Introduction to Multilevel Analysis

Tom Snijders

University of Oxford

University of Groningen

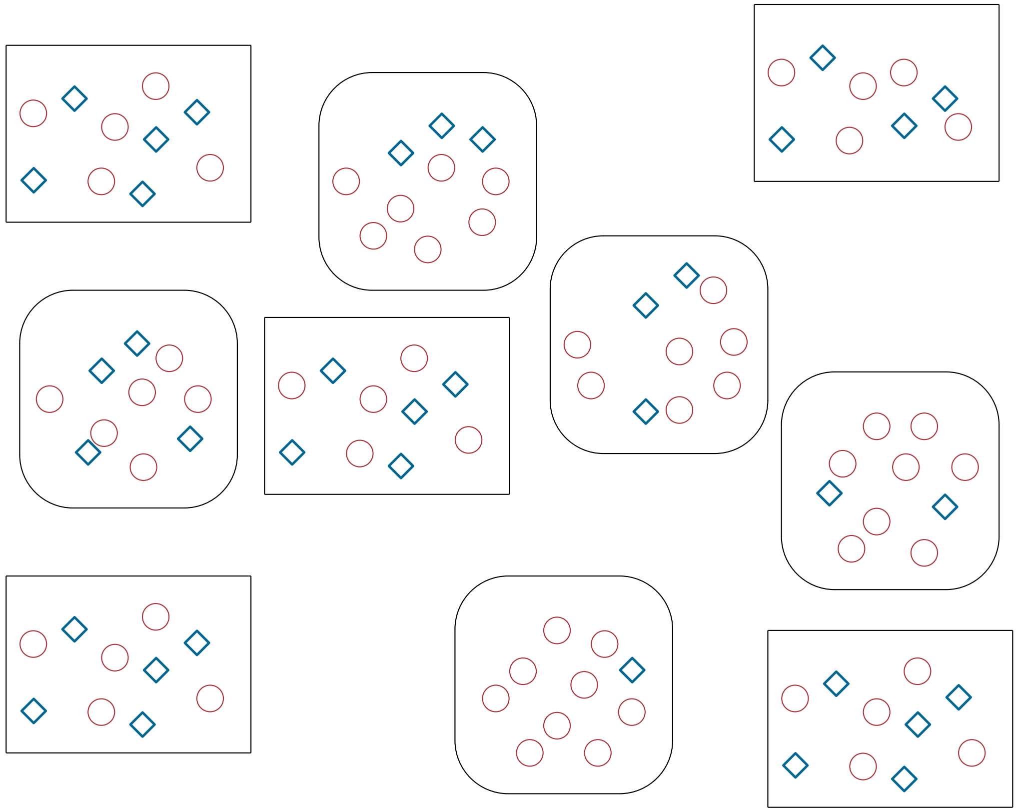
December 2009

Multilevel Analysis based on the

Hierarchical Linear Model (HLM)

is a kind of regression analysis / ANOVA for
situations with several, nested sources of unexplained variation.

Suitable for nested data sets where the dependent variable
is at the lowest (= most detailed) level.



Nested or clustered data: e.g., individuals in groups.

Examples of nesting structures:

∗ *persons in groups*
- ⇒ pupils in schools
- ⇒ employees in organizations
- ⇒ voters in municipalities

∗ *longitudinal or multivariate data*
- ⇒ measurements in individuals

∗ *meta-analysis*
- ⇒ subjects in studies (in research groups)

Terminology: *units*;
e.g., pupils are level-1 units, school level-2 units.

∗ *3 or more levels*

⇒ employees in departments in firms (in countries)

⇒ pupils in classrooms in schools

⇒ citizens in regions in countries

⇒ longitudinal data on employees in organizations

∗ *extensions*

⇒ relations between persons in groups (networks)

⇒ heteroscedastic models

⇒ discrete outcome variables .

Examples:

1. Multilevel analysis of the Demands-Control model.

2. Creation of cooperative norms in Internet Discussion Groups.

Continue theory.

Example based on

Nico W. Van Yperen & Tom A.B. Snijders,

"*Multilevel analysis of the Demands-Control model*"

J. of Occupational Health Psychology 5 (2000), 182-190.

260 employees in 31 groups (in a bank),
4-15 members per group.

Two types of population:

* population of groups

* population of employees.

Question:

How do Job Demands and Job Control
affect Psychological Health Symptoms,
and do these effects play on the group or individual level?

Continue theory. Continue example.

Example based on

Uwe Matzat,

"*Social Networks and Cooperation in Electronic Communities*"

Ph.D. thesis, ICS, University of Groningen, 2001.

and

"*Academic communication and Internet Discussion Groups: transfer of information or creation of social contacts?*"

Social Networks, 26 (2004), 221 – 255.

1059 researchers in
47 scientific Internet Discussion Groups,
4-15 members per group.

Two types of population:

* population of internet groups
* population of researchers.

Question:

On what does it depend
whether group norms for cooperative behavior are formed?

Continue theory. Continue example.

How to analyze such a question & data set?

¿ Forget about the groups, regression at individual level, disaggregate group variables to individual level

¿ Aggregate & analyse at group level

¿ ANCOVA with groups as a factor
= ‘fixed effects’ analysis

¿ Two-step procedure:
regression analysis within each group separately,
then use the estimated coefficients
as dependent variables in a between-group analysis.

! Two-level analysis

* Hierarchical Linear Model with
  random differences between individuals
  and
  random differences between groups.

* Random Coefficient Model
  Assumption: Normal distributions.

Basic points of the multilevel paradigm:

1. "*Differentiate levels of variables*":

   Variables may distinguish between individuals
   but also between groups.
   Most level-1 variables have an individual-level
   as well as a group-level aspect.

2. "*Groups differ*":

   The effect of X (individual-level explanatory variables)

   on Y (outcome variable)

   can differ from group to group.

   Group differences are interesting.

The distinction between individual-level and group-level aspects allows to solve problems related to

ecological and atomistic fallacies:

incorrect interpretation of macro-level relations as if they are micro-level relations and vice versa.

E.g. 1:

within-group differences in expertise can have very different effects *and meanings* from between-group differences in average expertise.

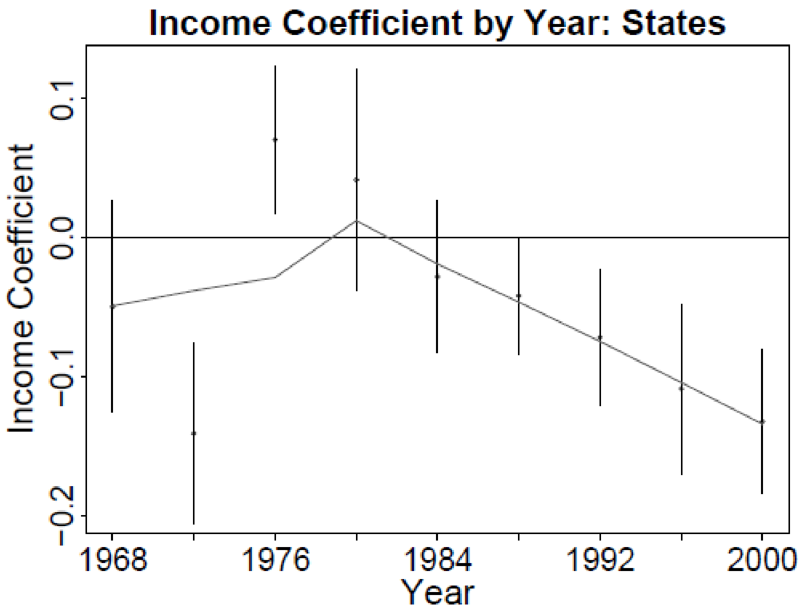
Example: Voting in America

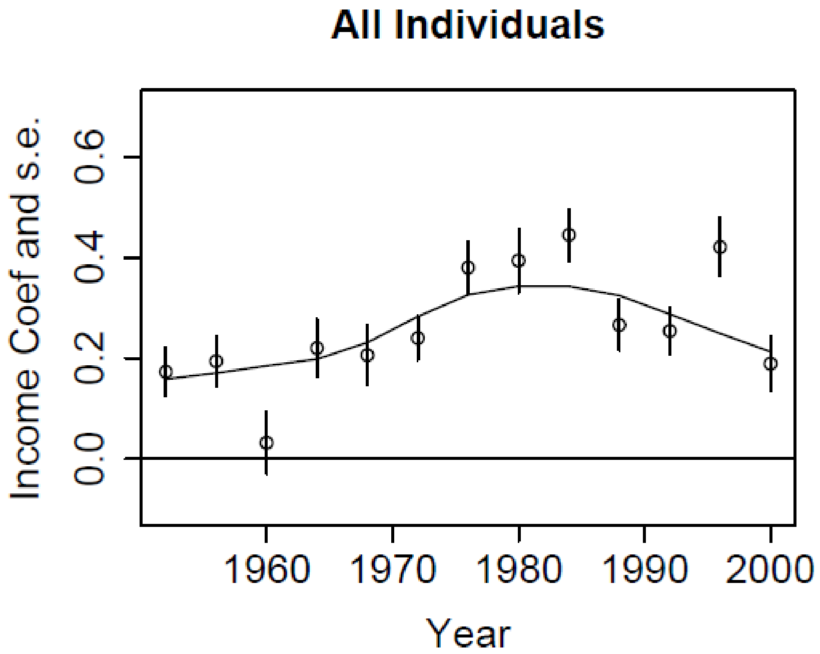
Andrew Gelman and colleagues did an interesting study on effects of income on voting in US presidential elections.

A. Gelman, B. Shor, J. Bafumi, D. Park. Rich state, poor state, red state, blue state: whats the matter with Connecticut? *Quart. J. Polit. Sci.* 2007: 345–367.

Observations :

Richer states have more Democrat votes;
richer individuals vote more Republican.





Coefficients of voting Republican on income;

top:
linear regression,
state averages;

bottom:
logistic regression,
individuals.

Source: Gelman et al., 2007.

Elaboration of the two points

"*1. Differentiate levels of variables*" :

Most variables exhibit within-group variability
and also between-group variability.
These can and should be distinguished.

Total variability
= *within-group variability* + *between-group variability*

If the groups have substantive meaning,
then these two components of variability
may well have substantively different meanings.

Methodological complication arises because of
spill-over of variability to higher levels:

Observed between-group variability
= *true variability* + *random variability*

where 'random variability' stands for the consequences
of the random (unmodeled) differences
between the lower-level units.

Even if there are no systematic between-group differences,
the within-group variability
plus random composition differences between groups
tends to lead to *observed* differences between groups.

Method: decompose variables:

Split variable Y with values Y_{ij}
(where i = individual, level 1; j = group, level 2)
into the *within-group deviation*

$$\tilde{Y}_{ij} = (Y_{ij} - \overline{Y}_{.j})$$

and the *group mean*

$$\overline{Y}_{.j} = \frac{1}{n_j} \sum_i Y_{ij} \ .$$

Note that

$$Y_{ij} = \tilde{Y}_{ij} + \overline{Y}_{.j} \ .$$

If the Y_{ij} variables are purely random with common variance σ^2, i.e., no systematic group differences between groups, the expected observed within- and between-group variances are

$$\text{within} \qquad \text{var}\left(\tilde{Y}_{ij}\right) = \frac{n_j - 1}{n_j}\,\sigma^2$$

$$\text{between} \qquad \text{var}\left(\overline{Y}_{.j}\right) = \frac{1}{n_j}\,\sigma^2 > 0\ .$$

The positive expected observed between-group variance is the spillover from level 1 to level 2.

If, in addition, there are systematic differences between the groups, then var$\left(\overline{Y}_{.j}\right)$ will be larger.

It is helpful to define a *model*
which makes explicit the separate contributions of the two levels:

$$Y_{ij} = \gamma_{00} + U_{0j} + R_{ij}$$

where U_{0j} is the *group effect* and R_{ij} is the *individual effect*.

The random variables U_{0j} and R_{ij} are assumed to be independent,
so that

$$\text{var}\big(Y_{ij}\big) = \text{var}\big(U_{0j}\big) + \text{var}\big(R_{ij}\big) \ .$$

This is a model for the *population*
which gives a convenient split between
the population within-group variance $\sigma_R^2 = \text{var}\,\big(R_{ij}\big)$
and the population between-group variance $\sigma_0^2 = \text{var}\,\big(U_{0j}\big)$.

The population between-group variance is 0
if there are no systematic between-group differences.

An interesting parameter is the *intraclass correlation coefficient*

= *true between-group variance divided by total variance*

= *fraction of population variance at group level*,

$$\rho_{\mathrm{I}} = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_R^2} \; .$$

This is also the correlation between the Y values
of two randomly chosen individuals
in *the same* randomly chosen group.

Continued examples:

1. Multilevel analysis of the Demands-Control model.

2. Creation of cooperative norms in Internet Discussion Groups.

Continue theory.

Continuation of the example: Demands-Control model.

Employees in work-groups in a bank.

What are the effects of Job Demands and Job Control on Psychological Health?

Variables:

* Psychological Health Symptoms (GHQ)
* Job Demands
* Job Control
* Self-Efficacy.

Percent of Variance

Variable	Level 2 *Group*	Level 1 *Employee*	Unreliability
GHQ	7	93	11
Job Demands	25	75	10
Job Control	37	63	6
Self-Efficacy	6	94	17

Unreliability is part of the level-1 variance,
and therefore can be subtracted from σ_R^2
to correct the intraclass correlation for unreliability.

Intraclass correlations, corrected for unreliability:

.08, .28, .39, .07

Like the individual variables,
also the correlations can be split up
into within-group and between-group correlations.

Again, it is important to distinguish
between *observed between-group correlations*
and *estimated population between-group correlations* (spillover).

Observed *between*-group correlations are influenced
not only by population *between*-group correlations
but also by population *within*-group correlations.

E.g.: correlations between Job Control and GQH.

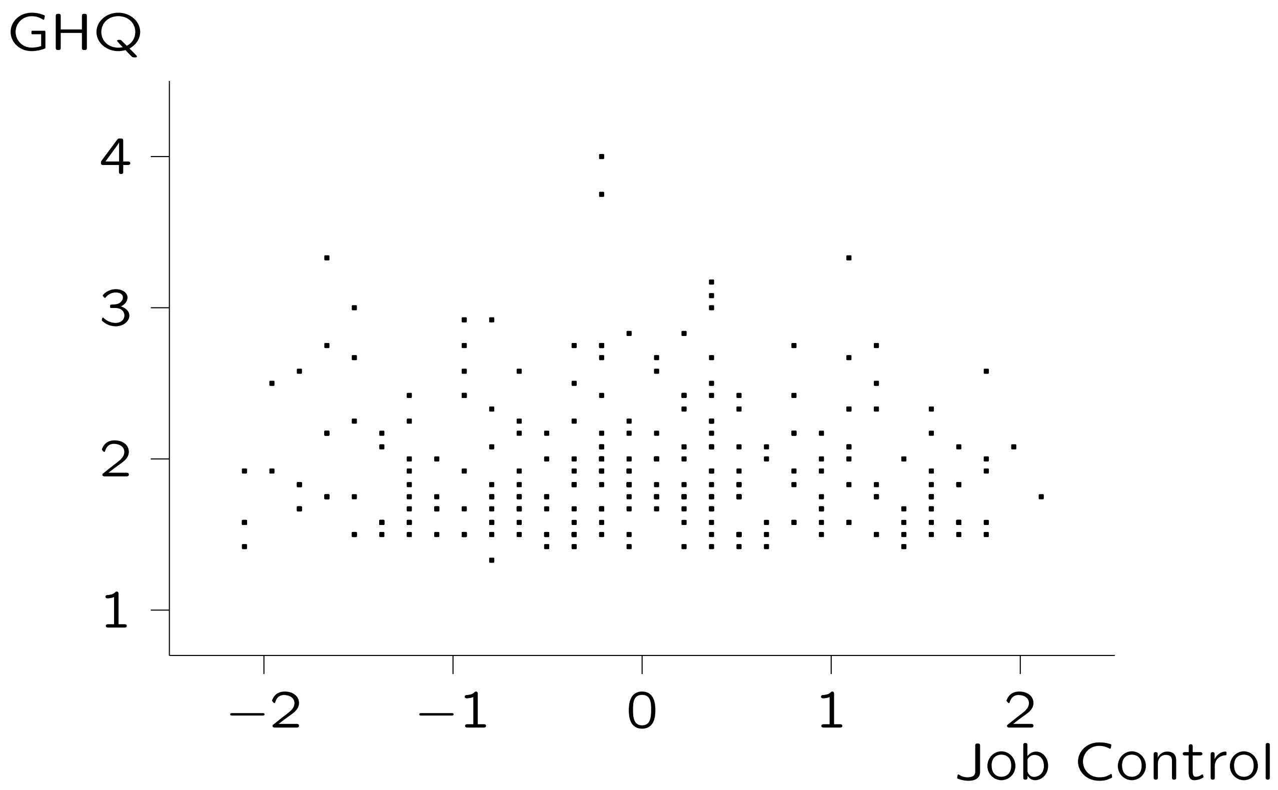
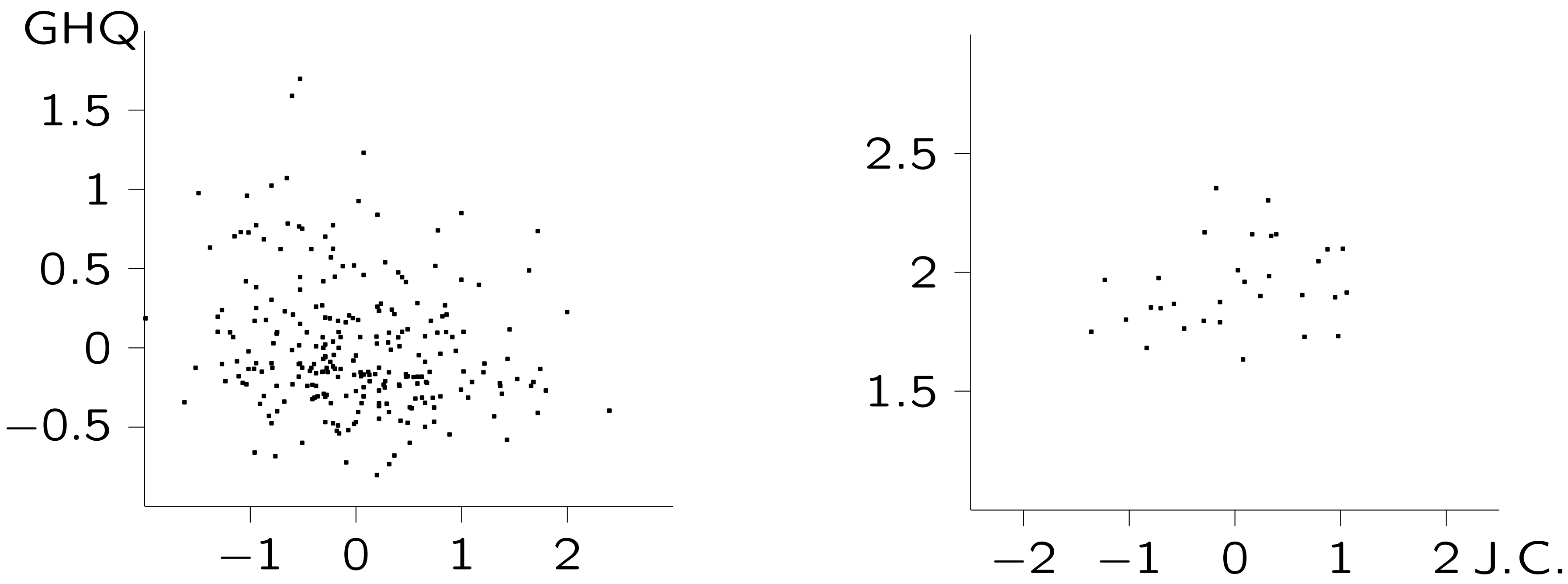


Figure 1. Job Control and GHQ.

Correlation −0.05.



Figures 2, 3. Job Control and GHQ:
within-group deviations (left) and group means (right).

Correlations −0.19 within groups, +0.31 between group means.

The model

$$
\begin{aligned}
X_{ij} &= \gamma_{00}^{X} + U_{0j}^{X} + R_{ij}^{X} \\
Y_{ij} &= \gamma_{00}^{Y} + U_{0j}^{Y} + R_{ij}^{Y}
\end{aligned}
$$

gives us the population correlations:

within-group correlations $\rho\left(R_{ij}^{X}, R_{ij}^{Y}\right)$

and

between-group correlations $\rho\left(U_{0j}^{X}, U_{0j}^{Y}\right)$.

The observed correlations are:

overall	−0.05
within groups	−0.19
between group means	0.31 .

Note that the negative within-group correlation by itself contributes negatively to the correlation between group means. Therefore there must be a large positive estimated population correlation between groups.

Calculations show

$$\hat{\rho}\left(R_{ij}^{X}, R_{ij}^{Y}\right) = -0.19$$

$$\hat{\rho}\left(U_{0j}^{X}, U_{0j}^{Y}\right) = 0.67.$$

Continue theory. Continue example.

Continuation of the example: Cooperative norms in Internet Discussion Groups.

1059 researchers in 47 scientific Internet Discussion Groups.

Question:

On what does it depend
whether group norms for cooperative behavior arise?

Variables:

* Perceived group norm on provision of answers.
* Embeddedness: perception of the discussion group as an integrated community.
* Personal contacts with list members outside the mailing list.
* Perceived own knowledge on list topic, compared to other members.
* Number of list questions in 2 months (control variable) (range 0–199).

Based on Coleman's theory of norms, it was investigated whether embeddedness leads to group norms for providing answers to questions.

First estimate the between-and within-group variances.

Variance components

Variable	Level 2 *Group*	Level 1 *Researcher*	Intraclass correlation
Norms	2.25	5.98	.27
Embeddedness	7.34	49.87	.13
Personal contacts	6.53	35.80	.15
Own knowledge	0.09	1.89	.05

Like the individual variables,
also the correlations can be split up
into within-group and between-group correlations.

Again, it is important to distinguish
between *observed between-group correlations*
and *estimated population between-group correlations*.

Observed *between*-group correlations are influenced
not only by population *between*-group correlations
but also by population *within*-group correlations.

E.g.: correlations between Embeddedness and Norms.

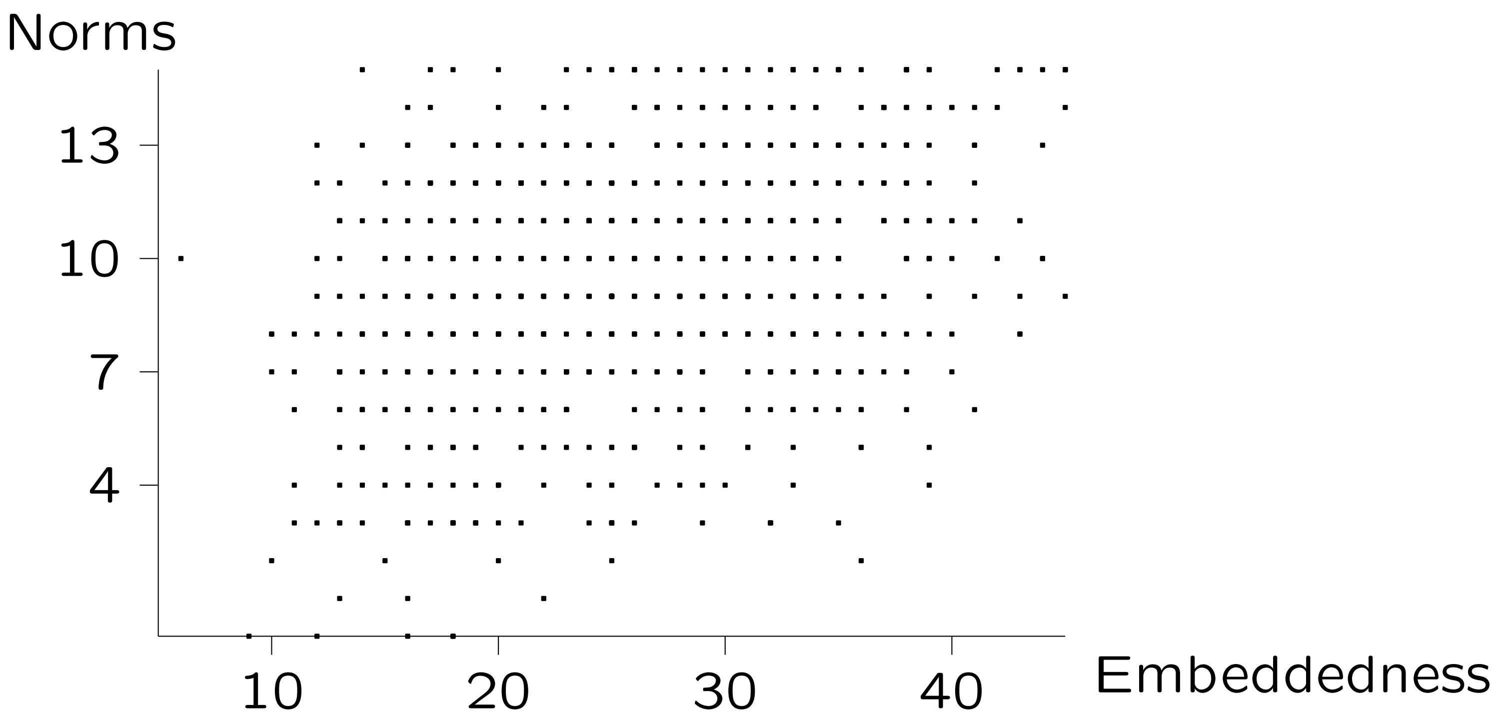
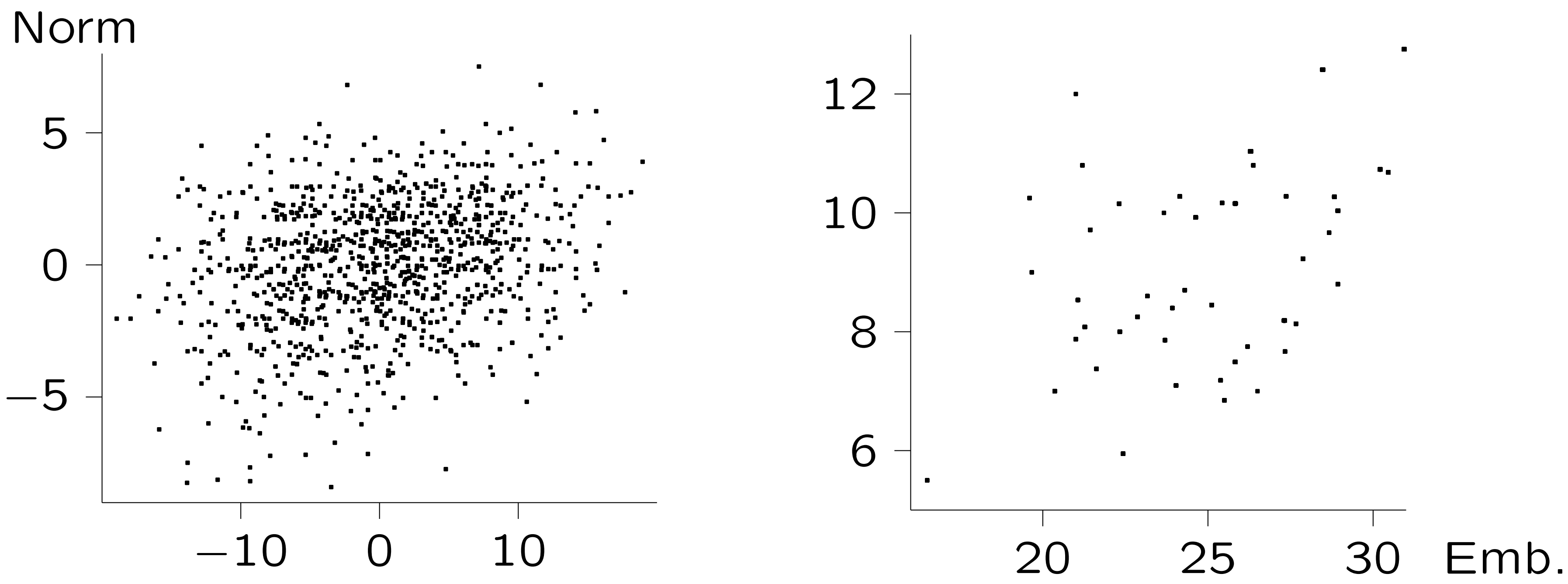


Figure 1. Embeddedness and norms.

Correlation 0.33.



Figures 2, 3. Embeddedness and norms:
within-group deviations (left) and group means (right).

Correlations 0.26 within groups, 0.61 between group means.

The model

$$X_{ij} = \gamma_{00}^{X} + U_{0j}^{X} + R_{ij}^{X}$$
$$Y_{ij} = \gamma_{00}^{Y} + U_{0j}^{Y} + R_{ij}^{Y}$$

gives us the population correlations:

within-group correlations $\rho\left(R_{ij}^{X}, R_{ij}^{Y}\right)$

and

between-group correlations $\rho\left(U_{0j}^{X}, U_{0j}^{Y}\right)$.

The observed correlations are:

overall	0.33
within groups	0.26
between group means	0.61 .

Calculations show

$$\hat{\rho}\left(R_{ij}^{X}, R_{ij}^{Y}\right) = 0.26$$

$$\hat{\rho}\left(U_{0j}^{X}, U_{0j}^{Y}\right) = 0.62.$$

In this case, there is very close correspondence between observed correlations and estimated population correlations.
This is not always so.

Continue theory. Continue example.

The total correlation between X and Y
is a combination of the
within-group and between-group correlations.

The dependence of total correlation
on population within- and between-group correlations
involves the intraclass correlations:

$$\rho(X_{ij}, Y_{ij}) = \underbrace{\sqrt{\rho_{\mathrm{I}}^X\, \rho_{\mathrm{I}}^Y}\, \rho\left(R_{ij}^X, R_{ij}^Y\right)}_{\Uparrow\ \text{within-groups correlation}} + \underbrace{\sqrt{(1-\rho_{\mathrm{I}}^X)(1-\rho_{\mathrm{I}}^Y)}\, \rho\left(U_{0j}^X, U_{0j}^Y\right)}_{\Uparrow\ \text{between-groups correlation.}}$$

Go back

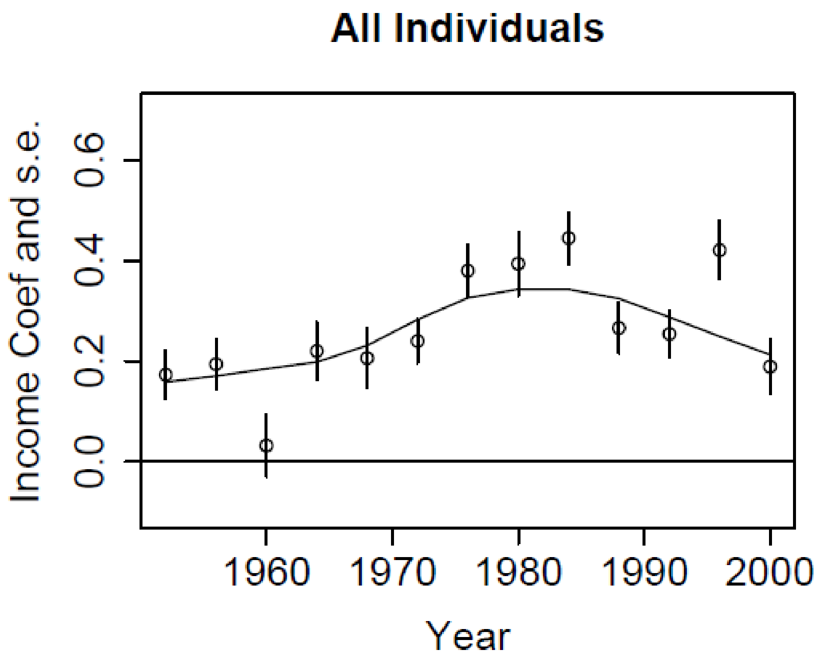
The danger of the ecological and atomistic fallacies:

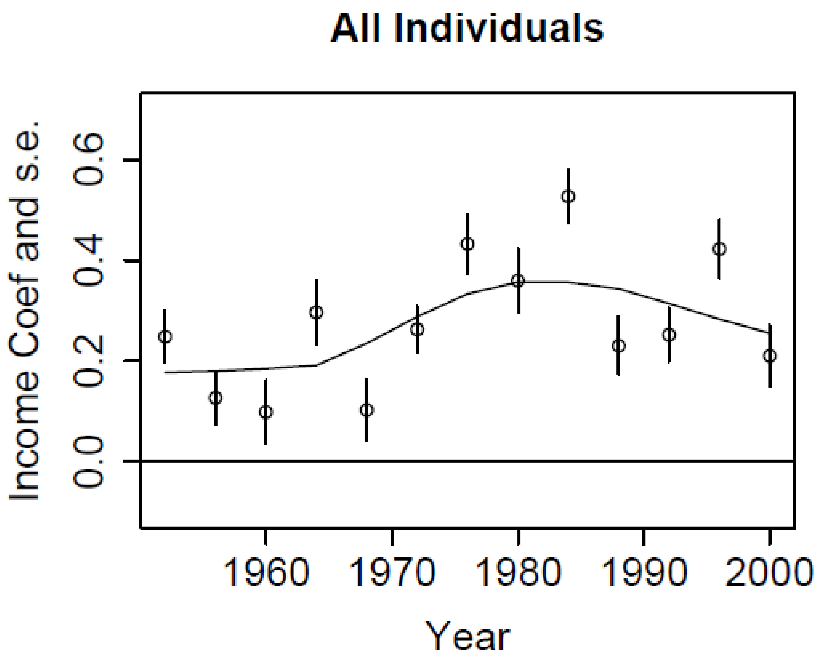
what happens between the groups can be
completely different from what happens within the groups.

To study this more in depth,
bivariate correlations are not enough,
a multivariate approach is better:

the Hierarchical Linear Model (*HLM*),

a regression-type model indicating in a multilevel situation
how a dependent variable Y depends on explanatory variable X.





Back to US voting.

Coefficient of voting Republican on income, for individuals, logistic regression;

top:
single-level analysis;

bottom:
multilevel analysis.

Source: Gelman et al., 2007.

In the *HLM*, the model used above

$$Y_{ij} = \gamma_{00} + U_{0j} + R_{ij}$$

is known as the *empty model*.
It gives a decomposition of the variability,
without a role for explanatory variables.

This model often is useful as a starting point
of the statistical analysis.

The simplest way to enter an explanatory variable
is the usual linear regression form:

$$Y_{ij} = \gamma_{00} + \gamma_{10}\, x_{ij} + U_{0j} + R_{ij}.$$

This models differs from a standard (“*OLS*”) regression model
by its more complicated *random part*,
incorporating *residuals at level 1 <u>and</u> level 2*, R_{ij} and U_{0j} .

It is called a *random intercept model*,
because the intercept $\gamma_{00} + U_{0j}$ of the regression line
differs randomly from group to group.

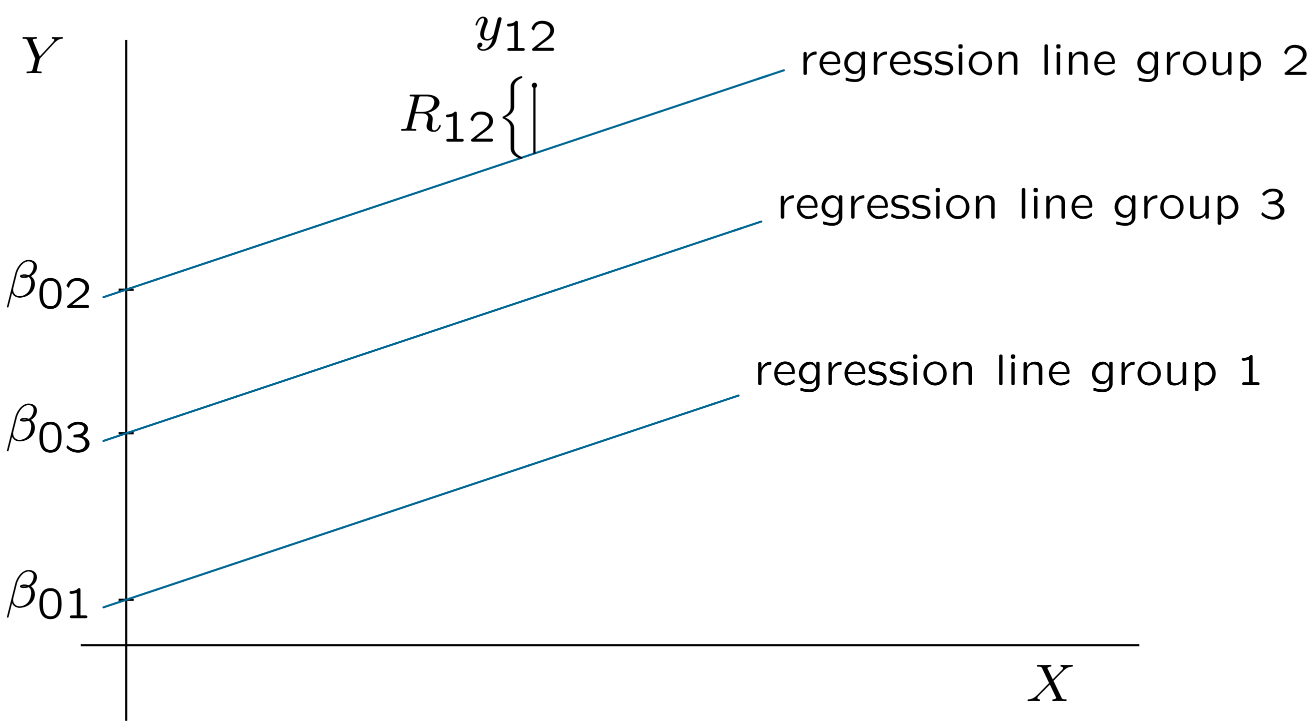


Figure 4. The random intercept model:
different parallel regression lines.
The point y_{12} is indicated with its residual R_{12} .

However, the model specification with one X variable
does not take into account that the
within-group effect of X on Y
may differ from the *between-group effect*.

Definition *within-group effect*:
expected difference on Y
between two individuals in the same group
who differ 1 unit on X.

Definition *between-group effect*:
expected differences on group mean $\overline{Y}$
between two groups
which differ 1 unit on group mean $\overline{X}$.

For a level-1 explanatory variable x_{ij},
use two variables in the fixed part:

$$x_{ij} \quad \text{and} \quad \overline{x}_{.j}$$

or

$$\tilde{x}_{ij} = (x_{ij} - \overline{x}_{.j}) \quad \text{and} \quad \overline{x}_{.j}.$$

The within-group regression coefficient is coefficient of $\tilde{x}_{ij}$,
between-group regression coefficient is coefficient of $\overline{x}_{.j}$

The HLM equation (still a random intercept model) now reads

$$\begin{aligned} Y_{ij} = \; & \gamma_{00} + \gamma_{10}\,\tilde{x}_{ij} + \gamma_{01}\,\overline{x}_{.j} && \textit{fixed part} \\ & + U_{0j} + R_{ij} && \textit{random part} \;. \end{aligned}$$

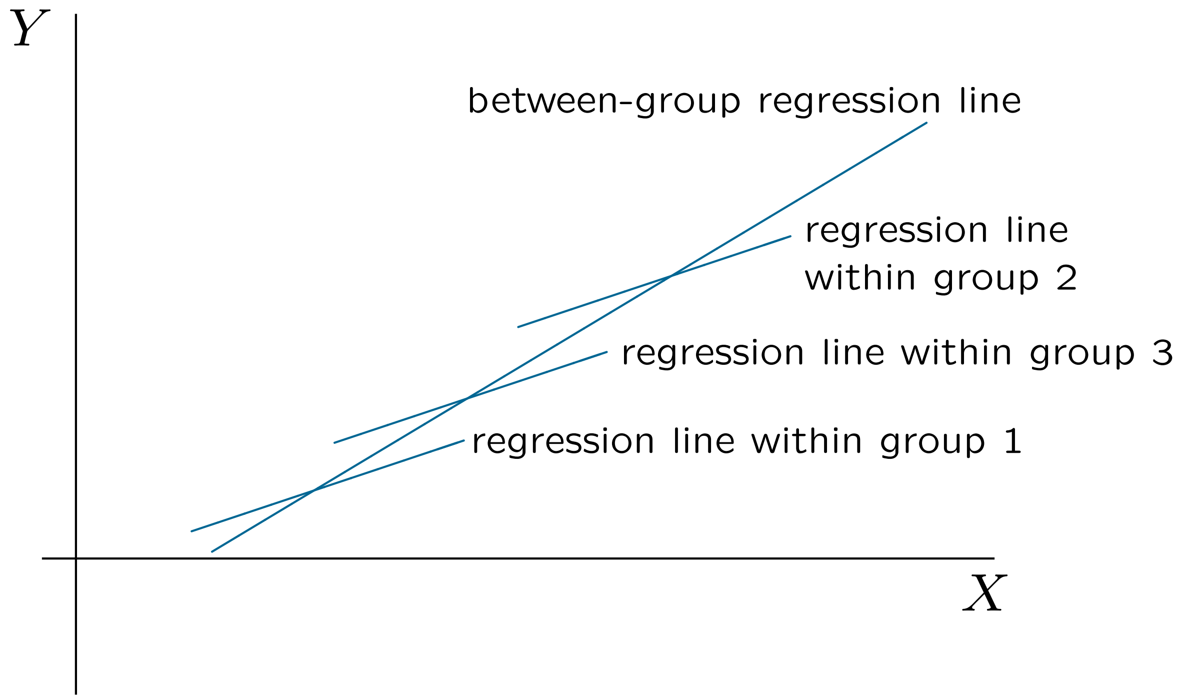
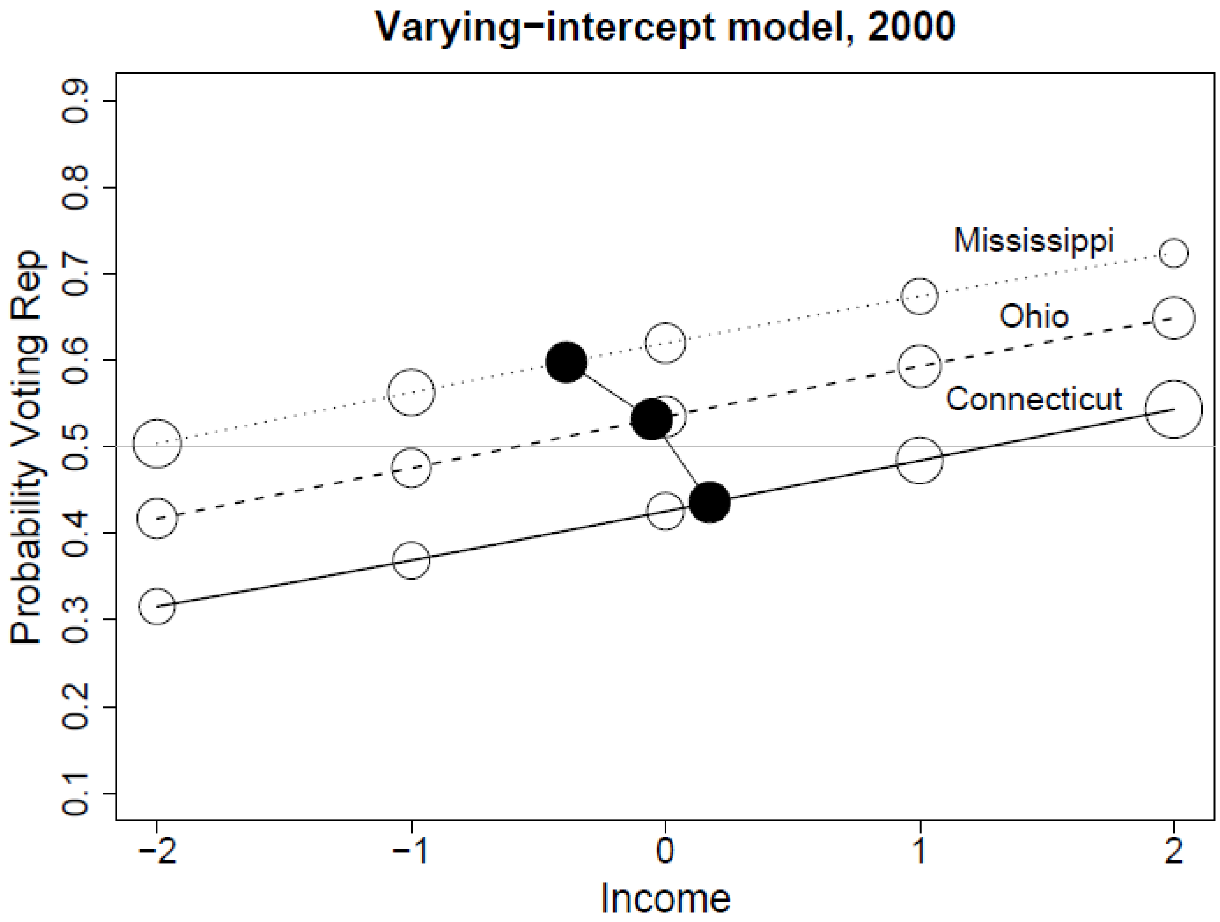


Figure 5. Different between-group and within-group regression lines.



Probability of supporting Bush as function of income
in three states.
within- and between-state regressions, random intercept model.

Source: Gelman et al., 2007.

Why do we need the *HLM* for testing the difference between within- and between-group coefficients?

As we are simultaneously testing
effects of individual-level and group-level variables,
we need models with *error terms at either level*.

Continued examples:

1. Multilevel analysis of the Demands-Control model.

2. Creation of cooperative norms in Internet Discussion Groups.

Continue theory.

Continuation of the example: Demands-Control model.

The empty model yields the following results.

	Empty model	
Fixed part		
constant	1.945	(.033)
Random part		
Level 2: σ_0^2	0.012	(.009)
Level 1: σ_R^2	0.182	(.017)
deviance	308.5	

Table 1. Parameter estimates (and standard errors) of empty model for GHQ.

Models for the effect of Job Control on GHQ
in various specifications yield the following results.

	Model 1		Model 2		Model 3	
Fixed part						
constant	1.944	(.036)	1.945	(.032)	1.945	(.032)
J.C. raw x_{ij}	−0.047	(.029)	−0.102	(.035)		
J.C. deviation $\tilde{x}_{.j}$					−0.102	(.035)
J.C. group mean $\overline{x}_{.j}$			0.183	(.059)	0.081	(.048)
Random part						
Level 2: σ_0^2	0.017	(.010)	0.010	(.008)	0.010	(.008)
Level 1: σ_R^2	0.177	(.017)	0.176	(.016)	0.176	(.016)
deviance	306.2		297.3		297.3	

Table 2. Parameter estimates (and standard errors) of various random intercept models for effect of Job Control on GHQ.

This shows that the effect of Job Control on GHQ (without taking into account other variables) is significant only if the split between the level-1 and level-2 variables is made, i.e., if between-group and within-group regressions are allowed to differ.

Models 2 and 3 are equivalent representations:

$$-0.102\, x_{ij} + 0.183\, \overline{x}_{.j} = -0.102\, \tilde{x}_{ij} + 0.081\, \overline{x}_{.j} \ .$$

Model 2 can be used for testing if the within-group and between-group regression coefficients are different; this means that the coefficient of $\overline{x}_{.j}$ is not 0.
This is supported here: $t = 0.183/0.059 = 3.10$, $p < 0.01$.

Continue theory. Continue example.

Continuation of the example:
Cooperative norms in Internet Discussion Groups.

Control variable: number of list questions, measure of activity.

	Empty model		Model 1	
Fixed part				
constant	9.047	(.247)	8.018	(.244)
$\sqrt{\text{list questions}}$			0.323	(.050)
Random part				
Level 2: σ_0^2	2.188	(.574)	0.922	(.285)
Level 1: σ_R^2	5.980	(.265)	5.981	(.265)
deviance	4984.65		4956.47	

Table 3. Parameter estimates (and standard errors)
of empty model and model with control variable for Norms.

	Model 2		Model 3		Model 4	
Fixed part						
constant	5.781	(.337)	4.187	(1.353)	4.187	(1.353)
$\sqrt{}$ list questions	0.289	(.046)	0.266	(0.049)	0.266	(0.049)
Emb. raw x_{ij}	0.093	(.010)	0.091	(0.010)		
Emb. deviation $\tilde{x}_{.j}$					0.091	(0.010)
Emb. group mean $\overline{x}_{.j}$			0.068	(0.056)	0.159	(0.055)
Random part						
Level 2: σ_0^2	0.784	(.247)	0.722	(0.233)	0.722	(0.233)
Level 1: σ_R^2	5.570	(.247)	5.575	(0.247)	5.575	(0.247)
deviance	4878.45		4877.05		4877.05	

Table 4. Parameter estimates (and standard errors) of various random intercept models for effect of Embeddedness on Norms.

Within-group coefficient: 0.091

Between-group coefficient: 0.159

Difference is not significant ($t = 0.068/0.056 = 1.21$)

This shows that the effect of Embeddedness on Norms (controlling for $\surd$ number of list questions) is significant, and the between-group and within-group regressions do not differ significantly.

Models 3 and 4 are equivalent representations:

$$0.091\, x_{ij} + 0.068\, \overline{x}_{.j} = 0.091\, \tilde{x}_{ij} + 0.159\, \overline{x}_{.j} \ .$$

Model 3 can be used for testing if the within-group and between-group regression coefficients are different; this means that the coefficient of $\overline{x}_{.j}$ is not 0, when controlling for x_{ij}.
This is not significant here: $t = 0.068/056 = 1.21$, $p > 0.10$.

Also adding the number of list contacts leads to the following.

	Model 5		Model 6	
Fixed part				
constant	3.448	(1.241)	3.444	(1.242)
$\sqrt{}$ list questions	0.234	(0.044)	0.235	(0.044)
Emb. raw	0.092	(0.011)		
Emb. deviation			0.091	(0.010)
Emb. group mean	0.143	(0.056)	0.236	(0.056)
L.C. raw	−0.009	(0.013)		
L.C. group mean	−0.171	(0.058)	−0.172	(0.058)
Random part				
Level 2: σ_0^2	0.517	(0.183)	0.518	(0.183)
Level 1: σ_R^2	5.577	(0.247)	5.580	(0.247)
deviance	4868.75		4869.32	

Table 5. Estimates (and standard errors) for various models for effect of Embeddedness and Contacts with list members (L.C.) on Norms.

This shows:

1. controlling for Contacts with list members,
   the within-group and between-group regression
   on Embeddedness *are* different;
   the between-group effect is larger
   than the within-group effect;

2. controlling for embeddedness,
   there is a negative effect of the Contacts with list members,
   but not within groups, only for the group mean.

Continue theory. Continue example.

Elaboration of the second point:

2. "Groups differ"

Method: regression coefficients (intercepts, slopes)
are regarded as latent group characteristics,
indicating the group-dependent relation between X and Y.

Recall that i indicates the individual (*level-1 unit*)

j indicates the group (*level-2 unit*).

This gives equations for both levels
(for simplicity, only one X variable is used)

$$Y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + R_{ij} \qquad \textit{individual level}$$

$$\left\{ \begin{array}{lcl} \beta_{0j} & = & \gamma_{00} + U_{0j} \\ \beta_{1j} & = & \gamma_{10} + U_{1j} \end{array} \right\} \qquad \textit{group level.}$$

Substitution yields:

$$\begin{aligned} Y_{ij} = {} & \gamma_{00} + \gamma_{10}\, x_{ij} && \textit{fixed part} \\ & + U_{0j} + U_{1j}\, x_{ij} + R_{ij} && \textit{random part .} \end{aligned}$$

In the model

$$\begin{aligned} Y_{ij} = &\ \gamma_{00} + \gamma_{10}\, x_{ij} && \textit{fixed part} \\ &+ U_{0j} + U_{1j}\, x_{ij} + R_{ij} && \textit{random part}\ . \end{aligned}$$

there are two kinds of parameters:

fixed part	
regression coefficients	γ_{00} γ_{10}
random part	
between-group var.-cov.	σ_0^2 σ_1^2 σ_{01}
residual individual variance	σ_R^2

Simplest case: the *random intercept model*,
defined by $\sigma_0^2 \geq 0$, $\sigma_1^2 = 0$,
with random variation only between
the intercepts $\beta_{0j} = \gamma_{00} + U_{0j}$;

the general case is the *random slopes model*, with $\sigma_1^2 \geq 0$:
random variation also between the slopes $\beta_{1j} = \gamma_{10} + U_{1j}$.

In the Random Intercept Model, the groups (level-2 units) are from a *population of parallel regression lines*.

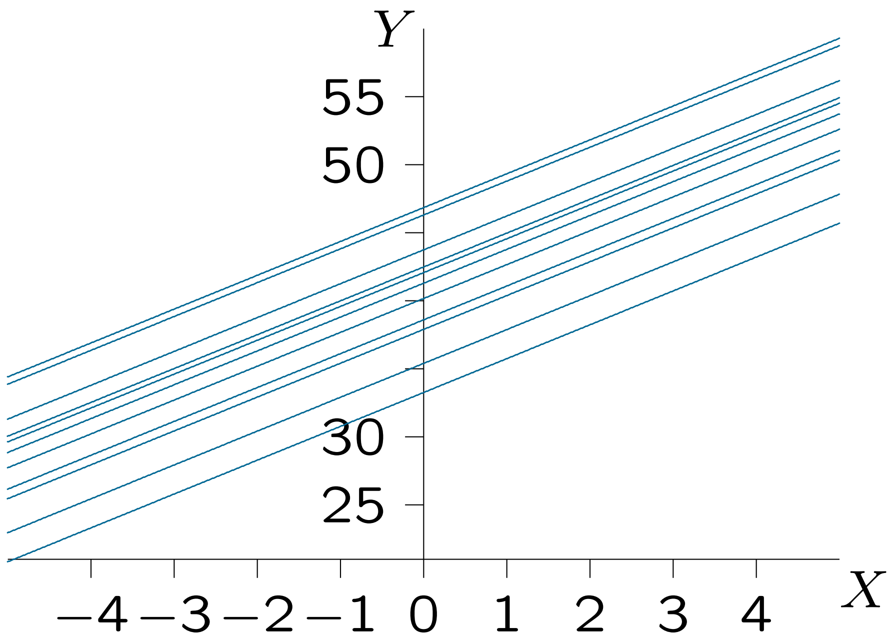


Figure 6. Random regression lines according to the random intercepts model of Table 4.2 in Snijders & Bosker (1999).

In the Random Slopes Model, the groups are from a *population of intersecting regression lines*.

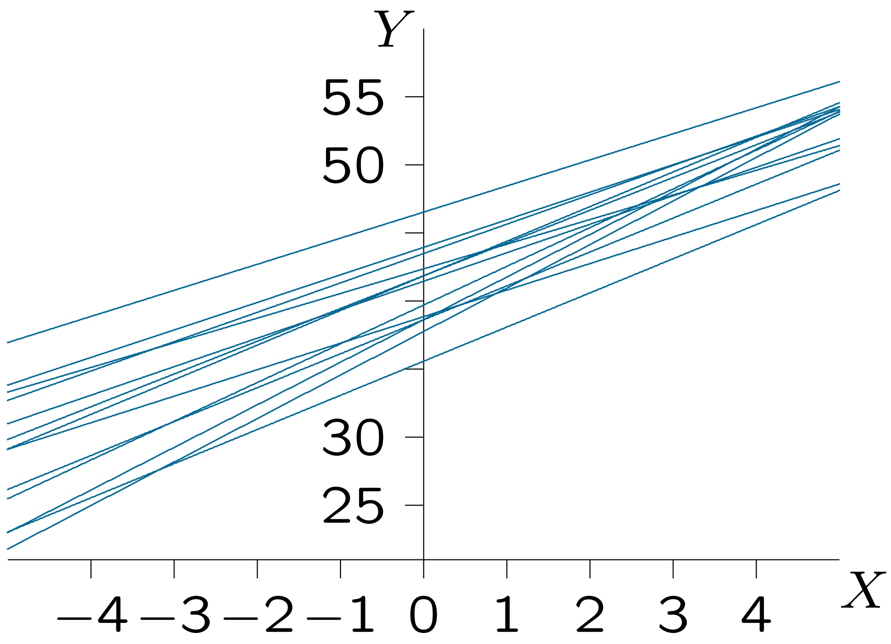
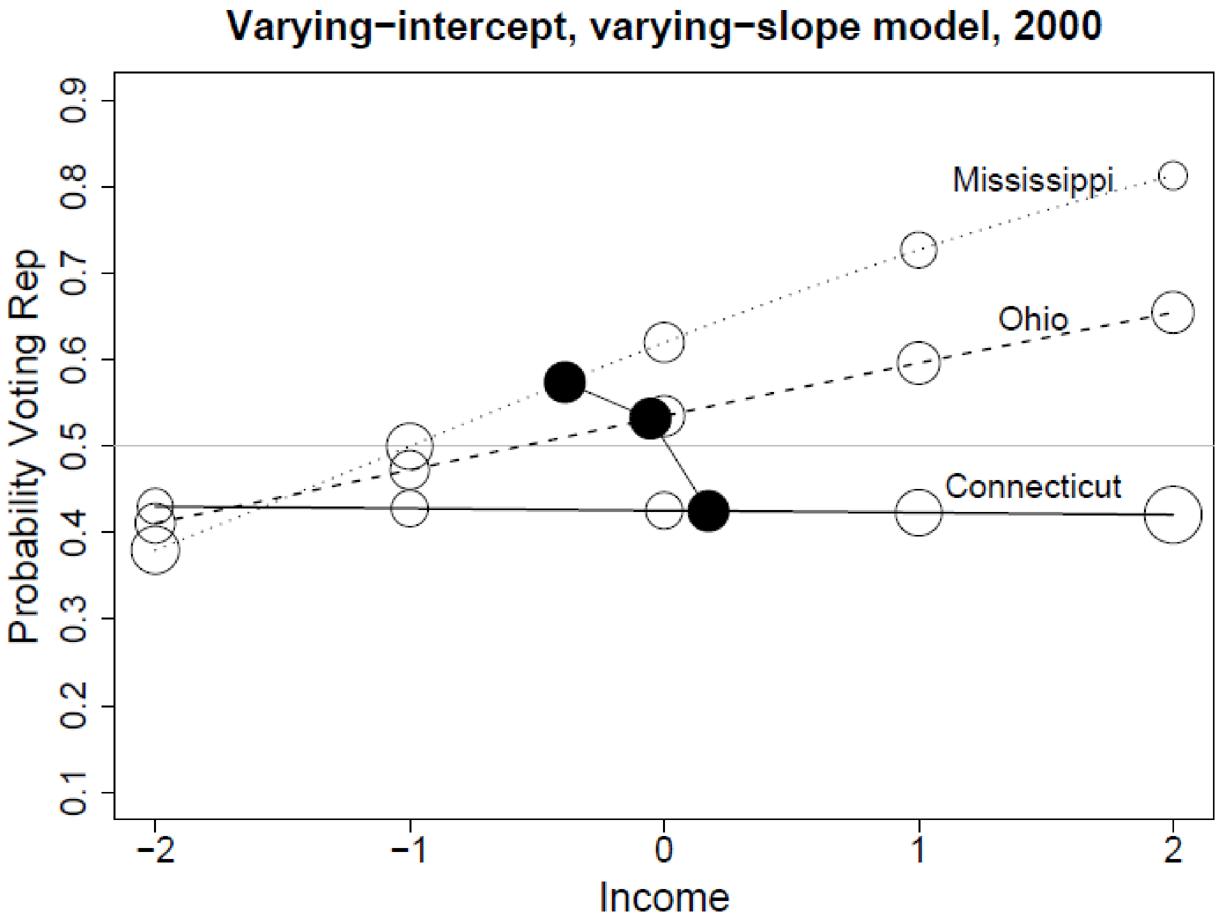
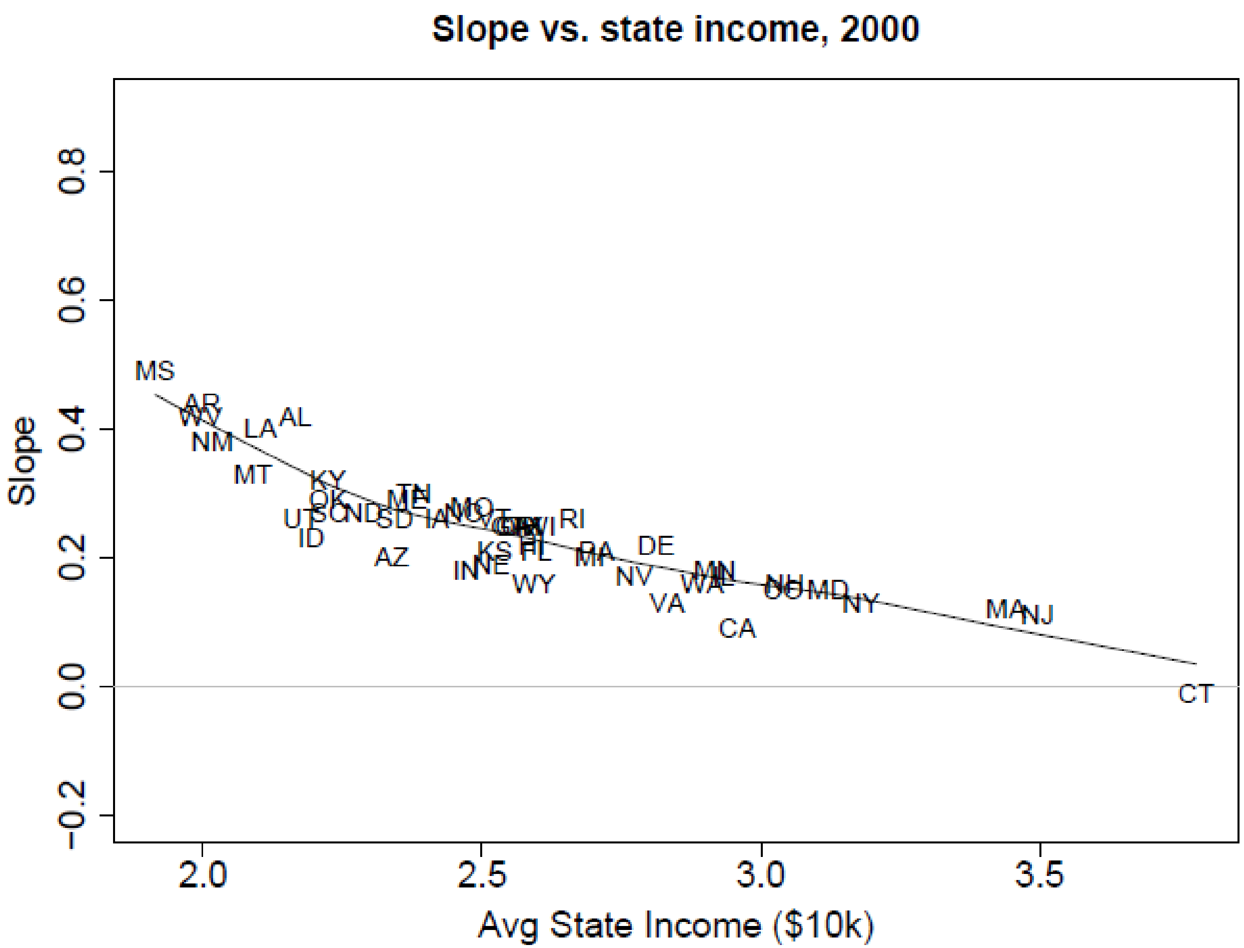


Figure 7. Random regression lines according to the random slopes model of Table 5.2 in Snijders & Bosker (1999).



Probability of supporting Bush as function of income
in three states.
within- and between-state regressions, random slope model.

Source: Gelman et al., 2007.



Coefficient of voting Republican by state
as a function of average state income.
Random slope model.

Together, this means that
for every level-1 explanatory variable,
the following choices must be made:

* random or only fixed effect?
* within- and between-group effects
  the same (only effect of x_{ij}) or different?

The distinction between the group mean $\overline{x}_{.j}$
and the within-group deviation $\tilde{x}_{ij}$
is a substance-matter distinction:

These variables have different meanings!

However, we are not accustomed to
deriving theoretical predictions about such level-specific effects.

This calls for new *multilevel theory*
and, as far as this is not fully developed,
for exploratory approaches
to the specification of level-specific effects.

Example:
O'Campo (2003), *American Journal of Epidemiology*, 157: 9–13.

Continued examples:

1. Multilevel analysis of the Demands-Control model.

2. Creation of cooperative norms in Internet Discussion Groups.
   Start Continue

Continue theory.

Continuation of the example:
Demands-Control model.

The *HLM* equation,
incorporating different within- and between group regressions,
as well as random slopes, is given by

$$\begin{aligned}
\text{GHQ}_{ij} &= \left(\gamma_{00} + U_{0j} + \gamma_{01}\,\overline{\text{JC}}_j\right) \\
&\quad + \left(\gamma_{10} + U_{1j}\right)\widetilde{\text{JC}}_{ij} + R_{ij} \\
&= \gamma_{00} + \gamma_{10}\,\widetilde{\text{JC}}_{ij} + \gamma_{01}\,\overline{\text{JC}}_j \qquad \textit{fixed} \\
&\quad + U_{0j} + U_{1j}\,\widetilde{\text{JC}}_{ij} + R_{ij} \qquad \textit{random}
\end{aligned}$$

	Model 4	
Fixed part		
constant	1.944	(.031)
J.C. deviation $\tilde{x}_{.j}$	−0.107	(.042)
J.C. group mean $\bar{x}_{.j}$	0.068	(.045)
Random part		
Level 2 :		
σ_0^2 intercept variance	0.010	(.008)
σ_1^2 J.C. slope variance	0.015	(.013)
σ_{01} intercept-slope covariance	−0.011	(.007)
Level 1 :		
σ_R^2 residual variance	0.166	(.016)
deviance	291.9	

Table 6. Parameter estimates (and standard errors) of random slope model for effect of Job Control on GHQ.

The random slope is significant
($\chi^2 = 297.3 - 291.9 = 5.4, d.f. \ = 2$,
p-value from χ^2 distribution may be divided by 2: $p < 0.05$).

Interpretation of slope variance:

Groups with low/high slope on $\widetilde{\text{JC}}$:

$$\begin{array}{rcl} \beta_{1j} = & -.107 \pm 2\sqrt{.0015} & : \ \text{values } -0.18 \text{ and } -0.03 \\ & \| & \\ & \gamma_{10} \pm 2\sqrt{\sigma_1^2} & \end{array}$$

Continue theory. Continue example.

Continuation of the example: Cooperative norms in Internet Discussion Groups.

The *HLM* equation for the effects of Embeddedness on Norms, incorporating different within- and between group regressions, as well as random slopes, and for convenience temporarily forgetting the other effects, is given by

$$\begin{aligned}
\text{NORMS}_{ij} &= \Big(\gamma_{00} + U_{0j} + \gamma_{01}\,\overline{\text{EMB}}_j\Big) \\
&\quad + \Big(\gamma_{10} + U_{1j}\Big)\,\widetilde{\text{EMB}}_{ij} + R_{ij} \\
&= \gamma_{00} + \gamma_{10}\,\widetilde{\text{EMB}}_{ij} + \gamma_{01}\,\overline{\text{EMB}}_j \quad \textit{fixed} \\
&\quad + U_{0j} + U_{1j}\,\widetilde{\text{EMB}}_{ij} + R_{ij} \qquad \textit{random}
\end{aligned}$$

	Model 7	
Fixed part		
constant	3.255	(1.236)
$\sqrt{\text{list questions}}$	0.241	(0.044)
Emb. deviation	0.105	(0.014)
Emb. group mean	0.244	(0.055)
L.C. group mean	−0.174	(0.058)
Random part		
Level 2 :		
σ_0^2 intercept variance	0.5202	(0.1830)
σ_1^2 Emb. slope variance	0.0022	(0.0015)
σ_{01} intercept-slope covariance	0.0067	(0.0120)
Level 1 :		
σ_R^2 residual variance	5.4788	(0.2460)
deviance	4864.93	

Table 7. Parameter estimates (and standard errors) of random slope model for effect of Embeddedness on Norms.

The random slope is 'almost significant'
($\chi^2 = 4869.32 - 4864.93 = 4.39, d.f. \ = 2$,
p-value from χ^2 distribution may be divided by 2: $p \approx 0.07$).

Interpretation of slope variance:

Groups with low/high slope on $\widetilde{\text{EMB}}$:

$$\begin{array}{rcl} \beta_{1j} = & \begin{array}{c} .105 \pm 2\sqrt{.0022} \\ \| \\ \gamma_{10} \pm 2\sqrt{\sigma_1^2} \end{array} & : \text{ values } .011 \text{ and } .199. \end{array}$$

Continue theory. Continue example.

Groups differ: part 3

The "*latent group characteristics*"

$\beta_{0j} = \gamma_{00} + U_{0j}$ and $\beta_{1j} = \gamma_{10} + U_{1j}$

can themselves be explained by group variables.

This means that the coefficients in the level-1 model become dependent variables in level-2 regression models: "*Intercepts and slopes as outcomes*".

This is not available in a 'fixed effects' analysis.

Explaining the group-dependent effect of X by the group variable Z leads to a $X \times Z$ interaction: "*cross-level interaction*".

(Level-1 and level-2 interactions can of course also be included.)

The $X \times Z$ cross-level interaction coefficient indicates how large is *the effect of* context variable Z *on the effect of* individual variable X.

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + R_{ij} \qquad \textit{level 1}$$

$$\left. \begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01} Z_j + U_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11} Z_j + U_{1j} \end{aligned} \right\} \qquad \textit{level 2 (latent)}$$

Substitution yields

$$\begin{aligned} Y_{ij} = \gamma_{00} + \gamma_{01} Z_j + \gamma_{10} X_{ij} + \gamma_{11} Z_j X_{ij} & \quad \textit{fixed} \\ + U_{01} + U_{1j} X_{ij} + R_{ij} & \quad \textit{random .} \end{aligned}$$

Continued examples:

1. Multilevel analysis of the Demands-Control model.

2. Creation of cooperative norms in Internet Discussion Groups.

Continue theory.

Continuation of the example: Demands-Control model.

In addition to Job Control,
also Job Demands was included in the model.
This was done in several ways:
group mean, within-group deviation,
and both in interaction with both Job Control variables.

Two of these four interactions are cross-level interactions.

	Model 5	
Fixed part		
constant	1.930	(.026)
J.C. deviation	−0.099	(.042)
J.C. group mean	0.060	(.042)
J.D. deviation	0.083	(.031)
J.D. group mean	0.158	(.045)
J.C. devi × J.D. devi	−0.135	(.036)
J.C. devi × J.D. group	−0.018	(.071)
J.C. group × J.D. devi	0.019	(.048)
J.C. group × J.D. group	0.034	(.080)
Random part		
σ_0^2 intercept variance	0.002	(.005)
σ_1^2 slope variance	0.017	(.012)
σ_{01} intercept-slope covariance	−0.010	(.006)
σ_R^2 residual variance	0.151	(.015)
deviance	258.5	

Table 8.
Parameter estimates (and standard errors) of random slope model for effect of Job Control and Job Demands on GHQ.

Conclusions:

Job Demands has significant positive effects
at the individual *and* the group level;

there is a significant interaction
between Job Control and Job Demands;
this is an interaction at the *individual* level;
neither of the cross-level interactions is significant;

the random slope remains significant
(it is not 'explained by a level-2 variable').

Other example: meta-analysis

Continue theory.

Continuation of the example: Cooperative norms in Internet Discussion Groups.

It was attempted to 'explain' the random slope on Embeddedness by a cross-level interaction of Embeddedness with $\sqrt{}$ number of list questions.

	Model 7	
Fixed part		
constant	3.585	(1.255)
$\sqrt{}$list questions	0.234	(0.044)
Emb. deviation	0.100	(0.013)
Emb. group mean	0.234	(0.056)
L.C. group mean	−0.178	(0.058)
Emb. dev. $\times$ $\sqrt{}$ list quest.	−0.0062	(0.0032)
Random part		
Level 2 :		
σ_0^2 intercept variance	0.530	(0.185)
σ_1^2 Emb. slope variance	0.0016	(0.0013)
σ_{01} intercept-slope covariance	0.0071	(0.0113)
Level 1 :		
σ_R^2 residual variance	5.476	(0.246)
deviance	4861.33	

Table 9. Parameter estimates (and standard errors) of random slope model for effect of Embeddedness on Norms.

Result:
with a higher number of list questions,
the effect of Embeddedness on Norms
becomes slightly weaker.

The test for a random slope of embeddedness now yields
$\chi^2 = 2.8, d.f. = 2$, p (divided by 2) > 0.10.

Other example: meta-analysis

Other example: productivity brainstorming groups

Continue theory.

Example 3. Meta-analysis of class size effects.

Example based on

H. Goldstein, M. Yang, R. Omar, R. Turner, & S. Thompson
"*Meta-analysis using multilevel models with an application to the study of class size effects*" *Applied Statistics* **49** *(2000), 399-412.*

9 studies 1986–1991, 36 classes, 15-30 pupils per class.

Three types of population:

* population of studies
* population of classes
* population of pupils.

Meta-analysis (the statistically formalized synthesis of literature)
traditionally proceeds by first testing
the homogeneity of the collected studies,
and, if homogeneity is supported, analyzing them jointly.

However, we should expect diversity, not homogeneity
of empirical research on a common research question:
different populations, instruments, etc.

This diversity can be taken seriously
by adopting a multilevel approach:
e.g., subjects within studies within research groups.

This permits the meta-analyst to split
the observed variability between studies
into *unreliable* (‘*error*’) variability and *true* variability.

However, data about individual subjects usually is not available.
This can be replaced by information about the
standard errors (*unreliability*) of the results.

This leads to models such as

$$Y_j = \sum_h \gamma_h x_{hj} + U_{0j} + U_{1j}\, x_{1j} + R_j \ ,$$

where Y_j is a standardized effect size estimator from study j,
R_j is the error in study j
with a known variance (*standard error*^{2}),
the x_{hj} are study variables and control variables,
and U_{0j}, U_{1j} are study-specific effects ('*true*' differences).

In the meta-analysis by Goldstein et al., first the 8 studies are used for which only aggregate information is available.
The dependent variable Y_j is an average, adjusted for pretest, of an achievement score standardized to S.D. = 1.0.

	Model 1		Model 2	
Fixed part				
constant	0.163	(0.028, 0.308)	0.224	(0.053, 0.393)
C.S., linear	−0.020	(−0.036, −0.004)	−0.048	(−0.072, −0.025)
C.S., quadratic			0.002	(0.001, 0.003)
Random part				
intercept var.	0.060	(0.028, 0.308)	0.067	(0, 0.135)
slope var. C.S.	0.0006	(0.028, 0.308)	0.0006	(0, 0.0010)
int.-slope cov.	−0.006	(−0.010, −0.001)	−0.006	(−0.013, −0.004)
residual var.	$1/n_j$		$1/n_j$	
deviance	−46.1		−54.1	

Table 10. Parameter estimates (and bootstrap 95 % confidence intervals) of meta-analysis of studies about effect of class size (C.S.) on achievement.

Conclusion: negative, but diverse, effect of class size.

Goldstein et al. continue their analysis by including
the 9th study which does have information on individual pupils.

They make a model which jointly incorporates
the aggregate data from the first 8 studies
and the individual data from the last study.
Results are similar.

Continue theory.

Up to this point, presentation focused on individuals in groups.

For *longitudinal data/repeated measures*
(and also multivariate data),
the multilevel approach provides a very flexible procedure:

∗ *Fixed occasion designs*

⇒ missing data on dependent variable can be simply handled by omitting these measurement occasions for these individuals

⇒ measurement-specific covariates can be used.

∗ *Longitudinal designs in general*
(with arbitrary measurement moments)

⇒ model a population of curves

⇒ different numbers & times of measurement for different individuals are no problem.

skip

What is a ‘level’ in the sense of multi*level* analysis?

As we saw above: the technical answer is
a population of random coefficients.

Alternative description: *a source of (unexplained) variation*.

In traditional ‘OLS’ regression analysis (OLS = Ordinary Least Squares)
there is only one source of variation:
the respondents / investigation subjects / cases in the analysis.

However, it is more realistic in the social sciences to acknowledge
multiple sources of variation for the studied phenomena.

Multiple actors:

pupils, teachers, principals, politicians;
patients, members of their social network,
general practitioners, specialists, hospital administrators;
employees, friends, colleagues, team leaders, employers;
etcetera.

Contexts:

for pupils:
classroom, family, school, neighborhood;
for patients:
family, neighborhood, region, general practice, hospital;
for employees:
department, organization;
generally:
region (e.g., labor market, presence of social problems); country;
etcetera.

Actors and contexts can be, *often are*, sources of variability.

In designing the analysis strategy,
determine the main of these sources.

(Often, contexts and actors are confounded;
e.g., teachers and classrooms.)

"Traditional" multilevel analysis
focuses on *nested sources of variability*.

This often is a good approximation but sometimes is insufficient:
e.g. schools – neighborhoods,
primary schools – secondary schools,
hospitals – neighborhoods,
general practitioners - specialists.

Crossed random coefficients can also be modeled.

When does it make sense to use the Hierarchical Linear Model for data with a nested structure?

(These are the same considerations that always apply to the use of random effect models.)

* Is it meaningful to consider the units (“cases”) at each level as a sample from a population?

* Are there unobserved effects that can be regarded as sources of unexplained variance?

* Are the residuals approximately normally distributed?

* Does the research question and available data imply the testing of effects of higher-level variables?

Rule of Thumb:

* less than 20 groups: random coefficients doubtful
* more than 20 groups: it may be sensible to use the HLM.

Relevant questions in this case:
Is the sample at the group level adequate?
Even if it is not a random sample, the question is whether this sample can be regarded as representative for some population (perhaps hypothetical) of higher-level units; and whether the residuals are exchangeable, which can be promoted by inclusion of relevant explanatory variables.
The usual caveats apply:
statistical modelling with small samples is risky.

Methodological message:

* Hierarchical Linear Model "automatically" accomodates for combination of within-group and between-group variability
  $\Rightarrow$ correct standard errors

* Differentiate between
  $\Rightarrow$ within-group regressions
  $\Rightarrow$ between-group regressions
  because these have different meanings
  (ecological and atomistic fallacies)

* Random slopes are interesting: unexplained context effects
  $\Rightarrow$ explanation possible through cross-level interactions.

For the historical consciousness

Traditional: 1950 – 1960

* random effects
* mixed models
* within- and between-group regressions, ecological fallacies.

Developments 1980 – 1990

(Goldstein; Bryk & Raudenbush; Aitkin & Longford; Laird & Ware; Mason, Wong & Entwisle)

* random slopes
* correlated random effects
* algorithms and software (VARCL, HLM, ML2)
* methodology for contextual & longitudinal analysis.

Consolidation since 1990

* more interpretations & applications

* MCMC algorithms for more complicated data structures

* latent variables (Rabe-Hesketh & Skrondal)

* further elaboration special topics:
  design, non-normal distributions, missing data, diagnostics, etc.

* textbooks

* more software, integration in mainstream software:
  MLwin, HLM, MPlus, Mixor/Mixreg, SAS, Stata, R, S-plus, SPSS.

Websites :

Multilevel Models Project :

`http://www.mlwin.com/`

Longitudinal and Multilevel Methods Project :

`http://www-personal.umich.edu/~rauden/`

Both can be reached from my homepage

`http://stat.gamma.rug.nl/snijders/multilevel`

Some textbooks:

* A. Bryk & S. Raudenbush, Sage 1992, 2002
* A. Gelman & J. Hill, CUP 2007
* H. Goldstein, Edward Arnold 1995, 2003
* J. Hox, LEA 2002
* J. Pinheiro & D. Bates, Springer 2000
* J. de Leeuw & E. Meijer, *Handbook of Multilevel Analysis*, Springer 2008
* T. Snijders & R. Bosker, Sage 1999
  with website `http://stat.gamma.rug.nl/snijders/mlbook1.htm`

Multilevel discussion list .