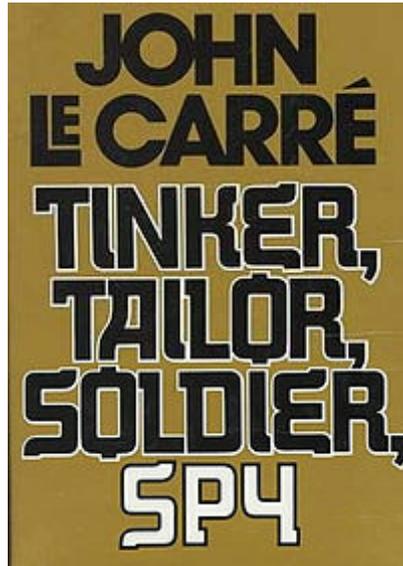


Tinker, Tailor, Soldier, Spy -

What is the relationship between parental occupations and filial educational attainment?



Professor Vernon Gayle (University of Stirling / ISER Essex University) &
Dr Paul Lambert (University of Stirling)

9th December 2009, Stirling

BEWARE!

WORK IN PROGRESS

Mainly exploratory analyses

The title is a play on the spy novel by John le Carré first published in 1974.

Pseudonym of David John Moore Cornwell, who taught French and German at Eton before joining MI5.

The novel's title is from the children's nursery rhyme 'tinker, tailor, soldier, sailor, rich man, poor man, beggar man thief' which are the code names of the five men suspected as being the mole in MI5.

The BBC adapted the novel for TV in 1979 with Alec Guinness playing the lead.

As a small boy I watched it with my legs crossed for an entire episode!

The overall motivation...

- Undertake a piece of work that locates within a wider sociological perspectives on social stratification
 - ‘inter’-generational relationships
- A long running empirical research theme within the sociology of education and the sociology of youth has been the relationship between parental social background (often measured by parental occupation) and educational attainment
- Historically the weight of evidence has indicated that educational attainment is stratified - typically, those from more advantaged social backgrounds generally achieve higher levels of attainment than their counterparts from less advantaged backgrounds
- An attempt to use some recently harmonised large-scale survey data
 - augment analyses with additional measures
- Begin to examine the effects of missing data (item non-response)
- Start to think about future analyses

Structure of Talk

- The sociology of education
 - Introduction to GCSEs
- The national data
- Youth Cohort Study of England & Wales
- Descriptive results
 - GCSE attainment measures
 - Parental occupational measures
- Modelling Results
 - A closer look at parental occupations
- Missing data & multiple imputations
 - A first attempt
- Conclusions
- Further work (if time)

General Certificate of Education

- General Certificate of Secondary Education (GCSE) introduced in the late 1980s
- The standard qualification for pupils in England and Wales in year 11 (aged 15/16)
- Usually a mixture of assessed coursework and examinations
- Generally each subject is assessed separately and a subject specific GCSE awarded
- It is usual for pupils to study for about nine subjects, which will include core subjects (e.g. English, Maths and Science) and non-core subjects
- GCSEs are graded in discrete ordered categories
- The highest being A*, followed by grades A through to G (A* from 1994)
- Arran Fernandez gained A* in Maths at age 8 !

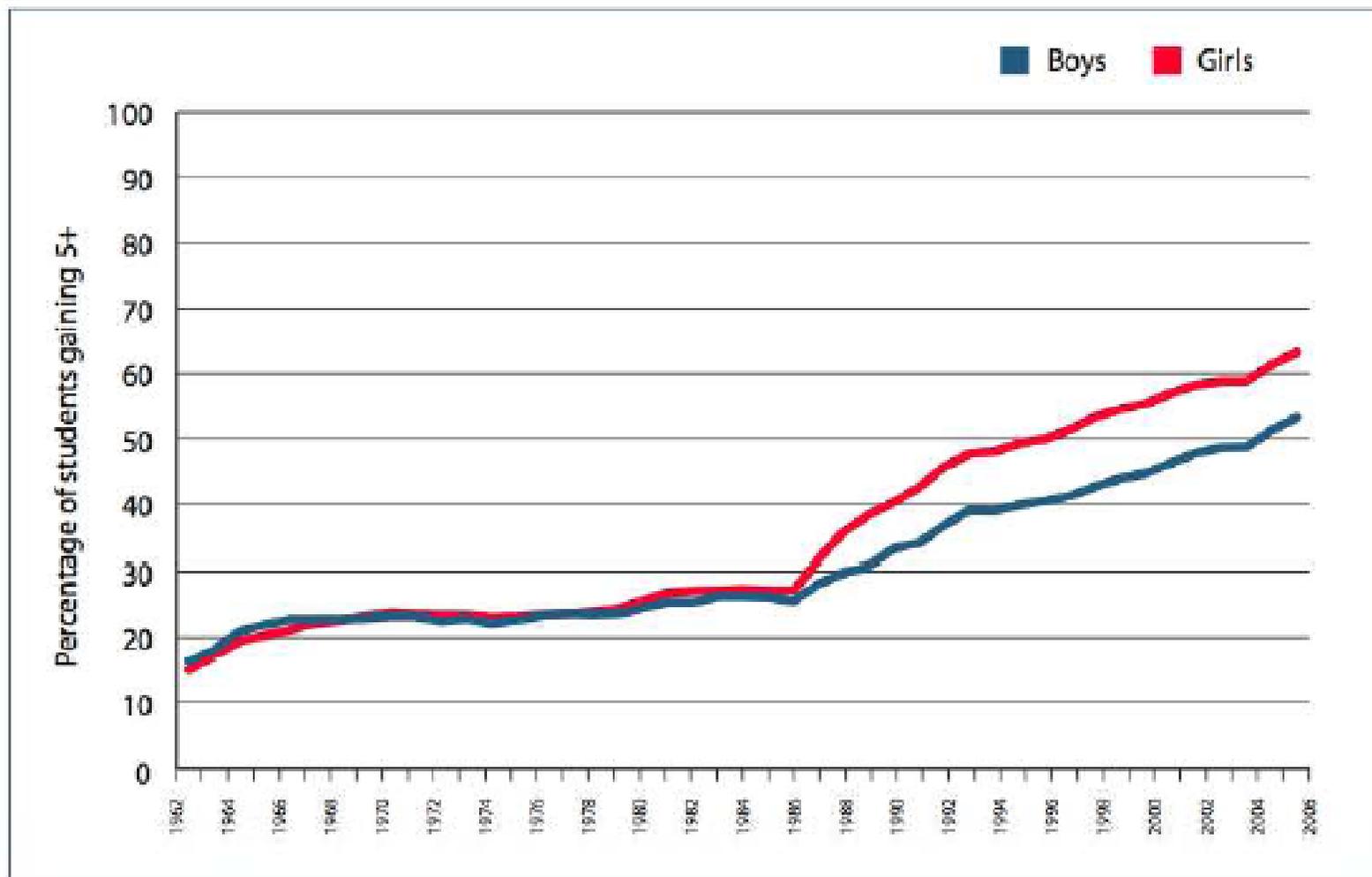
General Certificate of Education

- The *Education Reform Act 1988* led to rapid changes in the secondary school curriculum, and to the organisation, management and financing of schools
- A major change for pupils was the introduction of the General Certificate of Secondary Education (GCSE)
- GCSEs differed from the qualifications that they replaced
 - A new grading scheme was established and all pupils were entered for a common set of examinations
 - There were also changes in the content and format of examinations and assessment by coursework was introduced
- School league tables were published
- A newsworthy item each summer
 - Previously only teachers, parents and pupils knew when exam day was

Why explore GCSE attainment?

- GCSEs are public examinations and mark the first major branching point in a young person's educational career
- Poor GCSE attainment is a considerable obstacle which precludes young people from pursuing more advanced educational courses
- Young people with low levels of GCSE attainment are usually more likely to leave education at the minimum school leaving age and their qualification level frequently disadvantages them in the labour market
- Low levels of qualifications are also likely to have a longer term impact on experiences in the adult labour market
- Therefore, I argue that gaps in GCSE attainment are sociologically important

Figure 3-7 Percentage of School Leavers Achieving 5+ A-C (or Pass) O-level or A*-C GCSE by Gender (1962–2006)



Source: Department for Education Statistics of Education

Department for Education and Skills (2007) *Gender and Education: The evidence on pupils in England*

Year	All pupils %	Boys %	Girls %
1990	34.5	30.8	38.4
1991	36.8	33.3	40.3
1992	38.3	34.1	42.7
1993	41.2	36.8	45.8
1994	43.3	39.1	47.8
1995	43.5	39.0	48.1
1996	44.5	39.9	49.4
1997	45.1	40.5	50.0
1998	46.3	41.3	51.5
1999	47.9	42.8	53.4
2000	49.2	44.0	54.6
2001	50.0	44.8	55.4
2002	51.6	46.4	57.0
2003	52.9	47.9	58.2
2004	53.7	48.4	58.4
2005	56.3	51.4	61.4
2006	58.5	53.8	63.5
2007	60.3	55.8	65.0
2008	65.3	60.9	69.9

National Figures, % of pupils gaining
5+ GCSEs (grades A*-C)

Percentage aged 15 on roll in all English schools

Figures 1998 onwards are for GCSE & GNVQ equivalent

Data Sources DfE&E; DfES; DfCSF

Table 1 <http://www.bstubbs.co.uk/5a-c.htm#table1>

Youth Cohort Study of England & Wales (YCS)

- Major Longitudinal Study began Mid-1980s
- Designed to monitor behaviour of young people as they reach the minimum school leaving age and either stay on in education or enter the labour market
- Experiences of Education (qualifications); Employment; Training; Aspirations; Family; Personal characteristics & circumstances
- Nationally representative; Large sample size; Panel data (albeit short); Possible to compare cohorts (trends over time)
- Study contacts a sample from an academic year group (cohort) in the spring following completion of compulsory education
- The sample is designed to be representative of all Year 11 pupils in England & Wales
- Sample are tracked for 3 (sometimes 4) waves (called Sweeps) of data collection
- Growing up in the 1990s the GCSE era; Partly fills the gap left by the missing 198(2) birth cohort

Working with the YCS

- Documentation is very poor especially in the older cohorts – usually handwritten annotation on questionnaires (pdf) (*Compare this with the BHPS for example*)
- Changes in qualifications, educational policy etc adds data complications
- Changes in questions, measures, coding, timing etc, all add to the general confusion
- Recently available harmonized dataset

SN 5765 Title: *Youth Cohort Time Series for England, Wales and Scotland, 1984-2002*

Depositor(s):

Croxford, L., University of Edinburgh. Centre for Educational Sociology

Principal Investigator(s):

*Croxford, L., University of Edinburgh. Centre for Educational Sociology
Iannelli, C., University of Edinburgh. Centre for Educational Sociology
Shapira, M., University of Edinburgh. Centre for Educational Sociology*

Economic and Social Research Council Grant Number: R000239852

Official and YCS Surveys % of pupils gaining 5+ GCSEs (grades A*-C)

Year	All Pupils	YCS Pupils	Boys Official	Boys YCS	Girls Official	Girls YCS
1990	34.5	35.1	30.8	31.6	38.4	38.7
1993	41.2	42.3	36.8	38.1	45.8	46.7
1995	43.5	44.9	39.0	40.6	48.1	49.3
1997	45.1	46.2	40.5	41.9	50.0	50.6
1999	47.9	51.0	42.8	46.4	53.4	55.7

YCS Data Source: Dataset SN5765 (weighted data)

Why parental occupation

- Occupations is a key measure of social stratification
- Maps onto wider sociological conception of social class
- Why not income or wealth?
 - 16/17 year olds are being questioned
 - fluctuation in income and wealth
 - parents' location on the age/income distribution
- Occupation is a proxy
 - lifetime income
 - life chances (and opportunities)
 - life style & consumption patterns
 - (even correlates with health)

A proxy for income?

In this respect, we would argue that the use of socio-economic classifications in research is not simply to act as a proxy for income where income data themselves are unavailable. We use socio-economic classifications because they are measures designed to help us identify key forms of social relations to which income is merely epiphenomenal... It is also the case that socio-economic classifications are relatively more general and stable measures than income. Income is well known to fluctuate over the lifecycle; indeed panel data regularly reveals a high level of 'income churning' from year to year (for the UK see Jarvis and Jenkins 1997). What socio-economic classifications might reasonably be expected to proxy is the lifecycle/earnings profile.

(Rose and Pevalin 2003) *A Researcher's Guide to the National Statistics Socio-economic Classification*

Which measure of occupation?

Forty years ago, Bechhofer's review of the use of occupational information in sociology bemoaned the abundance of, and inconsistencies between, occupationally based social classifications, noting that *"..researchers are advised not to add to the already existing plethora of classifications without very good reason"* (1969 p.118)

However since that recommendation, the number of new classifications has increased steadily

We argue for the transparent use of classifications that have 'agreed' standards of measurement and can therefore be replicated and compared within and across analyses

- Various (unsystematic) parental occupation measures deposited with individual YCS cohorts
- NS-SEC (8 and 3 category) deposited with **SN 5765**
- RGSC, CAMSIS & Elias not in **SN 5765**
 - Derived from data using GEODE Resources
 - www.geode.stir.ac.uk
 - www.dames.org.uk/
- Simple dominance method
 - common in stratification research
 - father or mother whichever is higher on the relevant scheme
 - nurse mum and consultant dad = dad
 - nurse mum and hospital porter dad = mum

Brief illustration of parental occupational measures

- National Statistics Socio-economic Classification (NS-SEC)
 - Official classification
 - From 2001 used in all official statistics
 - Some Govt statisticians only speak NS-SEC
 - Various versions – below is 8 category
 - Not ordered categories !

1.1 Large employers & higher professionals

1.2 Higher Professionals

2 Lower managerial & professional

3 Intermediate occupations

4 Small employers & own account

5 Lower supervisory & technical

6 Semi-routine occupations

7 Routine occupations

Brief illustration of parental occupational measures

- Registrar General's Social Class (RGSC)
 - Official classification until 2001
 - Ordered categorical measure

Professional (I)

Intermediate (II)

Skilled Non-Manual (III_n)

Skilled Manual (III_m)

Partly Skilled (IV)

Unskilled (V)

Brief illustration of parental occupational measures

- CAMSIS: Social Interaction and Stratification Scale
 - www.camsis.stir.ac.uk/
 - Scale empirically derived from patterns of social interaction

Continuous scale with mean=50 s.d.=15

Brief illustration of parental occupational measures

- Elias' Skill Classification

Managerial & Professional

Middle

Working

Elementary

Simplified Illustration*

Job	NS-SEC	RGSC	CAMSIS (male)
Company Director Marketing (25+ employees)	Large employer & higher managers	Intermediate (II)	66
Judge	Higher professional	Professional (I)	86
Nurse	Lower professional	Intermediate (II)	52 (59 female)
Draughtsperson	Intermediate	Skilled Non-Manual (III _n)	59
Bricklayer (self employed)	Small employer & own account	Skilled Manual (III _m)	37
Painter & decorator (supervisor)	Lower supervisory & technical	Skilled Manual (III _m)	40
Tyre fitter	Semi-routine	Partly Skilled (IV)	42
Road Sweeper	Routine	Unskilled (V)	32

* Does not include Elias' skill measure

YCS Data

- YCS cohorts
 - *School leaving years 1990, 1993, 1995, 1997, 1999*
- Comprehensive school pupils
 - *Free schooling*
 - *No educational selection*
- Complete information on parental occupation and other measures (n=55120)

GCSE Outcome Measures

- 5+ GCSEs grades A*-C
 - Recognised official bench mark
 - Frequently used outcome measure in research
 - School league table measure
 - This measure is still published annually by The Department for Children, Schools and Families (see <http://www.dcsf.gov.uk/performance/tables/>)
 - Government target is 53% with 5+A*-C including Maths & English by 2011
- Number of GCSEs grades A*-C
- GCSE score (A/A*=7; G=1)
 - Capped at 84 point 12 GCSEs Grade A/A*
- Standardized GCSE score (A/A*=7; G=1)

Descriptive Results

(Table 1)

- Overall trend
 - Increasing proportions getting 5+GCSEs (A*-C)
 - Increasing mean number of A*-C grade GCSEs
 - Increasing mean GCSE points score
- Gender
 - Female pupils outperforming male pupils
- Ethnicity
 - Some groups doing better than white pupils (e.g. Indians)
 - Other groups doing worse (e.g. blacks)
- Parental Occupation
 - Observable gradient
 - Lower levels of GCSE attainment from those pupils with less occupationally advantaged parents

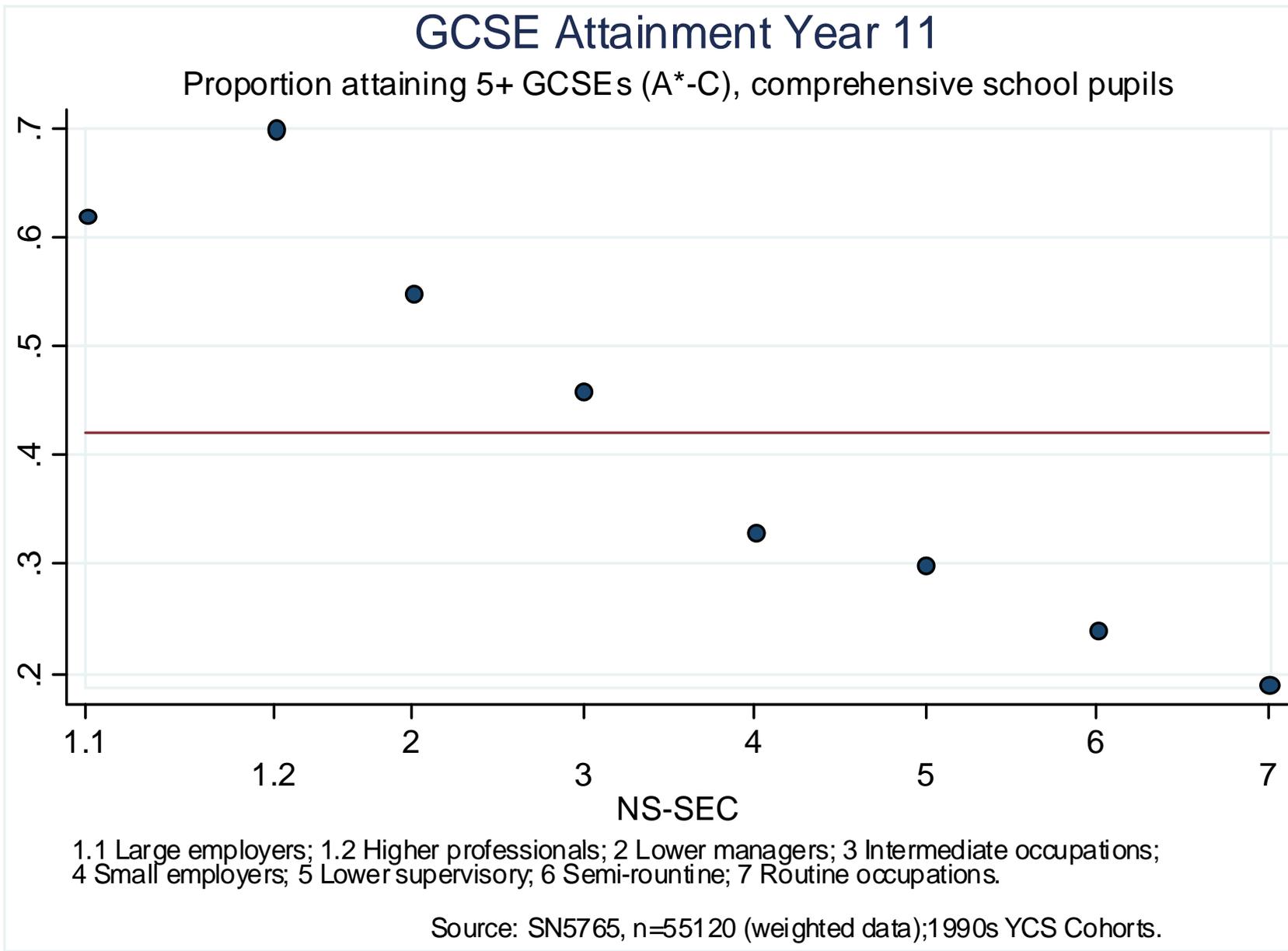
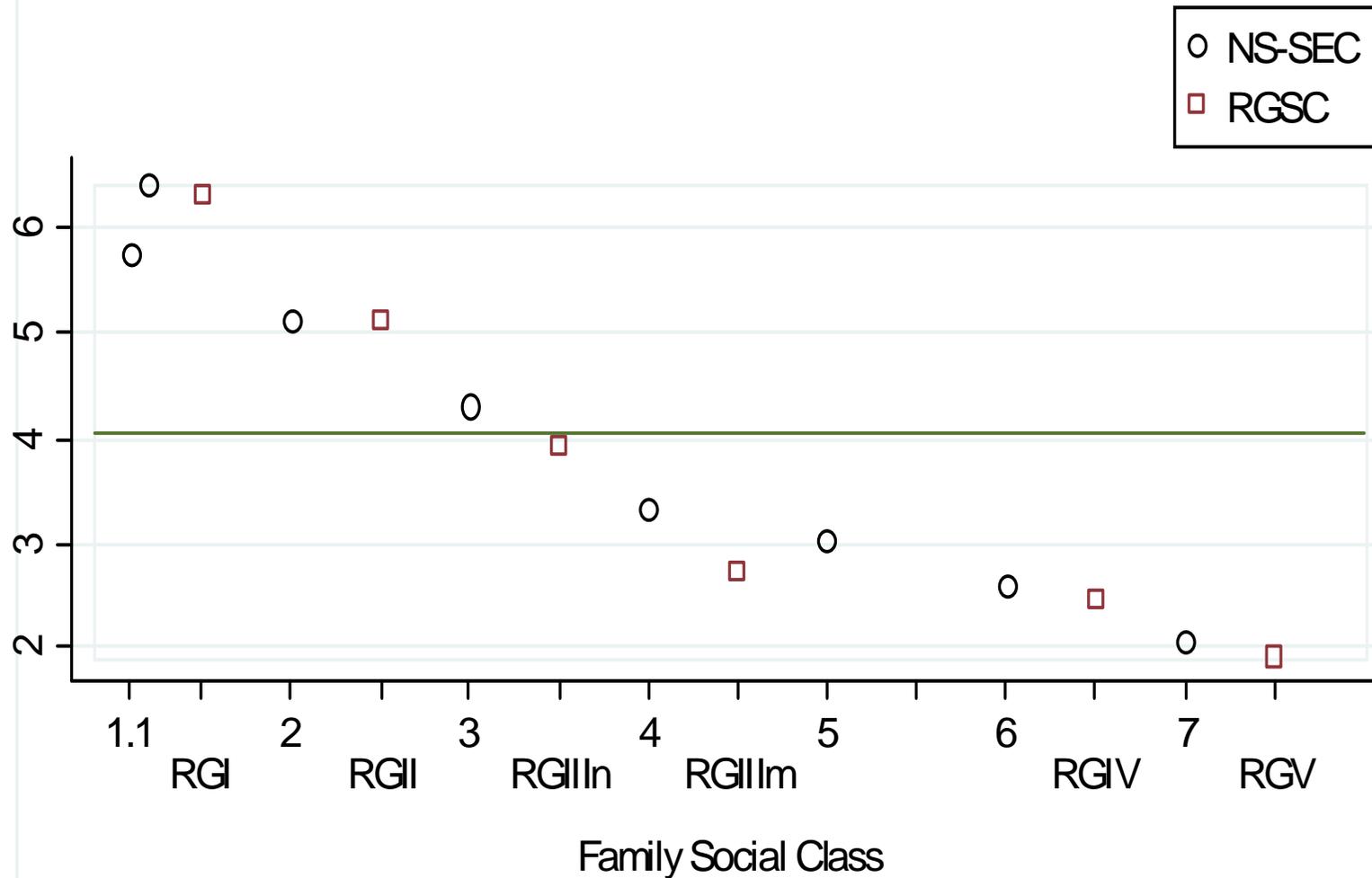


Figure 1

GCSE Attainment Year 11

Mean number of GCSEs (A*-C), comprehensive school pupils



1.1 Large employers; 1.2 Higher professionals; 2 Lower managers; 3 Intermediate occupations; 4 Small employers; 5 Lower supervisory; 6 Semi-routine; 7 Routine occupations.

RGI Professional; RGI Man & Tech; RGIIn Skilled Non-Manual; RGIIm Skilled Manual; RGIIV Partly Skilled; RGI Unskilled

Source: SN5765, n=55120 (weighted data); 1990s YCS Cohorts.

Figure 3

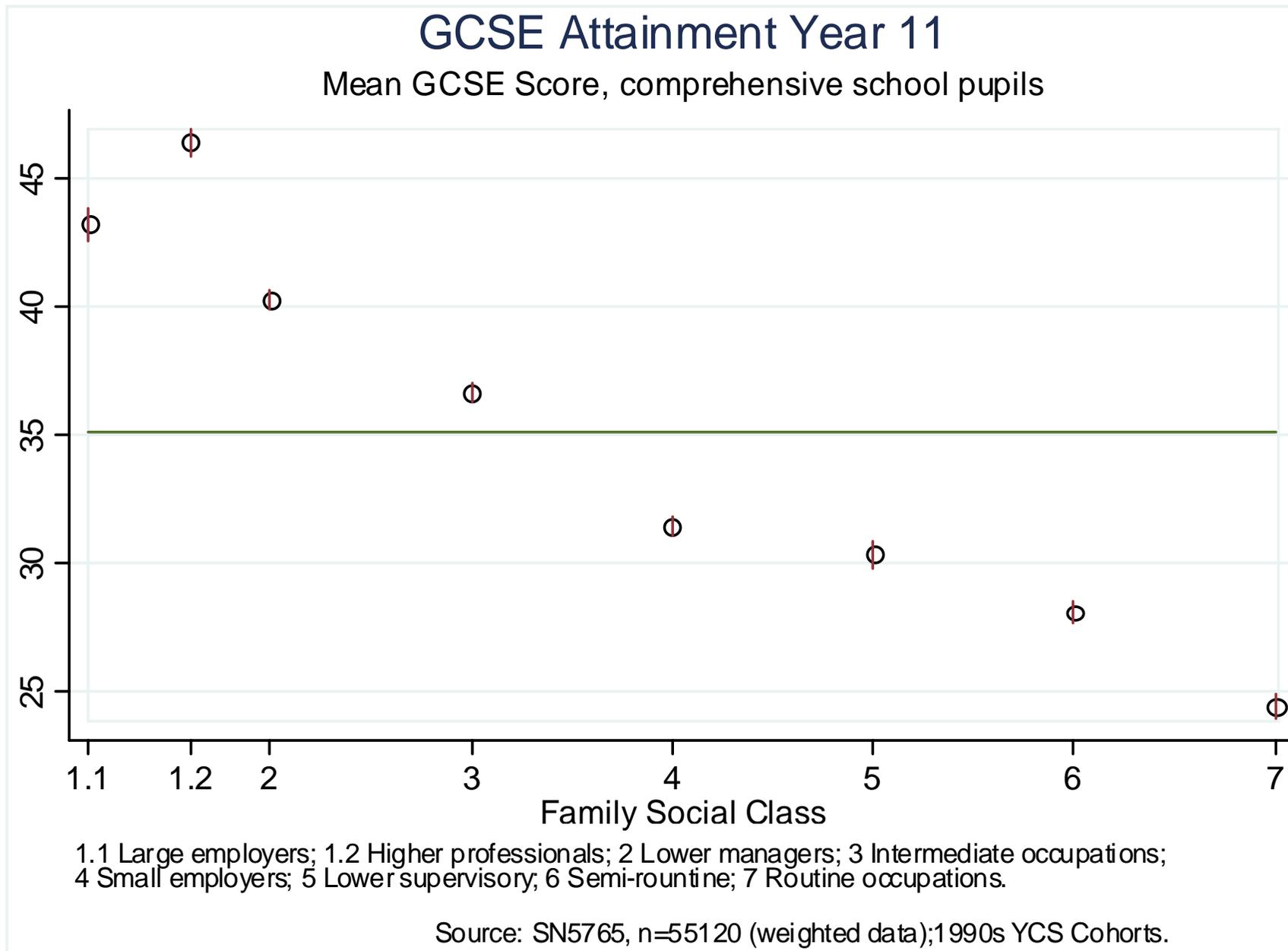


Figure 4

Modelling Results

(Tables 2 - 5)

Models (survey regression models)

- Logit 5+GCSEs (A*-C)
- Negative binomial (count) number of A*-C grade GCSEs
- Regression GCSE points score
- Regression GCSE points score (standardised across cohorts mean=0 s.d.=1)

- Gender
 - Female pupils outperforming male pupils
- Ethnicity
 - Some groups doing better than white pupils (e.g. Indians)
 - Other groups doing worse (e.g. blacks)

- Parental Occupation
 - Lower levels of GCSE attainment from those pupils with less occupationally advantaged parents whatever measure²⁸

Survey Regression model Year 11 GCSE score (all 1990s Cohorts)

Model fit summary statistics

	Adjusted R ²	BIC
Cohort + gender + ethnicity	.08	465195
+ NS-SEC 8 Cat	.20	457438
+ NS-SEC 3 Cat	.18	458613
+ RGSC	.19	457986
+ Camsis	.21	456761
+ Elias' Skill	.17	459082

Exploring at Occupational Unit Group (OUG) Level

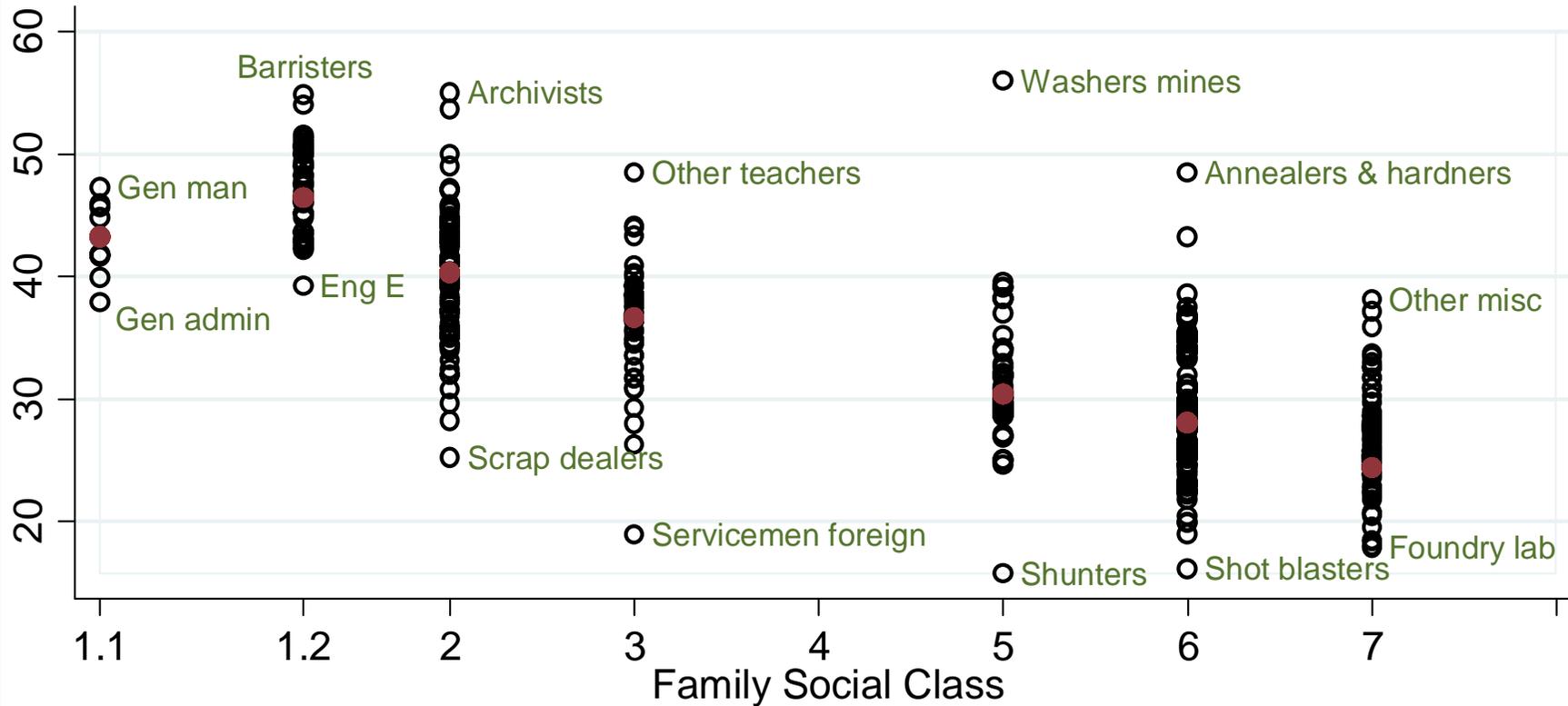
NS-SEC	No. of SOC90 Occupations*
1.1 Large Employers and higher managers	10
1.2 Higher professional occupations	38
2 Lower managerial and professional occupations	78
3 Intermediate occupations	42
5 Lower supervisory and technical occupations	41
6 Semi-routine occupations	88
7 Routine occupations	74
Total	371

* *Employees*

Possible interesting variations within NS-SEC categories

GCSE Attainment Year 11

Mean GCSE Score, Parents' SOC90



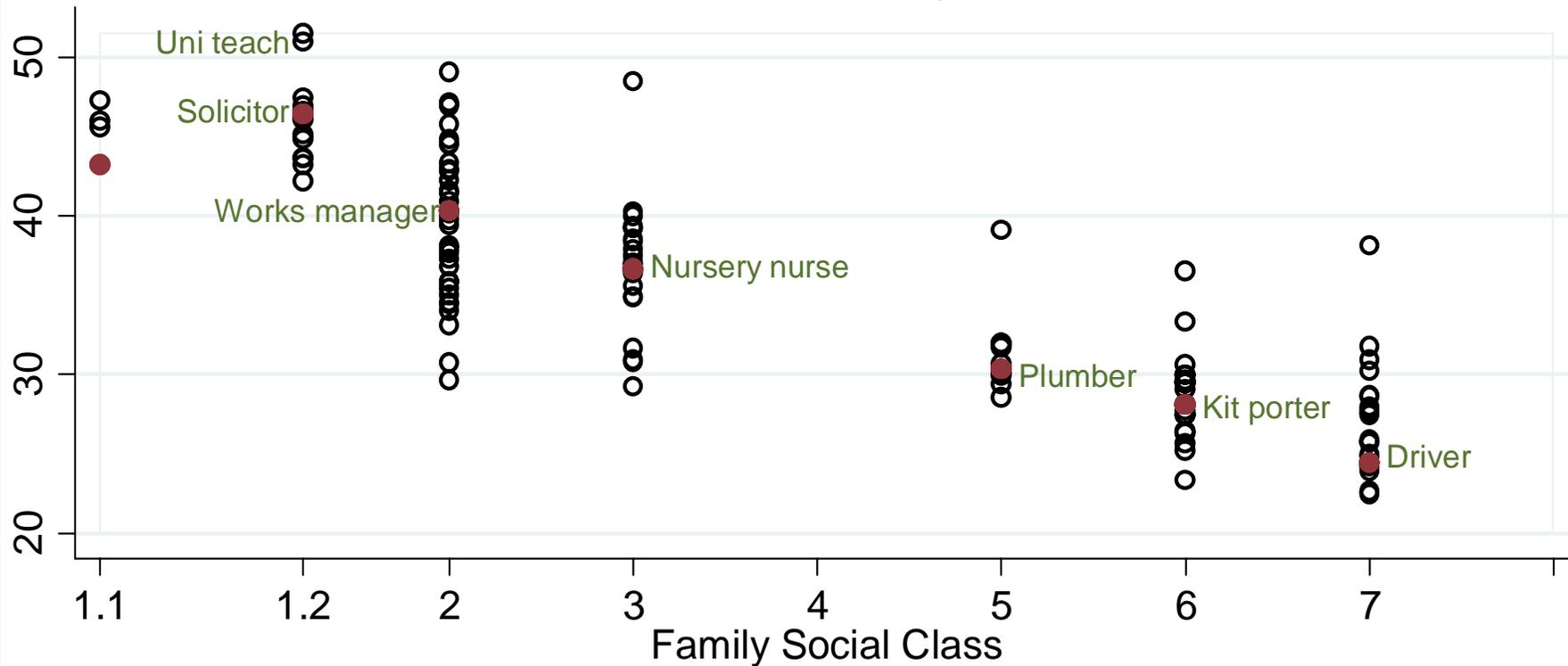
● Mean for NS-SEC Class

Source: 1990s YCS Cohorts; Comprehensive school pupils.
 369 SOC's; Pupils per SOC Mean 149; Min 1; Max 1836 (Nurses).

Figure 5

GCSE Attainment Year 11

Mean GCSE Score, Parents' SOC90 (large SOC groups)
Illustrations of occupations



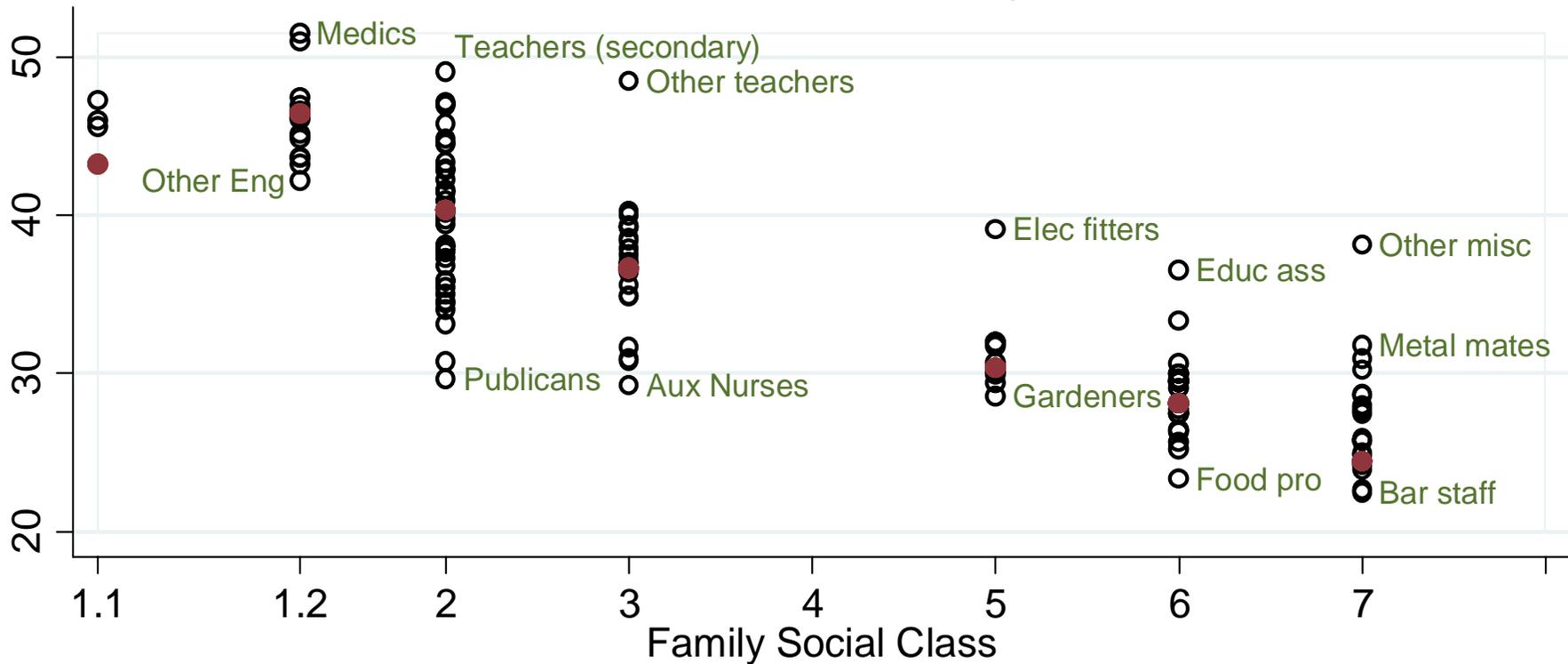
● Mean for NS-SEC Class

Source: 1990s YCS Cohorts; Comprehensive school pupils.
121 larger SOC's; Pupils per SOC Mean 380; Min 101; Max 1836 (Nurses).

Figure 6

GCSE Attainment Year 11

Mean GCSE Score, Parents' SOC90 (large SOC groups)
Illustrations of extreme occupations



● Mean for NS-SEC Class

Source: 1990s YCS Cohorts; Comprehensive school pupils.
121 larger SOC's; Pupils per SOC Mean 380; Min 101; Max 1836 (Nurses).

Figure 7

Parents' Dominant SOC90	Count	Prop 5+	No. A*-C	GCSE Score	Earnings	Earnings	% Grads	% Grad
		A*-C	GCSEs		Band Male	Band Female	Male	Female
219. Other engineers and technologists n.e.c.	387	0.57	5.3	42.18	£350-£400	£300-£350	37	47
242. Solicitors public	180	0.66	6.3	46.59	£700+	£700+	94	97
220. Medical practitioners	263	0.81	7.4	51.52	£700+	£700+	92	86
175. Publicans, innkeepers and club stewards	222	0.25	2.8	29.64	£250-£300	£150-£200	4	1
110. Production, works and maintenance managers	1053	0.55	5.1	40.13	£350-£400	£300-£350	22	20
233. Secondary (and middle school deemed secondary) education teachin	1320	0.78	7.0	49.05	£450-£500	£350-£400	85	71

McKnight & Elias (1998) Guide to the 371 Database

Earnings distributions in SOC OUG 'quindeciles?'

Distribution of employment by (highest) qualification SOC OUG

Further thinking required here...

Item non-response

- Parental occupation information is an important factor in GCSE attainment
- Remember... in the YCS young people aged 17 being asked about their parents
- NS-SEC 7,835 pupils with missing data (12%)
- Promising results from a paper at RC33 Nobel *et al.* 2008 reported an experiment with LSYPE cohort testing pupils with YCS question and interviewing their parents

Item non-response

- *Over 60% of young people correctly reported their parents' occupation at 4 digit OUG (disappointingly only approx. 74% get it at 1 digit level)*
 - Unit group 2111 Chemists
 - Minor group 211 Natural scientists
 - Sub major group 21 Science and engineering professionals
 - Major group 2 Professional occupations
- *Nobel et al. 2008 reported no significant social class pattern to the errors (using NS-SEC)! ³⁶*

Missing data & Multiple imputation

	number missing
GCSE Score	1135
NS-SEC (3 Categories)	7835
Ethnicity	1064
Gender	0
Cohort	0

Patterns of Missing Values

Y	NS-SEC	Ethnicity	Missing Vars	n
+	+	+	0	55145
+	.	+	1	6821
.	+	+	1	697
+	.	.	2	592
+	+	.	1	352
.	.	+	2	318
.	.	.	3	104
.	+	.	2	16

Creation of Multiple Imputed Datasets

Imputation by chained equations (ice in Stata)

- Regression model GCSE Score
- Ologit model of NS-SEC (3 Categories)
- Mlogit model of Ethnicity

- 10 imputed datasets
 - 20 cycles of regression switching

Analysis of Multiple Imputed Datasets

- Stata **mim** prefix to combine the results from the 10 different imputed data sets into a single output
- The promise of a (combined) analysis of datasets with 64045 pupils rather than 55145 pupils (complete cases)

	Complete Cases Analysis		Model of Multiple Imputed Datasets		
	Beta	s.e.	Beta	s.e.	FMI
NS-SEC (3 Classes)					
Managerial & Professional	0.00		0.00		
Intermediate Occupations	-7.46	0.15	-7.67	0.15	0.09
Routine Occupations	-13.85	0.17	-14.51	0.17	0.10
Constant	39.70	0.19	39.05	0.18	0.06
<i>n</i>	55145		64045		

The promise of a (combined) analysis of datasets with 64045 pupils rather than 55145 pupils (complete cases)

Substantive inference on NS-SEC appears similar....

	Complete Cases Analysis		Model of Multiple Imputed Datasets		
	Beta	s.e.	Beta	s.e.	FMI
Whites	0.00		0.00		
Black	-5.62	0.57	-7.31	0.50	0.10
Indian	3.60	0.44	3.08	0.41	0.04
Pakistani	-1.89	0.58	-3.61	0.48	0.15
Bangladeshi	0.42	1.04	-3.44	0.75	0.21
Other Asians	5.36	0.68	4.65	0.63	0.07
Other Groups	-0.36	0.70	-1.43	0.64	0.06
<i>n</i>	55145		64045		

The larger the value of FMI, the greater the loss of information (hence loss of precision) that has been induced in the estimated coefficient by the missing data.

The substantive inference is altered for Pakistani and Bangladeshi pupils & those within the other ethnic group...

Analysis of Multiple Imputed Datasets

- This might be an important first step
 - The negative effects of being Pakistani or Bangladeshi are potentially being underestimated in the complete case analysis

What is the relationship between parental occupations and filial educational attainment?

- Relatively strong (and persistent) association between a pupil's GCSE attainment and the occupations of their parents (net of cohort, gender and ethnicity)
 - Similar association with any GCSE measure
 - Similar association with any of the occupational based measures (e.g. NS-SEC, RGSC, Camsis, Elias Skill)
- The level of association is stronger than gender and ethnicity
 - The parental occupational gap is more striking
 - Ongoing concern about the gender gap
 - In educational circles and in public discourse (media fuelled)
- Contribution to wider debates within stratification research

Further Comments

- Labour Government want to ‘break the link between a child’s background and their chances of success’ (DfCSF 2009)
- Evidence is presented in a confused fashion
- Concepts are mixed up
 - Family income; Free School Meals; Social Class
Working class; Disadvantaged; Deprived ??

Some Methodological Thoughts

- More sophisticated modelling of missing data
 - e.g. imputation from more comprehensive statistical models
 - Weighted survey data beyond Stata at present
- Variation within socio-econ classifications
 - this requires more exploration

The Future is Brighter

(But not necessarily for me at age 42...)

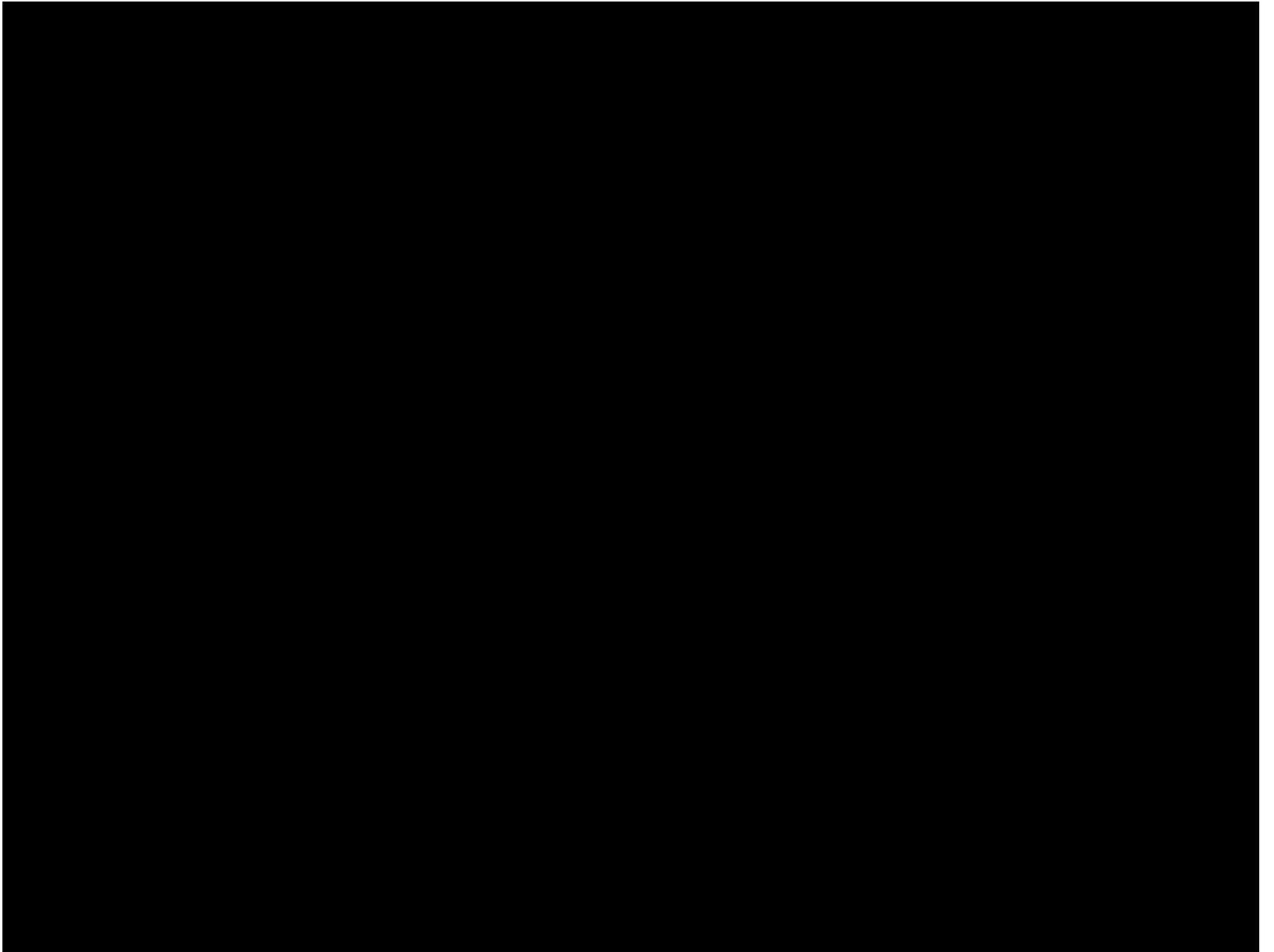
Still heavily constrained by data resources

The LSYPE is promising so are the newer birth cohorts

UKHLS – Estimated 1,000 new births per year

- Synthetic cohorts of school leavers
- Individual data (as early as rising 10 year olds)
- Linked (official) educational data
- Household data
- Income data
- Parental data (step-parents)

Consider what will be possible in 25 years time?



The Sociology of Education

Some studies that I was taught about when I was an undergraduate

Willis, P. (1977) Learning to Labour: How working class kids get working class jobs
Ethnographic study in West Midlands

Corrigan, P. (1979) Schooling the Smash Street Kids
Study of working class boys in Sunderland

Lacey, C. (1970) Hightown Grammar – The school as a social system
Study of a northern grammar school. A study of the 'disappointing performance of working class boys in grammar schools since the 1944 Education Act' (p.xi)

Rutter, M. et al. (1979) Fifteen Thousand Hours – Secondary schools and their effects on children
15,000 hours the average time spent at school, a study of 12 inner London schools
2,000 pupils tracked through secondary school

Douglas, J.W.B. (1964) The Home and the School: A study of ability and attainment in the primary school

Douglas, J.W.B. (1970) All our futures
Studies of the 1946 birth cohort

Rosenthal, R. and Jacobson, L. (1968) Pygmalion in the Classroom
Californian study – falsely informed teachers that a randomly selected 20% of pupils were gifted
They had pre and post intervention measures and noted that this group showed greater improvements in reading

A study that I have recently read

Wedge, P. and Prosser, H. (1973) Born to Fail?
Children's Bureau report on the striking differences in the lives of British children

A more recent example

Devine, F. (2004) Class Practices – How parents help their children get good jobs
A qualitative study of middle class parents in Britain and America

Models with Occupational Unit Group (OUG) Level

Linear regression with a large dummy-variable set

	areg	lsdv
n	55120	55120
R ²	0.234	0.234
Adjusted R ²	0.229	0.229
BIC	454889.53	458907.08

Identical parameter estimates and standard errors for cohorts, gender and ethnicity factors

areg is more parsimonious

The areg approach might be useful when parental occupation (at a very detailed level) is a control to be modelled rather than explicitly a variable of interest?

2000 – 5+ GCSE Passes (A*-C)

England	49%
Essex LEA	51%
Colchester High School	82%
Alderman Blaxill School	31%
My old comprehensive	25%

1997 School Performance Tables –
Examples of schools attended by notable social survey analysts

Name	Type	Year 11 n	% 5 A*-C	% 5 A*-G
England	-	-	45	84
St Bede's Redhill	VA Comp Mixed 11-18	245	69	96
St David's & St Katherine's (C of E)	VA Comp Mixed 11-18	187	14	70
Westcliffe High School for Boys	GM Selective Boys 11-18	130	96	97