Multilevel Models in Demography

Dr Maria Rita Testa

Habilitation lecture
Vienna University of Economics and Business
December 21, 2015

Elise Richter
1865-1943
Outline

1. Definitions and historical developments
2. Basics of the multilevel models
3. Applied examples and key methodological issues
4. My own research using multilevel models
Demography

The scientific study of population, with a central focus on the size, distribution and composition of population, and the processes of population change.

Multilevel models

Analysis of data collected at multiple levels of aggregation, such as surveys that collect data on children, their mothers and the communities where they live. Hierarchical system of regression equations
Development of multilevel models

1. **1980s**: publication of a number of research papers in the field of multilevel analysis

2. **1990s**: development of book-length treatments and software packages (MLn/MLwiN, HLM, VARCL)

3. **2000s**: multilevel procedures incorporated in standard statistical software packages (Stata, R, SAS)
Development of demography

1. **1662**: Statistical study of human population (John Graunt’s life table)
2. **1900**: Cross sectional analysis based on population censuses
3. **1950s**: Cohort analysis based on vital statistics and population register
4. **1970s**: Shift from aggregate- to individual-level analysis based on unit record data sample (surveys)
5. **1980s**: Event history analysis (based on retrospective information) with a multilevel approach
A synthesis between aggregate and individual-level approaches

Multilevel analysis: a new paradigm for social sciences

- Removes both the ecological and atomistic fallacy by considering aggregate characteristics as external factors that affect individuals’ behaviour and by incorporating the context into individual-level models.

- Transcends the opposition between holism and methodological individualism by allowing the examination of a large number of aggregation levels.

- Creates opportunity for a possible convergence of different disciplines in social sciences (new theory).

Source: Courgeau 2007
Different fields, different names

Design of experiments: variance components models
Statistics: mixed models (Harville, 1977), hierarchical linear models (HLM)
Econometrics: random coefficients models (Swamy 1972), random effects models for panel data
Biostatistics: mixed models for repeated measures (Laird and Ware, 1982), random effects models
Educational statistics: multilevel models (Cronbach 1976, Aitkin and Longford 1986)
Linear mixed models

Linear models with both fixed and random effects imposing a correlation structure on the observations

Generalized linear mixed models

Models for binary, categorical, count or survival data
One example in demography


They study contraceptive use in 15 countries using data from the World Fertility Survey (WFS) Explanatory variables: years of education and childhood residence (individual-level) and index of family planning efforts (country-level)
Using one-level model with clustered data

A standard logit model

\[ \text{logit}(\pi_{ij}) = \alpha + \beta x_{ij} \]

\( \pi_{ij} \) = Probability of having ever used contraception

\( i-th \) Woman aged 40-44

\( j-th \) Country; \( j=1, 2, \ldots 15 \)

\( x_{ij} \) Individual level predictor (urban/rural place of residence and years of education)

Problem:

Observations will not be independent if women within each country share unobserved characteristics. Consequence: biased standard errors (often underestimated, leading to type I error rates higher than the nominal level \( \alpha \))
Building multilevel models

**Step 1**: two-level empty model (random intercept)

\[
\text{logit}(\pi_{ij}) = \alpha + a_j
\]

The random effects are independent across countries and identically distributed with

\[a_j \sim N(0, \sigma_a^2)\]

In the study of Wong and Mason the random effect captures the variation across countries in the level of contraceptive use.
Building multilevel models

**Step 2**: two-level model (random intercept)

\[
\text{logit}(\pi_{ij}) = (\alpha + a_j) + \beta x_{ij}
\]

The *intra-class correlation coefficient* \((\rho)\) measures the share of country-level variance

\[
\rho = \frac{\sigma_a^2}{\sigma_a^2 + \frac{\pi^2}{3}}
\]
Step 3: two-level model (random slope)

\[ \logit(\pi_{ij}) = (\alpha + \alpha_j) + (\beta + b_j)x_{ij} \]

The random slope $b_j$
Allow the effect of an individual-level covariate to vary across countries.

Level-two residuals are assumed to have a bivariate normal distribution with unstructured variance-covariance matrix (intercept and slope can be correlated)

\[ \left( \begin{array}{c} \alpha_j \\ b_j \end{array} \right) \sim N \left( \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{array} \right) \right) \]
Fixed or random effects?

**Fixed effects models:**

i. Focus on within-group variation and therefore adjust implicitly for any omitted country-level confounder.

ii. Do not require assumptions on the random effects (exogeneity, homoscedasticity and normality)

iii. Can be used even with just a few countries (with large samples on each)

iv. Inference for the observed countries

**Random effects models:**

i. Capture both within- and between-group variation

ii. Are parsimonious

iii. Need a minimum number of countries even with a small sample size *(borrowing strength)*

iv. Inference for a population of countries
Building multilevel models

**Step 4: two-level model with contextual variable** \((z_j)\)

\[
\text{logit}(\pi_{ij}) = (\alpha_0 + \alpha_1 z_j + a_j) + (\beta + b_j) x_{ij}
\]

The **group mean** of an individual-level variable \(\bar{x}_j\) may be included in the model together with the individual-level predictor \(x_{ij} - \bar{x}_j\)

\[
\text{logit}(\pi_{ij}) = (\alpha + a_j) + \beta_B \bar{x}_j + \beta_W (x_{ij} - \bar{x}_j)
\]

\(B=\) between group variation
\(W=\) within group variation
Building multilevel models

**Step 5:** two-level model with cross-level interactions \((z_j)\)

\[
\text{logit}(\pi_{ij}) = (\alpha_0 + \alpha_1 z_j + a_j) + (\beta_0 + \beta_1 z_j + b_j)x_{ij}
\]

Rearranging:

\[
\text{logit}(\pi_{ij}) = (\alpha_0 + \alpha_1 z_j + \beta_0 x_{ij} + \beta_1 z_j x_{ij}) + (a_j + b_j x_{ij})
\]

- Fixed part
- Random part
Coefficients and probabilities

**Fixed coefficients**: interpreted in the usual way (odds ratios).

**Random coefficients**: capture residual variation across countries in levels and differentials in ever use of contraception

**Predicted probabilities**: subject-specific and population-average
Subject-specific: apply to an individual at given values of the *fixed* and *random* predictors.
Population-average: apply to an individual at given values of the *fixed* predictors but averaging over all random effects
Interpretation of fixed coefficients

Estimates of fixed coefficients

<table>
<thead>
<tr>
<th></th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>[FPE]</td>
<td>1.088</td>
</tr>
<tr>
<td>[Education]</td>
<td>1.239</td>
</tr>
<tr>
<td>[FPE*Education]</td>
<td>-0.995</td>
</tr>
<tr>
<td>[Urban]</td>
<td>1.633</td>
</tr>
</tbody>
</table>

- The odds of using contraception are higher in countries with higher level of family planning effort (9% higher odds per point of effort)
- In countries with no program, each year of education is associated with 24% higher odds of ever use (the education differential is smaller at higher level of effort)
- Women who grew up in urban areas have 63% higher odds of having used contraception than women who group up in rural areas
Interpretation of variance parameters

<table>
<thead>
<tr>
<th>Standard deviations of random effects</th>
<th>Standard deviation</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_a$ [Random intercept]</td>
<td>0.9017</td>
<td>2.4596</td>
</tr>
<tr>
<td>$\sigma_b$ [Random slope for education]</td>
<td>0.0681</td>
<td>1.0703</td>
</tr>
</tbody>
</table>

$\sigma_a \sigma_b$ can be interpreted as an ordinary logit coefficient, exponentiating it to obtain the odds ratio.

The odds ratio of ever use contraception in a country that is one standard deviation above the mean are 2.5 times the odds for comparable women in the average country.

The odds ratio per year of education in a country that is one standard deviation above the mean is 7% higher than in the average country.
Multilevel vs. single-level results

Odds ratios

One-level and two-level estimates for the complete immunization among children receiving any immunization

<table>
<thead>
<tr>
<th></th>
<th>Single-level model</th>
<th>Multilevel model</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Child aged 2, 3 or 4]</td>
<td>2.54</td>
<td>5.27</td>
</tr>
<tr>
<td>[Rural community]</td>
<td>0.56</td>
<td>0.35</td>
</tr>
<tr>
<td>[Proportion indigenous]</td>
<td>0.47</td>
<td>0.19</td>
</tr>
</tbody>
</table>

- **multilevel odds ratios are more extreme** (different from 1) than the single-level odds ratios

Multilevel vs. single-level results

Probabilities

---

**Urban community, Child aged 1**

**Urban community, Child aged 2, 3 or 4**

**Rural community, Child aged 1**

**Rural community, Child aged 2, 3 or 4**

Multilevel probabilities are more extreme (different from .5) than the single-level probabilities.
Why multilevel probabilities are more extreme than single-level probabilities?

- The single-level model fits overall population-averaged or marginal probabilities (the model implicitly averages over the level-two unobservables).
- The multilevel model fits cluster-specific or conditional probabilities for the individual families (explicit condition on the unobservables). Setting the random effects to zero we do not obtain the average population probabilities but the median population probabilities.
- More generally, the higher the cluster-level variance, the more the population-averaged and cluster-specific inferences diverge.

### How many countries are needed?

<table>
<thead>
<tr>
<th>Harmonized multi-country datasets</th>
<th>Number of countries per data round</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generation and Gender Surveys (GGP)</td>
<td>11</td>
</tr>
<tr>
<td>Eurobarometer (EB)</td>
<td>27</td>
</tr>
<tr>
<td>European Community Household Panel (ECHP)</td>
<td>15</td>
</tr>
<tr>
<td>European Quality of Life Survey (EQLS)</td>
<td>31</td>
</tr>
<tr>
<td>European Social Survey (ESS)</td>
<td>30</td>
</tr>
<tr>
<td>European Union Statistics on Income and Living Conditions (EU-SILC)</td>
<td>27</td>
</tr>
<tr>
<td>European Values Study (EVS)</td>
<td>45</td>
</tr>
<tr>
<td>International Social Survey Program (ISSP)</td>
<td>36</td>
</tr>
<tr>
<td>Luxembourg Income Study (LIS)</td>
<td>32</td>
</tr>
<tr>
<td>Survey of Health, Ageing and Retirement in Europe (SHARE)</td>
<td>14</td>
</tr>
</tbody>
</table>

How many countries are needed?

At least **30 countries** are needed for reliable results while using logit models. ML with adaptive Gaussian quadrature produces more accurate estimates.

Which **strategies** with too few level-two cases?

1) Supplement the regression-based modeling with descriptive analysis
2) Small sample corrections and bootstrapping (only linear models)
3) Use of Bayesian estimator

*Source:* Bryan and Jenkins 2015
**EB vs ML estimates**

**ML approach:** Maximum likelihood estimator provide estimates of the parameters that maximize the ‘likelihood function’, i.e., the function that describes the probability of observing the sample data given the specific values of the parameter estimates.

**EB approach:** The EB approach evaluates the posterior at the ML estimates of these parameters and then takes the mean or the mode. In the logit model the calculation requires numerical integration.

**Shrinkage estimates:** the Bayes residuals are closer to the mean than the corresponding ML residuals (biased) but they are also more precise.

**EB estimates** are a *compromise* between having separate estimates for each country and having a common estimate for all (with more weight to the separate estimates where there is more information).
Focusing on the regional level

Testa and Grilli (2006) consider regions as the second-level of unit to overcome the problem of limited number of countries (15 EU countries from 2001 EB data).

Random intercept proportional odds model for a given ideal family size among individuals desiring at least one child. Individuals aged 20-39

<table>
<thead>
<tr>
<th></th>
<th>All individuals</th>
<th>Individuals with children</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Mean number of children ]</td>
<td>1.89 ***</td>
<td>1.49***</td>
</tr>
<tr>
<td>[Mean number of children]^2</td>
<td>-0.63 *</td>
<td>-0.27</td>
</tr>
<tr>
<td>$\sigma_a^2$ [Regional-level variance]</td>
<td>0.14 *</td>
<td>0.18+</td>
</tr>
<tr>
<td>Individuals</td>
<td>4,932</td>
<td>2,223</td>
</tr>
<tr>
<td>Regions</td>
<td>72</td>
<td>72</td>
</tr>
<tr>
<td>Countries</td>
<td>16</td>
<td>16</td>
</tr>
</tbody>
</table>
Predicted probability of a given ideal family size among cohorts aged 20-39 according to the mean number of children ever born among older cohorts aged 40-60

Testa & Grilli 2006
Focusing on the contextual variable

“Multilevel models assume random sampling at all levels, while our survey design in fact does not use sampling at the country level. Even in such a circumstance, multilevel models could be useful because they allow the explicit inclusion of country-level explanatory variables and country-level residual variation (Hox, vand de Schoot, & Matthijsse, 2012).”

Testa 2014

Rationale for using random effects in the census case: parsimonious description of the observed variability among clusters, accounting for measurement error, generalizability in space and time
Focusing on the contextual variable

Testa (2014) (27 EU countries from 2011 EB data).

Random intercept proportional odds model for a given number of intended children. Childless women aged 20-45

<table>
<thead>
<tr>
<th></th>
<th>Empty model</th>
<th>Model with individual-level covariates</th>
<th>Model with individual- and country-level covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Medium education]</td>
<td>--</td>
<td>-0.06</td>
<td>-0.07</td>
</tr>
<tr>
<td>[High education]</td>
<td>--</td>
<td>0.34*</td>
<td>0.33+</td>
</tr>
<tr>
<td>[Country share of highly educated women]</td>
<td>--</td>
<td>--</td>
<td>0.02**</td>
</tr>
<tr>
<td>$\sigma^2_\alpha$ [Country-level variance]</td>
<td>0.19***</td>
<td>0.15***</td>
<td>0.11***</td>
</tr>
<tr>
<td>Individuals</td>
<td>3,332</td>
<td>3,332</td>
<td>3,332</td>
</tr>
<tr>
<td>Countries</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
</tbody>
</table>
Cross-country correlation between share of women highly educated and mean ultimately intended family size. EU-27. YEAR 2011

Source: Testa 2014
Cross country correlation between share of women highly educated and mean actual family size of highly educated women.

EU-27. YEAR 2011

Source: Testa 2014
**Individual or contextual effect?**


Random intercept proportional odds model for *uncertainty* in reaching a given intended family size. Individuals with one child aged 20-45

<table>
<thead>
<tr>
<th></th>
<th>Empty model</th>
<th>Model with individual-level covariates</th>
<th>Model with individual and country level covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Worsening of household financial situation in the previous 5 years]</td>
<td>0.14</td>
<td>0.11</td>
<td>0.00</td>
</tr>
<tr>
<td>[Country mean of worsening of household financial situation]</td>
<td>--</td>
<td>--</td>
<td>3.43*</td>
</tr>
<tr>
<td>$\sigma^2_\alpha$ [Country-level variance]</td>
<td>0.10*</td>
<td>0.09*</td>
<td>0.03</td>
</tr>
<tr>
<td>Individuals</td>
<td>1,015</td>
<td>1,015</td>
<td>1,015</td>
</tr>
<tr>
<td>Countries</td>
<td>27</td>
<td>27</td>
<td>27</td>
</tr>
</tbody>
</table>
Individual or contextual effect?

Share of people of reproductive ages (20-45) perceiving a worsening in household’s financial situation over the past five years (2006-2011).

EU27. EB data 2011

Source: Testa & Basten 2014
In principle, the extension of the two-level regression model to three and more levels is straightforward. The resulting models are difficult to follow from a conceptual point of view and difficult to estimate in practice.

The imagination of researchers “... can easily outrun the capacity of the data, the computer, and current optimization techniques to provide robust estimates.” (DiPrete and Forristal 1994, p.349)
Multilevel models are not a panacea

*Contextual fallacy*: the statistically significant effects of aggregate-level variables may be the results of a poor specification of individual-level relationships (Hauser 1974)

‘With large samples sizes of individuals within each country but only a small number of countries, estimates of the parameters summarizing country effects are likely to be unreliable’ (Bryan & Jenkins 2015)
Demographic studies using multilevel models

1. **Mills and Begall 2010**: look at the intention to have a third child in 24 countries using a random-intercept-ordered logit model which includes a cross-level interaction between a country-level gender gap index and sex of the previous child.

2. **NG et al 2006**: illustrate simulated ML using contraceptive prevalence data from Bangladesh.

3. **Pamuk et al. 2011** use multilevel logistic regression to study likelihood of infant death using a three-level logit model with random intercepts at the community and country levels.

4. **Pebley at 1996** use three-level logit models to study prenatal care and childhood immunization in Guatemala.

5. **Schmertmann et al 2013**: estimate total fertility for many small areas in Brazil using a combination of EB estimates to smooth age specific fertility with a Brass P/F correction adapted to work when fertility is declining rapidly.
Demographic studies using multilevel event history models

1. **Guo 1993** uses siblings data in a two-level model to estimate child survival in Guatemala

2. **Barber et al 2000**: use a discrete time three-level model to study the initiation of permanent contraceptive use in Nepal

3. **Lillard 1993**: was a pioneer in considering joint modeling of marriage and fertility

4. **Steele et al 1996**: study duration of contraceptive use in China using a multinomial version of a discrete hazards model to handle competing risks

5. **Steele et al 1999**: look at the impact of family planning service provision on contraceptive use in Morocco using multilevel event history models

6. **Steele et al 2006**: use multilevel/multiprocess models to study cohabitation and childbearing among young British women allowing for endogeneity of fertility decisions
Testa’s publications using multilevel


A micro-macro model of fertility

Cultural, institutional, social, economic context

Individual characteristics

Individual reproductive intentions & behaviour

Social outcome: Fertility rates

“Does education influence fertility intentions?”

The ReCap Project

The ReCap project aims to explain the dynamics between reproductive decision-making and human capital in a cross-country comparative approach by investigating the conditions under which a positive relationship between women’s educational level and...


