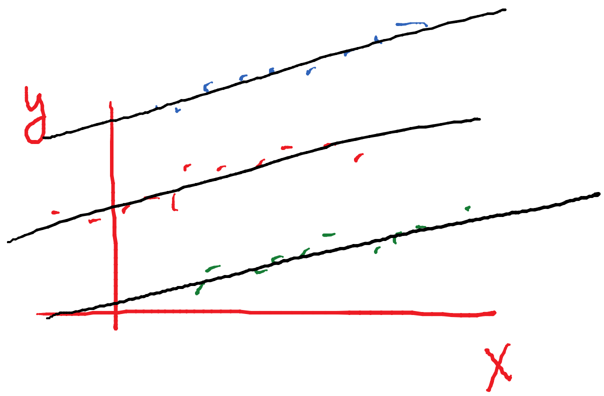
Introduction to Hierarchical Linear Modeling

Thursday, April 19, 2012
10:01 AM

- HLM is an approach for dealing with various forms of unit heterogeneity
- There are different forms of unit heterogeneity, some more complex than others
- HLM is a form of random effects model



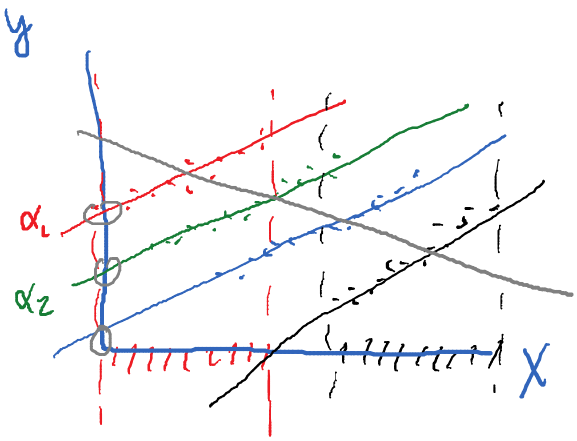
Review: Unit Heterogeneity and its Consequences

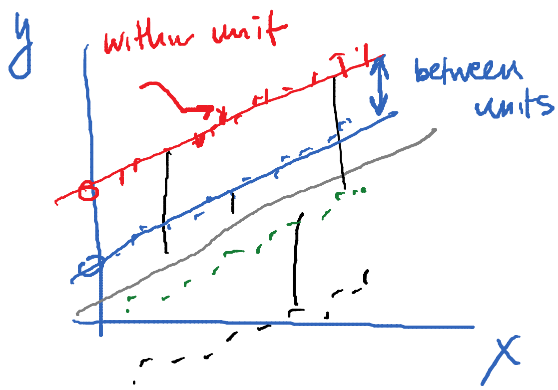
Thursday, April 19, 2012
2:38 PM

- The consequences of neglecting unit heterogeneity differ according to the nature of that heterogeneity
    - Key principle: unit heterogeneity correlated with X or not

bias: X and α are correlated

efficiency: X and α NOT correlated.





0

bias an issue

efficiency

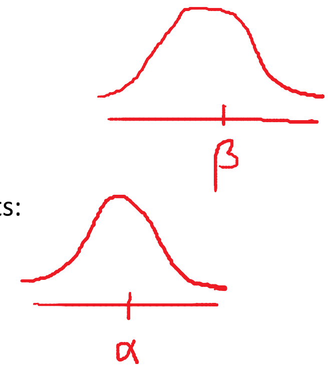
- The models from last time are designed to deal with a specific form of unit heterogeneity: unit-specific intercepts
    - Fixed effects: put in dummy variables per unit to estimate α for each unit
    - Random effects: assume unit effects uncorrelated with X and drawn from a common distribution
        - $y = \alpha^i + \beta X + \epsilon$ and $\alpha^i \sim \Phi(\bar{\alpha}, \sigma_\alpha^2)$

$\bar{a} = 0$

- But this is not the only form of unit heterogeneity that is possible...

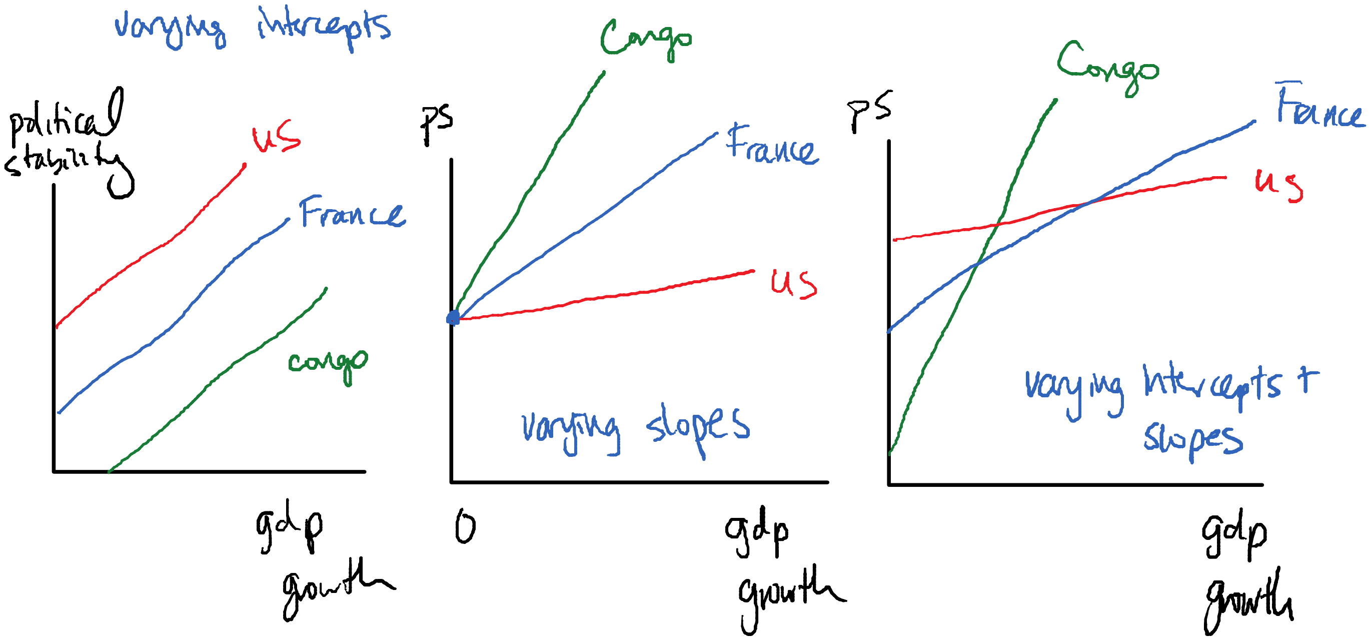
Varying Slopes and Intercepts

Thursday, April 19, 2012
11:21 AM



- Unit heterogeneity can also come in the form of varying slopes AND intercepts:
  - $y = \alpha^i + \beta^i X + \epsilon$
  - $\alpha^i \sim \Phi(\alpha, \sigma_\alpha^2), \beta^i \sim \Phi(\beta, \sigma_\beta^2)$
- What ~~do these models look like?~~

α = mean intercept
β = mean slope



- The random effects can, if you specify it so, be correlated with each other--or independent.

$$y = \alpha^i + X\beta^i + \varepsilon, \quad \begin{bmatrix}\alpha^i \\ \beta^i\end{bmatrix} \sim \Phi\left(\begin{bmatrix}\alpha \\ \beta\end{bmatrix}, \Sigma\right)$$

The "Hierarchy" in Hierarchical Models

Thursday, April 19, 2012
2:01 PM

A hierarchical model is typically structured so that there are multiple layers of random effect that are nested inside one another. So, for example:

50 US states

Regional — South, Midwest

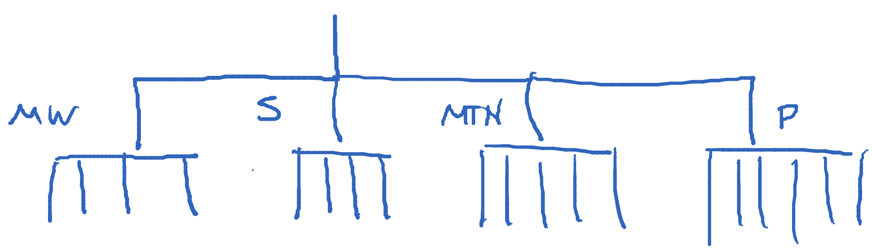
State — GA, AL   OH, IA, IN

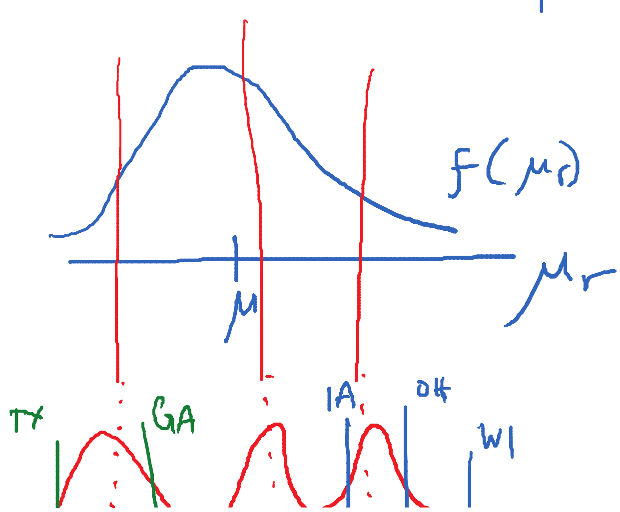
$$y \sim \alpha^i + X\beta + \varepsilon$$

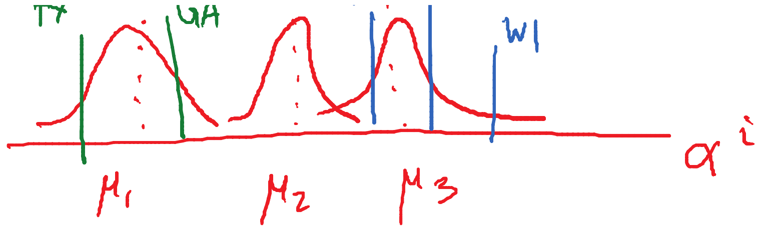
regional mean

$$\alpha_i \sim \Phi(\mu_r, \sigma_\alpha^2)$$ dist. of unit effects w/ common regional mean

$$\mu_r \sim \Phi(\mu, \sigma_r^2)$$ distribution of regional means







Not all mixed effects models are hierarchical models.

$$y = \alpha^i + \gamma^i + X\beta + \varepsilon$$ people

where $\alpha^i \sim \Phi(\alpha, \sigma_\alpha^2)$ country

$\gamma^i \sim \Phi(\gamma, \sigma_\gamma^2)$ year

The Consequences of Ignoring Hierarchy

Thursday, April 19, 2012
2:09 PM

What is the impact of ignoring the hierarchical structure of a data set?

Answers:

1) efficiency loss

gets us narrower SEs for $\hat{\beta}$ than we could achieve via pooling

   a. Note: "fixed effect" doesn't mean the same thing in the HLM context

"fixed effects": characteristics that are not random

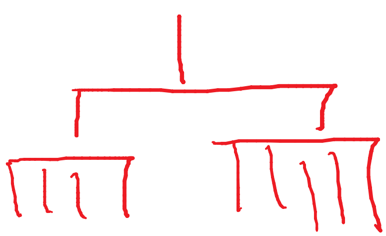
$$y \sim \boxed{\alpha^i} + X\beta + \varepsilon$$

$$y \sim \boxed{\alpha^i} + \boxed{X\beta^i} + \varepsilon$$

$$\beta^i \sim \phi(\beta, \sigma_\beta^2)$$

$$\alpha^i \sim \phi(\alpha, \sigma_\alpha^2)$$

2) neglect interesting structural aspects of panel heterogeneity



region
state

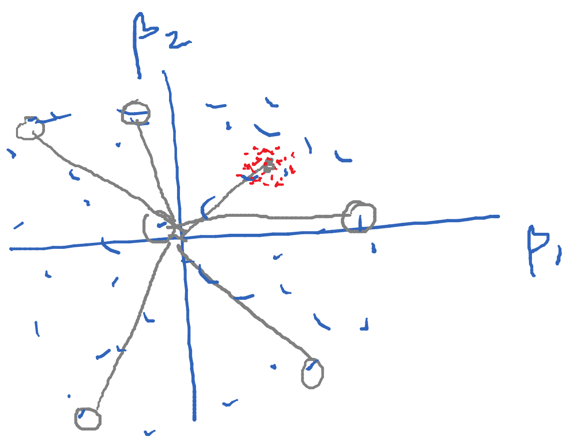
world
region
country

$\boxed{\text{MRP}}$

Bias-Variance Tradeoff

Thursday, April 19, 2012
2:52 PM

- RE models (including HLMs) are biased (and inconsistent) when X and ϵ are correlated, but efficient
- FE models are consistent but inefficient due to the incidental parameters problem
- Which to use?
- It depends… this can be illustrated with an example in R



Facts about HLM

Thursday, April 19, 2012
2:19 PM

A few facts about HLM estimates:

1) Estimates of unit heterogeneity will tend to compromise between the results of two estimators:

complete pooling $\rightarrow$ $\hat{y} \sim X\hat{\beta}$

no pooling $\rightarrow$ $\hat{y}_i = X_i \hat{\beta}_i$ $\quad i = 1 \ldots N$

N = # of units

↳ intercepts vary: LSDV/FE

↳ intercepts + slopes: separate regression for each panel

2) The similarity of HLM to these two estimators depends on the nature of the heterogeneity among units

if unit heterogeneity is small, HLM $\approx$ complete pooling.

if unit het. is large, HLM $\approx$ no pooling

HLM is most useful low/moderate heterogeneity.

3) The fewer the observations per unit, the closer the estimate for that unit is to the complete pooling estimator

	VA	GA	WI
T	8	53	61

borrowing strength

complete pooling

1) These models can get very complicated very quickly. Beware the proliferation of parameters and the difficulty of interpreting the resulting complicated models