

# *Modelling contextuality and heterogeneity with multilevel models*

Kelvyn Jones

School of Geographical Sciences, and  
Centre for Multilevel Modelling  
University of Bristol



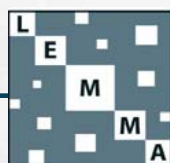
Edinburgh  
6 September 2010

# *Multilevel modelling as Realistically complex modelling*

*Statistical models as a formal framework of analysis with a complexity of structure that matches the system being studied*

- *“non-technical”*: introduction
- *diagrams and graphs*

**Four KEY Notions.....**



## Four KEY Notions

### 1 : Modelling data with a complex structure

A large range of structures that ML can handle routinely; eg houses nested in neighbourhoods

### 2: Modelling heterogeneity

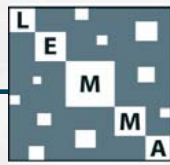
standard regression models 'averages', ie the general relationship  
ML **additionally** models *variances*; eg individual house prices vary from n'hood to neighbourhood

### 3: Modelling dependent data

potentially complex dependencies in the outcome over time, over space, over context; eg houses within a n'hood tend to have similar prices

### 4: Modelling contextuality: micro & macro relations

eg individual house prices depends on individual property characteristics **and** on neighbourhood characteristics



# ***STRUCTURES that generate dependencies***

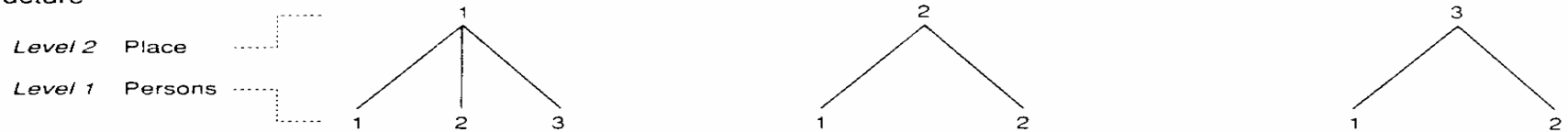
# Modelling data with complex structure

- Structures in the sample and/or in the population

1: Hierarchical structures : model all levels simultaneously

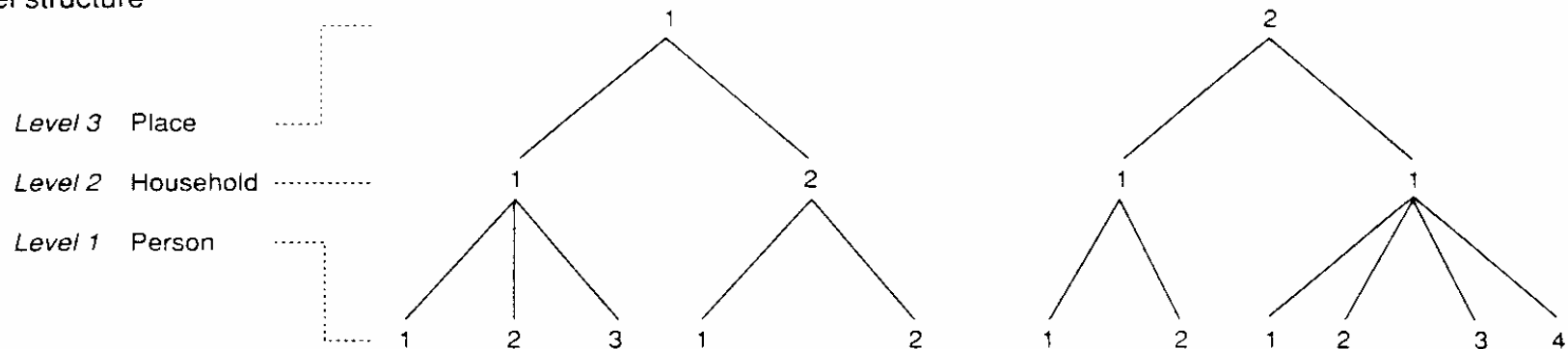
a) People nested within places: two-level model

Two-level structure



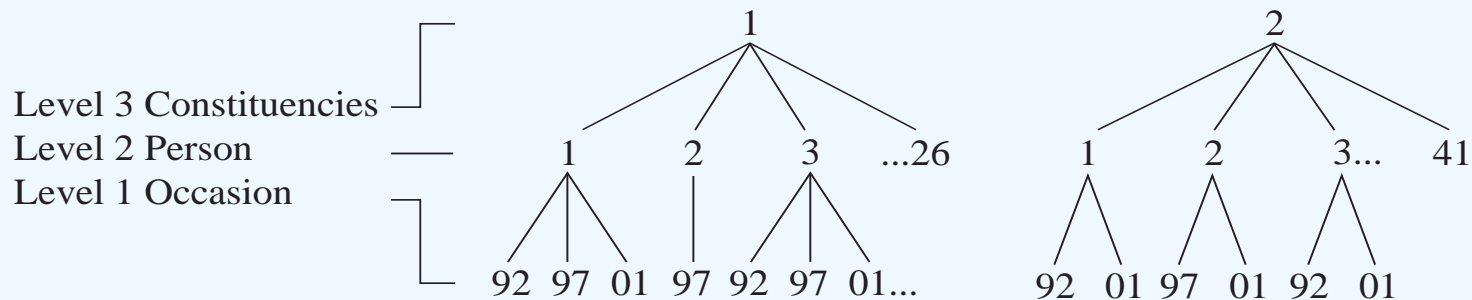
b) People nested within households within places: three-level model

Three-level structure

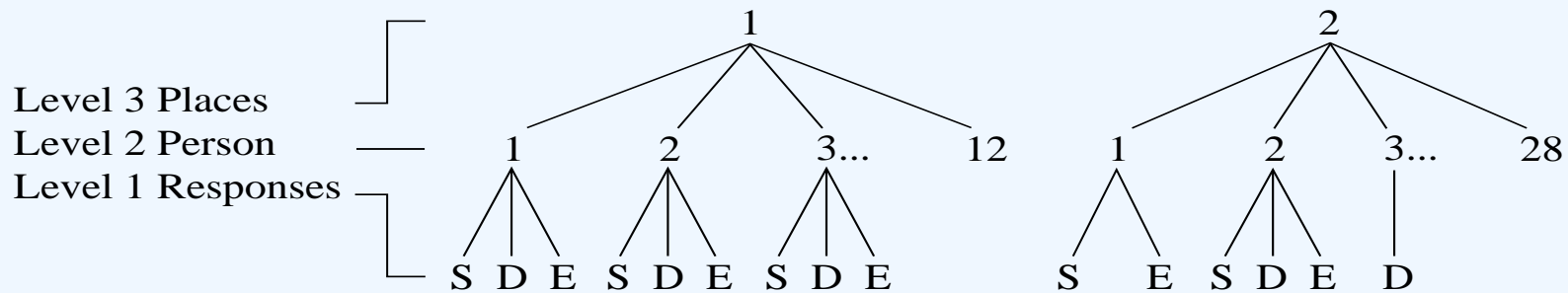


# 1: Hierarchical structures (continued)

## c) Repeated measures of voting behaviour at the UK general election (Panel)

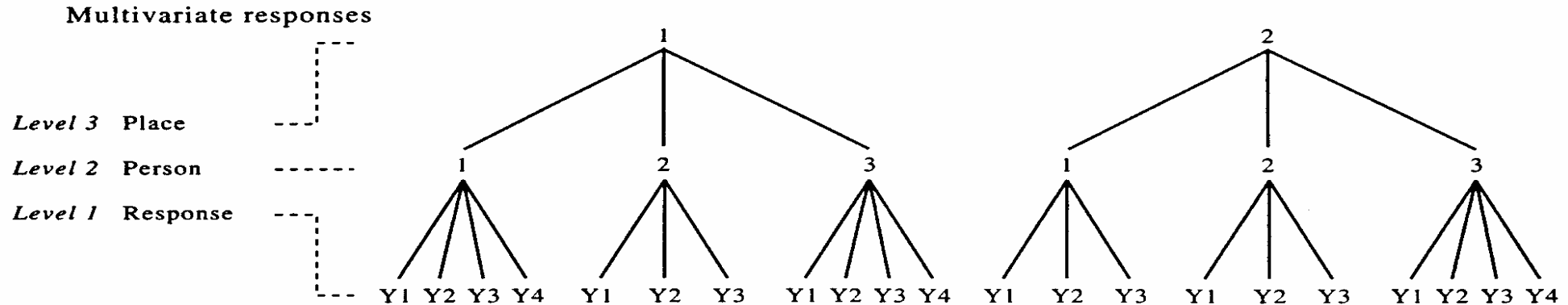


## d) Multivariate design for health-related behaviours



Where S is smoking, D is diet and E is exercise

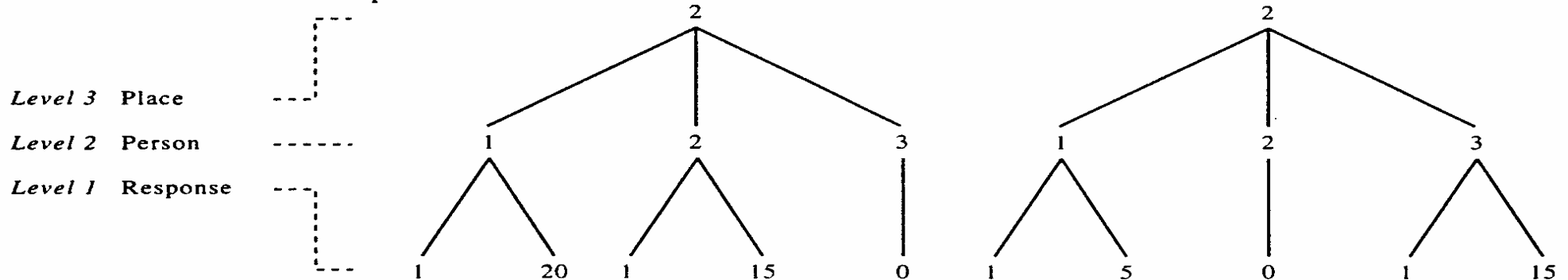
## Multivariate multilevel structures 2



- Handles extreme case of **matrix sampling design**; imbalance by design
- All respondents are asked a set of core questions but additional questions are asked of random subsets of the total sample
- Example: all pupils asked core set of core mathematics questions; but random subsets asked detailed questions on trigonometry, matrix algebra, calculus; all modeled simultaneously; borrows strength

## Multivariate multilevel structures 3

### Mixed multivariate responses



- **Structure:**

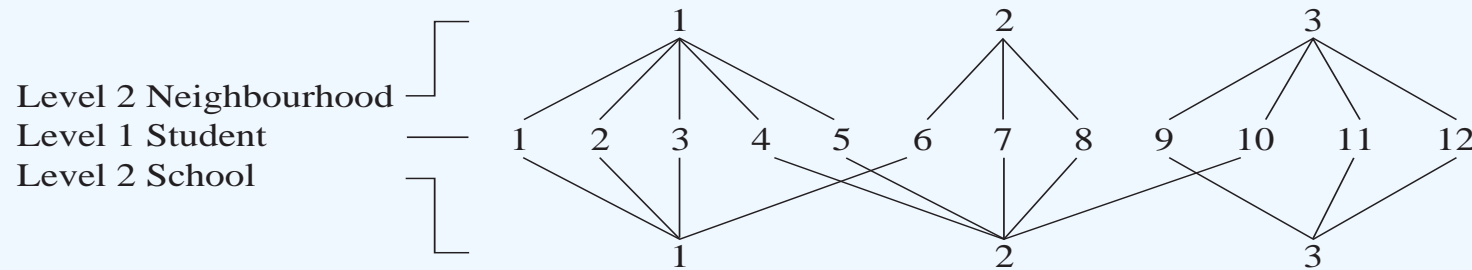
- Level-3: Places; Level-2: Individuals; Level-1: Two responses:
  - 1) smoke or not (a qualitative state)
  - 2) how many cigarettes they consume in a day (a quasi-continuous measure).....Imbalance; data with a spike at zero

- **Question**

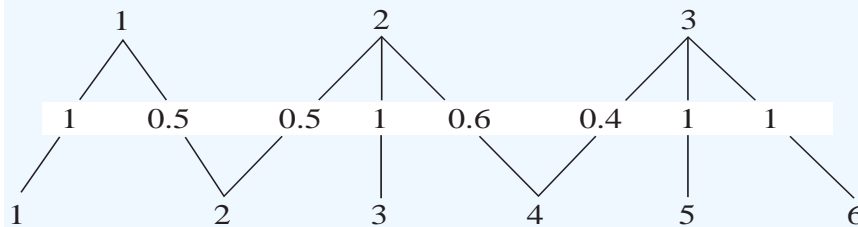
- is social class related to whether you smoke or not; but unrelated to the amount you smoke?
- Are places that have a high proportion of smokers also tend to have higher smoking intensity?

## 2: Non-Hierarchical structures

### a) cross-classified structure



### b) multiple membership with weights

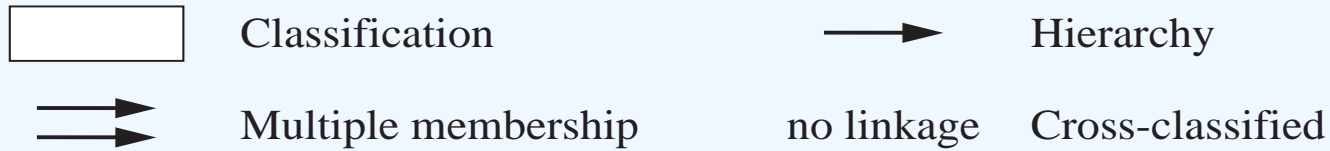


**Level 2 Teachers**

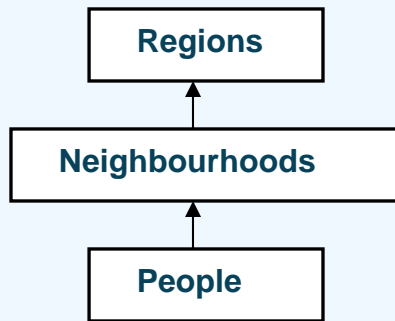
**Level 1 Pupils**

- So far unit diagrams now.....

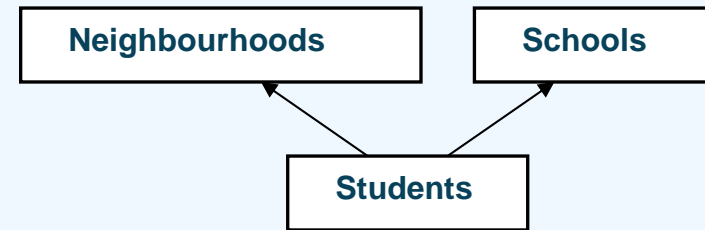
# CLASSIFICATION DIAGRAMS



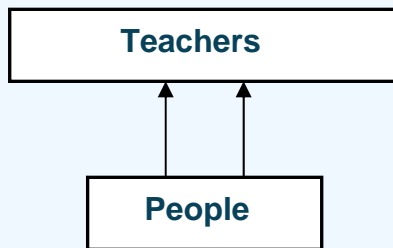
a) 3-level hierarchical structure



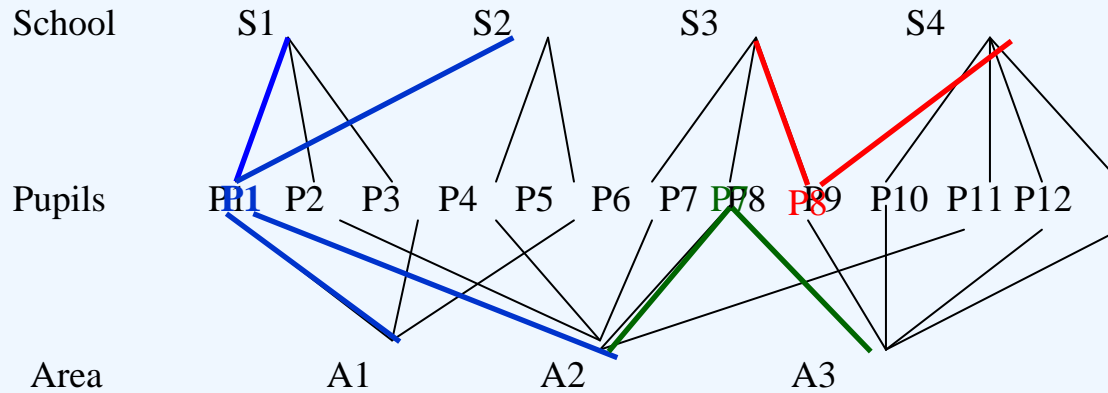
b) cross-classified structure



c) multiple membership structure



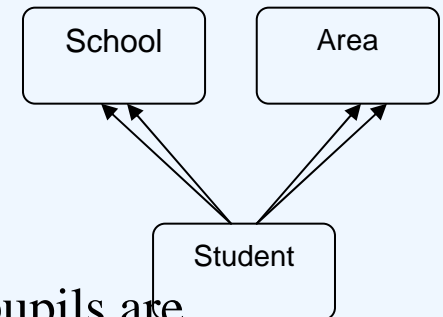
# Combining structures 1: crossed-classifications and multiple membership relationships



Pupil 1 moves in the course of the study from residential area 1 to 2 and from school 1 to 2

Pupil 8 has moved schools but still lives in the same area

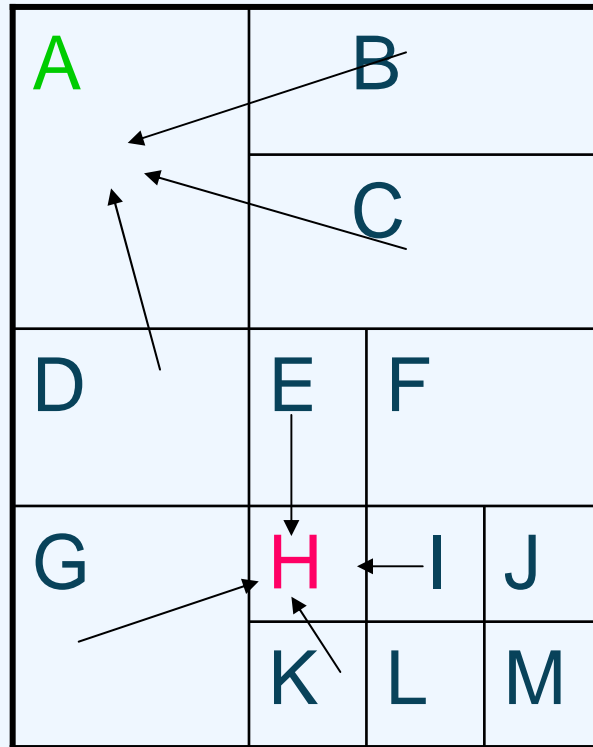
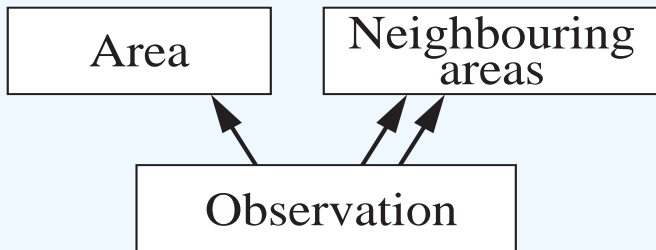
Pupil 7 has moved areas but still attends the same school



Now in addition to schools being crossed with residential areas pupils are *multiple members* of both areas and schools.

# Combining structures 2

## Spatial Models and GWR as combination of strict hierarchy and multiple membership



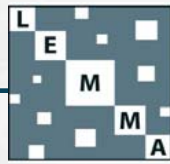
**Person in A**

Affected by A(**SH**) and B,C,D (**MM**)

**Person in H**

Affected by H(**SH**) and E,I,K,G (**MM**)

Multiple membership defined by common boundary; weights as function of inverse distance between centroids of areas



***MODELLING  
HETROGENEITY:  
varying relations, random  
effects & variance  
functions***

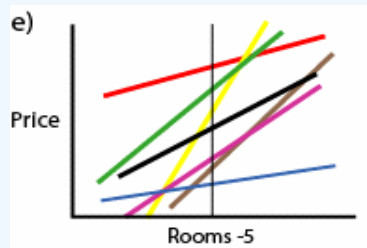
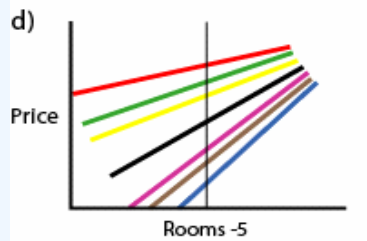
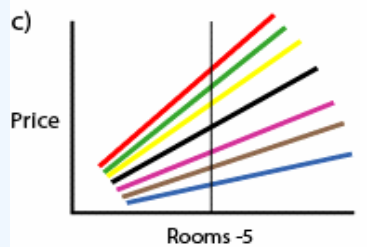
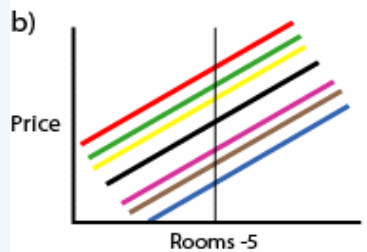
# VARYING RELATIONS

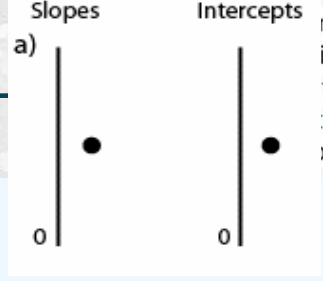
- **Single response:** house price
- **Single predictor**
  - size of house, number of rooms

Rooms	1	2	3	4	5	6	7	8
	-4	-3	-2	-1	0	1	2	3

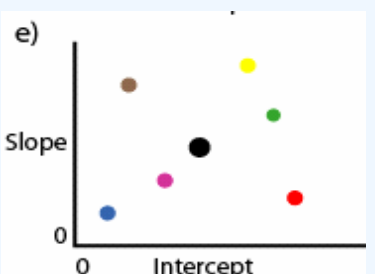
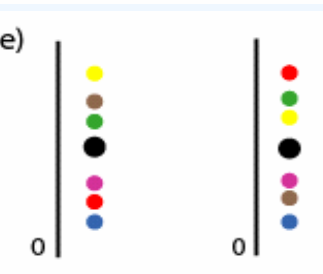
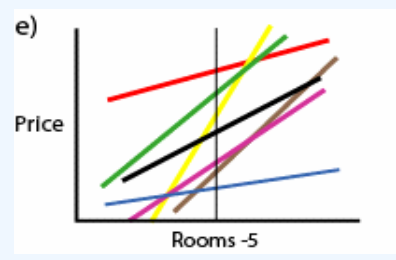
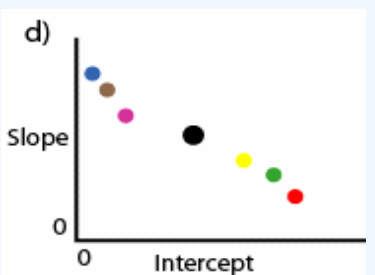
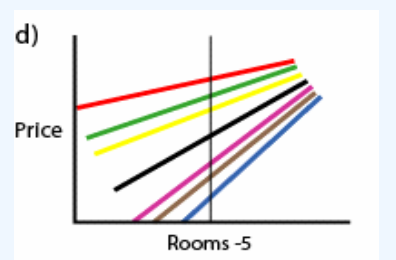
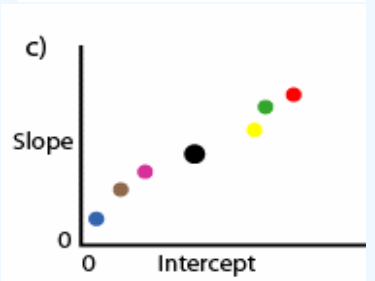
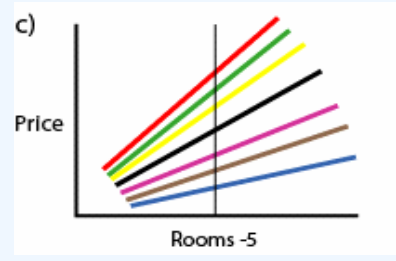
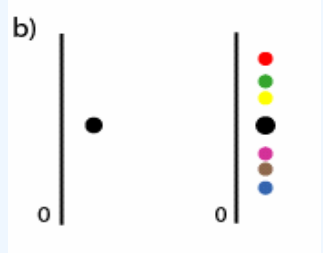
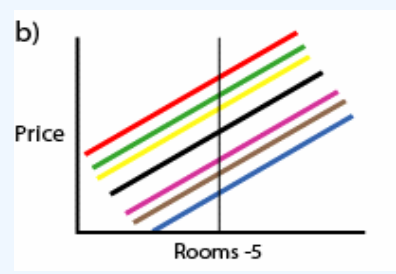
- **Two level hierarchy**
  - houses at level 1 nested within
  - neighbourhoods at level 2 are the contexts

**Set of characteristic plots.....**





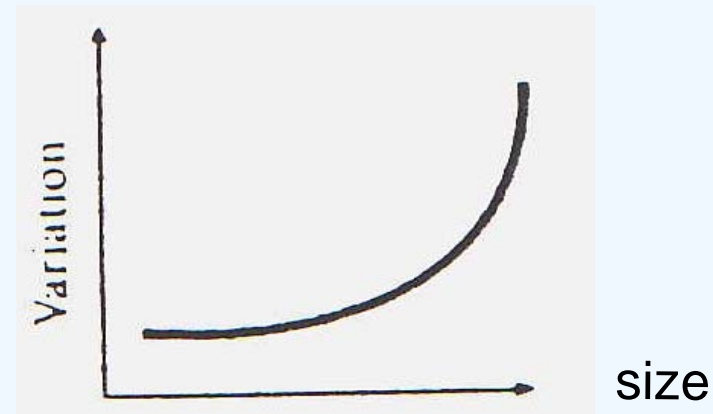
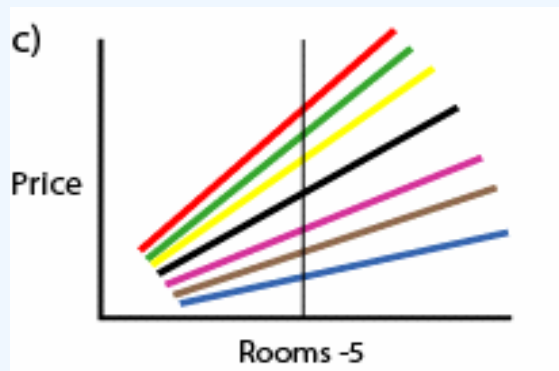
arning  
 ironment for  
 tilevel  
 :hodology and  
 lications



# What are Variance Functions?

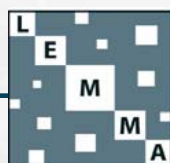
Functions that structuring the variance in terms of other variables

**Example:** between-district variance in price as a variance function of the size of the house



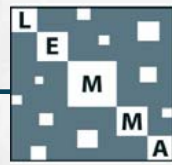
**Regression:** models many means (the fixed part) but single overall variance, called the error term and treated as disturbances

**Multilevel:** explicit modelling of (complex) variances, that is modelling heterogeneity



## WHY Use variance functions?

- Variance heterogeneity eg schools matter, but they matter most for the less able child- between school variation greatest for less able child.
- Impact heterogeneity: eg 2 interventions that reduce the mean of the problem, but one is preferred becomes it narrows the variance; ie more consistent in effect.
- Some research questions are about variances as well as means: segregation, inequality are all about differences *between* units; lots of such work remains descriptive
- First lesson in data analysis; middle **and** scatter around the middle; yet in modelling: homoscedastic assumption, ie all variation consigned to one term and labelled error!

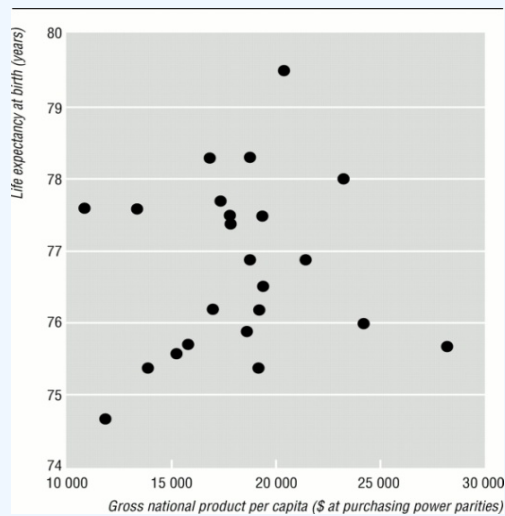


# ***SOME EXAMPLES .....***

# The problem of single level analysis I: aggregate analysis

Wilkinson (1996) *Unhealthy Societies*: relative income hypothesis

Life expectancy and GDP

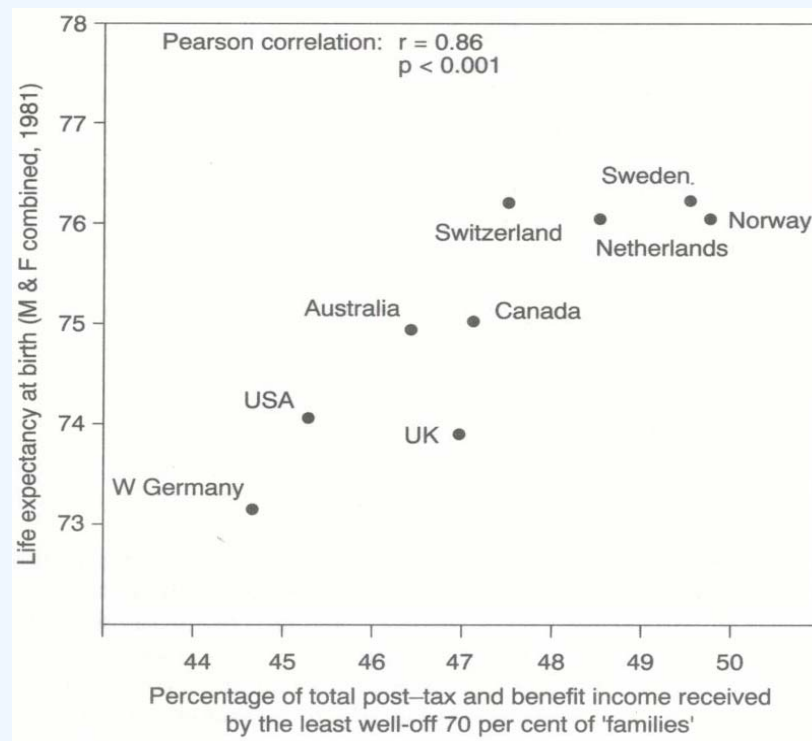


Source: Wilkinson, (US, 1996, p74)

**Places**

**People**

Life expectancy and Income Inequality

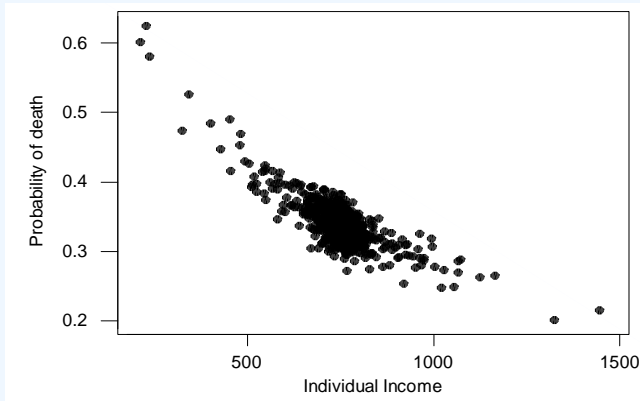


Wilkinson  
 (1966, 76)

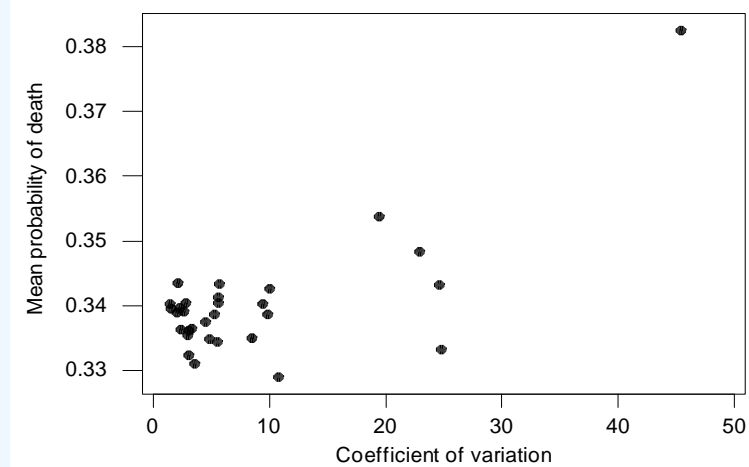
Analysis ONLY at the **aggregate scale**

# The problem of aggregate analysis (with MHJ & RJJ Health and Place, 2009)

Simulated **non-linear** relation  
 between income and mortality



**Artificial** relation between  
 mortality and inequality



## Multilevel analysis

$$\text{Probdeath}_{ij} \sim N(XB, \Omega)$$

$$\text{Probdeath}_{ij} = \beta_{0ij} \text{cons} + -20.752(0.473) \text{Loginc} - 6_{ij}$$

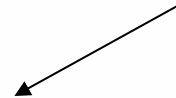
$$\beta_{0ij} = 46.583(0.294) + u_{0ij} + e_{0ij}$$

$$[u_{0ij}] \sim N(0, \Omega_u) : \Omega_u = [0.002(0.035)]$$

$$[e_{0ij}] \sim N(0, \Omega_e) : \Omega_e = [3.300(0.174)]$$

$$-2 * \text{loglikelihood(IGLS Deviance)} = 3024.323(750 \text{ of } 750 \text{ cases in use})$$

**NO between  
 country  
 Variation, BUT not  
 always the case!**



## The problem of single level analysis II: individual analysis

Robinson (1950) calculated the correlation between illiteracy and ethnicity in the USA.

2 scales of analysis for 1930 USA

- Individual: for 97 million people;
- States: 48 units

Level	Black illiteracy	Foreign-born illiteracy
Individual	0.20	0.11
State	0.77	-0.52

- **Robinson's conclusions**

“Ecological correlations are used simply because individual correlations are not available”

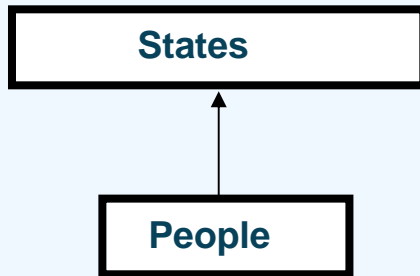
Places

“The paper will have served its purpose if it prevents future computation of meaningless correlations at the ecological level”

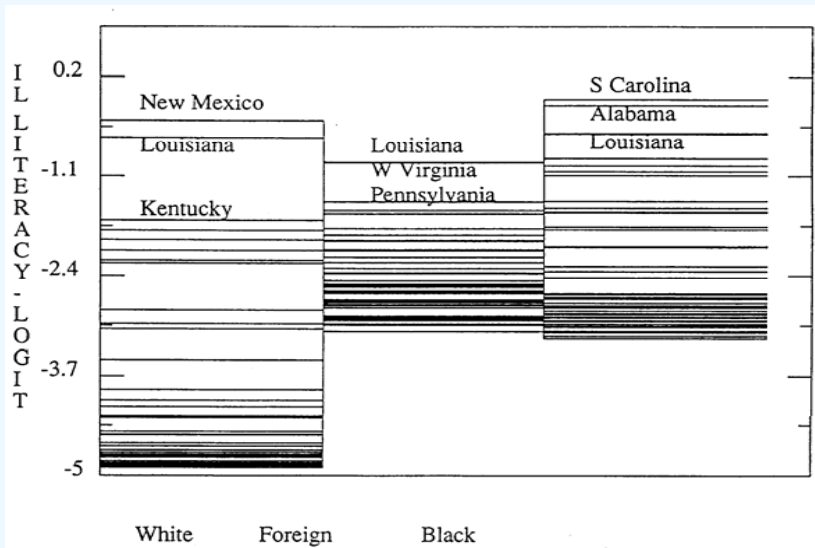
People

# What does an individual analysis miss (IJE, 2009)?

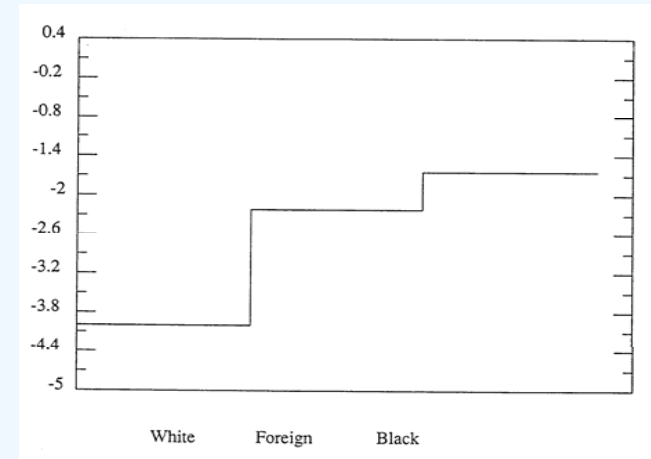
- Re-analysis as a two level model (97m in 48 States)



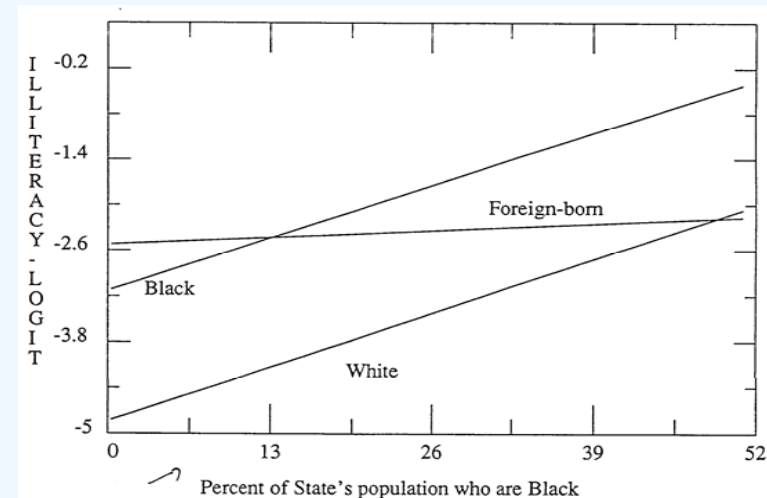
Does this vary from State to State?



Who is illiterate? Individual model



Cross-level interactions?



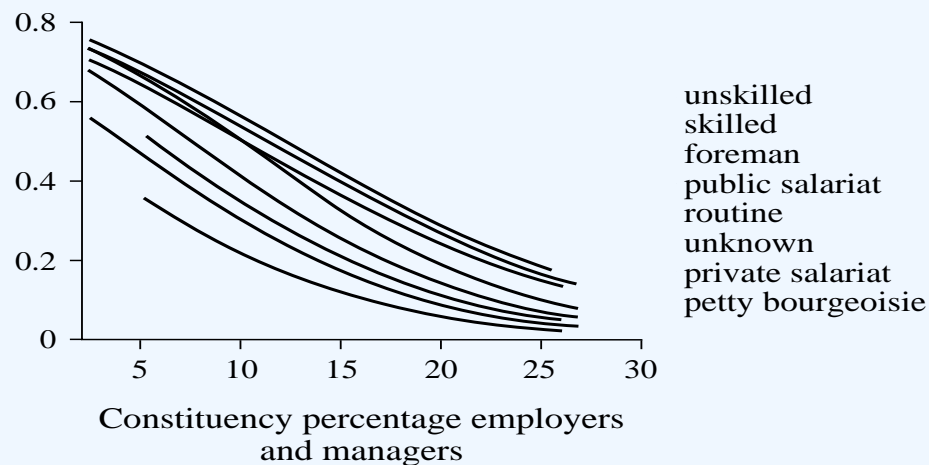
# Modelling complex, between place heterogeneity

- *'where a voter lives is of very little relevance (Rose & McAllister, 1990))*
- 1992 General Election; **Structure:** 2275 respondents nested within 218 constituencies; **Response:** vote Lab/Tory
- Substantial between constituency variance after taking account of individual age, sex, tenure, income, qualifications and class

EG: Probability (Labour) for a middle-aged woman with low qualifications, owner-occupier, unskilled working class, a 'middle' annual income, ranges from 0.22 in Nottingham East to 0.70 in Renfrew West.

## Cross level- interactions between individual class and 'place' elite class

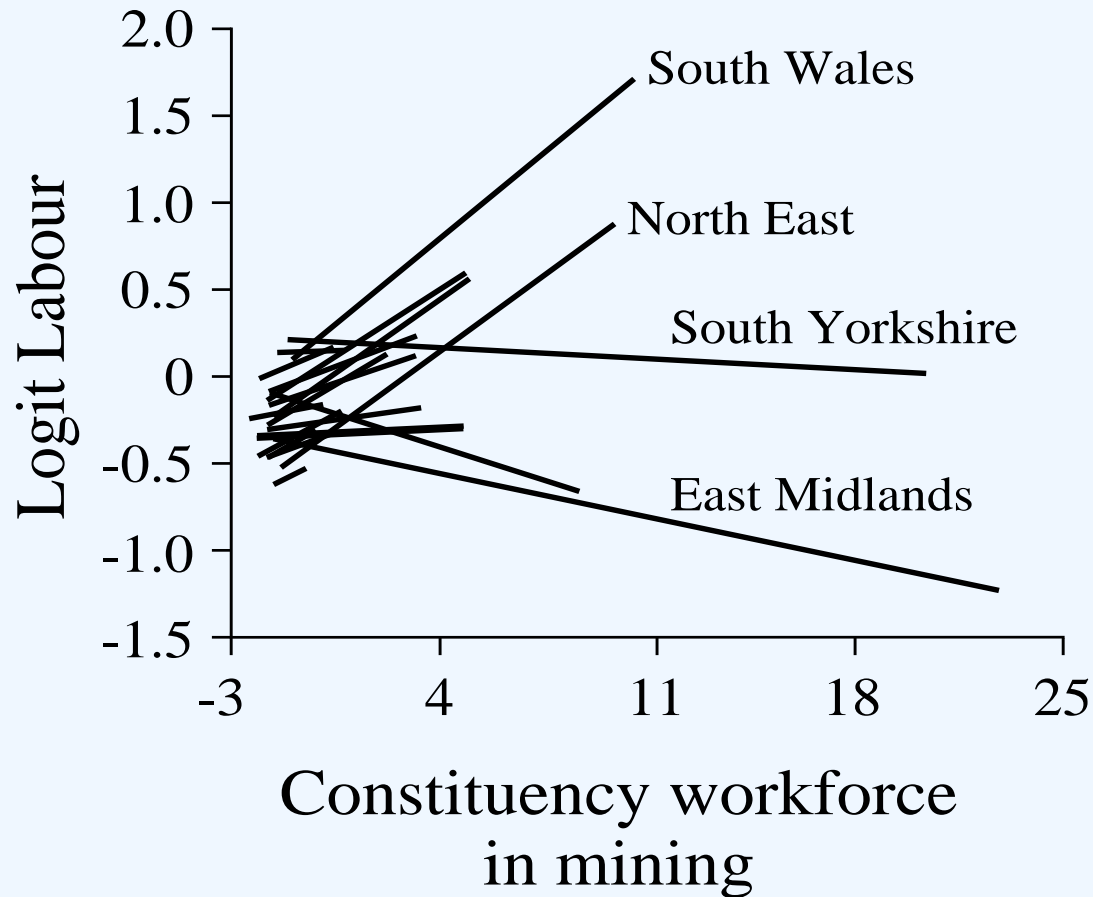
Prob (Labour)



# Heterogeneity continued: example of varying relations( with RJJ & CP, BJPS)

- **Response:** Voting for labour in 1987
- **Structure:** 3 levels strict hierarchy  
individuals within constituencies within regions
- **Predictors**
  - 1 age, class, tenure, employment status
  - 2 %unemployed, employment change, % in mining in 1981
- **Expectation:** coal mining areas vote for the left
- **Allow:** mining parameters for mining effect(2) to vary over region(3) in a 3-level logistic model

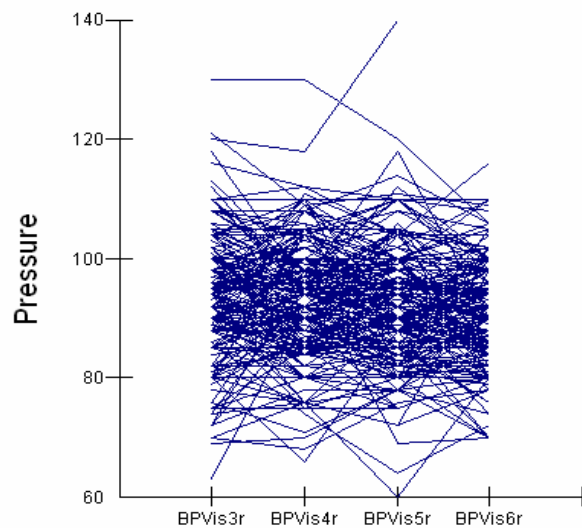
# Varying relations for Labour voting and % mining



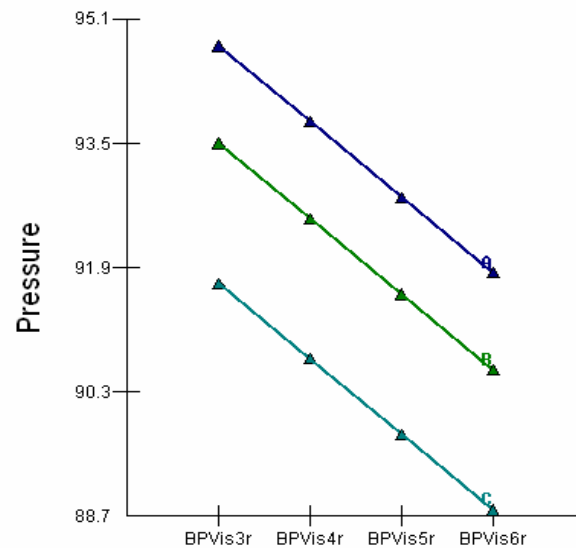
# Clinical trials: heterogeneity & contextuality

- How has diastolic BP changed with 3 different drug treatments; random allocation
- Structure: 4 post-treatment occasions for 287 patients nested in 29 treatment centres; ie longitudinal model

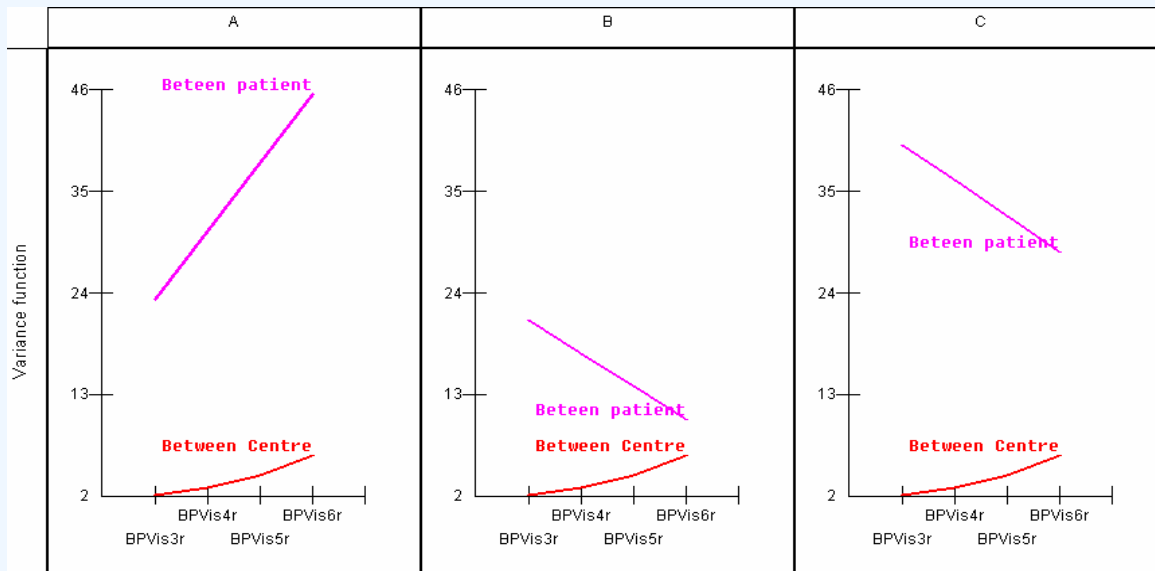
Post-treatment blood pressure for 287 patients



Linear treatment effects for 3 drugs for Person with Baseline BP of 102

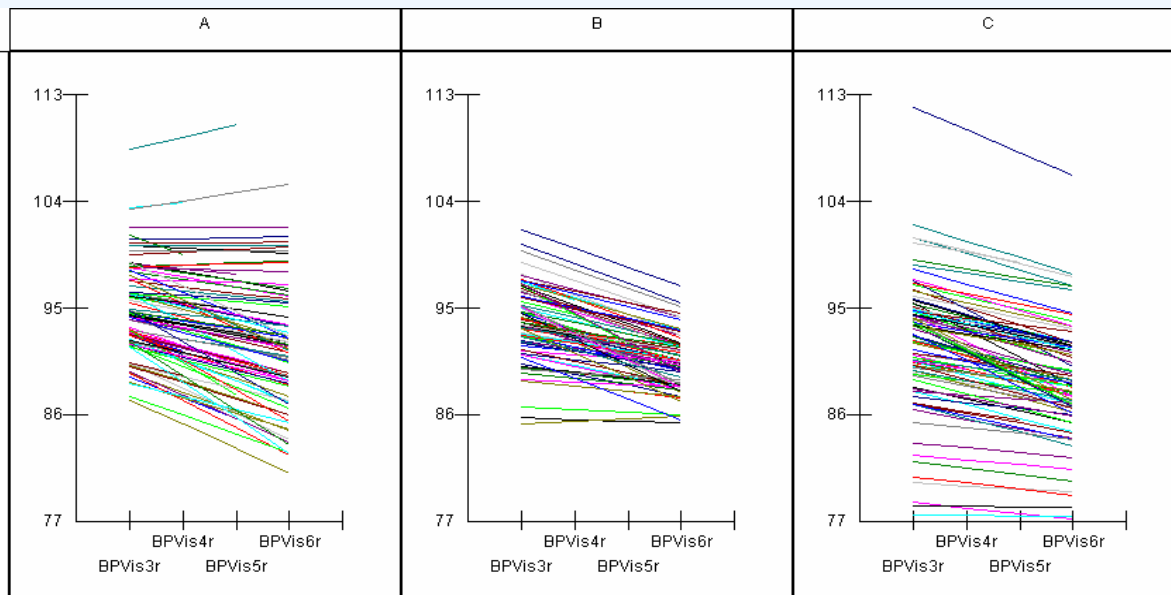


- Baseline average BP is 102
- Raw data and fixed effects
- All three drugs lower BP; no significant differences



## Variance functions

- Variance increases over time between centre
- Drug A: greater volatility around falling trend
- Drug B: most consistent and reduction over time
- Drug C: less consistent, but reduction over time



## Modeled patient Trajectories

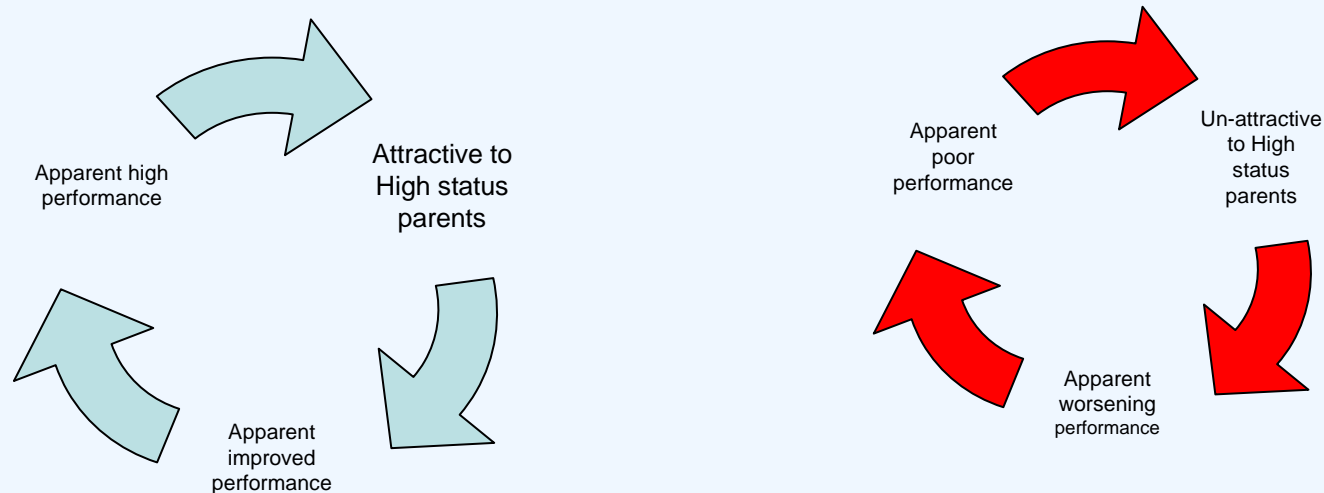
- Drug A: divergence in BP around falling mean
- Drug B: convergence in BP around falling mean
- Drug C: variable impact

# Modelling segregation

*Motivation: are we become a segregated society?*

## Virtuous and Vicious circles

Following 1988 Education Reform Act - choice, league tables, competition **expectation** of INCREASED segregation



Choice



increased polarization in terms of ability

Choice



increased polarization in terms of socio-economic background; poverty; ethnicity etc

# Research Questions

Eligibility for free school meals as a measure of poverty

- Has **school** FSM segregation increased?
- Has **LA** segregation increased?
- Has segregation been **differential** between types of schools and types of LA's

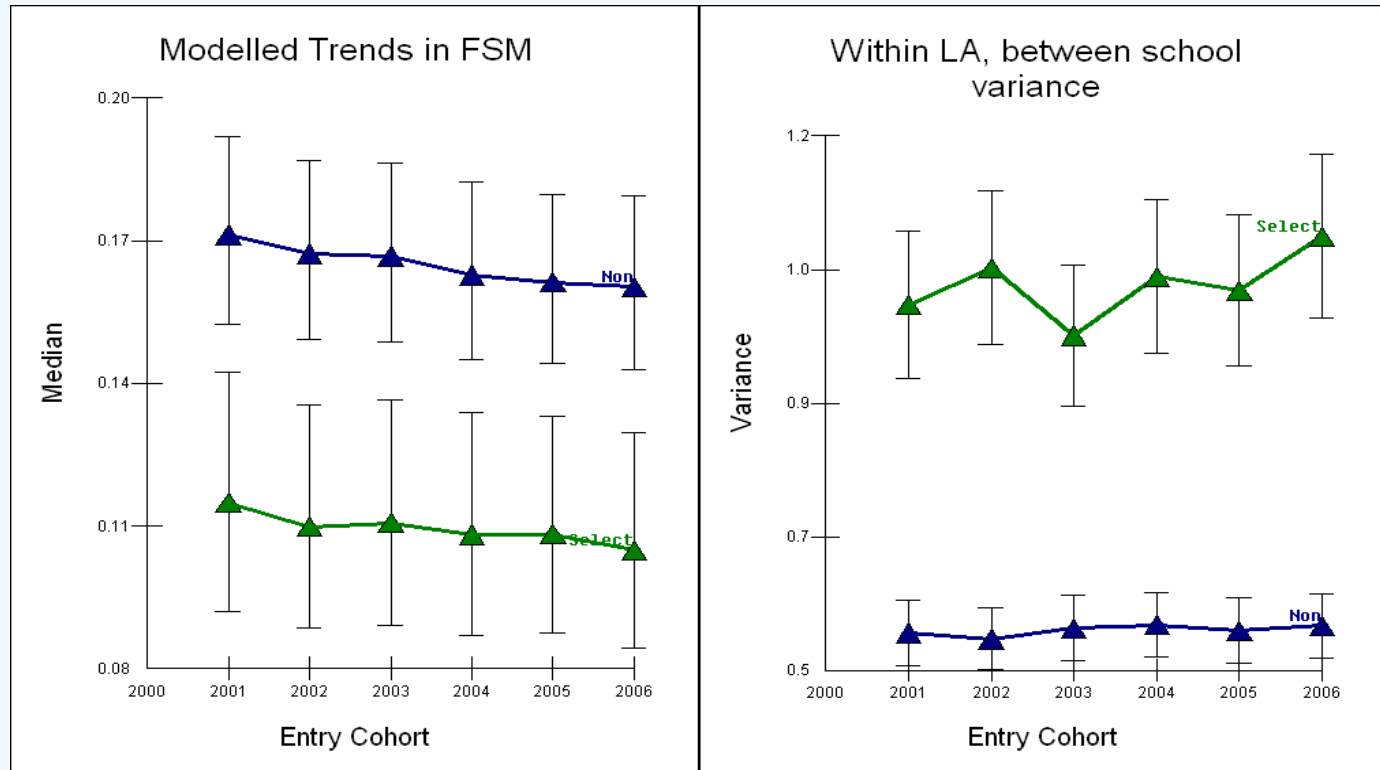
Approach: **model** the variance not calculate an Index; low variance = low segregation

## *FSM: the data*

- **Source:** Pupil Level Annual School Census
- **Outcome:** Proportion of intake Eligible for FSM
- **Intake:** Year 7 of the national curriculum in 2001-2006,

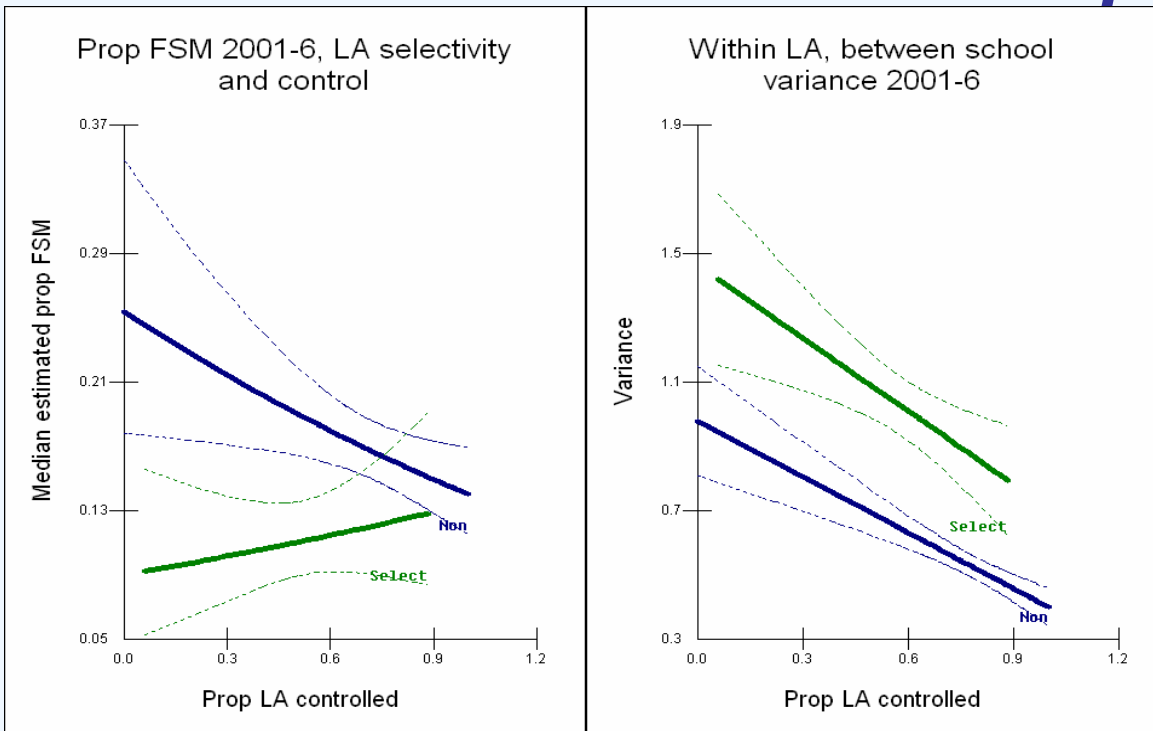
<b>Action</b>	<b>LA's</b>	<b>Schools</b>	<b>Cohorts</b>	<b>Pupils</b>
Complete data from PLASC; 2001-2006	148	5,615	26,178	3,587,459
Omit 'special' schools	148	4,088	20,952	3,536,152
Omit cohorts with less than 20 pupils	148	3,636	20,429	3,535,056
Omit schools without a new intake at aged 11 (ie middle schools) LA loss, eg IOW, Poole	144	3,076	17,695	3,271,010
Omit cohorts with implausible year-on-year differences (see next slide)	144	3,076	17,637	3,261,372

# Are LA's that are selective (Grammar/ Secondary) more segregated than totally Comprehensive systems?



- Pupils going to school in Selecting LA's are less likely to be in poverty
- Slight decline in poverty in both types of area
- Schools in Selecting areas are more segregated
- Slight evidence of an increase of area

# Is there more segregation in areas that are **Selective** and where less schools are under LA control in terms of admission policies?



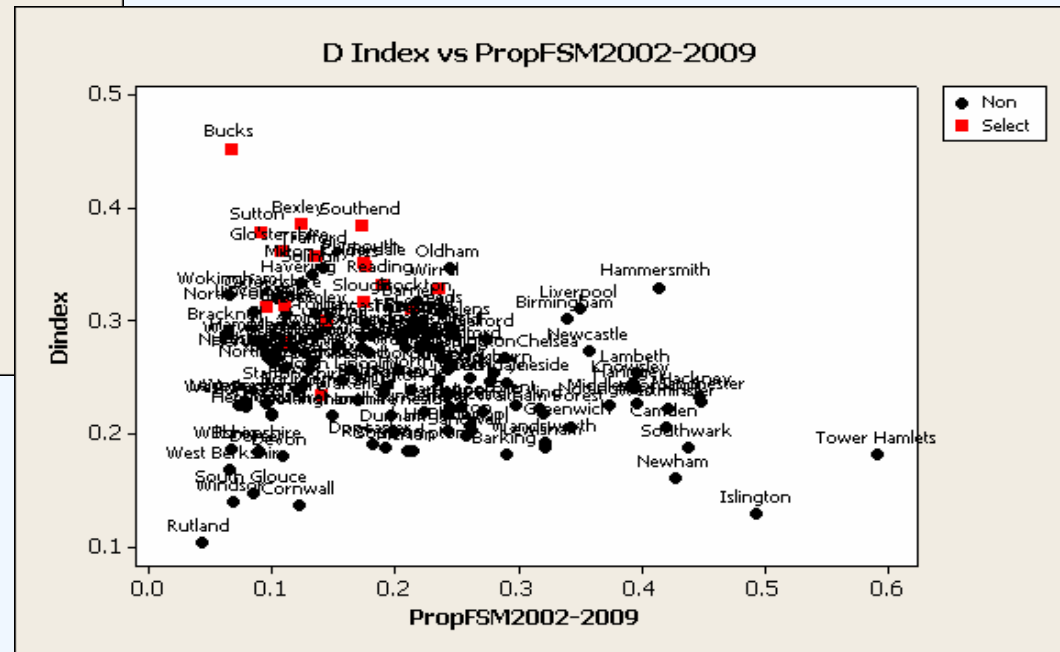
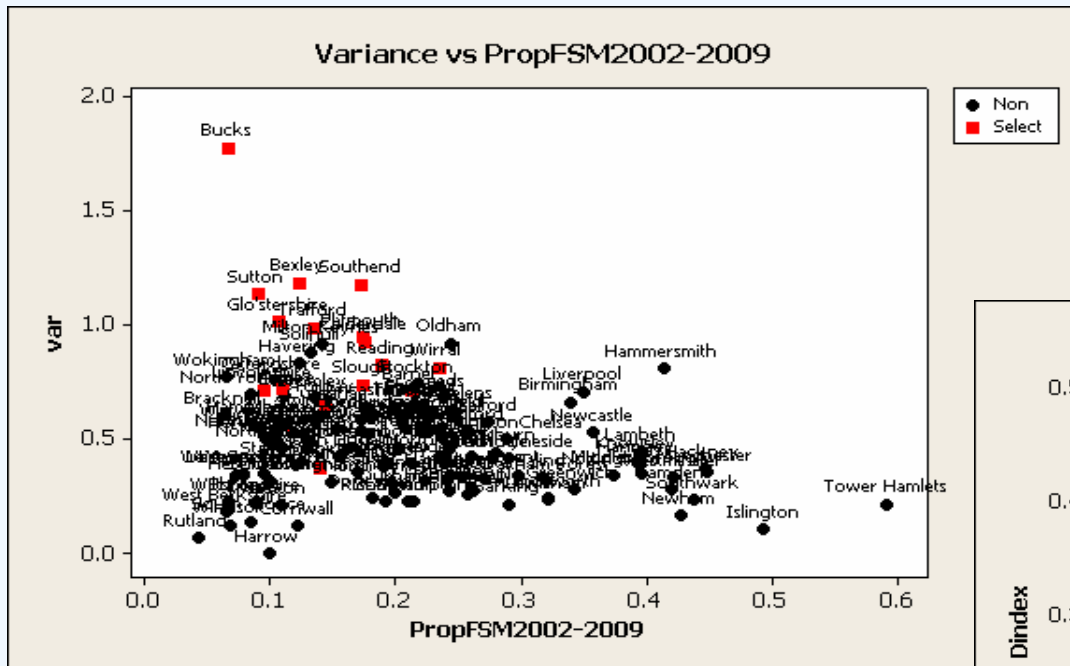
Variance function for Selective/Non-selective, **structured** by the proportion of pupils in an LA who go to Community or Voluntary Controlled schools (contra Voluntary Aided, Foundation, CTC's, Academies)

- Pupils going to school in Non-Selecting LA's with low LA control are more likely to be in poverty

- Schools in Selecting areas are more segregated
- Segregation decreases with greater LA control for both types of LA

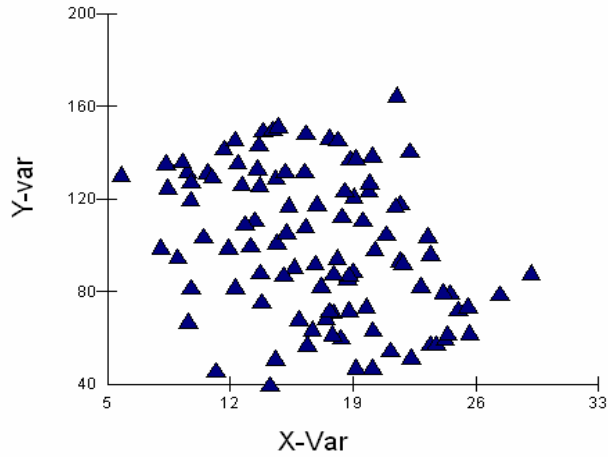
# Which of England's LA's have the most segregated school system?

Modelling 4.2million pupils; 144 Means and Variances and D Index

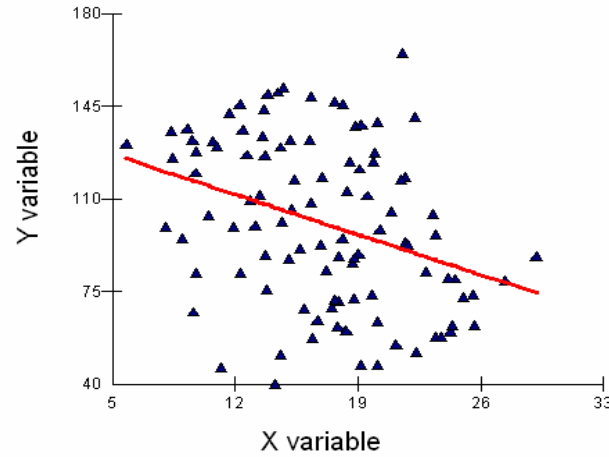


# Why should we use multilevel models?

What is the relation between Y and X?

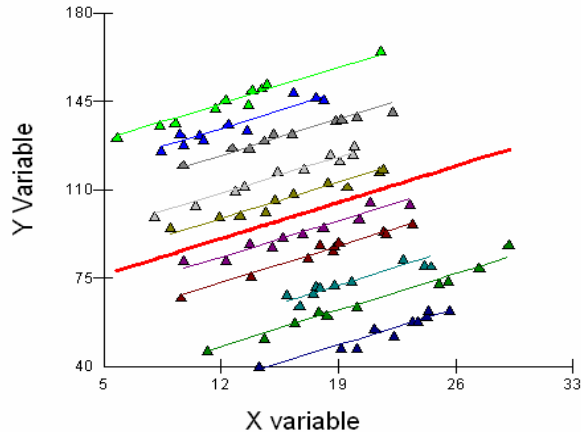


Negative: According to simple regression

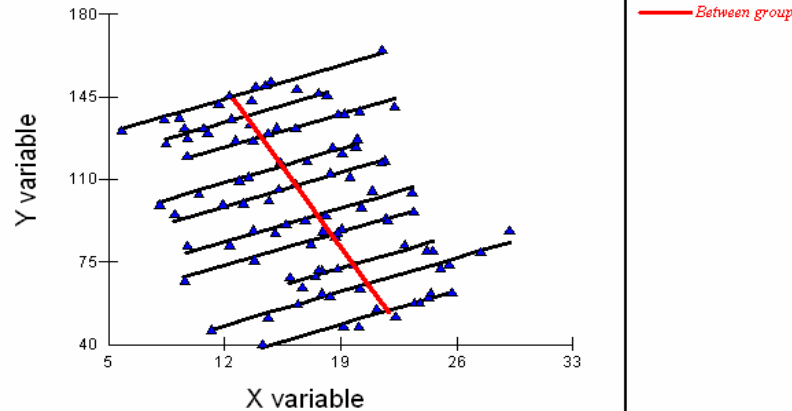


*Sometimes:  
 single level  
 models can be  
 seriously  
 misleading!*

Positive: according to a multilevel model



Positive and negative: within and between relations



# Finally !

“The complexity of the world is not ignored in the pursuit of a single universal equation, but the specific of people and places are retained in a model which still has a capacity for generalisation”

MLwiN UK-Academic; free download

<http://www.cmm.bristol.ac.uk/MLwiN/ordering/ac-uk.shtml>

<http://www.cmm.bristol.ac.uk/>

The screenshot shows a Mozilla Firefox browser window displaying the website for the Centre for Multilevel Modelling (CMM) at the University of Bristol. The browser's address bar shows the URL <http://www.cmm.bristol.ac.uk/>. The website header includes the University of Bristol logo and a search bar. A navigation menu contains links for [home](#), [contact](#), [team](#), [research](#), [learning](#), and [MLwiN](#). The main content area features a heading: "The Centre for Multilevel Modelling (CMM) is a research centre based at the University of Bristol". Below this, a paragraph states: "CMM researchers are drawn from the Graduate School of Education, The School of Veterinary Science and the the School of Geographical Sciences. CMM is part of the The Bristol Institute of Public Affairs (BIPA). Multilevel Modelling is a University Research Theme. We collaborate with a range of researchers working with multilevel models." The page is divided into several sections: "Our research >>" (describing the development of statistical methodology), "Learning >>" (with a sub-section for "On-line Course >>" detailing graduated modules), "Videos and voice-over presentations >>" (listing resources like "Residuals, Measuring Dependency, Covariance and Correlation Matrices"), "News >>" (with a search icon), "MLwiN software >>" (describing the center's own software and providing links for "Upgrade to latest version", "Download for free if you are a UK academic", "MLwiN FAQs", and "MLwiN user Forum"), and "Other CMM Software >>" (listing "Realcom" for developing models, "Realcom-Impute" for handling missing data, and "MLPowSim" for sample size/power calculations). The browser's taskbar at the bottom shows the Windows Start button and several open applications, including "AQmenLecture [Com...", "AQmen", and "Centre for Multilevel ...". The system tray shows the time as 17:33.

<http://www.cmm.bristol.ac.uk/learning-training/course.shtml>

learning and training - Mozilla Firefox

File Edit View History Bookmarks Yahoo! Tools Help

http://www.cmm.bristol.ac.uk/learning-training/course.shtml

Most Visited Getting Started Latest Headlines

Search Web Mail Shopping Personals My Yahoo! News Games Travel Finance Answers Sports Sign In

University of BRISTOL

Centre for Multilevel Modelling

home contact team research learning MLwiN

multilevel models | multilevel modelling support | multilevel modelling software | videos | LEMMA FAQs

left navbar

## LEMMA: Multilevel Modelling online course

Click here to register or log on

See [Course Topics](#)

Our on-line multilevel modelling course is now available: **LEMMA** (The Learning Environment for Multilevel Methodology and Applications) contains a set of graduated modules starting from an introduction to quantitative research progressing to multilevel modelling of continuous and binary data.

Each module has two integrated components:

1. A description of *concepts* and models; and
2. Instructions of *how to carry out analyses* in **MLwiN** and **Stata**
3. You will also find a discussion board and videos and interactive quiz questions throughout the modules to allow you to assess your understanding of the material.

[Log in or register](#) for the course.

Screenshot of on-line course - example - Module 5

Done

start AQmenLecture [Com... AQmen learning and training ... EN 17:36