Multilevel estimation of contextual effects

Philip S. Morrison.
Victoria University of Wellington

Philip.Morrison@vuw.ac.nz

Oceania Stata User Group Meeting.
The University of Sydney Business School
28-30 September 2016
What’s the problem?

The clustering of individuals in groups

There are two problems: statistical and substantive

**Statistical:** Clustering means your sample is not made up of independent (uncorrelated) individuals. Therefore you have fewer independent observations than you think. Without adjustment your standard errors are under estimated and the chances of Type 1 errors are higher.
Substantive: Conceptually, measurements of outcomes of micro-level processes on individuals may reflect the (macro) context in which the processes operate.

Examples:

• **Education**: learning takes place in classes in schools

• **Public health**: people grow up in neighbourhoods

• **Labour economics**: workers perform within firms

• **Management**: leadership operates within organisations

Admitting the presence of multiple levels means that your theory has to be articulated at two levels:

• How individuals behave in general (at the micro level)

• How individuals behave in specific contexts (the macro level)
What happens if we ignore context?


\[
\text{LIFE SATISFACTION} = a + b \text{ INCOME(log)} + e
\]

The general (micro) relationship

Contexts are countries and their wealth affects both the intercept and slope.

Context specific relationships

```
Life Satisfaction
```

```
Income (log)
```

```
Life Satisfaction
```

```
Income (log)
```

```
Rich Countries
```

```
Poor Countries
```

The Peggy Schyns example illustrates the value of applying the multilevel model. Among other things it tells us that the wellbeing returns to raising incomes are higher in low income countries. The general, micro-level model, is not general after all.

**Introductions to multilevel modelling**


**Advanced**

Other basic resources

Stata manual

https://www.stata.com/manuals13/me.pdf

Huber, C. Multilevel linear models in Stata, part 1: components of variance.

Stata YouTube


Module 7: Multilevel models for binary responses. George Leckie. Centre for Multilevel modelling

http://www.bristol.ac.uk/media-library/sites/cmm/migrated/documents/7-practicals-stata-sample.pdf

ESS EduNet. European Social Survey education
Learning multilevel analysis. Prof Kristen Ringdal. Contains Stata syntax

http://essedunet.nsd.uib.no/cms/topics/multilevel/
Contemporary approaches involve specifying the general model in terms of fixed effects and the context as a random variable. Hence ‘mixed’ (ME = mixed estimation).

Stata offers a suite of ME routines depending mainly on the way your dependent variable is measured.

Mixed Mixed-effects linear regression
Mixlm Mixed-effects generalized linear regression
Melogit Mixed-effects logistic regression
Meprobit Mixed-effects probit regression
Meologit Mixed-effects ordered logistic regression
Meoprobit Mixed-effects ordered probit regression
Count, multinomial and others
An application

Pride in the city*

Urban pride is an individual and collective response to living in a given city.

Unlike other emotions such as life satisfaction or happiness with which it is weakly positively correlated, pride involves stake holding; to be proud of something requires having an investment in its success either emotionally, financially, culturally or as a participant.

I specify a multilevel model based on responses to a five category survey question on how proud residents are in the ‘look and feel of their city’ drawing on over 6000 residents surveyed in 12 New Zealand cities in 2008.

* Adapted from Morrison, Philip.S. 2016 ‘Pride in the city’ REGIONS (in press as of 19 Oct 2016)  [http://region.wu.ac.at/]
Survey question:

Q: “On a scale of one to five where one is strongly disagree and five is strongly agree, rate your agreement with the statement “I feel a sense of pride in the way [my city] looks and feels.”


<table>
<thead>
<tr>
<th>Response</th>
<th>Frequency</th>
<th>Percent</th>
<th>Cumulative percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly disagree</td>
<td>82</td>
<td>1.34</td>
<td>1.34</td>
</tr>
<tr>
<td>Disagree</td>
<td>389</td>
<td>6.36</td>
<td>7.7</td>
</tr>
<tr>
<td>Neutral</td>
<td>1,803</td>
<td>29.48</td>
<td>37.18</td>
</tr>
<tr>
<td>Agree</td>
<td>2,763</td>
<td>45.17</td>
<td>82.34</td>
</tr>
<tr>
<td>Strongly agree</td>
<td>1,080</td>
<td>17.66</td>
<td>100</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>6,117</strong></td>
<td><strong>100</strong></td>
<td></td>
</tr>
</tbody>
</table>


*Note:* Excludes 21 respondents who did not know.
The location of the twelve cities included in the Quality of Life project.
New Zealand, 2008

1. Rodney
2. North Shore
3. Waitakere
4. Auckland
5. Manukau
6. Hamilton
7. Tauranga
8. Porirua
9. Hutt
10. Wellington
11. Christchurch
12. Dunedin
Beginning with the OLS model

Most studies of responses apply the conventional OLS ‘total’ regression model specified at the level of the \(i\)th individual in which the relationship between the outcome \(y\) and arguments \(X\) are described in terms of fixed parameters, \(\alpha\) and \(\beta\).

\[
y_i = \alpha_o + \beta X_i + \epsilon_i
\]

In such a model the random or allowed-to-vary element is captured by \(\epsilon\), the mean or expected value of which is assumed to be zero.

An accompanying assumption is that there is constant variability and no autocorrelation. The assumption is necessary if it is to be characterised by a single parameter \(\sigma^2_{\epsilon}\), the variance of the error term.
For illustration of the general relationship, lets assume that pride in the city can be ‘explained’ by the age of the resident:

\[ y = \text{pride} \times \text{age} \]

regress pride age

\[
\begin{align*}
\text{Source} & \quad | \quad \text{SS} & \quad \text{df} & \quad \text{MS} & \quad \text{Number of obs} = 6,117 \\
\hline
\text{Model} & \quad 38.0129991 & \quad 1 & \quad 38.0129991 & \quad \text{F(1, 6115) = 50.10} \\
\text{Residual} & \quad 4640.04814 & \quad 6,115 & \quad .758797734 & \quad \text{Prob > F = 0.0000} \\
\text{Total} & \quad 4678.06114 & \quad 6,116 & \quad .764889003 & \quad \text{R-squared = 0.0081} \\
\hline
\end{align*}
\]

\[
\begin{align*}
\text{pride} & \quad | \quad \text{Coef.} & \quad \text{Std. Err.} & \quad \text{t} & \quad \text{P>|t|} & \quad [95\% \text{ Conf. Interval}] \\
\hline
\text{age} & \quad .0044873 & \quad .000634 & \quad 7.08 & \quad 0.000 & \quad .0032445 \quad .0057302 \\
\text{cons} & \quad 3.515622 & \quad .0302126 & \quad 116.36 & \quad 0.000 & \quad 3.456395 \quad 3.574849 \\
\hline
\end{align*}
\]

Does this \textit{general} (micro-level) relationship apply to all cities?

Differences in the linear OLS relationship between urban pride and age across the 12 cities. New Zealand 2008.

The two parameters of the model both vary by city. Let's begin by assuming only intercepts vary.
The random intercepts model

Assume cities are sampled and treat the intercept as a random variable

(1) \[ y_i = \alpha_o + \beta X_i + \varepsilon_i \]

(2) \[ y_i = \alpha_{o_j} + \beta X_i + \varepsilon_{i_j} \]

We now have two subscripts, i = individual and j = city.

\( \alpha_{o_j} \) Indicates variability in the intercept from city to city, the ‘city effect’. We treat this as a ‘random effect’ and represent it as a variance.

\( \varepsilon_{i_j} \) Indicates presence of a second level variability....
Average levels of urban pride are allowed to vary from city to city. The average level of urban pride in city \( j \) is the sum of the city-wide average, \( \alpha_o \), and a varying difference \( u_j \).

The aim of the model is to estimate the fixed intercept, \( \alpha_o \), representing the average level of urban pride across the country, and the variance, \( \sigma^2 \mu \), which measures its inter-city variability about this average.

\[
(2) \quad a_{oj} = \alpha_o + u_j
\]

Combining the micro equation (above) and the macro equation of (2) produces the two-level mixed model in (3):

\[
(3) \quad y_{ij} = \alpha_o + \beta x_{ij} + (u_j + \varepsilon_{ij})
\]

The terms in bold denote the random part.
The initial step in applying the random coefficients model is to estimate the proportion of the variance attributable to differences among individuals and cities. In this null model.

\[ y_{ij} = \alpha_o + (\mu_j + \epsilon_{ij}) \]

The proportion of the variance attributable to individuals is

\[ \sigma^2_\epsilon / (\sigma^2_\epsilon + \sigma^2_\mu) \]

and the variation across cities

\[ \sigma^2_\mu / (\sigma^2_\epsilon + \sigma^2_\mu) \]

the \textit{intra-class correlation} \( \rho(\rho) \).

The intra-class correlation is a measure of the degree to which individuals share common experiences due to their residence in the same city.

If \( \rho \) is greater than zero then there is a case for applying a random coefficients model and its extension as a multilevel model.
Mixed-effects ML regression
Group variable: City

Number of obs = 6,117
Number of groups = 12

Obs per group:
min = 497
avg = 509.8
max = 535

Wald chi2(0) = .
Prob > chi2 = .

Log likelihood = -7698.2485

| pride | Coef.  | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|-------|--------|-----------|-------|------|---------------------|
| _cons | 3.716021 | 0.0612555 | 60.66 | 0.000 | 3.595963 3.83608 |

Random-effects Parameters

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>City: Identity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>var(_cons)</td>
<td>μj</td>
<td>0.0436127</td>
</tr>
<tr>
<td>var(Residual)</td>
<td>εij</td>
<td>0.7206042</td>
</tr>
</tbody>
</table>

LR test vs. linear model: chibar2(01) = 322.29 Prob >= chibar2 = 0.0000
estat icc  // Estimates intraclass correlations. Default is 95% conf. interval

Intraclass correlation

<table>
<thead>
<tr>
<th>Level</th>
<th>ICC</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>City</td>
<td>0.0570685</td>
<td>0.022694</td>
<td>0.0257981</td>
</tr>
</tbody>
</table>

display 4.36/(4.36+72.06)

.057

estat ic  // Gives ll(model), df, AIC and BIC

Akaike's information criterion and Bayesian information criterion

<table>
<thead>
<tr>
<th>Model</th>
<th>Obs</th>
<th>ll(null)</th>
<th>ll(model)</th>
<th>df</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>.</td>
<td>6,117</td>
<td>-7698.248</td>
<td>15402.5</td>
<td>3</td>
<td>15422.65</td>
<td>15422.65</td>
</tr>
</tbody>
</table>

Note: N=Obs used in calculating BIC; see [R] BIC note.
The random intercepts model

The random intercept model of equation 4 implies a different intercept term for each city, $\alpha + \mu_j$; $j = 1,\ldots,12$. These random intercepts are not estimated directly but we can use linear unbiased predictions (BLUPS) of their random effects as shown on the right. Recall that the mean level of pride is 3.71 on the urban pride 1-5 scale with a standard deviation of 0.874. At one extreme the City of Manukau has a half standard deviation measure lower than the grand mean, and Wellington City almost 0.4 higher.

‘Table 1’. Describing the arguments

Measures of stake holding and controls used in the modelling of urban pride.
New Zealand, 2008

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Controls</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>Female</td>
<td>0.53</td>
<td>0.50</td>
</tr>
<tr>
<td>Health</td>
<td>Health good or very good</td>
<td>0.61</td>
<td>0.49</td>
</tr>
<tr>
<td><strong>Emotional stakes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration</td>
<td>Resident in city 10 years +</td>
<td>0.70</td>
<td>0.46</td>
</tr>
<tr>
<td>Community</td>
<td>Sense of community</td>
<td>0.55</td>
<td>0.50</td>
</tr>
<tr>
<td><strong>Financial stakes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Owner</td>
<td>Home owner</td>
<td>0.62</td>
<td>0.49</td>
</tr>
<tr>
<td>Not employed</td>
<td>Not employed</td>
<td>0.26</td>
<td>0.44</td>
</tr>
<tr>
<td>Enough</td>
<td>Income meets everyday needs</td>
<td>0.87</td>
<td>0.34</td>
</tr>
<tr>
<td><strong>Cultural stake</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minority</td>
<td>Non-European</td>
<td>0.23</td>
<td>0.42</td>
</tr>
<tr>
<td><strong>Civic stakes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safe</td>
<td>Feel safe in central city</td>
<td>0.63</td>
<td>0.48</td>
</tr>
<tr>
<td>Clean</td>
<td>No rubbish noticed</td>
<td>0.49</td>
<td>0.50</td>
</tr>
<tr>
<td>Council</td>
<td>Confidence in council decisions</td>
<td>0.46</td>
<td>0.50</td>
</tr>
</tbody>
</table>

The distribution of urban pride. Stake holding fixed effects and city random effects. New Zealand, 2008

| Variable      | Description                              | Coef. | Std Err. | z     | P>|z| |
|---------------|------------------------------------------|-------|----------|-------|------|
| **FIXED EFFECTS** |                                          |       |          |       |      |
| **Controls**  |                                          |       |          |       |      |
| Female        | Female                                   | 0.10  | 0.02     | 4.91  | 0.00 |
| Health        | Health good or very good                 | 0.06  | 0.21     | 3.01  | 0.00 |
| **Emotional stakes** |                                       |       |          |       |      |
| Duration      | Resident in city 10 years +              | 0.11  | 0.02     | 4.78  | 0.00 |
| Community     | Sense of community                       | 0.24  | 0.02     | 11.22 | 0.00 |
| **Financial stakes** |                                       |       |          |       |      |
| Owner         | Home owner                               | 0.08  | 0.22     | 3.80  | 0.00 |
| Not employed  | Not employed                             | 0.06  | 0.02     | 2.45  | 0.01 |
| Enough        | Income meets everyday needs              | 0.10  | 0.03     | 3.17  | 0.00 |
| **Cultural stakes** |                                       |       |          |       |      |
| Minority      | Non-European                             | 0.20  | 0.03     | 7.44  | 0.00 |
| **Civic stakes** |                                       |       |          |       |      |
| Safe          | Feel safe in central city                | 0.21  | 0.02     | 9.37  | 0.00 |
| Clean         | No rubbish noticed                       | 0.23  | 0.02     | 11.20 | 0.00 |
| Council       | Confidence in council decisions          | 0.37  | 0.02     | 17.68 | 0.00 |
| Constant      |                                          | 2.80  | 0.07     | 37.36 | 0.00 |
| **RANDOM EFFECTS** |                                       |       |          |       |      |
|                | Estimate | Std Err. | Estimate | Std Err. |       |       |
| Cities        | Constant                                 | 0.04  | 0.02     |       |      |
| Residual      | 0.61   | 0.01     |       |      |

Number of cases: 5867
Log likelihood: -6897.12
LR test vs linear model: 348.72
Wald chi2 pr=0: 982.88
Df: 14
AIC: 13822.23
Intraclass correlation: 0.07

Source: Quality of Life Survey, 2008
Selected characteristics of the twelve New Zealand cities

<table>
<thead>
<tr>
<th>City</th>
<th>Pride</th>
<th>Population ('000)</th>
<th>Affluence</th>
<th>European</th>
<th>Council</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rodney District</td>
<td>3.56</td>
<td>89.56</td>
<td>0.10</td>
<td>0.95</td>
<td>0.30</td>
</tr>
<tr>
<td>North Shore City</td>
<td>3.90</td>
<td>205.61</td>
<td>0.13</td>
<td>0.77</td>
<td>0.44</td>
</tr>
<tr>
<td>Waitakere City</td>
<td>3.62</td>
<td>186.44</td>
<td>0.07</td>
<td>0.67</td>
<td>0.48</td>
</tr>
<tr>
<td>Auckland City</td>
<td>3.48</td>
<td>404.66</td>
<td>0.14</td>
<td>0.62</td>
<td>0.40</td>
</tr>
<tr>
<td>Manukau City</td>
<td>3.33</td>
<td>328.97</td>
<td>0.07</td>
<td>0.46</td>
<td>0.51</td>
</tr>
<tr>
<td>Hamilton City</td>
<td>3.83</td>
<td>129.25</td>
<td>0.07</td>
<td>0.76</td>
<td>0.57</td>
</tr>
<tr>
<td>Tauranga City</td>
<td>3.87</td>
<td>103.64</td>
<td>0.06</td>
<td>0.88</td>
<td>0.40</td>
</tr>
<tr>
<td>Porirua City</td>
<td>3.57</td>
<td>48.55</td>
<td>0.10</td>
<td>0.66</td>
<td>0.51</td>
</tr>
<tr>
<td>Lower Hutt City</td>
<td>3.61</td>
<td>86.93</td>
<td>0.09</td>
<td>0.75</td>
<td>0.47</td>
</tr>
<tr>
<td>Wellington City</td>
<td>4.12</td>
<td>179.47</td>
<td>0.17</td>
<td>0.81</td>
<td>0.50</td>
</tr>
<tr>
<td>Christchurch City</td>
<td>3.82</td>
<td>348.44</td>
<td>0.07</td>
<td>0.88</td>
<td>0.41</td>
</tr>
<tr>
<td>Dunedin City</td>
<td>3.88</td>
<td>118.68</td>
<td>0.05</td>
<td>0.92</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Source: Census of Population and Dwellings, 2006 and Quality of Life Survey, 2008

**Population** is drawn from the nearest population census (2006).

**Affluence** is the proportion of individuals with pre-tax incomes of over $70,000 per annum. **European** is the proportion of European in the city.

**Council** is the proportion of the city population who agree or strongly agree that the council makes decisions that are in the best interest of their city (aggregated from sample responses).
The multilevel model

In the urban pride case cities are contexts and as such their characteristics may influence the way the micro level arguments raise or lower urban pride.

I test three hypotheses:

1. whether the higher levels of urban pride exhibited by minorities rise as their share of the population increases,

   mixed pride female healthGVG duration community_sense owner not_employed enough i.minority safeCC no_rubbish conf_council ///
i.minority##c.EuropeanPr ///
|| City: EuropeanPr , mle

2. whether not having enough money lowers urban pride to a greater extent more affluent cities, and

3. whether the individuals’ support for council rises in cities where the overall support for council is higher and whether this context effect is greater for owners.
How the impact of minority status urban pride falls as the proportion of European in the city rises.

New Zealand, 2008

The a priori argument is that minorities will return higher levels of pride in cities where they make up a larger share of the population. The greater their proportion the greater the sense of identity and collective strength.

The focus in this case therefore is on the interaction of the level 2 variable ‘European’ and the individual or level 1 variable ‘minority’.

In the fixed effects results above, minorities return higher levels of urban pride than the European majority. Applying the interaction term exposes the fact that urban pride rises with the proportion of European. This rise is much slower in the case of minorities, and, as the dashed line shows, urban pride levels of minority and European converge as the European share grows. Context in this case clearly matters.


Note: With the fixed effects in the model, the addition of the cross-level term (minority x European) is $\beta = -0.710$ (SE=0.19; $z=-3.74$).
The effect of ‘not having enough money’ on urban pride by city affluence by housing tenure. New Zealand, 2008

When having enough money is interacted with city affluence separately for owners and renters, renters without enough money (typically younger residents) return higher levels of urban pride in more affluent cities: the solid line, right panel.

By contrast, owners without enough money (typically older residents), return lower levels of urban pride in more affluent cities (solid line, left panel).

Renters and owners who say they have enough money to meet daily needs return more urban pride in more affluent cities (the dashed lines). However, city affluence has a greater effect on the urban pride of homeowners (dashed line, left panel).


Note: With the same fixed and random effects as above adding the interaction of enough x owner x affluence term yields a coefficient of -3.72 and a standard error of 1.91 and a z of -1.95 and p>(z) of 0.052. The 95% confidence intervals are plotted.
The estimated relationship between urban pride and city wide support for Council among longer and shorter term residents. New Zealand, 2008

Those who see city councils acting in the interests of the majority return higher levels of urban pride. However this relationship may be affected by how long people have lived in the city.

The interaction of duration of residence (level 1) with support for Council (level 2), suggests that the positive relationship between urban pride and the city’s confidence in its council only applies to the longer term residents. The pride experienced by relative newcomers in their city appears unaffected by the confidence the city as a whole has for its council. The 95% confidence intervals are relatively wide in this case but with the fixed effects in the model the interaction between duration and Council is statistically significant.

Note: The estimate of the Council x duration interaction term is $\beta = 0.631$ (se= 0.32), $z = 1.97$. 
Summary

1. In the social sciences, context usually matters – statistically and substantively
2. Stata’s ME commands off most options non-specialist users will need.
3. Running the null model can act as a quick test for clustering
4. The urban pride example above illustrates the role of fixed effects (stake holding in this case) in the micro or level 1 model as well as how the characteristics of the context (the city) interact with level 1 arguments to alter patterns of urban pride.
5. Conceptualising multilevel models invites researchers to be more specific about the theory behind both the micro (level 1) model and the macro (level 2) model and the cross-level interactions.
6. At the end of the day it may be the way the multilevel model forces us to think about the theoretical role of context which is its greatest value.