Controversies in measuring school segregation

Rebecca Allen
Department of Quantitative Social Science
Institute of Education, University of London
rallen@ioe.ac.uk

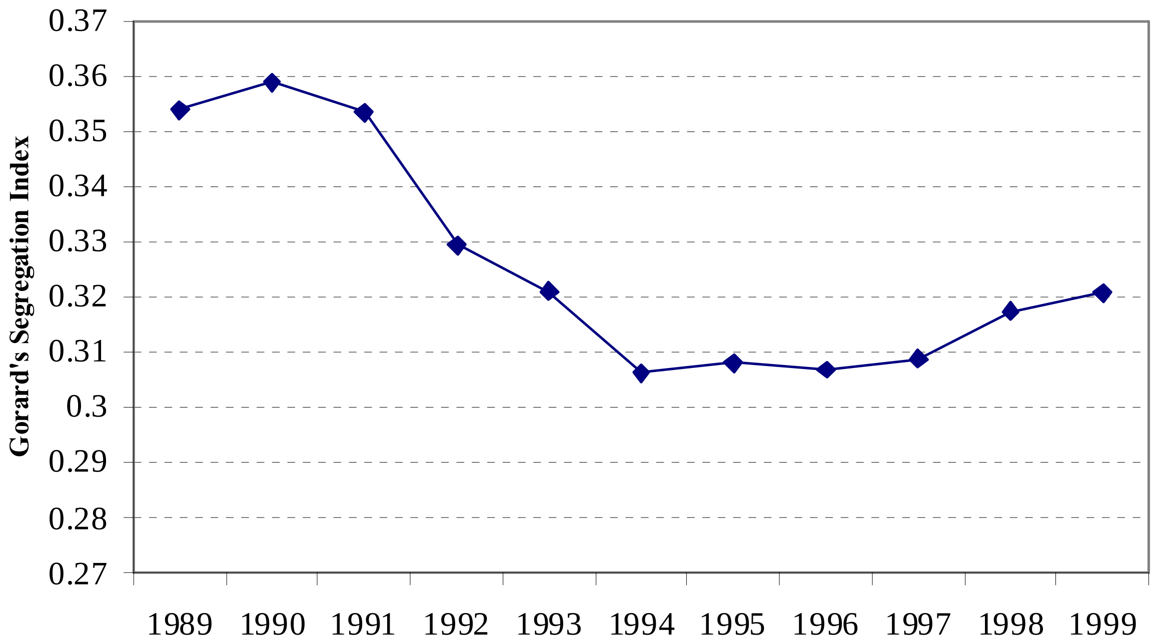
Presentation to NCRM, Oxford
30th June 2008

Introduction

- Segregation means separation, stratification, sorting
  - Unevenness or dissimilarity
  - Isolation or exposure
  - spatial measures: concentration, clustering, centralisation
- Why measure school segregation?
  - Descriptive statistic
  - Effects – segregation as one cause of inequalities
  - Causes – segregation as the outcome of a process
- Methodological developments
  - Progress over the past decade
  - Challenges resulting from availability of pupil-level data
  - Continuing controversies and unexplored avenues

Changes in school segregation – Gorard *et al.* (2003)

- Annual Schools Census (ASC) collected Free School Meals (FSM) take-up from 1989 onwards
- FSM eligibility *and* take-up were recorded from 1993
- Stephen Gorard, John Fitz and Chris Taylor used ASC to record changes in school segregation in England from 1989 onwards



Gorard, S., Taylor, C., & Fitz, J. (2003). Schools, Markets and Choice Policies. London: RoutledgeFalmer.

Gorard's Segregation Index (GS)

GS is an **absolute** index with clear meaning:

'proportion of FSM pupils that would have to exchange schools in order to achieve evenness'

$$GS = 0.5 * \sum_i |fsmshare_i - totalshare_i| \equiv D * (1 - p)$$

(where p is the overall FSM proportion in the area).

The Index of Dissimilarity is a **relative** index with meaning only relative to its fixed bounds of zero and one.

$$D = 0.5 * \sum_i |fsmshare_i - nonfsmshare_i|$$

Gorard, S., Taylor, C., & Fitz, J. (2003). Schools, Markets and Choice Policies. London: RoutledgeFalmer.

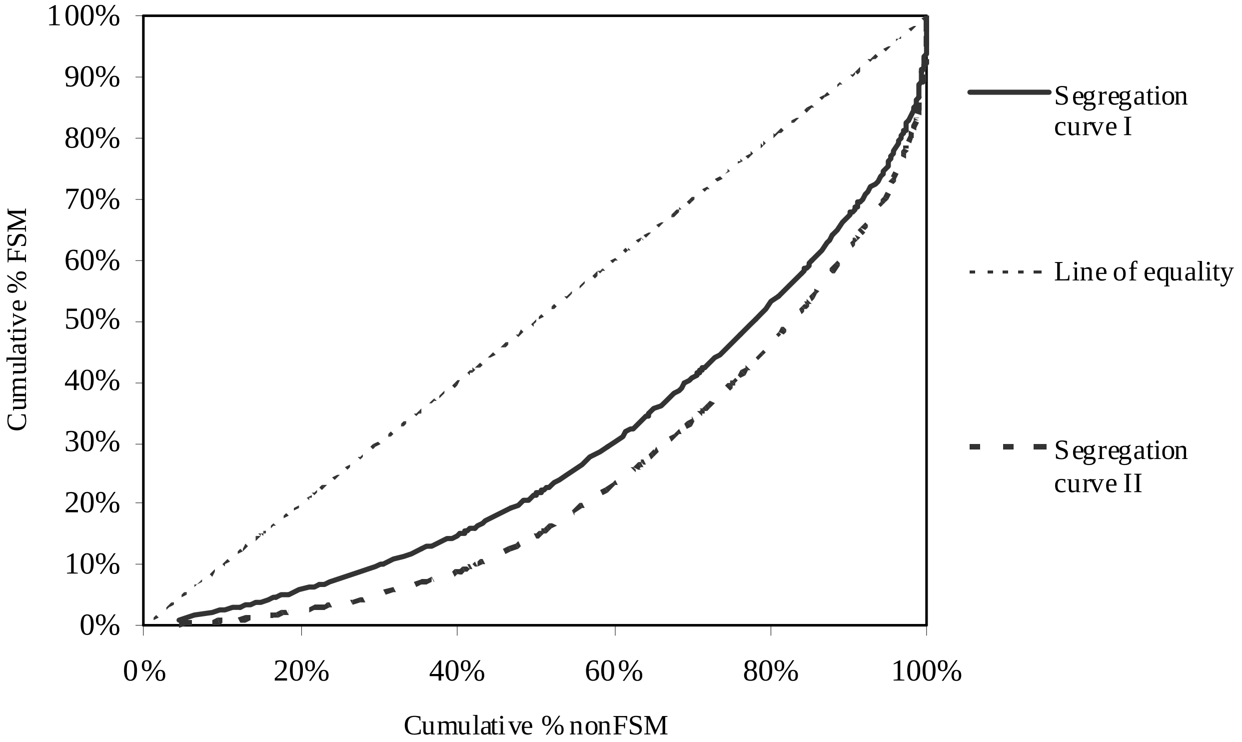
Does it matter which index is used?

1. The magnitude of the fall in segregation between 1989 and 1995 is 10% using GS and 5% using D
2. GS and D disagree on whether segregation actually fell or rose in an LEA between 1989 and 1995 in 35% of cases
3. If we placed LEAs in deciles according to their level of segregation, the 2 indices would disagree about which decile the LEA should be in 63% of the time

Allen, R., & Vignoles, A. (2007). What should an index of school segregation measure? Oxford Review of Education, 33 (5).

Unevenness as a segregation curve

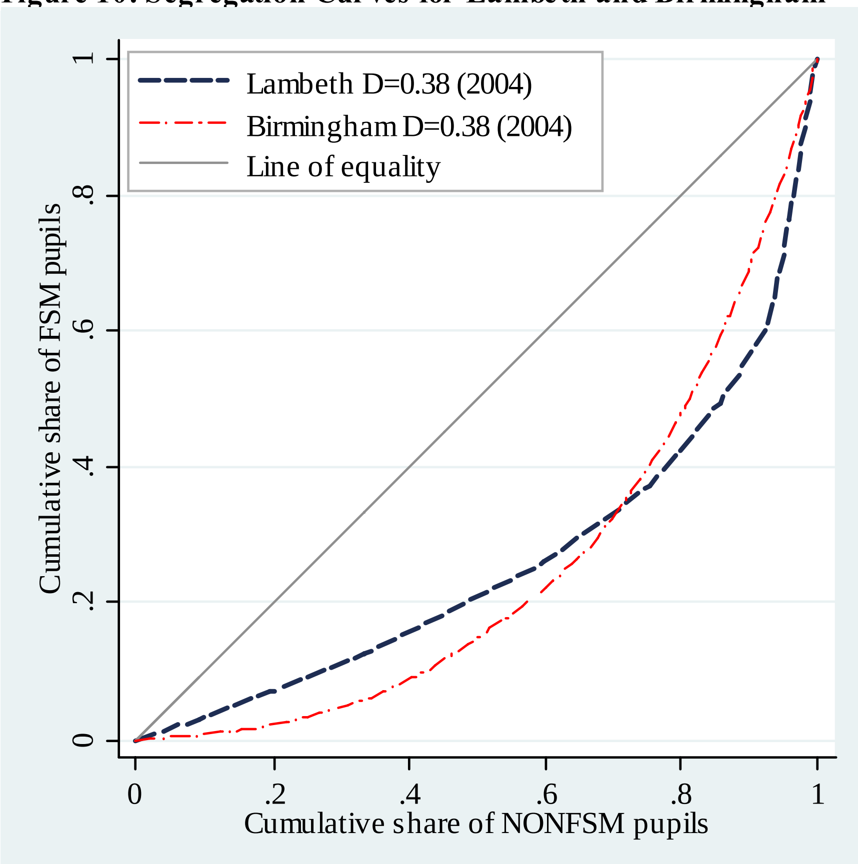
- Segregation curve plots the share of FSM pupils at each school against the share of NONFSM pupils
- Where curves do not cross we can identify whether one distribution of pupils is more uneven than another



Allen, R., & Vignoles, A. (2007). What should an index of school segregation measure? Oxford Review of Education, 33 (5).

Can we distinguish between different patterns of segregation?

Figure 10: Segregation Curves for Lambeth and Birmingham



- Same level of segregation but very different distributions of pupils across schools
- Segregation skew = $\log(O_{0.1}(x)/O_{0.9}(x))$
- Birmingham has concentrations of advantaged schools (skew = + 0.22)
- Lambeth has concentrations of disadvantaged schools (skew = - 0.20)

Allen, R., & Vignoles, A. (2007). What should an index of school segregation measure? Oxford Review of Education, 33 (5).

The desirability of fixed upper and lower bounds

- GS is not bounded by 0 and 1
- The upper bound is 1-p, i.e. GS can never display a value above 1-p

- Buckinghamshire: GS = 0.48; p = 6%; max possible value of GS = 0.94
- Tower Hamlets: GS = 0.11; p = 60%; max possible value of GS = 0.40

Allen, R., & Vignoles, A. (2007). What should an index of school segregation measure? Oxford Review of Education, 33 (5).

Non-symmetry of the index makes interpretation of changes difficult

- The value of FSM segregation is not the same as the value of NONFSM segregation using GS
- GS is capable of showing that FSM segregation is rising and NONFSM segregation is falling simultaneously
- Poole 1999-2004: GS_{FSM} *rose* by 10%; GS_{NONFSM} *fell* by 27%

Allen, R., & Vignoles, A. (2007). What should an index of school segregation measure? Oxford Review of Education, 33 (5).

Properties of GS – Compositional Variance

- What happens to GS when a set of NONFSM pupils 'switch' their status and become FSM pupils?
- Gorard claims GS is *'invariant to the change in scale from 1992 to 1993 in a way that other indices are not'*
- If there is a constant proportion increase in FSM, the most deprived schools in an area suffer disproportionately from the fall in NONFSM pupils

Allen, R., & Vignoles, A. (2007). What should an index of school segregation measure? Oxford Review of Education, 33 (5).

Implications of pupils arriving and leaving the area

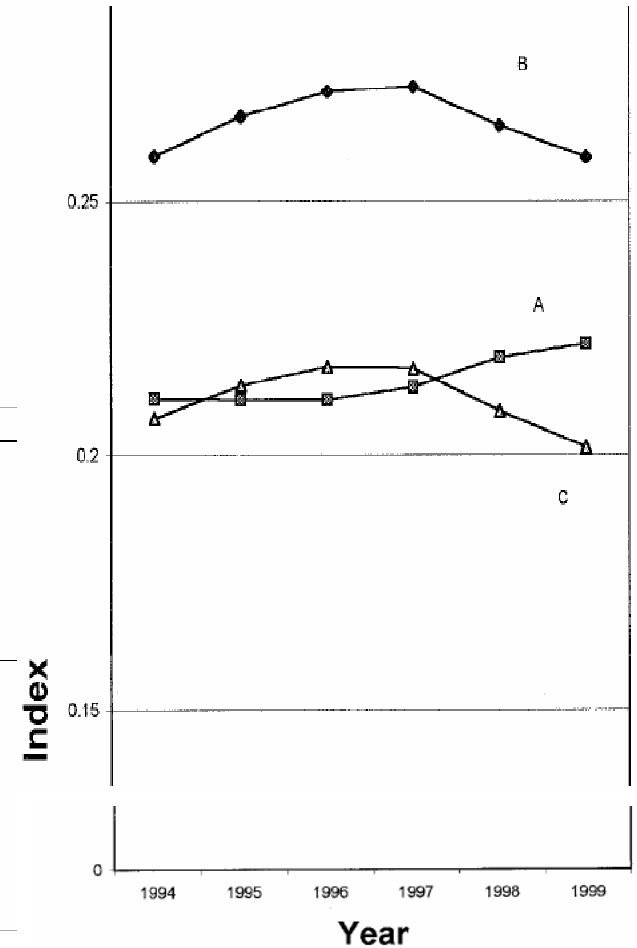
- Is compositional invariance really a desirable property?
- A large, but unresolved, literature exists on decomposing changes in the overall margin from other changes in segregation (Blackburn, Watts etc…)
- Implications for interpretation of longitudinal and cross-section situations
- Separate specific issue regarding instability of FSM characteristic over time

Allen, R., & Vignoles, A. (2007). What should an index of school segregation measure? Oxford Review of Education, 33 (5).

Segregation as isolation/exposure – Noden (2000)

- Isolation (I) = mean exposure of FSM pupils to FSM pupils

	Exposure to FSM pupils	Exposure to NONFSM pupils
FSM pupils	$\sum_s \frac{fsm}{FSM}\frac{fsm}{n}$ No segregation = $\frac{fsm}{n}$ Complete segregation = 1	$\sum_s \frac{fsm}{FSM}\frac{nonfsm}{n}$ No segregation = $\frac{nonfsm}{n}$ Complete segregation = 0
NONFSM pupils	$\sum_s \frac{nonfsm}{NONFSM}\frac{fsm}{n}$ No segregation = $\frac{fsm}{n}$ Complete segregation = 0	$\sum_s \frac{nonfsm}{NONFSM}\frac{nonfsm}{n}$ No segregation = $\frac{nonfsm}{n}$ Complete segregation = 1



Noden, P. (2000). Rediscovering the impact of marketisation: dimensions of social segregation in England's secondary schools, 1994 1999. British Journal of Sociology of Education, 21 (3), 371 390.

Dealing with sensitivity of FSM to the economic cycle

- One solution is to find a counterfactual to school segregation in the same time period
- How does current school segregation compare to current residential segregation (by wards) of the same pupils? (Burgess et al., 2007)
- How does current school segregation compare to a counterfactual simulation where all pupils are allocated to schools strictly on the basis of proximity? (Allen, 2007)

Estimating systematic segregation levels

- Allocation process with systematic tendency to assign groups *fsm* and *nonfsm* differently

$$\exists j: \quad P(unit = j \mid c = 1) \neq P(unit = j \mid c = 0)$$

- We cannot observe conditional probabilities, so we estimate them, but there are differences between observed D and D_{pop} (D for the underlying conditional probabilities rules)

$$\hat{P}(unit = j \mid c = 1) = \frac{\hat{P}(unit = j, c = 1)}{\hat{P}(c = 1)} = \frac{p_{jk} n_{jk}}{P_k N_k}$$

$$\hat{P}(unit = j \mid c = 0) = \frac{\hat{P}(unit = j, c = 0)}{\hat{P}(c = 0)} = \frac{(1 - p_{jk}) n_{jk}}{(1 - P_k) N_k}$$

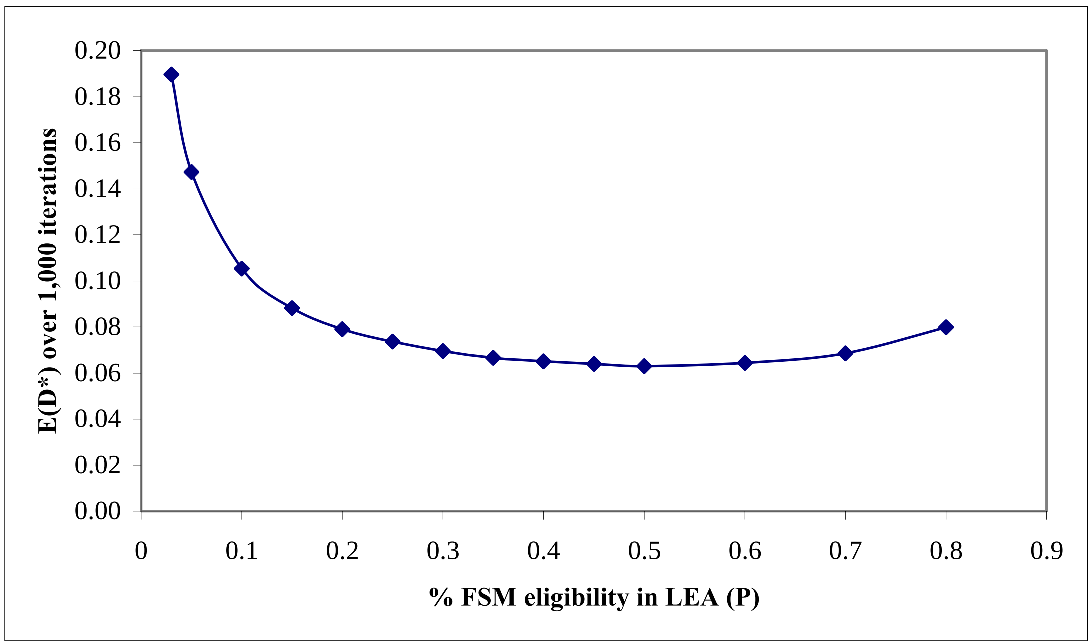
$$\hat{D}_k = \frac{1}{2} \sum_{j=1}^{J} \left| \hat{P}(unit = j \mid c = 1) - \hat{P}(unit = j \mid c = 0) \right|$$

$$\text{"Small Unit Bias"} = E\left[\hat{D}\right] - D_{pop}$$

Allen, R., Burgess, S. and Windmeijer, F. (2008). Correcting segregation indices for small unit bias. Unpublished Working Paper, CMPO, University of Bristol.

Small unit bias (1)

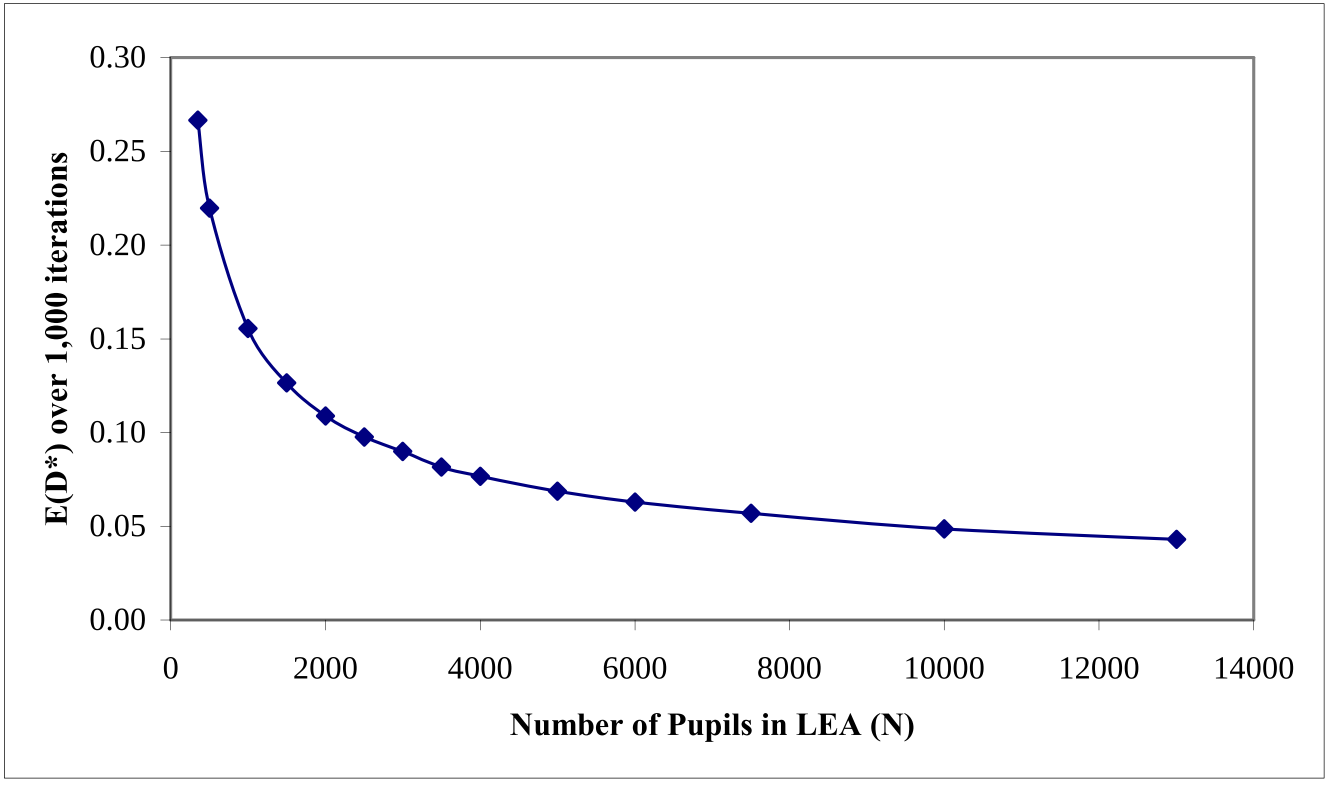
- How much segregation is there under random allocation (our null)?
- The value of D* (D under random allocation) depends on the ‘margins’:
  - P, the proportion FSM eligibility in the LEA
  - N, the number of pupils in the LEA
  - C, the number of schools in the LEA
- The graph shows E(D*) for a fictional LEA with 3,000 pupils, 20 schools, FSM eligibility varies



Allen, R. (2005). The Stratification Consequences of Proximity as the Sole Rule for Allocating Pupils to Secondary Schools in England. Unpublished MRes Dissertation, Institute of Education.

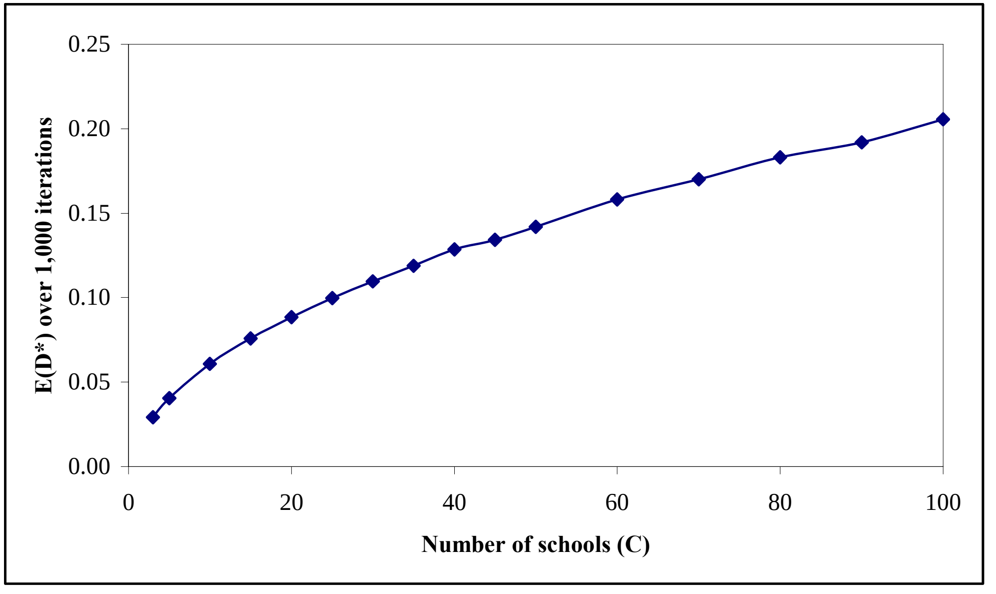
Small unit bias (2)

- The graph shows E(D*) for a fictional LEA with 20 schools, 15% FSM eligibility, number of pupils varies



Allen, R. (2005). The Stratification Consequences of Proximity as the Sole Rule for Allocating Pupils to Secondary Schools in England. Unpublished MRes Dissertation, Institute of Education.

Small unit bias (3)



- The graph shows E(D*) for a fictional LEA with 3,000 pupils, 15% FSM eligibility, number of schools varies

	Mean	S.D.	Min	Max
Number of pupils in LEA (N)	3,137	2,508	361	13,157
Proportion FSM eligibility in LEA (P)	15.6%	9.9%	3.1%	64.4%
Number of schools in LEA (C)	20.8	17.0	3	101
LEA average number of pupils in cohort per school (N/C)	151.0	25.3	101.3	231.6

Allen, R. (2005). The Stratification Consequences of Proximity as the Sole Rule for Allocating Pupils to Secondary Schools in England. Unpublished MRes Dissertation, Institute of Education.

Small unit bias correction (1)

- We infer the size of the small unit bias for any combination of N_k, P_k, J and D_{pop} by using the actual allocation of individuals to units as a set of estimated conditional allocation probabilities.
- We bootstrap the data by making repeated draws of size N_k, with expected minority group proportion P_k and expected unit sizes n_{jk}.

$$D_{pop} < \hat{D} < E\left[\hat{D}_{boot}\right]$$

We start with an upward bias

$$D_{pop} \approx \hat{D}_{abw} = 2\hat{D} - \hat{D}_{boot}$$

Our formula for small bias correction

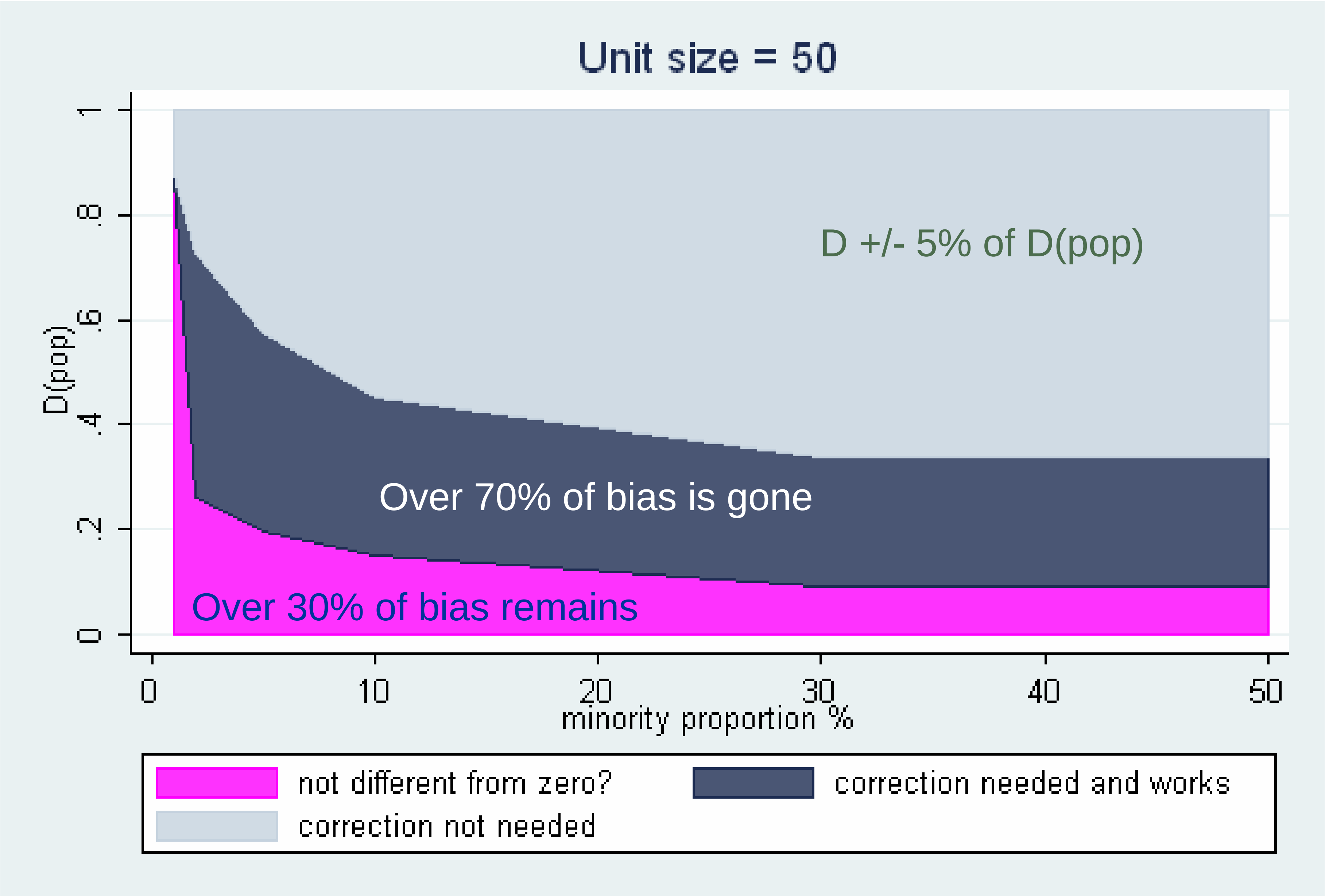
$$E\left[\hat{D}_{boot}\right] - E\left[\hat{D}\right] < E\left[\hat{D}\right] - D_{pop}$$

$$D_{pop} < E\left[\hat{D}_{abw}\right]$$

D_{abw} still has upward bias (but less than before)

Allen, R., Burgess, S. and Windmeijer, F. (2008). Correcting segregation indices for small unit bias. Unpublished Working Paper, CMPO, University of Bristol.

Small unit bias correction (2)



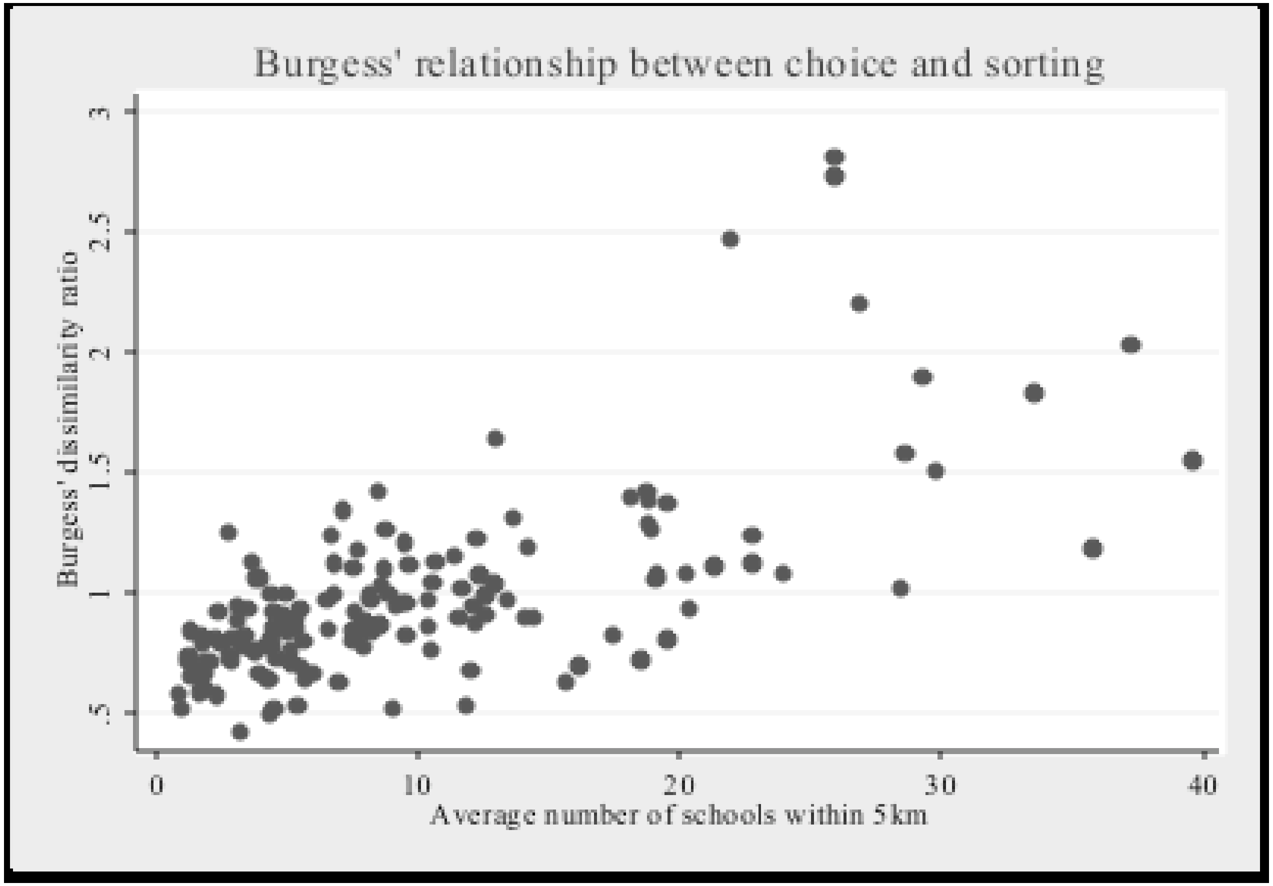
Is school choice associated with higher levels of post-residential sorting?

- Burgess *et al.* (2007) use cross-sectional data (pupils who were 11 in 2003/4) to attempt to establish a causal relationship between school choice and post-residential school segregation. These are the measures they use:
  - School choice: the LEA average number of competitor schools with a 10 minute drive-time zone (*choice*)
  - Post-residential segregation: a ratio of D for schools over D for wards in an LEA (*Dratio*)
- For segregation by disadvantage, measured by FSM eligibility, these are their findings (R-sq rises to 0.45 for only non-selective LEAs):

$$Dratio_i = 0.743_{(19.82)} + 0.034_{(9.58)} choice_i + \varepsilon_i \quad \left[R^2 = 0.39\right]$$

Burgess, S. et al. (2007). The Impact of School Choice on Sorting by Ability and Socio economic Factors in English Secondary Education. In L. Woessmann & P. Peterson, Schools and the Equal Opportunity Problem. MIT Press

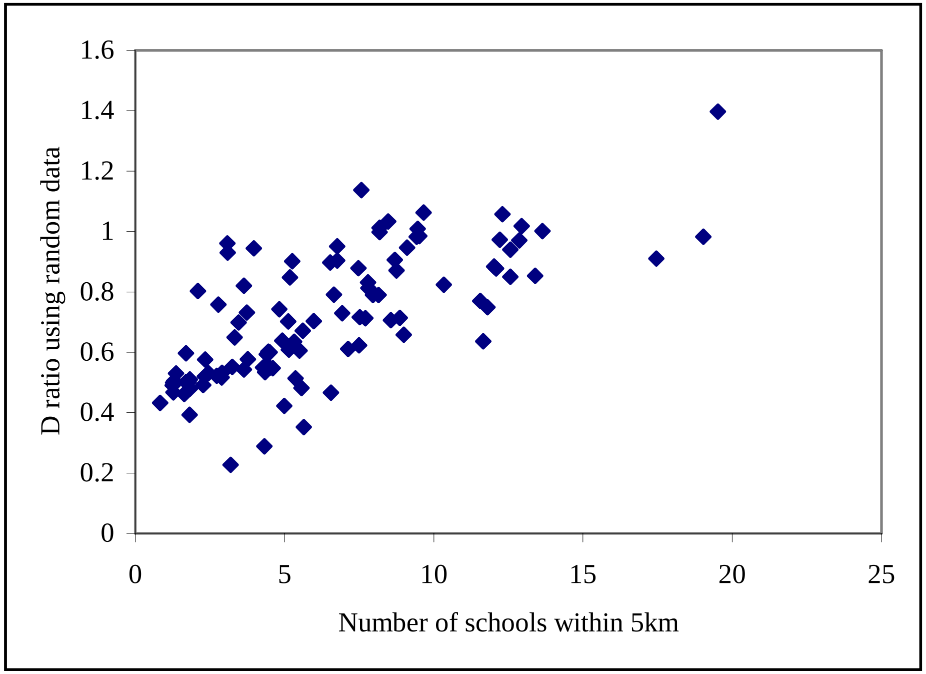
High population density LEAs have a higher school/residential segregation ratio



Note: this data is illustrative and not from Burgess et al. (2007)

Burgess, S. et al. (2007). The Impact of School Choice on Sorting by Ability and Socio economic Factors in English Secondary Education. In L. Woessmann & P. Peterson, Schools and the Equal Opportunity Problem. MIT Press

But the same relationship holds in randomly generated data…



- Taking each LEA in turn, pupils are randomly assigned FSM or NONFSM status, holding the LEA's FSM proportion constant. Then school and residential segregation are re-calculated.
- A ward cohort (average 85 pupils) is a smaller sub-unit than a school (average 150 pupils)
- In London, a ward is larger than average and a school is smaller than average so the school vs. ward size differential is smaller

Allen, R. (2005). The Stratification Consequences of Proximity as the Sole Rule for Allocating Pupils to Secondary Schools in England. Unpublished MRes Dissertation, Institute of Education.

Modelling approaches to segregation

- Why impose statistical models on the data?
  - Model based approach assumes an underlying process such that a suitable function of the parameters measures 'segregation'. This contrasts to traditional index construction that uses definitions based upon observed proportions.
  - Confidence intervals on segregation measures are established via the statistical model and are intended to reflect the uncertainty by which social processes cause segregation.
  - Some statistical models allow us to 'model' causes of segregation more explicitly (and in a single stage) compared to an indices approach.

Goldstein and Noden (2003)

- Intake 'cohorts' of children are nested within schools, schools are nested within areas
- Does underlying variation in the FSM proportion between schools and between areas change over time?
- Multilevel model:

$$\log[\pi_{jk}/(1-\pi_{jk})] = \text{logit}(\pi_{jk}) = (X\beta)_{jk} + v_k + u_{jk}$$

$$p_{jk} \sim bin(n_{jk}, \pi_{jk})$$

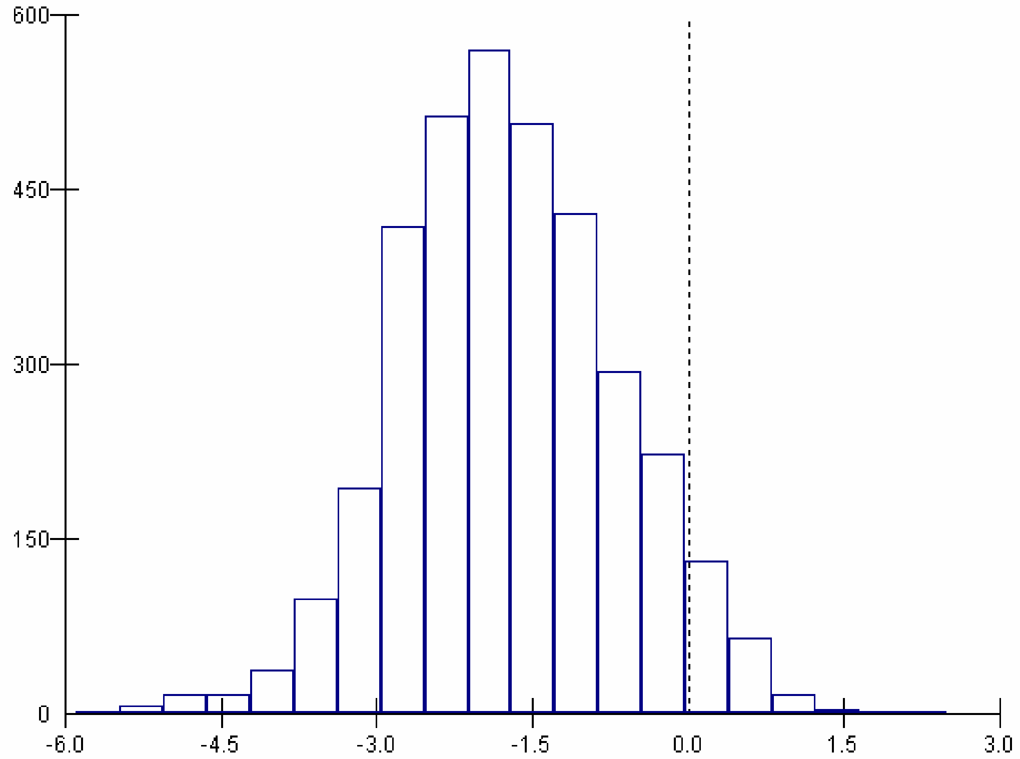
$$v_k \sim N(0, \sigma_v^2), \quad u_{jk} \sim N(0, \sigma_u^2)$$

- P_{jk} is observed proportion at any one time in j-th school in k-th area, is underlying probability which is decomposed into a school effect (u_{jk}) and an area effect (v_k). Interest lies in the variation between schools (σ^2_u) and areas (σ^2_v). If variation Normal then this is a complete summary of the data and avoids arbitrary index definitions.

Goldstein, H., & Noden, P. (2003). Modelling Social Segregation. Oxford Review of Education, 29, 225 237.

Observed FSM Proportions

- Distribution of observed logit(Π_{jk}) for all secondary schools in 1997 is normally distributed:



Goldstein, H., & Noden, P. (2003). Modelling Social Segregation. Oxford Review of Education, 29, 225–237.

Variance Estimates

TABLE I. Variance estimates (standard errors) for each year

Year	Between school	Between LEA	Total
1994	0.625 (0.016)	0.491 (0.066)	1.116
1995	0.636 (0.016)	0.522 (0.072)	1.158
1996	0.650 (0.016)	0.503 (0.064)	1.153
1997	0.660 (0.017)	0.498 (0.069)	1.158
1998	0.685 (0.017)	0.506 (0.068)	1.191
1999	0.691 (0.017)	0.506 (0.068)	1.197

TABLE II. Fitting a model for years 1994 and 1999

Fixed	Estimate	S.E.
Intercept 1994	−1.553	0.064
Intercept 1999	−1.585	0.066
Random		
σ^2_{u1} (School level variance 1994)	0.636	0.016
ρ_{u12} (Correlation between school residuals for 1994 and 1999)	0.95	
σ^2_{u2} (School level variance 1999)	0.707	0.017
σ^2_{v1} (LEA level variance 1994)	0.490	0.066
ρ_{v12} (Correlation between LEA residuals for 1994 and 1999)	0.98	
σ^2_{v2} (LEA level variance 1999)	0.508	0.069

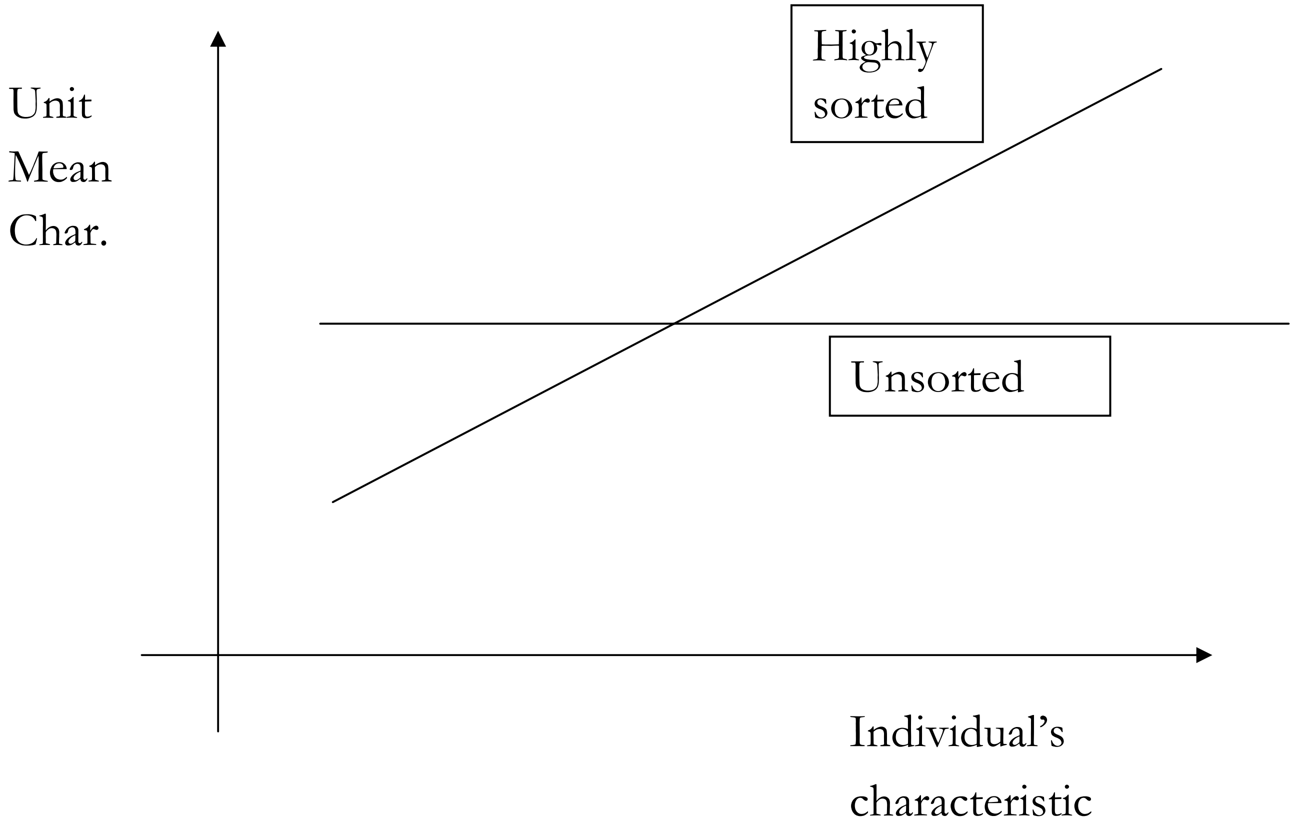
Hypothesis tests: $\sigma^2_{u1} = \sigma^2_{u2}$; $\chi^2_1 = 67.0$ $P < 0.001$. $\sigma^2_{v1} = \sigma^2_{v2}$; $\chi^2_1 = 0.87$ $P < 0.10$.

Goldstein, H., & Noden, P. (2003). Modelling Social Segregation. Oxford Review of Education, 29, 225 237.

From Variance in P to Segregation Measures

- Using model parameters we can derive expected values of any function of underlying school probabilities
- Hutchen's index is
$$-\frac{1}{\sum_j n_j}\sum_j\left(\frac{n_j(1-\pi_j)}{1-\pi}\right)\left(\sqrt{\frac{\pi_j(1-\pi)}{(1-\pi_j)\pi}}-1\right)$$
- Gorard index is
$$\frac{1}{2\pi\sum_j n_j}\sum_j n_j\,|\pi_j-\pi|$$
- These functions can be estimated by simulation from model parameters.

Burgess/Allen/Windmeijer's Matching Model of Pupils to Peer Groups



Burgess, Simon (2006) Presentation to the Royal Statistical Society. 3rd May 2006.

Burgess/Allen/Windmeijer – Set Up

- N individuals indexed by i
- Characterised by a variable, x_i,
- Overall mean of x is $\bar{\bar{x}}$
  and the overall standard deviation is σ.
- Individuals are assigned by a process to S units, indexed by s.
- Mean x in the particular unit s to which individual i assigned is denoted $\bar{x}_s$

Burgess, Simon (2006) Presentation to the Royal Statistical Society. 3rd May 2006.

Burgess/Allen/Windmeijer – Model

- Describe the outcome of the assignment process through the conditional density function:

  $$f\left(\bar{x}_s \mid x_i\right)$$

- Use estimated *f(. | .)* to characterise the degree of sorting.

$$\bar{x}_{s(i)} = \alpha + \beta . x_i + \varepsilon_i$$

- Linear model:

$$\beta = \frac{\sum_i \left(\bar{x}_{s(i)} - \bar{\bar{x}}\right)\left(x_i - \bar{\bar{x}}\right)}{\sigma^2}$$

Burgess, Simon (2006) Presentation to the Royal Statistical Society. 3rd May 2006.

Relation to Segregation Indices

- For dichotomous *x*, β is identical to an index called 'eta-squared'
  - Mean exposure of FSM to FSM pupils *minus* mean exposure of NONFSM to FSM pupils
  - Alternatively, it is the isolation index stretched (standardised) onto a 0-1 scale
- For continuous *x*, β is identical to the square of an index called the Neighbourhood Sorting Index (Jargowsky)
  - Variance partition coefficient = ratio of the between-school variance / total variance in *x*

Advantages of the Framework

- Natural way to introduce covariates:

$$\bar{x}_{s(i)} = \alpha + \beta . x_i + \boldsymbol{\gamma} . \mathbf{W_i} + \varepsilon_i$$

- Often a big issue.
  - e.g. Wilson, Massey and Denton, Jargowsky – segregation in US cities – race or class?
- Flexible way of considering segregation at different parts of the distribution – quantile regression.

Understanding differences in segregation

- Area differences in segregation:

$$\bar{x}_{s(i)} = \mu_a + \beta_a . x_i + \varepsilon_i$$

- But there may be variation within areas. Suppose factor Z_i available at aggregation *r*:

$$\bar{x}_{s(i)} = \mu_a + \beta\left(Z_{r(i)}\right)x_i + g\left(Z_{r(i)}\right) + \varepsilon_i \qquad a(i) > r(i) \geq i$$

- Link economic (or other) model of agents' behaviour directly to equation.

The Future…

- Estimation problems in statistical models of segregation
- ‘Causes’ of segregation via pupil, school and area characteristics
- Usefulness of reductionist ‘models’ of segregation, versus more explicit simulations of uncertainty surrounding the sorting process