The University Of Sheffield.

Department Of Economics.

Sheffield Economic Research Paper Series.

Class Setting and Children's Non-Cognitive Outcomes

Emily McDool

ISSN 1749-8368

SERPS no. 2018008

May 2018

www.sheffield.ac.uk/economics

Class Setting and Children's Non-Cognitive Outcomes

Emily McDool [1]

Department of Economics, University of Sheffield, UK

May 2018

Abstract

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. The effect of setting has long been debated; while the existing evidence identifies a negative effect on cognitive outcomes, especially for the low ability, little research has been undertaken to understand the impact of setting on non-cognitive outcomes. This paper provides the first evidence of setting on non-cognitive outcomes when adopting a nationally representative sample of primary aged pupils. Using Millennium Cohort Study data, Fixed Effects (FE) and Instrumental Variables (IV) methodologies are adopted to overcome potential unobserved heterogeneity and endogeneity. For boys, setting in maths negatively impacts non-cognitive outcomes, as measured by the Strengths and Difficulties Questionnaire. This effect is driven by a worsening of internalising behaviours. Little evidence is found for a significant impact of lowest set placement on non-cognitive outcomes.

JEL codes: J13, I21, J00

Key words: Educational economics, ability grouping, non-cognitive skills, children, primary schooling, teaching practices

Acknowledgements: This research was funded by the ESRC. I would like to thank Sarah Brown, Geraint Johnes, Steve McIntosh, Gurleen Popli and for the helpful comments on earlier drafts of this paper; and to the attendees of the 2018 Leuven Economics of Education Research conference for their input.

[1] Contact details of corresponding author: Department of Economics, 9 Mappin street., Sheffield , S1 4DT e.m.mcdool@sheffield.ac.uk +4401142223310

1. INTRODUCTION

Previous work has identified that childhood mental health and social and emotional wellbeing are key determinants of a wide range of economic and social later life outcomes (Goodman 2015; Layard et al. 2014; Richards and Abbott 2009). 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 (Clark et al. 2017; Layard et al. 2014; Frijters et al. 2011). Childhood emotional problems in particular are found to be stronger predictors of adult life satisfaction than cognitive skills.

Schooling assists in developing both cognitive and non-cognitive skills from childhood into adolescence; during this period, non-cognitive skills are likely to be more malleable than cognitive skills (Frijters et al. 2011). In primary school, children begin to develop and enhance their social and emotional skills and build the foundations for later life development. Despite the fundamental role that non-cognitive development plays in child progress and life outcomes (see for example Gutman & Schoon, 2013), research has predominantly focused on test scores as a measure of schooling outcomes and skills when evaluating the role of schooling experiences. Though growing, the literature on the impact of education or, more specifically, schooling on non-cognitive outcomes is somewhat limited.

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

[2] The primary education stage accommodates for children between the ages of 5 and 11. Primary education may be split into two stages; infant, known as Key stage 1 (KS1), which caters for children between foundation year and year 2 when pupils are aged between 5 and 7, and junior, referred to as Key stage 2 (KS2), which provides education to children in

cognitive outcomes remains inconclusive and rather limited; while some studies suggest that the cognitive outcomes of higher ability children are improved by setting, whilst lower ability pupils lose out (Hallam and Parsons 2014; Ireson 1999a; Suknandan and Lee 1998; Slavin 1988), other studies find an insignificant impact of setting on attainment (Whitburn 2001; Barker-Lunn 1970; Kulik and Kulik 1982, Ireson and Hallam 2005). Very little research has been undertaken to evaluate the impact of setting, and other ability grouping practises more generally, 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; Gamoran & Berends 1987). This is particularly so for low ability pupils whose motivation and self-esteem especially, suffer as a consequence of setting. There are a number of avenues through which setting may influence non-cognitive outcomes, including the adaptation of teaching which may be tailored to a narrower ability range; this may reduce the disengagement of pupils (Gamoran 2002; Ofsted 1998; House of Commons 2011). There are also positive effects associated with high ability peers (Kiss, 2013; Robertson and Symons, 2003; Bradley and Taylor, 2008) which are reduced or removed for lower ability children when segregating from higher ability children. With low ability sets being more likely to contain pupils from ethnic backgrounds and low socio-economic backgrounds, setting may reinforce the existing social gap in performance (Gamoran 2002). Additionally, by informing children of their relative ability by setting, may have damaging effects on low ability pupils' confidence and motivation (Kutnick et al., 2005).

This paper contributes to the limited research on ability grouping practises and schooling influences on non-cognitive outcomes by firstly exploring whether setting in mathematics influences non-cognitive outcomes. Ordinary Least Squares (OLS) and Fixed effects (FE) estimation approaches are adopted. Due to the potential issue of unobserved heterogeneity the FE model is favoured. In line with the existing literature, the paper additionally explores whether placement in the lowest set in particular, influences non-cognitive outcomes. An Instrumental Variables approach (IV) is adopted to overcome the potential endogeneity issue of set level by exploiting variation in school peer characteristics and the number of sets within schools. Data from the

year 3 to year 6, up to the age of 11. Upon completion of primary education at age 11, pupils begin the lower secondary stage of education which provides education to pupils aged between 11 and 16. In a similar manner to the primary stage, secondary education may be divided into two phases; Key stage 3 (KS3), which caters for pupils in year 7 to year 9, when aged between 11 and 14, and Key stage 4 (KS4), which relates to the final two years of lower secondary education when pupils are in year 10 and year 11, and are aged between 15 and 16.

Millennium Cohort Study (MCS) is utilised. This provides the non-cognitive outcomes of children at age 7 and 11, measured by the Strength and Difficulties Questionnaire (SDQ)[3].

The paper contributes to the existing literature in a number of ways; firstly, the paper examines 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 element 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. These issues are addressed by adopting FE and IV methodologies. Existing studies of setting largely estimate associations only. Additionally, the paper identifies whether a gender differential exists in both the impact of setting and set level placement since evidence suggests that the non-cognitive development and behaviour of girls and boys may differ (McNeish and Scott, 2014; Leadbeateret al. 1999). Whilst within the setting literature, gender differentials are infrequently considered, the peer effects literature provides some evidence that gender differentials exist in the response to the ability composition of peers (Lavy et al. 2012)[4]. The responses of girls and boys to setting may therefore be heterogeneous. The investigation of a gender differential in response to setting in maths may also be of interest to policy makers and researchers addressing the gender gaps in self-confidence and self-perceptions in maths (OECD 2014); these gaps ultimately lead to a decreased uptake of STEM subjects by women. Setting or de-setting in such subjects may provide a path to improve the non-cognitive skills of children which drive the gender gap.

The results provide some evidence for the impact of setting on non-cognitive outcomes. The FE analysis indicates that teacher reported non-cognitive problems are increased for children who are set between age 7 and age 11. This effect is identified for boys specifically who suffer from being set in terms of teacher reported internalising problems but also their parent reported externalising problems. The non-cognitive skills of girls, on the other hand, are insignificantly influenced by setting. When investigating whether the level of set

[3] The SDQ is a behavioural screening questionnaire, frequently used in the child development literature (See for example Gupta and Simonsen, 2010; Goodman, 1997).
[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.

placement impacts non-cognitive problems, the paper finds an insignificant influence of lowest set placement on outcomes; this result is robust to the inclusion of a control for ability and to the exclusion of Special Educational Needs (SEN) children from the sample, whose non-cognitive skills may differ from non-SEN children.

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 is discussed in section 3 with results following in section 4. The paper concludes in section 5.

2. SETTING AND CHILDREN'S NON-COGNITIVE OUTCOMES

The impact of setting is likely to encompass a multiplicity of complementary effects including peer effects, teaching influences and labelling effects. Peer effects are likely to play a role since when grouped by ability, students' classroom peers are limited to peers of a similar ability; low ability students are therefore segregated from higher attaining pupils thus removing the positive influence of the high ability on low ability children's progress (Kiss, 2013; Robertson and Symons, 2003; Bradley and Taylor, 2008). Since lower ability pupils are found to cause negative peer effects within the classroom (Lavy et al, 2011, Lavy et al., 2012), grouping by ability may benefit the high ability pupils and have negative consequences for the low ability in terms of attainment, behaviour and motivation. Carrell and Hoekstra (2010) estimate the impact of class peers from troubled families, who exhibit more disruptive behaviour, upon behaviour and reading and maths scores. The study finds that pupils exposed to these peers achieve lower academic outcomes and exhibit worse behaviour, relative to their siblings who were not exposed to peers experiencing domestic violence. The results may imply that setting could have undesirable effects upon the behaviour of pupils in lower sets where peers are more likely to misbehave. Oakes (1985) identifies that ability grouping produces a supportive peer environment in high ability groups but hostility and anger characterised peer interactions in low ability classes.

Ability grouping practises including setting 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 detraction from the class of pupils on both ends of the ability distribution, thereby reducing misbehaviour (House of Commons, 2011).

Alongside peer effects and teaching influences which may occur after setting pupils, the process of sorting pupils by ability may have damaging effects on the confidence, motivation and self-perceptions of low ability pupils who are implicitly or explicitly informed of their relative ability within their class. Conversely, high set placement may produce positive attitudes and expectations (Kutnick et al., 2005). Francis et al. (2017b) examines the impact of set level placement in English and maths on subject and general self- confidence. The study uses data from 139 secondary schools in England, providing data on over 11,500 students who are tracked between year 7 (aged 11/12) to year 8 (12/13). By comparing set pupils with those who experience mixed ability teaching for English and maths, the study identifies a significant positive relationship between perceived set placement and both subject confidence and general confidence; those in lower sets report lower levels of both subject and general confidence. The authors argue that the setting process, which labels pupils according to ability, causes a self-fulfilling prophecy whereby pupils behaves in accordance with their set label which reflects their set level.

In addition to Francis et al. (2017b), a limited literature examines the impact of class setting, and other forms of ability grouping upon non-cognitive outcomes. for example, self-concepts and self-esteem (Abadzi, 1985; Ireson and Hallam, 2009), pupil attitudes (Boaler, 1997; Ireson and hallam, 2001; Suknandan and Lee, 1998) and grade anxiety (Wang, 2014). Notably, Ireson (1999) examines the impact of setting in secondary schools within the UK upon non-cognitive outcomes. Whilst setting in maths and science is found to insignificantly affect self-esteem and self-concepts, setting in English improves the self-concepts of low attaining pupils, but lowers the self-concepts of higher attaining pupils. These results mirror the findings of a meta-analysis of 13 studies of setting and ability grouping by Kulik and Kulik (1992). More recently, Ireson and Hallam (2009) explore the relationship between setting and self-concepts, using data on 14-15 year olds from a stratified sample of 23 secondary schools in the UK. The study identifies that academic self-concepts, though not general self-concepts, are influenced by the extent of ability grouping within the school. Higher ability groups are found to have greater self-concepts than students in low-ability groups; this is true for English, mathematics and science.

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 a practise which 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 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. The study finds a positive influence of tracking on cognitive outcomes but an insignificant impact of track placement on non-cognitive skills.

Very few papers in the ability grouping literature have attempted to identify the causal effect of setting upon non-cognitive outcomes. This paper contributes to the existing literature by addressing the methodological issues of unobserved heterogeneity and endogeneity surrounding the measurement of the effect of setting and set placement on child outcomes to estimate a causal effect. In addition, panel data on a nationally representative sample of children in the UK is utilised which allows for individuals to be tracked overtime. Furthermore, the study focuses on setting amongst primary school children; despite the common implementation of setting in primary schools, especially for maths, few setting studies have focused on children of this age. The outcome measure of interest is also novel to this area of literature which, to date, has focused on a narrow range on non-cognitive skills.

3. DATA AND METHODOLOGY

The paper utilises data from the Millennium Cohort Study (MCS) which is a national longitudinal birth cohort study which initially followed 19,000 children born in the UK between September 2000 and January 2002. Six waves of the MCS are currently available, with data collected when respondents were aged 9 months, 3 years, 5 years, 7 years, 11 years and most recently at age 14 years. This paper utilises data from waves 4 (age 7) and 5 (age 11) which achieved samples of 14,043 and 13,469 children respectively. The MCS provides a wealth of information on social, economic and health aspects of the children's lives. Since observing children who recently attended primary school, the MCS providing a current reflection of the policies adopted within schools.

Alongside the self-completion questionnaires for the child respondent, the MCS collected data from respondents' main parent or carer[5] which provide data on the family context, the child's health, education and

[5] 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

income, employment and parenting activities. In addition, the fourth and fifth waves of the MCS provide the responses to postal self-completion questionnaires completed by the child's teacher to gain information on the child's abilities, behaviour, profile, parents, groupings alongside the characteristics of teacher and class and the move to secondary school.

The outcomes of interest are taken from the Strengths and Difficulties Questionnaire (SDQ) which is a behavioural screening questionnaire used broadly by psychologists, clinicians, educationalists and researchers[6] (see for example Gupta and Simonsen, 2010; Goodman, 1997). The questionnaire is suitable for children aged between 3 and 16 years and may be completed by parents or teachers (SDQ info, 2014a). Both the teacher and parent SDQ responses are observed in this study for comparative and robustness purposes. The SDQ comprises of 25 statements regarding the child's attributes or behaviour from five categories: emotional problems, conduct problems, hyperactivity/ inattention problems, peer relationship problems and prosocial behaviour (Gallop et al., 2013), with five questions for each category (see Table A1). The extent of the behaviours from each question must be rated on a 3-point scale from 'not true', 'somewhat true' or 'certainly true'. The SDQ is recoded so that a negative attribute that the child certainly exhibits is coded a higher value (equal to two) and coded zero when not exhibited. Reversely, a 'certainly true' response to positive behaviour questions is coded zero with vales of 1 or 2 given to somewhat true and not true responses respectively. A higher overall score therefore 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 (emotional, conduct, hyperactivity and peer problems); it therefore may equal a maximum of 40 and provides an overall indication of the behavioural problems. The internalising and externalising behaviour scores are also observed within the paper; the internalising behaviour score is the sum of emotional and peer problems and the externalising behaviour score is the sum of conduct and hyperactivity scores. All outcome measures are treated as continuous.

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,

[6] SDQ info (2016) reports that over 3,900 SDQ publications exist from across developmental, genetic, social, clinical and educational studies.

middle or lowest set. Within the questionnaires, a definition of class setting[7] and streaming[8] are provided thus reducing the potential problems 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 a week when aged 7 and 5.3 hours per week when aged 11, accounting for approximately one day of teaching time per week. Children are therefore taught within maths sets for a considerable proportion of their school week[9]. The impact of set level placement is observed only at age 11 (wave 5) since the evidence and data suggests that setting is more prevalent at age 11, with 65% of pupils being set, as shown in the raw data table 1[10]; 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 more than in wave 4, even after attrition.

Table 1: Setting at age 7 and age 11

	Age 7	**Age 11**
Set maths	1,218 (36.5%)	2,893 (63.4%)
Not set maths	2,134 (63.45%)	1,668 (36.6%)
Number of observations (N)	3,352	4,561

The analysis is based upon England and Wales since teacher responses are provided in these countries only for both waves. A sample of 7,913 observations (6,010 individuals) is achieved for the analysis of setting, which uses wave 4 and wave 5 (i.e. age 7 and 11 respectively). This analysis 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'. 294 observations are dropped from the sample due to being 'de-set' between age 7 and age 11. The descriptive statistics reflect the sample excluding these individuals. For the analysis of

[7] "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." NatCen (2008)
[8] "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" NatCen (2008)
[9] 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.
[10] These figures are in line with Hallam et al. (2003)

set level placement, the useable sample includes 4,779 individual observations (age 11). This sample includes all individuals whose set level is reported at age 11, regardless of their previous set experience [11].

Sampling weights are applied to take account of the sampling design of the MCS which involves oversampling within smaller countries, areas of high child poverty, and within England, areas of high proportions on ethnic minorities. Applying the weights provided by the MCS allows the panel to be restored, providing representative proportions on individuals from all areas.

The paper is concerned with answering two main questions: Are non-cognitive skills influenced by setting in maths? And does lowest set placement impact 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 fixed effects model is preferred due to the potential for 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 are likely to vary between individuals. 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 fixed effects model to be estimated:

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

SDQ_{it} denotes individual i's (=1,…,6010) SDQ score at time t (=age 7, age 11). β_0 denotes the intercept. Set_i is a dummy equalling 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 representing all stable individual characteristics of individuals. t denotes a time trend. Finally, ε_{it} indicates an individual and time specific error term. Control variables in the vector X and are described in detail below and in table A2. In addition to observing the total difficulties score, this analysis is also undertaken when internalising and

[11] This sample is smaller than the 4,561 observed at age 7 plus the 294 individual observations that are dropped in the previous analysis since, alongside set level, data is required on the employed instruments which do not feature in the previous analysis.

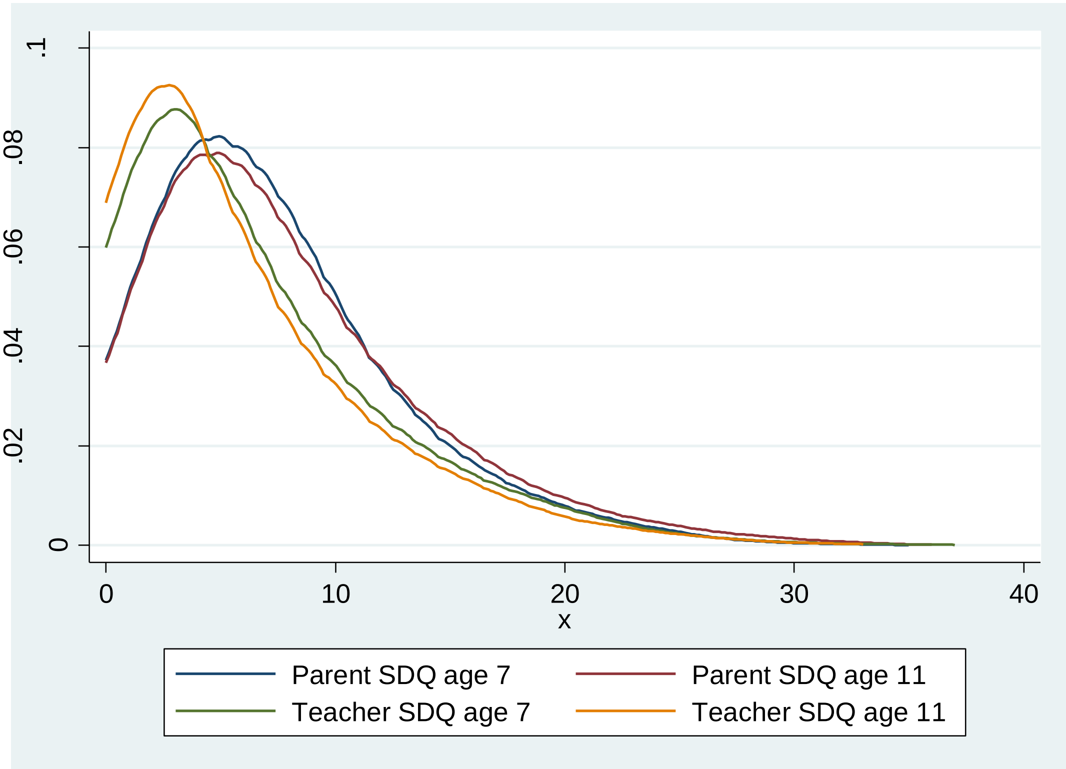
externalising problems are observed as the outcome variables of interest in order to identify the types of behaviour that children exhibit in response to being set in maths.

FE estimation relies upon variation in the outcome variable and the control variable of interest. Table 2 provides the descriptive statistics of setting and the outcome variables. Within individual variation in the SDQ measures are apparent over time; this variation is greatest for the teacher reported total difficulties score. The descriptive statistics also suggest that parents are likely to report higher difficulties scores; this is also apparent in figure 1 which plots the total difficulties score by wave and respondent.

Table 2: 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		

Figure 1: Density plot of SDQ outcome measures by age



In order to address the second question of this paper: Does placement in the lowest set influence non-cognitive outcomes, age 11 setting placement is observed and an instrumental variables (IV) approach is employed. An IV approach is taken due to the issue of endogeneity which is likely to arise due to reverse causality; while behaviour[12] may be influenced by set placement, the child's behaviour is likely to influence the level of the set in which they are placed. Children with worse behaviour or more behavioural issues are more likely to be placed in lower ability sets (Dunne et al. 2007; Boaler, 1997). OLS will therefore produce biased and inconsistent estimates of the impact of setting due to the violation of the OLS assumptions. To estimate the effect of placement in the lowest maths set upon behaviour, the following models are estimated.

First-stage regression:

$$Lowset_i = \beta_0 + \beta_1 SET + \beta_2 Z_i + \beta_i X_i + \varepsilon_i \tag{2}$$

Reduced form:

$$SDQ_i = \beta_0 + \beta_1 Set_i + \beta_2 Lowset_i + \beta_i X_i + \varepsilon_i \tag{3}$$

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 though 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':

$$Lowest\ set_i = \beta_1 Set_i + \beta_2 Lowset_i \tag{4}$$

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.

[12] 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.

The two major assumptions of the IV approach are that an instrument (Z) should be correlated with the endogenous variable $Lowset$ but should have no direct effect on the outcome variable once controlling for all X_i. Therefore, an instrument must correlate with age 11 maths set placement but must not directly influence non-cognitive skills, more specifically, the SDQ score which is likely to reflect child behaviour. 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[13]; this is likely to influence set placement since EAL are more likely to be overrepresented within lower sets and primary schools are often found to place EAL children in lower sets before language is improved (Dunne et al., 2007). The non-cognitive skills and behaviour of pupils are unlikely to be affected by the proportion of children in the class whose native language is not English since by age 10, EAL status is an insignificant predictor of a child's own social and behavioural outcomes (Sylva et al. 2007). Therefore, if an individual's own behaviour is unaffected by EAL status, peers' own behaviour is unlikely to be influenced by the status. 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 within the respondent's school year; this is also likely to be correlated with the likelihood of the pupil being placed within the lowest set since a higher number of sets within the school year makes placement within the lowest set more unlikely since there are more alternative sets. Additionally, 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 3 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

[13] 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.

and lowest amongst the high set. Whilst this could be a result of setting, this 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 significant relationship with the endogenous variable, low set placement, and an insignificant relationship with the non-cognitive skills, measured by the total difficulties score[14]. Further instrument tests are provided within the IV results table; it should be noted that 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.

Table 3: 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)
Par total SDQ	7.656	10.307	6.069
	(5.816)	(6.520)	(4.875)
Observations	4779	603	1385

Standard deviation in parentheses

Control variables (X_{it}) are entered into the models which analyse setting and set level to control for individual, family and teacher characteristics. A brief description of these controls is provided within table A2. Individual controls include Special educational needs (SEN) since SEN children characteristically perform worse than non-SEN children (Crawford and Vignoles, 2010; Kramarz et al. 2008). 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 identified as being a determinant of non-cognitive development since younger children may exhibit behavioural immaturity which may result in lower set placement due to 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

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

(Bohnert & Breslau, 2008), the existing literature also identifies a significant influence of birth order upon child behaviour and the behavioural roles adopted (Sulloway, 1996). Having a regular bed time has also been linked to behavioural difficulties during childhood (Kelly et al. 2013). Involvement in a religion is also added as a control since religion may influence the child's values, outlooks and attitudes, in turn potentially influencing non-cognitive skills and behaviour, as identified by Petts (2009). Participation in out-of-school activities has also been found to be related to SDQ (Chanfreau, 2015) and is therefore controlled for. Finally, gender is controlled for since girls due to differential non-cognitive development between girls and boys; for example, girls are more likely to exhibit internalising behaviours whilst boys are more likely to display externalising behaviours (McNeish and Scott, 2014; Leadbeateret al. 1999).

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 (Feinstein and Symons, 1999; Coulon et al. 2011). Additionally, attendance at parents evening alongside parent interest, as measured by the teacher, attempt to control for factors associated with parenting investments and style which are also associated with non-cognitive development and behaviour (Hernandez-Alava and Popli, 2017; Scott et al. 2010)

Teacher years and teacher tenure control for teacher experience and skills which may influence both the teacher's capability to assess pupil ability and set children accordingly, and their abilities in assisting learning and both cognitive and non-cognitive development. Relatedly, evidence suggests that teachers with many years of experience have significantly less control over student behaviour (Ritter and Hancock, 2007). Class size and mixed year group control for the number of peers alongside the presence of older, more mature peers. Whilst evidence suggests that in smaller classes children are more engaged in learning and exhibit less disruptive behaviour (Finn et al., 2003), a significant influence of mixed year group classes has been identified upon prosocial and aggressive forms of behaviour (McClellan and Kinsey, 2002). These variables may also pick up the effect of other school characteristics related to school size. The presence of disruptive peers and the proportion of excluded peers within the child's usual class controls take account of peer behaviour which may influence the child's own behaviour and conduct (Carrell and Hoekstra, 2010). The model also includes a control for whether the child is set in any other subject; data on setting practises is available for English in wave 4 (age 7) and for English and Science in wave 5 (age 11). Finally, neighbourhood deprivation, measured by the Index of Multiple deprivation (IMD) quartile[15] is controlled for; while evidence suggests that 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 variables are provided in table 4 for the full sample of individuals in wave 4 and 5, thereby reflecting characteristics at age 7 and age 11. In accordance with the methodology of the paper, the statistics are presented for the pooled sample as well as by gender to identify differentials in characteristics. Both the teacher and parents reported total difficulties scores, which are the main outcomes of interest, are higher for boys than for girls on average. Parents report greater non-cognitive problems than teachers for all measures of non-cognitive problems.

[15] 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.

Table 4: Descriptive statistics of all variables of interest by sample

	Main model	Females	Males
Explanatory variables in X			
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)
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			
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)
Prop. class EAL	10.782	10.617	10.950
	(21.358)	(21.120)	(21.600)
Observations	7913	3996	3917

Standard deviation in parentheses

4. RESULTS

4.1 Analysis of setting

The main results from the OLS and FE analysis are provided in Table 5; full results are available in Table A3. Whilst the OLS results indicate an insignificant influence of setting on outcomes in the pooled and boys sample, a negative impact of setting is identified for girls whose parent reported total difficulties increase as a result of being set, relative to girls who are not set for maths. The FE analysis suggests that being set for maths increases the children's non-cognitive problems; in the pooled sample, the teacher reported total difficulties score is 0.39 higher for children who are set relative to children who are not set, ceteris paribus. For the pooled sample, however, these effects are insignificant. For boys, setting increases the non-cognitive problems, measured by the SDQ score; 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 on non-cognitive problems through the FE results fail to provide evidence of a significant effect of being set.

Table 5: Effect of setting: OLS & Fixed effects results – 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	All N=7,913 (N=6,010)			
(Number of children N)	Male N= 3,917 (N=2,991)			
	Female N=3,996 (N=3,019)			

Notes: (i) Additional 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. (ii) Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The type of behaviour exhibited as a result of setting may be examined to identify whether internalising, externalising or a combination of both behavioural problems drive the result. The results, given in Table 6 indicate that setting increases both the parent reported externalising problems and the teacher reported internalising problems of children; being set increases the teacher internalising behaviours SDQ score by 0.28, ceteris paribus. This results appears to be driven by boys whose teacher reported internalising problems increase by 0.47 due to being set for maths. For girls, the results suggest that parent reported externalising behaviours drive the OLS results which suggest an increase in parent reported total difficulties. The results highlight the importance of analysing the effect of setting by gender since girls and boys do not respond to setting in a similar manner. This is somewhat unsurprising given the evidence on gender differences in non-cognitive skills and behaviour at this age (McNeish and Scott, 2014; Leadbeateret al. 1999), alongside the identified gender disparities in subject confidence and attitudes (see for example OECD, 2014). Whilst the effect of setting is likely to vary by gender, it is highly likely, given the existing literature, that this overall effect of setting reflects the differential impact of the level of set placement, as the IV analysis explores.

Table 6: Effect of setting: OLS & Fixed effects results – 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: (i) Additional 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. (ii) Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.2 Analysis of set level

The results from the IV analysis of set level, presented in Table 7, indicate that placement in the lowest set has a negative effect on the teacher reported total difficulties score for the pooled sample and for boys, thereby reducing non-cognitive problems. For girls, teacher reported non-cognitive problems are found to increase when placed in the lowest set. Parent reported non-cognitive problems are found to increase as a result of setting in all samples. Within the IV analysis however, an overall insignificant effect of setting and placement in the lowest maths set is identified in all analyses; this result is identified when the base group of comparison 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 for non-cognitive outcomes. This finding supports the existing evidence which finds an insignificant effect of ability grouping on a range of non-cognitive outcomes including self-esteem (Abadzi, 1985) and self-concepts (Ireson and Hallam, 2005). OLS results are given in Table A4 with full IV results in Table A5.

Table 7: Instrumental variable (IV) model main results

	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)
	0.247	-0.515	0.527	0.694	2.891	-0.375
Lowest set	(1.325)	(3.169)	(1.240)	(1.448)	(3.313)	(1.488)
	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) Additional controls include: 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. (ii) Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ (iii) [1] Sargan statistic for the overidentification test of all instruments with p-value given in parentheses (iv) [2] Weak identification test of Cragg-Donald Wald (v) [3] χ^2-statistic for the Anderson Canon under-identification test with p-values given in parentheses (vi) [4] coefficient of the number of maths sets instrument in the reduced form model, with standard errors given in parenthesis (vii) [5] coefficient of the proportion of EAL instrument in the reduced form model, with standard errors given in parenthesis (viii) [6] coefficient of the number of maths sets instrument in the first stage, with standard errors given in parenthesis (ix) [7] coefficient of the proportion of EAL instrument in the first stage, with standard errors given in parentheses.

4.3 Robustness checks

The FE analysis of the impact of set may include individuals who are set for maths, but also set for other subjects for example English or science, though the data suggests that setting in such subjects is much less likely. A binary variable is entered into the model to control for whether an individual is set in other subjects though one concern is that there may be a multiplicative effect of being set in each individual subject upon non-cognitive outcomes. To explore whether the results of the main analysis still hold when solely setting in maths is experienced, the analysis is restricted to individuals who are not set for other subjects. The results, given in Table 8, continue to signal a positive and significant effect of being set on the teacher reported non-cognitive problems in maths for boys; this effect is of much greater magnitude than the main analysis which controlled for but included children who were set in other subjects. For girls, an insignificant effect of setting on non-cognitive problems continues to be identified in the FE models. Based on conjecture alone, this larger 'pure' effect of setting in maths which is not contaminated by the effects of setting in other subjects maybe larger than the main analysis results 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 & Samuelson, 2015; Hargreaves et al. 2008). Men 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, for the low ability children, being informed of their relative ability and rank may impact their self-concepts, confidence and self-perceptions. For higher set boys, the increased difficulty of teaching and materials may challenge their prior perceptions of their ability and test their maths ability to a greater extent than in mixed ability classes.

One concern of the main IV analysis is that around 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; Kramarz et al. 2008), SEN may impact upon non-cognitive outcomes; using MCS data, Fauth et al. (2014) identify children with SEN encounter a greater increase in peer, hyperactivity and emotional problems over time, relative to non-SEN children. The behavioural trajectories of SEN children relative to non-SEN children are therefore likely to diverge over the primary school years. The results from the IV analysis indicate that SEN is a strong significant predictor of non-cognitive problems both reported by the teacher and a parent at age 11 (see Table A4). Although SEN is controlled for in the main IV model, there may be unobservable characteristics associated with SEN children that are not controlled for in the model which may influence the results for children without SEN[16]. SEN children may also respond differently to setting. In order to check the robustness of the results for non-SEN children, the IV analysis is undertaken when SEN children are dropped from the sample. The results, presented in Table 9 continue to indicate an insignificant effect of low set placement and setting in maths upon the non-cognitive outcomes of children, specifically those without SEN.

On further concern of the main analysis is the omission of ability from the model since ability may influence the effect of set placement on non-cognitive outcomes. The reason for this omission is due to ability being highly correlated with the set an individual is placed in. By including ability in the model, differences in behaviour between set and not set cannot be due to ability; this would mean separating out the effect of setting from the effect of ability so that the impact of set is not reflecting the ability level. This may reduce the potential for a confounding variable or omitted variable bias to arise. The MCS progress in maths[17] total raw score is used to control for maths ability. This 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 10 generally provide evidence of a negative relationship between maths ability the total SDQ score; children with higher ability have significantly fewer non-cognitive problems. Even after controlling for ability, setting and lowest set placement continues to have an insignificant effect on the non-cognitive outcomes of children as in the main model.

Overall, the main results and robustness checks identify an insignificant effect of being placed in the lowest set relative to non-set children; the non-cognitive behavioural problems of children in the lowest set and

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

[17] 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.

children who are not set for maths are insignificantly different, ceteris paribus. These findings fall in line with the existing literature on cognitive outcomes (Hallam and Parsons, 2014; Gamoran, 2002; Suknandan and Lee, 1998; Slavin, 1988; Ireson, 1999a).

Table 8: Fixed effect (FE) robustness results – exclusion of other subject set children

	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) Additional 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. (ii) Standard errors in parentheses; * p < 0.1, ** p < 0.05, *** p

Table 9: Robustness of IV results – exclusion of SEN children

	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^2-statistic [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) Additional controls include: 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. (ii) Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ (iii) [1] Sargan statistic for the overidentification test of all instruments with p-value given in parentheses (iv) [2] Weak identification test of Cragg-Donald Wald (v) [3] χ^2-statistic for the Anderson Canon under-identification test with p-values given in parentheses (vi) [4] coefficient of the number of maths sets instrument in the reduced form model, with standard errors given in parenthesis (vii) [5] coefficient of the proportion of EAL instrument in the reduced form model, with standard errors given in parenthesis (viii) [6] coefficient of the number of maths sets instrument in the first stage, with standard errors given in parenthesis (ix) [7] coefficient of the proportion of EAL instrument in the first stage, with standard errors given in parentheses

Table 10: Robustness of IV results – controls for Age 7 maths ability

	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 Ï‡^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) Additional controls include: 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. (ii) Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ (iii) [1] Sargan statistic for the overidentification test of all instruments with p-value given in parentheses (iv) [2] Weak identification test of Cragg-Donald Wald (v) [3] χ^2-statistic for the Anderson Canon under-identification test with p-values given in parentheses (vi) [4] coefficient of the number of maths sets instrument in the reduced form model, with standard errors given in parenthesis (vii) [5] coefficient of the proportion of EAL instrument in the reduced form model, with standard errors given in parenthesis (viii) [6] coefficient of the number of maths sets instrument in the first stage, with standard errors given in parenthesis (ix) [7] coefficient of the proportion of EAL instrument in the first stage, with standard errors given in parenthesis

5. DISCUSSION

Setting is an ability grouping practise which has been, and continues to be, adopted in primary schools across the UK. The school-level policy aims to improve teaching and pupil outcomes by narrowing the ability distribution of pupils in the class for a specific subject, allowing teaching to be more concentrated and tailored to the ability and needs of specific ability groups. 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 has examined the effect of setting in primary school on the non-cognitive outcomes of pupils at age 7 and 11, measured by the Strengths and Difficulties Questionnaire (SDQ). The effect of being set for maths between age 7 and age 11 was firstly examined by adopting OLS and fixed effect methodologies. Secondly, the effect of being placed in the lowest set upon non-cognitive outcomes was estimated when addressing issue of endogeneity by employing an IV approach.

Overall, the results suggest that the movement from maths teaching in a mixed ability class to an ability set has a negative impact upon non-cognitive problems; the overall impact of setting for all ability level pupils is therefore negative. This effect is found for boys whose teacher reported non-cognitive problems are increased by setting. This effect is due to an increase in teacher reported internalising problems, due to setting. This result is robust to the exclusion of children who are set in other subjects to obtain the independent effect of setting in maths.

The analysis of the effect of being placed in the lowest maths set identifies an insignificant effect of being placed in the lowest set 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. These findings are in line with Abadazi (1985) who finds an insignificant effect of setting on non-cognitive outcomes. The concerns behind setting largely relating 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 in itself is therefore interesting, especially since the evidence on setting and non-cognitive outcomes is, at present, very limited. If the non-cognitive development of lower set pupils is insignificantly influenced by setting, then the effects of setting upon cognitive outcomes, which are the primary concern of setting, may be purely focused on.

The findings also provide no evidence that setting in maths harms the non-cognitive development of girls which could otherwise be blamed for the gender gaps in confidence and self-esteem in maths.

Whilst future research on setting should continue to consider the potential for effects on non-cognitive outcomes, research on cognitive outcomes should consider the possible methodological issues associated with estimating the effect of setting. Issues such as unobserved heterogeneity and endogeneity been addressed in very few studies within the entirety of the ability grouping literature.

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. 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 during the pre-teen years.

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

Barker-Lunn, J. (1970) Streaming in the primary school: A Longitudinal Study of Children in Streamed and Non-streamed Junior Schools. (Slough, NFER)

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

Boaler, J. (1997) Setting, social class and survival of the quickest. British educational research journal 23: 575-595

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

Bradley, S. & Taylor, J. (2008) Do peers matter? Estimation of peer effects from pupil mobility between schools. Department of Economics. Lancaster.

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 Economic 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.

Clark, A., Fleche, S. & Lekfuangfu, W. (2017) The long-lasting effects of family and childhood on adult wellbeing: Evidence from British cohort data. CEP discussion paper No. 1493

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

Feinstein, L. & Symons, J. (1999) attainment in secondary school. Oxford Economic Papers. 51, 300–321.

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

Gallop, K., Rose, n. Wallace, E., Williams, R., Cleary, A., Thompson, A., Burston, K., Frere-smith, T. Dangerfield, P., & Tietz, S. (2013) Millennium cohort study fifth sweep (mcs5).

Gamoran, A. & Berends, M. (1987) The effects of stratification in secondary school: synthesis of survey and ethnographic research. Review of educational research, 57, 415-435.

Gamoran, A. (2002) Standards, inequality & ability grouping in schools. Centre for Educational Sociology Briefing 25.

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. Journal of Child Psychology and Psychiatry, 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

Gutman, L. & Schoon, L. (2013) The impact of non-cognitive skills on outcomes for young people. Education Endowment Foundation

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

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. Clark, H. & Hallam, S. (2001) Ability grouping in Education. London, sage.

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

Ireson, J., Hallam, S. Mortimore, P., Hack, S. Clark, H. Plewis, I. (1999b) Ability grouping in the secondary school: the effects on academic achievement and pupils' self-esteem. Retieved 24th March 2015 from http://www.leeds.ac.uk/educol/documents/00001359.htm

Kelly, Y., Kelly, A., & Sacker, A. (2013). Changes in bedtime schedules and behavioral difficulties in 7 year old children. Pediatrics. 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

Kramarz, F., Machin, S. & Ouazas, A. (2008) What Makes a Test Score? The Respective Contributions of Pupils, Schools, and Peers in Achievement in English Primary Education. IZA discussion paper No. 3866

Kulick, J. & Kulik, C. (1992) Meta –analytic findings on grouping programs. Gifted child Quarterly 36(2) 73-77

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. Paserman, D. & Schlosser, A. (2011) Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. The Economic Journal 122:208-237

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)

McNeish, Di & Scott, S. (2014) Women and girls at risk: Evidence across the life course. DMSS retrieved November 2017 from http://www.lankellychase.org.uk/assets/0000/2675/Women___Girls_at_Risk_-_Evidence_Review_040814.pdf

National centre for social Research (NatCen) (2008) Child of the New Century Age 7 Survey Teacher Questionnaire – England. November 2017 from http://www.esds.ac.uk/doc/6848/mrdoc/pdf/mcs4_teacher_england.pdf

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.

Richards. M & Abbott, R. (2009) Childhood mental health and life chances in post-war Britain: insights from three national birth cohort studies. Centre for Mental Health London.

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

Robertson, D. & Symons, J.(2003) Do peer groups matter? Peer groups versus schooling effects on Academic Attainment. Economica 70: 31-53

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.

Scott, S., Doran, M. Beckett, C. Harry, S. & Cartwright, S. (2010) How is parenting style related to child antisocial behaviour? Preliminary findings from the Helping Children Achieve Study. Department for Education Research Report DFE-RR185a

SDQ info (2014a) What is the SDQ? Retrieved from http://www.sdqinfo.com/a0.html November 2017

SDQ info (2017) Searchable database of SDQ publications. Retrieved November 2017 from http://www.sdqinfo.org/py/sdqinfo/f0.py

Slavin (1988) Synthesis of research on grouping in elementary and secondary schools. Educational leadership.

Sukhnandan, L. and Lee, B. (1998) Streaming, setting and grouping by ability: a review of the literature. Slough: NFER.

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.

Wang, L. (2014) All work and no play? The effects of ability sorting on students' non-school inputs, time use and grade anxiety. Economics of Education Review, 44 29-41

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

APPENDIX

Table A1: SDQ questions, sub-scales and measures

SCALE / CATEGORY	ATTRIBUTE	INCLUDED IN TOTAL DIFFICULTIES SCORE	INTERNALISING OR EXTERNALISING BEHAVIOUR
Emotional	Has many worries, often seems worried	✓	Internalising
	Often unhappy, downhearted, tearful	✓	Internalising
	Complains of headache / sickness	✓	Internalising
	Has many fears, is easily scared	✓	Internalising
	Nervous / clingy in new situations	✓	Internalising
Peer	Picked on or bullied by other children	✓	Internalising
	Often solitary, plays alone	✓	Internalising
	Has at least one good friend	✓	Internalising
	Generally liked by other children	✓	Internalising
	Gets on better with adults than children	✓	Internalising
Hyperactivity	Easily distracted, concentration wanders	✓	Externalising
	Sees tasks through to the end	✓	Externalising
	Constantly fidgeting or squirming	✓	Externalising
	Restless, overactive, cannot stay still	✓	Externalising
	Thinks things through before acting	✓	Externalising
Conduct	Often has temper tantrums	✓	Externalising
	Generally obedient	✓	Externalising
	Fights with or bullies other children	✓	Externalising
	Often lies or cheats	✓	Externalising
	Steals from home, school, elsewhere	✓	Externalising
Prosocial	Considerate of other people’s feelings		
	Shares readily with other children		
	Helpful if someone is hurt, upset or unwell		
	Kind to younger children		
	Often volunteers to help others		

Table A2: Variable definitions

VARIABLE	DESCRIPTION
DEPENDENT VARIABLES (Y)	
All dependent variables are taken from the Strengths and Difficulties Questionnaire (SDQ) which is completed by both the child's main parent and their teacher at age 7 and age 11	
Total SDQ	Total difficulties score – continuous variable with a maximum of 40
Internal SDQ	Internalising difficulties score – continuous variable with a maximum of 20
External SDQ	Externalising difficulties score – continuous variable with a maximum of 20
KEY INDEPENDENT VARIABLE	
Set maths	1 if child is set for mathematics, 0 otherwise
Low set maths	1 if the child is in the lowest maths set in Y6, 0 otherwise (other set, not set) This is the variable of interest for the IV analysis
Lowest maths set	1 if the child is set and in the lowest maths set; 0 otherwise
INSTRUMENTAL VARIABLES (Z)	
Proportion of EAL peers	Proportion of children in child's class who have a home with an additional language
Number of maths sets	Number of maths set in the child's school year
EXPLANATORY VARIABLES (X)	
Birth Order	A continuous variable indicating the birth order of the child.
Birth Weight	A continuous variable indicating the Birth weight of the child in lbs.
Regular bedtime	1 if the child has a regular bedtime, 0 otherwise.
DOB Autumn/ winter	1 if the child is born in the autumn or winter months, 0 otherwise.
Male	1 if the child is Male, 0 otherwise.
White	1 if the child is white, 0 otherwise.
SEN	1 if the child is classified as having special educational needs, 0 otherwise.
Religious non-participation	1 if the child does not attend religious service, 0 otherwise
Siblings in Household	Number of siblings living in the same household.
Parent degree	1 if one parent has a degree, 0 otherwise.
Married	1 if main parent is married or in a civil partnership, 0 otherwise.
Working Household	1 if at least 1 parent from the household is working, 0 otherwise.
Ln income	Natural logarithm of equalised income (income adjusted for size and composition).
Parent evening	1 if anyone has attended parents evening, 0 otherwise.
School club	1 if the child attends an after school club; 0 otherwise.
Mixed year group	1 if child's class contains mixed year group, 0 otherwise.
Class size	Number of children in the child's class.
Teacher tenure	Number of years the teacher has taught at the school.
Teacher years	Number of years as a teacher (teacher experience)
Disruptive peers	1 if teacher reports that there are children in child's class who prevent learning; 0 otherwise
Proportion of excluded peers	Proportion of child's class that have been excluded from school
Set for another subject	1 if set for another subject (English observed at age 7, English and science observed at age11)
IMD quartile 1	1 if IMD score (LSOA deprivation score) is in most deprived quartile; 0 otherwise

IMD quartile 2	1 if IMD score (LSOA deprivation score) is in second quartile; 0 otherwise
IMD quartile 3	1 if IMD score (LSOA deprivation score) is in third quartile; 0 otherwise
IMD quartile 4	1 if IMD score (LSOA deprivation score) is in least deprived quartile; 0 otherwise

Table A3: Full fixed effects model results

	Teacher total difficulties			**Parent total difficulties**		
	Pooled	Male	Female	Pooled	Male	Female
Set Maths	0.394	0.741^{*}	0.021	0.223	0.298	0.122
	(0.246)	(0.384)	(0.314)	(0.201)	(0.302)	(0.272)
Regular Bedtime	0.394	0.151	0.625	-0.169	-0.178	-0.056
	(0.342)	(0.505)	(0.465)	(0.280)	(0.397)	(0.402)
SEN	1.466^{***}	1.286^{***}	1.551^{***}	1.009^{***}	0.686^{**}	1.403^{***}
	(0.295)	(0.425)	(0.411)	(0.242)	(0.334)	(0.356)
Siblings HH	-0.353	-0.830^{**}	0.114	-0.117	-0.036	-0.128
	(0.217)	(0.334)	(0.279)	(0.177)	(0.262)	(0.241)
Married	0.388	0.741	0.179	-0.324	0.076	-0.560
	(0.387)	(0.632)	(0.477)	(0.317)	(0.496)	(0.413)
Working HH	-0.647	-0.962	-0.438	-0.413	-0.266	-0.620
	(0.443)	(0.723)	(0.542)	(0.362)	(0.568)	(0.469)
Log income	0.085	0.511	-0.357	0.420^{*}	0.329	0.531^{*}
	(0.266)	(0.402)	(0.354)	(0.218)	(0.315)	(0.307)
Parents Eve.	0.328	0.627	-0.029	-0.161	0.489	-0.698
	(0.460)	(0.743)	(0.572)	(0.376)	(0.584)	(0.495)
School club	-0.081	-0.449	0.158	0.083	0.194	-0.011
	(0.195)	(0.303)	(0.250)	(0.159)	(0.238)	(0.217)
Mixed year grp.	0.006	0.128	-0.079	-0.239	-0.385	-0.065
	(0.259)	(0.394)	(0.338)	(0.212)	(0.309)	(0.293)
Class size	-0.005	-0.022	0.005	-0.035^{**}	-0.053^{**}	-0.017
	(0.020)	(0.031)	(0.027)	(0.017)	(0.024)	(0.023)
Teacher tenure	0.005	-0.023	0.025	0.025^{*}	0.040^{*}	0.009
	(0.018)	(0.029)	(0.022)	(0.015)	(0.023)	(0.019)
Teach. years	-0.016	0.011	-0.039^{**}	-0.016	-0.027	-0.006
	(0.014)	(0.022)	(0.017)	(0.011)	(0.017)	(0.014)
Proportion of class peers excluded	0.056^{*}	0.145^{***}	0.006	0.009	0.050	-0.015
	(0.029)	(0.052)	(0.034)	(0.024)	(0.041)	(0.029)
Disruptive peers	0.747^{***}	1.056^{***}	0.500^{*}	-0.044	-0.071	0.009
	(0.198)	(0.307)	(0.255)	(0.162)	(0.241)	(0.221)
Set for English or Science	-0.203	-0.202	-0.204	-0.056	0.123	-0.207
	(0.167)	(0.257)	(0.216)	(0.137)	(0.202)	(0.187)
IMD quartile 1	1.353	1.714	0.946	-0.464	-0.700	-0.088
	(0.871)	(1.317)	(1.149)	(0.713)	(1.035)	(0.995)
IMD quartile 2	-1.678^{*}	-2.863^{**}	-0.127	-1.101	-2.376^{**}	0.525
	(0.896)	(1.324)	(1.215)	(0.734)	(1.040)	(1.052)
IMD quartile 3	1.168	1.696	0.987	-0.872	-0.477	-1.050
	(0.921)	(1.646)	(1.058)	(0.754)	(1.293)	(0.916)
IMD quartile 4	-0.916	-2.012	-0.562	-0.123	2.757^{*}	-1.841
	(1.153)	(2.021)	(1.348)	(0.944)	(1.588)	(1.167)
Wave	-1.197	-2.918^{*}	0.645	-1.469	-1.100	-1.905
	(1.163)	(1.755)	(1.546)	(0.952)	(1.379)	(1.338)
Constant	10.364^{***}	16.912^{***}	3.708	13.048^{***}	11.755^{***}	13.671^{***}
	(3.376)	(5.237)	(4.374)	(2.764)	(4.115)	(3.786)
Observations	7913	3917	3996	7913	3917	3996

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

Table A4: Full OLS results

	Teacher total difficulties			Parent total difficulties		
	Pooled	Male	Female	Pooled	Male	Female
Set Maths	-0.697^{***}	-0.574^{*}	-0.759^{***}	-0.365^{*}	-0.246	-0.448^{*}
	(0.183)	(0.293)	(0.223)	(0.200)	(0.302)	(0.265)
Low set	1.941^{***}	1.665^{***}	2.314^{***}	1.192^{***}	0.228	2.106^{***}
	(0.229)	(0.369)	(0.278)	(0.251)	(0.380)	(0.330)
Lowest set	1.244^{***}	1.091^{***}	1.554^{***}	0.826^{***}	-0.018^{***}	1.658^{***}
	(0.249)	(0.400)	(0.304)	(0.273)	(0.413)	(0.361)
Born A/W	-0.009	-0.438^{**}	0.430^{**}	-0.109	-0.126	-0.067
	(0.137)	(0.215)	(0.169)	(0.150)	(0.222)	(0.201)
Male	1.972^{***}			0.778^{***}		
	(0.139)			(0.153)		
White	0.549^{***}	0.482	0.672^{**}	0.182	0.261	0.187
	(0.207)	(0.322)	(0.261)	(0.227)	(0.332)	(0.310)
SEN	4.128^{***}	4.553^{***}	3.283^{***}	4.339^{***}	5.027^{***}	3.327^{***}
	(0.185)	(0.265)	(0.258)	(0.202)	(0.273)	(0.307)
Religion	-0.012	-0.067	0.005	0.296^{*}	0.470^{**}	0.164
	(0.143)	(0.227)	(0.176)	(0.157)	(0.234)	(0.209)
Birth Order	0.058	0.085	0.045	-0.171^{**}	-0.336^{***}	-0.023
	(0.079)	(0.125)	(0.098)	(0.087)	(0.129)	(0.116)
Birth Weight	-0.026	0.014	-0.081	-0.021	0.009	-0.061
	(0.053)	(0.083)	(0.065)	(0.058)	(0.086)	(0.077)
Siblings HH	-0.194^{**}	-0.359^{***}	-0.020	-0.185^{**}	-0.061	-0.258^{**}
	(0.085)	(0.133)	(0.107)	(0.094)	(0.137)	(0.128)
Parent degree	-0.342	-0.360	-0.386	-0.989^{***}	-0.521	-1.379^{***}
	(0.221)	(0.350)	(0.273)	(0.243)	(0.361)	(0.324)
Married	-0.558^{***}	-0.852^{***}	-0.285	-0.596^{***}	-0.809^{***}	-0.369^{*}
	(0.151)	(0.236)	(0.188)	(0.166)	(0.244)	(0.224)
Working HH	-1.413^{***}	-1.190^{***}	-1.600^{***}	-1.223^{***}	-1.244^{***}	-1.289^{***}
	(0.234)	(0.365)	(0.295)	(0.257)	(0.377)	(0.350)
Log income	-1.124^{***}	-1.602^{***}	-0.614^{**}	-1.779^{***}	-1.653^{***}	-1.763^{***}
	(0.248)	(0.386)	(0.312)	(0.272)	(0.398)	(0.372)
Parents Eve.	-0.628^{*}	-0.765	-0.388	-0.860^{**}	-0.541	-1.087^{**}
	(0.329)	(0.497)	(0.425)	(0.360)	(0.513)	(0.506)
School club	-0.059	0.018	-0.090	-0.101	-0.050	-0.153
	(0.140)	(0.222)	(0.171)	(0.153)	(0.229)	(0.203)
Mixed year grp.	0.264^{*}	-0.022	0.522^{***}	0.198	0.169	0.234
	(0.158)	(0.247)	(0.197)	(0.173)	(0.254)	(0.234)
Class size	-0.030^{**}	-0.062^{***}	0.009	-0.033^{**}	-0.066^{***}	0.013
	(0.013)	(0.020)	(0.017)	(0.015)	(0.021)	(0.021)
Teacher tenure	-0.011	-0.008	-0.015	0.014	0.053^{**}	-0.022
	(0.014)	(0.023)	(0.017)	(0.016)	(0.024)	(0.021)
Teach. years	-0.007	-0.006	-0.008	-0.007	-0.026	0.010
	(0.011)	(0.017)	(0.013)	(0.012)	(0.017)	(0.015)
Prop. peers excluded	0.047^{**}	0.059	0.023	0.058^{**}	0.078^{**}	0.034
	(0.023)	(0.037)	(0.027)	(0.025)	(0.038)	(0.032)
Disruptive peers	0.897^{***}	1.121^{***}	0.635^{***}	0.096	0.229	-0.066
	(0.148)	(0.234)	(0.183)	(0.162)	(0.241)	(0.217)
Set for Eng. or Sci.	0.193	0.005	0.315	0.191	0.073	0.233
	(0.177)	(0.282)	(0.217)	(0.194)	(0.291)	(0.258)
IMD quartile 1	-0.111	-0.319	0.172	-0.177	-0.114	0.013
	(0.245)	(0.376)	(0.313)	(0.268)	(0.388)	(0.372)
IMD quartile 2	-0.126	-0.153	-0.057	-0.209	-0.547	0.178
	(0.231)	(0.356)	(0.293)	(0.253)	(0.367)	(0.348)
IMD quartile 3	-0.439^{*}	-0.487	-0.331	-0.701^{***}	-0.766^{**}	-0.524

	(0.239)	(0.373)	(0.301)	(0.262)	(0.384)	(0.357)
IMD quartile 4	-0.696***	-0.799**	-0.580*	-0.757***	-0.730*	-0.733*
	(0.258)	(0.395)	(0.330)	(0.283)	(0.408)	(0.392)
Constant	18.184***	26.140***	11.492***	29.257***	29.075***	28.091***
	(2.472)	(3.863)	(3.103)	(2.710)	(3.984)	(3.691)
N	4779	2361	2418	4779	2361	2418

Table A5: Full IV model results

	Teacher total difficulties			Parent total difficulties		
	Pooled	Male	Female	Pooled	Male	Female
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)
Regular Bedtime	0.058	0.040	0.105	-1.059***	-1.310***	-0.850**
	(0.231)	(0.359)	(0.291)	(0.252)	(0.376)	(0.349)
Born A/W	-0.058	-0.523*	0.387**	-0.116	0.029	-0.152
	(0.150)	(0.273)	(0.176)	(0.164)	(0.285)	(0.211)
Male	1.931***			0.772***		
	(0.149)			(0.163)		
White	0.573***	0.566	0.662**	0.185	0.109	0.168
	(0.209)	(0.361)	(0.261)	(0.229)	(0.378)	(0.313)
SEN	4.429***	5.031***	3.608***	4.379***	4.161***	3.969***
	(0.435)	(0.973)	(0.459)	(0.475)	(1.017)	(0.551)
Religion	0.023	-0.008	0.040	0.301*	0.364	0.233
	(0.150)	(0.255)	(0.180)	(0.164)	(0.266)	(0.216)
Birth Order	0.074	0.104	0.065	-0.169*	-0.371***	0.017
	(0.082)	(0.131)	(0.101)	(0.090)	(0.137)	(0.121)
Birth Weight	-0.030	-0.001	-0.078	-0.021	0.036	-0.054
	(0.053)	(0.088)	(0.065)	(0.058)	(0.092)	(0.078)
Siblings HH	-0.198**	-0.339**	-0.042	-0.186**	-0.097	-0.303**
	(0.086)	(0.139)	(0.110)	(0.094)	(0.145)	(0.132)
Parent degree	-0.375*	-0.415	-0.415	-0.993***	-0.421	-1.437***
	(0.226)	(0.367)	(0.275)	(0.247)	(0.383)	(0.329)
Married	-0.565***	-0.826***	-0.313	-0.597***	-0.854***	-0.425*
	(0.151)	(0.242)	(0.191)	(0.165)	(0.253)	(0.229)
Working HH	-1.444***	-1.257***	-1.626***	-1.227***	-1.122***	-1.342***
	(0.238)	(0.389)	(0.296)	(0.260)	(0.406)	(0.355)
Log income	-1.166***	-1.618***	-0.691**	-1.785***	-1.624***	-1.915***
	(0.255)	(0.388)	(0.325)	(0.278)	(0.406)	(0.390)
Parents Eve.	-0.628*	-0.726	-0.417	-0.860**	-0.614	-1.144**
	(0.329)	(0.504)	(0.426)	(0.359)	(0.527)	(0.511)
School club	-0.068	0.004	-0.101	-0.102	-0.025	-0.175
	(0.140)	(0.224)	(0.171)	(0.153)	(0.234)	(0.205)
Mixed year grp.	0.214	-0.094	0.470**	0.191	0.300	0.130
	(0.170)	(0.285)	(0.206)	(0.186)	(0.298)	(0.247)
Class size	-0.032**	-0.061***	0.005	-0.033**	-0.067***	0.004
	(0.014)	(0.020)	(0.018)	(0.015)	(0.021)	(0.022)
Teacher tenure	-0.012	-0.008	-0.016	0.014	0.053**	-0.023
	(0.014)	(0.023)	(0.017)	(0.016)	(0.024)	(0.021)
Teach. years	-0.007	-0.007	-0.008	-0.007	-0.024	0.010
	(0.011)	(0.017)	(0.013)	(0.011)	(0.018)	(0.015)
Prop. peers excluded	0.041*	0.051	0.017	0.057**	0.091**	0.021
	(0.024)	(0.040)	(0.028)	(0.026)	(0.042)	(0.034)
Disruptive peers	0.883***	1.106***	0.615***	0.094	0.256	-0.106
	(0.149)	(0.236)	(0.184)	(0.163)	(0.247)	(0.221)
Set for Eng. or Sci.	0.175	-0.027	0.301	0.189	0.129	0.204
	(0.179)	(0.289)	(0.217)	(0.195)	(0.302)	(0.260)
IMD quartile 1	-0.136	-0.309	0.111	-0.180	-0.133	-0.108
	(0.247)	(0.377)	(0.321)	(0.270)	(0.394)	(0.384)
IMD quartile 2	-0.162	-0.217	-0.092	-0.214	-0.432	0.108

	(0.235)	(0.378)	(0.295)	(0.257)	(0.395)	(0.354)
IMD quartile 3	-0.463*	-0.527	-0.363	-0.704***	-0.693*	-0.586
	(0.241)	(0.381)	(0.302)	(0.263)	(0.398)	(0.363)
IMD quartile 4	-0.710***	-0.854**	-0.574*	-0.759***	-0.630	-0.722*
	(0.259)	(0.410)	(0.329)	(0.283)	(0.429)	(0.395)
Constant	18.735***	26.255***	12.537***	29.330***	28.867***	30.160***
	(2.576)	(3.873)	(3.332)	(2.813)	(4.049)	(3.997)
N	4779	2361	2418	4779	2361	2418

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