# Unit 3:
# Multilevel Panel and Growth Models
# I. Introduction

PS2701-2019

Longitudinal Analysis

Week 9

Professor Steven Finkel

# The Role of “Time” in Panel Models

- Thus far, we have included “time" in panel models in order to disentangle the effects of independent variables or “treatments” from the effects of the passage of time, or from events that take place at particular times that affect all units, regardless of “treatment” status
- In First Difference/Treatment Effects models, “time” captured changes in the dependent variable that occurred for the control group, and we assumed the treatment group would have shown the same change in the absence of treatment. This identified the “treatment effect” as the “difference in differences” between treatment/control
- In the “Two-way” FE/RE/Hybrid models, we included dummy variables for each wave in order to capture changes the occurred from wave to wave for all units, separate from the effects that may have been produced by changes in the independent variables
- In SEM versions of these models: allow the intercept for Y to change for different waves

- More complex variants of the “time as a control” or “time as a nuisance” variable strategy:
  - With multiwave data, include a separate time trend in Y for treatment and control groups in order to test/control for the “parallel trends” assumption and estimate treatment effects
  - As T increases, also important to include time as a control for “non-stationarity” in the series, such that both Y and X may exhibit trends (i.e. changing means) that could result in the well-known problem of “spurious regression”. [Example: democracy and economic development both increased globally in the post-WW II period, and a spurious causal connection between the two could exist without taking the common trends into account].
  - Wang and Maxwell (2015) and Curran and Bauer (2011) apply this logic in recommending the inclusion of “time" in panel models in order to control for trends in both Y and in X, which will then allow the proper (unconfounded) estimation of “within” and “between” effects of the independent variables

- An alternative conception of "time": it may represent **substantively interesting "growth" (either positive or negative)** in a dependent variable. That is, Y at time *t* is related directly to TIME, or due to the **progression** or **growth** of the unit through time.
- Increasingly popular in many areas of social sciences, e.g. psychology, education, biostatistics
  - Children's acquisition of vocabulary or motor skills as function of time
  - Delinquency behaviors in adolescents and young adulthood
  - Sexual behaviors post-puberty
  - Recovery from depression, post-therapeutic treatment
- In all of these examples, we can talk about ***growth*** over time, and we could model the outcome for each individual as a positive or negative function of time itself. As months past puberty increase, amount of sexual activity increases, etc. This is a very intuitive and very direct way to talk about change.

- Modeling “growth” has traditionally been conducted in the “multilevel” panel tradition, and so we will use growth models a way to deepen our general understanding of **multilevel panel models,** the third major analytic framework for panel analysis (econometrics and SEM being the other two)
- **NOTE:** The dependent variable in growth (multilevel) models *need not be* continuous, though this is how we will introduce the topic .
  - In the examples above, in fact, the DV might be conceived (no pun intended) as the occurrence of sexual behaviors of some kind, or in the second example, as a frequency or count of the number of delinquent acts at a given point in time. In the last example, we can model the “growth” in some ordinal measure of non-depression (so to speak), i.e., recovery, as a function of time since a treatment for an experimental and control group.
  - We can model all of these kinds of growth processes in non-continuous dependent variables within the framework of **“Multilevel Generalized Linear Mixed Models,” (or its SEM equivalent of “Generalized SEM”)** which we will cover as time permits. For now, we will stick to continuous outcomes to talk about the basic ideas of the approach and basic modeling and estimation issues (which are complicated enough in the linear case!)

# Growth Models in Political Science

- Relatively few, but promising method since many DVs in our discipline also involve growth
  - Plutzer, "Becoming a Habitual Voter," 2001 *American Political Science Review*: Models growth in the likelihood of voting among people entering the electorate. The model includes each election after which individual is eligible, so can talk of *developmental* process of vote. Models influences on the initial vote decision and then the likelihood of subsequent voting as a function of time and other variables.
  - Finkel, Pérez-Liñán, Seligson, "The Effects of US Foreign Assistance on Democracy Building, 1990-2003," 2007 *World Politics*: Models the growth in democracy among countries world-wide from 1990-2003. Using Polity IV and Freedom House measures, it plots trajectories of democratic growth as a function of time, and then models the factors determine that determine the initial 1990 starting point and then the growth rates over time.
  - Paxton, Hughes and Painter, 2010 *European Journal of Political Research* model growth in women's representation 1975-2000 in 110 countries as a function of the existence of gender quotas, overall growth in democracy, and other factors that vary over countries and/or over time

- Applications to other areas of political science:
  - Growth in system or regime support among individuals, either in new or advanced democracies
  - Growth in party identification in new democracies
  - Growth in party support during a political campaign
  - Growth in budgets, revenues, political violence, trade, democracy, human rights violations, etc. etc.

# The Basic Growth Model

- We can start to model this idea of growth very simply with:

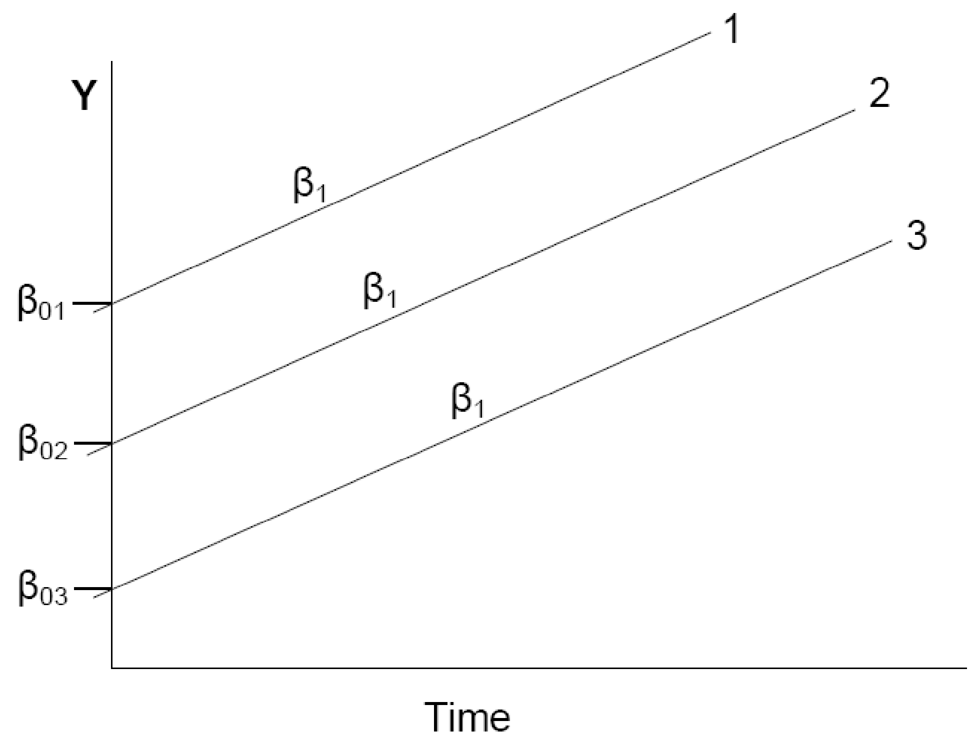
$$(1) \qquad Y_{it} = \beta_0 + \beta_1 Time_i + \varepsilon_{it}$$

- which models Y for a particular unit at a particular point in time as a function of TIME. The regression coefficient would be interpreted in accordance with the unit of time (years, months, days, etc.) that was chosen for the given model.
- Note on subscripts: Time formally has an *i* subscript because different individuals are observed at different waves. In a balanced design you would not need the *i* because everyone would have the same values for Time at all times. You also could put a *t* in Time *i* subscript but this seems superfluous, since the value of Time at any *t* is already *t*. I will omit the *i* subscript from here on, but it is implicitly there.

- But this model only contains one of the two main ideas from this approach. In fact it is a model that applies equally to every unit, as every case in the sample (and population) is assumed to be governed by the same parameters ($\beta_0$ and $\beta_1$) in determining Y. This is not realistic in the panel situation, due to…
- **HETEROGENEITY** of the units! This is a big reason why we do panels in the first place
  - We need at least to take into account the fact that observations generally will be higher or lower on Y due to unmeasured stable factors (previously labeled $U_i$); alternatively,
  - We have also said that we include unit-level heterogeneity to model the fact that the *same* unit is being observed over time, and so the observations are not independent. We model the clustered nature of the data so as to make correct inferences (i.e. correct standard errors of estimated parameters) and to model possible differences in the outcomes and the effects of variables on the outcomes over time.

# How to Incorporate Heterogeneity into the Growth Model?

- One way: different units are likely to have different "starting" points or intercepts as well, exactly as in the models we've looked at with fixed or random effects in the non-growth model context.




GRAPH 1:

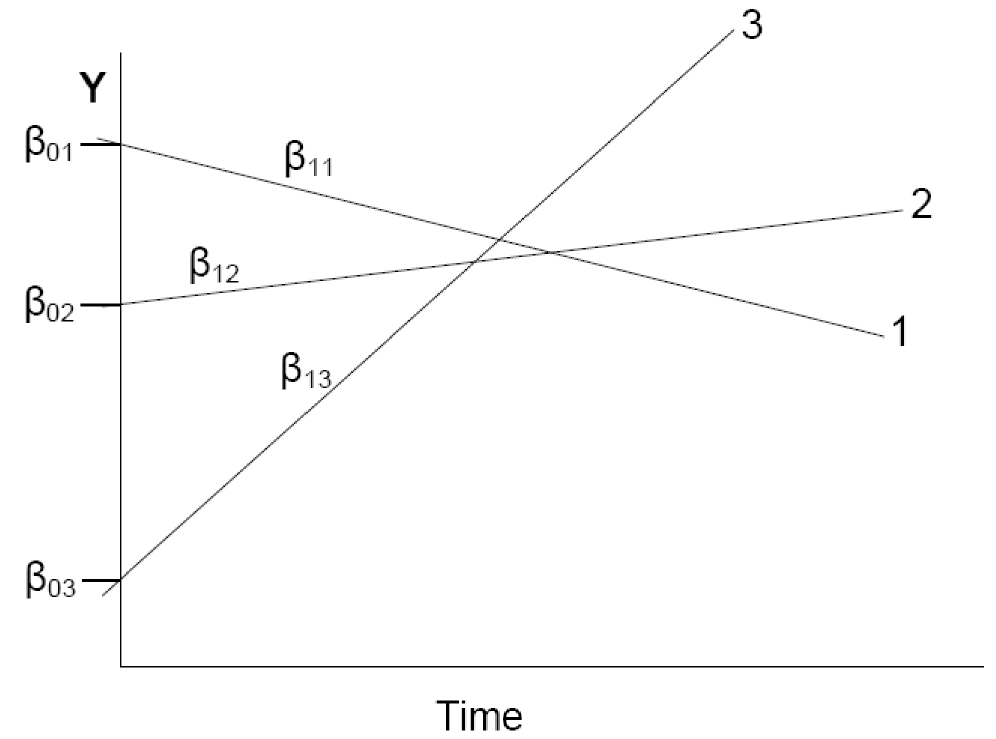
COMMON TIME TREND, HETEROGENEITY IN INTERCEPTS

- Some cases might start higher or lower than others, and stay that way regardless of the overall effect of TIME. In the FE tradition, we have a dummy variable that stands for this unobserved unit-specific effect, while in the RE tradition, we assume a normal distribution for the unit effects, and a random draw from this distribution makes a given case higher or lower at all times from the common intercept value.

$$(2) \qquad Y_{it} = \beta_{0i} + \beta_1 Time + \varepsilon_{it} \text{ and } \beta_{0i} = \beta_{00} + U_i$$

- Where $\beta_{0i}$ is the intercept for a given case, separate from the influence of TIME, and is composed of the common intercept plus the individual unit effect
- This is old news to us!!!!

- But growth models --- and the more general multilevel framework --- also allow another kind of heterogeneity, unit-level differences in the ***slopes or rates of change*** over time as well. That is, in growth models we allow for possible differences in $\beta_1$ across units, so that some units may grow at a faster rate than others.




GRAPH 2:

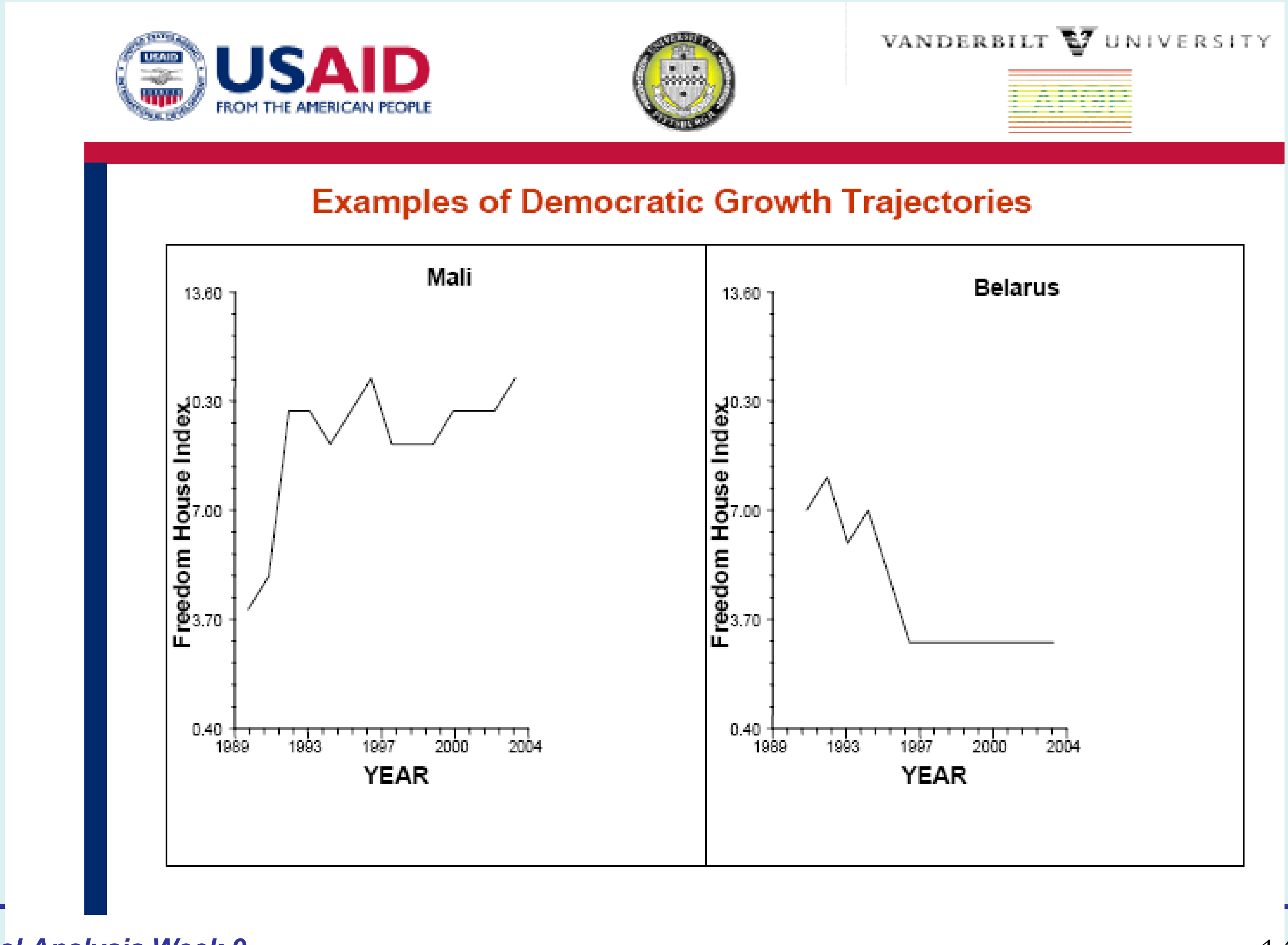
HETEROGENEITY IN INTERCEPTS AND TIME-TREND SLOPES

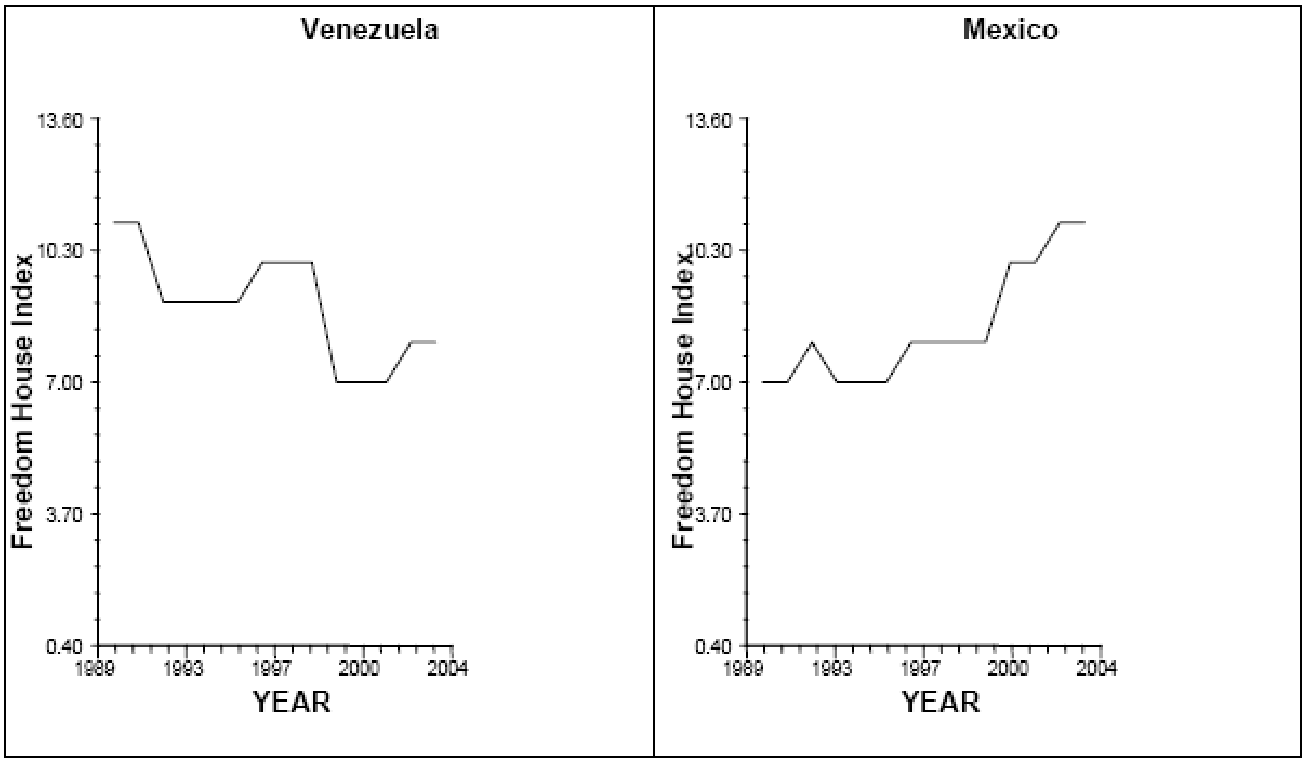
- In terms of the above examples: Some children pick up vocabulary at faster rates than others, some countries increase in democracy at faster rates than others, some adolescents develop delinquent behaviors more rapidly over time than others, some people change more strongly in their support for a candidate during a particular campaign than others. All of these processes necessitate heterogeneity in the TIME effect and not only in the INTERCEPT effect.

$$(3) \qquad Y_{it} = \beta_{0i} + \beta_{1i} Time + \varepsilon_{it}$$

- where each case now has its own intercept as well as its own slope for growth over time. Each unit has its own "growth trajectory"
- This is the basic structure of the growth model: individual units start at different points and change at different rates on some dependent variable of interest. This structure accounts for unit-level clustering with the different intercepts – which push cases up or down in general compared to other cases – and allows different rates of change from case to case as well.

# Examples: Finkel/ Pérez-Liñán/Seligson (2007) on growth in democracy





Venezuela
Mexico
13.60
10.30
7.00
3.70
0.40
Freedom House Index
1989
1993
1997
2000
2004
YEAR

# Modeling Unit-Specific Growth Trajectories

- Next step: use the pooled data for all cases to **model** the differential growth trajectories for different units, that is, to determine *what unit-level factors predict the units' rates of change and their "starting points" or growth trajectory intercepts*.
- Examples:
  - Children with educated parents learn vocabulary more quickly than children with non-educated parents
  - Countries with higher levels of ethnic fractionalization democratize less quickly than others
  - Patients who were treated with one type of therapy recovered from depression more quickly than patients treated with another type
  - Children who come from divorced households are more likely to show increases in delinquent behavior activities over the course of adolescence and early adulthood than children from stable households
- So we model change *within* units over time, and then account for the rates of change *between* units with a series of independent variables

# Key Features of the Growth Model

- Aside from focusing on modeling unit-specific trajectories, two aspects are of most interest:
  - It is a **multilevel** or a **hierarchical** model. The data comprise a hierarchical or "nested" structure, with analysis and modeling taking place at different levels of the hierarchy; and
  - It is a **"mixed"** model -- with some elaboration of equation (3) above, it will be seen as containing both "fixed" and "random" effects (though the former term is used in a *slightly* different sense than what we have talked about so far)
- So the growth model is one type of multilevel, or hierarchical mixed model as applied to longitudinal data
  - There are also multilevel or mixed models for non-longitudinal data
  - There are also multilevel or mixed longitudinal models that are *not* explicitly growth models
  - In fact, *all* of the models we've looked at can fit into the multilevel framework!
- Basic multilevel structure of panels: Time of measurement or "wave" (Level 1) is nested within individual units (Level 2)

# Multilevel Models are Ubiquitous in Social Sciences, Panels or Not

- In **education research**, children (Level 1) are nested within schools (Level 2), nested within districts (Level 3).
- **Survey research**: respondents (Level 1) are (hypothetically) nested within blocks (Level 2) which are nested within Primary Sampling Units (Level 3) which are nested within, for example, U.S. states (Level 4)
- **Cross-national survey research**: respondents (Level 1) are nested within countries (Level 2).. Example: World Values, LAPOP Surveys from one wave
- **Repeated Surveys on the Same Cross-Sectional Unit (Trend Analysis)**: respondents (Level 1) are nested within Time (Level 2), since in this kind of design *different individuals* from the same aggregate unit (country, state) are interviewed at multiple points in time. This is the *reverse* of panel data, where time of measurement (Level 1) is nested within individuals (Level 2). Example: Pooled NES Election data, 1948-2012.
- **Repeated Surveys on Multiple Cross-Sectional Units:** Individuals (Level 1) nested within country-years (Level 2) nested within country (Level 3). Example: World Values, LAPOP Surveys from multiple waves

# The Growth Model as a "Multilevel" Model

- Two different levels of analysis
- ***Level 1: Change at Individual Level*** –That is, ***Intra-Individual*** Growth Over Time as represented by equation (3):

$$(3) \qquad Y_{it} = \beta_{0i} + \beta_{1i} Time + \varepsilon_{it}$$

- Level 1 is the occasion or "wave" of measurement for all units
- Each unit has its own Level 1 intercept and own slope
- By the way: slope is *linear* in this model but need not be limited to the linear case. Could have quadratic, polynomial patterns of change over time (e.g. political science faculty productivity as function of time: up before tenure, dip post-tenure, up once the kids are out of the house…)

- ***Level 2: Unit Level Factors that Account for Inter-Individual Differences in the Level 1 Parameters*.**
  - At Level 2 we have unit-level variables that determine the $\beta_{0i}$ and the $\beta_{1i}$ parameters at Level 1. What characteristics of individual units cause intercepts to be higher or lower? What characteristics of individuals units cause slopes to be higher or lower?
- We can write out one possible Level 2 model as:

$$(a) \quad \beta_{0i} = \beta_{00} + \beta_{01}X_{1i} + \beta_{02}X_{2i}$$

$$(4)$$

$$(b) \quad \beta_{1i} = \beta_{10} + \beta_{11}X_{1i} + \beta_{12}X_{2i}$$

  - $X_1$ and $X_2$ as Level 2 variables predicting:

    in (a): why the intercept $\beta_0$ at Level 1 in equation (3) is higher or lower for some Level 2 units than others, and

    in (b) why the slope $\beta_1$ at Level 1 in equation (3) is higher or lower for some Level 2 units than others.
- Important point: We conceive of the causal processes as occurring at levels that are nested in some hierarchical fashion, and we model processes at each level of the hierarchy

- In the youth delinquency longitudinal example: $X_1$ could be a dummy for parents divorce, and $X_2$ could be how often the child moved before age 13. These would be the "Level 2 explanatory variables".
- Hypotheses:
  - Children from divorced households are more likely to start adolescence at higher levels of delinquency (prediction: a positive effect of $\beta_{01}$ in 4a)
  - Children who moved more often before age 13 are more likely to start adolescence at higher levels of delinquency (prediction: a positive effect of $\beta_{02}$ in 4a)
  - Children from divorced households increase in delinquent behaviors over time at a faster rate (prediction: a positive effect of $\beta_{11}$ in 4b); and
  - Children who move more often before age 13 increase in delinquent behaviors over time at a faster rate (prediction: a positive effect of $\beta_{12}$ in 4b)
- **NOTE:** The variables that are used to predict the intercept value across units need not be the same variables that are used to predict the slope. Theory should be the guide as always (though the default is usually including them in both Level 2 equations).

# The Growth Model as a "Mixed" Model

- Other key aspect of growth models: the prediction of Level 1 coefficients may have ***random*** components, such that the Level 2 variables predict them imperfectly, with some amount of unexplained random variation.
- In the Level 2 models so far in equation (4) there were no error terms; in keeping with the conceptual discussion above, we can also include a random disturbance term in the level 2 equations. This would mean that we can account for ***some*** of the heterogeneity in the intercept and slopes with the independent variables specified at Level 2, but that there is still some random or unexplained variation that remains. This would be represented by a "random effect" in the level 2 equations capturing the unexplained, random variation.

- We can thus modify the Level 2 equations:

$$(a)\quad \beta_{0i} = \beta_{00} + \beta_{01}X_{1i} + \beta_{02}X_{2i} + \zeta_{0i}$$

(5)

$$(b)\quad \beta_{1i} = \beta_{10} + \beta_{11}X_{1i} + \beta_{12}X_{2i} + \zeta_{1i}$$

  where

  $\zeta_{0i}$ is the random portion of the Level 1 intercept that is unexplained by Level 2 variables;

  and

  $\zeta_{1i}$ is random (unexplained) portion of the Level 1 slope
- So the growth model is a kind of random effects model, with random effects for both intercepts and slopes (here, the slope for the variable "TIME"). So it is what we have dealt with before in terms of a random intercept, along with a new random effect term in the slope as well
- **NOTE**: Brüderl and Ludwig (2015) suggest simply putting in a dummy variable for each case in the level 2 equations and turn it into a "fixed effect" growth model (XTFEIS – "fixed effects individual slopes")

- Why a "mixed" model? Can see this by plugging in equation (5) into the Level 1 growth equation (3):

$$(3) \qquad Y_{it} = \beta_{0i} + \beta_{1i} Time_i + \varepsilon_{it}$$

$$(6) \qquad Y_{it} = \underbrace{(\beta_{00} + \beta_{01} X_{1i} + \beta_{02} X_{2i} + \zeta_{0i})}_{\beta_{0i}} + \underbrace{(\beta_{10} + \beta_{11} X_{1i} + \beta_{12} X_{2i} + \zeta_{1i})}_{\beta_{1i}} * Time_i + \varepsilon_{it}$$

- And rearranging to yield:

$$(7) \qquad Y_{it} = \beta_{00} + \beta_{01} X_{1i} + \beta_{02} X_{2i} + \beta_{10} * Time_i + \beta_{11} X_{1i} * Time_i + \beta_{12} X_{2i} * Time_i + \ (\zeta_{0i} + \zeta_{1i} * Time_i + \varepsilon_{it})$$

{ FIXED PART OF THE MODEL } + {RANDOM PART}

- We call all the coefficients before the parentheses in (7) **"FIXED EFFECTS,"** and we call all the coefficients within the parentheses **"RANDOM EFFECTS."** The model is thus "MIXED" in the sense of containing both fixed and random effects

- **FIXED** effects are those coefficients that do not vary across individuals --- everyone in the sample gets the same β. Fixed effects are the "non-stochastic" part of the model. This is a slightly different sense of fixed effects from what we have discussed so far in the course – but only slightly. If we had N-1 dummy variables to represent the unit effect in equation (7), we would simply have added N-1 "fixed effects," because everyone in the sample/population still gets the same β (though only one dummy variable is non-zero per case). So the fixed effects models for the Unit Effect ($U_i$) that we considered earlier are one type of "FIXED EFFECT" in the more general sense we are considering here.
- **RANDOM** effects are those that *do* vary across individuals, as can be seen from the *i* subscript associated with the ζ and ε terms in equation (7). The random effects represent the stochastic, unexplained part of the model, the part that differs by a random amount for each individual case or unit. In this model there are three random effects: a random part of the unit's intercept; a random part of the slope for the TIME variable for each unit; and a random idiosyncratic error at each point in time.
- We can thus describe the growth model as a **Multilevel Mixed Model with Random Intercepts and Random Time Trends.**

- It is very useful to get comfortable with both ways of looking at the model: one way with separate equations for Level 1 and Level 2 processes, as in (3) and (5), and the other way as the full MIXED equation of (7). They are **identical** and you can move back and forth between the two kinds of specifications. In some software programs (e.g. MPLUS, HLM), you program the two level equations separately, while in other software (e.g. STATA, SPSS, SAS, R), you program the MIXED model directly.

$$(3) \qquad Y_{it} = \beta_{0i} + \beta_{1i} Time + \varepsilon_{it}$$

$$(5) \quad \begin{array}{ll} (a) & \beta_{0i} = \beta_{00} + \beta_{01} X_{1i} + \beta_{02} X_{2i} + \zeta_{0i} \\ (b) & \beta_{1i} = \beta_{10} + \beta_{11} X_{1i} + \beta_{12} X_{2i} + \zeta_{1i} \end{array}$$

Separate Level 1 and
Level 2 Equations

$$(7) \qquad Y_{it} = \beta_{00} + \beta_{01} X_{1i} + \beta_{02} X_{2i} + \beta_{10} * Time_i + \beta_{11} X_{1i} * Time_i + \beta_{12} X_{2i} * Time_i + \quad (\zeta_{0i} + \zeta_{1i} * Time_i + \varepsilon_{it})$$

Full “Mixed” Model

- Hierarchical/Multilevel panel models don't *need* a time trend. You can still have randomly varying slopes for *any* independent variable that varies over time.
- Example from last unit: Democracy as a function of USAID Democracy Assistance at a given point in time
  - Level 1: (8) $$Y_{it} = \beta_{0i} + \beta_{1i} X_{1it} + \varepsilon_{it}$$
  - Level 2: (9) $$\beta_{0i} = \beta_{00} + \beta_{01} Z_1 + \zeta_{0i}$$ $$\beta_{1i} = \beta_{10} + \beta_{11} Z_1 + \zeta_{1i}$$
  - At Level 2, we think that the country's level (intercept) of democracy, *and* the effect of AID on democracy in that country, will depend on the country's Human Development ($Z_1$) as measured, for example, by the UNDP index. So we might hypothesize Human Development to *positively* affect the intercept and (perhaps) *negatively* affect the slope, such that AID "works" better in tougher economic/social contexts.
  - This is called a "Random Coefficient Model" (RCM) where the effect of **any** independent variable (X) may be determined by both fixed and random components

- We can also combine this RCM model with a growth model by including TIME as an additional independent variable:

(10) Level 1: $$Y_{it} = \beta_{0i} + \beta_{1i} * Time_i + \beta_{2i} X_{1it} + \varepsilon_{it}$$

(11)Level 2:
$$\begin{aligned} \beta_{0i} &= \beta_{00} + \beta_{01} Z_1 + \zeta_{0i} \\ \beta_{1i} &= \beta_{10} + \beta_{11} Z_1 + \zeta_{1i} \\ \beta_{2i} &= \beta_{20} + \beta_{21} Z_1 + \zeta_{2i} \end{aligned}$$

- So we have fixed and random components for the intercept, fixed and random components for the time trend slope, and fixed and random components for the slope of the AID variable at Level 1. In this model we would call AID (X) a "time-varying covariate."
- Note: The effects of time-varying covariates **need not necessarily** have random component associated with them, but they may. This is your modeling choice. The same thing goes for the effect of Time, but the standard growth model *does* have a random component for the time trends at Level 2.

- **YOU DON'T NEED ANY INDEPENDENT VARIABLES AT LEVEL 2 AT ALL. AS LONG AS YOU INCLUDE RANDOM EFFECTS, YOU WOULD STILL HAVE A "MIXED MODEL"**

(12) Level 1: $Y_{it} = \beta_{0i} + \beta_{1i} X_{1it} + \varepsilon_{it}$

(13) Level 2: $\beta_{0i} = \beta_{00} + \zeta_{0i}$
$\beta_{1i} = \beta_{10} + \zeta_{1i}$

- There are no variables predicting the Level 1 parameters, but there are random effects for both the intercept and the slope associated with $X_1$.
  - Every unit has the fixed overall or average "intercept" ($\beta_{00}$) plus some random error $\zeta_{0i}$.
  - Every unit has the fixed overall or average "slope" for X ($\beta_{10}$) plus some random error $\zeta_{1i}$.
- MIXED VERSION OF THIS MODEL:

(14) $Y_{it} = \beta_{00} + \beta_{10} X_{1it} + (\zeta_{0i} + \zeta_{1i} X_{1it} + \varepsilon_{it})$

with the fixed portion representing the average intercept for the sample and the average slope, and the random portion representing random deviations from the average value for a given unit.

- And if we did not have a random effect in this model for the slope at Level 2 (i.e. no $\zeta_{1i}$ in (13b)), we would arrive at the following mixed model:

$$(15) \qquad Y_{it} = \beta_{00} + \beta_{10} X_{1it} + \zeta_{0i} + \varepsilon_{it}$$

- **DOES THIS LOOK FAMILIAR? IT IS THE BASIC RANDOM INTERCEPT PANEL MODEL!!!**
- How about this?

$$\text{Level 1: } Y_{it} = \beta_{0i} + \beta_{1i} X_{1it} + \varepsilon_{it}$$

$$(16) \qquad \text{Level 2: } \beta_{0i} = \beta_{00} + \beta_{01} \bar{X}_{1i} + \zeta_{0i}$$

$$\beta_{1i} = \beta_{10}$$

$$\text{Mixed: } \quad Y_{it} = \beta_{00} + \beta_{10} X_{1it} + \beta_{01} \bar{X}_{1i} + \zeta_{0i} + \varepsilon_{it}$$

- It is the random effects **hybrid model** with cluster-level (Level 2) means of the time-varying independent variables included as additional predictors of the Level 2 intercept (See Bell&Jones!)

- We can even eliminate the independent variable X altogether, and we arrive at a mixed model that is:

$$(17) \qquad Y_{it} = \beta_{00} + \zeta_{0i} + \varepsilon_{it}$$

- This model says that Y at a given point in time is equal to an overall intercept, or a "GRAND MEAN" ($\beta_{00}$), plus a random unit effect $\zeta_{0i}$ , plus an idiosyncratic unit-time error term $\varepsilon_{it}$. This is the first RE model that Rabe-Hesketh and Skrondal discuss in the book, *Multilevel and Longitudinal Modeling Using Stata*.

- **SO: ALL PANEL MODELS CAN BE EXPRESSED IN THE MULTILEVEL ("MIXED") FRAMEWORK!!!**

# SUMMARY

- Modeling in the multilevel tradition consists of specifying random intercepts to account for the clustered observations and unit-level heterogeneity in exactly the same way we did in units 1 and 2. In addition, we can also include:
  - The effects of time to capture intra-unit growth
  - Random effects for the effects of time across units (Hierarchical growth models)
  - Random effects for the slopes of other time-varying independent variables (RCM models)
  - Cluster means for time-varying independent variables as predictors of the Level 2 intercept (as in the "RE-Hybrid Model") or slope