Study of Early Education and Development (SEED): Impact Study on Early Education Use and Child Outcomes up to Age Three

Technical Annex to the Main Report

June 2017: revised November 2021

Edward Melhuish, Julian Gardiner & Stephen Morris

University of Oxford

Acknowledgments

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The purpose of the technical report

This report is an adjunct to the SEED Research Report "Study of Early Education and Development (SEED): Impact Study on Early Education Use and Child Outcomes up to age three: Research report (Melhuish, Gardiner and Morris, 2017)"[1]. This Technical Report gives further details of the analyses given in the Research Report as well as the results of some additional analyses. It is intended to be read in conjunction with the Research Report.

The scope of the report

The research sought to address two main objectives:

1. To explore the impact of introducing a policy of free early education for disadvantaged two-year-olds on take-up of early education for two- to three-year-old children, in the year following the introduction of the policy.

2. To study the associations between the amount of differing types of early childhood education and care (ECEC) and child development, as well as associations between child development and aspects of the home environment.

As a result, this technical report, in conjunction with the Research Report, aims to address two main questions:

1. Has the introduction of 570 hours per year of funded childcare for disadvantaged two year olds led to an increase in the use of childcare among these children? This question is addressed in Chapter 3.

2. What influence does the amount of Early Childhood Education and Care (ECEC) which children receive between ages two and three have on their cognitive and socio-emotional outcomes measured at age three? This question is addressed in Chapter 4.

In addition, it aims to investigate the influence of home environment and demographic factors on children's cognitive and socio-emotional outcomes; this material is covered in Chapter 5.

---

[1] Available here at:

Chapter 2: The SEED longitudinal study: design and methodology

Study design

Design constraints

In this section a number of constraints faced in the design of the SEED study are considered. This chapter is to be read in conjunction with Chapter 2 of the Research Report. Key decisions regarding the design and implementation of the policy were made before explicit consideration was given to evaluation and how amenable the design and rollout of the policy were to research. It was not possible to influence the timing of the policy change, how eligibility was defined, and whether the policy changes were restricted in some way to create a control group.

All two-year-olds were in range of the policy change from September 2013 as long as both age and family income criteria were met. There was no scope for randomising eligibility. Moreover, the introduction of the policy change was imminent at the time the research team was engaged. Due to the reliance on primary data collection, this late engagement meant that the study team were unable to specify earlier, pre-policy change. Pre-policy primary data collection would have brought considerable benefits. Due to budgetary considerations it was not possible to sample older cohorts of ineligible children to act as controls, as well as cohorts directly affected by the policy changes that were to be the focus of data collection.

Research design

The basic features of the sampling and data collection are illustrated in Figure 1. The population of two-year-olds is divided into birth cohorts based on the school-terms in which their birth date falls. Each cohort is further sub-divided by a proxy indicator of family income derived from Child Benefit claim records, benefits data and Tax Credit records.

The cohorts depicted in Figure 1 were chosen as those from which samples were to be drawn and outcome measures obtained. Variation in eligibility for the policy among the target population creates the potential to identify contrasts in eligibility and therefore to study the effects of these contrasts.

Figure 1: SEED longitudinal, multi-cohort research design.
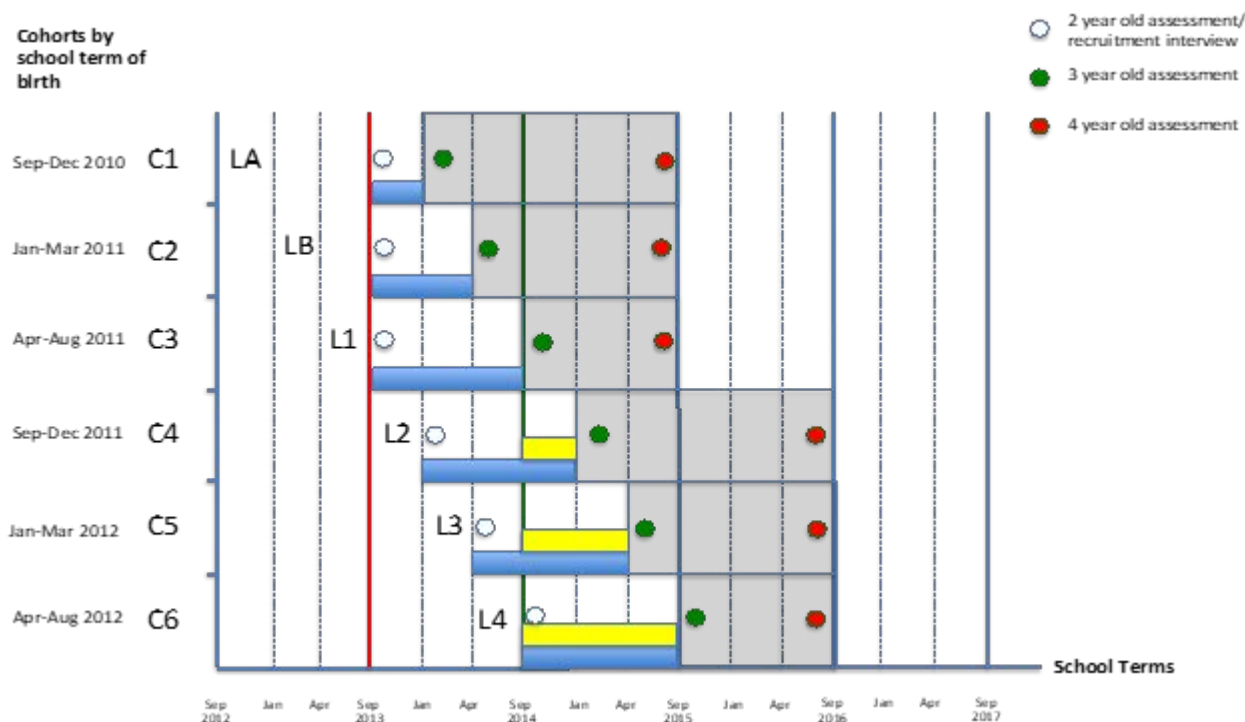
Note: Blue bars indicate the number of terms of eligibility for ECEC prior to children's third birthday for 20% most disadvantaged. Yellow bars indicate the number of terms of eligibility for the >20% to 40% moderately disadvantaged. See below for explanation of LA to L4.

One source of variation in eligibility was through identifying 'transition cohorts'. These are cohorts qualifying for support around the point in time when the policy change came into force. This means that instead of being eligible for three terms of early education support between the ages of two and three, which would be typical, they instead would be eligible for only one or two terms. Comparing outcomes across these transition cohorts is a way to provide an insight into the impact (if any) of the policy change. This variation can be considered to represent different dosages in eligibility. As a result the analysis compared outcomes at age three for low-income families who were eligible for one as opposed to three terms of early education support, or two as opposed to three terms.

An alternative strategy is comparing outcomes among the low-income group of children who qualified for early education support at two years from September 2013, to those among children from the income group just immediately above them, who did not qualify for support until September 2014 – this group is referred to as the moderately disadvantaged group. The complication in this approach is that some movement between income groups over the course of the study is anticipated. But if this movement is minimal then analysis of effectiveness may be obtained through such a comparison.

The vertical axis of Figure 1 depicts the various cohorts of two year olds by their term of birth around two policy changes. For children from the lowest income group (termed 'most disadvantaged' in this report), they became eligible for funded early years education in the term following their second birthday from September 2013 (red vertical

line). By contrast, children from the next lowest income group (termed 'moderately disadvantaged' in this report) become eligible for funded early education from September 2014 (green vertical line).

Figure 1 captures the consequences of these two policy changes for the birth cohorts, defined by school-term of birth, displayed on the vertical axis. Cohorts are labelled C1 through to C6. The pattern of eligibility among the most disadvantaged by cohort is demonstrated by the horizontal blue bars. First consider Cohort C1, born in the school term September to December 2010. This cohort became eligible for funded early education in the term following their second birthday, i.e. in the school term January-April 2013. However, the policy change did not come into force until September 2013; therefore the first term that this cohort were eligible, in official policy terms, was the term September-December 2013. Thus cohort C1 in Figure 1 can receive only one term of funded early education before they turn three and thereby qualify automatically under pre-existing policy for funded early education, represented by the <u>blue</u> horizontal bar.

Likewise cohort C2, born the school term January to March 2011, turned two during the term January to March 2013, and therefore qualified for support from April 2013. Here again, this cohort could not actually benefit from funded early education until the policy change came into effect in September 2013 and thus could only qualify for two terms of support before turning three. The first of the cohorts to qualify for a full three terms of funded early education was cohort C3, born April to August 2011.

Thus, in summary, the blue horizontal bars shows that the lowest income children in C1 qualify for one term of funded early education, C2 two terms, and C3 three terms, and so on, until the policy was fully rolled out and all subsequent low-income cohorts qualify for the full three terms of funded early education.

The same pattern among cohorts can also be seen around the introduction of funded early education among children in the moderately disadvantage group. Their eligibility is represented by the <u>yellow</u> horizontal bars. This policy change came into force around September 2014. In cohort C4, while the most disadvantaged group (<u>blue</u> horizontal bar) is eligible for the full three terms of funded early education the next income group – or moderately disadvantage group - represented by the horizontal <u>yellow</u> bar - is only eligible for one term of funded early education. This is because the official policy change affecting this cohort came into force from September 2014. In the subsequent cohorts C5 and C6, the moderately disadvantaged groups due to their terms of birth and the official policy funding rules qualified for two and three terms of support respectively – again represented by the horizontal yellow bars in Figure 1.

Sampling methodology and data collection

Having described why the cohorts were chosen, this section describes how these cohorts were sampled, and how measures are taken on the variables of interest.

The research design aims to facilitate the evaluation of the policy as well as having a wider objective in providing the basis for a long-term study of early childhood education and care (ECEC). Hence data that permit testing links proposed between setting characteristics, child and home characteristics and developmental outcomes was required.

Sampling individual children

Further details on the sample design for this study can be found in Speight et al., (2015). The key points are summarised here.

Sampling lists for children within the birth cohorts of interest are depicted in Figure 1 along with the timing of their compilation, by the labels 'LA to L4'. Taking cohort C1 as an example; List LA was run in the autumn term 2012, the term these children turned two years. Child Benefit records were used to identify children within cohort C1[2]. The compiled list is then matched to benefit records to identify the most disadvantaged children in the bottom 20% of household income, the next quintile or moderately disadvantage group, and those with household incomes above 40 percentile point of the household income distribution.

Speight et al (2015:61) elucidate the sampling criteria as:
1. The *20% most disadvantaged families* had a parent in receipt of one of[3]:
    a. Income-based Jobseeker's Allowance (JSA-IB);
    b. Income-related Employment Support Allowance (ESA-IR);
    c. Income Support (IS);
    d. Guaranteed element of State Pension Credit (PC with Guarantee Credit);
    e. Child Tax Credit only (not in receipt of an accompanying Working Tax Credit award) with household gross earnings of less than £16,190.
2. The *moderately disadvantaged group (20-40%)* had a parent in receipt of Working Tax Credits with household gross earnings of less than £16,190[4].

---

[2] Note there is some under-coverage of higher income groups in these lists due to changes in the eligibility criteria for receipt of Child Benefit. This under-coverage is discussed in more detail in (Speight et al., 2015).

[3] The full DfE eligibility criteria from September 2013 are: (i) All 2-year-olds who are looked after by their local authority; (ii) 2-year-olds whose family receives one of the following are also eligible: income support; income-based Jobseeker's Allowance (JSA); income-related Employment and Support Allowance (ESA); support through part 6 of the Immigration and Asylum Act; the guaranteed element of State Pension Credit; Child Tax Credit (but not Working Tax Credit) and have an annual income not over £16,190; the Working Tax Credit 4-week run on (the payment you get when you stop qualifying for Working Tax Credit) or Universal Credit.

[4] From September 2014, the eligibility criteria include two-year-olds who meet any one of the following criteria: eligibility criteria also used for free school meals; if their families receive Working Tax credits and have annual gross earnings of no more than £16,190 per year; if they have a current statement of special educational needs (SEN) or an education, health and care plan; if they attract Disability Living Allowance; if they are looked after by their local authority; or if they have left care through special guardianship or

3. The *not disadvantaged group (>40%)* had parents not in receipt of any of the qualifying benefits or tax credits."

Effectively this process aims to identify those children who would become eligible for funded early education in the term following their second birthday, bearing in mind that the lists were run in the term during which the cohort members turned two years. The following two lists, LB and L1, identify members of the cohort C1 again and through examining their benefit and tax credit records assess whether their eligibility for funded early education has changed since list LA was compiled.

The eligibility rule used to determine the list-compilation process is based on the understanding that once a child becomes eligible they remain so, even if their household income subsequently rises above the 20% threshold; however, were a child's family to see their income decline they will become eligible for support. Thus the cohort C1's family income status is re-assessed in the compilation of lists LB and L1 to determine whether children previously identified as being in a higher income group have subsequently become eligible as a result of their family's income falling. It is worth remembering that all children regardless of their family income become eligible for support in the term after they turn three years. Hence in the lists L2-L4, cohort C1 does not appear and their eligibility is not determined again. By the time L2 is compiled the cohort C1 is eligible for funded early education regardless of their household income.

This process, for identifying children within cohort C1 is replicated for cohorts C2-C6.

Sample selection

A three-stage clustered sample design was used for this study, with sampling from Child Benefit records (Speight et al. 2015). First, postcode districts were designated primary sampling units (PSUs). In the second stage of sampling groups of postal sectors were identified within each PSU and designated as Secondary Sampling Units (SSUs). Finally, a fixed number of eligible families were selected for interview within each SSU.

A three-stage approach was adopted in order to generate a highly clustered sample of children, but also a sample of ECEC settings within the SSUs that the sampled children were likely to use. Thus with each SSU all setting were sampled and quality measures taken. Thus measurement of setting quality could be carried out such that there was a high probability that such settings were those attended by children in the sample, and setting quality measures could subsequently be linked to children's sample records.

As Speight et al (2015: 60) elaborate further:

through an adoption or child arrangements order.

"In practice the sampling was done in three stages:

- 111 PSUs were selected in proportion to a weighted sum of the number of eligible families within each PSU (with weights calculated to reflect the final desired proportions of the three disadvantage groups, see below).
- Three SSUs were selected within each PSU in proportion to a weighted sum of the number of eligible families within each SSU.
- Five or six families in each disadvantage group were selected within each SSU in proportion to their weights."

The income groups above were sampled to form approximately equal numbers in the sample. "As the three groups were not of equal size in the population, a weighted sampling approach was used to create as close to an equal probability sample as possible, with weights equal to the ratio of the desired proportion (one third) to the population proportion in each cohort" (Speight, et al, 2015: 60).

A note on causality

Where associations are found between children's outcomes and their use of early childhood education and care (ECEC), the possible causal pathways which may account for these associations need to be considered. Four possible pathways are shown in Figure 2:

1. Simple causation: ECEC usage influences children's developmental outcomes.
2. Reverse causation: child development factors (outcomes) influence children's ECEC usage.
3. Confounding: other unknown factors influence both the ECEC usage and the outcomes.
4. Mediated causation: ECEC usage influences children's outcomes via unobserved mediating factors.

Figure 2: Four possible causal pathways linking ECEC use and children's outcomes.
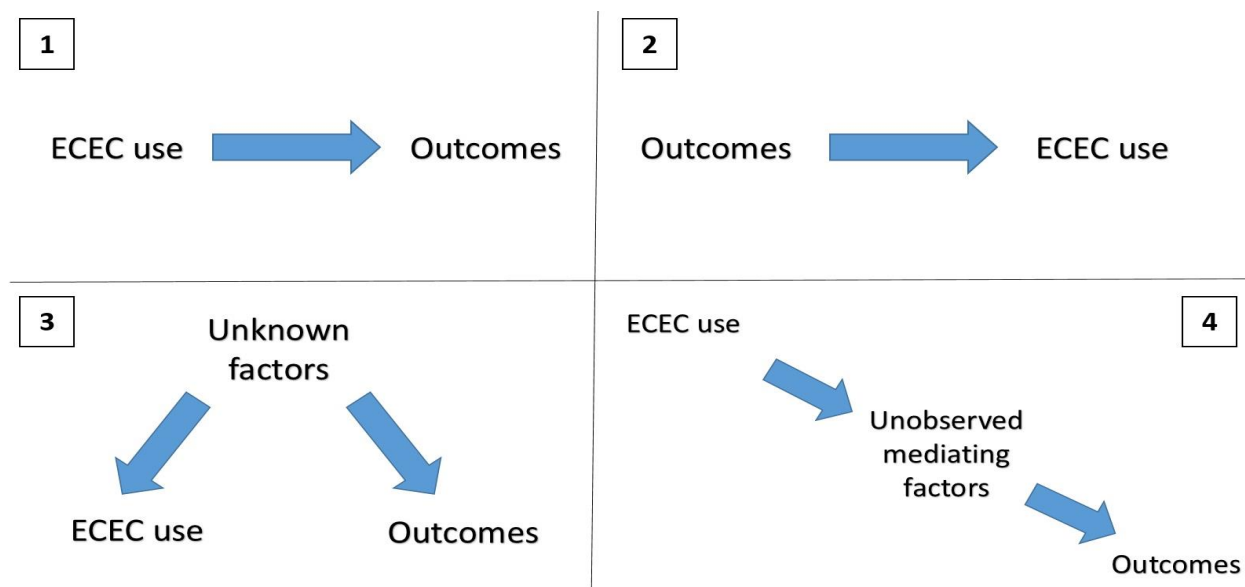
It is possible that all four of these causal pathways are present to some extent. However, there is good reason to believe that some of these pathways are more likely than others to account for associations between ECEC use and children's outcomes.

Reverse causation

Whilst it is probable that parents' decisions about childcare use are influenced sometimes by their children's abilities and behaviour, it is unlikely that this will be the dominant factor behind associations between ECEC use and children's outcomes found across a whole population. In general, it is suggested that parental decisions on ECEC use will be driven by pre-existing beliefs about what are the best childcare arrangements for children and also by family needs, e.g., the need for day care to allow parents to return to work, these beliefs and needs being independent of children's cognitive and behavioural outcomes.

Confounding

Models of child outcomes in terms of ECEC use control for many demographic and home environment factors that might otherwise confound the relationship between ECEC use and children's outcomes. Whilst the existence of other confounding factors not controlled for cannot be ruled out (e.g., mother's personality), the existence of such additional confounders is unlikely to explain the large number of significant associations that exist between ECEC use and children's cognitive and socio-emotional outcomes.

Causation and mediated causation

If, as suggested, reverse causation, confounding and unknown factors are unlikely to explain associations between ECEC use and outcome variables, then it may be cautiously concluded that any associations found are likely to result from causation of the outcomes by exposure to ECEC (module 1 in figure 3). In general this causation will be via mediating factors not been directly observed. For example, the time which a child spends interacting with peers in a given environment or the nature of the ECEC provider / child relationship.

Chapter 3: Introducing funded early education from two years olds for disadvantaged families – effects of policy change

Introduction

Whether the policy change affected the take-up of childcare among the eligible groups of disadvantaged families was investigated using three analytical approaches:

1. Discontinuity analysis.
2. Multiple linear regression.
3. Difference-in-differences analysis.

Discontinuity analysis

Introduction

The first approach to assessing the effects of official eligibility for early education on take-up and use of early education is a regression discontinuity analysis.

The logic of the discontinuity approach as applied in this study can be best understood with reference to Figure 3.

Figure 3: SEED longitudinal, multi-cohort research design
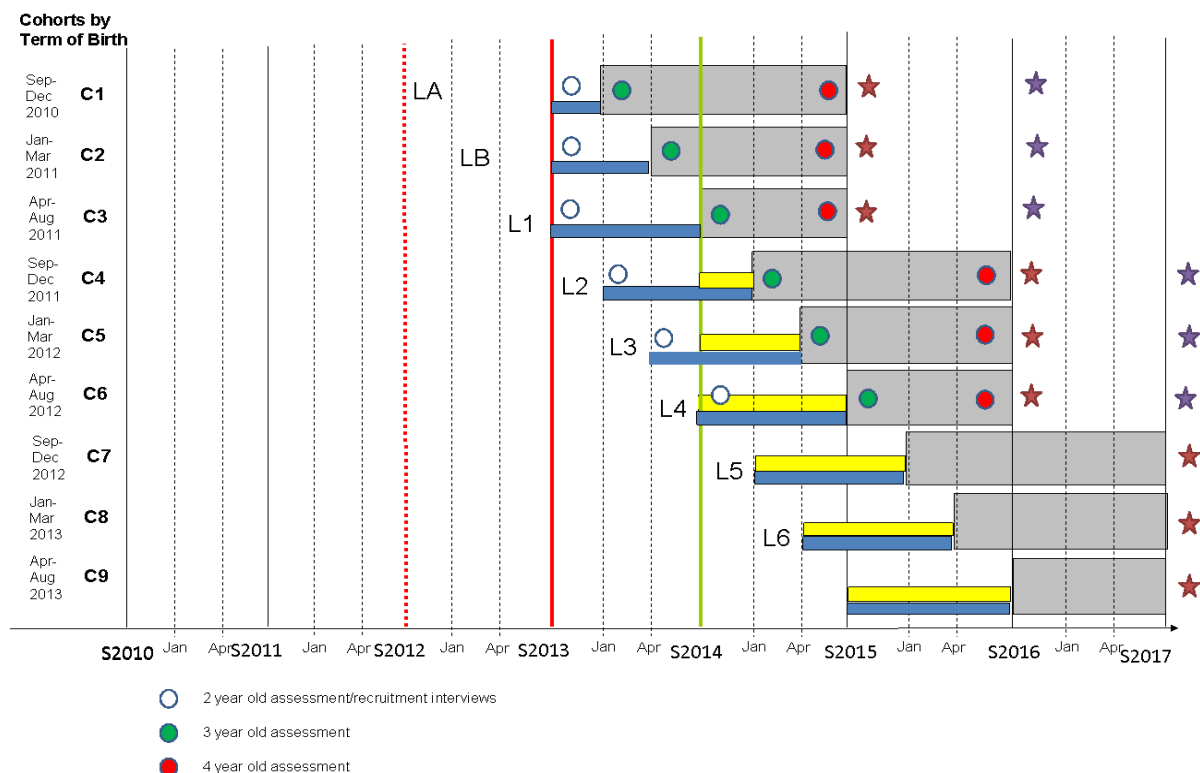
The vertical axis on this chart represents the school term in which a given cohort was born. For example, Cohort C1 were born in the term September to December 2010, likewise cohorts C2 and C3 were born in the terms January to March 2011 and April to August 2011 respectively. Within each cohort, children are further grouped by family income into most disadvantaged (roughly the bottom income quintile), moderately disadvantaged (roughly the second income quintile), and other (all other quintiles).

The purpose of this analysis was to evaluate the effects of variations in official eligibility for funded early education on take-up and use of early education for the 'most' and 'moderately disadvantaged' groups. Looking at the 'most disadvantaged' group by way of illustration, children in this group became eligible for free early education from 1st September 2013. This means that the cohort C1 that turned two in the term September to December 2012 would be eligible for one term of support. This is indicated in Figure 3 by the length of the horizontal blue bar for this cohort. Cohorts C2 and C3 would be eligible for two and three terms of support respectively, and this is indicated by the horizontal blue bars in Figure 3 for these cohorts.

Bearing these features of the data in mind, we can see that by comparing use and take-up among the most disadvantaged children in cohort C1 with use and take-up among the most disadvantaged children in the cohort C2, we are able to assess the effects of being eligible for two terms of support rather than one for the most disadvantaged group. Likewise, a similar comparison between take-up and use in cohorts C2 and C3 allows us to determine the effect of receiving three terms of support relative to two terms.

Comparing cohorts formed on the basis of entire school terms, however, may not result in a fair test of the effects of different term-based official levels of entitlement for early education. Children from cohort C1 may differ in ways besides being eligible for one term of support rather than two, to children in cohort C2. For example, we know that children born earlier in the school year tend on average to performed better in developmental tests than those born later by virtue of the fact that they are older. In general, older children within a school year may also be more likely to take-up early education than those that are younger. Thus a simple comparison of all children in both cohorts may not enable us to isolate the effect of differences in eligibility on take-up and use from all the other ways in which the two cohorts may differ from each other.

One way to potentially avoid this problem is to focus on making comparisons between children born just either side of a cohort boundary, a boundary that will also coincide with differences in eligibility. The argument is that children, for example, born at the end of December 2010 will be very similar to those children born at the beginning of January 2011. They will have reached similar levels of maturity by the time the policy change for most disadvantaged groups comes into force and we would not expect them to perform differently in developmental tests. More importantly, in the absence of the official policy change, there would be no reason to expect take-up and use of early education to differ among children born just either side of the 31st December cohort boundary. They are, however, different in one important respect; those born at the end of December 2010 were eligible for one term's support, whereas those born in early January 2011 qualified

for two terms of support. This means that the analyses should be able to conduct a fair test of the effect of being eligible for one as opposed to two terms of support by comparing take-up and use of early education for those born *just* either side of the December 31st cohort boundary, but not too far away, after the policy change on 1ˢᵗ September 2013. Children born either side of a cohort boundary will have different levels of eligibility 'as if' at random and as such these data can be treated as a natural experiment.

Table 1 makes explicit the comparisons that can be made on a similar basis for the most and moderately disadvantaged groups. Cross-referencing Table 1 with Figure 3 makes clear the cohort boundaries within the SEED data that capture variation in the level of eligibility for the most and moderately disadvantaged groups across the sample.

Although Table 1 refers to cohorts, comparisons are not made across entire cohorts. The analyses aimed to compare children born around the cohort boundary points, who are likely to have taken-up early education at similar rates in the absence of the policy change. Thus the effect of differences in eligibility on take-up of early education after the policy change came into force, either side of a boundary point, should result in an estimate of the effect of the change in official eligibility for these groups that is free from bias.

Table 1: Contrasts in official eligibility for early years education across cohorts within the SEED sample

Income group	Cohort Boundary	Comparison	Contrast in Eligibility
Most disadvantaged	December 31 2010	C1 v C2	One versus two terms
Most disadvantaged	March 31 2011	C2 v C3	Two versus three terms
Moderately disadvantaged	August 31 2011	C3 v C4	None versus one term
Moderately disadvantaged	December 31 2011	C4 v C5	One versus two terms
Moderately disadvantaged	March 31 2012	C5 v C6	Two versus three terms

The sample

The sample for analysis constitutes families for whom data was available at Wave 1 (two-year-olds) and Wave 2 (three-year-olds). The discontinuity analysis was performed on samples from both the most and moderately disadvantaged groups. Table 2 gives the size of relevant cohort samples disaggregated by the disadvantage groups.

Table 2: Breakdown of sample by cohort and disadvantage group.

Cohort	Term of birth	Disadvantage group			Total
		Most	**Moderate**	**Not disad.**	
1	Sep-Dec 2010	170	304	329	803
		21.17%	37.86%	40.97%	100.00%
2	Jan-Mar 2011	186	297	331	814
		22.85%	36.49%	40.66%	100.00%
3	Apr-Aug 2011	179	307	298	784
		22.83%	39.16%	38.01%	100.00%
4	Sep-Dec 2011	247	277	291	815
		30.31%	33.99%	35.71%	100.00%
5	Jan-Mar 2012	220	241	266	727
		30.26%	33.15%	36.59%	100.00%
6	Apr-Aug 2012	208	199	233	640
		32.50%	31.09%	36.41%	100.00%
Total		1210	1625	1748	4583
		26.40%	35.46%	38.14%	100.00%

In total 4,583 cases provide data at Waves 1 and 2; with 1,210 children from most disadvantaged families and 1,625 from moderately disadvantaged families. The circumstances of families were reassessed at termly intervals up to the point the cohort turned three, roughly between a child's 2nd and 3rd birthdays. Although the official policy rules implied that once a child was deemed most disadvantaged it continued to be eligible for funded early education regardless of whether its family's income subsequently rose, those families within a cohort whose income fell such that they qualified for out-of-work, means-tested benefits, could subsequently qualify, even after initially being assessed as ineligible. Families that subsequently became eligible due to a deterioration in their income after their child reached two years but before they turned three are not included in the analysis presented here. Thus the sample constructed for the discontinuity analysis was that where children were assessed around their second birthday, found to be a most disadvantaged child, and therefore eligible from first September 2013 for one term or more's worth of support. It does not include children who subsequently became eligible.

The discontinuity analysis requires that a 'bandwidth' was defined around each cohort boundary. As explained above, analyses do not compare take-up and use of early education across whole cohorts but among those born immediately either side of a cohort boundary. As such the bandwidth within which comparisons will be made around the boundary needs to be determined. On the basis of a battery of statistical tests undertaken to explore the assumption that official eligibility for support was assigned 'as if' at random around the cohort boundary points, the analysis that follows selects cases born one month either side of each cohort boundary point – this was therefore the bandwidth identified for analysis. Thus the tests conducted around the boundary point 31st December 2010 involve comparing take-up rates for those born during December 2010 and eligible for one term's support, with those born during the month of January

2011 and eligible for two term's support. Likewise around the boundary cut point of 31st March 2011, we compare take-up among those born in March 2011 with those born in April 2011, and so on. Table 3 below sets out the sample sizes for each group and how these relate to the cohorts and variations in official eligibility.

Table 3: Size of monthly comparison groups for discontinuity analysis

Cohort	Month of birth	Income group		Contrast in official Eligibility (number of terms)
		Most disadvantaged	**Moderately disadvantaged**	
1	Dec 2010	40		One v Two
2	Jan 2011	71		
2	Mar 2011	51		Two v three
3	Apr 2011	37		
3	Aug 2011		59	Zero v one
4	Sep 2011		69	
4	Dec 2011		66	One v two
5	Jan 2012		91	
5	Mar 2012		79	Two v three
6	Apr 2012		45	

Limiting the analysis to the sample defined by the bandwidth means that for any estimate obtained, the assumption of 'as if' random, should be a reasonable one and therefore any estimates should be free from bias and reliable[5]. However, this advantage comes at a price. Due to the restricted nature of the samples used in the analysis, the difference in take-up rates and use either side of the boundary was formally the average effect of intention to treat in the region of the cut point or boundary from the perspective of central government. This means the analyses cannot be certain that these estimate of effectiveness apply to cases further away from the boundary points. Therefore the results should be interpreted with caution and used only with qualification due to their limited generalizability. Secondly, the sample sizes used for comparison purposes are quite small. Thus in comparing the effect on use of early education for those in the most disadvantaged group eligible for one term of support relative to two terms, the total sample size for the analysis was 111 (n=40+71). This means that differences in take-up between those born either side of the December 2010 boundary would have to be quite large to be distinguishable from chance variations[6]. This weakness in the analysis should be kept in mind.

---

[5] In other words, in the absence of the policy change we expect take-up and usage around the boundary points to be equal in expectations.
[6] We estimate an effect size of 0.55 for a test of mean difference at the December 31st 2010 threshold with a total sample size of n=111. This is quite large, suggesting that the risk of a Type II statistical error is not insubstantial.

Statistical tests

The assumption that the assignment of official eligibility for early education was 'as if' random in the region of the cohort boundary implies that a simple comparison of take-up or use either side of the boundary provides an unbiased estimate of effects in the region of the boundary point. Furthermore, the analysis involves nothing more complex than comparison of means and percentages and the reporting of p-values (Dunning, 2012, p. 122). The results reported below present mean differences. However, a separate test of statistical significance is presented that is not related directly to these differences. Due to non-normality in the data, the Wilcoxon Mann Whitney rank sum test is reported. Put simply, this test examines whether the distributions of weekly usage of early education *differs* in groups either side of the various cohort boundaries[7]. P-values at or below 0.05 imply differences in the two distributions at the 95 per cent level. Where outcomes are dichotomous, effects are computed as take-up percentages and p-values calculated using Fisher's exact test.

Testing assumptions

The main assumption upon which the identification of causal effects using a discontinuity approach rests is that assignment to different levels of eligibility is 'as if' random. In this section the statistical tests testing this assumption are described, as well as what might be done if it appears the assumption of 'as-if' randomised is implausible. These test results are presented later.

Following Dunning (2012) **balance-tests** will explore the degree of similarity on the basis of measures collected at baseline, between those born either side of the boundaries at 31st December 2010 and 31st March 2011, and so on, implementing different bandwidths. If the assumption of 'as-if' randomised is plausible, groups either side of these boundary points should look similar to each other. Measures collected at baseline should be carefully considered in terms of their suitability for such tests given that a number of baseline interviews were conducted late, after the policy changes were made for the most disadvantaged groups.

Placebo-tests test for the presence of discontinuities in the take-up of early education, at different points where no policy change occurred, that is at points where there is no cohort boundary and therefore no associated policy contrast. So for example, no large effect on take-up and use of early education should expected at a threshold of 31st October 2010, because children both either side of this date are both eligible for one term's worth of support. If this is the case, then results support the assumption that eligibility for one versus two terms worth of support, or two versus three terms, is assigned 'as if' random.

---

[7] Technically, the test is whether the samples either side of the boundary points are drawn from the same underlying population.

If the assumption of 'as-if' random does not appear to hold, then simple comparison of means across cohort boundaries capturing policy changes will produce biased estimates. In this case, a 'model-based' approach in order to estimate discontinuity effects may be required; for example, the use of local-linear regression models. This approach will be necessary where data points close to the cohort boundaries are relatively sparse and bandwidths therefore need to be wide in order to obtain reasonable precision, or where there is a non-zero derivative in the potential outcomes at the cohort boundary points (Dunning, 2012, p. 159). In this case identification of an impact would rely on the relationship between age and take-up of funded early education between two and three years as being relatively smooth up to and beyond each threshold in the absence of the policy change. The average effect of treatment will be defined at limit of the cohort boundary, rather than in than in the region of the cohort boundary (Dunning, 2012, pp. 128-133).

Results

The results of the discontinuity analyses are given in Table 4 (formal ECEC usage) and Table 5 (total ECEC usage).

Table 4: Discontinuity analysis; formal ECEC usage.

Disadvantage group	Contrast	Difference in mean formal ECEC usage	p[1]	Change in % using formal ECEC	p[2]
Most disadvantaged	1 to 2	+0.823	0.616	-2.7%	0.767
	2 to 3	+2.813	0.097	+15.4%	0.084
Moderately disadvantaged	0 to 1 term	+0.897	0.573	+5.8%	0.300
	1 to 2	-1.721	0.159	-3.0%	0.612
	2 to 3	+4.655	0.006	+12.0%	0.083

p[1] is the p-value from a non-parametric Wilcoxon rank-sum test for differences in means.
p[2] is the p-value from Fisher's exact test for difference in proportions.

Table 5: Discontinuity analysis; total ECEC usage.

Disadvantage group	Contrast	Difference in mean total ECEC usage	p[1]	Change in % using ECEC	p[2]
Most disadvantaged	1 to 2 terms	+1.383	0.722	-3.8%	0.743
	2 to 3 terms	+2.208	0.277	+9.5%	0.228
Moderately disadvantaged	0 to 1 term	+0.155	0.711	+0.5%	1.000
	1 to 2 terms	-0.308	0.676	+3.2%	0.494
	2 to 3 terms	+2.525	0.076	+3.2%	0.710

p[1] is the p-value from a non-parametric Wilcoxon rank-sum test for differences in means.
p[2] is the p-value from Fisher's exact test for difference in proportions.

There is evidence that the amount of formal ECEC used increases in the moderately disadvantaged group between those eligible for 3 rather than 2 terms of funded ECEC (Table 4), $p < 0.01$.

Placebo tests

Tables 6 to 7 report results from Placebo tests, which involve comparing take-up and use of early education around cohort boundaries where no or very small differences in take-up would be expected, as there is no variability in official eligibility. Put simply if the assumption enabling comparison of take-up either side of cohort boundaries to get an unbiased estimate where there are variations in eligibility is a fair one, the effects reported in Tables 6 to 7 should be small and ideally not reach statistical significance, i.e. there should be high p-values reported.

As can be seen in the tables below, this is broadly the case, suggesting that the assumption that official eligibility is assigned 'as if' random is plausible. There are two instances in which effects reach statistical significance in Table 6 and one in Table 7. This gives pause for thought. However, due to the number of tests performed and due to a lack of clear pattern across these analyses it can be tentatively concluded that these results are broadly supportive of the main identifying assumption underpinning the discontinuity analysis.

Table 6: Placebo tests; formal ECEC use.

Disadvantage group	Cohort boundary	Contrast	Difference in mean formal ECEC usage	p[1]	Change in % using formal ECEC	p[2]
Moderately disadvantaged	Dec 31 2010	n/a	+0.009	0.924	-0.3%	1.000
	Mar 31 2011	n/a	+1.951	0.129	+13.1%	0.031
Most disadvantaged	Aug 31 2011	n/a	+0.898	0.560	+12.2%	0.027
	Dec 31 2011	n/a	-0.251	0.972	-2.8%	0.754
	Mar 31 2012	n/a	+1.492	0.457	+2.4%	1.000

p[1] is the p-value from a non-parametric Wilcoxon rank-sum test for differences in means.
p[2] is the p-value from Fisher's exact test for difference in proportions.

Table 7: Placebo tests; total ECEC use.

Disadvantage group	Cohort boundary	Contrast	Difference in mean total ECEC usage	p[1]	Change in % using any ECEC	p[2]
Moderately disadvantaged	Dec 31 2010	n/a	-1.719	0.202	+1.3%	0.759
	Mar 31 2011	n/a	+0.826	0.836	+8.3%	0.096
Most disadvantaged	Aug 31 2011	n/a	+0.160	0.898	+12.2%	0.027
	Dec 31 2011	n/a	-1.031	0.858	-0.1%	1.000
	Mar 31 2012	n/a	+2.889	0.157	-0.1%	1.000

p[1] is the p-value from a non-parametric Wilcoxon rank-sum test for differences in means.
p[2] is the p-value from Fisher's exact test for difference in proportions.

Balance tests

A range of balance tests were performed on the data (see Table 8). These compared the groups either side of the cohort boundaries, within the bandwidths, to see if we can find systematic differences between them. If we can, then this suggests that our assumption that official eligibility is 'as if' randomly assigned lacks plausibility.

The following variables were examined.
- Child's sex.
- Child's ethnicity.
- Child's birthweight.
- Child's birth order.
- Number of sibs living in household.
- Maternal age at birth of child.
- Whether the child comes from a couple or lone parent household.
- Whether anyone is working in the household.

- Total household income.
- Index of Multiple Deprivation (IMD).
- Type of accommodation tenure.
- Mother's highest qualification.
- Highest parental Socio-Economic Status (SES).

In total 13 tests were conducted at each boundary point to examine whether differences in means or proportions of variables reached statistical significance. For continuous variables t-tests were performed, and for categorical variables chi-squared tests (see Table 8). For convenience, the p-values are in ascending order. Given the large number of tests, these p-values give no reason to reject the as-if random assumption.

Table 8: Summary of balance test results.

Disadvantage group	Cohort transition	p-values from balance tests (sorted)
Most disadvantaged	C1-C2	0.048 0.075 0.103 0.134 0.147 0.162 0.242 0.375 0.621 0.867 0.959 0.980 1.000
	C2-C3	0.119 0.194 0.242 0.300 0.318 0.336 0.406 0.443 0.545 0.560 0.584 0.653 0.874
Moderately disadvantaged	C3-C4	0.130 0.227 0.254 0.334 0.418 0.446 0.462 0.630 0.646 0.666 0.756 0.848 1.000
	C4-C5	0.021 0.322 0.347 0.371 0.400 0.452 0.509 0.519 0.550 0.658 0.712 0.749 0.875
	C5-C6	0.132 0.134 0.142 0.142 0.186 0.212 0.246 0.469 0.474 0.497 0.587 0.673 0.848

Regression analysis

Introduction

A second strategy for estimating the effects of eligibility for the prescribed official policy on take-up of early education takes the form of estimating a series of multiple regression models. Data from all cohorts are pooled in these analyses. Four measures of take-up of early education, outlined in the main text of this report, form the dependent variables. In order to evaluate the effects of different levels of eligibility for support from the perspective of the policy change, a hierarchical linear model of the following form is estimated where dependent variables modelled are continuous. Where dependent variables are dichotomous, logistic regression models are estimated that have a similar structure. These models are sometimes referred to as multi-level or mixed models.

The equations below represent the simplest form of model for ease of exposition:

$$Y_{ijkt} = \beta_{0jkt} + \beta_1 ONE_{ijkt} + \beta_2 TWO_{ijkt} + \beta_3 THREE_{ijkt} + \varepsilon_{ijkt} \ldots\ldots\ldots\ldots[1]$$

$$\beta_{0jkt} = \gamma_{00kt} + U_{0jkt} \ldots\ldots\ldots\ldots[2]$$

$$\gamma_{00kt} = \gamma_{000t} + V_{00kt} \ldots\ldots\ldots\ldots[3]$$

$$\gamma_{000t} = \gamma_{0000} + R_{000t} \ldots\ldots\ldots\ldots[4]$$

Equation [1] represents the model at the child level, or level 1. Three independent variables are included in the model – though in practice models such models will include a range of control variables that are discussed below. The independent variables capture eligibility according to the official policy for one, two or three terms of funded early education, with no eligibility the omitted category against which the effects are contrasted. Hence parameters β_1 β_2 and β_3 represent the effects of being eligible for one, two and three terms worth of support relative to zero eligibility. These effects are the average effect of intention to treat from the perspective of central government policy.

The random components of the model are in equations [2], [3] and [4]. The subscript *'i'*, *'j'*, *'k'* and *'t'* index different levels, namely individual pupils, primary sampling units, strata and region. The model is a random intercepts model. Slope coefficients are fixed. For example, equation [2] reveals that we allow the mean score at PSU level to vary around the mean at strata level, and so on. This model structure is required to reflect the complexity of the sample design and to ensure correct statistical inferences are made.

The final model for estimation in its simplest form is:

$$Y_{ijkt} = \gamma_{0000} + \beta_1 ONE_{ijkt} + \beta_2 TWO_{ijkt} + \beta_3 THREE_{ijkt} + R_{000t} + V_{00kt} + U_{0jkt} + \varepsilon_{ijkt}$$

All multi-level models were fitted in the software package R. In the multivariate models, the following control variables were included:

- Child's sex.
- Child's ethnicity.
- Child's birthweight.
- Child's birth order.
- Number of sibs living in household.
- Maternal age at birth of child.
- Whether the child comes from a couple or lone parent household.
- Whether anyone is working in the household.
- Total household income.
- Index of Multiple Deprivation (IMD).
- Disadvantage group.
- Type of accommodation tenure.
- Mother's highest qualification.
- Highest parental Socio-Economic Status (SES).

We do not report the coefficients on these control variables in the main text, as the chief concern is the evaluation of variations in eligibility.

Results

The results of the regression models are shown in Tables 9 to 12.

Models are considered for four outcome variables:

1. Formal ECEC use as a continuous variable (Table 9).
2. Whether there was any formal ECEC used (Table 10).
3. Total ECEC use as a continuous variable (Table 11).
4. Whether there was any ECEC used (Table 12).

For each outcome there are three models:

1. Univariate model of outcome by terms of eligibility for funded childcare.
2. Multivariate model of outcome by terms of eligibility for funded childcare, controlling for demographic covariates.
3. Multivariate model of outcome by terms of eligibility for funded childcare, controlling for demographic covariates and trend in childcare usage over time in each disadvantage group.

Table 9: Regression models; formal ECEC usage.

	Model 1 - univariate		Model 2 - controlled		Model 3 - controlled with linear trends	
	Coef.	p	Coef.	p	Coef.	p
Constant	+12.69	<0.001	+6.97	<0.001	+7.56	<0.001
1 term	-1.39	0.007	+0.52	0.366	-0.68	0.487
2 terms	-2.24	<0.001	+0.22	0.711	-1.50	0.207
3 terms	-0.63	0.085	+3.03	<0.001	+0.48	0.734
Linear trends						
Most disadvantaged					+0.39	0.097
Moderately disadvantaged					+0.61	0.078
Not disadvantaged					+0.56	<0.001

Sample Size = 4,583

Table 10: Regression models; any formal ECEC use.

	Model 1 - univariate		Model 2 - controlled		Model 3 - controlled with linear trends	
	OR	p	OR	p	OR	p

Constant	11.96	<0.001	6.53	0.004	5.79	0.020
1 term	0.93	0.683	1.42	0.122	1.61	0.203
2 terms	0.42	<0.001	0.71	0.089	0.84	0.690
3 terms	1.11	0.472	2.17	<0.001	2.75	0.062
Linear trends						
Most disadvantaged					0.97	0.769
Moderately disadvantaged					0.94	0.645
Not disadvantaged					1.06	0.357

Sample Size = 4,583

Table 11: Regression models; total ECEC usage.

	Model 1 - univariate		Model 2 - controlled		Model 3 - controlled with linear trends	
	Coef.	p	Coef.	p	Coef.	p
Constant	+17.65	<0.001	+10.49	<0.001	+9.61	<0.001
1 term	-2.78	<0.001	+0.93	0.170	+0.45	0.697
2 terms	-2.81	<0.001	+1.69	0.015	+0.76	0.583
3 terms	-3.19	<0.001	+3.64	<0.001	+1.48	0.371
Linear trends						
Most disadvantaged					+0.65	0.018
Moderately disadvantaged					+0.38	0.343
Not disadvantaged					+0.67	<0.001

Sample Size = 4,583

Table 12: Regression models; any ECEC use.

	Model 1 - univariate		Model 2 - controlled		Model 3 - controlled with linear trends	
	OR	p	OR	p	OR	p
Constant	20.77	<0.001	17.70	<0.001	18.29	0.002
1 term	0.81	0.358	1.60	0.094	1.52	0.378
2 terms	0.40	<0.001	0.87	0.582	0.81	0.685
3 terms	0.73	0.048	2.03	0.012	1.82	0.354
Linear trends						
Most disadvantaged					1.01	0.901
Moderately disadvantaged					1.03	0.866
Not disadvantaged					1.07	0.427

Sample Size = 4,583

The results are similar across the models of all four outcomes.

The univariate models show a reduction in childcare usage for those eligible for funded

childcare. As this does not control for any of numerous demographic confounders this result is of limited value.

The multivariate models show an increase in ECEC use for eligible children, especially for those eligible for 3 terms of funded childcare. These models do take account of demographic covariates, however they fail to take account of the general upward trend in ECEC usage over time, which is confounded with the number of terms of eligibility.

The final models include time-trends in ECEC usage for each disadvantage group. These models fail to show significant effects of terms of eligibility on ECEC use. We conclude that the final regression models give no clear evidence for the impact of the policy on the take up of ECEC.

Difference in differences approach

Introduction

The previous two approaches to estimating the effects of the official policy change on take-up of early education looked at results for both the most and moderately disadvantage groups. Looking at the moderately disadvantage group we can see that comparing take-up and use of early education across cohorts C3 and C6 (see Figure 3) should in principle enable us to investigate the effects of being eligible a full three terms of support (Cohort C6) against being eligible for no support (Cohort C3).

The comparison of take-up and use in these cohorts is made difficult by support made available by local authorities and how this differs, often substantially, from official policy. As such any comparison of these two cohorts would only be able to provide an estimate of the average intention to treat in line with the official policy. These departures from official policy may mask any underlying effect, making it difficult to find if it is present.

A second problem with simply comparing take-up and use among moderately disadvantaged groups in cohorts C3 and C6 is that other factors possibly drive take-up; for example, longer-run and pre-existing trends. Furthermore, other policy changes or events occurring over the period of time under consideration may also determine use of early education. However, if we assume that these events and trends are common to both the moderately disadvantage group and the most disadvantaged group, then we may use the latter as a control group. Thus we can use the most disadvantaged groups in cohorts C3 and C6 as a control group for the moderately disadvantaged group.

To make this clearer, consider the following:

$$\left(Y_{C6,M} - Y_{C3,M}\right) - \left(Y_{C6,L} - Y_{C3,L}\right)$$

'Y' represents weekly early education usage between a child's 2nd birthday and the Wave 2 survey (which in most cases took place in the term following the child's 3rd birthday). The subscript C6 or C3 refers to the cohort, and the letters 'M' and 'L' medium and most disadvantaged groups respectively. The expression in the first brace represents the difference between weekly early education use for the moderately disadvantage groups in cohort C6 with that used in cohort C3. Not that under the former outcome children qualified for three terms of support, whereas under the latter no support.

The second brace represents change in use of early education for the same period for the most disadvantaged group. At both time points this group qualified for three terms of support. Taking the difference in outcomes over the period concerned removes the common effect of three terms of support *assuming this effect is relatively fixed over time*, with any remaining difference capturing underlying trends in early education use. Assuming residual change over time is common to both disadvantaged groups in the absence of policy change, then subtracting the first difference from the second yields an impact of the policy change controlling for background trends.

The analysis above takes the form of estimating the following linear regression model:

$$Y_{it} = \beta_0 + \beta_1 Q_i + \beta_2 I_i + \beta_3 Q_i I_i + \varepsilon_{it}$$

Here Q_i is coded '1' if the child is in the moderately disadvantaged group, '0' if in the most disadvantaged group. The variable I_i is coded '1' if the child is observed in a period of time after moderately disadvantaged groups would qualify for three months of support, zero otherwise. The effect of the policy is captured by the coefficient β_3, i.e. the average effect of intention to treat under central government policy. In our analysis we include extra covariates in the regression model to take account of confounders.

Such an approach relies on three further assumptions:

1. That cohort characteristics are relatively fixed and that individuals do not attempt to manipulate their income to make themselves eligible for support
2. That confounding differences between eligible and ineligible groups are observed and we can control for them, and/or they are unobserved but fixed over time.
3. That the effects of being eligible for three terms of support may differ for income groups but that these differences are stable or time invariant.

One further complication of this analysis is movement between disadvantage groups. If we consider cohort C3 it is possible that both higher income children and moderately disadvantaged children move into the most disadvantaged group and become eligible through doing so. This would effectively contaminate our results. In the results reported in the main body of this report such movement is ignored and higher income groups are completely excluded from the analysis. Initial exploration of the scale of this problem suggests that it is likely to be small and therefore the effect on results trivial.

The sample for difference-in-differences analysis

The numbers in the sample for this analysis are shown in the Table 13 below.

Table 13: Sample for difference in differences analysis

	Most disadvantaged	Moderately disadvantaged	Total
Cohort 3	179	307	486
Cohort 6	208	199	407
Total	387	506	893

The difference in differences analysis occurred in two ways. First a simple comparison of mean use per week, both for total and formal early education, before and after the policy change for control and treatment groups was carried out. Second, four separate multiple regressions were estimated, two with use of total early education per week as dependent variable and two with use of formal early education per week as dependent variable. The first of each pair of models contained no control variables, while the second contained a range of control variables (results from the more elaborate models are not reported as results were unaffected by inclusion of control variables). Regression models provide standard errors and p-values for estimated effects. All results are in two tables, the first for usage of formal early education the second for total early education.

Results

The results of the difference-in-differences analyses are shown in Table 14 (formal ECEC usage) and Table 15 (total ECEC usage).

Table 14: Difference in differences analysis – formal ECEC usage.

	Pre-policy change - September 2013- August 2014	Post-policy change – September 2014- August 2015	Difference
Low-income (Control Group)	11.50	12.51	1.01
Middle-income (Treatment	12.75	14.46	1.71
Impact			0.70
p-value (obtained from multiple regression model with no covariates)			0.5572
Sample size = 893			

Table 15: Difference in differences analysis – total ECEC usage.

	Pre-policy change - September 2013- August 2014	Post-policy change – September 2014- August 2015	Difference
Low-income (Control Group)	13.43	14.54	1.11
Middle-income (Treatment	16.54	18.34	1.80

Impact			0.69
p-value (obtained from multiple regression model with no covariates)			0.6133
Sample size = 893			

Both models show a positive effect, suggestive of an effect of the policy change in increasing ECEC usage. However, in neither case does the effect achieve statistical significance.

Chapter 4: The relationship between early childhood education and care (ECEC) aged two to three and children's cognitive and socio-emotional outcomes at age three

Models of children's age three outcomes in terms of ECEC usage age two to three

Introduction

A breakdown of children's ECEC usage in the first three years of life by type and level of use is shown in Table 16.

Table 16: Breakdown of ECEC use by age group and type (3 categories).

Age range (years)	Mean use (hours per week)	Formal Group ECEC		Formal Individual ECEC		Informal Individual ECEC	
		N	%	N	%	N	%
0 to 1	≤ 2	4064	88.68	4349	94.89	3635	79.31
	>2-5	161	3.51	67	1.46	308	6.72
	>5-15	267	5.83	124	2.71	457	9.97
	>15-25	58	1.27	33	0.72	95	2.07
	>25-35	19	0.41	6	0.13	44	0.96
	>35	14	0.31	4	0.09	44	0.96
1 to 2	≤ 2	3475	75.82	4203	91.71	3160	68.95
	>2-5	131	2.86	31	0.68	135	2.95
	>5-15	317	6.92	88	1.92	524	11.43
	>15-25	255	5.56	104	2.27	329	7.18
	>25-35	178	3.88	71	1.55	239	5.21
	>35	227	4.95	86	1.88	196	4.28
2 to 3	≤ 2	1207	26.34	4181	91.23	3062	66.81
	>2-5	539	11.76	45	0.98	317	6.92
	>5-15	1789	39.04	160	3.49	723	15.78
	>15-25	666	14.53	121	2.64	325	7.09
	>25-35	233	5.08	41	0.89	88	1.92
	>35	149	3.25	35	0.76	68	1.48

ECEC use over 1 year is expressed as the mean weekly usage over the 38 weeks of the school terms. Sample size = 4583.

Our principal interest was in the effect of ECEC use aged two to three on children's outcomes at age three. Models in terms of ECEC usage aged one to three are given in Appendix C.

Initial models

Initial models were fitted in terms of the amount of formal group, formal individual and informal individual ECEC used aged two to three. Models controlled for home environment and demographic covariates. Because of the clustered sample design, mixed-effects models were used with random effects fitted for government region, for stratum within government region and for primary sampling unit within stratum. Models were fitted to multiply imputed data. Results of the initial models are shown in Table 17.

Table 17: Results of models of child outcomes in terms of ECEC use age two to three years.

Outcome	Type of ECEC					
	Formal Group		**Formal Individual**		**Informal Individual**	
	Coef.	p	Coef.	p	Coef.	p
Naming Vocabulary	+0.022	0.439	+0.106	<0.001 ***	+0.099	<0.001 ***
Picture Similarities	+0.014	0.630	+0.043	0.139	-0.031	0.315
SDQ Hyperactivity	+0.062	0.133	-0.002	0.968	-0.019	0.684
SDQ Emotional Symptoms	-0.118	0.022 *	-0.060	0.196	+0.017	0.733
SDQ Conduct Problems	+0.116	0.005 **	+0.084	0.049 *	-0.016	0.687
SDQ Peer Problems	-0.199	<0.001 ***	-0.023	0.641	+0.077	0.189
SDQ Prosocial	+0.109	0.010 **	+0.046	0.247	-0.020	0.694
Behavioural Self-regulation	+0.070	0.070	+0.102	0.024 *	-0.020	0.641
Emotional Self-regulation	-0.100	0.024 *	-0.016	0.678	+0.058	0.234
Co-operation	+0.051	0.247	+0.067	0.090	+0.024	0.624

Models control for home environment and demographic covariates. Standardized model coefficients are given along with p values: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Model coefficients represent the difference in the standardized outcome produced by a change of 2 standard deviations in the ECEC covariate, controlling for all other covariates. Models are fitted to multiply imputed data. Sample size = 4583.

This gives the detailed results which are summarized in simplified form in Chapter 4 in the Research Report.

Detail models

Where there were significant effects of ECEC use in the initial models, we proceeded to fit detail models in terms of specific levels of ECEC usage aged two to three. The usage bands analysed are detailed in Table 18.

The results of the detail models are given in Table 19, and summarized in Figures 4 to 10. Each figure shows the difference in the named outcome (e.g., Naming Vocabulary), in standard deviations, between five levels of ECEC use (hours per week) for formal group, formal individual and informal individual ECEC use, respectively, compared to a

baseline of children with two hours or less of that type of ECEC.[8] This is represented by the dotted horizontal line. Circles indicate the scores for each category based on hours per week and the vertical lines show 95% confidence intervals for the difference in scores. Statistically significant effects are shown in bold.

Table 18: Analysis of ECEC use aged two to three by level of usage band. ECEC use over a year is expressed as mean weekly usage over the 38 weeks of the school terms.

Type of ECEC	Usage level	N	%	Mean	SD
Formal group	≤ 2	1207	26.3	0.50	0.54
	>2-5	539	11.8	3.61	0.92
	>5-15	1789	39.0	10.01	2.75
	>15-25	666	14.5	19.12	2.83
	>25-35	233	5.1	29.34	2.70
	>35	149	3.3	42.27	7.23
Formal individual	≤ 2	4181	91.2	0.01	0.14
	>2-5	45	1.0	3.71	0.76
	>5-15	160	3.5	10.02	2.88
	>15-25	121	2.6	19.71	2.95
	>25-35	41	0.9	29.78	2.78
	>35	35	0.8	41.42	5.58
Informal individual	≤ 2	3062	66.8	0.13	0.40
	>2-5	317	6.9	3.42	0.89
	>5-15	723	15.8	9.19	2.81
	>15-25	325	7.1	19.52	2.72
	>25-35	88	1.9	29.36	2.79
	>35	68	1.5	45.31	8.74

Sample size = 4583.

---

[8] Standard Deviation is a standardised measure of the spread of data values. In this example the standardised units are used so the effects are comparable for variables measured on different scales. See Technical Report Appendix B for outcome variable summary statistics including standard deviations.

Table 19: Results of models of child outcomes in terms of levels of ECEC use between ages two and three. All effects are relative to a baseline group using ≤ 2 hours ECEC per week.

Outcome	Usage level (mean hours per week)	Type of ECEC					
		Formal Group		Formal Individual		Informal Individual	
		Coef.	p	Coef.	p	Coef.	p
Naming Vocabulary	>2-5	+0.088	0.060	+0.075	0.574	+0.069	0.199
	>5-15	+0.046	0.175	+0.134	0.063	+0.122	0.002 **
	>15-25	+0.066	0.136	+0.192	0.021 *	+0.158	0.003 **
	>25-35	+0.086	0.190	+0.081	0.560	+0.137	0.162
	>35	-0.010	0.903	+0.493	0.001 **	+0.169	0.134
SDQ Emotional Symptoms Scale	>2-5	-0.051	0.446	-0.068	0.703	-0.064	0.342
	>5-15	-0.093	0.113	-0.010	0.916	+0.002	0.973
	>15-25	-0.130	0.106	-0.105	1.000	+0.034	0.657
	>25-35	-0.164	0.094	-0.112	0.555	+0.157	0.251
	>35	-0.298	0.029 *	-0.276	0.267	-0.044	0.800
SDQ Conduct Problems Scale	>2-5	-0.047	0.421	-0.001	0.994	+0.005	0.938
	>5-15	-0.013	0.807	+0.185	0.052	-0.086	0.056
	>15-25	+0.058	0.281	-0.012	0.908	+0.010	0.881
	>25-35	+0.122	0.153	+0.281	0.175	+0.059	0.617
	>35	+0.347	<0.001 ***	+0.383	0.149	-0.016	0.920
SDQ Peer Problems Scale	>2-5	-0.072	0.204	+0.020	0.901	-0.018	0.782
	>5-15	-0.128	0.017 *	+0.041	0.684	-0.003	0.953
	>15-25	-0.266	<0.001 ***	-0.124	0.332	+0.054	0.518
	>25-35	-0.354	<0.001 ***	-0.165	0.417	+0.275	0.029 *
	>35	-0.390	<0.001 ***	+0.035	0.894	+0.224	0.294
SDQ Prosocial Scale	>2-5	+0.092	0.212	-0.073	0.660	+0.013	0.843
	>5-15	+0.054	0.297	-0.008	0.928	+0.051	0.257
	>15-25	+0.133	1.000	+0.188	0.060	-0.007	0.924
	>25-35	+0.234	0.010 **	+0.041	0.828	-0.215	0.101
	>35	+0.158	0.149	+0.084	0.682	-0.014	0.935
Behavioural Self-regulation Scale	>2-5	+0.091	0.120	-0.089	0.615	+0.035	0.572
	>5-15	+0.059	0.222	+0.023	0.782	+0.015	0.747
	>15-25	+0.103	1.000	+0.281	0.012 *	+0.026	0.759
	>25-35	+0.142	0.109	+0.134	0.512	-0.194	0.105
	>35	+0.106	0.267	+0.370	0.116	-0.056	0.755
Emotional Self-regulation Scale	>2-5	+0.022	0.749	-0.030	0.849	+0.042	0.524
	>5-15	-0.004	0.936	-0.152	0.109	+0.065	0.214
	>15-25	-0.026	0.701	+0.156	0.126	+0.070	1.000
	>25-35	-0.064	0.470	-0.111	0.553	+0.042	0.725
	>35	-0.366	0.001 **	-0.155	0.511	+0.169	0.306

Notes on Table 19: Level of ECEC use is mean hours per week over the 38 weeks of the school terms. Models controlled for home environment and demographic covariates. Standardized model coefficients are given along with p values: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Model coefficients represent the difference in the standardized outcome between children with a given level of ECEC use and the baseline (≤ 2 hours per week) group, controlling for all other covariates. Mixed-effects models were fitted to multiply imputed data. Sample size = 4,583.

Figure 4: Association of ECEC and Naming Vocabulary.
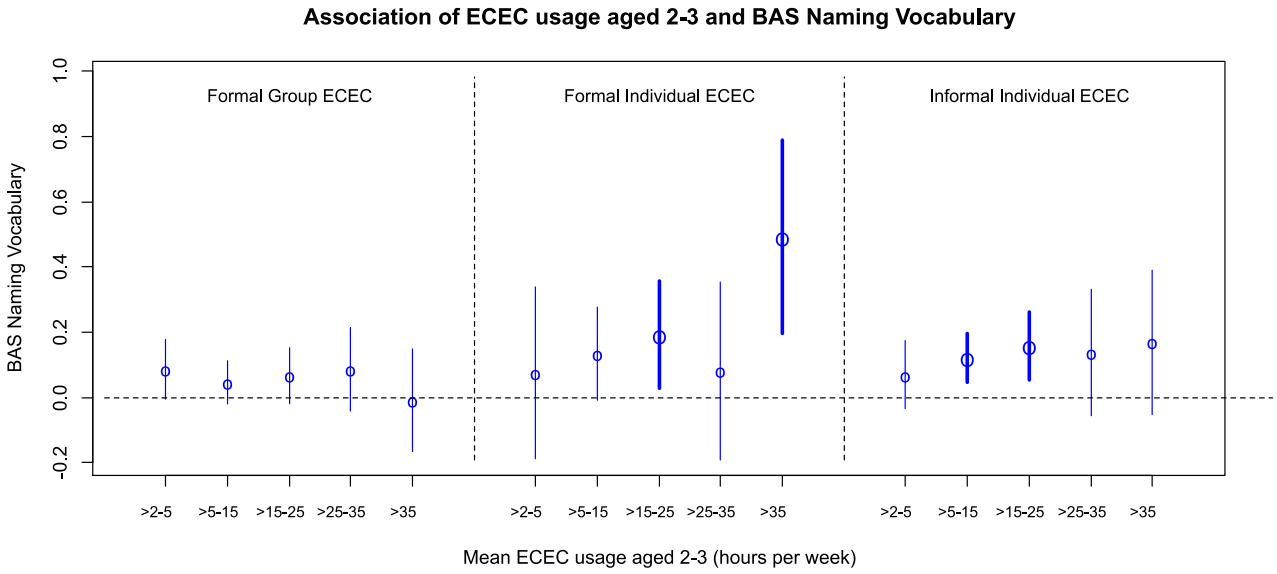
Sample size = 4,583
Statistically significant effects are shown in bold.

Figure 5: Association of ECEC and Emotional Symptoms.
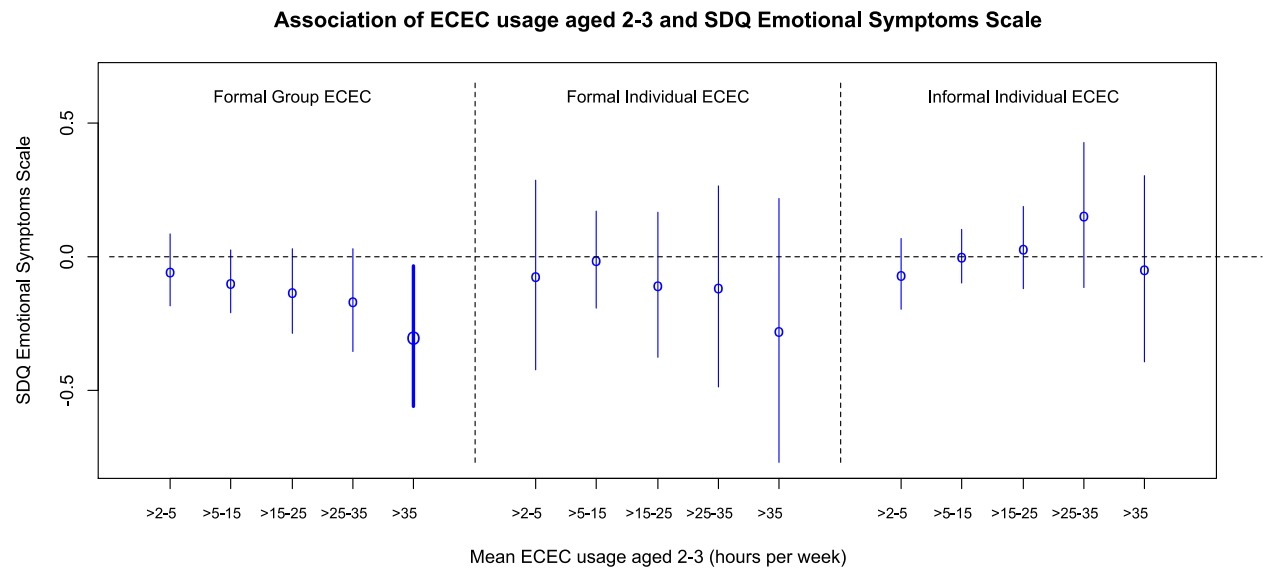
Sample size = 4,583
Statistically significant effects are shown in bold.

Figure 6: Association of ECEC and Conduct Problems.

Association of ECEC usage aged 2-3 and SDQ Conduct Problems Scale
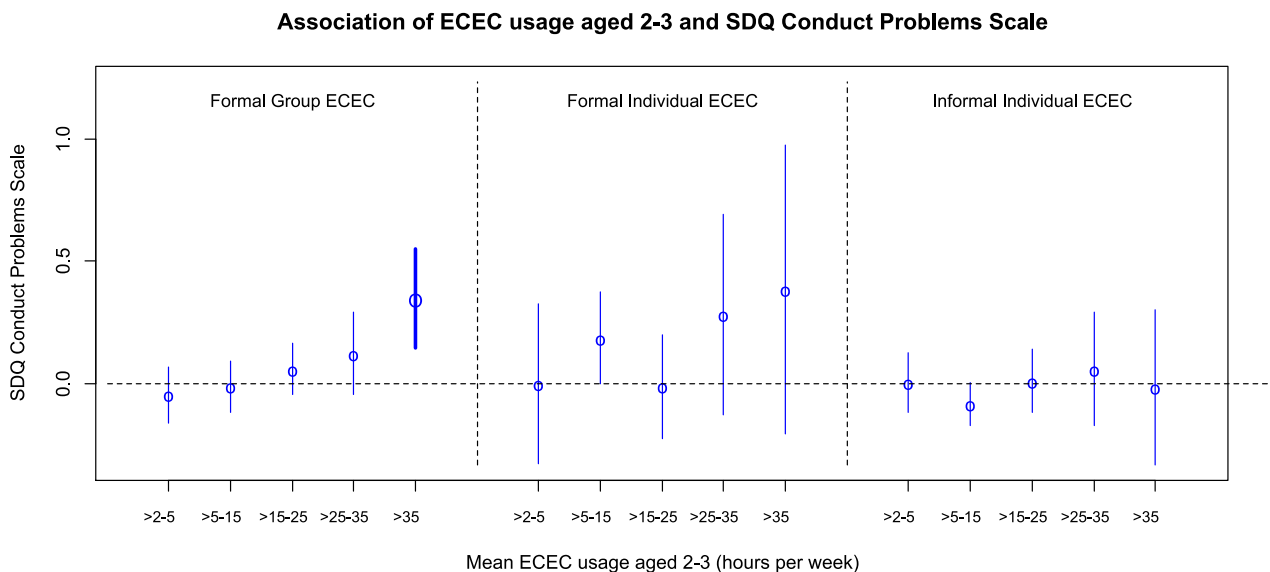
Sample size = 4,583
Statistically significant effects are shown in bold.

Figure 7: Association of ECEC and Peer Problems.

Association of ECEC usage aged 2-3 and SDQ Peer Problems Scale
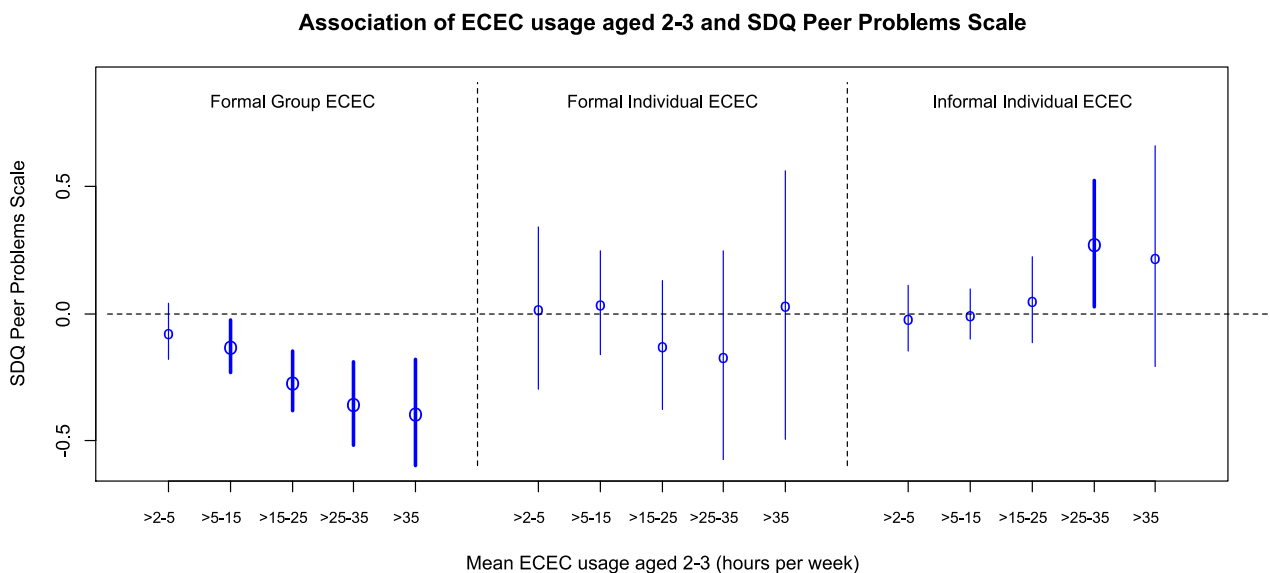
Sample size = 4,583
Statistically significant effects are shown in bold.

Figure 8: Association of ECEC and Prosocial Behaviour.

Association of ECEC usage aged 2-3 and SDQ Prosocial Scale
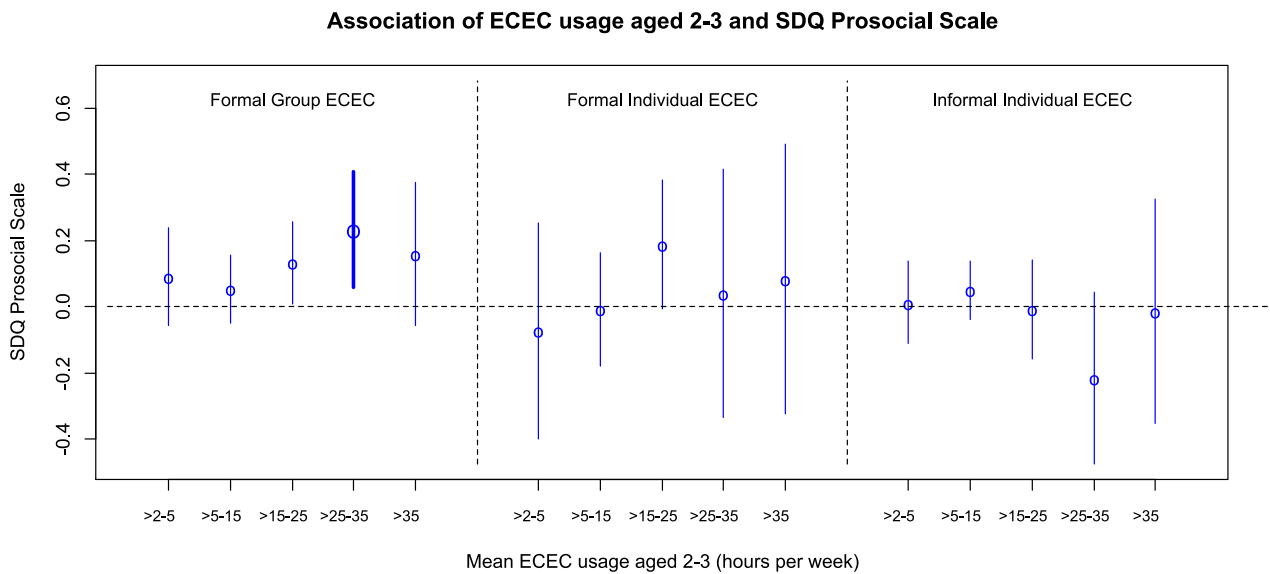
Sample size = 4,583
Statistically significant effects are shown in bold.

Figure 9: Association of ECEC and Behavioural Self-regulation.

Association of ECEC usage aged 2-3 and Behavioural Self-Regulation Scale
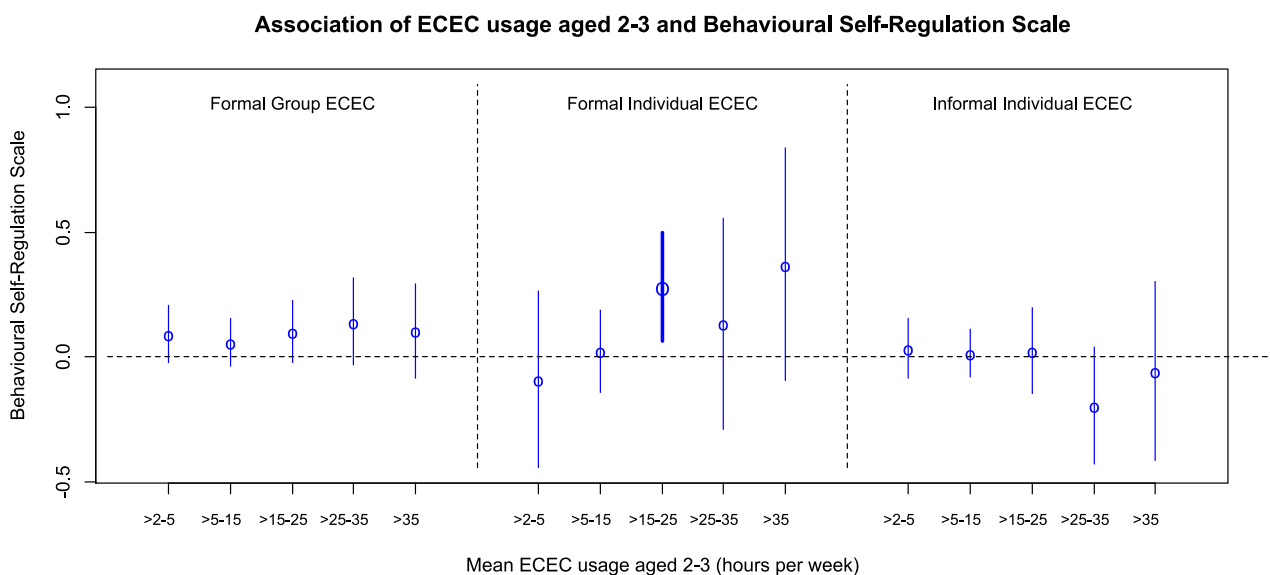
Sample size = 4,583
Statistically significant effects are shown in bold.

Figure 10: Association of ECEC and Emotional Self-regulation.

Association of ECEC usage aged 2-3 and Emotional Self-Regulation Scale
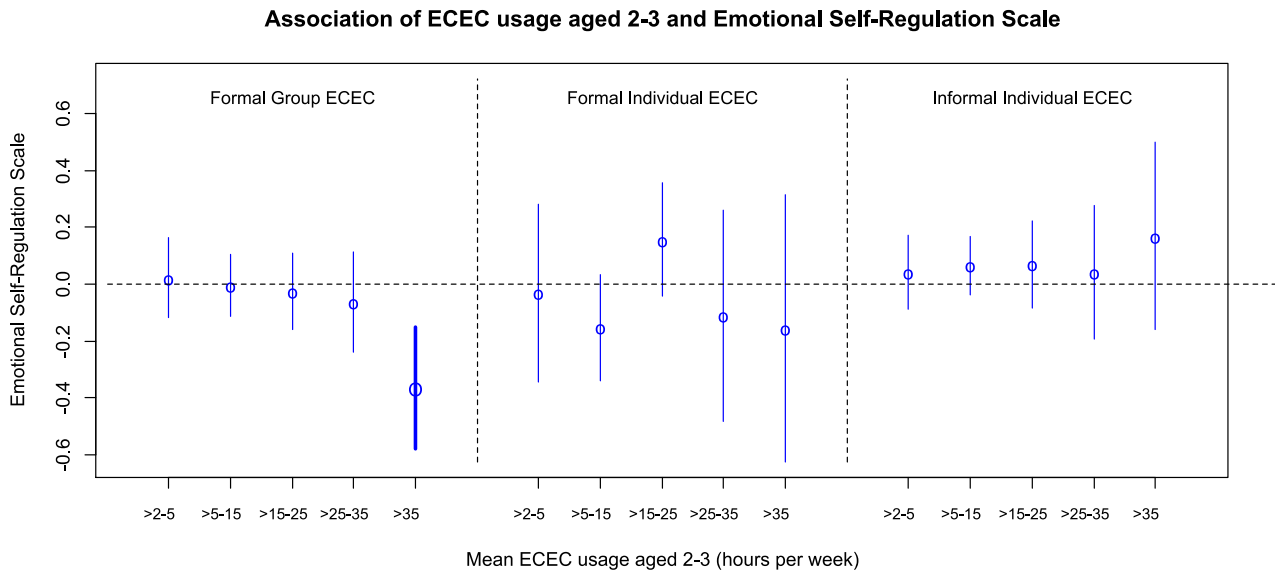
Sample size = 4,583
Statistically significant effects are shown in bold.

Multiple imputation

All the regression models were fitted to multiply imputed data. The imputation model included all outcome variables, home environment variables, demographic covariates and ECEC usage data. Missing data were imputed using the Amelia II package (Honaker 2010). The imputation model assumes a multivariate normal distribution for the complete data (missing and observed). Binary, categorical and ordinal variables are incorporated into this distribution using appropriate transformations. Ten imputations were generated, and models fitted to each imputed data set. Model results were consolidated using Rubin's Rules (Rubin 1987), with degrees of freedom found using Hesterberg (Hesterberg 1998).

Subject to the assumption that the missing data are missing at random (MAR), the results of the models fitted to the multiply imputed data are unbiased. The MAR assumption entails that the probability that an observation is missing is determined by the observed data rather than by unobserved factors. This assumption cannot be proved. However, we suggest that given the wide range of childcare, demographic and home environment variables included in the imputation models this assumption is reasonable.

The results of complete cases models are unbiased only if the missing data are missing completely at random (MCAR). The MCAR assumption entails that the probability that an observation is missing does not depend on any other variables (observed or unobserved). This is unlikely to be true. For example, the probability that a child has missing values for the variables derived from the ECEC provider assessment is likely to be higher for children who spend less time in ECEC.

In light of the above, we use the results from analysis of the multiply imputed data as our main analysis results presented in the Research Report. The results from complete cases models are given in Appendix D.

Investigating the high formal group ECEC use children

Examining the age at which formal group ECEC was first used

Children with above 35 hours per week of formal group ECEC over the 38 weeks of the school terms had poorer outcomes than the reference group for the outcomes Conduct Problems and Emotional Self-regulation; see Table 19, Figure 6 and Figure 10.

A comparison of this high formal group ECEC use group with other children showed that these children were more likely than other children to have started formal group ECEC early in life, see Table 20.

Table 20: Breakdown of sample by formal group ECEC usage aged two to three and age at which formal group ECEC usage started.

Age started formal group ECEC	Formal group ECEC usage aged 2 to 3			
	≤ 35 hours per week		> 35 hours per week	
	N	%	N	%
Age 0-1	526	11.9	86	57.7
Age 1-2	570	12.9	43	28.9
Age 2-3	2868	64.7	20	13.4
All	4434	100.0	149	100.0

We fitted models of Conduct Problems and Emotional Self-regulation, comparing children with greater than 35 hours of formal group ECEC per week with a control group with ≤ 2 hours per week of formal group ECEC, the greater than 35 hours per week group being broken down by the age at which formal group ECEC was first used; see Table 21.

Table 21: Results of models of child outcomes in terms of formal group ECEC usage aged two to three and age at which formal group ECEC use started.

Formal group ECEC usage	N	Outcome			
		SDQ Conduct Problems Scale		Emotional Self-regulation Scale	
		Coef.	p	Coef.	p
Control group: ≤ 2 hours per week	1207	0.000 (reference level)		0.000 (reference level)	
> 35 hours per week: first used aged 0-1	86	+0.325	0.013 *	-0.399	0.007 **
> 35 hours per week: first used aged 1-2	43	+0.311	0.084	-0.266	0.131
> 35 hours per week: first used aged 2-3	20	+0.512	0.065	-0.434	0.193

Models controlled for home environment and demographic covariates. Standardized model coefficients are given along with p values: * = p < 0.05, ** = p < 0.01, *** = p < 0.001. Model coefficients represent the difference in the standardized outcome between children and the baseline (≤ 2 hours per week) group, controlling for all other covariates. Models are fitted to multiply imputed data.

The deleterious effects of high formal group ECEC use on the outcomes were statistically significant only for children who had started formal group ECEC in the first year of life. However, the numbers of high formal group ECEC use children who had started formal group ECEC later were low. Therefore the failure to find significant effects for these groups should be interpreted with caution.

The demographics of the high formal group ECEC use children

We compared the 149 children with high formal group ECEC use aged two to three (greater than 35 hours per week) with all other children on demographic variables. The results are given in Table 22 (continuous variables) and in Tables 23 and 24 (categorical variables).

Table 22: Comparison of demographic factors between high formal group ECEC use children and other children; continuous variables.

Variable	All other children (N = 4434)		High formal ECECE use children (N = 149)		p-value from test for difference in means
	Mean	SD	Mean	SD	
Birth weight	3.34	0.628	3.39	0.661	0.540
Maternal age	29.2	5.98	30.8	5.50	<0.001 ***
Home learning environment	23.8	6.93	23.9	5.92	0.706
Household chaos	8.04	2.32	7.42	2.02	0.003 **
Kessler psychological distress	9.37	3.95	8.85	2.99	0.582
Limit Setting	2.62	0.723	2.64	0.644	0.564
Parent / child closeness	14.4	1.36	14.7	0.835	0.005 **
Parent / child conflict	13.4	4.81	12.7	4.60	0.081

p-values are from non-parametric Wilcoxon rank-sum tests for difference in means.

Table 23: Comparison of demographic factors between high formal group ECEC use children and other children; categorical variables (part 1).

Variable	Level	All other children (N = 4434)		High formal ECECE use children (N = 149)		p-value from test for difference in proportions
		N	%	N	%	
Region	The North	1165	26.3	42	28.2	0.669
	The Midlands	905	20.4	30	20.1	1.000
	East of England	495	11.2	16	10.7	0.976
	London	511	11.5	30	20.1	0.002 **
	The South	1358	30.6	31	20.8	0.013 *
Sex	Female	2116	47.7	70	47.0	0.924
	Male	2318	52.3	79	53.0	0.924
Ethnic group	White	3699	83.5	114	76.5	0.033 *
	Asian	296	6.7	7	4.7	0.430
	Black	175	3.9	17	11.4	<0.001 ***
	Mixed / other	261	5.9	11	7.4	0.561
Number of siblings	0	1350	30.4	67	45.0	<0.001 ***
	1	1866	42.1	66	44.3	0.650
	2+	1218	27.5	16	10.7	<0.001 ***
Birth order	1	1907	43.0	76	51.0	0.064
	2	1539	34.7	58	38.9	0.329
	3+	988	22.3	15	10.1	<0.001 ***
Parental status	Couple	3253	73.4	108	72.5	0.885
	Lone parent	1181	26.6	41	27.5	0.885
Work status	No one working	1062	24.0	12	8.1	<0.001 ***
	Someone working	3372	76.0	137	91.9	<0.001 ***
Family income	Less than £10,000	645	15.7	13	8.8	0.032 *
	£10,000 to £20,000	1008	24.5	16	10.9	<0.001 ***
	£20,000 to £40,000	1333	32.4	39	26.5	0.157
	Greater than £40,000	1124	27.3	79	53.7	<0.001 ***

Table 24: Comparison of demographic factors between high formal group ECEC use children and other children; categorical variables (part 2).

Variable	Level	All other children (N = 4434)		High formal ECECE use children (N = 149)		p-value from test for difference in proportions
		N	%	N	%	
Accommodation	Owner-occupier	1869	42.2	82	55.0	0.002 **
	Renting	2386	53.9	61	40.9	0.002 **
	Living rent free	173	3.9	6	4.0	1.000
Disadvantage group	20% most disadvantaged	1198	27.0	12	8.1	<0.001 ***
	20%-40% most disadvantaged	1562	35.2	63	42.3	0.092
	60% least disadvantaged	1674	37.8	74	49.7	0.004 **
IMD	1 = least deprived	812	18.3	31	20.8	0.506
	2	776	17.5	36	24.2	0.047 *
	3	866	19.5	32	21.5	0.629
	4	907	20.5	29	19.5	0.848
	5 = most deprived	1073	24.2	21	14.1	0.006 **
Social class	Not working	200	4.5	0	0.0	0.014 *
	Routine	346	7.8	2	1.3	0.006 **
	Semi-routine	707	15.9	7	4.7	<0.001 ***
	Small employer or own account workers	363	8.2	1	0.7	0.001 **
	Lower supervisory	318	7.2	5	3.4	0.104
	Intermediate occupations	594	13.4	30	20.1	0.025 *
	Lower managerial	1253	28.3	51	34.2	0.135
	Professional or managerial	652	14.7	53	35.6	<0.001 ***
Mother's qualifications	None or don't know	565	12.7	5	3.4	0.001 **
	GCSE D-G	269	6.1	1	0.7	0.010 *
	GCSE A-C	1110	25.0	20	13.4	0.002 **
	A levels or further education	1147	25.9	35	23.5	0.577
	Degree	865	19.5	57	38.3	<0.001 ***
	Higher degree	478	10.8	31	20.8	<0.001 ***

Comparing the families using high levels (greater than 35 hours per week) of formal group ECEC with the rest of the sample, we found several differences. High formal group ECEC use children tend to have older mothers and higher levels of parent / child closeness. They are more likely to come from families with lower levels of household chaos. They are also more likely than other children to come from the Black ethnic group. They tend to have fewer siblings than other children. They are more likely to come from working families and from higher income families. Their parents are more likely to be professionals and to be highly qualified.

We also compared the particular types of formal group ECEC used aged two to three for the high formal group ECEC use children and other children; see Table 25.

Table 25: Percentage breakdown of type of formal group ECEC used aged two to three for three groups of children.

Type of ECEC	Controls N = 4434	High use 2-3 N = 63	High use 2-3 and early start N = 86
Nursery school	27.6%	26.5%	28.7%
Nursery attached to primary / infant	10.1%	6.9%	0.7%
Day nursery	39.3%	58.1%	66.6%
Maintained nursery	0.6%	0.0%	0.0%
Pre-school or playgroup	22.2%	8.5%	4.1%
SEN day school, nursery or unit	0.2%	0.0%	0.0%

High use 2-3 = children with > 35 hours per week of formal group ECEC aged 2-3.
High use 2-3 and early start = children with > 35 hours per week of formal group ECEC aged 2-3 and started using formal group ECEC before age 1.
Controls = all other children.

There are also differences between the high use (greater than 35 hours per week) group and other children in the type of formal group ECEC used aged two to three, with higher use of day nurseries and lower use of nurseries attached to primary / infant schools and of pre-schools / playgroups. These differences are more pronounced for children with high formal group ECEC use and an early start in formal group ECEC.

Finally, we compared the outcomes for children with high formal group ECEC use aged two to three (greater than 35 hours per week) with those of all other children; see Table 26. These comparisons do not control for demographic or home environment covariates.

Table 26: Comparison of outcomes between high formal group ECEC use children and other children.

Outcome	All other children (N = 4434)		High formal ECECE use children (N = 149)		p-value from test
	Mean	SD	Mean	SD	
Naming Vocabulary	51.09	11.99	52.23	10.65	0.181
Picture Similarities	47.90	9.45	48.53	10.05	0.466
SDQ Hyperactivity Scale	3.00	2.65	3.23	2.75	0.438
SDQ Emotional Symptoms Scale	2.31	1.78	1.78	1.21	<0.001 ***
SDQ Conduct Problems Scale	1.26	1.86	1.85	2.22	0.005 **
SDQ Peer Problems Scale	1.66	1.88	1.00	1.46	<0.001 ***
SDQ Prosocial Scale	6.74	2.43	7.11	2.12	0.185
Behavioural Self-regulation Scale	7.20	2.06	7.59	1.99	0.071
Emotional Self-regulation Scale	7.61	2.21	6.74	2.60	<0.001 ***
Co-operation Scale	7.81	2.14	7.72	2.00	0.479

We see that whilst these children have poorer outcomes for Conduct Problems and Emotional Self-regulation, they also have significantly lower levels of Peer Problems and Emotional Symptoms than other children.

Binary models

The effect on the outcomes of exposure to all formal ECEC (formal group and formal individual ECEC) aged two to three was also modelled as a binary variable. Because of the small number of children who had no formal ECEC use aged two to three we used children with ≤ 2 hours per week of formal ECEC aged two to three as the control group rather than those with no formal ECEC use.

Table 27: Regression model of three year old outcomes in terms of binary formal ECEC use (aged two to three years)

Outcome	Coef.	p
Naming Vocabulary	+0.085	0.009 **
Picture Similarities	+0.018	0.611
SDQ Hyperactivity Scale	+0.023	0.680
SDQ Emotional Symptoms Scale	-0.116	0.050
SDQ Conduct Problems Scale	+0.028	0.533
SDQ Peer Problems Scale	-0.146	0.004 **
SDQ Prosocial Scale	+0.087	0.116
Behavioural Self-regulation Scale	+0.092	1.000
Emotional Self-regulation Scale	-0.016	0.753
Co-operation Scale	+0.082	0.076

Models control for informal individual ECEC use, home environment and demographic covariates. Standardized model coefficients are given along with p values: * = p < 0.05, ** = p < 0.01, *** = p < 0.001. Model coefficients represent the difference in the standardized outcome between children with >2 hours per week formal group ECEC aged 2 to 3 and children with ≤ 2 hours per week formal group ECEC aged 2 to 3, controlling for all other covariates. Models are fitted to multiply imputed data. Sample size = 4583.

Multivariate mixed-effects linear regression models were fitted in terms of the binary formal ECEC aged two to three usage variable (see Table 27). Models controlled for informal individual ECEC aged two to three and demographic and home environment variables.

Children with >2 hours per week formal ECEC had significantly higher Naming Vocabulary Scores and significantly lower Peer problems scores than the control group (≤ 2 hours per week formal ECEC).

Testing for interactions

Interaction with disadvantage group

We tested for interactions between disadvantage group and the effects of ECEC usage in the initial models of child outcomes.

Table 28: p-values from tests for interactions between ECEC usage and disadvantage group in the initial models.

Outcome	Type of ECEC		
	Formal group	Formal individual	Informal individual
BAS Naming Vocabulary	0.339	0.959	0.849
BAS Picture Similarities	0.345	0.976	0.806
SDQ Hyperactivity Scale	0.815	0.693	0.201
SDQ Emotional Symptoms Scale	0.698	0.732	0.631
SDQ Conduct Problems Scale	0.847	0.776	0.591
SDQ Peer Problems Scale	0.481	0.765	0.132
SDQ Prosocial Scale	0.633	0.734	0.382
Behavioural Self-regulation Scale	0.831	0.839	0.306
Emotional Self-regulation Scale	0.901	0.863	0.848
Co-operation Scale	0.646	0.877	0.484

Sample size = 4583.

There were no significant interactions.

Models stratified by disadvantage group

Because of the particular interest in differential effects by disadvantage group, we nevertheless proceeded to fit separate models by disadvantage group. However, it should be remembered that any difference in the effects of ECEC use on the outcomes between disadvantage groups are not statistically significant.

Table 29: Results of models of child three year old outcomes in terms of ECEC usage between ages two and three; models stratified by disadvantage group (part 1).

Outcome	Group	Type of ECEC					
		Formal Group		Formal Individual		Informal Individual	
		Coef.	p	Coef.	p	Coef.	p
BAS Naming Vocabulary	All children	+0.022	0.439	+0.106	<0.001 ***	+0.099	<0.001 ***
	Most disadvantaged	-0.027	0.692	+0.101	0.373	+0.071	0.328
	Moderately disadvantaged	+0.074	0.129	+0.140	<0.001 ***	+0.117	0.018 *
	Not disadvantaged	-0.008	0.851	+0.094	0.016 *	+0.095	0.019 *
BAS Picture Similarities	All children	+0.014	0.630	+0.043	0.139	-0.031	0.315
	Most disadvantaged	+0.105	0.149	+0.061	0.619	+0.047	0.569
	Moderately disadvantaged	+0.043	0.390	+0.059	0.170	-0.051	0.317
	Not disadvantaged	-0.050	0.303	+0.035	0.422	-0.041	0.367
SDQ Hyperactivity Scale	All children	+0.062	0.133	-0.002	0.968	-0.019	0.684
	Most disadvantaged	+0.042	0.640	-0.022	0.884	-0.100	0.374
	Moderately disadvantaged	+0.064	0.337	+0.001	0.986	-0.085	0.257
	Not disadvantaged	+0.070	0.194	-0.011	0.849	+0.045	0.370
SDQ Emotional Symptoms Scale	All children	-0.118	0.022 *	-0.060	0.196	+0.017	0.733
	Most disadvantaged	-0.086	0.361	+0.063	0.722	-0.029	1.000
	Moderately disadvantaged	-0.155	0.026 *	-0.077	1.000	-0.013	0.868
	Not disadvantaged	-0.097	0.145	-0.053	0.405	+0.057	0.393
SDQ Conduct Problems Scale	All children	+0.116	0.005 **	+0.084	0.049 *	-0.016	0.687
	Most disadvantaged	+0.080	0.380	+0.151	0.382	-0.063	0.579
	Moderately disadvantaged	+0.123	0.079	+0.095	0.129	-0.015	0.822
	Not disadvantaged	+0.127	0.014 *	+0.062	0.211	-0.001	0.992

Models control for home environment and demographic covariates. Standardized model coefficients are given along with p values: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Model coefficients represent the difference in the standardized outcome produced by a change of 2 standard deviations in the ECEC covariate, controlling for all other covariates. Models are fitted to multiply imputed data. Sample size = 4583.

Table 30: Results of models of child three year old outcomes in terms of ECEC usage between ages two and three; models stratified by disadvantage group (part 2).

| Outcome | Group | Type of ECEC | | | | | |
| | | Formal Group | | Formal Individual | | Informal Individual | |
		Coef.	p	Coef.	p	Coef.	p
SDQ Peer Problems Scale	All children	-0.199	<0.001 ***	-0.023	0.641	+0.077	0.189
	Most disadvantaged	-0.147	0.094	-0.033	0.841	-0.022	0.837
	Moderately disadvantaged	-0.262	<0.001 ***	-0.037	0.583	+0.036	0.706
	Not disadvantaged	-0.174	0.001 **	-0.017	0.766	+0.139	1.000
SDQ Prosocial Scale	All children	+0.109	0.010 **	+0.046	0.247	-0.020	0.694
	Most disadvantaged	+0.038	0.691	+0.074	1.000	+0.032	0.759
	Moderately disadvantaged	+0.125	0.060	+0.028	0.609	+0.023	0.758
	Not disadvantaged	+0.125	1.000	+0.070	0.188	-0.065	0.216
Behavioural Self-regulation Scale	All children	+0.070	0.070	+0.102	0.024 *	-0.020	0.641
	Most disadvantaged	+0.026	0.782	+0.160	0.372	+0.048	1.000
	Moderately disadvantaged	+0.099	0.089	+0.097	0.096	+0.029	0.659
	Not disadvantaged	+0.070	0.218	+0.113	0.063	-0.070	0.175
Emotional Self-regulation Scale	All children	-0.100	0.024 *	-0.016	0.678	+0.058	0.234
	Most disadvantaged	-0.064	0.519	-0.040	0.813	+0.107	0.337
	Moderately disadvantaged	-0.104	0.176	-0.012	0.852	+0.069	0.344
	Not disadvantaged	-0.111	0.039 *	-0.015	0.751	+0.043	0.409
Co-operation Scale	All children	+0.051	0.247	+0.067	0.090	+0.024	0.624
	Most disadvantaged	+0.011	1.000	+0.035	0.807	+0.085	0.420
	Moderately disadvantaged	+0.062	0.351	+0.056	0.332	+0.056	0.482
	Not disadvantaged	+0.055	0.380	+0.088	0.096	-0.010	0.849

Models control for home environment and demographic covariates. Standardized model coefficients are given along with p values: * = p < 0.05, ** = p < 0.01, *** = p < 0.001. Model coefficients represent the difference in the standardized outcome produced by a change of 2 standard deviations in the ECEC covariate, controlling for all other covariates. Models are fitted to multiply imputed data. Sample size = 4583.

Interactions with region and IMD

For the purpose of this analysis, government office regions were aggregated as shown in Table 31.

Table 31: Breakdown of sample by region.

Region	N	%	Government Office Regions
The North	1207	26.3	North East, North West, Yorkshire and the Humber
The Midlands	935	20.4	East Midlands, West Midlands
East of England	511	11.1	East of England
London	541	11.8	London
The South	1389	30.3	South East, South West

The results of tests for interactions between the effects of ECEC use and region are shown in Table 32.

Table 32: p-values from tests for interactions between ECEC usage and region in the initial models.

Outcome	Type of ECEC		
	Formal group	Formal individual	Informal individual
BAS Naming Vocabulary	0.934	0.314	0.066
BAS Picture Similarities	0.634	0.492	0.392
SDQ Hyperactivity Scale	0.423	0.896	0.905
SDQ Emotional Symptoms Scale	0.931	0.658	0.931
SDQ Conduct Problems Scale	0.352	0.809	0.867
SDQ Peer Problems Scale	0.696	0.840	0.910
SDQ Prosocial Scale	0.948	0.791	0.906
Behavioural Self-regulation Scale	0.994	0.997	0.939
Emotional Self-regulation Scale	0.610	0.996	0.739
Co-operation Scale	0.916	0.604	0.983

There were no significant interactions with region.

The results of tests for interactions between the effects of ECEC use and Index of Multiple Deprivation (IMD) are shown in Table 33.

Table 33: p-values from tests for interactions between ECEC usage and IMD in the initial models.

Outcome	Type of ECEC		
	Formal group	Formal individual	Informal individual
BAS Naming Vocabulary	0.901	0.527	0.624
BAS Picture Similarities	0.358	0.705	0.272
SDQ Hyperactivity Scale	0.813	0.782	0.723
SDQ Emotional Symptoms Scale	0.931	0.989	0.886
SDQ Conduct Problems Scale	0.850	1.000	0.630
SDQ Peer Problems Scale	1.000	0.983	0.844
SDQ Prosocial Scale	0.924	1.000	0.399
Behavioural Self-regulation Scale	0.844	0.946	0.857
Emotional Self-regulation Scale	0.943	0.887	0.811
Co-operation Scale	0.981	0.969	0.599

There were no significant interactions with IMD.

Chapter 5: The effects of home environment on child outcomes

Introduction

The models of children's age three cognitive and socio-emotional outcomes measured at Wave 2 control for home environment and demographic covariates measured at Wave 1 (age two).

The 6 home environment variables controlled for were:

1. Home Learning Environment (HLE) index (i.e. home activities that allow learning opportunities for the child; e.g., child read to, taken to library, painting/drawing, play with letters/numbers, songs/rhymes).
2. Household disorder (CHAOS scale including confusion, hubbub and disorder scale).
3. Parent's Psychological Distress (using the Kessler scale).
4. Limit Setting (i.e. how often parents use various measures to set limits when a child is naughty).
5. Parent/child Closeness (i.e. affectionate bond, child seeks comfort, child shares feelings).
6. Parent/child Conflict (i.e. parent-child struggles, child easily angry with parent).

The demographic variables controlled for were:

1. Birth weight.
2. Maternal age at birth of child.
3. Child gender ("Child is female").
4. Whether parenting is from a couple or a lone parent ("Lone parent").
5. Whether anyone is working in the household ("Working household").
6. Birth order of child (birth order = 1 / birth order = 3+).
7. Number of siblings (number of sibs = 0 / no. of sibs = 2+).
8. Household income (annual income below £10,000 / annual income above £40,000).
9. Deprivation (least deprived quintile of IMD / most deprived quintile of IMD).
10. SEED disadvantage group (60% least disadvantaged / 20% most disadvantaged).
11. Accommodation (renting / owner occupier).
12. Maternal education (no qualifications / higher degree).
13. Parental socio-economic class (not working / professional or managerial).

The effects of the home environment variables are given in Table 34. This gives the detailed results which are summarized in simplified form in Table 10, Chapter 5 in the Research Report.

The relative effect sizes of the ECEC, home environment and demographic factors are shown in Figures 11 to 20.

The results of the tests for interactions between the effects of ECEC use and Home Learning Environment are given in Table 35. The results of separate models of BAS Picture Similarities for children with high and low HLE scores are given in Table 36.

The effects of the home environment covariates on children's age three outcomes

Table 34: Results of models of child outcomes in terms of home environment variables.

Outcome	Covariate												
	Hine Learning Environment		Household CHAOS		Parent's Psychological Distress		Limit Setting		Parent / child closeness		Parent / child conflict	
	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p	Coef.	p
Naming Vocabulary	+0.295	<0.001 ***	+0.037	0.224	-0.066	0.022 *	+0.198	<0.001 ***	+0.107	<0.001 ***	-0.123	<0.001 ***
Picture Similarities	+0.226	<0.001 ***	-0.036	0.261	-0.028	0.377	+0.114	<0.001 ***	+0.023	0.429	-0.049	0.150
SDQ Hyperactivity Scale	-0.131	1.000	+0.064	0.172	+0.016	0.745	-0.082	0.066	-0.092	0.044 *	+0.109	0.045 *
SDQ Emotional Symptoms Scale	+0.043	0.311	+0.082	0.061	+0.093	0.072	-0.161	<0.001 ***	-0.076	0.139	-0.050	0.279
SDQ Conduct Problems Scale	-0.036	0.420	+0.036	0.485	-0.008	0.873	+0.088	0.036 *	-0.014	0.783	+0.130	0.006 **
SDQ Peer Problems Scale	-0.067	0.098	+0.074	0.154	+0.056	0.225	-0.188	<0.001 ***	-0.157	0.008 **	+0.102	0.041 *
SDQ Prosocial Scale	+0.119	0.007 **	-0.113	0.018 *	-0.056	0.222	+0.185	<0.001 ***	+0.106	0.023 *	-0.076	0.220
Behavioural Self-regulation Scale	+0.132	0.006 **	-0.074	0.104	-0.021	0.687	+0.226	<0.001 ***	+0.110	0.087	-0.059	0.278
Emotional Self-regulation Scale	+0.008	0.830	-0.042	0.382	-0.028	0.539	-0.019	0.711	+0.041	0.355	-0.145	0.027 *
Co-operation Scale	+0.081	0.080	-0.127	0.011 *	-0.002	0.972	+0.075	0.080	+0.098	0.073	-0.080	0.210

Models controlled for formal group, formal individual and informal individual ECEC use aged 2 to 3 and demographic covariates. Standardized model coefficients are given with p values: * = p < 0.05, ** = p < 0.01, *** = p < 0.001. Model coefficients represent the difference in the standardized outcome produced by a change of 2 standard deviations in the home environment covariate, controlling for all other covariates. Models are fitted to multiply imputed data. Sample size = 4583.

Relative effect size of ECEC, home environment and demographic factors on children's age three outcomes

The relative sizes of such effects on child outcomes associated with formal group, formal individual and informal individual ECEC use aged two to three, home environment variables and demographic factors were investigated[9]. The results are summarized in Figures 11 to 20. Plots show the effect size of all statistically significant variables in descending order of size. Early education use is shown in red, home environment factors in blue and demographic factors in green.

Cognitive Outcomes

Naming Vocabulary

Figure 11: Comparing effect sizes for Naming Vocabulary in terms of formal group, formal individual and informal individual ECEC use aged two to three and home environment and demographic covariates.

Relative effect sizes of factors on BAS Naming Vocabulary
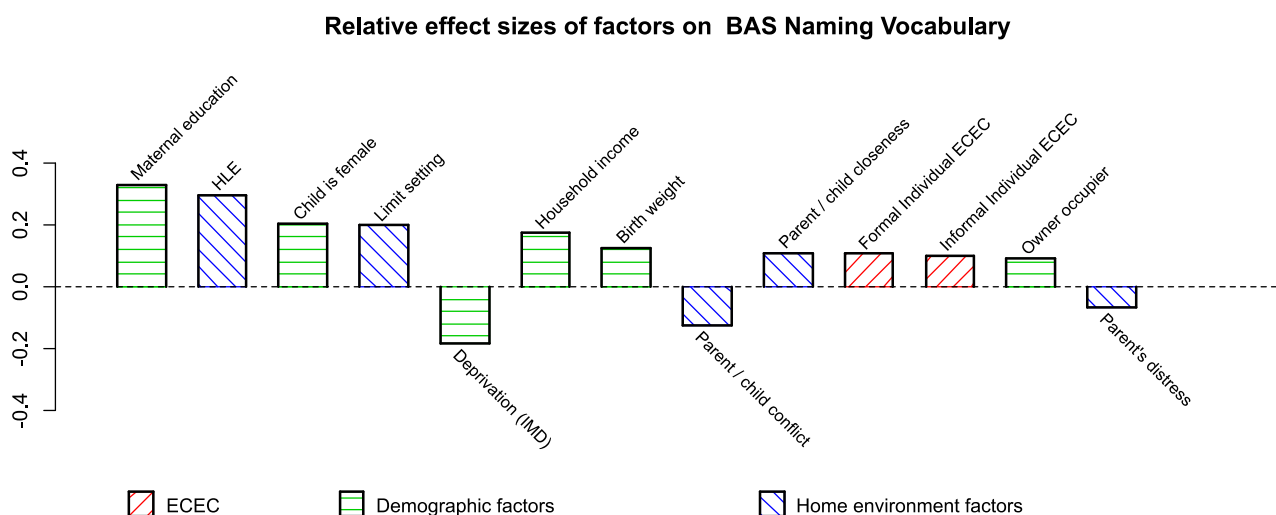
Sample size = 4,583.

Note: All statistically significant effects are shown in decreasing order of absolute size.

HLE = Home Learning Environment.

---

[9] Analysis also controlled for child's ethnic group, but because of the small sizes of most of the ethnic groups ethnicity effects were omitted from the results.

Picture Similarities

Figure 12: Comparing effect sizes for Picture Similarities in terms of formal group, formal individual and informal individual ECEC use aged two to three and home environment and demographic covariates.

Relative effect sizes of factors on BAS Picture Similarities

Sample size = 4,583.
Note: All statistically significant effects are shown in decreasing order of absolute size.

Socio-emotional outcomes

Hyperactivity

Figure 13: Comparing effect sizes for Hyperactivity in terms of formal group, formal individual and informal individual ECEC use aged two to three and home environment and demographic covariates.

Relative effect sizes of factors on SDQ Hyperactivity Scale

Sample size = 4,583.
Note: All statistically significant effects are shown in decreasing order of absolute size.

Emotional Symptoms

Figure 14: Comparing effect sizes for Emotional Symptoms in terms of formal group, formal individual and informal individual ECEC use aged two to three and home environment and demographic covariates.

Relative effect sizes of factors on SDQ Emotional Symptoms Scale

Sample size = 4,583.

Note: All statistically significant effects are shown in decreasing order of absolute size.

Conduct Problems

Figure 15: Comparing effect sizes for Conduct Problems in terms of formal group, formal individual and informal individual ECEC use aged two to three and home environment and demographic covariates.

Relative effect sizes of factors on SDQ Conduct Problems Scale

Sample size = 4,583.

Note: All statistically significant effects are shown in decreasing order of absolute size.

Peer Problems

Figure 16: Comparing effect sizes for Peer Problems in terms of formal group, formal individual and informal individual ECEC use aged two to three and home environment and demographic covariates.

Relative effect sizes of factors on SDQ Peer Problems Scale

Sample size = 4,583.

Note: All statistically significant effects are shown in decreasing order of absolute size.

Prosocial Behaviour

Figure 17: Comparing effect sizes for Prosocial Behaviour in terms of formal group, formal individual and informal individual ECEC use aged two to three and home environment and demographic variables.

Relative effect sizes of factors on SDQ Prosocial Scale

Sample size = 4,583.

Note: All statistically significant effects are shown in decreasing order of absolute size.

Behavioural Self-regulation

Figure 18: Comparing effect sizes for Behavioural Self-regulation in terms of formal group, formal individual and informal individual ECEC use aged two to three and home environment and demographic variables.

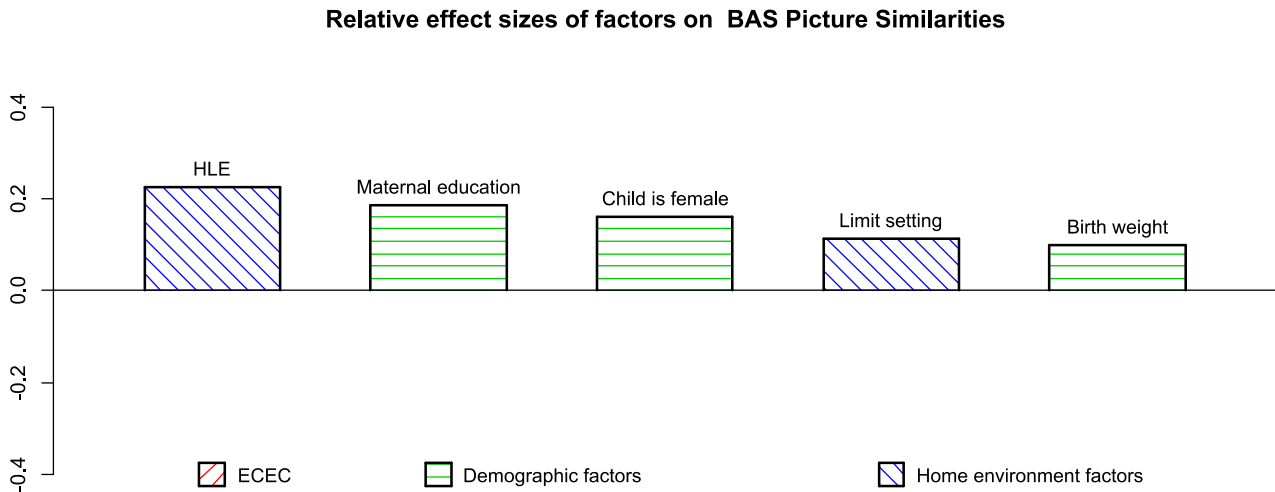**Relative effect sizes of factors on Behavioural Self-Regulation Scale**
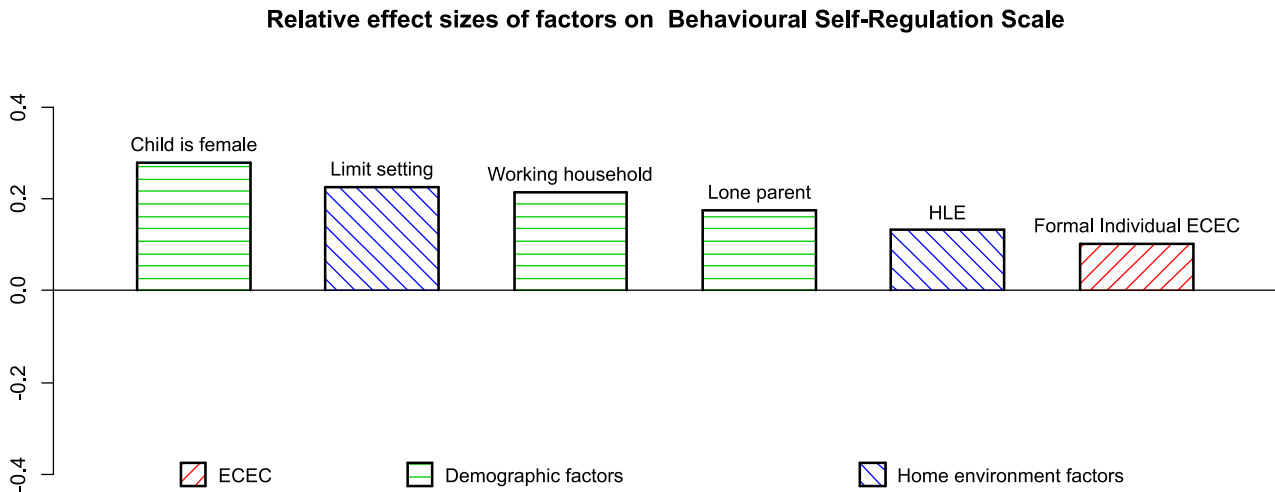
Sample size = 4,583.
Note: All statistically significant effects are shown in decreasing order of absolute size.

Emotional Self-regulation

Figure 19: Comparing effect sizes for Emotional Self-regulation in terms of formal group, formal individual and informal individual ECEC use aged two to three and home environment and demographic variables.

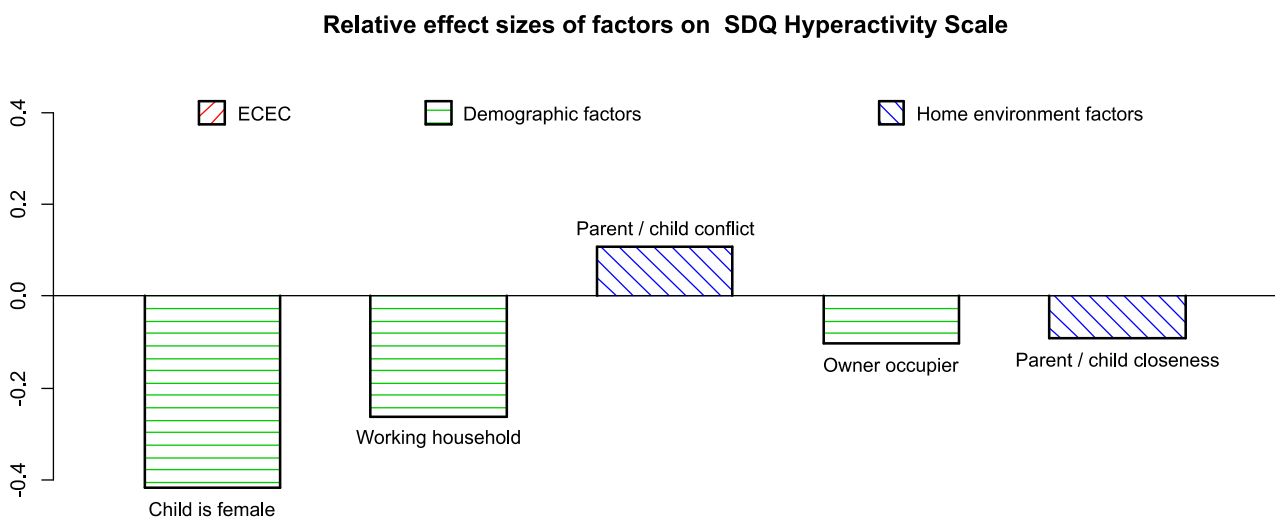**Relative effect sizes of factors on Emotional Self-Regulation Scale**
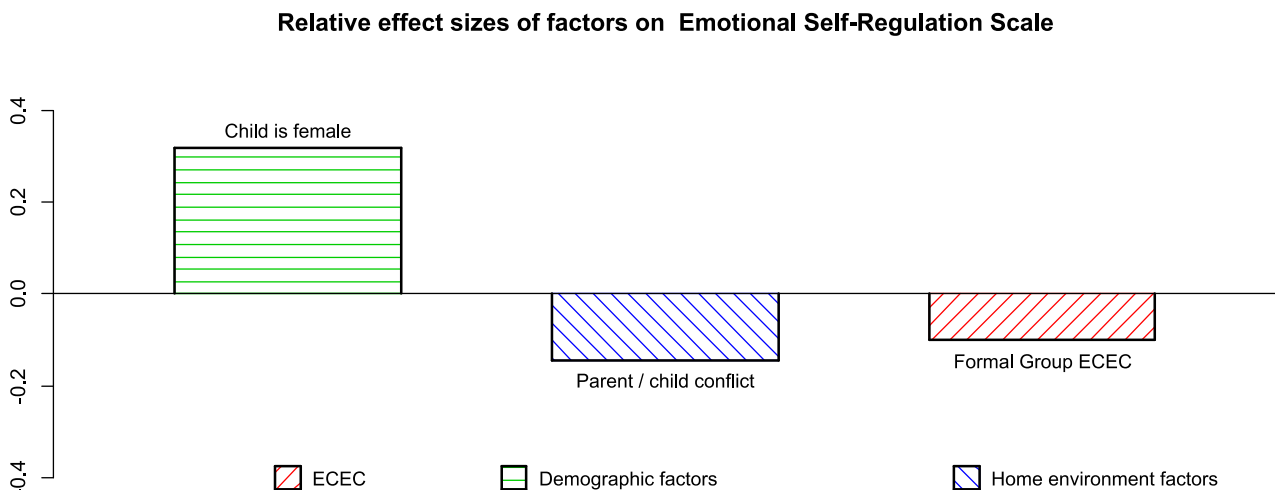
Sample size = 4,583.
Note: All statistically significant effects are shown in decreasing order of absolute size.

Co-operation

Figure 20: Comparing effect sizes for Co-operation in terms of formal group, formal individual and informal individual ECEC use aged two to three and home environment and demographic variables.
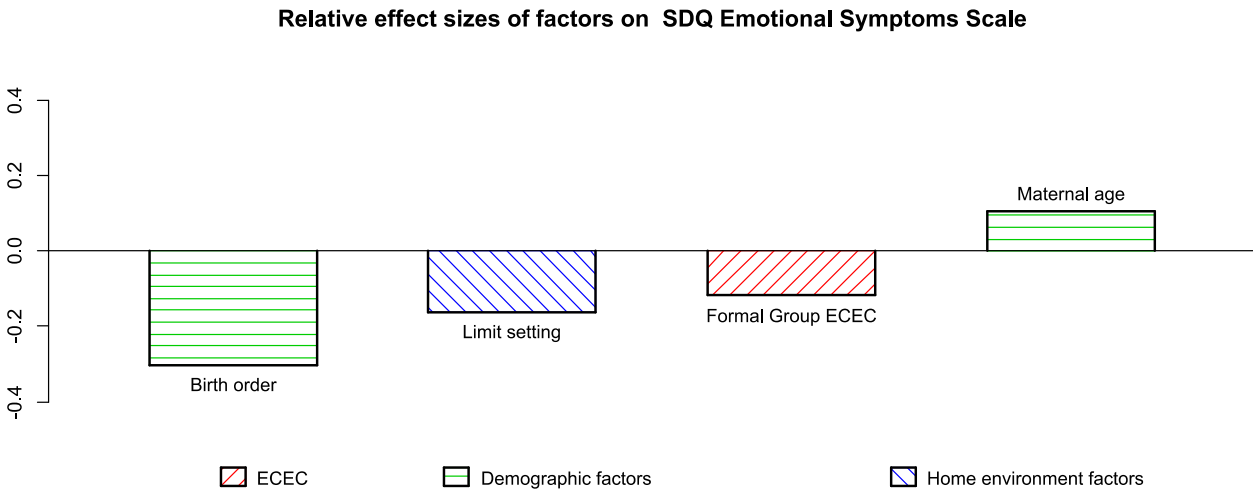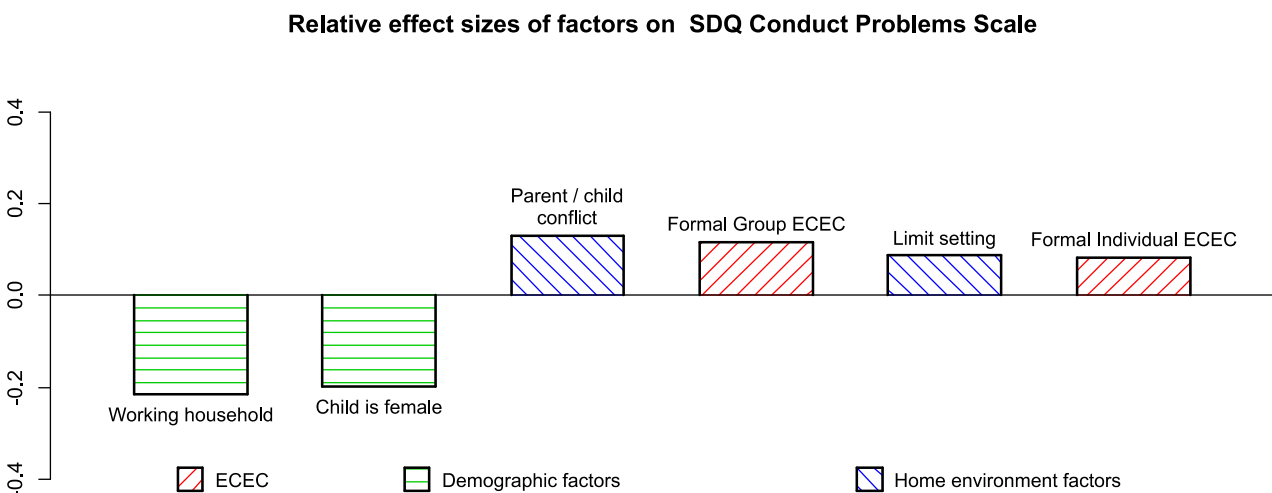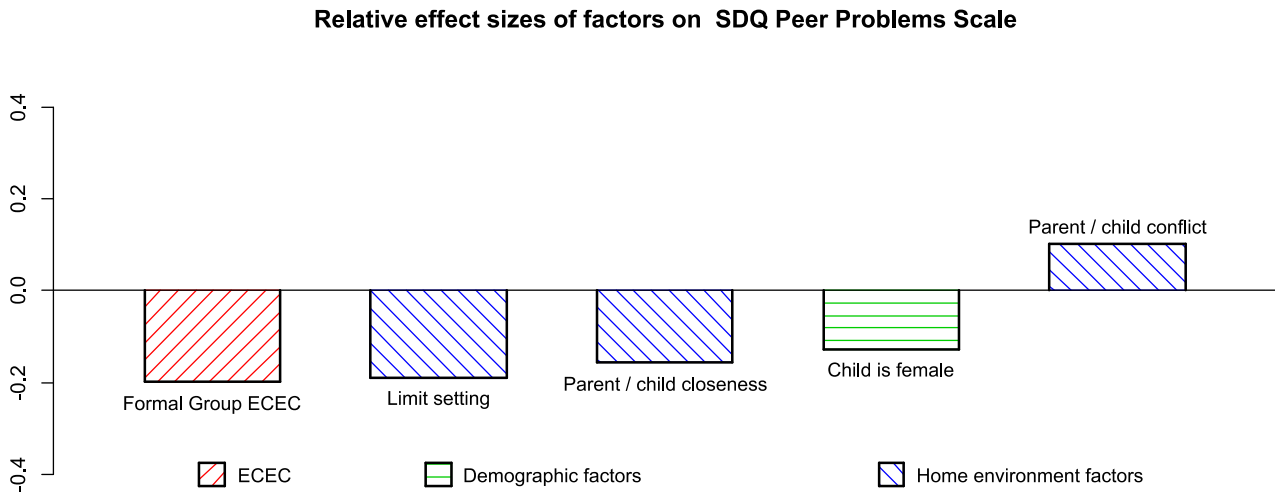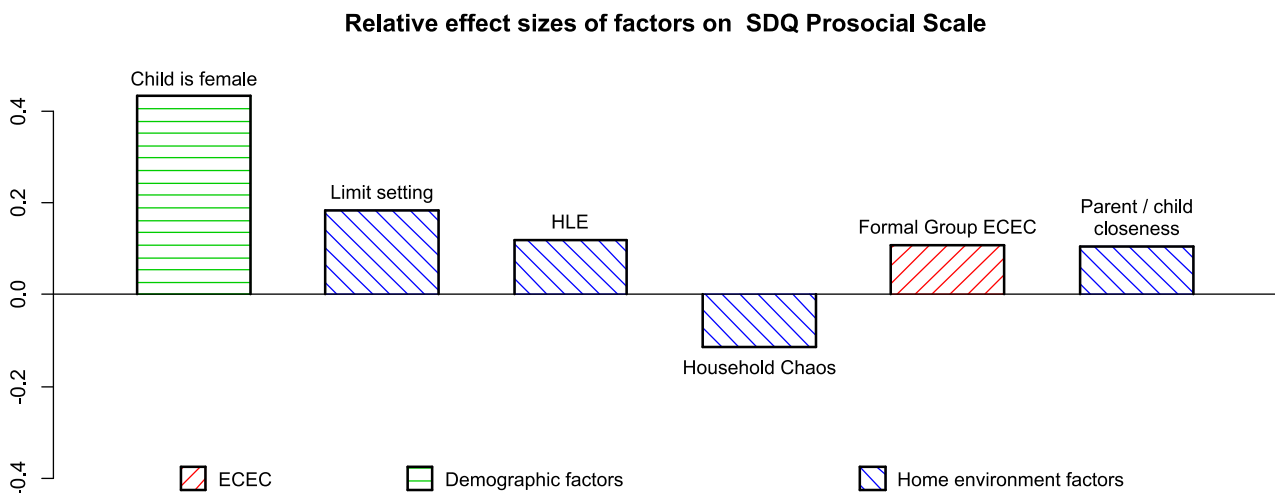
Relative effect sizes of factors on Co-operation Scale
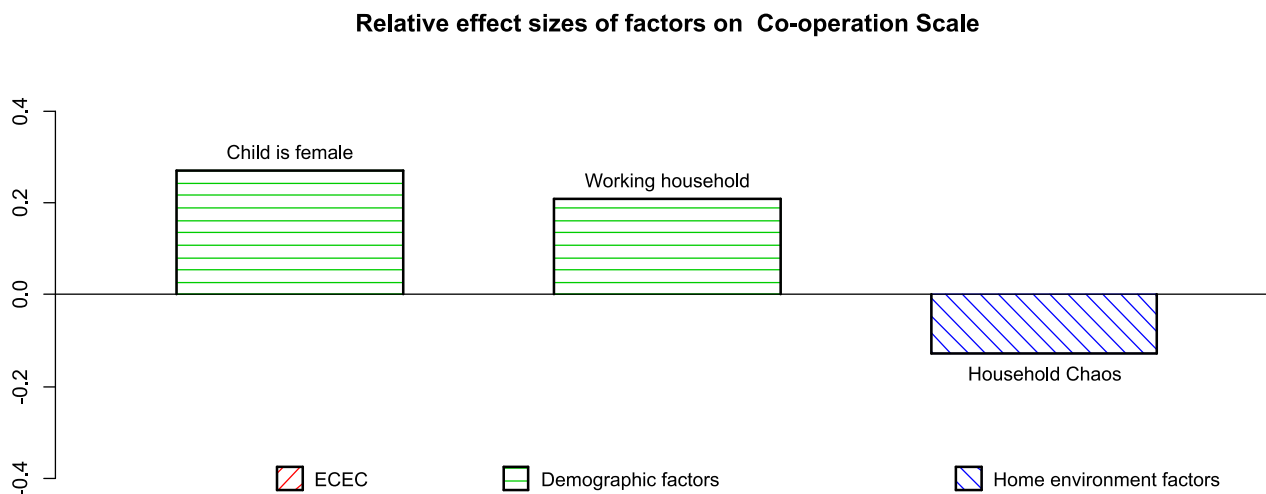
Sample size = 4,583.

Note: All statistically significant effects are shown in decreasing order of absolute size.

Interaction between ECEC use and the Home Learning Environment

Findings have shown that both ECEC use and Home Learning Environment Index were associated with child outcomes. We hypothesised that there may be an interaction between ECEC use and the Home Learning Environment Index: specifically, that the effect of ECEC use on the outcomes would be smaller when the Home Learning Environment Index score was high and the effect would be larger when the Home Learning Environment score was low. This may be characterised as a saturation effect; i.e. children already experiencing a rich home learning environment may have received enough "learning opportunities" and thus may derive less benefit from time in an ECEC setting than those whose home learning environment was less rich.

We tested for interactions between the effects of ECEC use and Home Learning Environment Index. Two models were fitted:

Model A, of outcome variable in terms of
- Formal group ECEC use
- Formal individual ECEC use
- Informal individual ECEC use
- Home learning environment index
- Other home environment and demographic covariates

Model B, of outcome variable in terms of
- Formal group ECEC use
- Formal individual ECEC use
- Informal individual ECEC use
- Home Learning Environment Index
- The interaction between Home Learning Environment Index and ECEC use
- Other home environment and demographic covariates

The ECEC use in the interaction term was each of formal group ECEC use, formal individual ECEC use and informal individual ECEC use in turn.

An ANOVA test comparison of Models A and B (described above) provides a test of whether there is a significant interaction between Home Learning Environment Index and a given type of ECEC usage. The results of these tests are given in Table 35.

Table 35: p-values from ANOVA tests comparing models with and without interactions between Home Learning Environment Index and each type of ECEC

Outcome	Type of ECEC		
	Formal group	Formal individual	Informal individual
Naming Vocabulary	0.342	0.749	0.707
Picture Similarities	0.144	0.011 *	0.372
SDQ Hyperactivity Scale	0.717	0.596	0.604
SDQ Emotional Symptoms Scale	0.485	0.655	0.693
SDQ Conduct Problems Scale	0.546	0.604	0.464
SDQ Peer Problems Scale	0.692	0.645	0.511
SDQ Prosocial Scale	0.649	0.503	0.629
Behavioural Self-regulation Scale	0.756	0.666	0.562
Emotional Self-regulation Scale	0.555	0.502	0.669
Co-operation Scale	0.619	0.643	0.717

Significant p-values are marked: * = p < 0.05.

There was a significant interaction between formal individual ECEC use (with childminders) and Home Learning Environment Index in the model of the outcome BAS Picture Similarities score. To investigate this interaction further, we fitted separate models for children with high and low Home Learning Environment Index scores. The results are given in Table 36.

A positive association was seen between Picture Similarities score and time in formal individual ECEC, but only for children with lower Home Learning Environment Index scores. This is in accord with the "saturation hypothesis".

Table 36: Results of models of child outcomes in terms of ECEC use between age two and three

Outcome	Group	Type of ECEC					
		Formal Group		Formal Individual		Informal Individual	
		Coef.	p	Coef.	p	Coef.	p
Picture Similarities	All children	+0.014	0.630	+0.043	0.139	-0.031	0.315
	Low HLE (≤	+0.065	0.114	+0.096	0.013 *	-0.028	0.493
	High HLE (>	-0.062	0.181	-0.026	0.570	-0.043	0.373

Models controlled for home environment and demographic covariates. Standardized model coefficients are given with p values: * = p < 0.05, ** = p < 0.01, *** = p < 0.001. Model coefficients represent the difference in the standardized outcome for a change of 2 standard deviations in the ECEC covariate, controlling for all covariates. Models are fitted to multiply imputed data. Sample size = 4583 (all children), = 2542 (HLE ≤ 25), = 2039 (HLE > 25).

References

Dunning, T. (2012). *Natural experiments in the social sciences: A design-based approach*. Cambridge, UK: Cambridge University Press.

Hesterberg, T., *Combining multiple imputation t, chi-square, and F inferences . Research report No 75*. Seattle, WA; MathSoft, 1998.
http://www.uvm.edu/~dhowell/StatPages/Missing_Data/tech75-mi-inference.pdf

Honaker J, King G, Blackwell M. Amelia II (R package) 2010. http://cran.r-project.org/web/packages/Amelia/index.html.

Melhuish, E., Gardiner, J., and Morris, S. (2017). Study of Early Education and Development (SEED): Impact Study on Early Education Use and Child Outcomes up to Age Three: Research report. London: DfE.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Chichester: Wiley.

Speight, S., Maisey, R., Chanfreau, J., Haywood, S., Lord C., and Hussey, D. (2015). *Study of Early Education and Development: Baseline survey of families. Research Report, July 2015*. DFE-RR480. ISBN: 978-1-78105-509-0. London: DfE.
http://www.seed.natcen.ac.uk/media/5645/Study_of_early_education_and_development_survey_of_families.pdf

Appendix A: Histograms of ECEC usage between age two and the Wave 2 survey

Figures 21 and 22 display histograms showing the frequencies with which different weekly amounts of formal and total ECEC were observed in the data for all children in the SEED sample.

Figure 21: Histogram of hours of formal early education used per week between a child's 2nd birthday and the Wave 2 survey.

Formal ECEC use between age 2 and the Wave 2 survey

Hours per week over the 38 weeks of the school terms
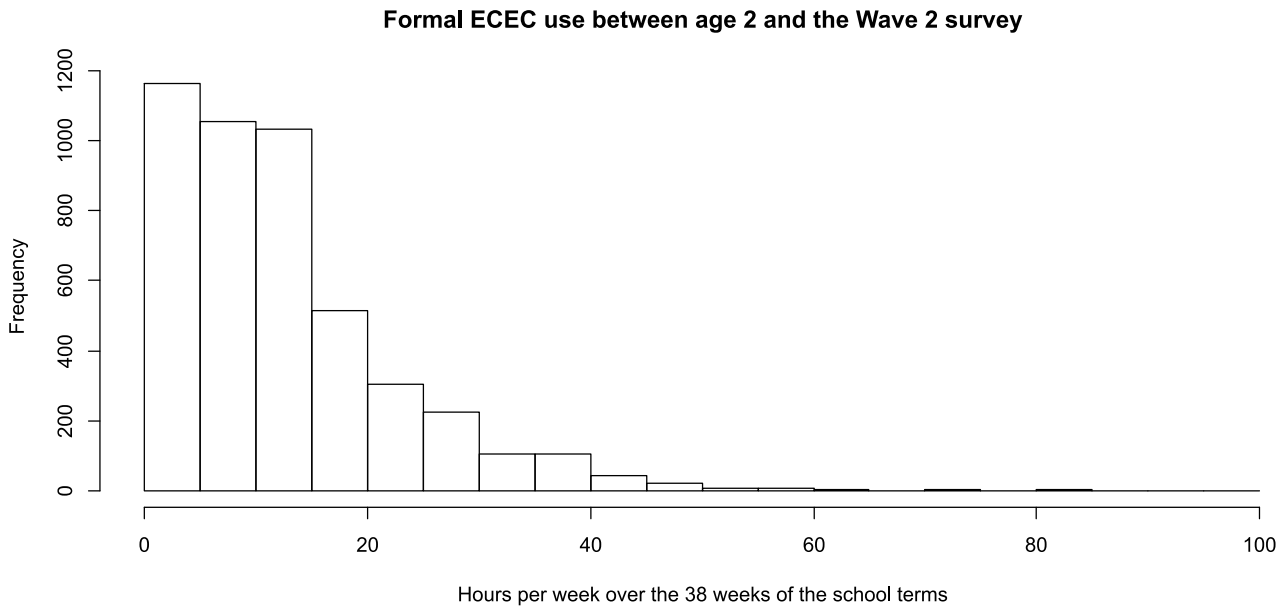
Figure 22: Histogram of hours of total early education used per week between a child's 2nd birthday and the Wave 2 survey.

Total ECEC use between age 2 and the Wave 2 survey

Hours per week over the 38 weeks of the school terms
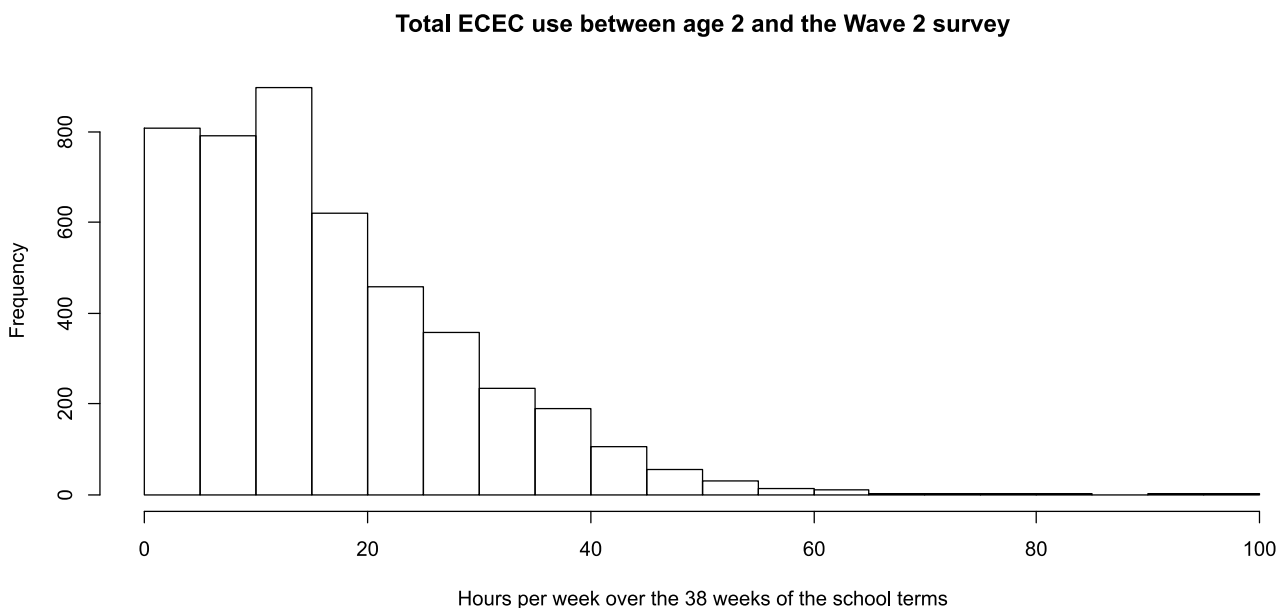
Appendix B: Variable summaries

A breakdown of the demographic variables is given in Tables 37 and 38. Summary statistics for the home environment variables are given in Table 39. Summary statistics for the British Ability Scales are given in Table 40. Summary statistics for the socio-emotional outcomes are available in Table 47 in Appendix E.

Table 37: Summary of the demographic covariates (part 1).

Variable	Level	N	%
Child's sex	Male	2397	52.30
	Female	2186	47.70
Child's ethnic group	White	3813	83.25
	Asian	303	6.62
	Black	192	4.19
	Mixed / other	272	5.94
Child's birth weight	≤ 3 kg	1024	22.43
	3-4 kg	3002	65.75
	>4 kg	540	11.83
Birth order	1	1983	43.27
	2	1597	34.85
	3+	1003	21.89
Maternal age at birth of child	≤ 25 years	1292	28.64
	25-29 years	1009	22.37
	29-34 years	1281	28.40
	>34 years	929	20.59
Number of sibs living in household (Wave 1)	0	1417	30.92
	1	1932	42.16
	2+	1234	26.93
Couple or loan parent household (Wave 1)	Couple	3361	73.34
	Lone parent	1222	26.66
Anyone working in household (Wave 1)	Someone working	3509	76.57
	No one working	1074	23.43
Household annual income (Wave 1)	Less than £10,000	658	14.36
	£10,000 to £20,000	1024	22.34
	£20,000 to £40,000	1372	29.94
	Greater than £40,000	1203	26.25
	Don't know / don't want to say	326	7.11
Index of multiple deprivation (Wave 1)	1 = least deprived	843	18.39
	2	812	17.72
	3	898	19.59
	4	936	20.42
	5 = most deprived	1094	23.87

Table 38: Summary of the demographic covariates (part 2).

Variable	Level	N	%
Deprivation group (Wave 1)	20% most deprived	1210	26.40
	20%-40% most deprived	1625	35.46
	60% least deprived	1748	38.14
Type of accommodation tenure (Wave 1)	Owner-occupier	1951	42.63
	Renting	2447	53.46
	Living rent free	179	3.91
Mother's highest qualification (Wave 1)	None / don't know	570	12.44
	GCSE D-G	270	5.89
	GCSE A-C	1130	24.66
	A levels / further education	1182	25.79
	Degree	922	20.12
	Higher degree	509	11.11
Highest parental socio-economic status (Wave 1)	Professional / managerial	705	15.39
	Lower managerial	1304	28.46
	Intermediate occupations	624	13.62
	Small employer / own account	364	7.94
	Lower supervisory	323	7.05
	Semi-routine	714	15.58
	Routine	348	7.59
	Not working	200	4.36

Table 39: Summary statistics for the home environment variables

Variable	Range	Mean	Median	SD
Home learning environment	0-35	23.85	25.00	6.90
Household chaos	4-18	8.02	8.00	2.31
Parent's Kessler psychological distress	6-30	9.35	8.00	3.92
Limit Setting scale	1-5	2.62	2.57	0.72
Parent / child closeness	3-15	14.43	15.00	1.35
Parent / child conflict	7-35	13.39	13.00	4.80

SD = Standard Deviation

Table 40: Summary statistics for BAS scales.

Variable	Range	Mean	Median	SD
BAS Naming Vocabulary (Verbal Ability)	27-80	51.13	54.00	11.95
BAS Picture Similarities (Non-verbal Ability)	24-80	47.92	49.00	9.47

Appendix C: Modelling effects associated with ECEC up to age three

The correlations between formal group, formal individual and informal individual ECEC before age one, aged one to two, and two to three are shown in Table 41.

Table 41: Correlations between ECEC use variables (Kendall's τ).

| | | Age 0 to 1 | | | Age 1 to 2 | | | Age 2 to 3 | | |
		Formal group	Formal individual	Informal individual	Formal group	Formal individual	Informal individual	Formal group	Formal individual	Informal individual
Age 0 to 1	**Formal group**	+1.000	-0.034	+0.102	+0.600	-0.051	+0.043	+0.281	-0.030	+0.053
	Formal individual		+1.000	+0.063	-0.046	+0.657	+0.021	-0.046	+0.498	+0.029
	Informal individual			+1.000	+0.048	+0.028	+0.628	+0.007	+0.034	+0.480
Age 1 to 2	**Formal group**				+1.000	-0.096	+0.018	+0.425	-0.084	+0.048
	Formal individual					+1.000	+0.011	-0.076	+0.723	+0.012
	Informal individual						+1.000	-0.015	+0.027	+0.652
Age 2 to 3	**Formal group**							+1.000	-0.118	-0.006
	Formal individual								+1.000	+0.028
	Informal individual									+1.000

In some instances, the correlations between ECEC usage age one to two and aged two to three are sufficiently high to present difficulties in separating the effects of these ECEC usage variables in regression models. Since childcare usage before age one is generally low (see Table 16) we therefore developed models in terms of mean ECEC use aged one to three.

This approach is also supported by a comparison of models with separate effects of ECEC aged one to two and aged two to three and models with effects of ECEC aged one to three using the AIC criterion; see Table 42. The AIC values are lower for the models with effects for ECEC use aged one to three. We conclude that there is no evidence that models with separate effects of ECEC aged one to two and two to three have greater explanatory power than models with a single effect of mean ECEC use aged one to three.

The results of regression models of the outcome variables in terms of formal group, formal individual and informal individual ECEC use aged one to three are given in Table 43. Models control for home environment and demographic covariates. Models are fitted to multiply imputed data.

Table 42: Comparing AIC values for two treatments of ECEC use age one to three years.

Outcome	Model A	Model B
Naming Vocabulary	12049.8	12033.5
Picture Similarities	12650.7	12636.7
SDQ Hyperactivity Scale	12729.4	12720.6
SDQ Emotional Symptoms Scale	13141.5	13126.1
SDQ Conduct Problems Scale	12993.1	12980.4
SDQ Peer Problems Scale	12940.5	12925.8
SDQ Prosocial Scale	12733.7	12717.1
Behavioural Self-regulation Scale	12839.7	12825.4
Emotional Self-regulation Scale	12939.5	12930.7
Co-operation Scale	12915.9	12902.8

The "A" models control separately for mean ECEC use aged 1 to 2 and mean ECEC use aged 2 to 3. The "B" models control for mean ECEC use aged 1 to 3. Models are in terms of formal group, formal individual and informal individual ECEC use. All models control for home environment and demographic covariates. Lower AIC values indicate better model fit.

Table 43: Results of models of child outcomes in terms of ECEC use between age one and three

Outcome	Type of ECEC					
	Formal Group		Formal Individual		Informal Individual	
	Coef.	p	Coef.	p	Coef.	p
Naming Vocabulary	+0.014	0.638	+0.105	<0.001 ***	+0.088	0.002 **
Picture Similarities	-0.023	0.455	+0.047	0.107	-0.032	0.305
SDQ Hyperactivity Scale	+0.026	0.502	+0.001	0.987	-0.010	0.804
SDQ Emotional Symptoms Scale	-0.115	0.020 *	-0.055	0.185	+0.019	0.688
SDQ Conduct Problems Scale	+0.104	0.010 **	+0.079	0.035 *	+0.004	0.912
SDQ Peer Problems Scale	-0.211	<0.001 ***	-0.029	0.506	+0.068	0.200
SDQ Prosocial Scale	+0.127	0.002 **	+0.045	0.249	-0.007	0.858
Behavioural Self-regulation Scale	+0.093	0.025 *	+0.095	0.020 *	-0.011	0.783
Emotional Self-regulation Scale	-0.080	0.055	-0.017	0.640	+0.033	0.487
Co-operation Scale	+0.067	0.106	+0.049	0.234	+0.013	0.769

Models control for home environment and demographic covariates. Standardized model coefficients are given along with p values: * = p < 0.05, ** = p < 0.01, *** = p < 0.001. Model coefficients represent the difference in the standardized outcome produced by a change of 1 standard deviation in the ECEC covariate, controlling for all other covariates. Sample size = 4583.

Where there were significant effects of ECEC in the initial models, further "detail models" were fitted in terms of specific levels of ECEC use. The results of these models are shown in Table 44.

Notes **on** Table 44
Level of ECEC use is mean hours per week over the 38 weeks of the school terms. Models controlled for home environment and demographic covariates. Standardized model coefficients are given along with p values: * = p < 0.05, ** = p < 0.01, *** = p < 0.001. Model coefficients represent the difference in the standardized outcome between children with a given level of ECEC use and the baseline (≤ 2 hours per week) group, controlling for all other covariates. Sample size = 4583.

Table 44: Analysis of ECEC use by level of usage band. ECEC use aged one to three is expressed as mean weekly usage over the 38 weeks of the school terms.

Outcome	Usage level (mean hours per week)	Type of ECEC					
		Formal Group		Formal Individual		Informal Individual	
		Coef.	p	Coef.	p	Coef.	p
Naming Vocabulary	>2-5	-0.028	0.443	+0.058	0.591	-0.067	0.201
	>5-15	+0.051	0.136	+0.168	0.020 *	+0.128	0.002 **
	>15-25	+0.053	0.327	+0.163	0.047 *	+0.116	0.020 *
	>25-35	+0.046	0.488	+0.167	0.149	+0.143	0.049 *
	>35	-0.068	0.393	+0.381	0.004 **	+0.071	0.431
SDQ Emotional Symptoms Scale	>2-5	-0.085	1.000	-0.050	1.000	-0.030	0.666
	>5-15	-0.099	0.095	+0.006	0.954	-0.029	0.584
	>15-25	-0.187	0.022 *	-0.112	0.335	-0.014	0.843
	>25-35	-0.181	0.038 *	-0.098	0.510	+0.088	0.363
	>35	-0.238	0.052	-0.213	0.303	+0.048	0.760
SDQ Conduct Problems Scale	>2-5	-0.037	0.459	-0.007	0.952	-0.009	0.894
	>5-15	+0.003	0.951	+0.089	0.394	-0.004	0.936
	>15-25	+0.094	0.157	+0.043	0.651	-0.090	0.143
	>25-35	+0.102	0.264	+0.340	0.023 *	+0.090	0.359
	>35	+0.222	0.035 *	+0.182	0.362	+0.060	0.673
SDQ Peer Problems Scale	>2-5	-0.110	0.050 *	+0.117	0.386	-0.013	0.838
	>5-15	-0.169	0.001 **	+0.003	0.974	+0.005	0.931
	>15-25	-0.258	<0.001 ***	-0.125	0.290	+0.022	0.738
	>25-35	-0.367	<0.001 ***	-0.021	0.887	+0.114	0.272
	>35	-0.358	<0.001 ***	-0.049	0.822	+0.206	0.288
SDQ Prosocial Scale	>2-5	+0.038	0.426	-0.049	0.707	+0.025	0.734
	>5-15	+0.078	0.100	+0.001	0.990	+0.036	0.484
	>15-25	+0.201	0.002 **	+0.127	0.226	+0.045	0.460
	>25-35	+0.189	0.035 *	+0.049	0.737	-0.024	0.792
	>35	+0.221	0.045 *	+0.192	0.335	-0.087	0.550
Behavioural Self-regulation Scale	>2-5	+0.058	0.265	-0.012	0.930	+0.016	0.811
	>5-15	+0.067	0.215	+0.057	0.513	+0.033	0.532
	>15-25	+0.129	0.118	+0.178	0.107	+0.023	0.752
	>25-35	+0.187	0.029 *	+0.145	0.296	-0.050	0.642
	>35	+0.167	0.148	+0.339	0.085	-0.067	0.603
Emotional Self-regulation Scale	>2-5	+0.026	0.604	+0.022	0.869	+0.029	0.656
	>5-15	-0.003	0.959	-0.056	0.542	+0.044	0.419
	>15-25	-0.045	0.522	+0.002	0.985	+0.127	0.101
	>25-35	-0.063	0.509	-0.143	0.282	-0.053	0.579
	>35	-0.208	0.057	+0.047	0.790	+0.060	0.683

Appendix D: Models of children's age three outcomes in terms of ECEC usage age two to three: Results of complete cases analysis

Models in terms of amount of ECEC used aged two to three.

We fitted complete cases models of children's three year old outcomes in terms of ECEC use aged two to three. Models controlled for demographic and home environment variables. These models are the same as the models discussed in Chapter 4 of the Research Report, only rather than fitting the models to multiply imputed data the models are fitted to complete cases data (that is, any observations with missing values for any variables are omitted.

Initial models

We analysed the child outcome variables in terms of the amount of ECEC use in three categories: formal group ECEC, formal individual ECEC and informal individual ECEC. Models controlled for home environment and demographic covariates. Model results are shown in Table 45.

Table 45: Results of models of child outcomes in terms of ECEC use between age two and three. Complete cases models.

| Outcome | N | Type of ECEC | | | | | |
| | | Formal Group | | Formal Individual | | Informal Individual | |
		Coef.	p	Coef.	p	Coef.	p
Naming Vocabulary	4232	+0.008	0.785	+0.108	<0.001 ***	+0.094	0.001 **
Picture Similarities	4149	+0.014	0.658	+0.040	0.175	-0.022	0.489
SDQ Hyperactivity	2229	+0.064	0.147	-0.032	0.519	-0.010	0.834
SDQ Emotional Symptoms	2229	-0.115	0.013 *	-0.103	0.048 *	+0.007	0.884
SDQ Conduct Problems	2229	+0.134	0.003 **	+0.036	0.481	-0.018	0.705
SDQ Peer Problems	2229	-0.192	<0.001 ***	-0.027	0.592	+0.087	0.064
SDQ Prosocial	2229	+0.119	0.007 **	+0.075	0.131	-0.004	0.932
Behavioural Self-regulation	2227	+0.075	0.092	+0.111	0.029 *	-0.022	0.635
Emotional Self-regulation	2227	-0.098	0.030 *	+0.011	0.829	+0.076	0.108
Co-operation	2227	+0.046	0.303	+0.094	0.066	+0.035	0.461

Models control for home environment and demographic covariates. Standardized model coefficients are given along with p values: * = p < 0.05, ** = p < 0.01, *** = p < 0.001. Model coefficients represent the difference in the standardized outcome for a change of 2 standard deviations in the ECEC covariate, controlling for all covariates. Models are fitted to complete cases. N = sample size.

Detail models

Where there were significant effects in the initial models, we re-fitted the models replacing the continuous ECEC use covariates with multilevel factors modelling specific

levels of ECEC use. The results of the detail models are shown in Table 46 and summarized graphically in Figures 23 to 29.

Table 46: Results of models of child outcomes in terms of levels of ECEC use between ages two and three. Complete cases models.

Outcome	N	Usage level (mean hours per week)	Type of ECEC					
			Formal Group		Formal Individual		Informal Individual	
			Coef.	p	Coef.	p	Coef.	p
Naming Vocabulary	4232	>2-5	+0.092	0.063	+0.036	0.796	+0.066	0.236
		>5-15	+0.047	0.198	+0.113	0.128	+0.123	0.002 **
		>15-25	+0.077	0.100	+0.209	0.018 *	+0.139	0.013 *
		>25-35	+0.067	0.329	+0.085	0.575	+0.156	0.125
		>35	-0.033	0.687	+0.498	0.001 **	+0.160	0.179
SDQ Emotional Symptoms Scale	2229	>2-5	-0.005	0.956	-0.041	0.839	-0.047	0.565
		>5-15	-0.089	0.168	+0.019	0.872	+0.034	0.581
		>15-25	-0.140	0.072	-0.213	0.147	+0.040	0.647
		>25-35	-0.156	0.145	-0.154	0.572	+0.165	0.301
		>35	-0.287	0.021 *	-0.382	0.255	-0.028	0.891
SDQ Conduct Problems Scale	2229	>2-5	-0.035	0.670	+0.039	0.843	-0.049	0.543
		>5-15	+0.023	0.713	+0.199	0.091	-0.135	0.023 *
		>15-25	+0.024	0.749	-0.139	0.331	+0.044	0.610
		>25-35	+0.128	0.218	+0.116	0.662	+0.044	0.780
		>35	+0.437	<0.001 ***	+0.446	0.173	+0.001	0.996
SDQ Peer Problems Scale	2229	>2-5	-0.086	0.290	+0.128	0.517	-0.010	0.902
		>5-15	-0.110	0.082	+0.072	0.541	+0.034	0.562
		>15-25	-0.286	<0.001 ***	-0.149	0.295	+0.058	0.499
		>25-35	-0.389	<0.001 ***	-0.179	0.498	+0.415	0.008 **
		>35	-0.363	0.003 **	+0.087	0.790	+0.100	0.610
SDQ Prosocial Scale	2229	>2-5	+0.116	0.143	-0.181	0.347	+0.058	0.454
		>5-15	+0.023	0.708	-0.045	0.697	+0.081	0.160
		>15-25	+0.143	0.053	+0.284	0.041 *	-0.004	0.962
		>25-35	+0.279	0.006 **	-0.068	0.793	-0.295	0.051
		>35	+0.144	0.222	+0.168	0.598	+0.165	0.387
Behavioural Self-regulation Scale	2227	>2-5	+0.136	0.092	-0.217	0.265	+0.113	0.155
		>5-15	+0.081	0.197	-0.047	0.686	+0.038	0.514
		>15-25	+0.103	0.168	+0.340	0.016 *	+0.018	0.831
		>25-35	+0.129	0.209	-0.148	0.572	-0.334	0.030 *
		>35	+0.150	0.210	+0.533	0.099	+0.054	0.780
Emotional Self-regulation Scale	2227	>2-5	+0.024	0.771	-0.096	0.625	+0.109	0.171
		>5-15	-0.024	0.705	-0.219	0.062	+0.099	0.095
		>15-25	+0.030	0.695	+0.261	0.068	+0.039	0.651
		>25-35	-0.034	0.741	-0.142	0.589	+0.136	0.380
		>35	-0.427	<0.001 ***	-0.145	0.656	+0.216	0.269

Notes on Table 46: Level of ECEC use is mean hours per week over the 38 weeks of the school terms. Models controlled for home environment and demographic covariates. Standardized model coefficients are given along with p values: * = p < 0.05, ** = p < 0.01, *** = p < 0.001. Model coefficients represent the

difference in the standardized outcome between children with a given level of ECEC use and the baseline (≤ 2 hours per week) group, controlling for all other covariates. N = sample size.

Figure 23: Association of ECEC and Naming Vocabulary; complete cases model.
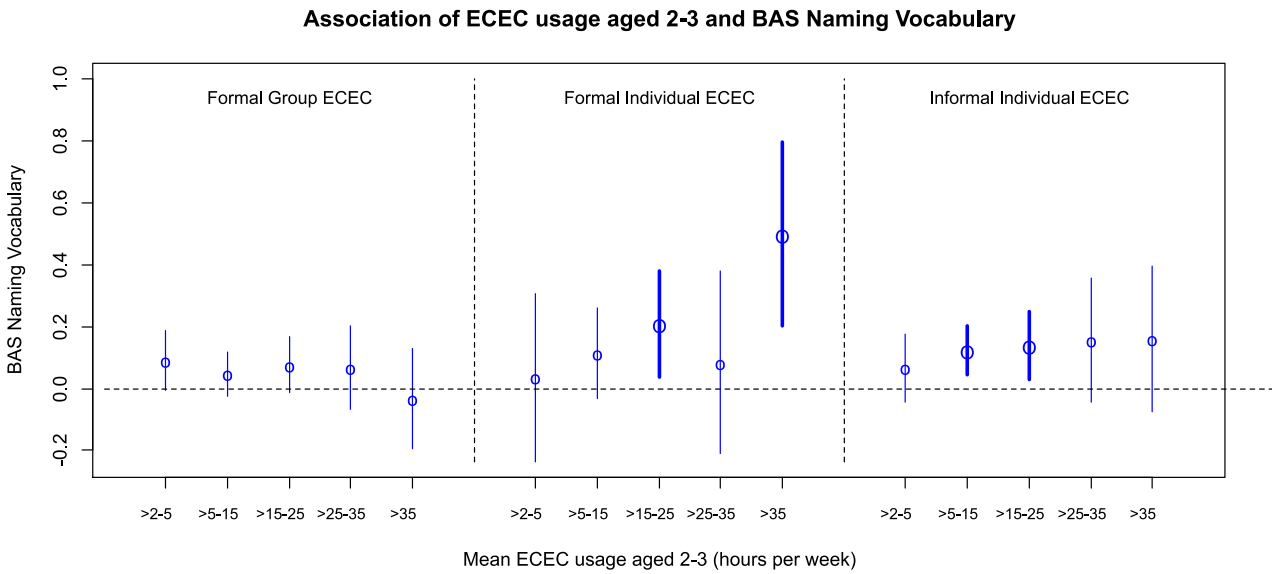
Association of ECEC usage aged 2-3 and BAS Naming Vocabulary

Model is in terms of formal group, formal individual and informal individual ECEC. Point estimates are shown ("O") with vertical lines giving a 95% confidence interval. The horizontal line represents the ≤ 2 hours of ECEC per week reference level. Statistically significant effects are shown in bold.

Figure 24: Association of ECEC and SDQ Emotional Symptoms Scale; complete cases model.
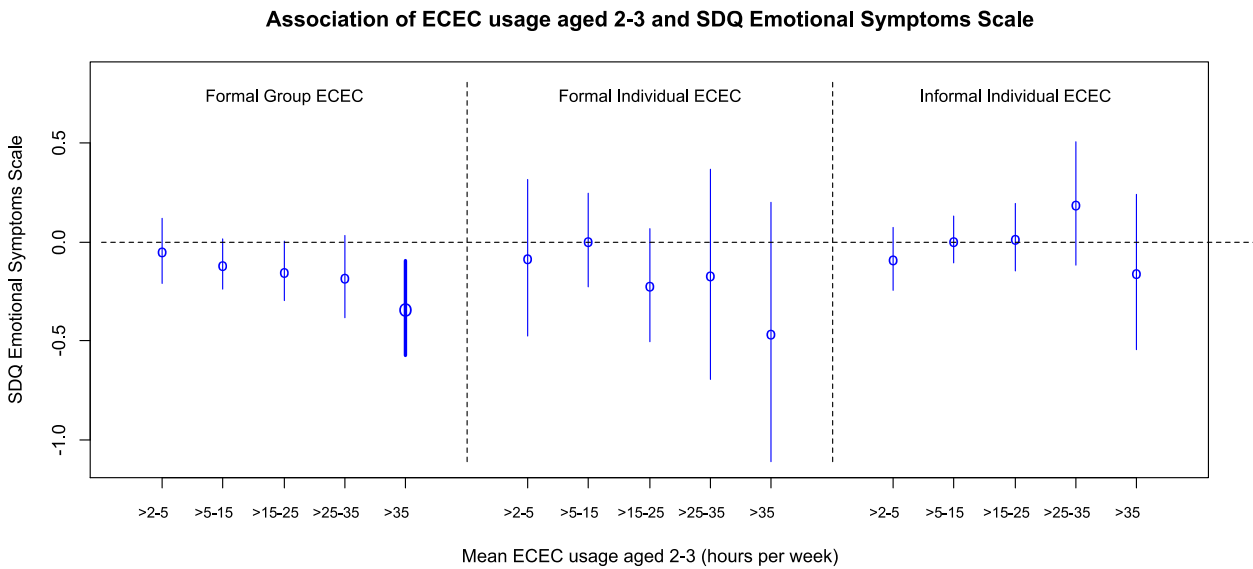
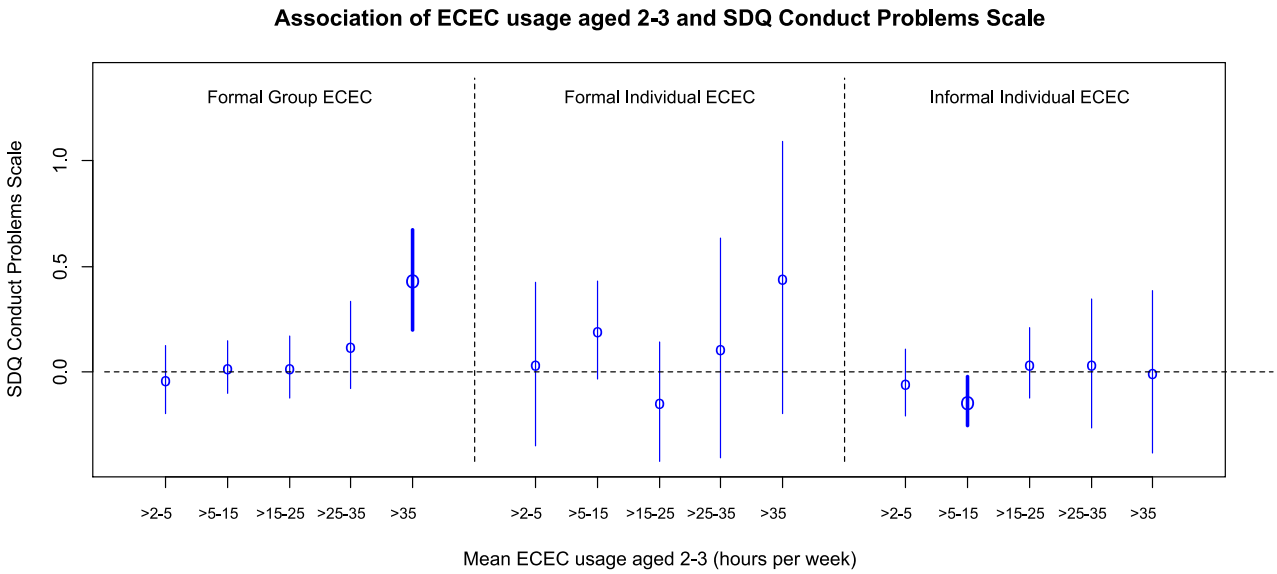**Association of ECEC usage aged 2-3 and SDQ Emotional Symptoms Scale**

Model is in terms of formal group, formal individual and informal individual ECEC. Point estimates are shown ("O") with vertical lines giving a 95% confidence interval. The horizontal line represents the ≤ 2 hours of ECEC per week reference level. Statistically significant effects are shown in bold.

78

Figure 25: Association of ECEC and SDQ Conduct Problems Scale; complete cases model.

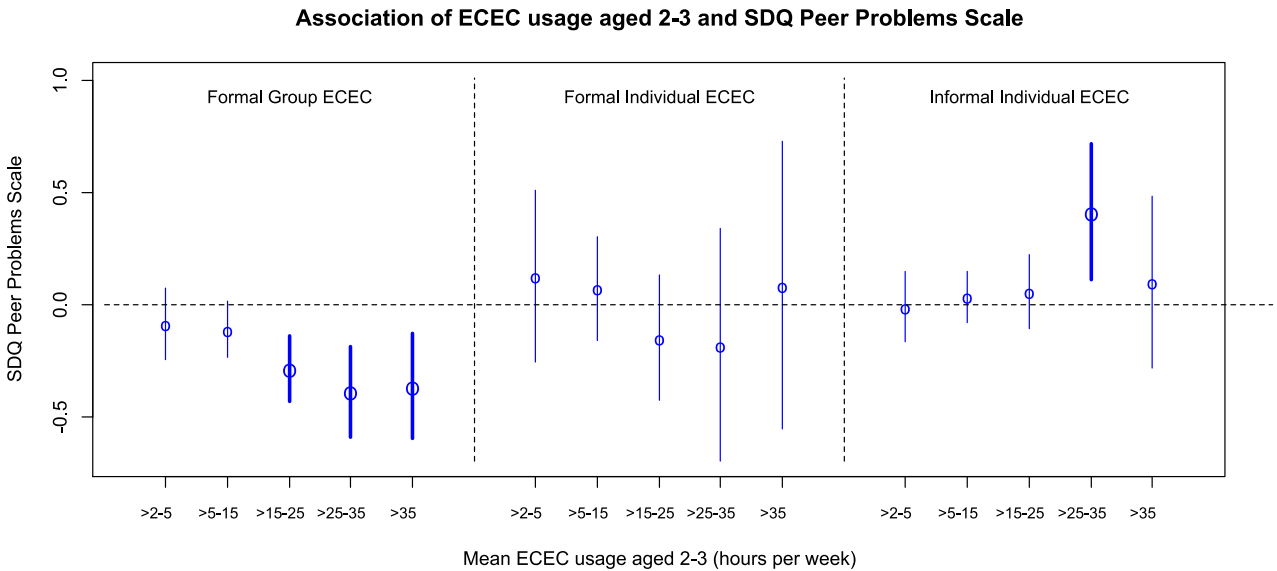Association of ECEC usage aged 2-3 and SDQ Conduct Problems Scale

Model is in terms of formal group, formal individual and informal individual ECEC. Point estimates are shown ("O") with vertical lines giving a 95% confidence interval. The horizontal line represents the ≤ 2 hours of ECEC per week reference level. Statistically significant effects are shown in bold.

Figure 26: Association of ECEC and SDQ Peer Problems Scale; complete cases model.

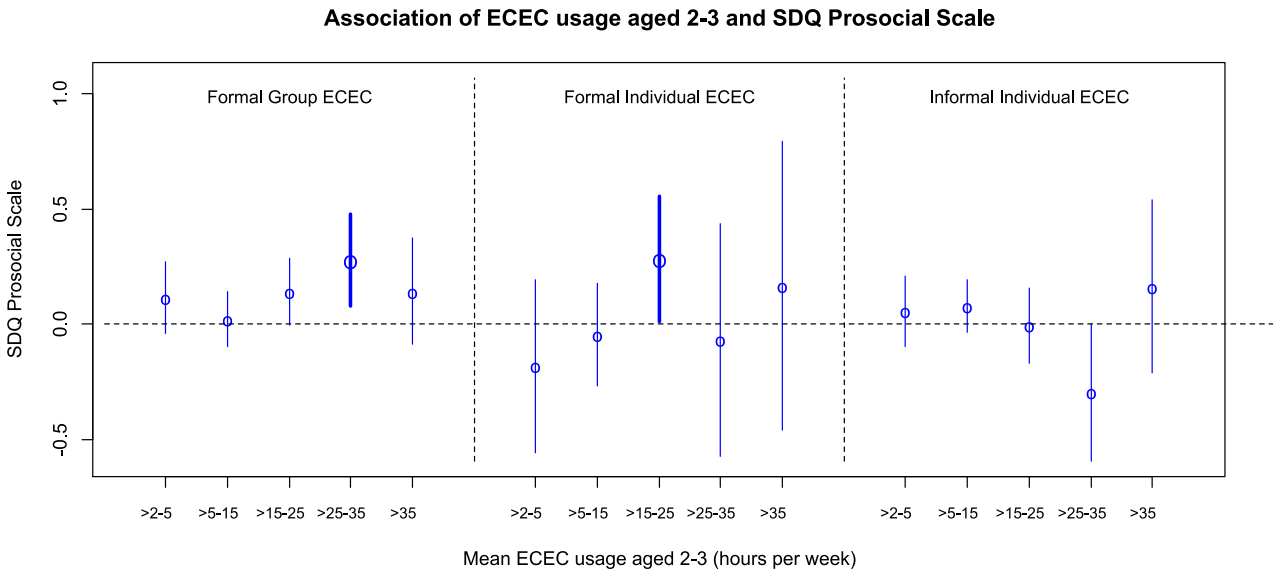Association of ECEC usage aged 2-3 and SDQ Peer Problems Scale

Model is in terms of formal group, formal individual and informal individual ECEC. Point estimates are shown ("O") with vertical lines giving a 95% confidence interval. The horizontal line represents the ≤ 2 hours of ECEC per week reference level. Statistically significant effects are shown in bold.

Figure 27: Association of ECEC and SDQ Prosocial Scale; complete cases model

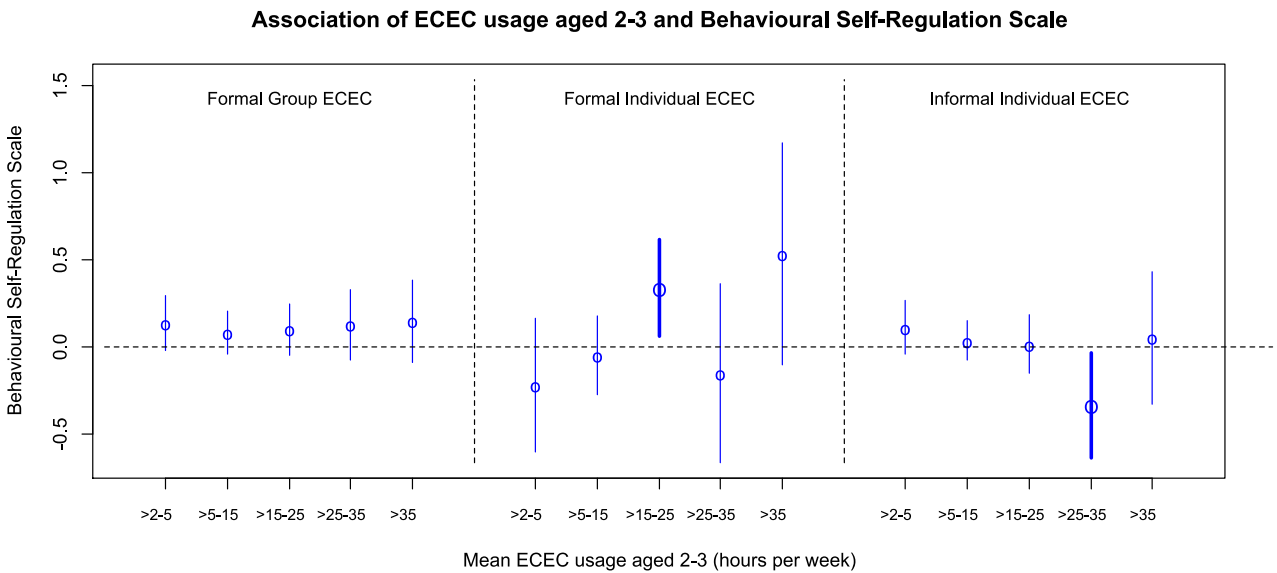Association of ECEC usage aged 2-3 and SDQ Prosocial Scale

Model is in terms of formal group, formal individual and informal individual ECEC. Point estimates are shown ("O") with vertical lines giving a 95% confidence interval. The horizontal line represents the ≤ 2 hours of ECEC per week reference level. Statistically significant effects are shown in bold.

Figure 28: Association of ECEC and Behavioural Self-regulation Scale; complete cases model.

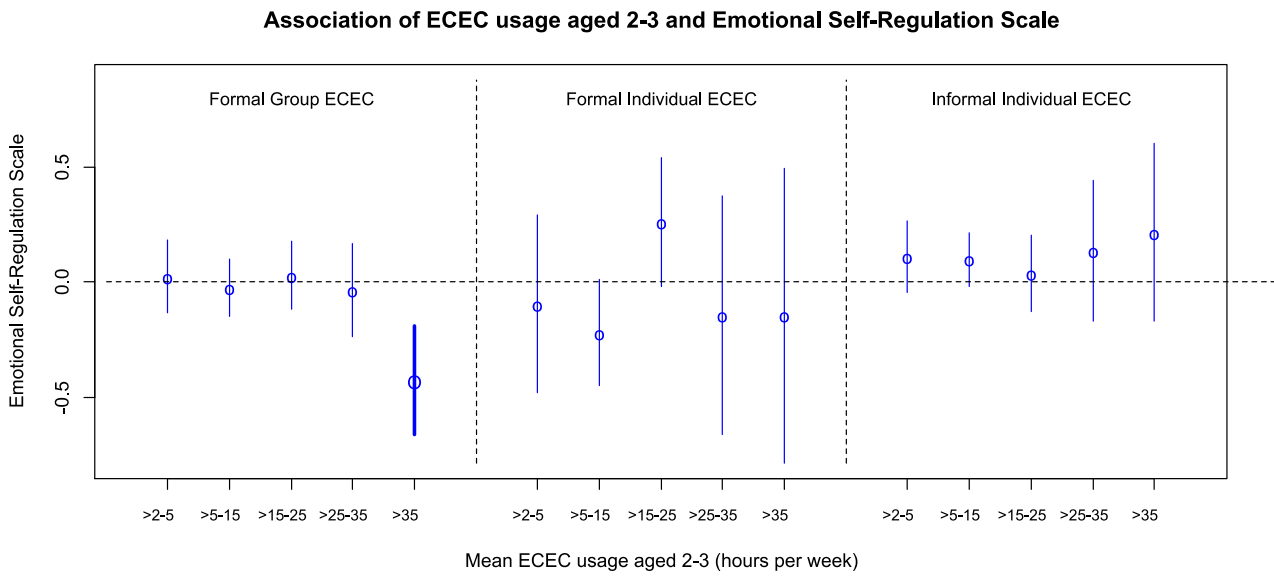Association of ECEC usage aged 2-3 and Behavioural Self-Regulation Scale

Model is in terms of formal group, formal individual and informal individual ECEC. Point estimates are shown ("O") with vertical lines giving a 95% confidence interval. The horizontal line represents the ≤ 2 hours of ECEC per week reference level. Statistically significant effects are shown in bold.

Figure 29: Association of ECEC and Emotional Self-regulation Scale; complete cases model.

Model is in terms of formal group, formal individual and informal individual ECEC. Point estimates are shown ("O") with vertical lines giving a 95% confidence interval. The horizontal line represents the ≤ 2 hours of ECEC per week reference level. Statistically significant effects are shown in bold.

Appendix E: Results using parental assessment of socio-emotional scales

Summary statistics for ECEC provider assessed and parentally assessed socio-emotional scales are shown in Table 47.

We fitted initial models of parentally assessed socio-emotional outcomes in terms of Early Childhood Education and Care (ECEC) between ages two and three. Models were in terms of formal group ECEC, formal individual ECEC and informal individual ECEC. Models controlled for home environment and demographic covariates. Models were fitted to multiply imputed data. The results of the initial models are given in Table 48.

Where there were significant effects of ECEC in the initial models we fitted further detail models with separate effects for different levels of mean weekly ECEC use aged two to three. The results of these detail models are summarized in Table 49 and represented graphically in Figures 30 to 34.

SDQ Emotional Symptoms Scale
Parentally assessed SDQ Emotional Symptoms Scale shows a negative association with formal individual ECEC use (Table 48). The detail model shows a specific benefit for children having > 35 hours of formal group ECEC per week (see Table 49 and Figure 30).

SDQ Peer Problems Scale
Parentally assessed peer problems scores are lower for children who have higher levels of formal group and formal individual ECEC use (see Tables 47 and 48 and Figure 31). High levels of informal individual ECEC (>25-35 hours per week) are associated with higher parentally assessed Peer Problems scores (see Table 49 and Figure 31).

SDQ Prosocial Scale
Parentally assessed SDQ Prosocial Scale shows a positive association with formal group ECEC use aged two to three (Table 48). Children having an average of > 15 hours per week formal group ECEC use show significantly higher parentally assessed SDQ Prosocial scores (Table 49, Figure 32).

Behavioural Self-regulation Scale
Parentally assessed Behavioural Self-regulation is positively associated with group ECEC use (Table 48). Children having an average of > 15 hours per week formal group ECEC exhibit significantly higher Behavioural Self-regulation (Table 49, Figure 33).

Co-operation Scale
Parentally assessed Co-operation Scale shows a positive association with formal group ECEC use (Table 48). Children having an average of >35 hours per week formal group

ECEC aged two to three have significantly higher Co-operation scores than the reference group (Table 49, Figure 34).

Table 47: Summary statistics for the ECEC provider and parentally assessed SDQ scales and additional related scales.

SDQ and additional related scales	Range	ECEC provider assessment			Parent assessment			p-value of t test
		Mean	Median	SD	Mean	Median	SD	
Prosocial Behaviour	0-10	6.75	7.00	2.42	7.81	8.00	1.79	<0.001 ***
Hyperactivity	0-10	3.01	2.00	2.65	3.76	4.00	2.21	<0.001 ***
Emotional Symptoms	0-10	1.29	1.00	1.76	1.47	1.00	1.47	<0.001 ***
Conduct Problems	0-10	1.28	0.00	1.87	2.37	2.00	1.85	<0.001 ***
Peer Problems	0-10	1.63	1.00	1.87	1.49	1.00	1.55	0.005 **
Behavioural Self-regulation	0-10	7.22	7.00	2.05	7.32	7.00	1.75	0.055
Emotional Self-regulation	0-10	7.58	8.00	2.23	6.42	7.00	2.08	<0.001 ***
Co-operation	0-10	7.81	8.00	2.14	7.51	8.00	1.72	<0.001 ***

SD = Standard Deviation. The p-value is from a t-test for a difference in means between the parent and ECEC provider assessed scales.

Table 48: Results of models of parentally assessed socio-emotional outcomes in terms of ECEC use between age two and three

| Outcome	Type of ECEC					
	Formal Group		Formal Individual		Informal Individual	
	Coef.	p	Coef.	p	Coef.	p
SDQ Hyperactivity Scale	-0.029	0.311	+0.022	0.439	+0.027	0.342
SDQ Emotional Symptoms Scale	-0.016	0.594	-0.059	0.042 *	+0.014	0.636
SDQ Conduct Problems Scale	+0.022	0.435	+0.022	0.419	+0.033	0.226
SDQ Peer Problems Scale	-0.225	<0.001 ***	-0.094	0.001 **	+0.077	0.010 **
SDQ Prosocial Scale	+0.152	<0.001 ***	+0.040	0.165	-0.019	0.518
Behavioural Self-regulation Scale	+0.110	<0.001 ***	+0.029	0.316	-0.020	0.502
Emotional Self-regulation Scale	+0.046	0.104	-0.004	0.869	+0.007	0.812
Co-operation Scale	+0.109	<0.001 ***	+0.013	0.649	-0.030	0.294

Models control for home environment and demographic covariates. Standardized model coefficients are given along with p values: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Model coefficients represent the difference in the standardized outcome produced by a change of 2 standard deviations in the ECEC covariate, controlling for all other covariates.

Table 49: Results of models of parentally assessed socio-emotional outcomes in terms of ECEC use between age two and three

| Outcome	Usage level (mean hours per week)	Type of ECEC					
		Formal Group		Formal Individual		Informal Individual	
		Coef.	p	Coef.	p	Coef.	p
SDQ Emotional Symptoms Scale (parent assessment)	>2-5	-0.028	0.574	-0.170	0.230	-0.040	0.482
	>5-15	-0.027	0.443	-0.052	0.502	-0.003	0.933
	>15-25	-0.088	0.061	-0.005	0.951	+0.053	0.356
	>25-35	+0.024	0.726	-0.200	0.180	+0.187	0.069
	>35	-0.241	0.005 **	-0.303	0.060	+0.071	0.544
SDQ Peer Problems Scale (parent assessment)	>2-5	-0.037	0.447	+0.004	0.979	-0.004	0.943
	>5-15	-0.119	<0.001 ***	-0.049	0.527	-0.011	0.792
	>15-25	-0.261	<0.001 ***	-0.237	0.008 **	+0.081	0.159
	>25-35	-0.302	<0.001 ***	-0.305	0.041 *	+0.235	0.023 *
	>35	-0.483	<0.001 ***	-0.204	0.206	+0.224	0.055
SDQ Prosocial Scale (parent assessment)	>2-5	0.000	0.993	-0.011	0.940	+0.050	0.375
	>5-15	+0.035	0.328	+0.019	0.810	+0.012	0.760
	>15-25	+0.123	0.009 **	+0.209	0.019 *	-0.093	0.106
	>25-35	+0.246	<0.001 ***	+0.074	0.621	-0.051	0.625
	>35	+0.337	<0.001 ***	-0.062	0.700	+0.075	0.522
Behavioural Self-regulation Scale (parent assessment)	>2-5	-0.029	0.557	+0.241	0.093	+0.061	0.286
	>5-15	+0.002	0.963	+0.037	0.639	-0.007	0.859
	>15-25	+0.098	0.039 *	+0.032	0.723	-0.126	0.029 *
	>25-35	+0.168	0.018 *	+0.072	0.632	-0.074	0.476
	>35	+0.222	0.010 *	+0.104	0.525	+0.173	0.143
Co-operation Scale (parent assessment)	>2-5	+0.010	0.835	-0.029	0.832	-0.092	0.097
	>5-15	+0.053	0.133	+0.053	0.481	-0.065	0.105
	>15-25	+0.088	0.055	+0.058	0.506	-0.088	0.118
	>25-35	+0.110	0.108	+0.036	0.805	-0.176	0.081
	>35	+0.271	0.001 **	-0.055	0.728	+0.125	0.274

Level of ECEC use is mean hours per week over the 38 weeks of the school terms. Models control for home environment and demographic covariates. Standardized model coefficients are given along with p values: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Model coefficients represent the difference in the standardized outcome between children with a given level of ECEC use and the baseline (≤ 2 hours per week) group, controlling for all covariates.

Figure 30: Association of ECEC and SDQ Emotional Symptoms Scale (parent assessment).
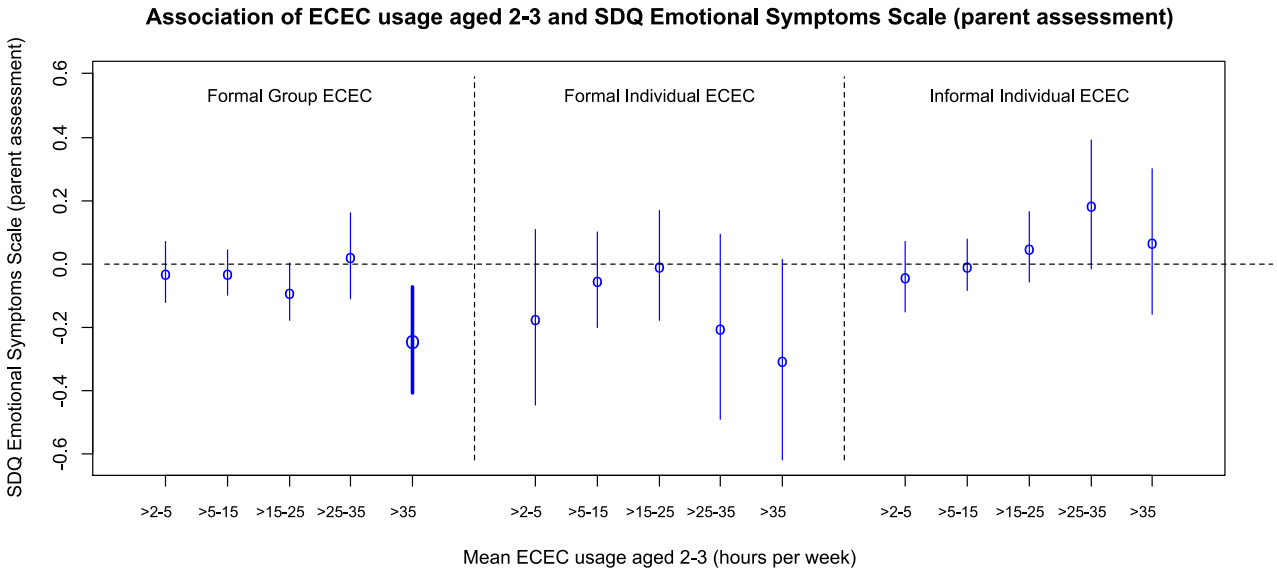
Association of ECEC usage aged 2-3 and SDQ Emotional Symptoms Scale (parent assessment)
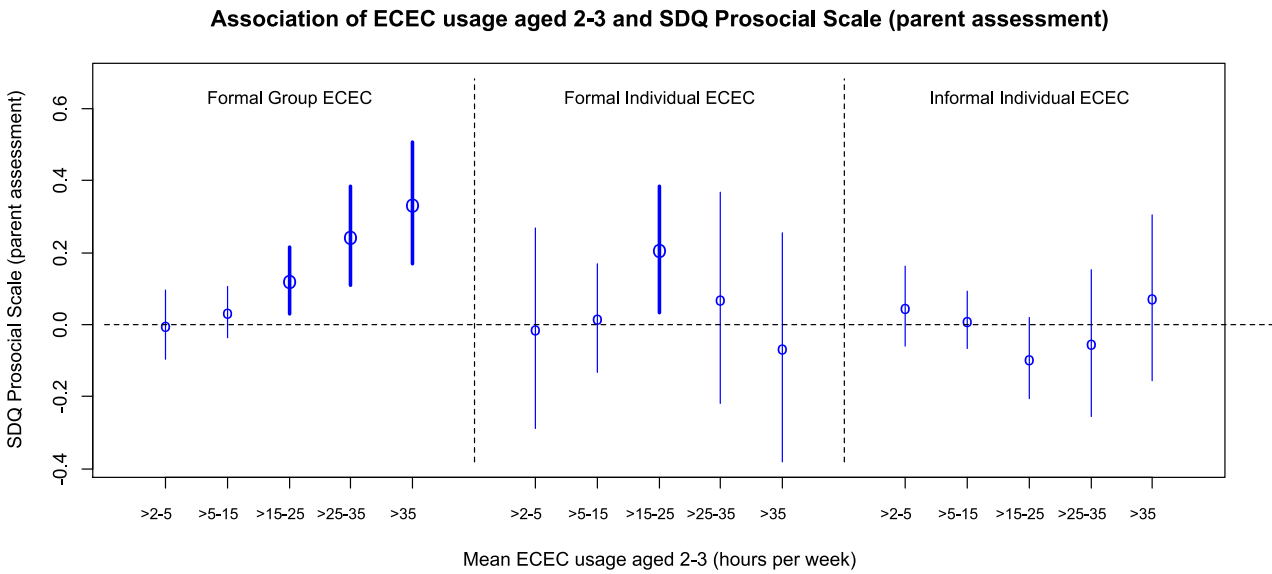
Model is in terms of formal group, formal individual and informal individual ECEC. Point estimates are shown ("O") with vertical lines giving a 95% confidence interval. The horizontal line represents the ≤ 2 hours of ECEC per week reference level. Statistically significant effects are shown in bold.

Figure 31: Association of ECEC and SDQ Peer Problems Scale (parent assessment).
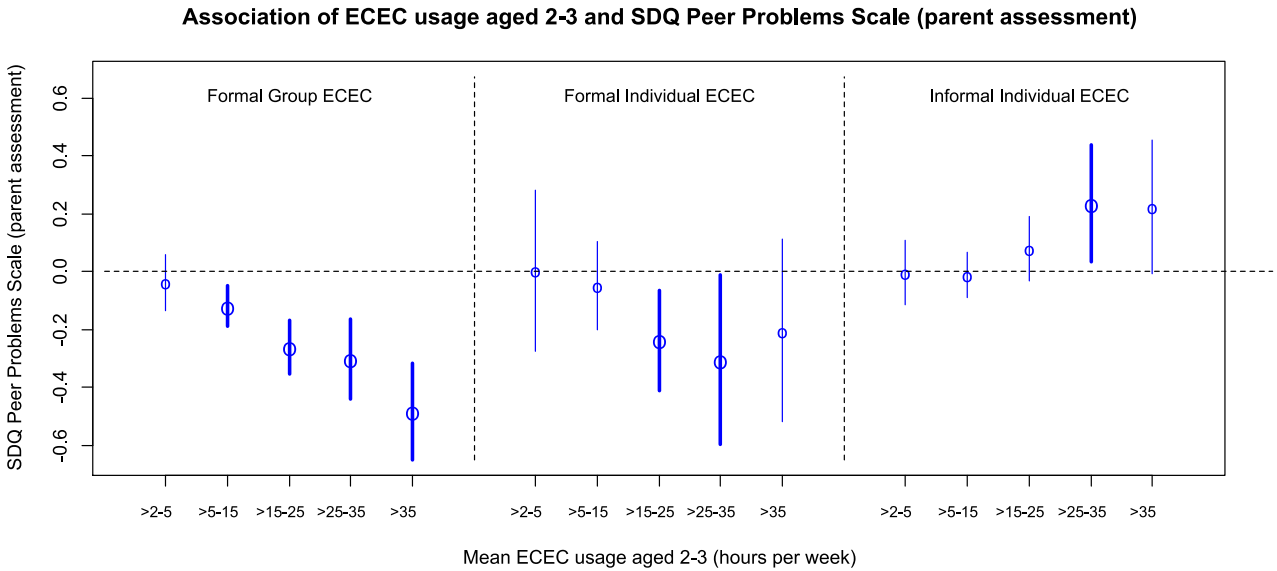
Association of ECEC usage aged 2-3 and SDQ Peer Problems Scale (parent assessment)
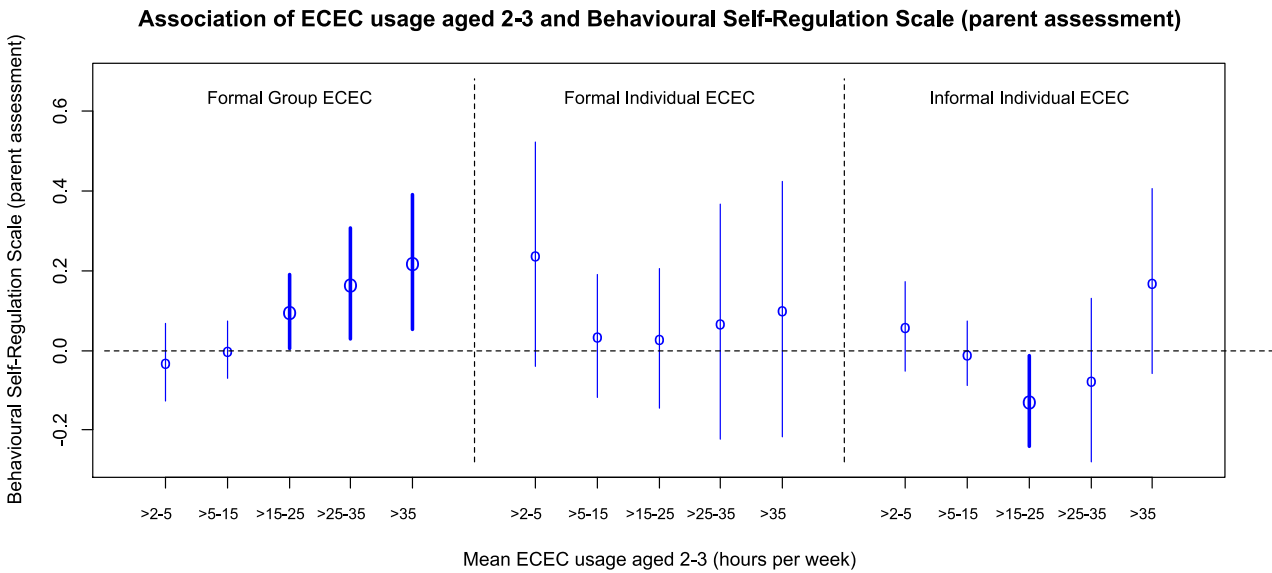
Model is in terms of formal group, formal individual and informal individual ECEC. Point estimates are shown ("O") with vertical lines giving a 95% confidence interval. The horizontal line represents the ≤ 2 hours of ECEC per week reference level. Statistically significant effects are shown in bold.

Figure 32: Association of ECEC and SDQ Prosocial Scale (parent assessment).

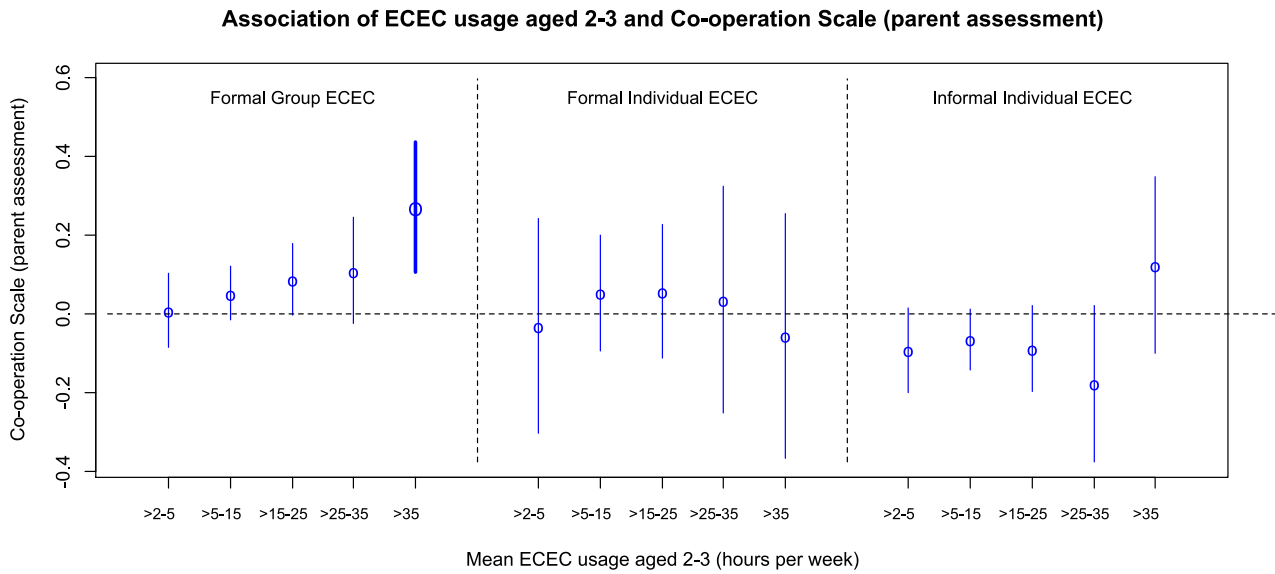Association of ECEC usage aged 2-3 and SDQ Prosocial Scale (parent assessment)

Model is in terms of formal group, formal individual and informal individual ECEC. Point estimates are shown ("O") with vertical lines giving a 95% confidence interval. The horizontal line represents the ≤ 2 hours of ECEC per week reference level. Statistically significant effects are shown in bold.

Figure 33: Association of ECEC and Behavioural Self-regulation Scale (parent assessment).

Association of ECEC usage aged 2-3 and Behavioural Self-Regulation Scale (parent assessment)

Model is in terms of formal group, formal individual and informal individual ECEC. Point estimates are shown ("O") with vertical lines giving a 95% confidence interval. The horizontal line represents the ≤ 2 hours of ECEC per week reference level. Statistically significant effects are shown in bold.

Figure 34: Association of ECEC and Co-operation Scale (parent assessment).

Association of ECEC usage aged 2-3 and Co-operation Scale (parent assessment)

Model is in terms of formal group, formal individual and informal individual ECEC. Point estimates are shown ("O") with vertical lines giving a 95% confidence interval. The horizontal line represents the ≤ 2 hours of ECEC per week reference level. Statistically significant effects are shown in bold.

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

Any enquiries regarding this publication should be sent to us at:
ey.analysisandresearch@education.gov.uk or www.education.gov.uk/contactus

This document is available for download at www.gov.uk/government/publications