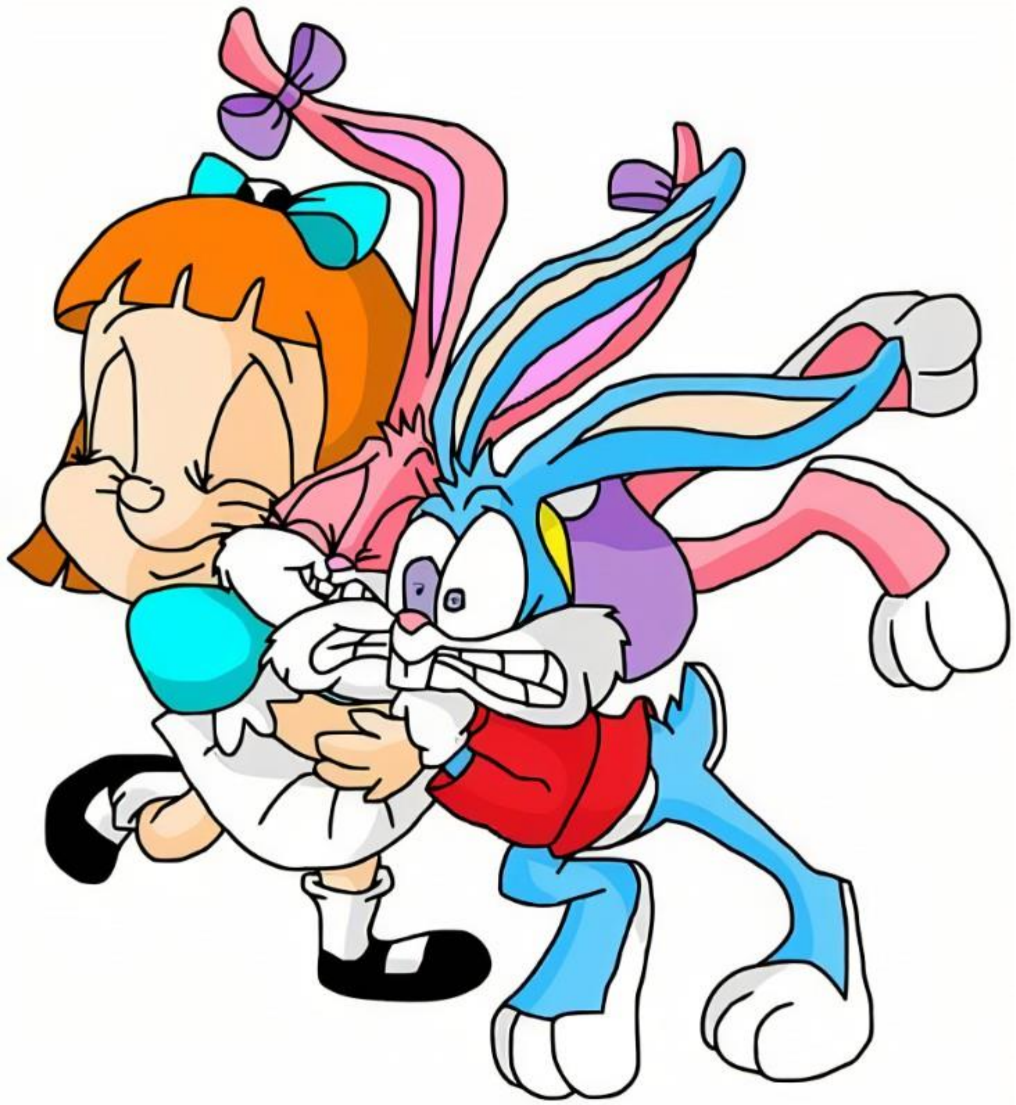
RANDOM EFFECTS

FELIPE BUCHBINDER



I Love FIXED EFFECTS, BUT…

… it has a technical, mathematical, problem:



WHEN $\mathrm{COV}(U_i; X_{it}) = 0$, IT'S POSSIBLE TO FIND ANOTHER ESTIMATE FOR $\boldsymbol{\beta}$ THAT'S ALSO UNBIASED BUT HAS A SMALLER VARIANCE

COMPETITIVE, HEH?

If we already have an unbiased estimate for $\boldsymbol{\beta}$, why do we care if there's another one that has a smaller standard error?

THIS ESTIMATOR IS THE RANDOM EFFECTS ESTIMATOR

Random Effects is better than Fixed Effects when (and only when) $\mathrm{Cov}(U_i; X_{it}) = 0$

START FROM THE GENERAL EQUATION OF PANEL DATA

$$\mathbf{Y}_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \mathbf{U}_{\mathrm{i}} + \epsilon_{it}$$

IF WE TREAT THIS AS A USUAL REGRESSION, $\mathrm{U_I}$ WILL BE ABSORBED IN THE ERROR TERM

Composition Error (ν_{it})

$$\mathbf{Y}_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \boxed{\mathbf{U_i} + \epsilon_{it}}$$

THIS LEADS TO HETEROSKEDASTICITY

$$\mathrm{Cov}(\nu_{it};\nu_{is}) = \mathrm{Cov}(U_i + \epsilon_{it}; U_i + \epsilon_{is})$$

$$= \underbrace{\mathrm{Cov}(U_i;U_i)}_{\sigma_U^2} + \underbrace{\mathrm{Cov}(U_i;\epsilon_{is})}_{0} + \underbrace{\mathrm{Cov}(\epsilon_{it};U_i)}_{0} + \underbrace{\mathrm{Cov}(\epsilon_{it};\epsilon_{is})}_{\sigma_\epsilon^2 \cdot \mathbb{I}(t=s)}$$

$$= \begin{cases} \sigma_U^2 + \sigma_\epsilon^2 & \text{if} \quad t = s \\ \sigma_U^2 & \text{if} \quad t \neq s \end{cases}$$

THERE ARE TRADITIONAL WAYS OF DEALING WITH HETEROSKEDASTICITY. WHY NOT USE THEM?

Rather than treating this as a problem of unobserved heterogeneity, we can simply treat this as a problem of heteroskedasticity!


LET’S DO IT!

OLS REFRESHER

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

$$\boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}; \boldsymbol{\Omega})$$

$$\boldsymbol{\Omega} = \sigma^2 \mathbf{I} = \begin{bmatrix} \sigma^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma^2 \end{bmatrix}$$

$$\boldsymbol{\beta} = \arg\min \boldsymbol{\epsilon}^{\mathbf{T}} \boldsymbol{\epsilon}$$

OLS REFRESHER

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

$$\boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}; \boldsymbol{\Omega})$$

$$\boldsymbol{\Omega} = \sigma^2\mathbf{I} = \begin{bmatrix} \sigma^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma^2 \end{bmatrix}$$

$$\boldsymbol{\beta} = \arg\min \boldsymbol{\epsilon}^{\mathrm{T}} \boldsymbol{\Omega}^{-1} \boldsymbol{\epsilon}$$

Claim:
Putting Ω^{-1} here doesn't change $\boldsymbol{\beta}$

PROOF

$$\boldsymbol{\Omega} = \sigma^2\mathbf{I}$$
$$\therefore$$
$$\boldsymbol{\Omega}^{-1} = \sigma^{-2}\mathbf{I}$$

$$\arg\min \boldsymbol{\epsilon}^{\mathbf{T}}\boldsymbol{\Omega}^{-1}\boldsymbol{\epsilon} =$$
$$\arg\min \boldsymbol{\epsilon}^{\mathbf{T}}\sigma^{-2}\mathbf{I}\boldsymbol{\epsilon} =$$
$$\arg\min \boldsymbol{\epsilon}^{\mathbf{T}}\mathbf{I}\boldsymbol{\epsilon} =$$
$$\arg\min \boldsymbol{\epsilon}^{\mathbf{T}}\boldsymbol{\epsilon} =$$
$$\boldsymbol{\beta}$$

OLS IN A NEW WAY

$$\boldsymbol{y} = \mathbf{X\beta} + \boldsymbol{\epsilon}$$

$$\boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}; \boldsymbol{\Omega})$$

$$\boldsymbol{\Omega} = \sigma^2 \mathbf{I} = \begin{bmatrix} \sigma^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma^2 \end{bmatrix}$$

$$\boldsymbol{\beta} = \arg\min \boldsymbol{\epsilon}^{\mathbf{T}} \boldsymbol{\Omega}^{-1} \boldsymbol{\epsilon}$$

GENERALIZED LEAST SQUARES (GLS) REGRESSION

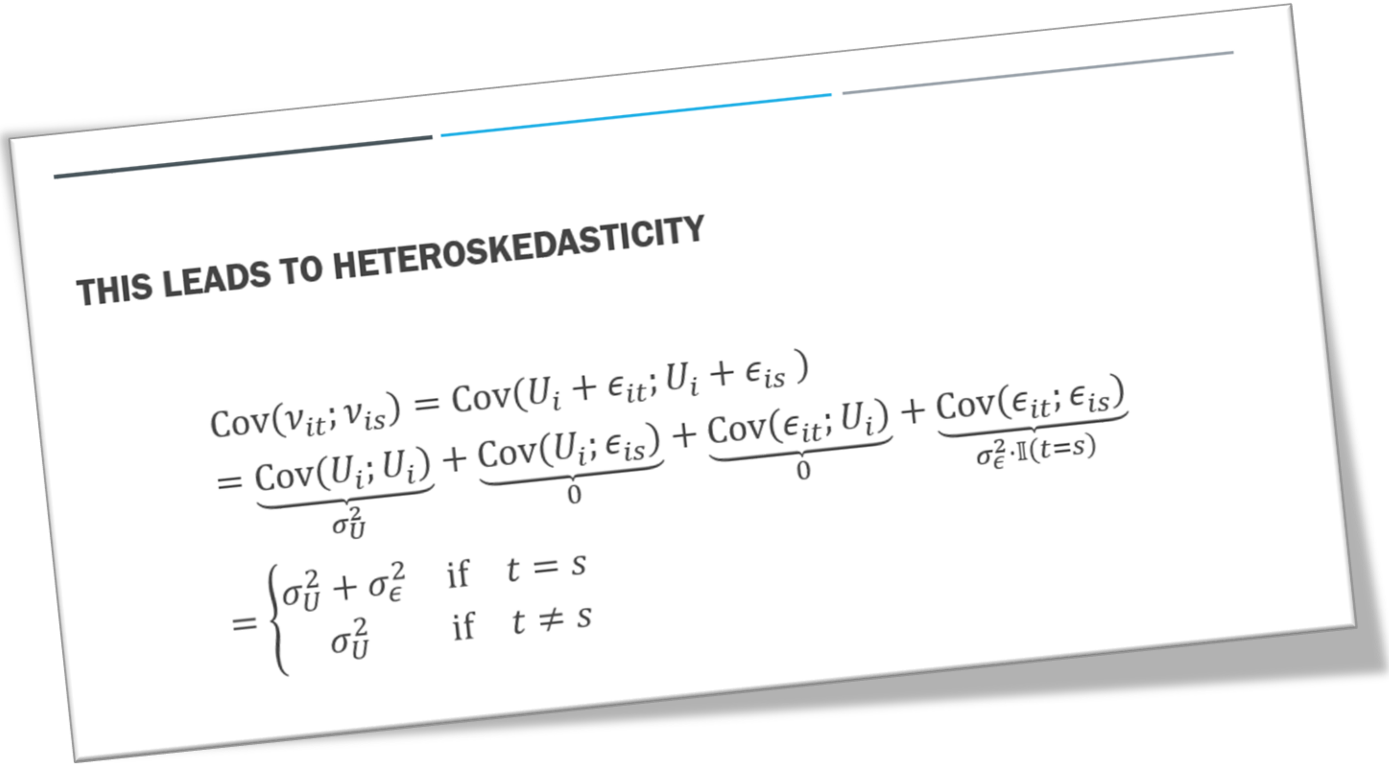
$$\boldsymbol{y} = \mathbf{X\beta} + \boldsymbol{\epsilon}$$

$$\boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}; \mathbf{\Omega})$$

$$\mathbf{\Omega} = \text{any covariance matrix}$$

$$\mathbf{\beta} = \arg\min \boldsymbol{\epsilon}^{\mathbf{T}} \mathbf{\Omega}^{-\mathbf{1}} \boldsymbol{\epsilon}$$

COVARIANCE MATRIX OF THE COMPOSITION ERROR



$$\boldsymbol{\Omega} = \begin{bmatrix} \sigma_U^2 + \sigma_\epsilon^2 & \cdots & \sigma_U^2 \\ \vdots & \ddots & \vdots \\ \sigma_U^2 & \cdots & \sigma_U^2 + \sigma_\epsilon^2 \end{bmatrix}$$

THE RANDOM EFFECTS MODEL

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

$$\boldsymbol{\epsilon} \sim \mathrm{N}(\mathbf{0}; \boldsymbol{\Omega})$$

$$\boldsymbol{\Omega} = \begin{bmatrix} \sigma_U^2 + \sigma_\epsilon^2 & \cdots & \sigma_U^2 \\ \vdots & \ddots & \vdots \\ \sigma_U^2 & \cdots & \sigma_U^2 + \sigma_\epsilon^2 \end{bmatrix} = \sigma_U^2 \mathbf{1} + \sigma_\epsilon^2 \mathbf{I}$$

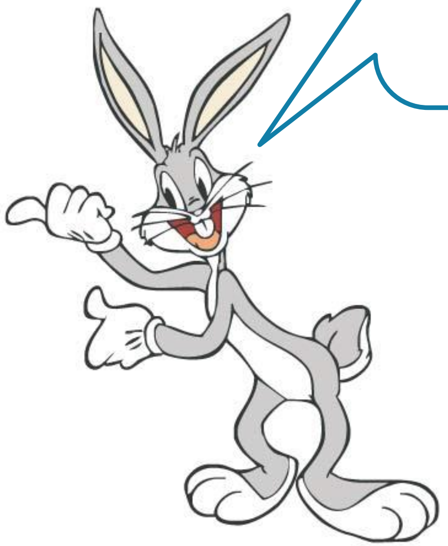
$$\boldsymbol{\beta} = \arg\min \boldsymbol{\epsilon}^{\mathrm{T}} \boldsymbol{\Omega}^{-1} \boldsymbol{\epsilon}$$

SOLUTION TO THE GLS PROBLEM

$$\boldsymbol{\beta} = \arg\min \boldsymbol{\epsilon}^{\mathbf{T}}\boldsymbol{\Omega}^{-\mathbf{1}}\boldsymbol{\epsilon}$$
$$= (\boldsymbol{y} - \mathbf{X}\boldsymbol{\beta})^{\mathbf{T}}\boldsymbol{\Omega}^{-\mathbf{1}}(\boldsymbol{y} - \mathbf{X}\boldsymbol{\beta})$$
$$\Downarrow$$
$$\mathbf{X}^{\mathbf{T}}\boldsymbol{\Omega}^{-\mathbf{1}}\mathbf{y} - \mathbf{X}^{\mathbf{T}}\boldsymbol{\Omega}^{-\mathbf{1}}\mathbf{X}\boldsymbol{\beta} = \mathbf{0}$$
$$\therefore$$
$$\boldsymbol{\beta} = \left(\mathbf{X}^{\mathbf{T}}\boldsymbol{\Omega}^{-\mathbf{1}}\mathbf{X}\right)^{-\mathbf{1}}\mathbf{X}^{\mathbf{T}}\boldsymbol{\Omega}^{-\mathbf{1}}\mathbf{y}$$

SOLUTION TO THE GLS PROBLEM

$$\boldsymbol{\beta} = \arg\min \boldsymbol{\epsilon}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\boldsymbol{\epsilon}$$
$$= (y - \mathbf{X}\boldsymbol{\beta})^{\mathrm{T}}\boldsymbol{\Omega}^{-1}(y - \mathbf{X}\boldsymbol{\beta})$$
$$\Downarrow$$
$$\mathbf{X}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{y} - \mathbf{X}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{X}\boldsymbol{\beta} = \mathbf{0}$$
$$\therefore$$
$$\boldsymbol{\beta} = \left(\mathbf{X}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{y}$$



Hey, Doc.!
Do you get the OLS coefficient when $\Omega = \sigma^2\mathbf{I}$?
You should…

SPECIAL CASES OF THE GLS EQUATIONS

- Matrix Ω enables us to model more complex patterns of variance and covariance between residuals
- If $\Omega = \sigma^2 \mathrm{I}$ (diagonal and all terms equal), the estimate for β becomes equivalent to the estimate of the classical least-squares regression
- If Ω is diagonal but not all terms are equal, we have heteroskedasticity
- In this case, the objective function can be thought of as a sum of *standardized* residuals squares
- If Ω is all-but diagonal, we have serial correlation (not necessarily an AR(1))

$$\mathbb{E}(\boldsymbol{\epsilon}|\mathbf{X}) = 0$$
$$\mathbb{V}(\boldsymbol{\epsilon}|\mathbf{X}) = \boldsymbol{\Omega}$$
$$\widehat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T \boldsymbol{\Omega}^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$$
$$\widehat{\boldsymbol{\beta}} = \left(\mathbf{X}^\mathrm{T} \boldsymbol{\Omega}^{-1} \mathbf{X}\right)^{-1} \mathbf{X}^\mathrm{T} \boldsymbol{\Omega}^{-1} \mathbf{y}$$

My way of thinking about

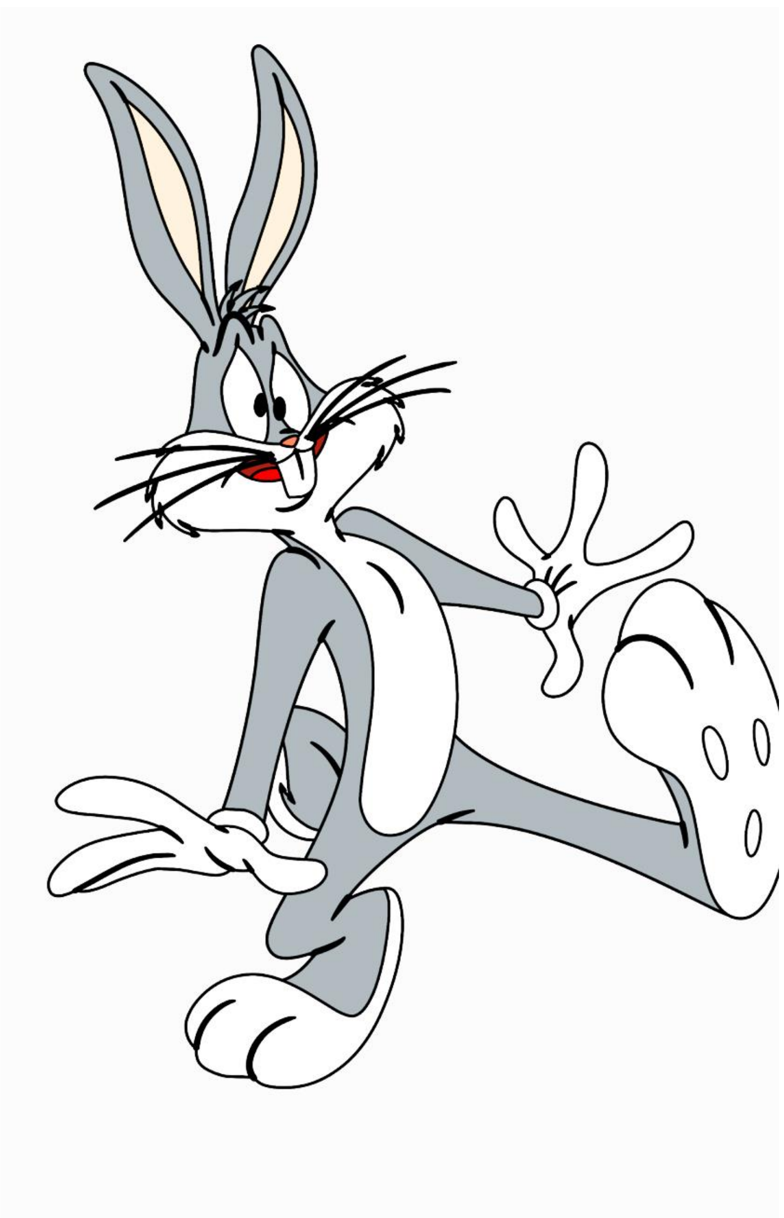
GENERALIZED LEAST SQUARES (GLS) REGRESSION

This minimizes the residual's norm using Euclidian metric...

$$\beta_{OLS} = \arg\min \boldsymbol{\epsilon}^{\mathbf{T}} \boldsymbol{\epsilon}$$

$$\beta_{GLS} = \arg\min \boldsymbol{\epsilon}^{\mathbf{T}} \boldsymbol{\Omega}^{-\mathbf{1}} \boldsymbol{\epsilon}$$

This minimizes the residual's norm using Mahalanobis metric...



A PROBLEM WITH GLS IS THAT Ω IS UNKNOWN!

- Any ideas?

FEASIBLE GENERALIZED LEAST SQUARES

- Feasible Generalized Least Squares (FGLS) solves the problem of Ω being unknown by starting from an estimate of $\boldsymbol{\Omega}$ and iterating over successive estimates until convergence.

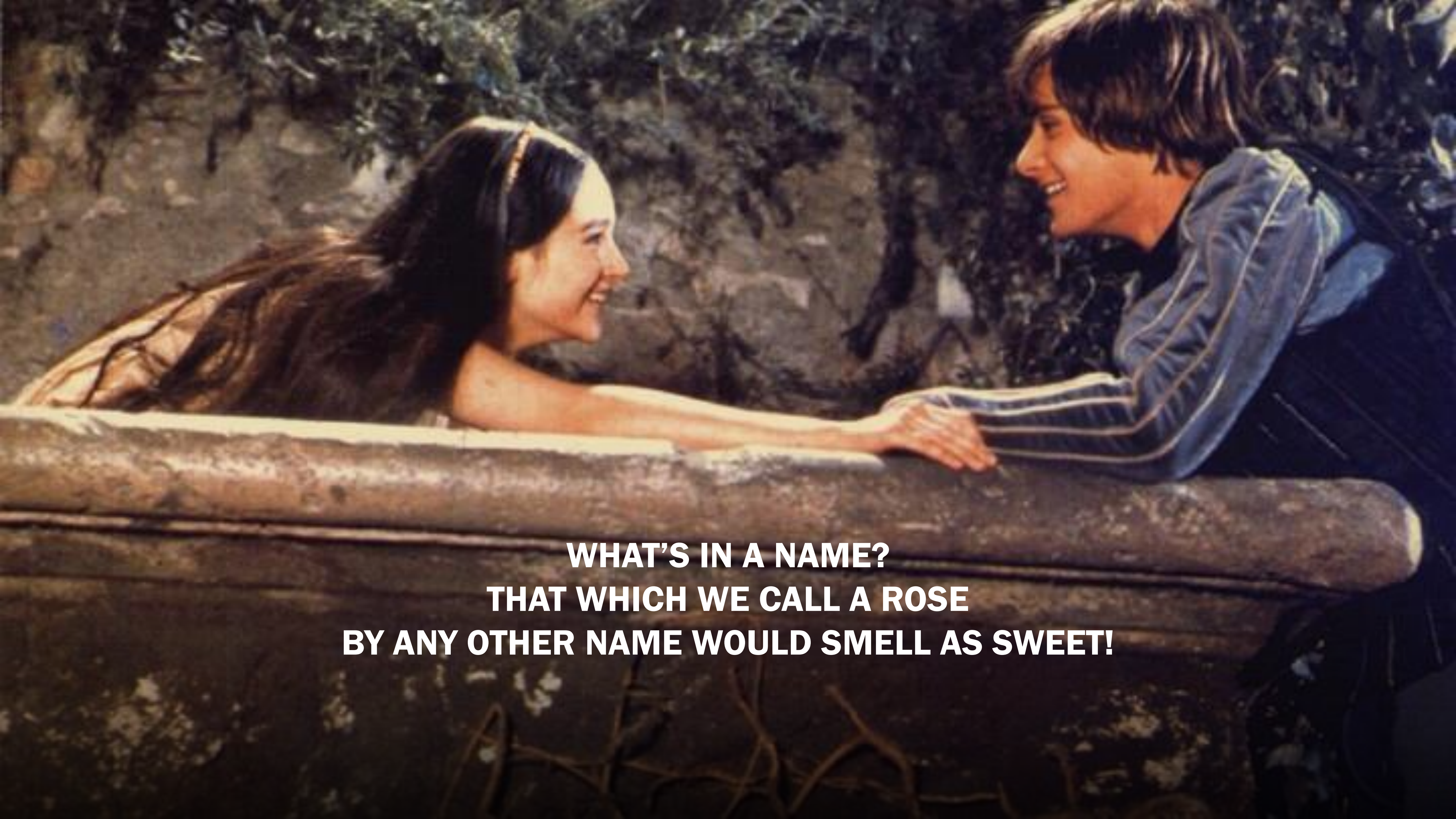
IS INVESTMENT DETERMINED BY COMPANY VALUE? THE GRUNFELD DATASET

#Random Effects model

fe <- plm(invest ~ value + capital, index=c(“firm”, “year”), data=Grunfeld, model=“random”)

	Dependent variable:	
	invest	
	Fixed Effects (1)	Random Effects (2)
value	0.110***	0.109***
	(0.011)	(0.010)
capital	0.310***	0.308***
	(0.017)	(0.016)
Constant		-53.944**
		(25.697)
Observations	220	220
R2	0.767	0.770
Adjusted R2	0.753	0.768
F Statistic	340.079*** (df = 2; 207)	726.428***
Note:		*p<0.1; **p<0.05; ***p<0.01

WHAT’S “RANDOM” IN RANDOM EFFECTS?


WHAT’S IN A NAME?
THAT WHICH WE CALL A ROSE
BY ANY OTHER NAME WOULD SMELL AS SWEET!

“The crucial distinction between fixed and random effects is whether the unobserved individual effect embodies elements that are correlated with the regressors in the model, not whether these effects are stochastic or not” (William H. Greene)

“The crucial distinction between fixed and random effects is whether the unobserved individual effect embodies elements that are correlated with the regressors in the model, not whether these effects are stochastic or not” (William H. Greene)

$$Cov(U_i; X_{it}) = 0?$$
$$Cov(U_i; X_{it}) \neq 0?$$

FIXED EFFECTS OR RANDOM EFFECTS?

- Essentially, this is tantamount to testing whether $\text{Cov}(U_i; X_{it}) = 0$ or not.
- Indirectly, this is done by a Hasuman test.
- The Hausman test helps us decide whether we should use a Random Effects model or a Fixed Effects model. It's hypotheses are:

$$\begin{cases} H_0\text{: The covariance is zero, so you should use Random Effects} \\ H_a\text{: The covariance is } not \text{ zero, so you should use Fixed Effects} \end{cases}$$

So if $p < \alpha$, use Fixed Effects, and if $p > \alpha$, use Random Effects.

KEY TAKEAWAYS

1. When the unobserved effect U_i is uncorrelated with X_{it}, Random Effects yield better estimates than Fixed Effects.
2. The Hausman test helps us decide if we should use a Fixed or a Random effects model. Its null hypothesis is that we should use a Random Effects model.