5th ESRC Research Methods Festival

2-5 July 2012

St Catherine's College, Oxford

King's College London

University of London

Institute of Psychiatry at the Maudsley

Cross measure calibration methods and how they can enhance our analytical potential

Andrew Pickles, King's College London

Chris Roberts, University of Manchester

Damian Farnell, University of Glamorgan

MRC Medical Research Council

How do you compare over time when we keep changing the instrument?

- Within participant change: Language development of children with Specific Language Impairment – internal calibration and application of SEM methods

- Between cohort change: trends in adolescent mental health – external calibration and application of multiple imputation, regression calibration and SEM methods.

Developmental Trajectories in Specific Language Impairment (SLI)

- SLI is a heterogeneous disorder with a variety of language and related problems
- A number of studies have focused on outcomes, however few have examined developmental language growth patterns and how this may inform the classification (subgrouping) of SLI
- Heterogeneous nature may lead to different developmental trajectories with differing associated symptomatology
- Manchester Language Study – cohort of children in special language schools followed from age 6.

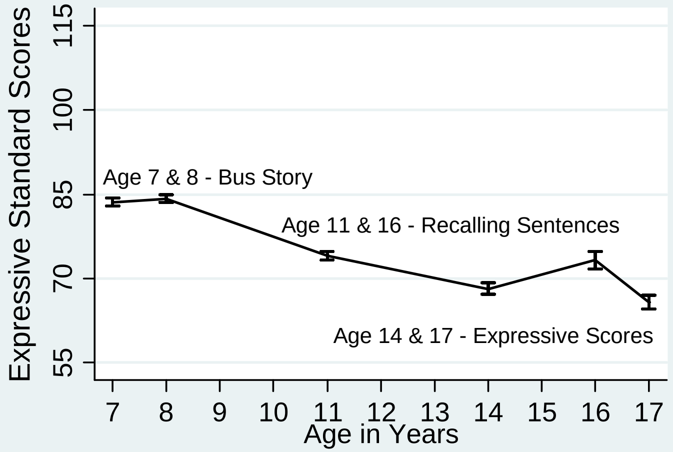
Manchester Language Study

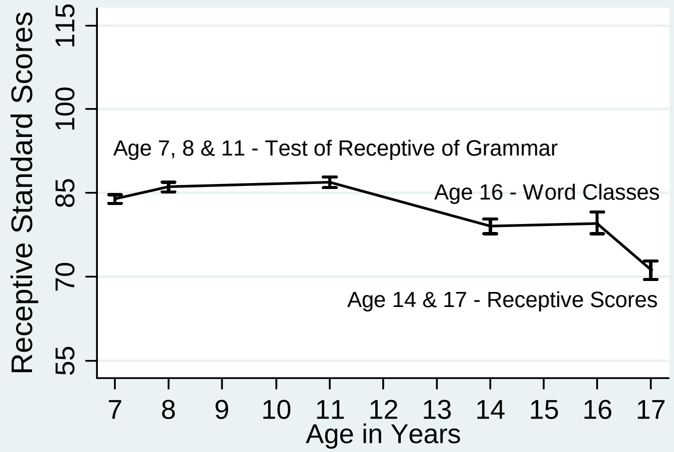
Table 1. Receptive language, expressive language, and nonverbal skills measures for each time point

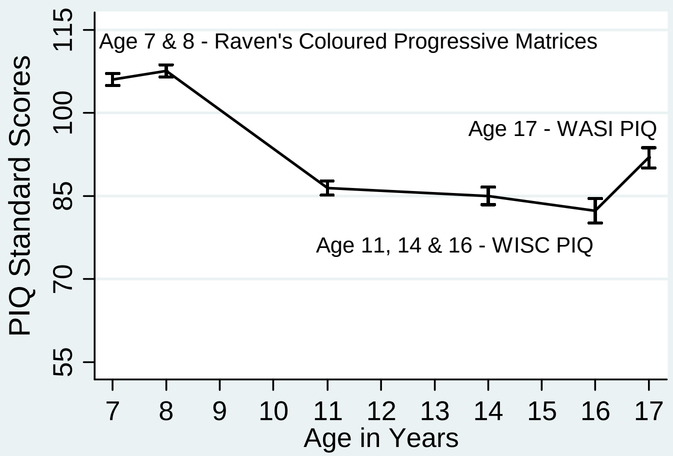
	7	*8*	*11*	*14*	*16*	***17***
Receptive Language	TROG	TROG	TROG	CELF-3 RLC	CELF-R WC	CELF-4 RLI
Expressive Language	Bus Story	Bus Story	CELF-R RC	CELF-3 ELC	CELF-R RC	CELF-4 ELI
Nonverbal Skills	Raven's	Raven's	WISC-III	WISC-III	WISC-III	WASI

Key: ELC = Expressive Language Composite; ELI = Expressive Language Index;
RLC = Receptive Language Composite; RLI = Receptive Language Index;
RC = Recalling Sentences Subtest; WC = Word Classes Subtest

Figure 1. Expressive language, receptive language and PIQ ability (in standard score format) from age 7 to 17 (whole sample means with standard error bars)





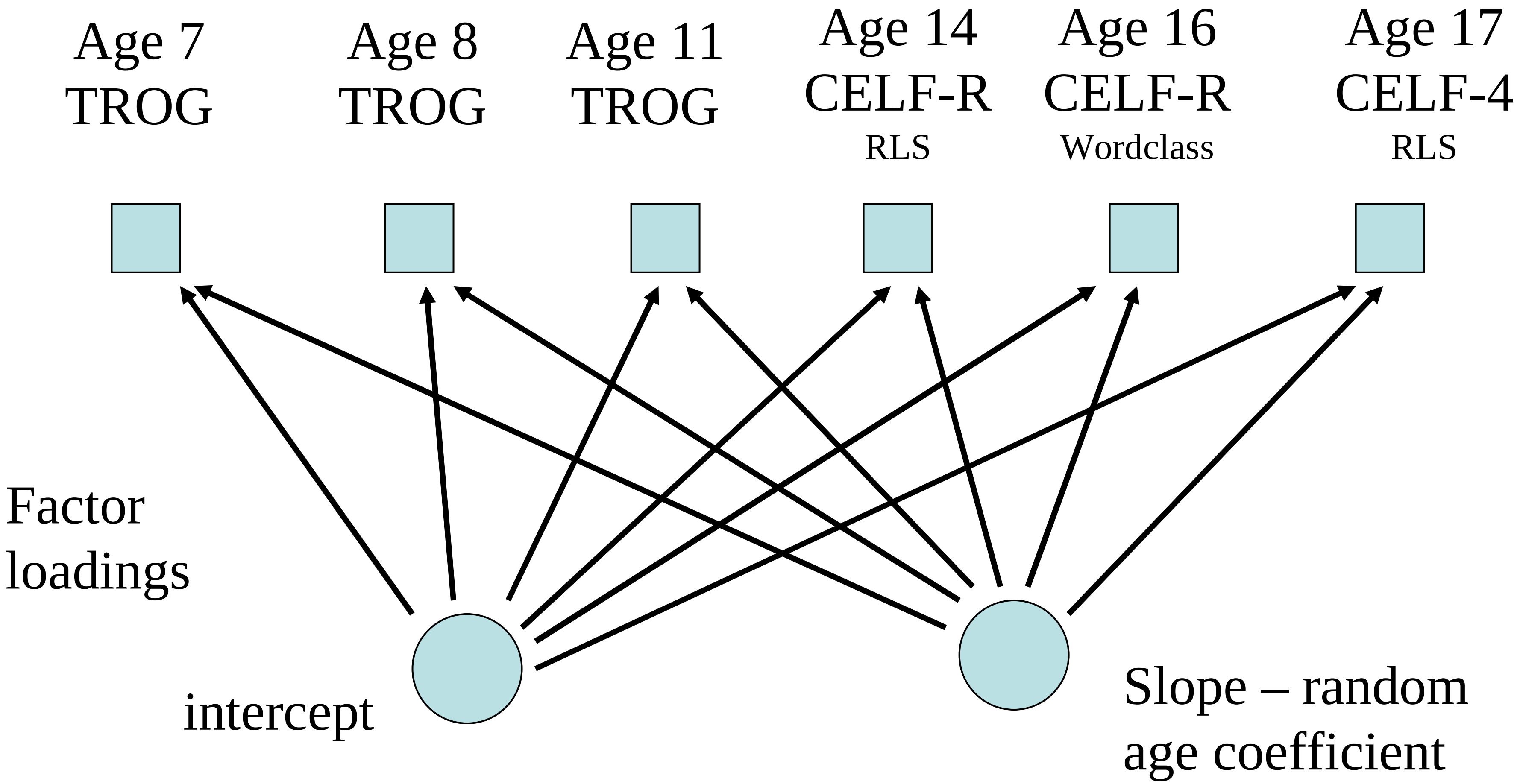


Problems of standard scores

Often span a more modest range than do raw scores (floor and ceiling effects)

Standard scores from different tests are not calibrated against each other – standardization populations are smaller and more different than you might expect.

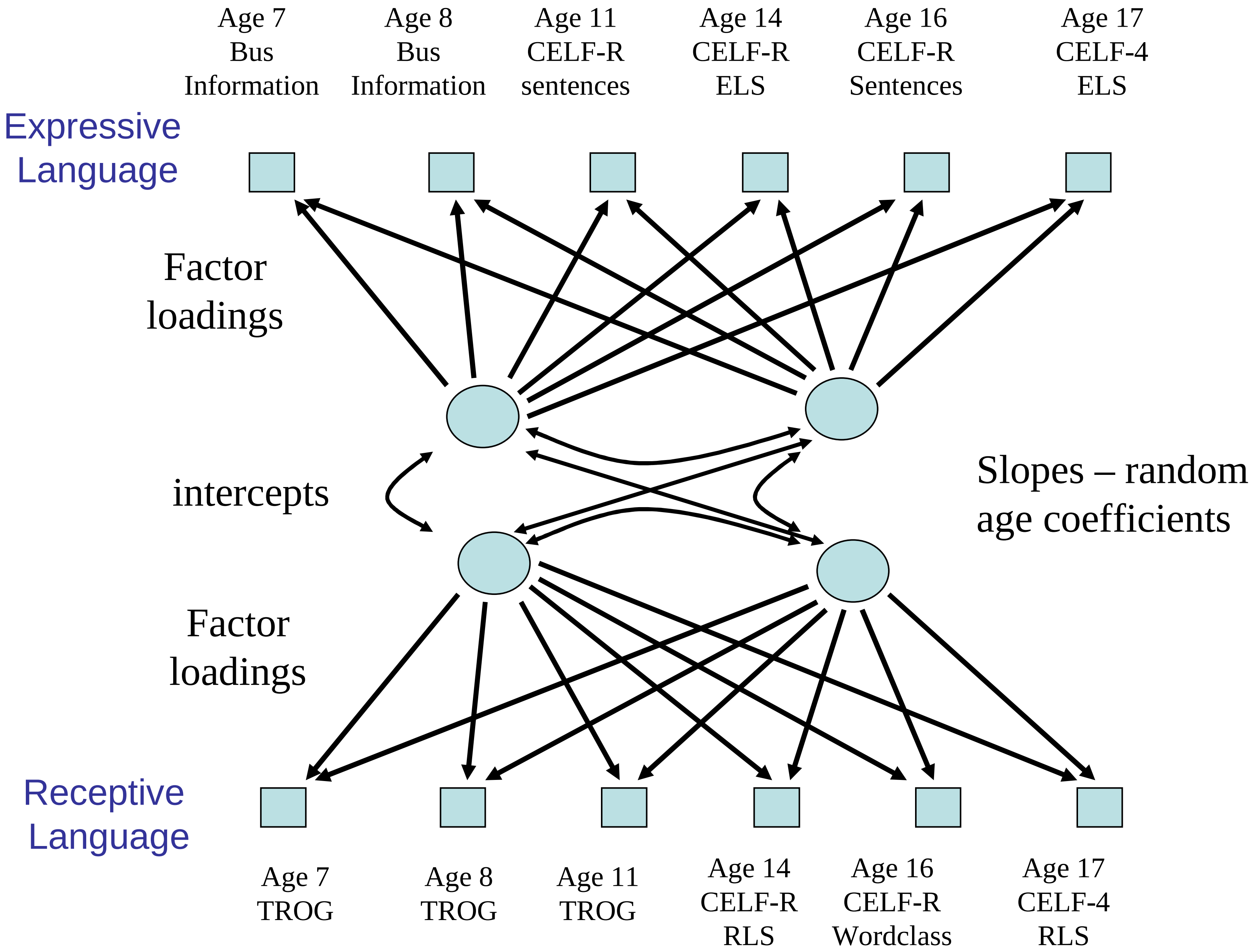
Receptive Language-scaled growth curve



Impose factor loading and error variance constraints where the same measure is being used at different occasions.

Use model estimates to generate standardized scores

”scaled” growth curve



Random coefficient models in GLLAMM

- One covariate multiplies each latent variable,

$$\eta_m^{(l)} z_{m1}^{(l)} \quad (\lambda_{m1}^{(l)} = 1)$$

- e.g. Latent growth curve model for individuals j (level 2) observed at times t_{ij}, $i = 1, \cdots, n_j$ (level 1)

Linear predictor: $\nu_{ij} = \beta_1 + \beta_2 t_{ij} + \eta_{1j}^{(2)} + \eta_{2j}^{(2)} t_{ij}$

β_1, β_2 : mean intercept and slope
$\eta_{1j}^{(2)}$, $\eta_{2j}^{(2)}$: random deviations of unit-specific intercepts and slopes from their means

Generalized random coeff. model in GLLAMM

$$\nu = \mathbf{x}'\boldsymbol{\beta} + \sum_{l=2}^{L}\sum_{m=1}^{M_l} \eta_m^{(l)} \mathbf{z}_m^{(l)\prime} \boldsymbol{\lambda}_m^{(l)}$$

For identification, $\lambda_{m1}^{(l)} = 1$

- Fixed part: $\mathbf{x}'\boldsymbol{\beta}$ as usual
- Random part:
  - $\eta_m^{(l)}$ is mth latent variable at level l, $m = 1, \cdots, M_l$, $l = 2, \cdots, L$
    Can be a factor or a random coefficient
  - $\mathbf{z}_m^{(l)}$ are variables and $\boldsymbol{\lambda}_m^{(l)}$ are parameters
  - Unless regressions for the latent variables are specified, latent variables at different levels are independent whereas latent variables at the same level may be dependent

Discrete latent variables in GLLAMM

- Linear predictor in two-level models:

$$\nu_{ij} \;=\; \mathbf{x}'_{ij}\boldsymbol{\beta} + \sum_{m=1}^{M} \eta_{jm}\mathbf{z}'_{mij}\boldsymbol{\lambda}_m, \qquad \lambda_{m1} = 1$$

- Latent variable vector $\boldsymbol{\eta}_j$ for unit j with discrete values (or locations) $\boldsymbol{e}_c, c\!=\!1,\cdots,C$ in M dimensions
- Units in same latent class share the same value or location $\boldsymbol{e}_c$
- Probability that unit j is in latent class c is $\pi_c = \frac{\exp(\varrho)}{1+\exp(\varrho)}$
- Two parameterizations:
  1. **non-centered**: $\boldsymbol{e}_c$, C locations freely estimated
  2. **centered**: $\widetilde{\boldsymbol{e}}_c$, $C-1$ locations estimated, last location determined by constraint $\sum_c \pi_c \widetilde{\boldsymbol{e}}_c = 0$
     Allows mean structure to be modeled using $\mathbf{x}'_{ij}\boldsymbol{\beta}$

Language Model Specification

- Discrete trajectory classes located in 4 dimensions (2 intercept x 2 slope)
  - allows random effects to be correlated across expressive and receptive.
  - Increase number of classes and select "best-fit" model

- 6 receptive measures using 4 tests
  - 4 intercept factor loadings constrained equal to corresponding slope factor loadings
  - 4 measurement error variances
- 6 expressive measures using 4 tests (2 near parallel)
  - 3 intercept factor loadings constrained equal to corresponding slope factor loadings
  - 3 measurement error variances

`gllamm` model for joint expressive and receptive language trajectory classes

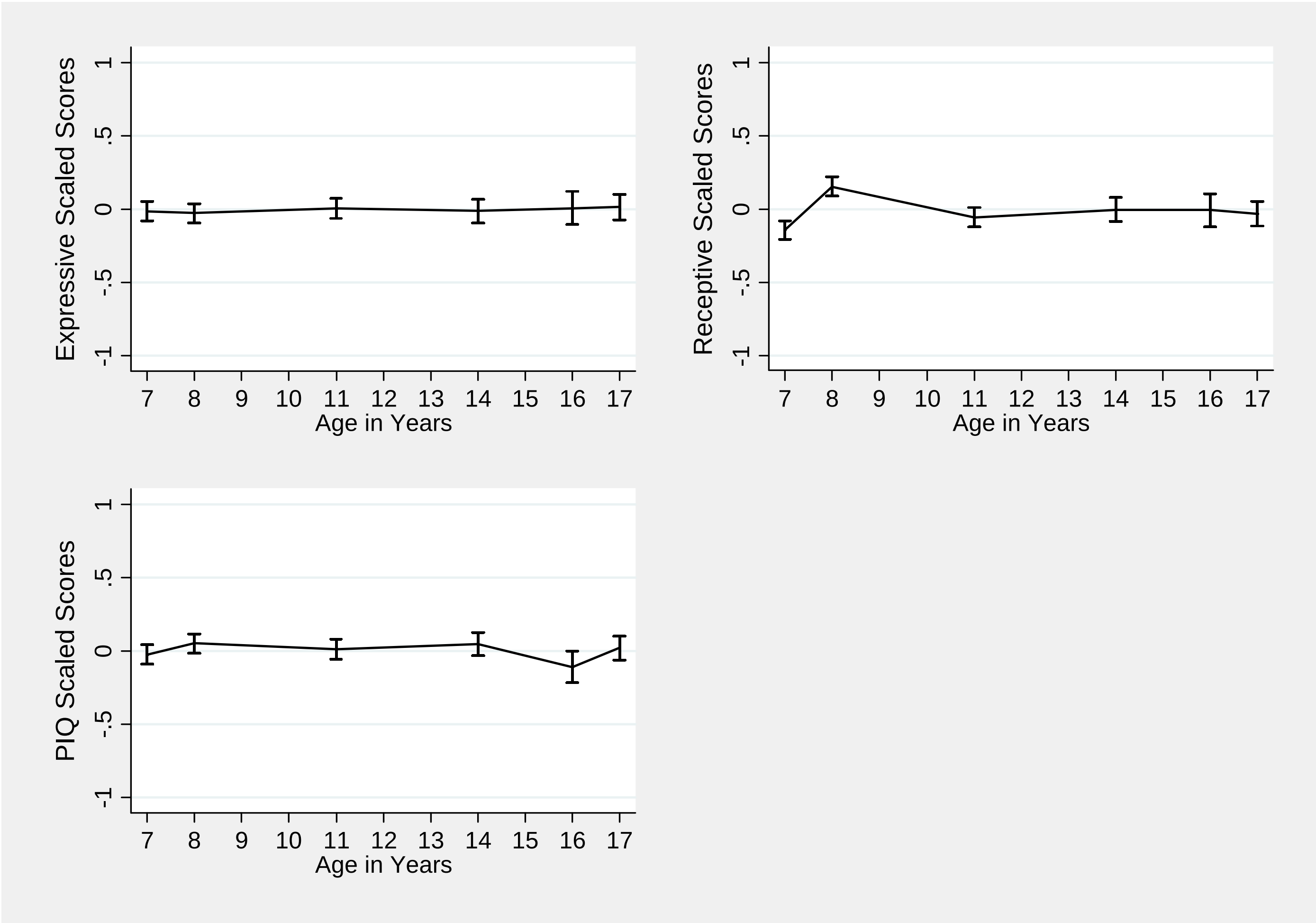
```
eq het: e1 e2 e5 r1 r2 r3 r4  ! Eqn for log std dev of measurement error
eq inte: e1 e2 e5             ! Eqn for expressive intercept factor loadings
eq intr: r1 r2 r3 r4          ! Eqn for receptive intercept factor loadings
eq line: ageye1 ageye2 ageye5        !Eqn for exp linear slope factor loadings
eq linr: ageyr1 ageyr2 ageyr3 ageyr4 !Eqn for rec linear slope factor loadings
cons def 1 [fid1_1l]e2 = [fid1_2l]ageye2 !Constraints for intercept and slope
cons def 2 [fid1_1l]e5 = [fid1_2l]ageye5 !   factor loadings equal
cons def 3 [fid1_3l]r2 = [fid1_4l]ageyr2
cons def 4 [fid1_3l]r3 = [fid1_4l]ageyr3
cons def 5 [fid1_3l]r4 = [fid1_4l]ageyr4

gllamm y e1 e2 e3 e4 r1 r2 r3 r4 ageye1 ageye2 ageyr1, i(fid) nrf(4) /*
 */ eqs(inte line intr linr) s(het) nip(6) cons(1 2 3 4 5) iter(40)/*
 */ nocons trace
```

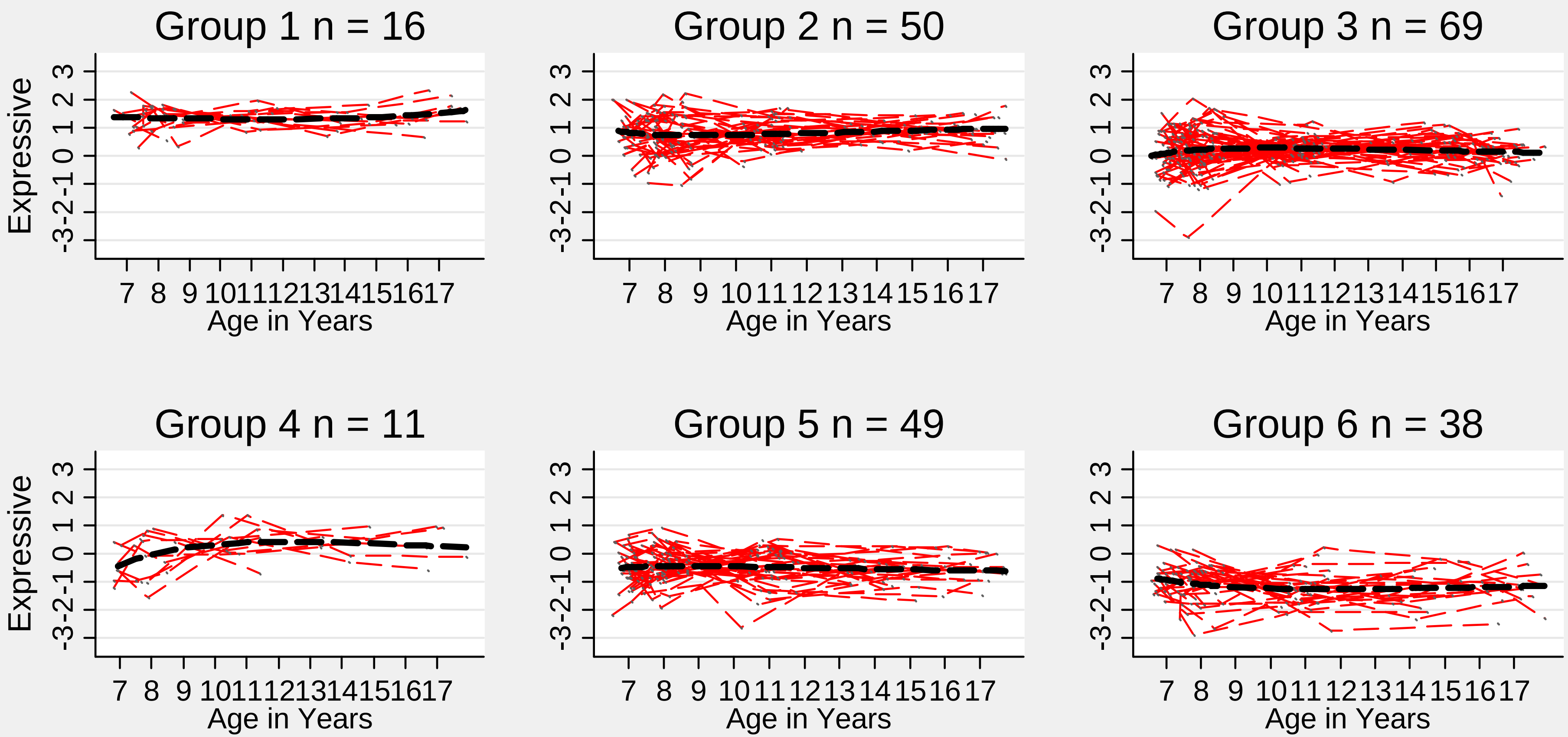
Classification for categorical latent variables

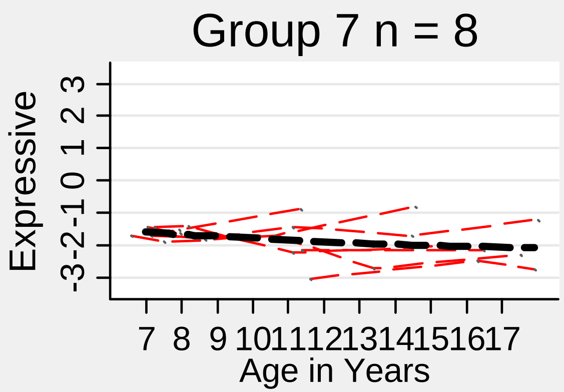
- Units are usually assigned to latent class with largest posterior probability, often called Maximum Aposteriori (MAP) or Empirical Bayes Modal (EBM)
- Posterior probabilities:

$$\Pr(c \mid y_j) = \frac{\pi_c \prod_{i=1}^{I} \pi_{ij|c}^{y_{ij}} (1 - \pi_{ij|c})^{1-y_{ij}}}{\sum_{c=1}^{C} \pi_c \prod_{i=1}^{I} \pi_{ij|c}^{y_{ij}} (1 - \pi_{ij|c})^{1-y_{ij}}}$$

Figure 2. Expressive language, receptive language and PIQ ability (in scaled score format) from age 7 to 17 (whole sample means with standard error bars)

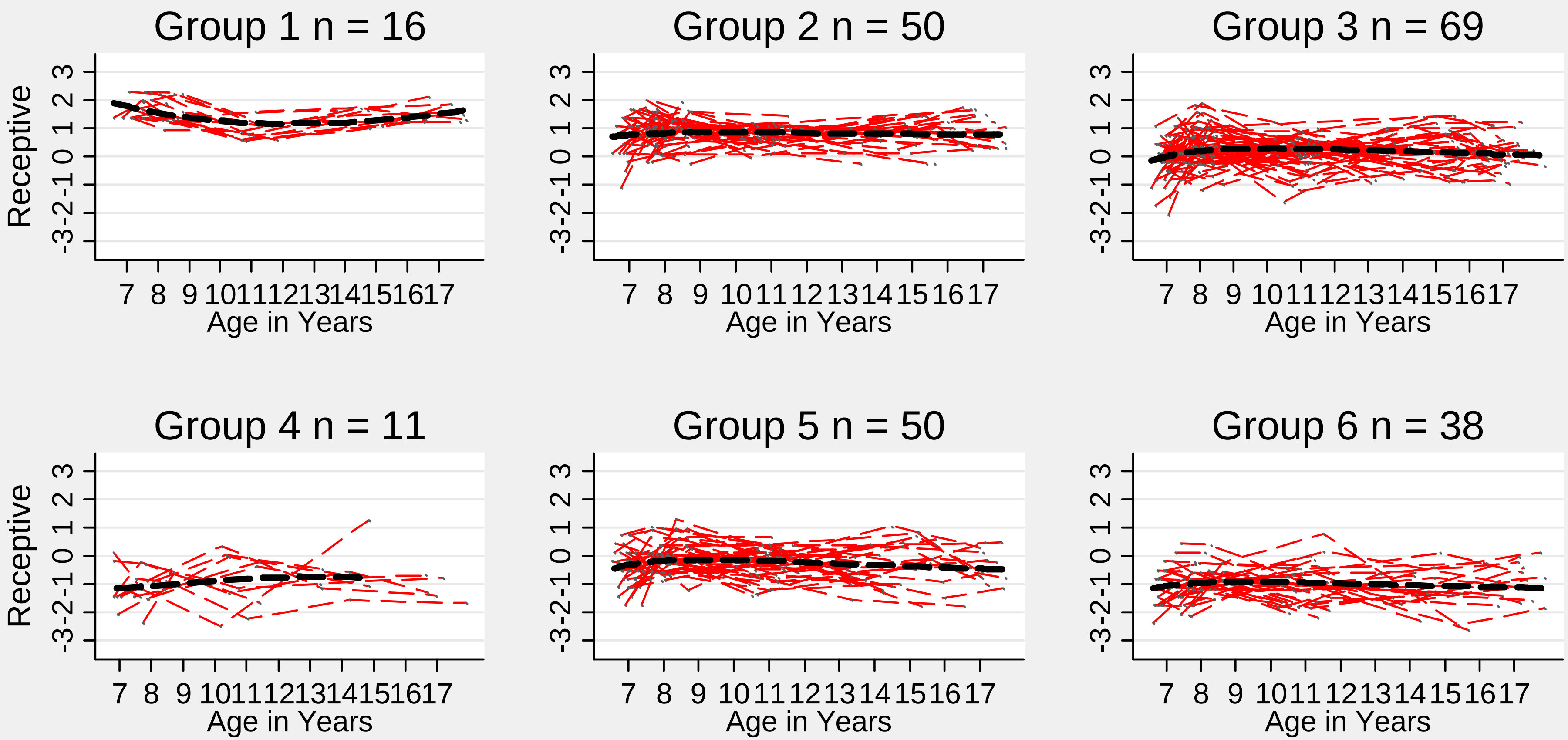


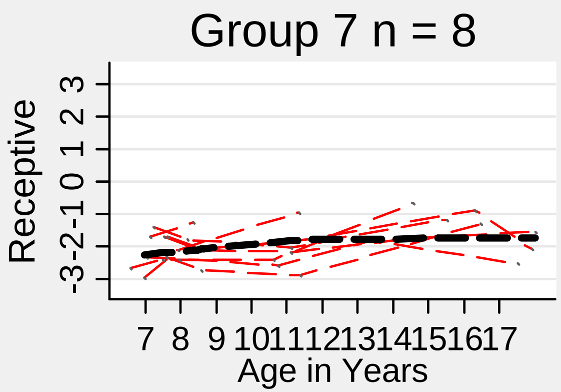




At each age point (7,8,11,14,16 and 17) the **Expressive language** score standardized to the mean (0) of the entire SLI population.

No differences in the developmental trajectory – the same relative level of expressive language is maintained

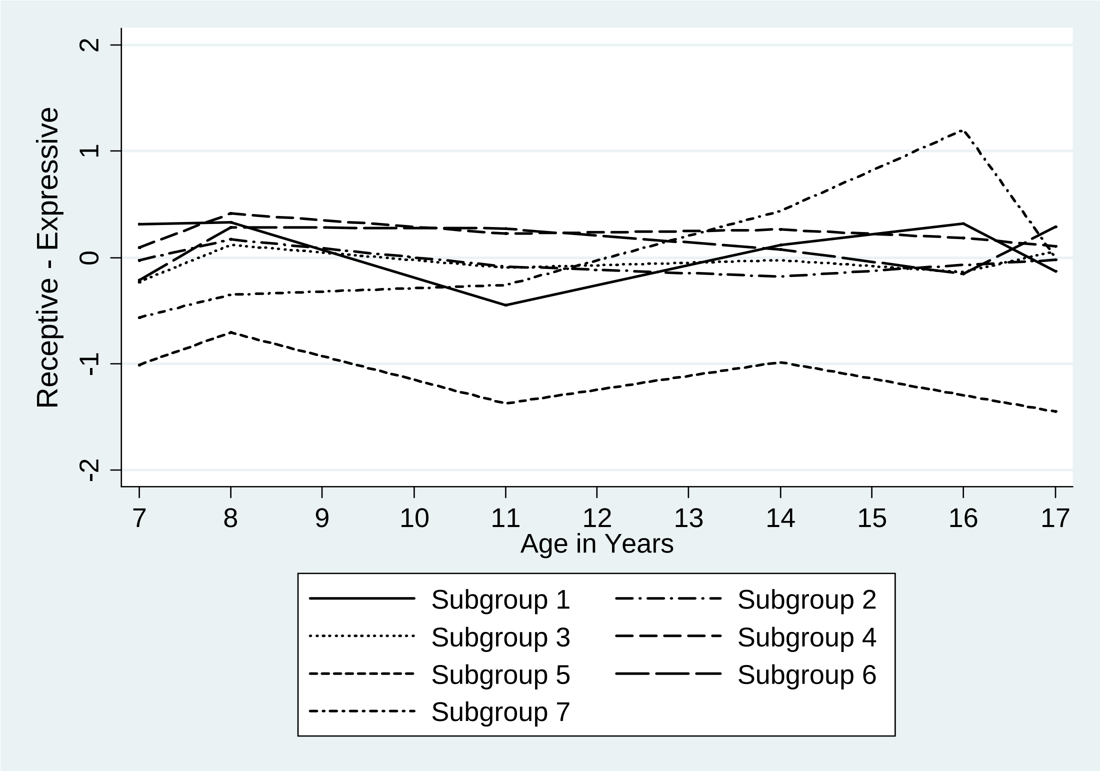




At each age point (7,8,11,14,16 and 17) the **Receptive language** score standardized to the mean (0) of the entire SLI population.

No differences in the developmental trajectory – the same relative level of receptive language is maintained

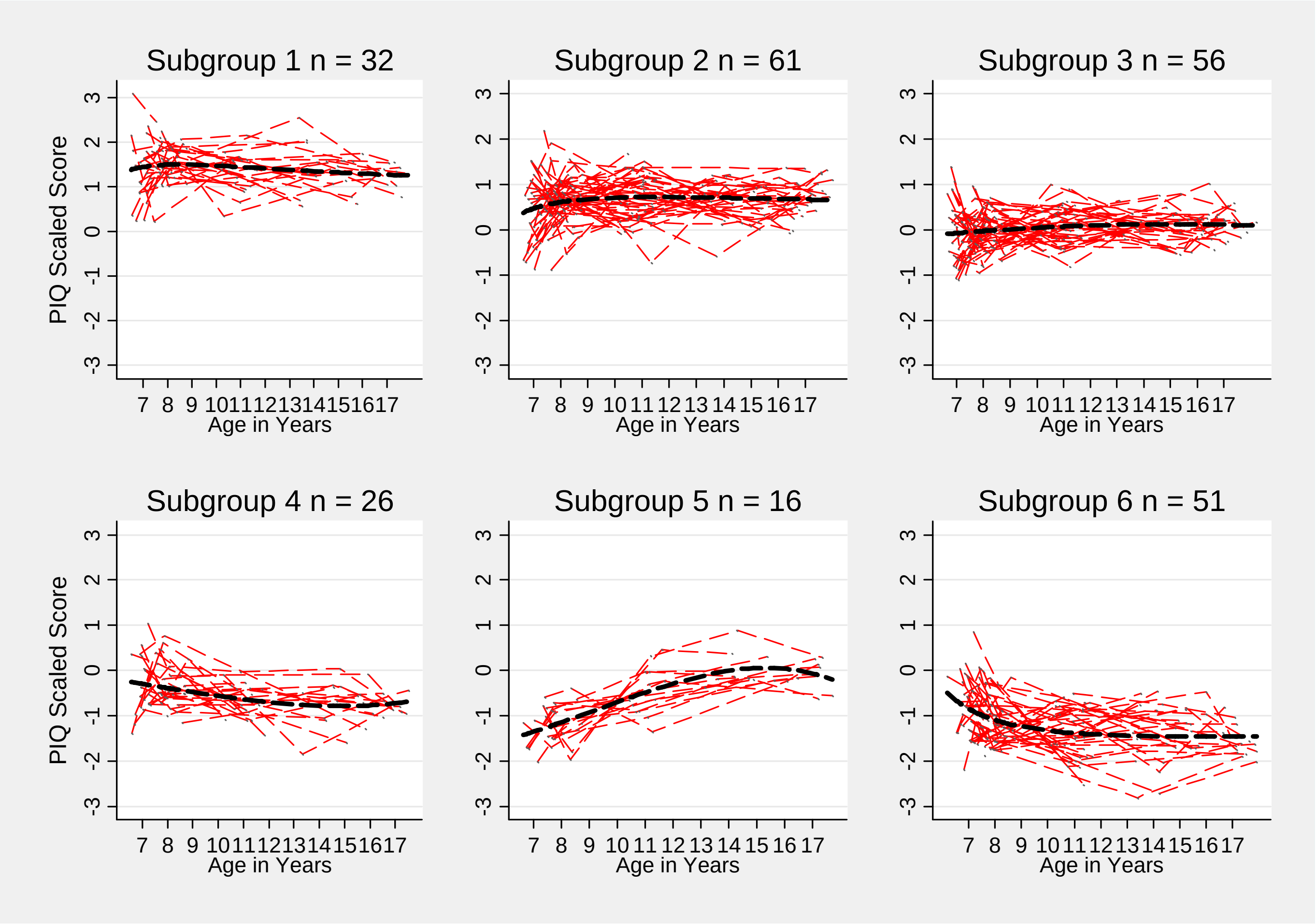
Receptive-expressive discrepancy



Expressive and Receptive largely continue in tandem except for one small group

Performance IQ trajectories

Figure 6. Individual developmental trajectories and average developmental trajectories of PIQ ability in PIQ subgroups 1 to 6



SLI Study Conclusions

- Expressive and receptive language development of children with language impairment is one of remarkably homogeneity with children retaining their relative rank orderings in both language domains.

- The little naturalistic variation in development is not encouraging as to the prospects for bringing about change through intervention after age 6.

- By contrast, and surprisingly, trajectories of performance IQ were more heterogeneous.

Trends in child mental health study

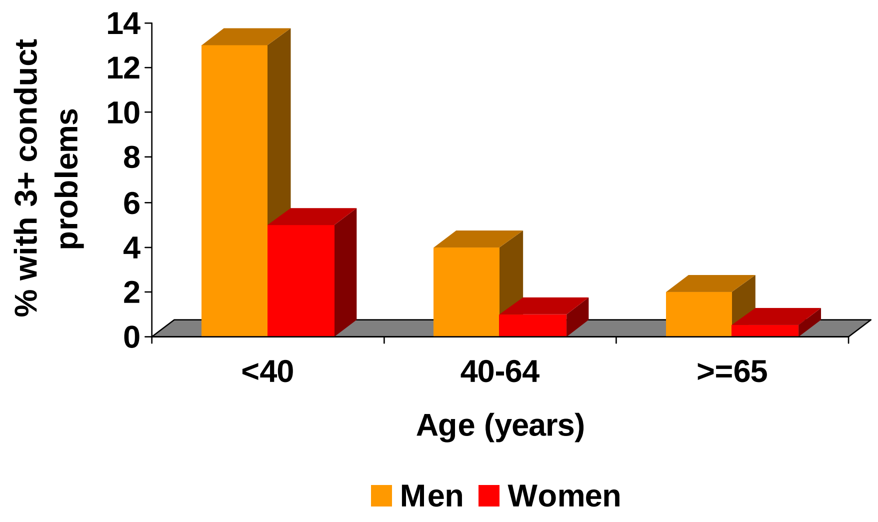
- Evidence on trends

- Specific Application
  - The TRENDS Data Set and original findings
  - Regression, SEM & and MI
  - Comparison of Methods
- Some simulation
- Further work on trends
- Other Applications

Possible artefacts in reports

Retrospective report

- selective mortality/ institutionalization
- effects of memory & recall
- changes in 'psychological -mindedness'
- general reporting bias



Prospective Approach

- Changes in definitions, completeness and coverage of administrative recording
- Self, parent and teacher reports from national cohort studies

TRENDS Data Set

	Questionnaire Administered	Design	Number of Participants (male, female)	Reliability (Cronbach's α)
National Child Development Study (NCDS) (Fogelman 1983)	Rutter-A	Longitudinal	aged 15: 10,499 (5371, 5128)	0.76
1970 British Cohort study (BCS70) (Butler & Golding 1986)	Rutter-A	Longitudinal	aged 15: 7293 (3533, 3760)	0.79
1999 British Child and Adolescent Mental Health Survey (B-CAMHS99) (Meltzer et al. 2000).	Strengths and Difficulties (SDQ)	Cross-Sectional	aged 15: 868 (439, 429)	0.71
Calibration Sample	Both SDQ and Rutter A	N/A	All ages: 380 (203, 177)	SDQ: 0.88 Rutter A: 0.79

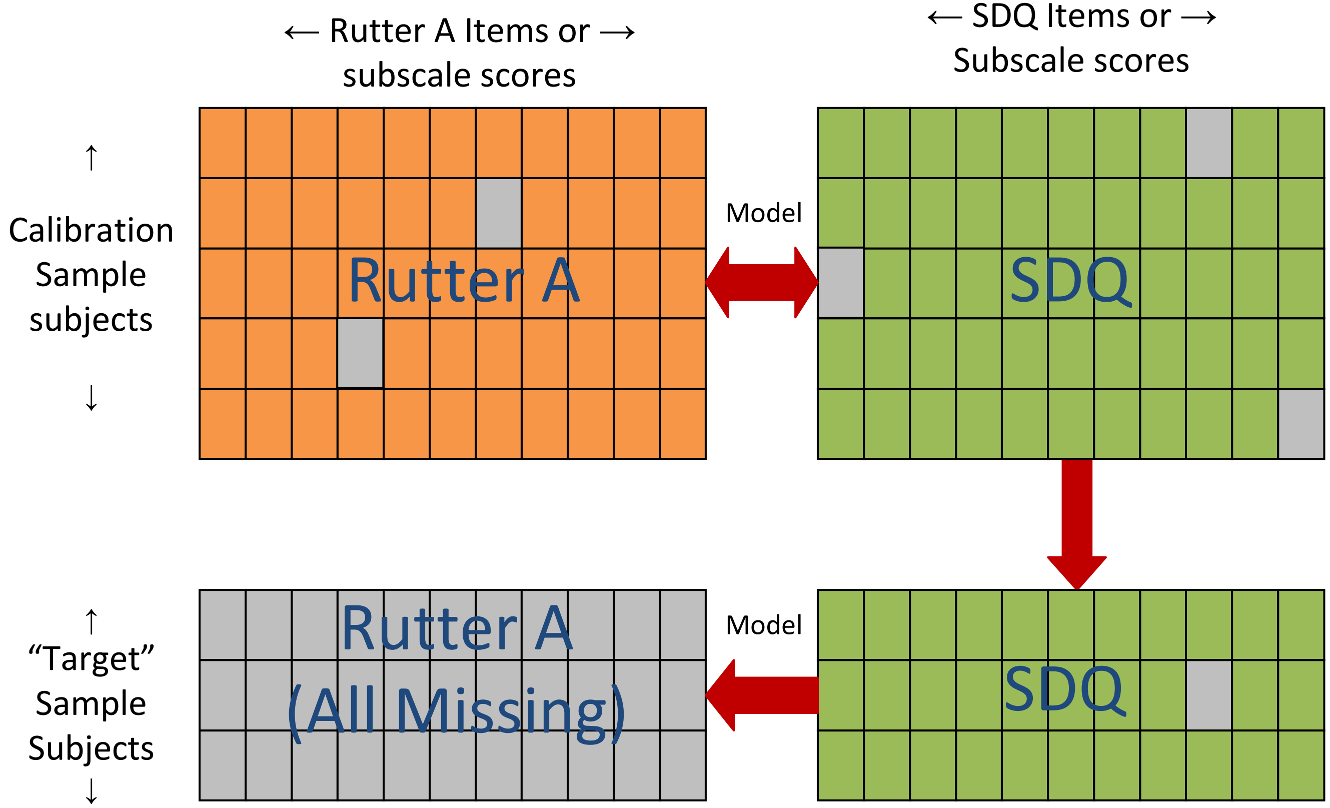
See: Collishaw & Pickles *et al.*, Journal of Child Psychology and Psychiatry 45:8 (2004), pp 1350–1362 for more information about these data sets.

Connection to Missing Data Problems

- Missing Completely at Random (MCAR): missingness of a variable independent of any variables
- Missing At Random (MAR): missingness of a variable independent of its values though possibly dependent on other variables
- Missing Not At Random (MNAR): missingness of a variable depends on its values

- Internal Calibration Samples: generally MCAR
- External Calibration Samples: generally MAR or NMAR

Basic Problem: Infer the 1999 Rutter-A Data via Linear Regression, Multiple Imputation, & SEM



Original Method: Calibration

Predicted items/subscales should possess all the variability and inter-item/scale associations as original A-scale items:

- problems of overfitting
  - use Bayesian approach
  - use pragmatic approach

For each measure of interest fit ordinal logistic regression to predict a Rutter-A output item/scale from a set of input SDQ predictors that consisted of:

- any closely matching input items
- relevant sub-scale scores
- overall scale score

Done separately for boys and girls

Original Method: Multiple Imputation

To reflect the uncertainty in our prediction equation, we first sampled the estimated coefficients and thresholds of our ordinal logistic regressions by drawing values from a multivariate normal distribution defined by the estimated parameter covariance matrix.

$$\beta_{imp} = \sim N(\beta, \Sigma_{\beta})$$

We then used these β_{imp} to predict the probability of each feasible response value for each individual (e.g. 0, 1 or 2). One of these values was then picked with probability equal to this estimated response probability.

Repeated 20 times to produce 20 B-CAMHS99 datasets with Rutter A scale measures (4 times the “then rule of thumb” of 5)

Rubin, D. 1987, *Multiple Imputation for Nonresponse in Surveys* (J. Wiley & Son.)

Original Method: Multiple Imputation

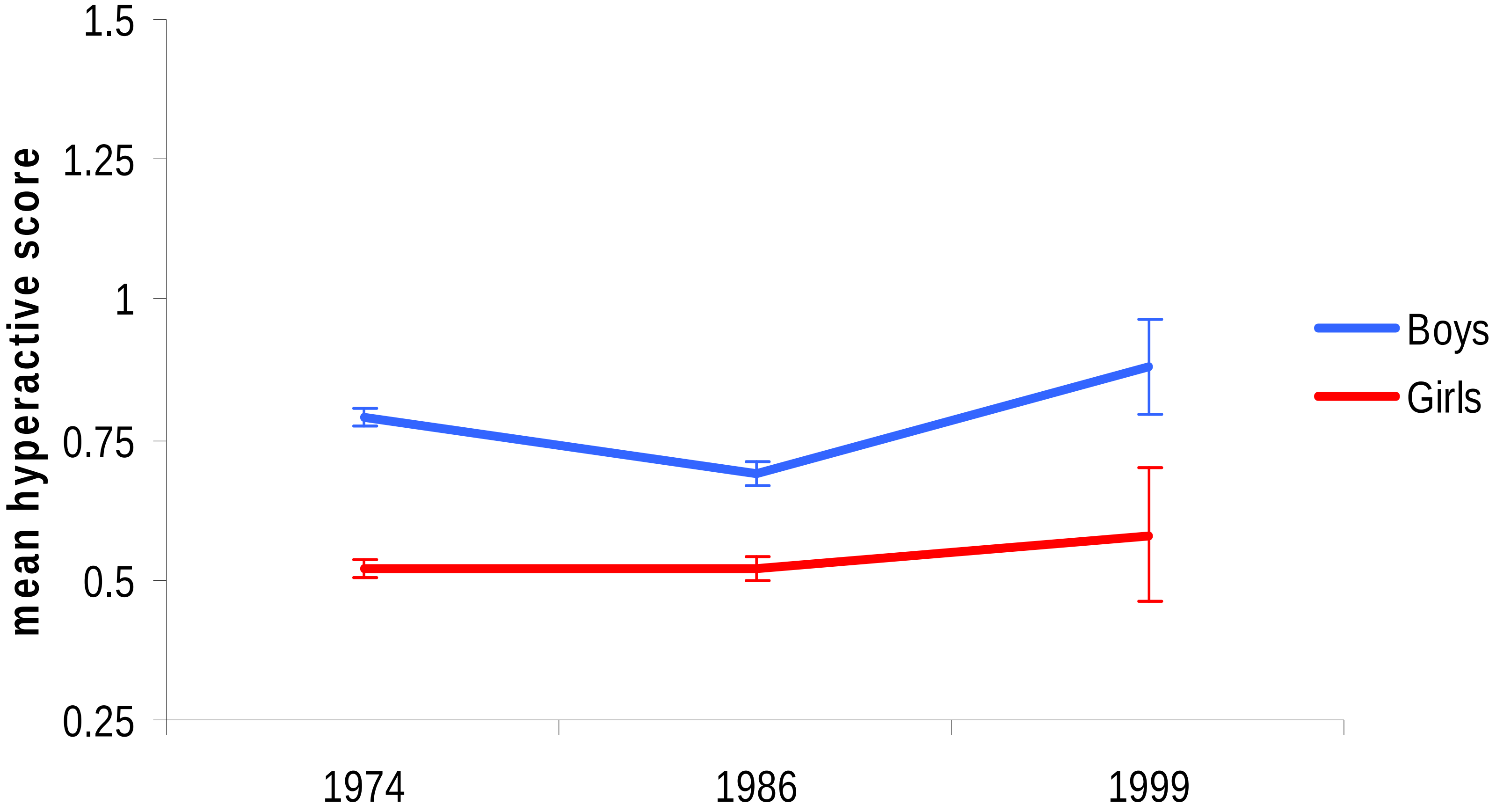
- In this way the ‘made-up’ measures properly reflected behavioural variation as reported by the SDQ but the extent to which these datasets differed one from another properly reflected the uncertainty as to what value each of those ‘made-up’ values should be.

Analyse each of the 20 (m) datasets (where data the same for NCDS, BCS-70 but may differ for B-CAMHS99) and use Rubin’s Rule

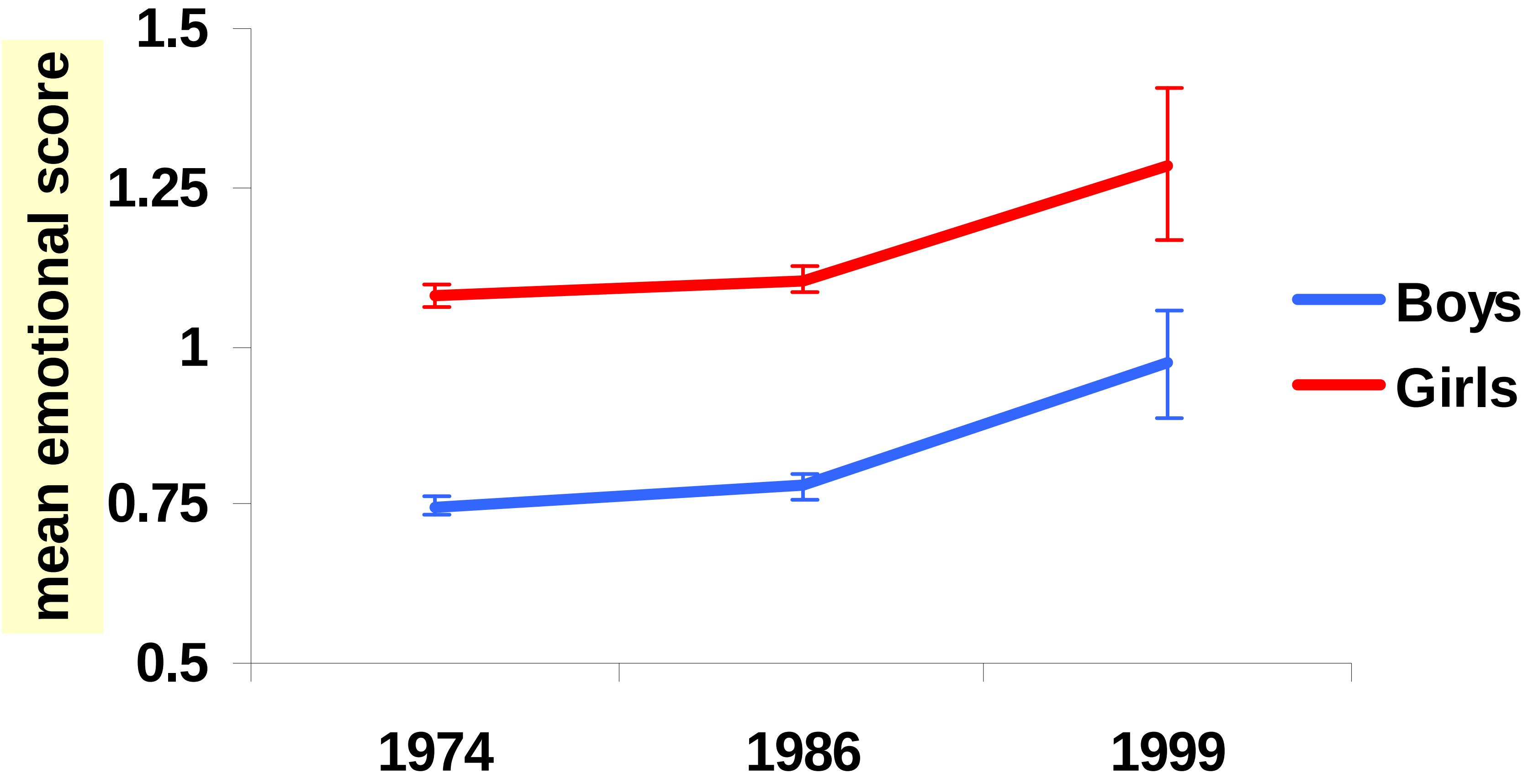
Parameter estimate = mean of 20 estimates

Estimated parameter variance =
Estimated mean variance + (1+m-1) estimated between dataset variance

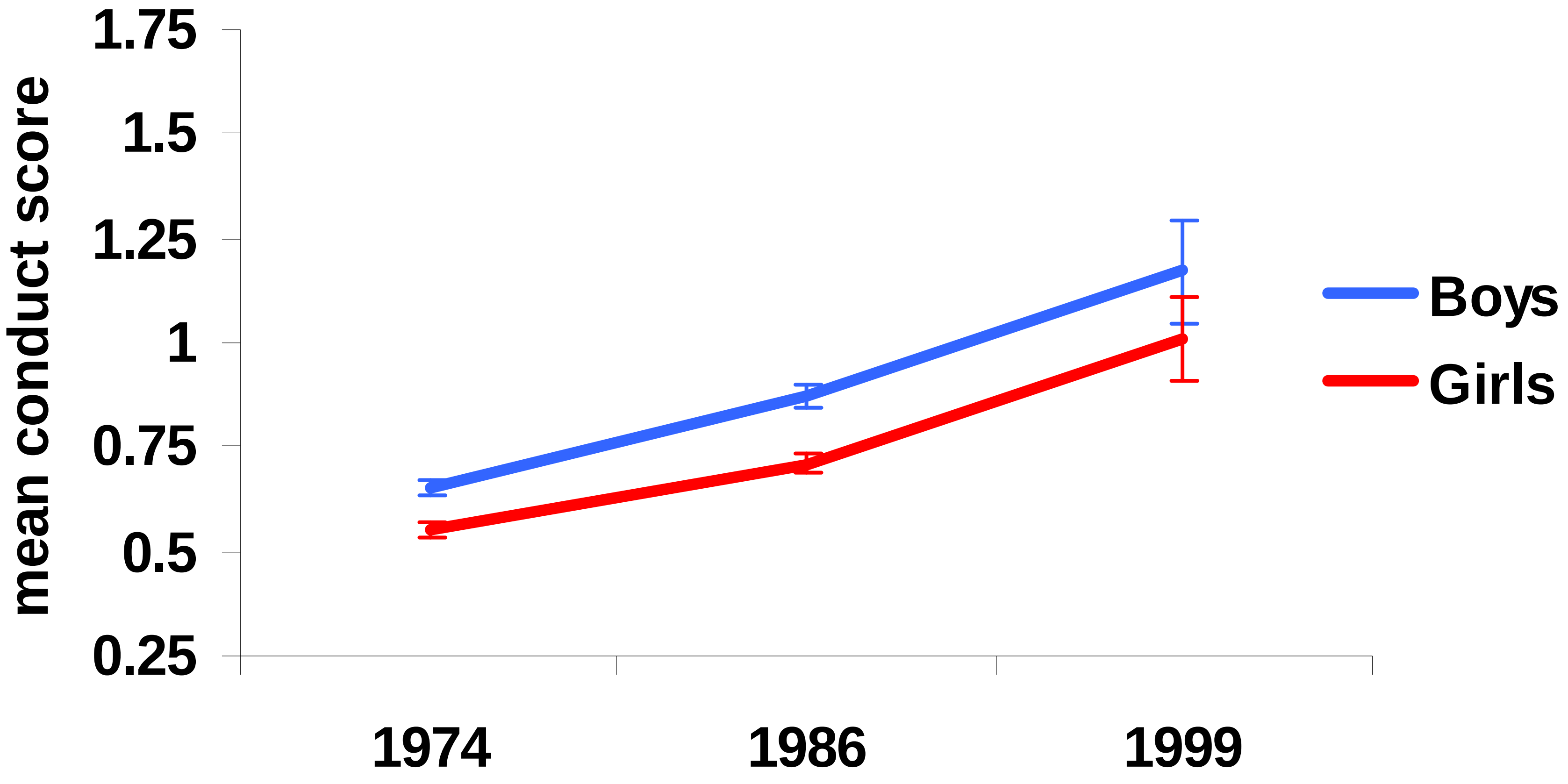
Time trends in adolescent hyperactive problems



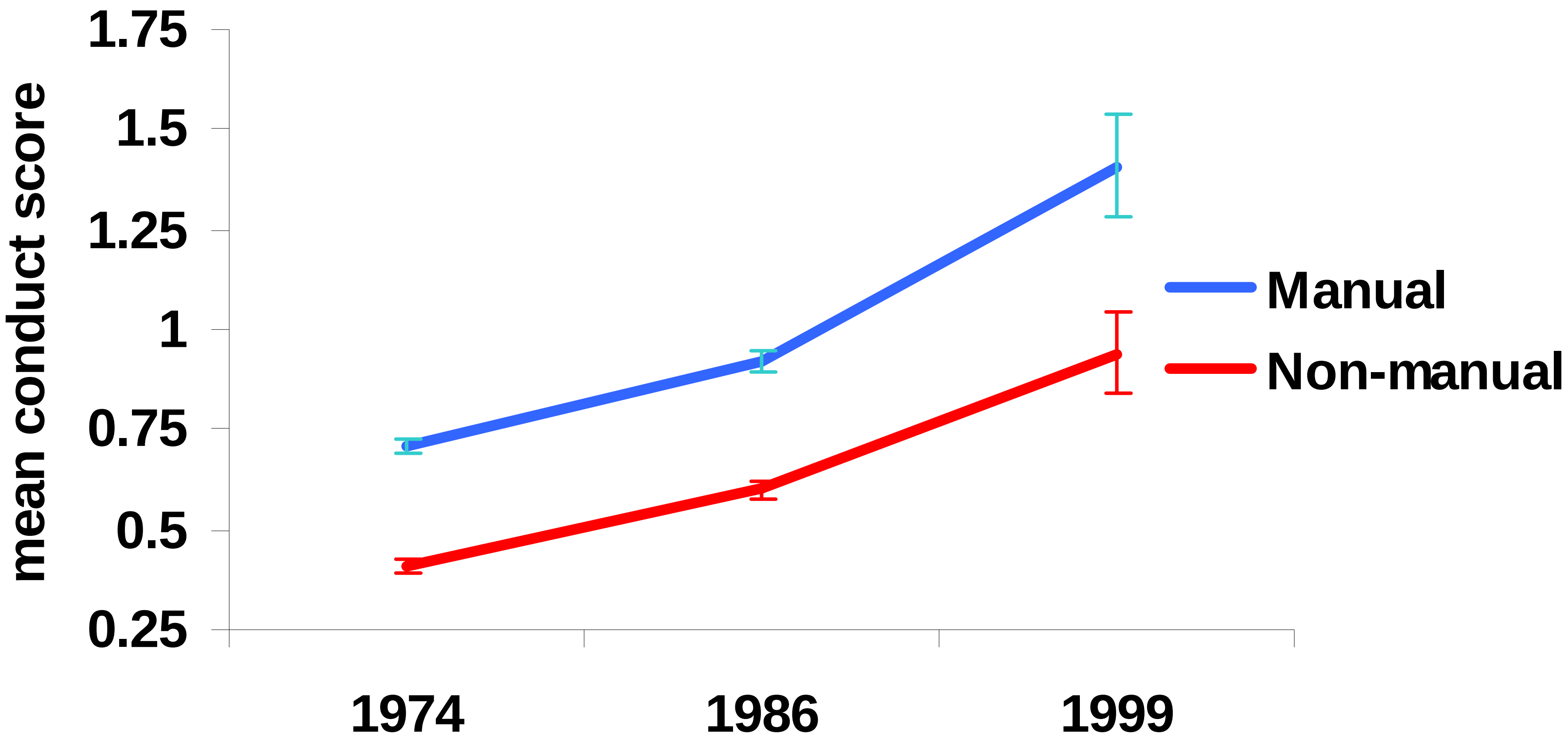
Time trends in adolescent emotional problems



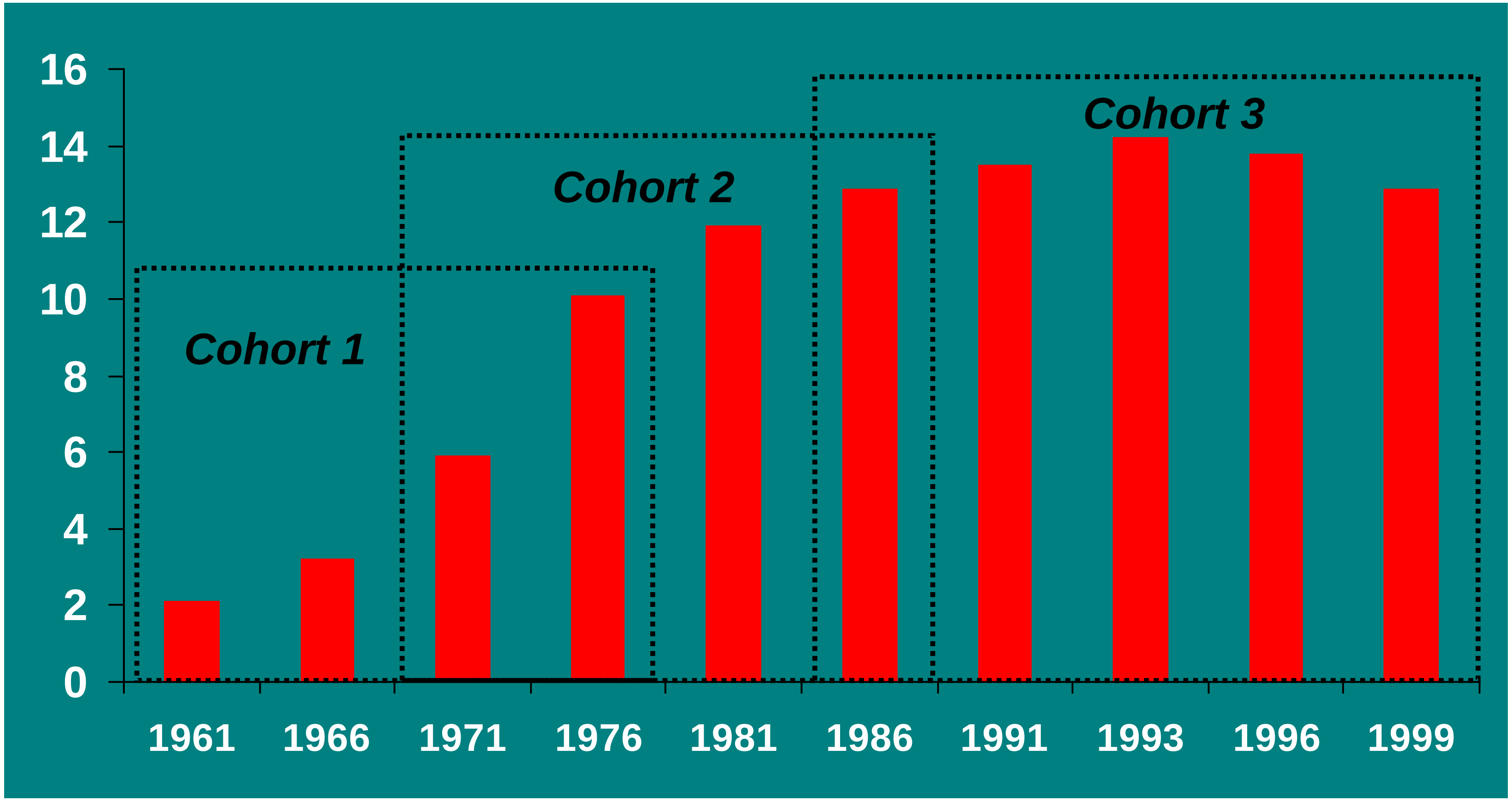
Time trends in adolescent conduct problems



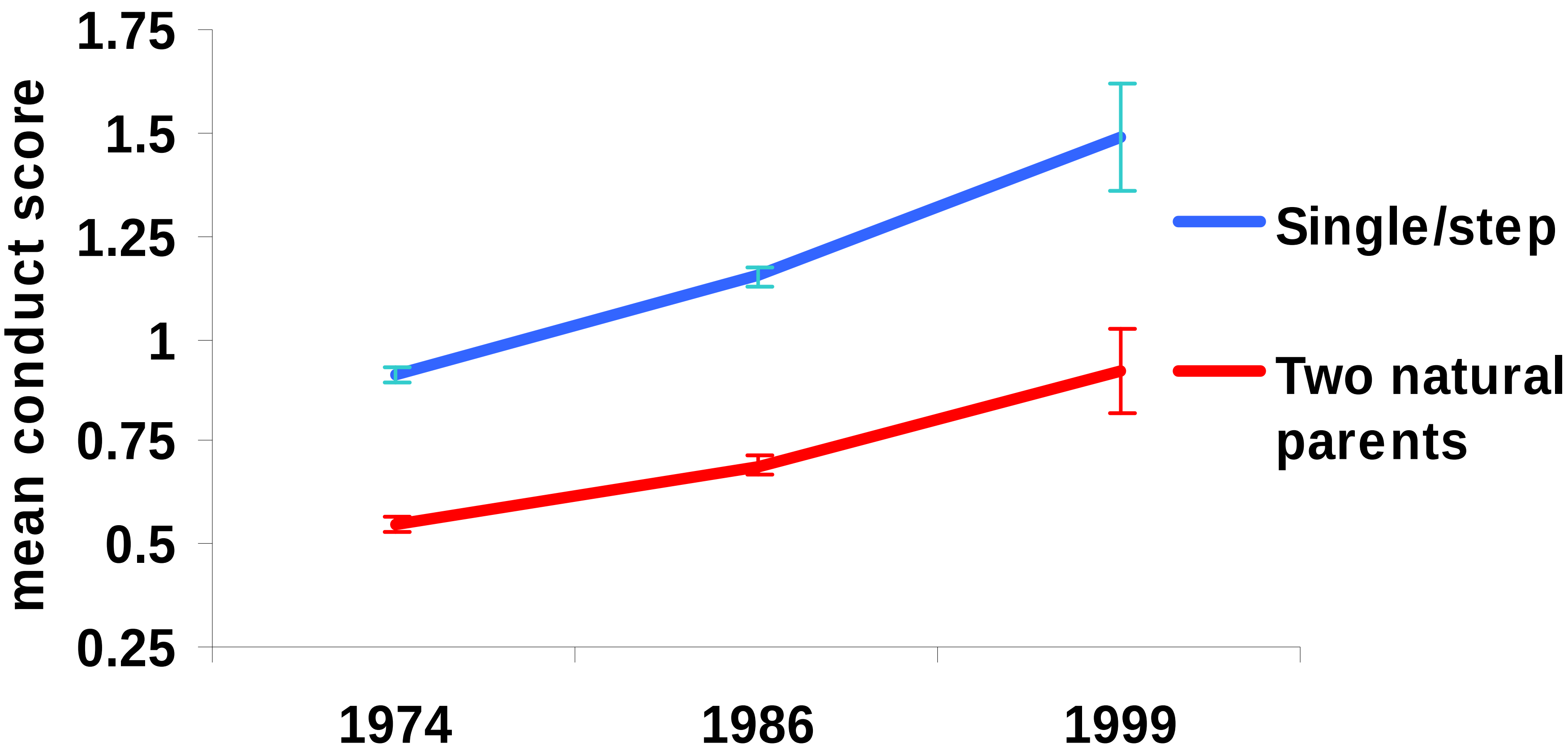
Trends in conduct problems: by social class



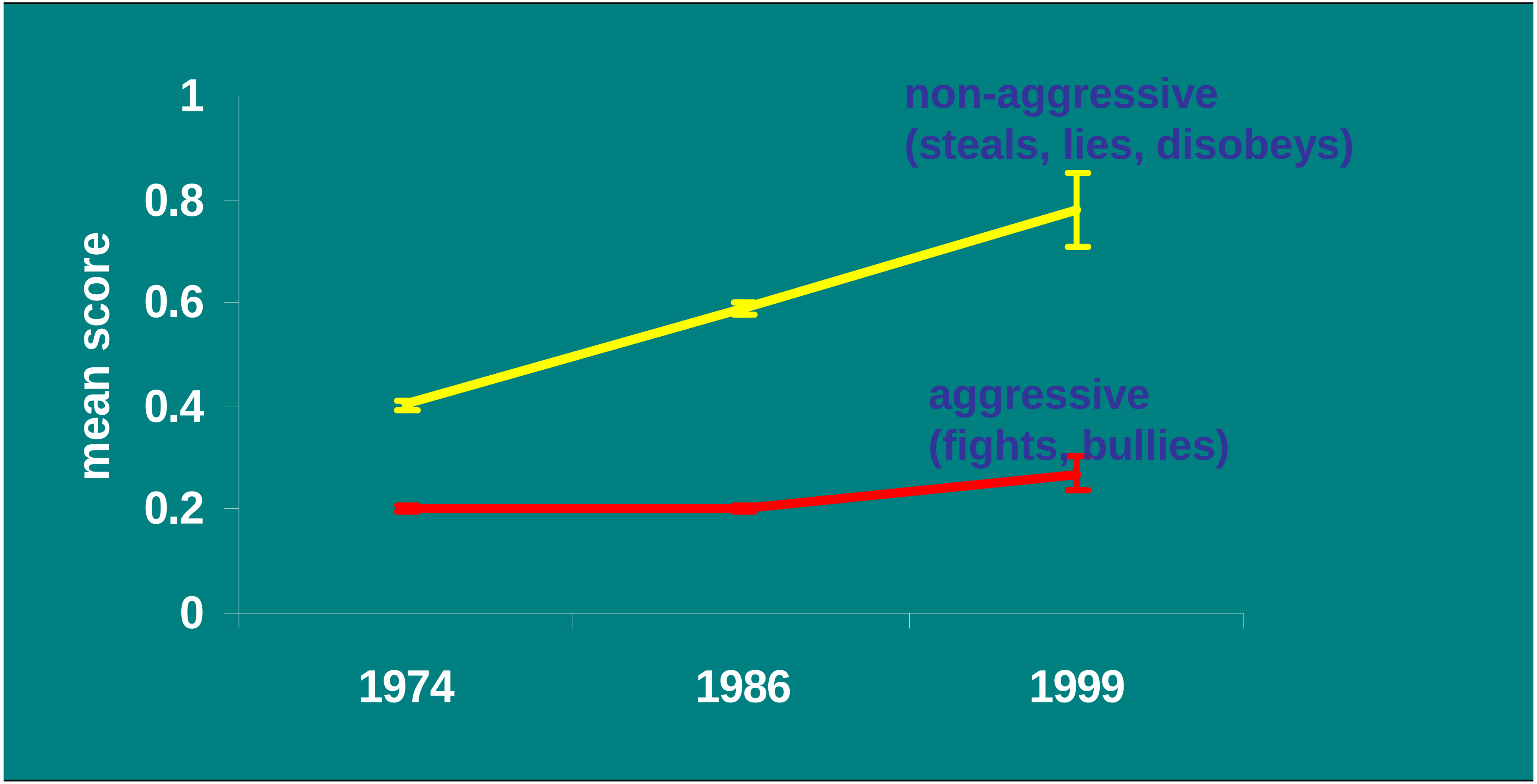
Divorce rate per 1,000 married population 1961-1999 (England & Wales)



Trends in conduct problems: by family type



Trends in aggressive and non-aggressive problems



Can we believe these results?

- comparison with alternative estimation methods

- comparison with parallel questions

- validation against additional criteria

Linear Regression

$$y_i^{c,t} = a + bx_i^{c,t} + \varepsilon_i^{c,t}$$

Calibration Sample: $\bar{x}^c = \frac{1}{n^c}\sum_{i=1}^{n^c} x_i^c$ and $\bar{y}^c = \frac{1}{n^c}\sum_{i=1}^{n^c} y_i^c$

"Target" Sample: $\bar{x}^t = \frac{1}{n^t}\sum_{i=1}^{n^t} x_i^t$ and $\bar{y}^t = \frac{1}{n^t}\sum_{i=1}^{n^t} y_i^t$

$$\bar{y}^t = \bar{y}^c + (\bar{x}^t - \bar{x}^c)\frac{\rho_{xy}\sigma_y}{\sigma_x}$$

This is a “one-step” process

Standard Errors of Predicted Mean

$$\mathrm{Var}\left(\overline{y}\right) = s^2 \left(\frac{1}{n_c} + \frac{(\overline{x}^c - \overline{x}^t)^2}{\sum_{i=1}^{n_c} (x_i^c - \overline{x}^c)^2} \right) = \sigma^2$$

Does not include any "variability" in $\overline{x}^t$
i.e., this assumes

$$\mathrm{Var}\left(x^t\right) = 0$$

Total Variance

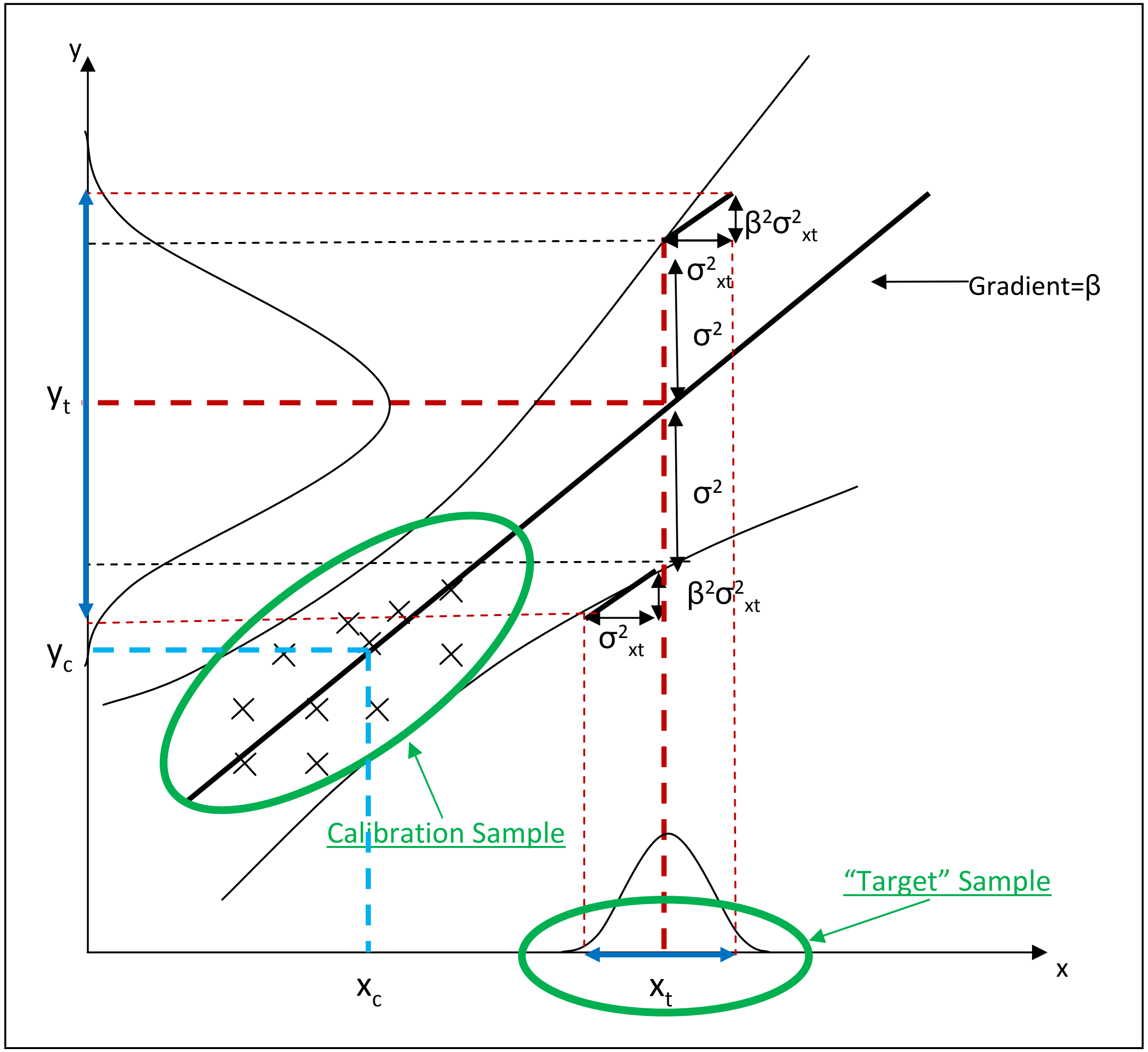
$$\mathrm{Var}\left(\bar{y}^t\right) = \mathrm{E}\left(\mathrm{Var}\left(\bar{y}^t \middle| y^c\right)\right) + \mathrm{Var}\left(\mathrm{E}\left(\bar{y}^t \middle| y^c\right)\right)$$

$$\mathrm{E}\left(\bar{y}^t / y^c\right) \quad = \quad a + b\bar{x}^t \quad \rightarrow \quad \mathrm{Var}\left(\mathrm{E}\left(\bar{y}^t \middle| y^c\right)\right) = \frac{b^2}{n^t}\mathrm{Var}\left(x^t\right)$$

$$\mathrm{Var}\left(\bar{y}^t \middle| y^c\right) \quad = \quad \sigma^2 \quad \rightarrow \quad \mathrm{E}\left(\mathrm{Var}\left(\bar{y}^t \middle| y^c\right)\right) = \sigma^2$$

$$\mathrm{Var}\left(\bar{y}^t\right) = s^2\left(\frac{1}{n^c} + \frac{(\bar{x}^c - \bar{x}^t)^2}{\sum_{i=1}^{n_c}(x_i^c - \bar{x}^c)^2}\right) + \frac{b^2\mathrm{Var}(x^t)}{n^t}$$

Or Graphically



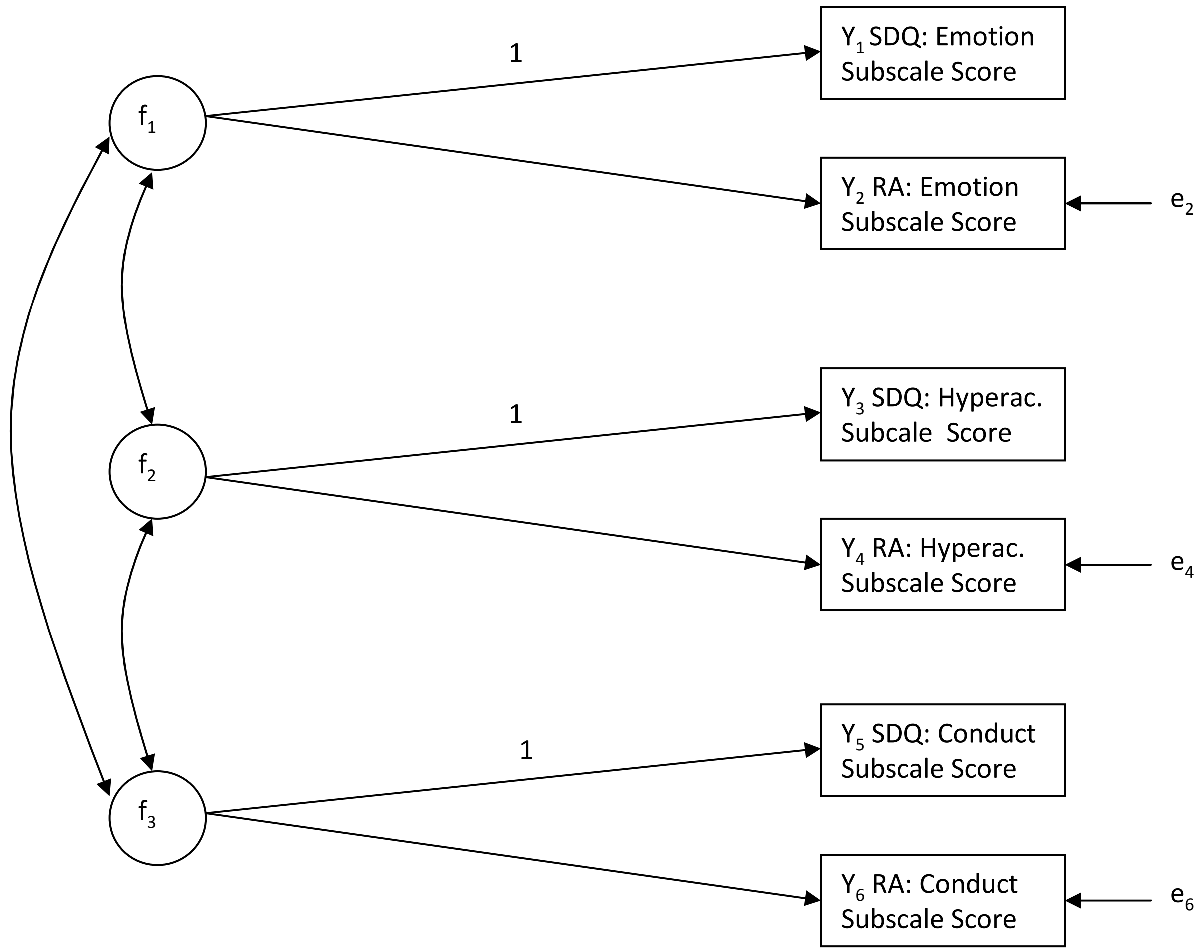
Multiple Regression

$$y = a + \sum_{j=1}^{l} b_j x_{ji} + \varepsilon_i \qquad X = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{l1} \\ 1 & x_{12} & x_{22} & & x_{l2} \\ 1 & x_{13} & x_{23} & & x_{l3} \\ \vdots & & & & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{ln} \end{bmatrix} \qquad X_0 = \begin{bmatrix} x_{01} \\ x_{02} \\ \vdots \\ x_{0n} \end{bmatrix}$$

$$\sigma^2 = s^2 X_0' \left[X'X \right]^{-1} X_0$$

$$\text{total variance} \quad \rightarrow \quad \mathrm{Var}\left(\bar{y}^t\right) \equiv \sigma^2 + \mathrm{Var}\left(a + \sum_{j=1}^{l} b_j \bar{x}_j \right) = \sigma^2 + \frac{1}{n_t} \sum_{j=1}^{l} \sum_{k=1}^{l} b_j b_k \mathrm{Cov}\left(x_j, x_k\right)$$

Structural Equation Modelling



SEM: Maximum Likelihood

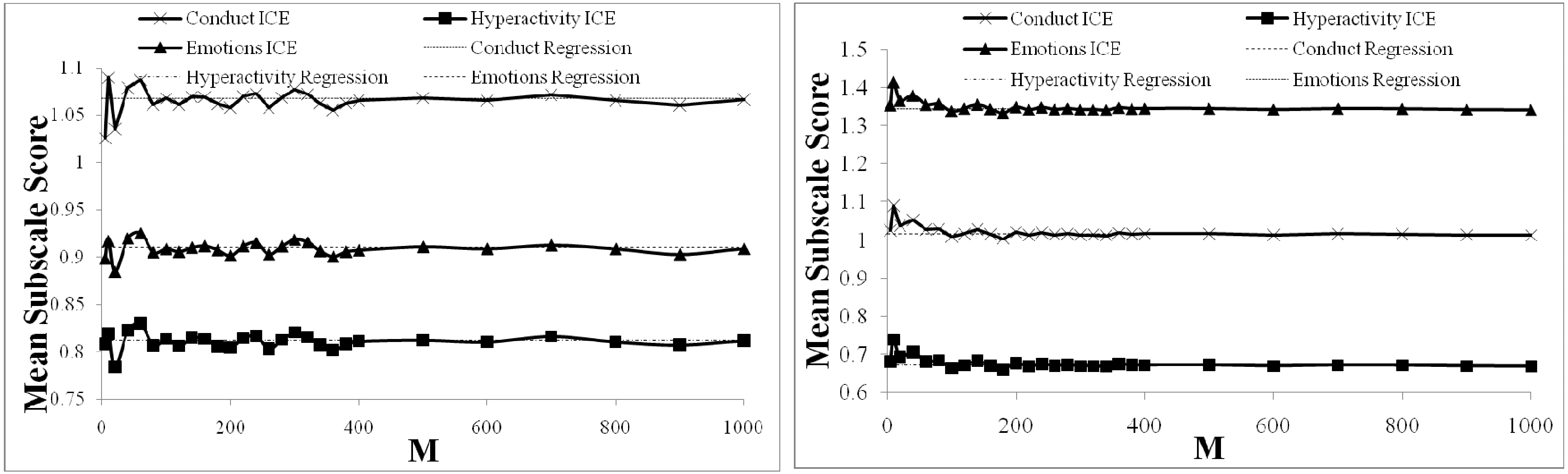
- Mplus v.5 used here to solve for SEM above.
- Maximum likelihood used to infer properties of missing data
- Estimates of mean subscale scores may be found and their standard errors directly.
- As for linear regression, a “one-step” process.

Results: Convergence of MI compared to Linear Regression & SEM

Method	Females			Males		
	Conduct	Hyperactivity	Emotions	Conduct	Hyperactivity	Emotions
"Simple Regression"	1.0153 (0.081) [0.103]	0.6707 (0.075) [0.090]	1.3444 (0.076) [0.094]	1.0675 (0.077) [0.102]	0.8120 (0.065) [0.086]	0.9100 (0.070) [0.084]
SEM (1)	1.015 (0.096)	0.671 (0.079)	1.344 (0.087)	1.067 (0.097)	0.812 (0.088)	0.910 (0.074)
ICE MI (1) M=5	1.0257 (0.1432)	0.6800 (0.1220)	1.3541 (0.1304)	1.0257 (0.1432)	0.8085 (0.1209)	0.8986 (0.1228)
ICE MI (1) M=20	1.0380 (0.1061)	0.6927 (0.0913)	1.3655 (0.0961)	1.0355 (0.1404)	0.7839 (0.1165)	0.8847 (0.1194)
ICE MI (1) M=100	1.0084 (0.1218)	0.6641 (0.1073)	1.3379 (0.1114)	1.0678 (0.1237)	0.8135 (0.1026)	0.9088 (0.1044)
ICE MI (1) M=500	1.0160 (0.1307)	0.6713 (0.1157)	1.3450 (0.1196)	1.0680 (0.1260)	0.8123 (0.1059)	0.9108 (0.1041)
ICE MI (1) M=2000	1.0148 (0.1266)	0.6702 (0.1118)	1.3439 (0.1159)	1.0666 (0.1275)	0.8112 (0.1069)	0.9096 (0.1059)

Mean Rutter-A scores (standard error of the mean in brackets) for the conduct, hyperactivity, and emotion subscales in 1999 for 15-year subjects. ICE=Multiple Imputation; SEM=Structural Equation Model (MPlus). Results for "usual" standard error in curved brackets and of total variance in square brackets.

Convergence of MI



Convergence of mean imputed subscale scores for via ICE (1): (left) females; (right) males

Conclusions for TRENDS Dataset

- Results support those calculations of AP
  - An increase in conduct problems in both sexes
  - An increase in emotional problems in girls
  - Mixed evidence for a rise in hyperactivity
- Biases may well have occurred. However, our results (looking at mean scores + also using reweighting) suggest that bias is arguably not strong in this case.
- Multiple Imputation takes much longer to “converge” for regression/calibration type problems than the “rule of thumb” of 5 imputations.
- Multiple Imputation consistently gave the largest standard errors.

Replication without calibration by using parallel questions

- Cohort 4: The 2002 and 2003 Health Surveys for England (Department of Health, 2003; National Statistics, 2004)

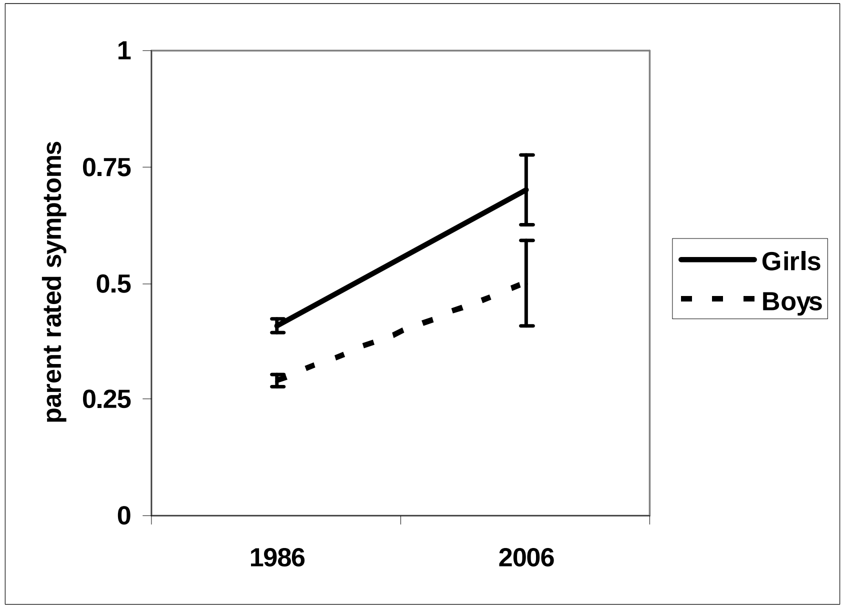
- 1401 children born 1st April 1988 to 31st March 1990 (mean age = 17.1 years, sd = 0.57 years)

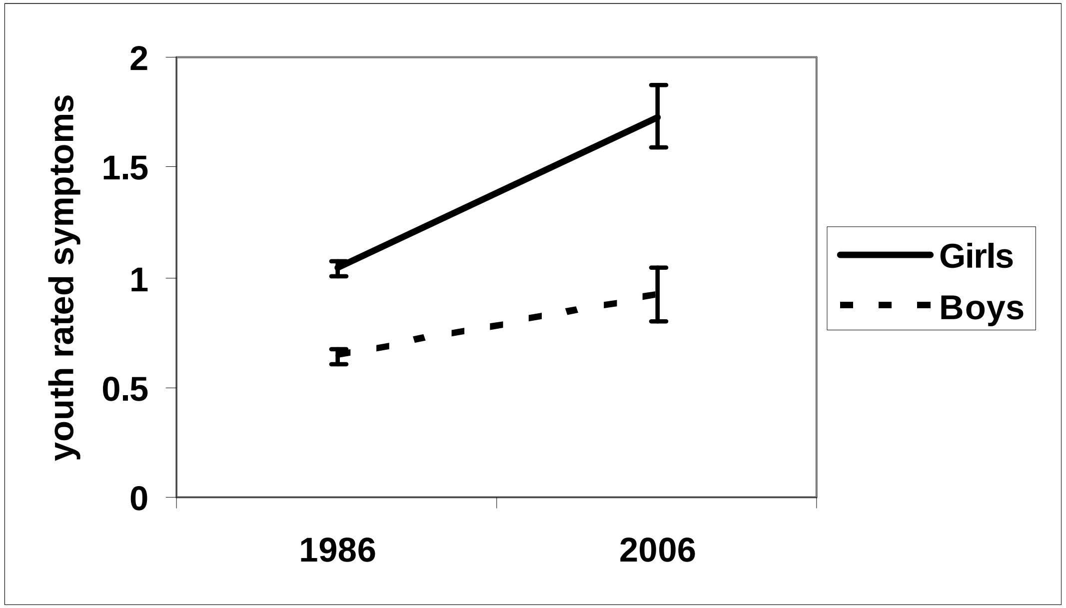
- Surveyed in 2006 with same questions and scales as BCS 1970 birth cohort 1986 survey

- 715 adolescents and 737 parents (86% mothers, 14% fathers) responded to the 2006 survey

- Weighting to make comparable to general population

Trends based on identical questions

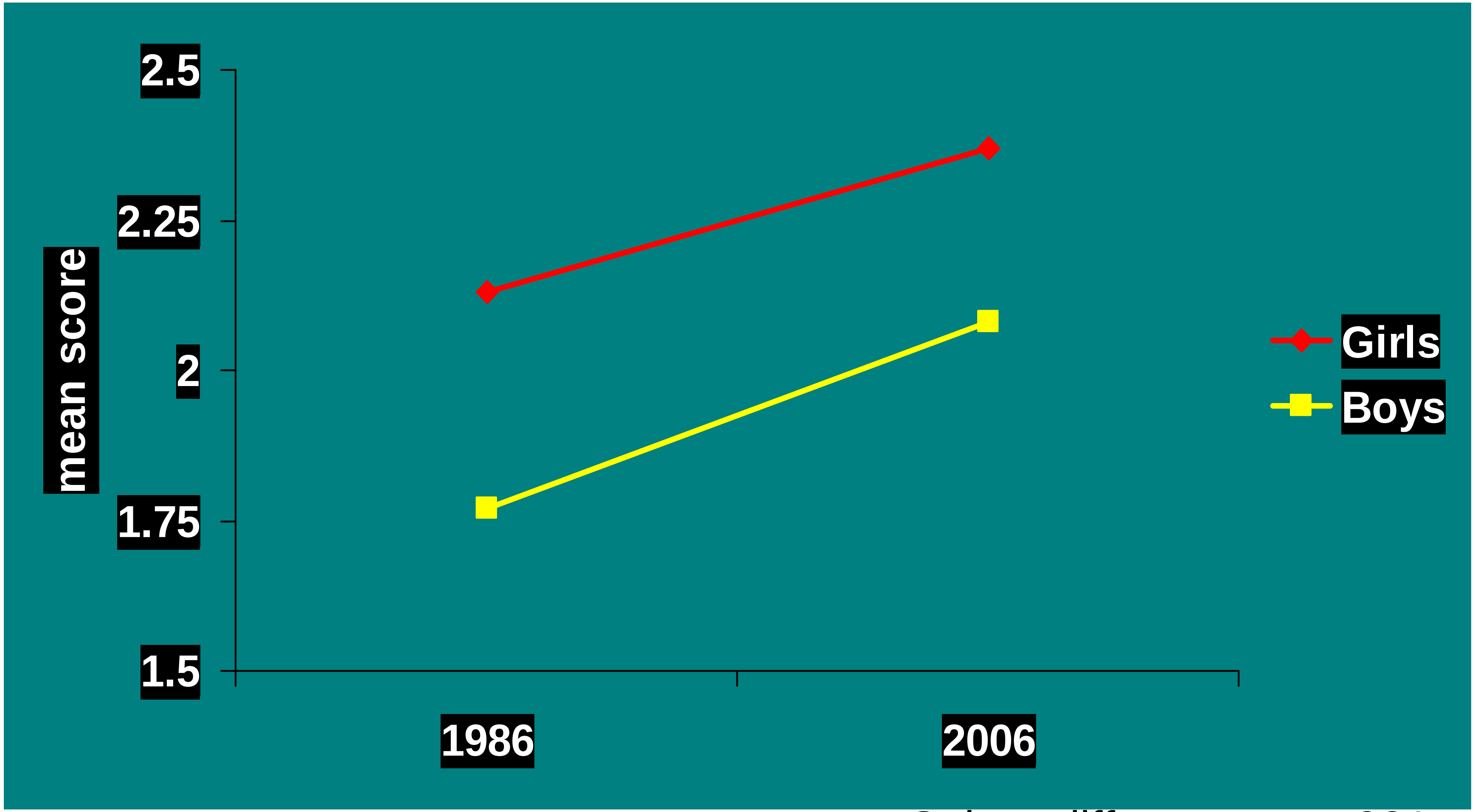




Parent

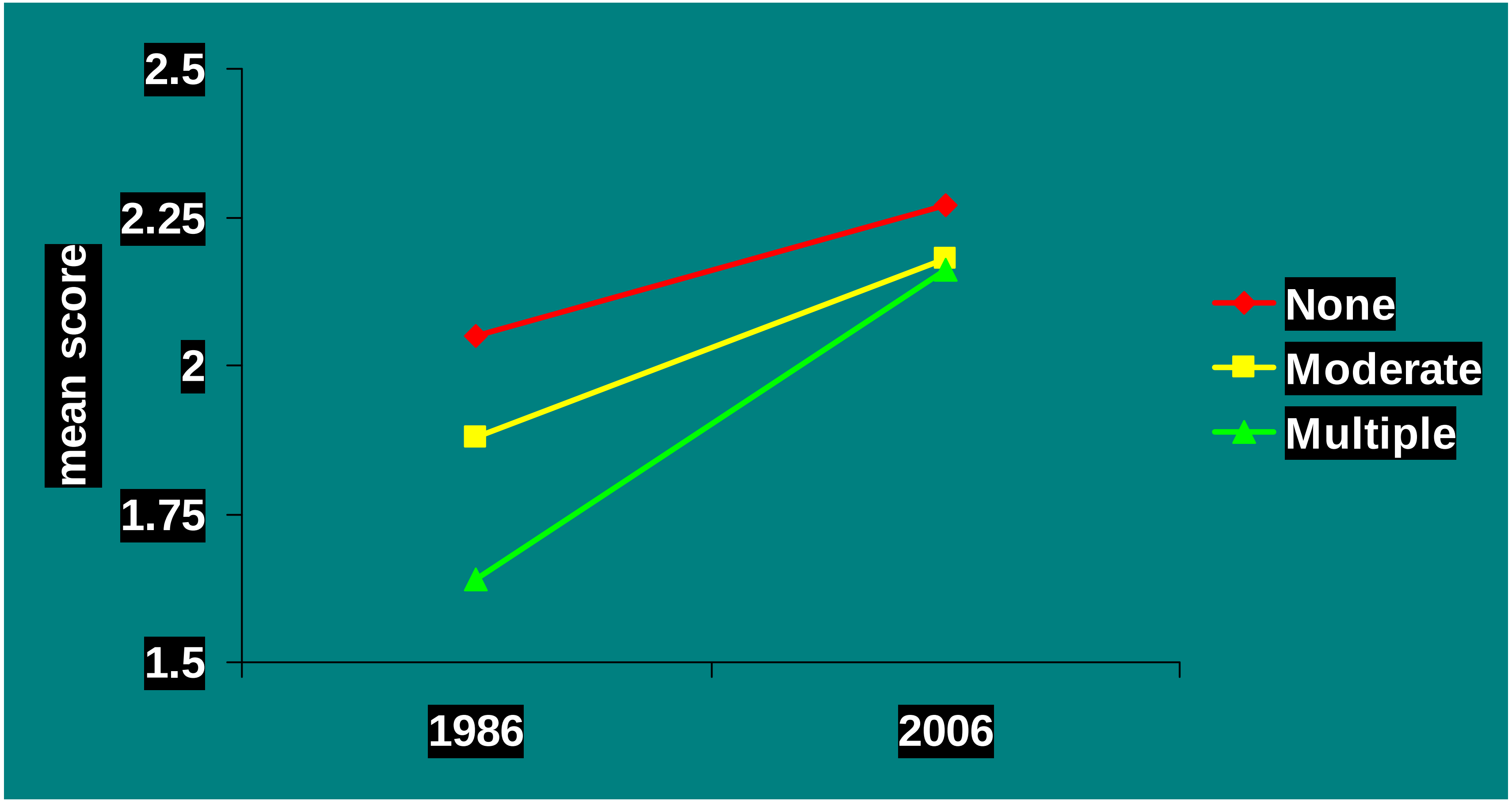
Self Report

Time trends in parental monitoring

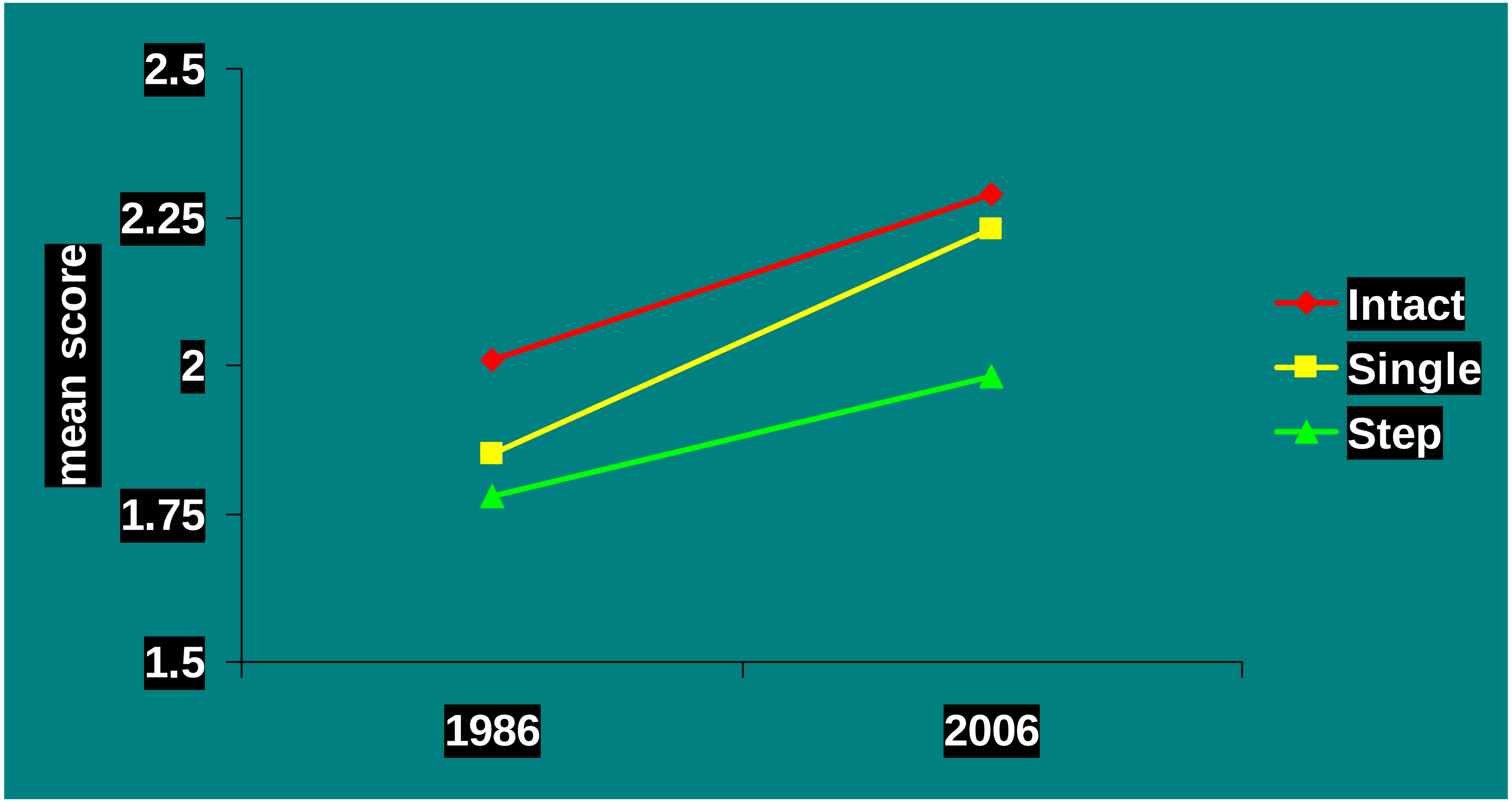


Cohort difference p < .001

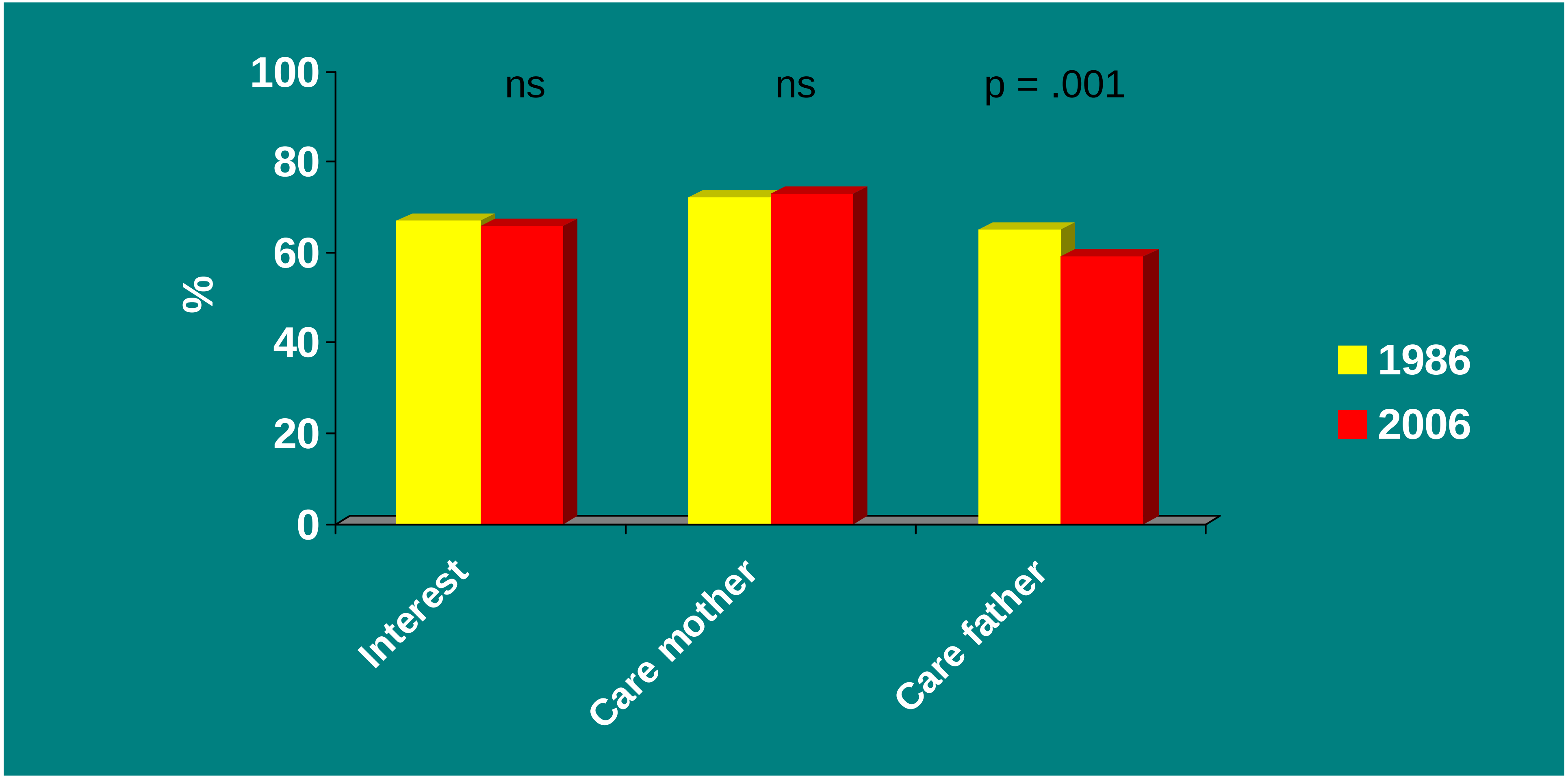
Monitoring over time by level of social disadvantage



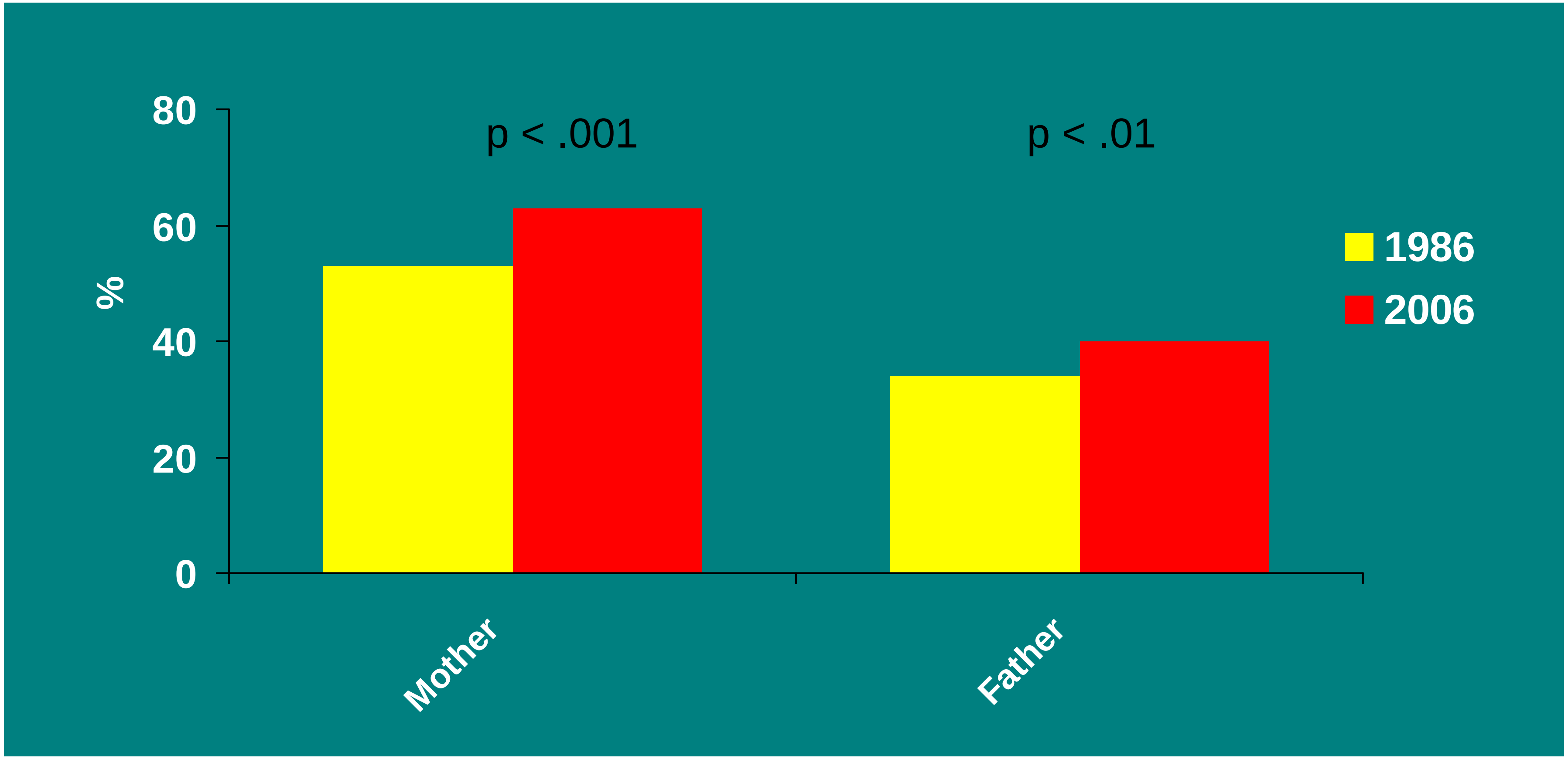
Monitoring over time by family type



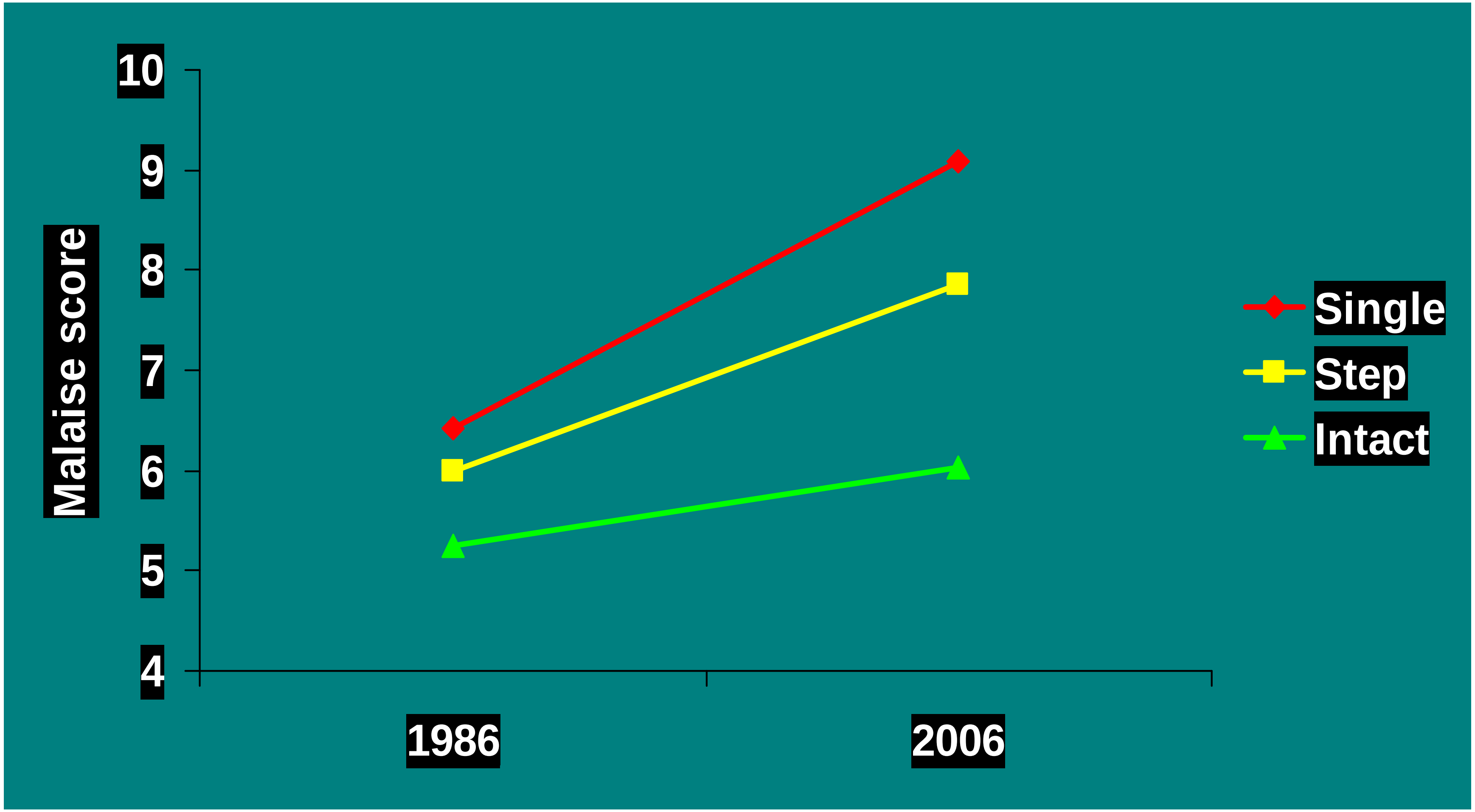
Parental interest and child cares what parents thinks



Quality time with parents (some/most days)



Trends in parental mental health



Main effects, cohort and family type, p < .001; Interaction, p < .01

Literature

Biostatistics (2009), **10**, 2, *pp.* 335–351
doi:10.1093/biostatistics/kxn041
Advance Access publication on November 27, 2008

Bayesian graphical models for regression on multiple data sets with different variables

C. H. JACKSON*
MRC Biostatistics Unit, Institute of Public Health, Forvie Site, Robinson Way, Cambridge CB2 0SR, UK
chris.jackson@mrc-bsu.cam.ac.uk

N. G. BEST, S. RICHARDSON
Department of Epidemiology and Public Health, Imperial College London, London, UK

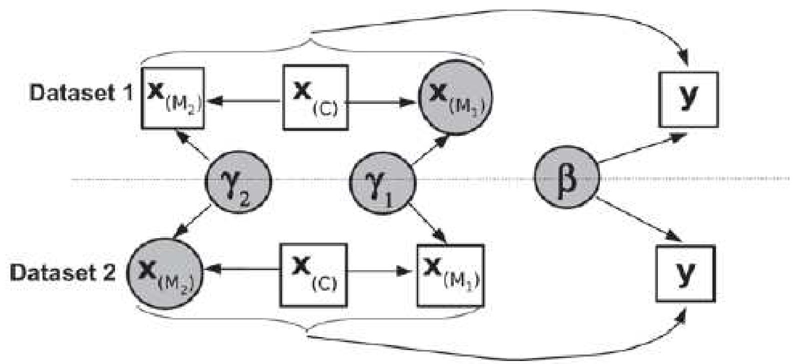


Fig. 1. General model for regression of $\mathbf{y}$ on $\mathbf{x}$ using a combination of data sets with different observed covariates. Circles represent unknown quantities and squares represent observed data. Covariates $\mathbf{x}_{(M_1)}$ missing in data set 1 are predicted from a regression fitted using the observed values of $\mathbf{x}_{(M_1)}$ in data set 2 and variables $\mathbf{x}_{(C)}$ common to both. Covariates $\mathbf{x}_{(M_2)}$ missing in data set 2 are predicted in a similar way using information from data set 1.

Developmental Psychology
2008, Vol. 44, No. 2, 365–380

0012-1649/08/$12.00 DOI: 10.1037/0012-1649.44.2.365

Pooling Data From Multiple Longitudinal Studies: The Role of Item Response Theory in Integrative Data Analysis

Patrick J. Curran, Andrea M. Hussong, Li Cai, and Wenjing Huang
University of North Carolina at Chapel Hill

Laurie Chassin
Arizona State University

Kenneth J. Sher
University of Missouri at Columbia and the Midwest Alcoholism Research Center

Robert A. Zucker
University of Michigan

Table 1
Item Content, Scale Source, Study Source, Proportion Endorsed, and Item Response Theory (IRT) Item Parameters for the 21 Internalizing Symptomatology Items

Item content summary	Scale(s)	Study source(s)	Proportion endorsed	IRT discrimination	IRT severity
1. Hopeless about future	BSI	MLS/AHBP	.16	2.19	1.26
2. Scared for no reason	BSI	MLS/AHBP	.08	2.00	1.86
3. Blue	BSI	MLS/AHBP	.34	1.92	0.55
4. No interest in things	BSI	MLS/AHBP	.24	1.65	1.00
5. Terror/panic	BSI	MLS/AHBP	.03	2.09	2.43
6. Restless	BSI	MLS/AHBP	.21	1.15	1.41
7. Cries a lot	CBCL	MLS/AFDP	.17	1.41	1.47
8. Might think/do something bad	CBCL	MLS	.17	1.31	1.56
9. Have to be perfect	CBCL	MLS/AFDP	.36	1.06	0.64
10. No one loves me	CBCL	MLS/AFDP	.10	2.70	1.53
11. Feel guilty	CBCL	MLS/AFDP	.15	1.45	1.60
12. Unhappy/sad/depressed	CBCL	MLS/AFDP	.25	2.96	0.79
13. Worried	CBCL	MLS/AFDP	.35	1.83	0.52
14. Others out to get me	CBCL	MLS	.11	1.65	1.79
15. Suspicious	CBCL	MLS	.39	1.21	0.49
16. Lonely	CBCL/BSI	MLS/AFDP/AHBP	.32	2.13	0.60
17. Worthless/inferior	CBCL/BSI	MLS/AFDP/AHBP	.13	3.26	1.27
18. Nervous/tense	CBCL/BSI	MLS/AFDP/AHBP	.44	1.52	0.22
19. Fearful/anxious	CBCL/BSI	MLS/AFDP/AHBP	.19	1.80	1.21
20. Self-conscious/easily embarrassed	CBCL/BSI	MLS/AHBP	.39	1.51	0.43
21. Thinks about killing self	CBCL/BSI	MLS/AHBP	.06	2.04	2.07

Note. Total sample size for proportion endorsed and IRT parameters, N = 1,827; for the MLS, N = 512; for the AFDP, N = 830; for the AHBP, N = 485. For items 16 through 19, AFDP items were drawn from the CBCL, AHBP items were drawn from the BSI, and MLS items were drawn from both the CBCL and BSI. BSI = Brief Symptom Inventory (Derogatis & Spencer, 1982);MLS = Michigan Longitudinal Study; AHBP = Alcohol and Health Behavior Project; AFDP = Adolescent/Adult Family Development Project; CBCL = Child Behavior Checklist (Achenbach & Edelbrock, 1981).

Psychological Methods
2009, Vol. 14, No. 2, 126–149

1082-989X/09/$12.00 DOI: 10.1037/a0015857

Modeling Life-Span Growth Curves of Cognition Using Longitudinal Data With Multiple Samples and Changing Scales of Measurement

John J. McArdle
University of Southern California

Kevin J. Grimm
University of California, Davis

Fumiaki Hamagami
University of Virginia

Ryan P. Bowles
Michigan State University

Table 1
Summary of Available Data From Multiple Testing Occasions for Three Longitudinal Studies

Age (years)	Berkeley Growth (n = 61)	Guidance–Control (n = 206)	Bradway–McArdle (n = 111)
2–5½			SB-L, SB-M (111)
6	1916 SB (60)	1916 SB (205)	
7	1916 SB (47), SB-L (8)	1916 SB (204)	
8	SB-L (51)	SB-L (187)	
9	SB-L (53)	SB-L (94), SB-M (98)	
10	SB-M (53)	SB-L (102), SB-M (88)	
11	SB-L (48)	SB-L (77)	
12	SB-M (50)	SB-L (90), SB-M (43)	
13–14	SB-L (42)	SB-L (82), SB-M (97)	SB-L (111)
15		SB-M (51)	
16	WB-I (48)		
17	SB-M (44)		
18	WB-I (41)	WB-I (157)	
21	WB-I (37)		
25	WB-I (25)		
29			WAIS, SB-L (110)
36	WAIS (54)		
40		WAIS (156)	WAIS, SB-LM (48)
53	WAIS-R (41)	WAIS-R (118)	WAIS (53)
63			WAIS, WJ-R (48)
67			WAIS, WJ-R (33)
72	WAIS-R, WJ-R (31)		

Note. Available sample size for specific tests is contained in parentheses. SB-L, SB-M, SB-LM = Stanford–Binet Forms L, M, and LM; WB-I = Wechsler–Bellevue Intelligence Scale Form I; WAIS = Wechsler Adult Intelligence Scale; WAIS-R = Wechsler Adult Intelligence Scale–Revised; WJ-R = Woodcock–Johnson Psycho-Educational Battery–Revised.

Collaborators, references & funders

Collishaw, S., Maughan, B., Goodman, R. & Pickles, A. (2004) Time trends in adolescent mental health. *Journal of Child Psychology and Psychiatry. 45, 1350-1362.*

Goodman, R. Iervolino, A., Collishaw, S., Pickles, A. & Maughan, B. (2007) Seemingly minor changes to a questionnaire can make a big difference to mean scores: a cautionary tale. *Social Psychiatry and Psychiatric Epidemiology. 42(4):322-327*

Collishaw, S., Goodman, R., Pickles, A. & Maughan, B. (2007) Modelling the contribution of changes in family life to time trends in adolescent conduct problems*. Social Science & Medicine* 65, 2576-87

Collishaw, S., Maughan, B., Natarajan, L. & Pickles A (2010) Trends in adolescent emotional problems in England: a comparison of two national cohorts twenty years apart. Journal of Child Psychology and Psychiatry, 51, 885-994.

Conti-Ramsden G, St Clair MC, Pickles A, Durkin K. (2012) Developmental Trajectories of Verbal and Nonverbal Skills in Individuals with a History of SLI: From Childhood to Adolescence. J Speech Lang Hear Res. 2012 May 4. [Epub ahead of print]

Funded by Nuffield Foundation, MRC G0802307 and ESRC