Article:

McDool, E. orcid.org/0000-0002-3530-7921 (2019) Ability grouping and children's non-cognitive outcomes. Applied Economics. ISSN 0003-6846

https://doi.org/10.1080/00036846.2019.1705239

Ability Grouping and Children's Non-Cognitive Outcomes

Emily McDool[1]

December 2019

The value of ability grouping is often debated despite being adopted in primary and secondary schools across the UK for the past 80 years. Setting is one form of ability grouping which is widely adopted in English schools; it involves dividing pupils from the same cohort into classes according to ability in a specific subject. While the existing evidence identifies a negative effect on cognitive outcomes, especially for low ability pupils, little research has been undertaken to understand the impact of setting on non-cognitive outcomes. This paper provides the first evidence of the effect of setting on non-cognitive outcomes when utilising a nationally representative sample of primary-aged pupils and adopting fixed effects and instrumental variables methodologies. For boys, setting in maths negatively impacts non-cognitive outcomes, driven by a worsening of internalising behaviours. No evidence of a significant impact of lowest set placement on non-cognitive outcomes is identified.

JEL codes: I20 I21, J00

Keywords: Education, ability grouping, non-cognitive skills, behaviour

---

[1] Email: e.m.mcdool@sheffield.ac.uk
Department of Economics, 9 Mappin street, Sheffield, S1 4DT, UK
I would like to thank Sarah Brown, Geraint Johnes, Steve McIntosh, Damon Morris and Gurleen Popli for the helpful comments on earlier drafts of this paper; and to the attendees of the 2018 Leuven Economics of Education Research conference and the 2018 International Conference on Education Economics for their input.
The data that support the findings of this study are openly available from the UK Data Service at http://doi.org/10.5255/UKDA-SN-7464-4 and http://doi.org/10.5255/UKDA-SN-6411-7

1. INTRODUCTION

Childhood mental health and non-cognitive skills are key determinants of a wide range of economic and social later life outcomes (Goodman 2015; Layard et al. 2014). Even after accounting for socio-economic status and ability, childhood conduct and emotional problems have persistent effects on adulthood outcomes including educational attainment, economic activity and life satisfaction. Childhood emotional problems in particular are stronger predictors of adult life satisfaction than cognitive skills (Layard et al. 2014; Frijters et al. 2011). Schooling assists in developing both cognitive and non-cognitive skills from childhood into adolescence; during this period, non-cognitive skills are more malleable than cognitive skills (Frijters et al. 2011). Despite the fundamental role that schooling plays in non-cognitive development, research has predominantly focused on test scores as a measure of schooling outcomes and skills when evaluating the role of schooling experiences. Though expanding, the literature on the impact of schooling on non-cognitive outcomes is limited.

One feature of schooling that specifically attempts to improve the cognitive development of children but in doing so simultaneously influences non-cognitive skills is ability grouping. The value of ability grouping practises have long been debated despite being adopted within both primary and secondary schools[2] across the UK for the past 80 years (Francis et al., 2017a). With recommendations dating back to the 1960s, setting is one such practise. Setting involves dividing pupils in the same year group into classes according to measured or perceived ability for teaching in a given subject. Setting was widely encouraged by the 1997 Labour government, leading to a growth in its incidence in schools. Current policy, however, provides little guidance on the implementation of ability grouping practises which continue to be implemented in both primary and secondary schools, especially in mathematics. The evidence on the impact of setting on cognitive outcomes remains inconclusive and limited; while

[2] The UK primary education stage accommodates children between age 5 and 11 years; it is split into two stages; Key stage 1 (KS1) which caters for pupils aged 5-7 years, and Key stage 2 (KS2), which provides education to children aged 7-11 years. After completing primary education, pupils attend the lower secondary stage of education which is also divided into two stages; Key stage 3 (KS3), which caters for pupils aged 11-14, and Key stage 4 (KS4), which is the final two years of lower secondary education catering for pupils aged 14-16.

some studies find that the cognitive outcomes of higher ability children are improved by setting, whilst lower ability pupils lose out (Hallam and Parsons 2014; Ireson et al. 1999), other studies find an insignificant impact of setting on attainment (Whitburn 2001; Ireson and Hallam 2005). Very little research has been undertaken to evaluate the impact of setting, and other ability grouping practises, upon non-cognitive outcomes. The limited evidence suggests that setting influences academic self-concepts, self-confidence, motivation and self-esteem (Francis et al 2017b; Ireson and Hallam 2009). This is particularly so for low ability pupils whose motivation and self-esteem suffer as a consequence of setting[3].

This paper contributes to the limited research on ability grouping practises and schooling influences on non-cognitive outcomes by firstly exploring the extent to which setting in mathematics influences non-cognitive outcomes. Fixed Effects (FE) estimation is adopted to address the potential issue of unobserved heterogeneity, are adopted. The paper additionally explores the extent to which the level of set placement influences non-cognitive outcomes. An Instrumental Variables approach (IV) is adopted to overcome the potential endogeneity issue. The paper contributes to the existing literature in a number of ways; firstly, by examining the impact of setting in maths in primary school on non-cognitive outcomes, using a nationally representative sample of primary aged pupils in the UK. Whilst very few papers have examined the effect of setting in primary school, none have done so with a large representative panel data set which allows children to be tracked over time. The panel nature of the data allows for variation in individuals' setting experiences over time to be exploited. Secondly, the paper attempts to identify the causal effect of setting on non-cognitive outcomes by overcoming the methodological issues associated with estimating the effect of setting on outcomes, namely unobserved heterogeneity and endogeneity. Existing studies of setting observe associations only. Furthermore, the paper explores whether the responses to setting are heterogeneous by gender. Evidence suggests that a gender differential exists in the non-cognitive development and behaviour

[3] Avenues through which setting may influence non-cognitive outcomes are discussed in section 2

of primary aged children (Leadbeater al. 1999), while the peer effects literature suggests that gender differentials exist in the response to the ability composition of peers (Lavy et al. 2012)[4]. The investigation of a gender differential in response to setting in maths may be of interest to policy makers and researchers addressing the gender gaps in self-confidence and self-perceptions in maths (OECD, 2014), which ultimately lead to a decreased uptake of STEM subjects by women.

The results provide evidence that setting significantly impacts non-cognitive outcomes. The FE analysis indicates that teacher and parent reported non-cognitive problems are increased for children who are set between age 7 and 11. This effect is driven by boys whose teacher reported internalising[5] problems suffer from being set. The non-cognitive skills of girls are insignificantly influenced by setting. When investigating whether the level of set placement impacts non-cognitive problems, an insignificant influence of lowest set placement is identified.

The paper is structured as follows: a discussion of the evidence on setting and children's' non-cognitive outcomes follow in section 2; the data and methodology are discussed in section 3 with results following in section 4. Section 5 provides a discussion of the findings before concluding in section 6.

2. SETTING AND CHILDREN'S NON-COGNITIVE OUTCOMES

The impact of setting is likely to encompass a multiplicity of effects including peer effects, teaching influences and labelling effects. Peer effects play a role since students' classroom peers are limited to pupils of a similar ability; this segregation may limit the positive influence of higher attaining pupils (Kiss, 2013). Similarly, the negative behavioural and motivational influence of lower ability peers may be exacerbated in groups of low ability pupils and limited for high ability pupils (Lavy et al., 2012). Carrell and Hoekstra (2010) find that pupils exposed to disruptive peers achieve lower

[4] Lavy et al. (2012) identifies that while girls benefit in terms of their age 14 test scores from high academic ability peers, boys do not.

[5] Externalising behaviours are targeted towards others or are an outward expression or manifestation. Internalising behaviour, describes more inward behaviours and expressions.

academic outcomes and exhibit worse behaviour than their siblings who were not exposed to such peers. Setting has undesirable effects upon the behaviour of pupils in lower sets where peers are more likely to misbehave. In lower sets, ability grouping produces peer interactions characterised by hostility and anger; conversely, a supportive peer environment is produced in high ability groups (Oakes, 1985).

The process of sorting pupils may have damaging effects on the confidence, motivation, self-perceptions and ultimately, the behaviour of low ability pupils who are implicitly or explicitly informed of their relative ability among their peers. Conversely, high set placement may produce positive attitudes and expectations (Kutnick et al., 2005). Francis et al. (2017b) identifies a positive relationship between perceived set placement and both subject confidence and general confidence [6]. The authors argue that the setting process causes a self-fulfilling prophecy whereby pupils behave in accordance with their set label and level. Similarly, when exploring the relationship between setting and self-concepts in secondary school pupils in the UK, Ireson and Hallam (2009) identify that higher ability groups have greater self-concepts than students in low-ability groups in English, mathematics and science

Ability grouping practises are also likely to involve a change in teaching strategy or approach, allowing teachers to narrow their instruction according to the ability of the class (Ofsted, 1998). This focussed learning environment may reduce the likelihood of disinterest from pupils at both ends of the ability distribution, thereby reducing misbehaviour (House of Commons, 2011).

Closely related to ability grouping is tracking which involves separating pupils by ability into different types of schools, thus influencing the peers and schooling experiences of pupils. Tracking is implemented in the compulsory-schooling systems of multiple countries including Germany and Austria. The tracking literature predominantly focuses on the effect on student performance in terms

[6] The paper examines the impact of set level placement in English and maths on subject and general self-confidence among 11-12 year olds in secondary schools.

of cognitive skills. Borghans et al. (2015), however, also considers the effect on non-cognitive skills such as extraversion, conscientiousness and school motivation, when estimating the effect of being placed in a high ability track for pupils at the margin. A positive influence of tracking on cognitive outcomes but an insignificant impact of track placement on non-cognitive skills is identified.

3. DATA AND METHODOLOGY

This paper utilises data from the Millennium Cohort Study (MCS) which is a national longitudinal birth cohort study of children born in the UK between September 2000 and January 2002[7]. The MCS provides a wealth of information on social, economic and health aspects of children's lives. This paper utilises data from waves 4 (age 7) and 5 (age 11) which achieved samples of 14,043 and 13,469 children respectively. Responses to the main parent or carer[8] and child's school teacher questionnaires are utilised[9]; this provides key information on the child's abilities and schooling experiences alongside individual and family characteristics.

The Strengths and Difficulties Questionnaire (SDQ) provide the outcomes of interest. SDQ is a behavioural screening questionnaire suitable for children aged between 3 and 16 years, used broadly by psychologists, clinicians, educationalists and researchers (Gupta and Simonsen, 2010; Goodman, 1997). Both the teacher and parent SDQ responses are observed in this study for comparative and robustness purposes. The SDQ comprises of 25 statements[10] regarding the child's attributes or behaviour from five categories: emotional, conduct, hyperactivity/ inattention and peer relationship problems and prosocial behaviour, with five questions for each category (see Table A1). The SDQ is recoded so that a higher overall score indicates greater behavioural problems. The total difficulties score is the main outcome of interest; this is a sum of the problems within the first four categories

[7] Since observing children who recently attended primary school, the MCS provides a current reflection of the policies adopted within schools

[8] Responses to the parent interview are provided by the main parent or carer; in most cases this is the natural mother: 97.4% in 2008 & 97.5% in 2012

[9] Teacher responses are available in England and Wales only

[10] The extent of the behaviours from each question must be rated on a 3-point scale from 'not true', 'somewhat true' or 'certainly true'.

(emotional, conduct, hyperactivity and peer problems)[11] and provides an overall indication of behavioural problems. In order to identify the types of behaviour that children exhibit, the internalising (emotional and peer problems) and externalising behaviour (conduct and hyperactivity) scores are also observed.

The main independent variables of interest are associated with class setting in mathematics; this information is provided by the teacher who is asked whether the child is set for maths and the level of the set: highest, middle or lowest set. Within the questionnaires, a definition of class setting[12] and streaming[13] are provided thus reducing the potential problem in varying definitions of class setting across schools and teachers. The MCS data indicates that children are taught maths for an average of 5 hours per week when aged 7 and 5.3 hours per week when aged 11, accounting for approximately one day of teaching per week which is a considerable proportion of teaching time[14]. The impact of set level placement is observed only at age 11 (wave 5) since the data suggests that setting is more prevalent at age 11, with 63% of pupils being set, as opposed to 37% at age 7[15]; setting at age 11 is therefore more likely to be due to widely adopted school polices and less determined by other school characteristics. The teacher questionnaire achieves a higher response rate in wave 5 than in wave 4, leading to a larger useable sample of children in wave 5 even after attrition.

A sample of 7,913 observations (6,010 individuals) is achieved for the analysis of setting, which uses observed children at age 7 and 11, and is limited to only individuals who are not set at age 7 and are subsequently set for maths at age 11; the estimated effect therefore represents the effect of setting rather than 'de-setting'[16]. For the analysis of set level placement, the useable sample includes 4,779

[11] The total difficulties score may equal a maximum of 40; this and all outcome measures are treated as continuous.
[12] The questionnaire states: "Some schools group children from different classes by ability for certain subjects only and they may be taught in different ability groups for different subjects. We refer to this as setting."
[13] The questionnaire states: "Some schools group children in the same year by general ability and they are taught in these groups for most or all lessons. We refer to this as streaming. … Other schools do not group children by ability between classes. Sometimes this may be because there are not multiple classes in the year"
[14] The impact of setting in English was also considered and investigated but the results provided few interesting results, possibly due to the lesser time spent in English sets, relative to maths, and the smaller sample of pupils set for English.
[15] These Figures are in line with Hallam et al. (2003)
[16] 294 observations are dropped from the sample due to being 'de-set'

individual observations which includes all individuals whose set level is reported at age 11, regardless of their previous set experience[17].

The paper is concerned with answering the following questions: 1. How are the non-cognitive skills of primary-aged pupils influenced by being set for maths, as opposed to mixed ability teaching? And 2. To what extent does lowest set placement impact upon non-cognitive skills? To firstly identify whether being set for maths influences non-cognitive outcomes, a fixed effects (FE) methodology is employed. OLS results are provided, however, the FE model is preferred due to the potential unobserved heterogeneity. There are likely to be unobserved characteristics, which are related to or determine an individual's non-cognitive skills which may be intrinsic characteristics or personality traits that vary between individuals. For instance, adaptability, impulsivity, tendencies to 'self-minimize[18]' and the big five personality traits have been found to be determined by early childhood and remain relatively stable across the lifespan (Nave et al. 2010; Hampson and Goldberg, 2006). Unobserved individual effects may therefore be correlated with the regressors. By adopting a FE approach the effect of a change in setting upon non-cognitive skills is estimated when controlling for school-level, individual and family characteristics.

The FE model to be estimated:

$$SDQ_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Set_{it} + a_i + \lambda t + \varepsilon_{it} \qquad (1)$$

SDQ_{it} denotes individual i's (=1,…,6010) SDQ score at time t (=age 7, 11). β_0 denotes the intercept. Set_i equals one when the individual is set and zero otherwise. X_{it} denotes a vector of school, teacher, individual and family characteristics of individual i at time t. a_i is a set of fixed parameters

[17] Since data is required on the employed instruments which do not feature in the previous analysis, this sample is smaller than the 4,561 observations at age 7 plus the 294 individual observations that are dropped in the main analysis.
[18] Self minimizing is defined as "tends to minimize own importance; humble; never brags or shows off; seeks out or is content with less important tasks or positions."

representing all stable individual characteristics of individuals. t denotes a time trend. Finally, ε_{it} indicates an individual and time specific error term.

FE estimation relies upon variation in the outcome variable and the control variable of interest. As Table 1 indicates, within-individual variation in the SDQ measures are apparent; this variation is greatest for the teacher reported total difficulties score. Teacher reported scores are more likely to change over the observed period than parent scores possibly since the child's teacher may change over time. The descriptive statistics also suggest that parents are likely to report higher difficulties scores[19], possibly since children may amend their behaviour at home and at school (Mitchell and Shepard, 2011) which substantiates the utilisation of both teacher and parent reported behaviours.

Table 1 Descriptive statistics of dependent variable and main independent variable of interest

Variable	Mean	Std. dev	Within Std. dev	Within min	Within Max
Set maths	0.520	0.500	0.219	0.020	1.020
Teacher total	5.750	5.530	1.846	-7.750	19.250
Parent total	7.505	5.611	1.496	-6.495	21.505
Teacher internalising	2.275	3.016	1.192	-4.925	10.075
Parent internalising	2.985	3.018	0.966	-5.105	10.985
Teacher externalising	3.175	3.627	1.129	-4.825	11.175
Parent externalising	4.520	3.558	0.950	-2.480	11.520
Number of children (n):			6,010		
Number of observations (N):			7,913		
Average waves observed:			1.32		

An IV approach is employed to estimate the extent to which lowest set placement impacts non-cognitive skills, due to the issue of endogeneity which is likely to arise due to reverse causality; while behaviour[20] may be influenced by set placement, set placement may be influenced by behaviour. Children with worse behaviour or behavioural issues are more likely to be placed in lower ability sets

[19] This is also apparent in **Error! Reference source not found.**

[20] The non-cognitive outcomes of interest are taken from the SDQ which is a behavioural screening questionnaire and is therefore highly likely to reflect child behaviour which is observable by class teachers.

(Dunne et al. 2007). To estimate the effect of placement in the lowest maths set upon behaviour, the following models are estimated.

First-stage regression:

$$Lowset_i = \beta_0 + \beta_1 SET + \beta_2 Z_i + \beta_i X_i + \varepsilon_i \quad (2)$$

Reduced form:

$$SDQ_i = \beta_0 + \beta_1 Set_i + \beta_2 Lowset_i + \beta_i X_i + \varepsilon_i \quad (3)$$

SDQ_i denotes individual i's (=1,…,4,779) SDQ score. SET is a binary variable indicating whether the individual is set or not; $Lowset$ is also a binary variable which indicates whether the individual is in the lowest maths set. X_i indicates a vector of characteristics discussed in detail below and in Table A2. The coefficient on SET indicates the impact of being set for maths; this will specifically relate to being set in the mid or high level set, relative to not being set.

The addition of the SET and $Lowset$ coefficients is referred to as 'lowest set':

$$Lowest\ set_i = \beta_1 Set_i + \beta_2 Lowset_i \quad (4)$$

Lowest set indicates the impact of being set and being placed in the lowest maths set, relative to not being set. The combination of the two coefficients allows for the 'pure' effect of being placed in the lowest maths set to be observed, over and above the effect of being set. The standard errors for the addition of the two coefficients are correspondingly estimated.

The two major assumptions of the IV approach are that the instrument (Z) should be correlated with the endogenous variable $Lowset$ (age 11 maths set placement) but should have no direct effect on the outcome variable (non-cognitive skills indicated by the SDQ score) once controlling for all X_i. The first instrument adopted is the proportion of children who are from English as an additional language

(EAL) homes in the child's normal class[21]; this is likely to influence set placement since EAL are more likely to be overrepresented within lower sets. Primary schools often place EAL children in lower sets (Dunne et al., 2007). The non-cognitive skills and behaviour of pupils are unlikely to be affected by the proportion of EAL children in the class since, by age 10, EAL status is an insignificant predictor of a child's own social and behavioural outcomes (Sylva et al. 2007). The behaviour of EAL peers is therefore insignificantly different to non-EAL peers. Furthermore, a comprehensive study by Geay et al. (2012) identifies an insignificant effect of non-native English speakers in the classroom on the performance of their classroom peers. In a similar vein, Ohinata and Ours (2016) find no evidence for negative peer effects of immigrant children in primary schools. The proportion of EAL children within a class of a given school may also be considered fairly random since primary classes are often constructed by date of birth.

The second instrument is the number of maths sets in the respondent's school year; this is also correlated with the likelihood of the pupil being placed within the lowest set since a higher number of alternative sets in the school year makes placement in the lowest set less likely. The number of sets that the school allocates to each school year for maths is unlikely to directly influence the behaviour of the child. It seems reasonable to assume that the size of the school outside of the classroom has little influence on the SDQ scores of children, especially when considering the lack of evidence of school size effects. Factors such as class size, which may be correlated and could influence outcomes are controlled for within the model. The mean values of the instrumental variables are provided in Table 2 for the pooled sample, the lowest set and, for comparison, the highest set sample. The mean SDQ scores are also given by set level and show that non-cognitive problems are highest within the low set children and lowest amongst the high set. Whilst this could be a result of setting, SDQ could also be a determinant of set level, thus encouraging the adoption of the IV methodology. Both the proportion of EAL students and the number of maths sets have a negative and

[21] The proportion of EAL children is calculated using the number of children with EAL in the class that the child is normally taught in (i.e. not their maths set) and the class size.

significant relationship with the endogenous variable, lowest set placement, and an insignificant relationship with non-cognitive skills, measured by the total difficulties score[22] [23].

Table 2 Mean values and standard deviation of instrument and outcomes by set level

	All sets	Lowest set	Highest set
Instruments:			
Prop. class EAL	10.738	10.939	12.507
	(21.003)	(20.240)	(22.745)
No. maths sets	2.216	2.816	2.929
	(1.269)	(0.955)	(0.965)
Outcomes:			
Teach. total SDQ	5.504	8.469	3.762
	(5.465)	(5.966)	(4.173)
Parent total SDQ	7.656	10.307	6.069
	(5.816)	(6.520)	(4.875)
Observations	4,779	603	1,385

Standard deviation in parentheses

Control variables (X_{it}) include a range of individual, family and teacher characteristics[24]. Individual controls include Special Educational Needs (SEN) since SEN children characteristically perform worse than non-SEN children (Crawford and Vignoles, 2010). Ethnicity is also controlled for since evidence suggests that the teacher reported incidence of abnormal and borderline behavioural problems amongst children varies by ethnicity (Popli & Tsuchiya, 2014). Similarly, the season of birth is a determinant of non-cognitive development since younger children exhibit behavioural immaturity, resulting in perceived lower ability by teachers (Campbell, 2013). Birth order and birth weight are additionally controlled for; whilst evidence suggests that birth weight is correlated with the susceptibility to issues such as anxiety, depression and aggressive outbursts amongst school children (Bohnert & Breslau, 2008), birth order significantly influences child behaviour and the behavioural roles adopted (Sulloway, 1996). Having a regular bed time has also been linked to

[22] The reduced form and first stage regression results are presented in the lower panel of Table 7.

[23] Further instrument tests are provided in the IV results Table; in the main sample and girls sample, the instruments perform well in the rigorous testing procedure. For boys, the proportion of EAL has a negative, zero effect on the endogenous variable i.e. in the first stage regression.

[24] A brief description of these controls is provided within Table A2.

behavioural difficulties during childhood (Kelly et al. 2013). Involvement in a religion is controlled for since religion may influence the child's values, outlooks and attitudes, in turn potentially influencing behaviour (Petts, 2009). Gender is controlled for due to differential non-cognitive development; for example, girls are more likely to exhibit internalising behaviours whilst boys are more likely to display externalising behaviours (Leadbeater al. 1999). Finally, Chanfreau (2015) finds that participation in out-of-school activities is similarly related to SDQ.

Family characteristic controls include household income and parental education which are proxies of parental skills which have been found to determine their children's cognitive and non-cognitive skills (Coulon et al. 2011). Additionally, attendance at parents evening alongside parent interest, as measured by the teacher, attempt to control for factors associated with parental investments and style which are also associated with non-cognitive development and behaviour (Hernandez-Alava and Popli, 2017)

Teacher years and tenure control for experience and skills which may influence both the teacher's capability to assess pupil ability and set accordingly, and their abilities in assisting cognitive and non-cognitive development. Evidence suggests that a negative relationship exists between teacher experience and student behaviour control (Ritter and Hancock, 2007). Class size and mixed year group variables control for the number of peers alongside the presence of older, more mature peers[25]. Whilst in smaller classes children are more engaged in learning and exhibit less disruptive behaviour (Finn et al., 2003), mixed year group classes significantly influence prosocial and aggressive forms of behaviour (McClellan and Kinsey, 2002). The presence of disruptive peers and the proportion of excluded peers within the child's usual class controls hold constant peer behaviour which may influence the child's own behaviour and conduct (Carrell and Hoekstra, 2010). Whether the child is set in any other subject is also controlled for; data on setting practises is available for English at age 7 and 11 and Science at age 11. Finally, neighbourhood deprivation, measured by the Index of

[25] These controls are also associated with other school characteristics related to school size.

Multiple deprivation (IMD) quartile[26] is controlled for; while neighbourhood deprivation and characteristics influence children's outcomes (Goux and Maurin, 2007), peer characteristics, both in and out of school, are likely to be correlated with neighbourhood characteristics.

The mean and standard deviation of control and outcome variables are provided for the pooled and gender samples in Table 3 for individuals at age 7 and 11. Teacher and parent reported non-cognitive problems are higher for boys than for girls on average. Parents report greater non-cognitive problems than teachers for all measures.

Table 3 Descriptive statistics of independent and dependent variables by sample

Explanatory variables in X	Main model	Females	Males
Set maths	0.520	0.517	0.522
	(0.500)	(0.500)	(0.500)
Birth Order	0.861	0.852	0.871
	(0.990)	(0.982)	(0.998)
Birth Weight	6.958	6.813	7.107
	(1.330)	(1.322)	(1.322)
Regular Bedtime	0.906	0.909	0.903
	(0.292)	(0.288)	(0.296)
Born A/W	0.506	0.492	0.520
	(0.500)	(0.500)	(0.500)
Male	0.495	0.000	1.000
	(0.500)	(0.000)	(0.000)
White	0.854	0.855	0.852
	(0.354)	(0.352)	(0.355)
SEN	0.211	0.152	0.271
	(0.408)	(0.359)	(0.444)
Religion	0.502	0.466	0.538
	(0.500)	(0.499)	(0.499)
Siblings HH	1.475	1.453	1.498
	(1.037)	(1.034)	(1.040)
Parent degree	0.123	0.121	0.125
	(0.328)	(0.327)	(0.330)
Married	0.582	0.582	0.581
	(0.493)	(0.493)	(0.493)
Working HH	0.871	0.872	0.869
	(0.336)	(0.334)	(0.337)
Log income	8.346	8.371	8.320
	(2.187)	(2.186)	(2.189)
Parents Eve.	0.955	0.958	0.952
	(0.207)	(0.201)	(0.214)

[26] The IMD is a measure of deprivation within each Lower Super Output Area (LSOA). The IMD considers seven aspects of deprivation including: Income, employment, health deprivation and disability, education skills and training, barriers to housing and services, crime and the living environment. Each LSOA contains on average 1,200 people.

School club	0.361	0.374	0.348
	(0.480)	(0.484)	(0.476)
Mixed year grp.	0.247	0.241	0.253
	(0.431)	(0.428)	(0.435)
Class size	26.347	26.490	26.200
	(5.242)	(5.093)	(5.386)
Teacher tenure	8.350	8.417	8.282
	(7.062)	(7.246)	(6.870)
Teach. years	13.449	13.573	13.323
	(9.595)	(9.657)	(9.532)
Prop peers excluded	0.806	0.754	0.859
	(3.217)	(3.331)	(3.096)
Disruptive peers	0.349	0.334	0.365
	(0.477)	(0.472)	(0.482)
Set for Eng. or Sci.	0.522	0.523	0.521
	(0.500)	(0.500)	(0.500)
IMD quartile 1 (most deprived)	0.191	0.190	0.192
	(0.393)	(0.392)	(0.394)
IMD quartile 2	0.257	0.264	0.249
	(0.437)	(0.441)	(0.432)
IMD quartile 3	0.269	0.278	0.259
	(0.443)	(0.448)	(0.438)
IMD quartile 4 (least deprived)	0.199	0.190	0.207
	(5.530)	(4.765)	(5.986)
Outcome variables Y			
Teach total SDQ	5.750	4.584	6.939
	(5.530)	(4.765)	(5.986)
Par total SDQ	7.505	6.793	8.231
	(5.611)	(5.221)	(5.896)
Teach internal. SDQ	2.575	2.485	2.668
	(3.016)	(2.885)	(3.142)
Par internal. SDQ	2.985	2.922	3.050
	(3.018)	(2.910)	(3.123)
Teach external. SDQ	3.175	2.100	4.271
	(3.627)	(2.835)	(4.000)
Par external. SDQ	4.520	3.871	5.181
	(3.558)	(3.243)	(3.739)
Observations	7,913	3,996	3,917

Standard deviation in parentheses

4. RESULTS

4.1 Analysis of setting

Table 4 provides the main results[27] from OLS and FE analysis[28]. The results from the preferred FE specification suggest that for boys, setting increases non-cognitive problems; the teacher reported total difficulties score increases as a result of being set for maths by 0.74, ceteris paribus. For girls, the OLS results also signal a positive influence of being set on non-cognitive problems, though the FE results fail to provide evidence of a significant effect.

Table 4 Effect of setting on total difficulties scores

	Teacher total difficulties		Parent total difficulties	
	OLS	FE	OLS	FE
All	0.081	0.394	0.054	0.223
	(0.114)	(0.246)	(0.118)	(0.201)
Male	0.060	0.741*	-0.181	0.298
	(0.176)	(0.384)	(0.174)	(0.302)
Female	0.112	0.021	0.271*	0.122
	(0.145)	(0.314)	(0.161)	(0.272)
Observations NT (Number of children N)	All N=7,913 (N=6,010) Male N= 3,917 (N=2,991) Female N=3,996 (N=3,019)			

Notes: Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The type of behaviour exhibited as a result of setting is also examined to identify whether changes in internalising or externalising behavioural problems result from setting children by ability. Table 5 indicates that setting increases both parent reported externalising problems and teacher reported internalising problems; being set, as opposed to mixed ability teaching, increases teacher internalising behaviours SDQ score by 0.28, ceteris paribus. This negative influence on behaviour is driven by boys whose teacher reported internalising problems increase by 0.47 due to being set for maths. The FE results continue to suggest that girls are insignificantly impacted by setting.

[27] Full results in appendix

[28] Time invariant controls include: regular bedtime, logged equalized weekly family income, parents evening attendance, school club attendance, special educational needs (SEN), married parents, working household, household siblings, class size, mixed year group class, teacher years taught, disruptive class peers, proportion of excluded classroom peers, teacher tenure at child's school, set for another subject, IMD quartile, wave.

Table 5 Effect of setting on internalising and externalising scores

	Teacher Internalising		Parent Internalising		Teacher Externalising		Parent Externalising	
	OLS	FE	OLS	FE	OLS	FE	OLS	FE
All	0.008	0.283*	0.015	-0.023	0.073	0.111	0.039	0.246*
	(0.067)	(0.161)	(0.067)	(0.128)	(0.075)	(0.150)	(0.076)	(0.129)
Males	-0.014	0.471**	-0.073	0.069	0.074	0.270	-0.108	0.229
	(0.098)	(0.237)	(0.097)	(0.185)	(0.121)	(0.246)	(0.113)	(0.198)
Females	0.034	0.089	0.093	-0.122	0.077	-0.068	0.178*	0.244
	(0.092)	(0.220)	(0.093)	(0.178)	(0.088)	(0.180)	(0.101)	(0.168)
Observations NT (Number of children N)	All N=7,913 (N=6,010) Male N= 3,917 (N=2,991) Female N=3,996 (N=3,019)							

Notes: Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

This analysis may include individuals who are set for maths alongside other subjects, for example English or science, though the data indicates that setting in such subjects is less likely. Though the main model controls for whether an individual is set in other subjects, one concern is that there may be a multiplicative effect of being set in each individual subject upon non-cognitive outcomes. As a robustness check, the analysis is restricted to individuals who are not set for other subjects to explore whether the results still hold when setting solely in maths is experienced. The results in Table 6 continue to signal a positive and significant effect of being set on the teacher reported non-cognitive problems for boys; this effect is of much greater magnitude than the main analysis. For girls, an insignificant effect of setting on non-cognitive problems continues to be identified in the FE models. Based on conjecture, the size of the estimated effects may be unequal due to differences in confidence and self-perceptions across different subjects. It is well-documented that boys have greater confidence, more positive attitudes and attach greater importance to maths than girls (Samuelson and Samuelson, 2015; Hargreaves et al. 2008). Males are also likely to overestimate their performance in maths (Bench et al. 2015). The impact of being set may therefore be large for boys since self-concepts, confidence and self-perceptions are impacted by knowledge of relative ability and rank for low ability pupils. For higher set boys, the increased difficulty of teaching and materials may challenge prior perceptions of ability and maths skill and ability to a greater extent than mixed ability classes.

Table 6 Effect of setting excluding alternative subject set pupils

	Teacher total difficulties		Parent total difficulties	
	OLS	FE	OLS	FE
All	-0.265	2.568	-0.166	0.282
	(0.182)	(1.639)	(0.192)	(1.208)
Males	-0.230	4.480*	-0.273	-0.026
	(0.289)	(2.590)	(0.286)	(1.776)
Females	-0.274	0.766	-0.087	0.010
	(0.226)	(2.301)	(0.258)	(1.678)
Observations NT (Number of children N)	All N=3,785 (N=3,614) Male N= 1,878 (N=1,790) Female N=1,907 (N=1,824)			

Notes: (i) Time invariant controls from previous FE analyses also included (ii) Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$

4.2 Analysis of set level

The results from the IV analysis of the impact of lowest set placement are presented in Table 7[29]. The results indicate an overall insignificant effect of both setting and placement in the lowest maths set in all analyses; this result holds when the base group is both children who are not set for maths and children who are set in other sets. The results therefore provide little evidence to support the theoretical arguments against setting due to the negative consequences on the non-cognitive outcomes of children who placed in the lowest sets.

[29] All results from IV models additionally control for: birth order, birth weight, regular bedtime, logged equalized weekly family income, parents evening attendance, school club attendance, special educational needs (SEN), married parents, working household, household siblings, religious participation, parental education, gender, autumn/winter date of birth, white, class size, mixed year group class, teacher years taught, disruptive class peers, proportion of excluded classroom peers, teacher tenure at child's school, whether set for English or Science, IMD quartile.

Table 7 IV estimation of effect of lowest set placement

	Teacher total difficulties			Parent total difficulties		
	All	Males	Females	All	Males	Females
Low set	0.652	-0.411	0.987	1.021	3.990	-0.520
	(1.698)	(4.082)	(1.578)	(1.855)	(4.268)	(1.892)
Set maths	-0.405	-0.104	-0.460	-0.327	-1.098	0.145
	(0.423)	(0.967)	(0.415)	(0.462)	(1.011)	(0.498)
Lowest set	0.247	-0.515	0.527	0.694	2.891	-0.375
	(1.325)	(3.169)	(1.240)	(1.448)	(3.313)	(1.488)
Sargan statistic [1]	0.074	0.009	0.392	0.435	0.015	1.064
	(p=0.786)	(p=0.923)	(p=0.531)	(p=0.509)	(p=0.903)	(0.302)
Cragg-Donald Wald F-Statistic [2]	43.850	9.576	38.573	43.850	9.576	38.573
Anderson Canon LM, χ^2-statistic [3]	86.581	19.199	75.518	86.581	19.199	75.518
	(p=0.000)	(p=0.000)	(p=0.000)	(p=0.000)	(p=0.000)	(p=0.000)
IV reduced form: No. maths sets [4]	-0.086	-0.042	-0.107	-0.061	-0.122	-0.006
	(0.064)	(0.103)	(0.079)	(0.070)	(0.106)	(0.093)
IV reduced form: Prop. EAL [5]	0.001	0.000	0.003	0.003	0.001	0.007
	(0.004)	(0.006)	(0.005)	(0.004)	(0.006)	(0.006)
IV first stage: No. maths sets [6]	-0.049***	-0.034***	-0.063***	-0.049***	-0.034***	-0.063***
	(0.006)	(0.008)	(0.008)	(0.006)	(0.008)	(0.008)
IV first stage: Prop. EAL [7]	-0.001***	-0.000	-0.001***	-0.001***	-0.000	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations NT:	All N=4,779 Male N= 2,361 Female N=2,418					

Notes: (i) Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ (ii) [1] Sargan statistic for the overidentification test of all instruments with p-value given in parentheses (iii) [2] Weak identification test of Cragg-Donald Wald (iv) [3] χ^2-statistic for the Anderson Canon under-identification test with p-values given in parentheses (v) [4] coefficient of the number of maths sets instrument in the reduced form model, with standard errors given in parenthesis (vi) [5] coefficient of the proportion of EAL instrument in the reduced form model, with standard errors given in parenthesis (vii) [6] coefficient of the number of maths sets instrument in the first stage, with standard errors given in parenthesis (viii) [7] coefficient of the proportion of EAL instrument in the first stage, with standard errors given in parentheses. More formally, the reduced form and first stage regressions attempt to test that the instruments satisfy the following conditions: Instrument should be correlated with the endogenous variable (first stage) but is uncorrelated with the residual i.e. uncorrelated with unobservable factors influencing the dependent variable (reduced form). The first stage residual is insignificant in the reduced form model.

One potential concern of the main IV analysis is that 20% of the sample are classified as having Special Educational Needs (SEN). Whilst SEN children characteristically perform worse than non-SEN children (Crawford and Vignoles, 2010), SEN may impact upon non-cognitive outcomes[30] thus influencing the response to setting. The results from the IV analysis indicate that SEN is a strong significant predictor of non-cognitive problems. Although SEN is controlled for in the main IV model, there may be unobservable characteristics associated with SEN and non-cognitive behaviours[31]. Table 8 presents the IV results when SEN children are dropped from the sample. Results continue to indicate an insignificant effect of lowest set placement and setting in maths upon the non-cognitive outcomes of children, specifically those without SEN.

[30] Fauth et al. (2014) find that children with SEN encounter a greater increase in peer, hyperactivity and emotional problems over time, than non-SEN children

[31] These characteristics are likely to be controlled for in FE analysis as time invariant unobserved individual characteristics.

Table 8 IV estimation of effect of lowest set placement excluding SEN pupils

	Teacher total difficulties			Parent total difficulties		
	All	Males	Females	All	Males	Females
Low set	-0.501	0.505	-1.310	-0.932	1.986	-2.174
	(1.632)	(3.908)	(1.630)	(1.807)	(3.961)	(1.972)
Set maths	-0.094	0.009	-0.100	0.109	0.042	0.101
	(0.295)	(0.554)	(0.348)	(0.326)	(0.561)	(0.422)
Lowest set	-0.595	0.515	-1.410	-0.823	2.027	-2.072
	(1.405)	(3.443)	(1.376)	(1.556)	(3.489)	(1.666)
Sargan statistic [1]	0.217	0.337	1.729	0.731	0.236	0.777
	(p=0.641)	(p=0.561)	(p=0.189)	(p=0.392)	(p=0.627)	(p=0.378)
Cragg-Donald Wald F-Statistic [2]	50.266	12.439	38.182	50.266	12.439	38.182
Anderson Canon LM, χ^2statistic [3]	98.586	24.865	74.495	98.586	24.865	74.495
	(p=0.000)	(p=0.000)	(p=0.000)	(p=0.000)	(p=0.000)	(p=0.000)
IV reduced form: No. maths sets [4]	-0.038	0.022	-0.067	0.001	0.028	-0.018
	(0.064)	(0.106)	(0.078)	(0.071)	(0.107)	(0.094)
IV reduced form: Prop. EAL [5]	0.002	-0.004	0.008	0.004	0.002	0.007
	(0.004)	(0.006)	(0.005)	(0.004)	(0.006)	(0.006)
IV first stage: No. maths sets [6]	-0.051***	-0.037***	-0.062***	-0.051***	-0.037***	-0.062***
	(0.005)	(0.008)	(0.008)	(0.005)	(0.008)	(0.008)
IV first stage: Prop. EAL [7]	-0.001***	-0.000	-0.001**	-0.001***	-0.000	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations NT:	All N=43,826 Male N= 1,739 Female N=2,087					

Notes: (i) Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ (ii) [1] Sargan statistic for the overidentification test of all instruments with p-value given in parentheses (iii) [2] Weak identification test of Cragg-Donald Wald (iv) [3] χ^2-statistic for the Anderson Canon under-identification test with p-values given in parentheses (v) [4] coefficient of the number of maths sets instrument in the reduced form model, with standard errors given in parenthesis (vi) [5] coefficient of the proportion of EAL instrument in the reduced form model, with standard errors given in parenthesis (vii) [6] coefficient of the number of maths sets instrument in the first stage, with standard errors given in parenthesis (viii) [7] coefficient of the proportion of EAL instrument in the first stage, with standard errors given in parentheses. More formally, the reduced form and first stage regressions attempt to test that the instruments satisfy the following conditions: Instrument should be correlated with the endogenous variable (first stage) but is uncorrelated with the residual i.e. uncorrelated with unobservable factors influencing the dependent variable (reduced form). The first stage residual is insignificant in the reduced form model.

One further concern is the omission of ability from the main model since ability may influence the non-cognitive responses to set placement. Ability is initially omitted as it is highly correlated with set placement, thus by controlling for ability, differences in behaviour between set and not set pupils cannot be due to ability; this would mean separating out the setting effect from the ability effect so that the impact of set is not reflecting ability level. This may reduce the potential for a confounding variable bias to arise. However, the ability of children in the lowest set is likely to vary between schools since setting is dependent upon the relative ability of a pupil in the ability distribution of their cohort within their school. The effect of setting is re-estimated when holding ability, which determines set level differently in different schools, constant. The MCS progress in maths score [32] controls for and provides an exogenous measurement of maths ability which does not involve or inform the child's school teacher, who determines the set of the child and reports their non-cognitive behaviours. Ability is measured at age 7 before the child's set is observed at age 11. The results presented in Table 9 generally provide evidence of a negative relationship between maths ability and the total SDQ score; children with higher ability have significantly fewer non-cognitive problems. Even after controlling for ability, setting and lowest set placement continue to have an insignificant effect on the non-cognitive outcomes of children as in the main results.

[32] The MCS progress in maths score provides a measurement of mathematical ability which indicates progress in relation to the National Curriculum in the UK. The test undertaken by the MCS respondents is a reduced version of the National Foundation for Educational Research standard Progress in Maths (PiM) test. The test is undertaken at age 7 in the fourth wave of the MCS and involves a series of 'paper and pencil' calculation exercises covering a number of mathematical topics (Connelly, 2013). The total raw score is used within this study as a control for ability, this simply represents the number of correct answers given on the test.

Table 9 IV estimation of effect of lowest set placement with maths ability controls

	Teacher total difficulties			Parent total difficulties		
	All	Males	Females	All	Males	Females
Low set	1.129	0.079	1.676	2.407	4.225	1.351
	(1.656)	(3.482)	(1.684)	(1.797)	(3.623)	(1.974)
Set maths	-0.514	-0.205	-0.648	-0.428	-0.828	-0.209
	(0.396)	(0.809)	(0.416)	(0.430)	(0.842)	(0.487)
Lowest set	0.615	-0.126	1.029	1.980	3.397	1.142
	(1.316)	(2.743)	(1.350)	(1.429)	(2.854)	(1.583)
Maths ability	-0.188***	-0.225***	-0.155**	-0.189***	-0.115	-0.259***
	(0.049)	(0.079)	(0.061)	(0.053)	(0.083)	(0.072)
Sargan statistic [1]	0.289	0.047	0.307	0.817	0.380	0.661
	(p=0.591)	(p=0.829)	(p=0.580)	(p=0.366)	(p=0.537)	(p=0.416)
Cragg-Donald Wald F-Statistic [2]	49.134	13.678	36.557	49.134	13.678	36.557
Andersson Canon LM, χ^2-statistic [3]	96.775	27.390	71.733	96.775	27.390	71.733
	(p=0.000)	(p=0.007)	(p=0.021)	(p=0.000)	(p=0.000)	(p=0.002)
IV reduced form: No. maths sets [4]	-0.107	-0.054	-0.138*	-0.058	-0.078	-0.041
	(0.067)	(0.108)	(0.081)	(0.072)	(0.110)	(0.095)
IV reduced form: Prop. EAL [5]	0.002	0.001	0.002	0.003	0.003	0.004
	(0.004)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)
IV first stage: No. maths sets [6]	-0.054***	-0.043***	-0.064***	-0.054***	-0.043***	-0.064***
	(0.006)	(0.008)	(0.008)	(0.006)	(0.008)	(0.008)
IV first stage: Prop. EAL [7]	-0.000*	-0.000	-0.001	-0.000*	-0.000	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations NT:	All N=4,359 Male N= 2,119 Female N=2,240					

Notes: (i) Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ (ii) [1] Sargan statistic for the overidentification test of all instruments with p-value given in parentheses (iii) [2] Weak identification test of Cragg-Donald Wald (iv) [3] χ^2-statistic for the Anderson Canon under-identification test with p-values given in parentheses (v) [4] coefficient of the number of maths sets instrument in the reduced form model, with standard errors given in parenthesis (vi) [5] coefficient of the proportion of EAL instrument in the reduced form model, with standard errors given in parenthesis (vii) [6] coefficient of the number of maths sets instrument in the first stage, with standard errors given in parenthesis (viii) [7] coefficient of the proportion of EAL instrument in the first stage, with standard errors given in parentheses. More formally, the reduced form and first stage regressions attempt to test that the instruments satisfy the following conditions: Instrument should be correlated with the endogenous variable (first stage) but is uncorrelated with the residual i.e. uncorrelated with unobservable factors influencing the dependent variable (reduced form). The first stage residual is insignificant in the reduced form model.

5. DISCUSSION

The results provide concern for policy makers and stakeholders, such as parents and teachers, interested in the effects of ability grouping. For boys, experiencing setting, as opposed to mixed ability teaching, for cognitive development purposes, produces harmful effects on non-cognitive outcomes including emotional and peer skills. Based on the previous literature, a combination of multiplicative harmful effects are likely to be at play for the lowest set. By being informed of their relative ability, general and subject confidence may be negatively affected, especially if maths ability and set placement are overestimated, as is common amongst males (Bench 2015; Francis 2017b). Set placement also highlights a child's ability to peers, potentially further influencing self-esteem and self-concepts. Additionally, since motivation and behaviour are correlated with set level and pupils behave in accordance with their set label (Francis et al. 2017b), lowest set pupils' behaviour may deteriorate once set. For higher set boys, the increased difficulty of classes may impact upon confidence, which may be especially harmful when ability prior to setting is overestimated.

Since the analysis initially evaluates the impact of being set for children from all sets, the effects by ability are difficult to disentangle thus providing rationale for analysing the impact of set level. An insignificant effect of lowest set placement, relative to mixed ability maths teaching is identified[33]. The results are reassuring for policy makers concerned purely with the influence of setting on the non-cognitive development of low set children. The findings concur with Borghans et al. (2015) who identify an insignificant impact of track placement level on non-cognitive skills. This study does however differentiate itself from existing studies of ability grouping and non-cognitive outcomes since it attempts to identify a causal effect of set placement rather than observing associations only. It is possible that studies which identify a significant association between set level and behaviour do so due to a confounding variable related to both set level and non-cognitive skills; classroom

[33] These findings are in line with the existing evidence which finds an insignificant effect of ability grouping on a range of non-cognitive outcomes including self-esteem (Abadzi, 1984) and self-concepts (Ireson and Hallam, 2005)

disruptive behaviour for example influences set level decisions while simultaneously determining and being determined by non-cognitive development. Furthermore, the study utilizes a large, representative longitudinal data set of children from a range of schools across England rather than a single school or region as they experience changes in teaching approach over time.

The findings highlight the importance of analysing the effect of setting by gender, which is an infrequent practise in existing studies of ability grouping. The response of girls and boys to setting is heterogeneous, possibly due to gender differences and the malleability of non-cognitive skills and behaviour during childhood, alongside gender disparities in subject confidence and attitudes. Unlike girls, boys are less likely to struggle with 'maths anxiety' and may be more influenced by the actual placement. Furthermore, the effect of peer quality, which is likely to change as a result of setting, is heterogeneous by gender; Lavy et al. (2012) identifies that while girls gain from academically bright peers boys fail to benefit.

The results also interestingly highlight that the negative effect of setting for boys is apparent in the school environment, as reported by the teacher but are not projected at home since parental reported behaviour is insignificantly influenced. While parents may not recognise behavioural changes, it is possible for children to alternatively exhibit behaviour disorders either only at school or at home (Mitchell and Shepard, 2011). It is therefore extremely valuable to analyse the effect of schooling policies on non-cognitive development in both environments since problems may manifest in alternate settings. Very few studies in this literature have done so to date.

Overall, the findings suggest that additional measures should be taken to limit the harmful effects of changing teaching strategies from mixed ability teaching to ability sets. The findings suggest that lowest set placement insignificantly influences non-cognitive skills in the final year of primary school. It is possible that once set, the benefits of setting for the lowest set pupils, such as tailored teaching, which improves esteem and confidence through comprehensible classes, balance the negative features of setting, including the segregation of pupils, thus leading to an overall insignificant effect of lowest set placement on non-cognitive skills.

6. CONCLUSION

Setting is an ability grouping practise which continues to be adopted in primary schools across England. The school-level policy aims to improve pupil outcomes by narrowing the ability distribution of pupils in the class for a specific subject, allowing teaching to be tailored to pupils' needs. This policy has not gone without opposition, with peer effects, segregation and pupil self-confidence being at the forefront of the arguments against the policy. A major concern of setting relates to the non-cognitive outcomes of pupils, such as their emotional, conduct and behavioural development, which may suffer as a consequence of pursuing cognitive improvements. This paper examines the effect of setting in primary school on the non-cognitive outcomes of pupils, measured by the Strengths and Difficulties Questionnaire (SDQ). The effect of being set for maths between age 7 and 11 is firstly examined by adopting OLS and fixed effect methodologies. Secondly, the effect of lowest set placement upon non-cognitive outcomes is estimated when addressing the issue of endogeneity by employing an IV approach. Whilst this research contributes to the ongoing debate surrounding ability grouping and setting more specifically, it also contributes to the economics literature on the determinants of non-cognitive skills and development in primary school.

Overall, the results suggest that non-cognitive skills are negatively impacted by being set by ability after experiencing mixed ability teaching for maths. This effect is driven by boys whose teacher reported non-cognitive problems are increased by setting, specifically internalising problems.

The analysis of the effect of being placed in the lowest maths set identifies an insignificant effect upon non-cognitive outcomes. This result is identified in the pooled sample alongside the sub-sample analysis of boys and girls. The results therefore provide evidence that the non-cognitive development of lowest set children is not significantly harmed by set placement. The concerns behind setting largely relate to the non-cognitive consequences for lower ability children, in terms of their emotional and conduct difficulties and their behaviour and self-esteem. The insignificant effect of lowest set placement is therefore interesting, especially since the evidence on setting and non-cognitive outcomes is currently limited. While the results present little concern for the impact of setting on the lowest set pupils, the results do promote interventions within primary schools to reduce the impact of changing children from mixed ability teaching to set ability groups in the later primary years. Research on setting should continue to consider these potential effects on non-cognitive outcomes.

REFERENCES

Abadzi, H. (1984) Ability Grouping Effects on Academic Achievement and Self-Esteem in a Southwestern School District. *The Journal of Educational Research* 77(5): 287-292

Bench, A., Lench, H. Liew, J., Miner, K. & Flores, S. (2015) Gender Gaps in Overestimation of Math Performance. *Sex roles* 72(11-12):536-546

Bohnert, K. & Breslau, N. (2008) Stability of psychiatric outcomes of low birth weight: A longitudinal investigation. *Archives of General Psychiatry* 65(9):1080-1086.

Borghans, L., Korthals, R., & Schils, T. (2015). *The effect of track placement on cognitive and non-cognitive skills.* Unpublished Manuscript. Maastricht University

Campbell, T. (2013) *In-school ability grouping and the month of birth effect: preliminary evidence from the Millennium Cohort Study*. Centre for Longitudinal Studies (CLS) working paper 2013/1

Carrell, S. & Hoekstra (2010) Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids. *American Economic Journal: Applied Economics* 2(1): 211–228

Chanfreau, J. (2015) Out of school activities in the primary years: analysis of Millennium Cohort Study (Powerpoint presentation) Presented at Centre for Longitudinal Studies Cohort Studies Research Conference 2015 17th March 2015.

Coulon, A. Meschi, E. & Vignoles, A. (2011) Parents' skills and children's cognitive and non-cognitive outcomes. *Education Economics*, 1-24

Crawford, C. & Vignoles, A. (2010) *An analysis of the educational progress of children with special educational needs.* DoQSS Working Paper No. 10-19, London: Institute of Education.

Dunne, M. Humphreys, S. & Sebba, J. (2007) *Effective Teaching and Learning for Pupils in Low Attaining Groups.* Department for Children, Schools and Families report DCSF_RR011

Fauth, R. Parsons, S.& Platt, L. (2014) *Convergence or divergence? A longitudinal analysis of behaviour problems among disabled and non-disabled children aged 3 to 7 in England*. Institute of Education working paper no. 14-13

Finn, J., Pannozzo, G., & Achiles, M. (2003) The 'why's' of class size: student behaviour in small classes. *Review of Educational Research*, 73(3):321-368

Francis, B., Archer, L., Hodgen, J., Pepper, D. Taylor, B. & Travery, M. (2017a) Exploring the relative lack of impact of research on 'ability grouping' in England: a discourse analytic account. *Cambridge Journal of Education,* 47(1):1-17

Francis, B., Connolly, P., Archer, L., Hodgen, J., Mazenod, A., Pepper,D., Sloan, S., Taylor, B., Tereshchenko, A. & Travery, M. (2017b) Attainment Grouping as self-fulfilling prophesy? A mixed methods exploration of self-confidence and set level among Year 7 students. *International Journal of Educational Research,* 86: 96-108

Frijters, P., Johnston, D. & Shields, M. (2013) Does childhood predict adult life satisfaction? Evidence from British Cohort Survey. *The Economic Journal*. 124: F688-F719

Geay, C., McNally, S. Telhaj, S. (2013) Non-native speakers of English in the Classroom: What Are the Effects on Pupil Performance? *The Economic Journal* 123(570): F281–F307

Goodman, A., Joshi, H. Nasim, B. & Tyler, C. (2015) *Social and emotional skills on childhood and their long-terms effects on adult life*. London: Early Intervention Foundation.

Goodman R (1997) The Strengths and Difficulties Questionnaire: A Research Note. J*ournal of Child Psychology and Psychiatr,* 38, 581-586.

Gupta, N. & Simonsen, M. (2010) Non-cognitive child outcomes and universal high quality child care. *Journal of Public Economics* 94(1-2):30-43

Hallam, S. & Parsons. (2014) The impact of streaming on attainment at age seven: evidence from the Millennium Cohort Study. *Oxford Review of Education,* 40(5): 567-589H

Hallam, S.Ireson, J. Lister, V. & Chaudhury (2003) *Ability grouping practices in the primary school: A survey*. Institute of Education

Hampson, S., Goldberg, L. (2006) A first large cohort study of personality trait stability over the 40 years between elementary school and midlife. *Journal of Personality and Social Psychology*. 1:763–779.

Hargreaves, M. Homes, M. & Swinnerton, B. (2008) A comparison of performance and attitudes in mathematics amongst the 'gifted'. Are boys better at mathematics or do they just think they are? Assessment in Education: Principles, *Policy & Practise* 15(1):19-38

Hernández-Alava, M., & Popli, G. (2017). Children's Development and Parental Input: Evidence from the UK Millennium Cohort Study. *Demography*, 54(2), 485–511

House of Commons (2011) Behaviour and Discipline in Schools. HC 516-II

Ireson, J. & Hallam, S. (2005) Pupils' liking for school: Ability grouping, self-concept and perceptions of teaching. *British Journal of Educational Psychology* 75: 297–311

Ireson, J. & Hallam, S. (2009) Academic self-concepts in adolescence: relations with achievement and ability grouping in schools. *Learning and instruction* 19:201-213

Ireson, J. Mortimore, P., & Hallam, S. (1999). The common strands of pedagogy and their implications in Ireson, J. & Hallam, S. Ability grouping in Education (pp 30-39, pp46-51).

Kelly, Y., Kelly, A., & Sacker, A. (2013). Changes in bedtime schedules and behavioral difficulties in 7 year old children. *Paediatrics*. 132(5): e1184-e1193

Kiss, D. (2013) Impact of peer achievement and peer heterogeneity on own achievement growth. *Economics of Education Review* (37) 58-65

Kutnick, P., Sebba, J., Blatchford, P., Galton, M., and Thorp, J. (2005) *The effects of pupil grouping: literature review*. Department for Education and Skills.

Lavy, V. Silva, O. and Weinhardt, F. (2012) The Good, the Bad and the Average: Evidence of Ability Peer Effects in Schools. *Journal of Labor Economics* 30(2):367-414

Layard, R., Clark, R. Cornaglia, F., Powdthavee, N. & Vernoit, J. (2014) What predicts a successful life? A life-course model of Well-being. *The Economic Journal* 124: F720-F738

Leadbeater, B. Kuperminc, G. & Hertzog, C. (1999) A Multivariate Model of Gender Differences in Adolescents' Internalizing and Externalizing Problems. *Developmental Psychology* 35(5):1268-1282

McClellan, E., & Kinsey, J. (2002). Children's social behavior in relation to participation in mixed-age or same-age classrooms. *Early Childhood Research and Practice,* 1(1)

Mitchell, S. & Shepard, M. (2011) A comparative study of children's behaviour at home and at school. *British Journal of Educational Psychology* 36(3), 248-254

Nave, C., Sherman, R., Funder, D., Hampson, S., & Goldberg, L. (2010). On the Contextual Independence of Personality: Teachers' Assessments Predict Directly Observed Behavior after Four Decades. *Social psychological and personality science*, 3(1), 1-9.

Oakes, J. (1985) *Keeping track: How schools structure inequality*. New Haven: Yale University Press.

OECD (2014) *Are boys and girls equally prepared for life*? OECD publishing.

Office for Standards In Education (1998) *Setting in primary schools: A report from the Office of Her Majesty's Chief Inspector of Schools*. London, HMSO

Ohinata, A. & Ours, J. (2016) Quantile Peer Effects of Immigrant Children at Primary Schools. *Labour* 30(2):135-157

Popli, G. & Tsuchiya, A. (2014) *Sons and Daughters: Parental Beliefs and Child Behaviour (Evidence from the UK Millennium Cohort Study*). Working Papers 2014013. The University of Sheffield, Department of Economics.

Ritter, J. & Hancock, D. (2007) Exploring the relationship between certification sources, experience levels, and classroom management orientations of classroom teachers. *Teaching and Teacher education* 23(7): 1206-1216

Samuelson, M. & Samuelson, J. (2015) Gender differences in boys' and girls' perception of teaching and learning mathematics. *Open review of Educational Research* 3(1):18-34.

Sulloway, F. (1996) *Born to rebel: Birth order, family dynamics, and creative lives*. New York, NY, US: Pantheon Books

Sylva, K., Melhuish, E., Siraj-Blatchford, I., & Taggart, B. (2007) *Promoting equality in the early years, Report to the Equalities Review*. London: Cabinet Office.

Whitburn, J. (2001) Effective classroom organisation in primary schools: mathematics. *Oxford Review of Education*, 27 (3), 411-428.