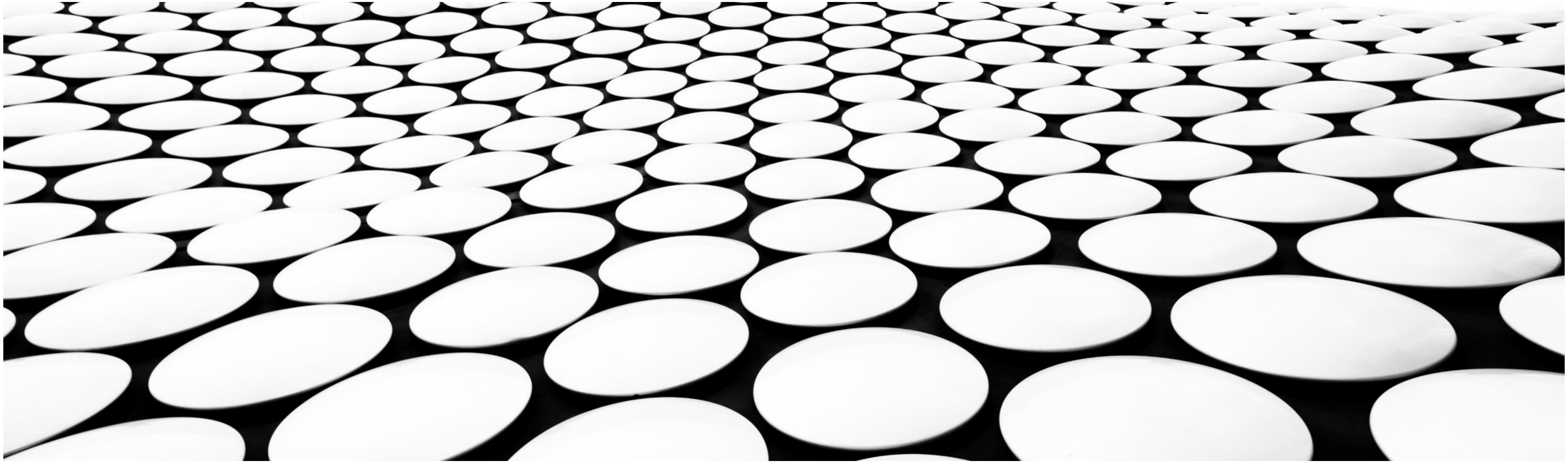
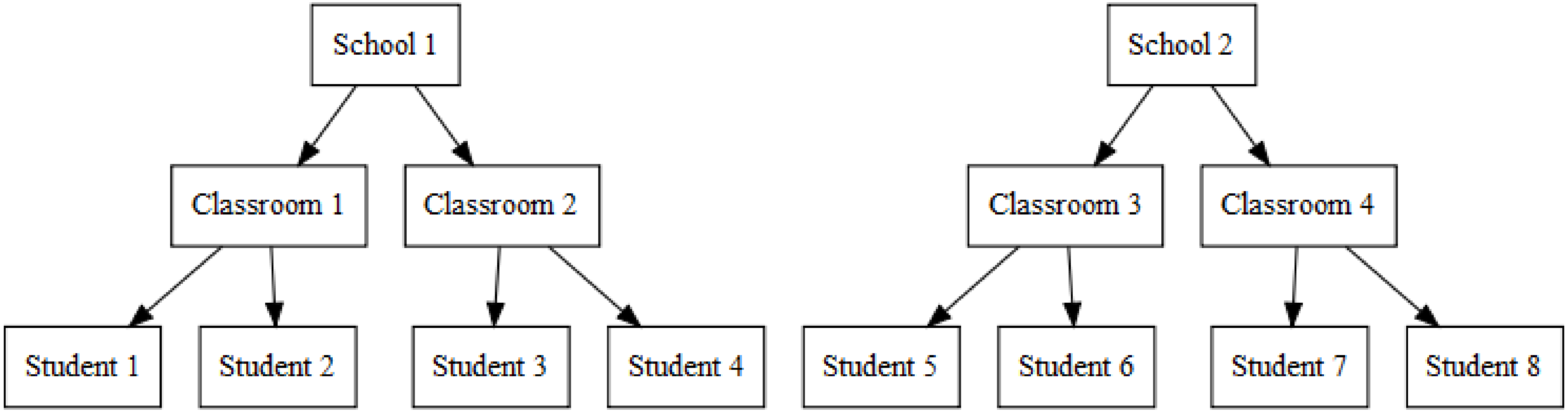
MULTI-LEVEL REGRESSION MODELS

FELIPE BUCHBINDER



PANEL DATA REGRESSION DOESN'T HAVE TO BE ABOUT TIME

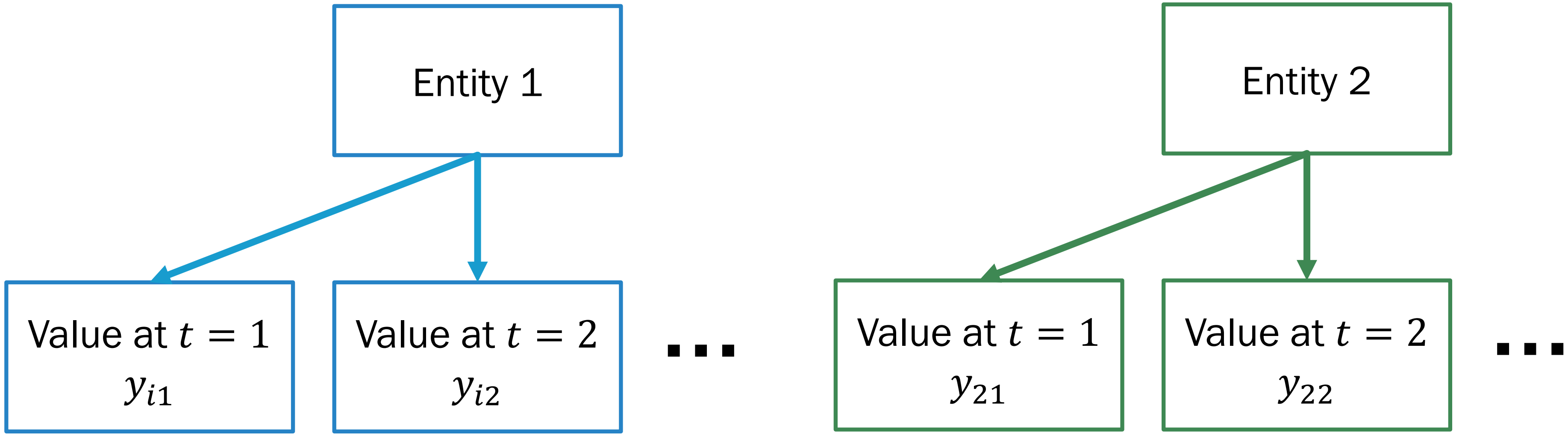


MANY RELEVANT PROBLEMS IN SOCIAL SCIENCES CAN BE TACKLED WITH PANEL DATA EVEN THOUGH THEY HAVE NOTHING TO DO WITH THINGS CHANGING IN TIME.

A MENTAL IMAGE FOR PANEL DATA

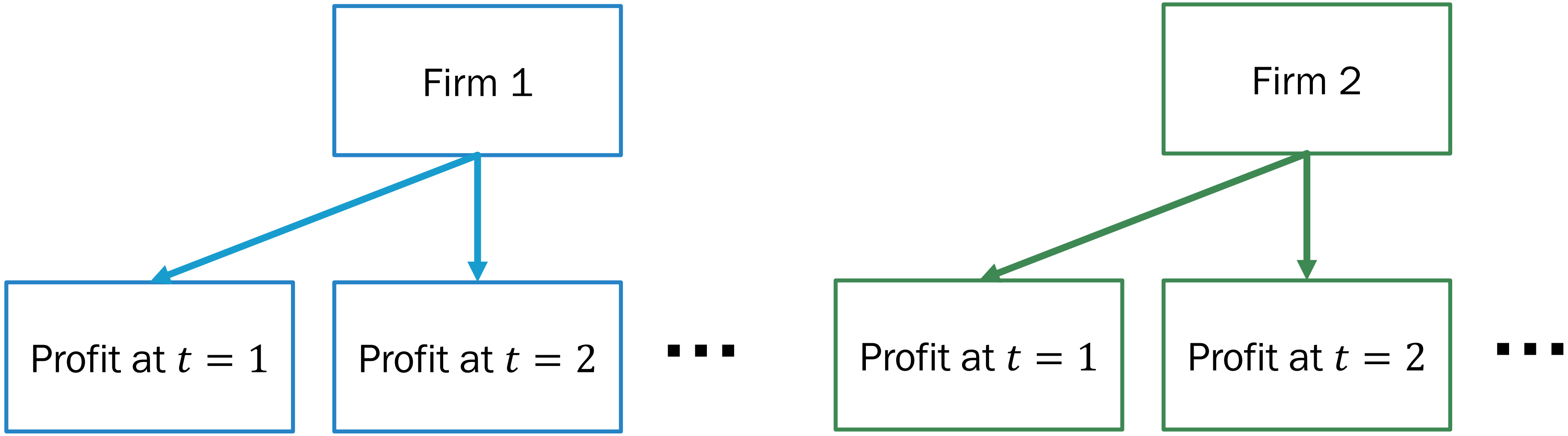
(USING TIME)

Many entities. Each entity produces one value for each time period



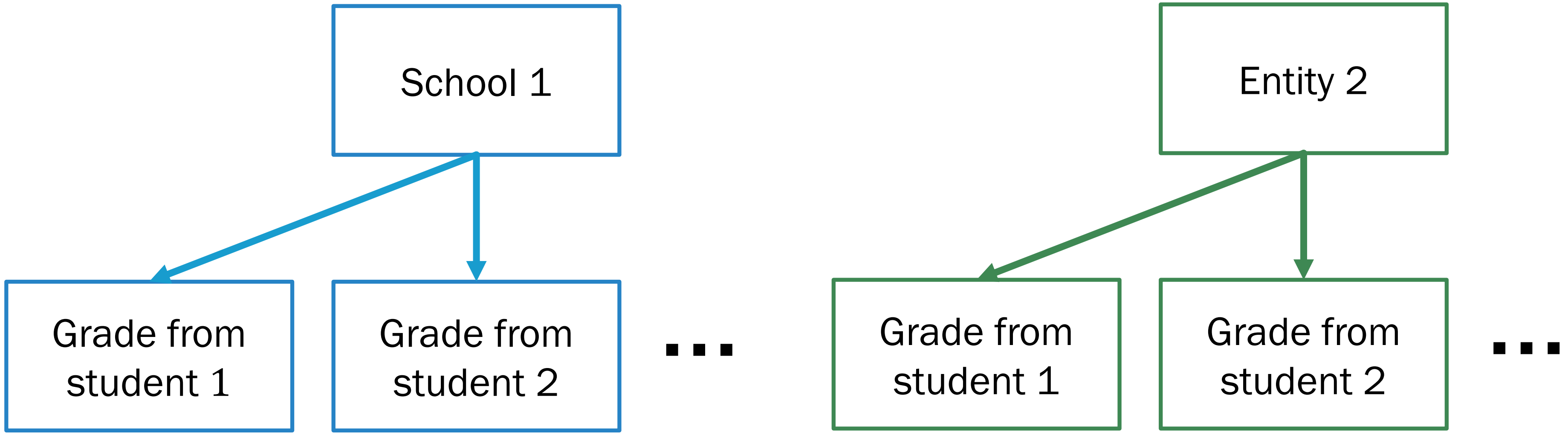
FOR EXAMPLE…
(USING TIME)

Many firms. Each firm produces one value for profit at each quarter



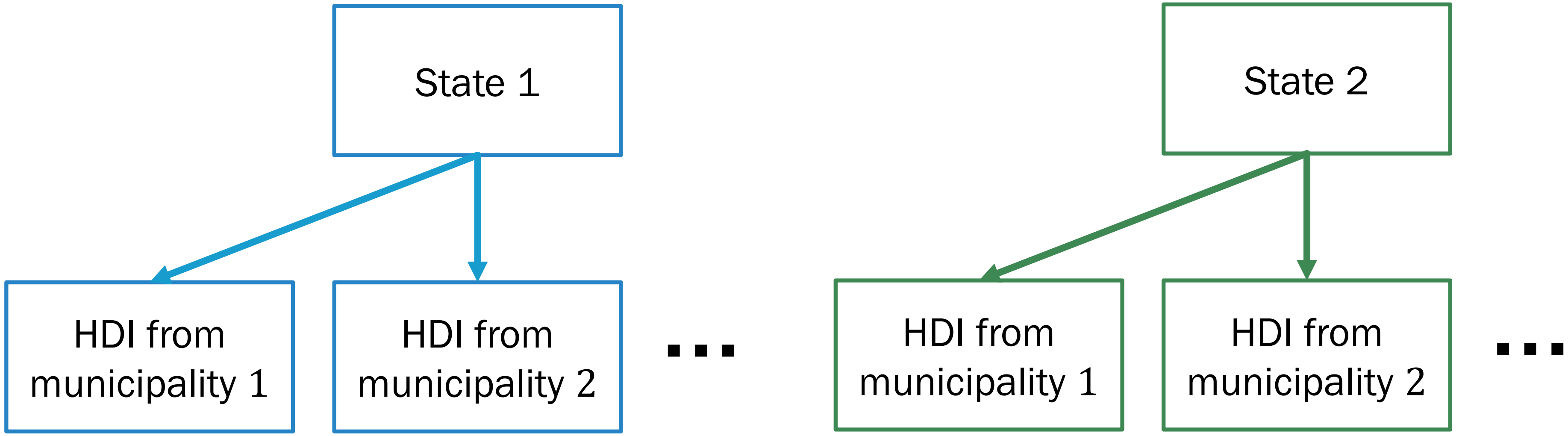
AN ANALOGOUS PROBLEM
(THAT DOESN'T USE TIME)

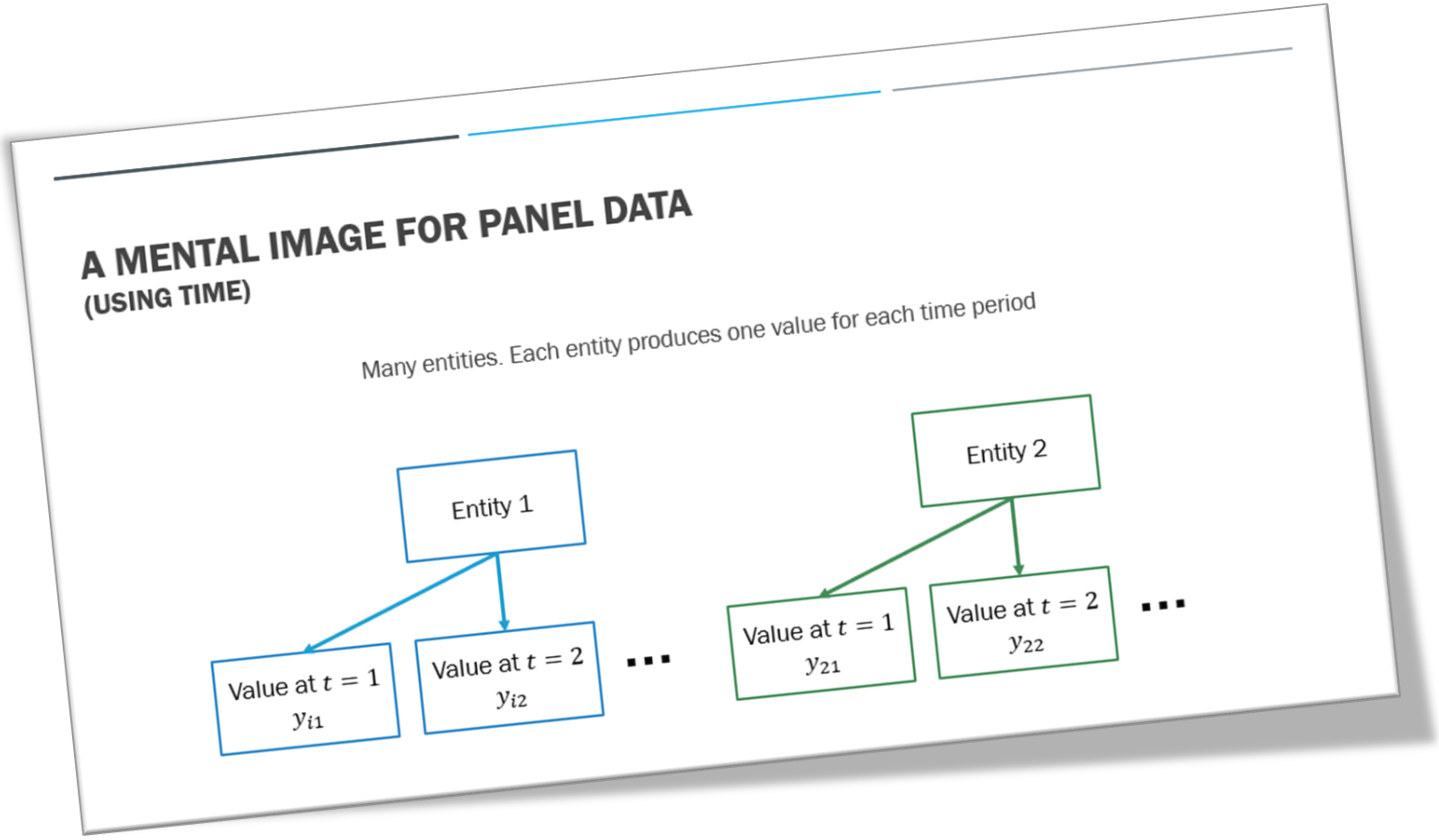
Many schools. Each school produces one SAT grade for each student



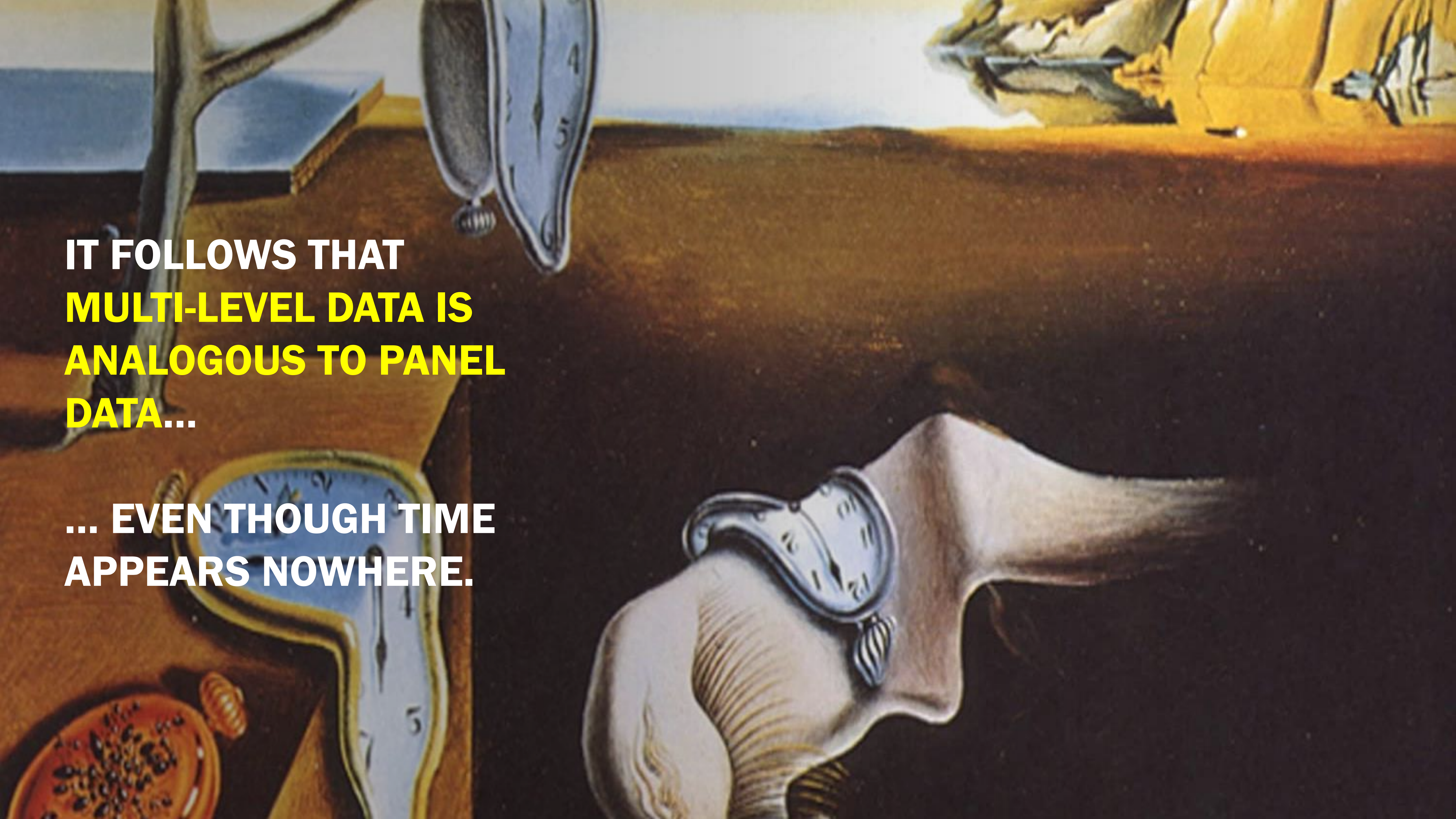
ANOTHER ANALOGOUS PROBLEM
(THAT DOESN'T USE TIME)

Many states. Each state produces one HDI for each municipality





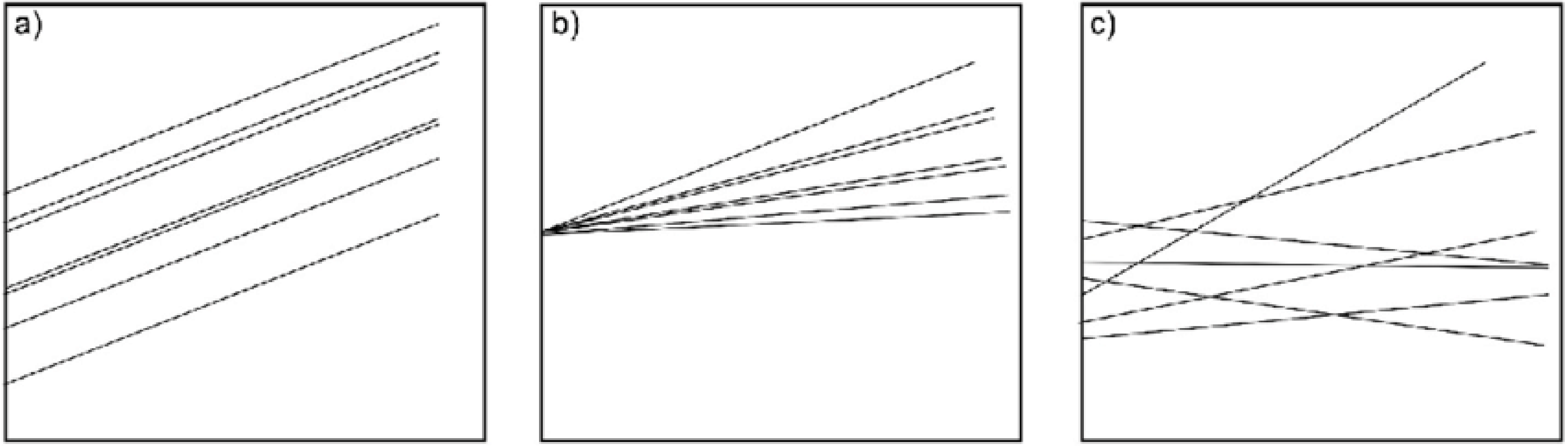
Panel data is analogous to multi-level data where level-1 is time and Level-2 is entity.



IT FOLLOWS THAT
MULTI-LEVEL DATA IS ANALOGOUS TO PANEL DATA...

... EVEN THOUGH TIME APPEARS NOWHERE.

3 TYPES OF MULTI-LEVEL REGRESSION



Random Intercepts

Random Slopes

Random Intercepts & Slopes

3 TYPES OF MULTI-LEVEL REGRESSION

Random Intercepts is the most common, so let's focus on this one.

Random Intercepts

AN EXAMPLE WILL HELP…

- Suppose we wish to estimate the distribution of test scores for students at J different high schools.
- In each school j, where $j = 1, \dots, J$, suppose we test a random sample of n_j students.
- Let Y_{ij} be the test score from the i-th student at school j, with $i = 1, \dots, n_j$.

THE RANDOM INTERCEPTS MODEL

Level 1 equation: $Y_{ij} = \beta_{0j} + \beta_1 X_{ij} + \eta_{ij}$

Level 2 equation: $\beta_{0j} = \gamma_0 + \gamma_1 Z_j + \nu_j$

THE RANDOM INTERCEPTS MODEL

Test score from student i at school j

Student's characteristics

Level 1 equation: $Y_{ij} = \beta_{0j} + \beta_1 X_{ij} + \eta_{ij}$

Level 2 equation: $\beta_{0j} = \gamma_0 + \gamma_1 Z_j + \nu_j$

Each school (j)
has it's own intercept

Same slope for all students

THE RANDOM INTERCEPTS MODEL

School's characteristics

Level 1 equation: $Y_{ij} = \beta_{0j} + \beta_1 X_{ij} + \eta_{ij}$

Level 2 equation: $\beta_{0j} = \gamma_0 + \gamma_1 Z_j + \nu_j$

Intercept of school j

MULTI-LEVEL REGRESSION SOLVES THIS PROBLEM BY USING A BAYESIAN APPROACH

Level 1 (students): $Y_{ij}|\mu_j, \sigma_j^2 \sim \mathcal{N}(\mu_j; \sigma_j^2)$

Level 2 (schools): $\mu_j|\mu_0, \tau^2 \sim \mathcal{N}(\mu_0; \tau^2)$

μ_0 and τ^2 are unknown and have to be estimated using Maximum Likelihood or Bayesian methods.

MULTI-LEVEL REGRESSION ARRIVES AT THE FOLLOWING SOLUTION FOR THE EFFECT OF EACH SCHOOL ON STUDENT'S GRADES:

$$\mu_j | Y_{ij} \sim \mathcal{N}\left(\frac{\frac{n_j}{s_j^2}\overline{y}_j + \frac{1}{\hat{\tau}^2}\overline{y}_{\text{all}}}{\frac{n_j}{s_j^2} + \frac{1}{\hat{\tau}^2}} \; ; \; \frac{1}{\frac{n_j}{s_j^2} + \frac{1}{\hat{\tau}^2}} \right)$$

MULTI-LEVEL REGRESSION ARRIVES AT THE FOLLOWING SOLUTION FOR THE EFFECT OF EACH SCHOOL ON STUDENT'S GRADES:

Suppose all schools have the same σ_j^2. What happens if n_j is very large? What happens if it is very small?

Suppose all schools have the same n_j. What happens if σ_j^2 is very large? What happens if it is very small?

$$\mu_j | Y_{ij} \sim \mathcal{N}\left(\frac{\frac{n_j}{s_j^2}\bar{y}_j + \frac{1}{\hat{\tau}^2}\bar{y}_{\text{all}}}{\frac{n_j}{s_j^2} + \frac{1}{\hat{\tau}^2}} \; ; \; \frac{1}{\frac{n_j}{s_j^2} + \frac{1}{\hat{\tau}^2}} \right)$$

EFFECTS OF SCHOOL SIZE

Large n_j $(n_j \to +\infty)$

$$\mu_j | Y_{ij} \sim \mathcal{N}\left(\bar{y}_j; 0\right)$$

If there's a lot of data about a school, multi-level regression uses only this data to make inferences about it and ignores data of other schools.

Small n_j $(n_j \to 0)$

$$\mu_j | Y_{ij} \sim \mathcal{N}(\bar{y}_{\text{all}}; \hat{\tau}^2)$$

If there's not much data about a school, multi-level regression uses information from other schools to infer about it.

EFFECTS OF SCHOOL HETEROGENEITY

Large s_j^2 ($s_j^2 \rightarrow +\infty$)

$\mu_j | Y_{ij} \sim \mathcal{N}(\bar{y}_{\text{all}}; \hat{\tau}^2)$

Small s_j^2($s_j^2 \rightarrow 0$)

$\mu_j | Y_{ij} \sim \mathcal{N}\left(\bar{y}_j; 0\right)$

The multi-level model shrinks estimates with high variance towards the grand mean.

SINCE THERE'S AN ANALOGY BETWEEN PANEL DATA AND MULTI-LEVEL DATA, COULD I SOLVE THIS PROBLEM USING PANEL DATA METHODS?

THE RANDOM INTERCEPTS MODEL

Level 1 equation: $Y_{ij} = \beta_{0j} + \beta_1 X_{ij} + \eta_{ij}$

Level 2 equation: $\beta_{0j} = \gamma_0 + \gamma_1 Z_j + \nu_j$

EXTREMELY ADVANCED & DIFFICULT MATH

Level 1 equation: $Y_{ij} = \beta_{0j} + \beta_1 X_{ij} + \eta_{ij}$

Level 2 equation: $\beta_{0j} = \gamma_0 + \gamma_1 Z_j + \nu_j$

$$\therefore$$

$$Y_{ij} = \gamma_0 + \beta_1 X_{ij} + \gamma_1 Z_j + \epsilon_{ij}$$

NOW, COMPARE…

Model	Equation
Random Intercept Multi-Level Regression	$Y_{ij} = \gamma_0 + \beta_1 X_{ij} + \gamma_1 Z_j + \epsilon_{ij}$
Panel Data's Fundamental Equation	$Y_{it} = \beta_0 + \beta_1 X_{it} + U_i + \epsilon_{it}$

CHALLENGE QUESTION

Is a Multi-Level Random Intercepts model equivalent to a Fixed Effects or to a Random Effects model?

RANDOM EFFECTS

SINCE BOTH REQUIRE THAT CHARACTERISTICS OF STUDENTS AND SCHOOLS BE UNCORRELATED

THIS MAKES PANEL DATA BETTER THAN MULTI-LEVEL REGRESSION, BECAUSE IF SCHOOL CHARACTERISTICS *DO* CORRELATE WITH THE ERROR TERM, WE CAN STILL USE FIXED EFFECTS AND IT WILL WORK JUST FINE.