Title: The Impact of Pre-school on Adolescents' Outcomes: Evidence from a Recent English Cohort

Author: Patricia Apps Silvia Mendolia Ian Walker

PII: S0272-7757(13)00131-3
DOI: http://dx.doi.org/doi:10.1016/j.econedurev.2013.09.006
Reference: ECOEDU 1447

To appear in: *Economics of Education Review*

Received date: 23-10-2012
Revised date: 10-9-2013
Accepted date: 23-9-2013

Please cite this article as: Apps, P., Mendolia, S., & Walker, I., The Impact of Pre-school on Adolescents' Outcomes: Evidence from a Recent English Cohort, *Economics of Education Review* (2013), http://dx.doi.org/10.1016/j.econedurev.2013.09.006

The Impact of Pre-school on Adolescents' Outcomes: Evidence from a Recent English Cohort

Patricia Apps
Sydney Law School
The University of Sydney
Level 3, Law School Building
Eastern Avenue, Camperdown Campus
NSW 2006, Australia
Email: patricia.apps@sydney.edu.au

Silvia Mendolia
Corresponding author
School of Economics
University of Wollongong
Northfields Avenue
North Wollongong
NSW 2522, Australia
Email: silvia_mendolia@uow.edu.au
Tel. +61 02 42981322

Ian Walker
Department of Economics
Lancaster University Management School
Lancaster University
Bailrigg
Lancaster
LA1 4YX, UK
Email: ian.walker@lancaster.ac.uk

The Impact of Pre-school on Adolescents' Outcomes: Evidence from a Recent English Cohort

1. Introduction

This paper analyses the role of pre-school education on child outcomes in later life. This research question has received increasing attention from economists, and is motivated by the strong interest in childcare and early education policy in many developed countries. Governments have devoted significant amounts of resources towards developing early-childhood policies, with the objective of improving the well-being of children, either through the direct effects of early education on children, or through the impact of childcare on maternal employment and income, or both. Pressure to expand childcare even more widely persists. For example, in the US, there have been recent expansions to HeadStart and to the childcare subsidy for low income working parents, and there as been a recent Presidential commitment to better access. There was intense political debate in Germany in the run up to the 2013 elections. In the UK, recent announcements to provide tax breaks for childcare expenses incurred by working parents follow previous expansions of free time-limited pre-school for 4 and 3 year olds, and implicit subsidies to childcare for parents in receipt of welfare to work programmes.

This paper sheds light on the effects of formal pre-school arrangements, provided in an institutional setting, prior to compulsory education, on various outcomes, both cognitive and non-cognitive. While the importance of the formation of cognitive skills for lifetime welfare has been long established, it is only recently that attention has been drawn to non-cognitive outcomes and aspects of social behaviours. It is now recognized that non-cognitive skills can be shaped in early life and that there is an element of choice in social behaviours; and that both have important long term consequences. We add to the existing literature, by conducting an analysis of a recent and very rich data-set of adolescents up to early adulthood to provide evidence of long-term effects in a wide variety of areas of children's lives.

Our analysis here is based on the Longitudinal Study of Young People in England (LSYPE), a very rich study of a large cohort of children, born in 1990, selected through their schools and interviewed for the first time at age 14-15 in 2004. A great deal of information is collected about the child and her/his family, and seven waves of data are available. Milestone cognitive outcomes are merged into the data from national administrative records.

We consider the effect of attending pre-school (before the child entered primary schooling at around age 5) on a variety of outcomes in adolescence and early adulthood using both conventional regression methods, with and without controlling for a very rich set of child and family characteristics, and matching methods. Ordinary Least Squares, to control for observable confounders, is well known to lead to biased estimates of the causal effects in the face of neglected heterogeneity. It is not clear what direction this bias might be: pre-school users might be better or worse children/parents in unobservable ways. A partial solution to this problem is to attempt to better match users and non-users in observational data. Thus, we adopt Propensity Score Matching to estimate the effects of attending pre-school school on children's outcomes and we compare the magnitude of these effects with those of other important variables, such as maternal education, marital status at birth, month of birth of the child, etc. Propensity Score Matching has been used in various recent papers that investigate the determinants of child well-being in the recent economics literature (see for example Ruhm, 2008; Berger, Hill, & Waldfogel, 2005; and Goodman & Sianesi, 2005).

Our paper contributes to the existing literature in three principal ways. First, we produce new evidence, based on the large and recent cohort of children born in 1990 and followed for seven years from 2004. Second, we take into consideration a wide variety of outcomes, including measures of cognitive and non-cognitive development, and we are able to follow the children in our sample until the age of 21. A major problem with much of the existing literature is that it measures the effect on short term outcomes and there is a fear that such effects might "fade". Part of the contribution of this paper derives from our ability to investigate long-term outcomes. In particular, we investigate the effect of childcare on adolescent outcomes that are effectively permanent (like educational achievement). Such educational outcomes are known to have important effects on lifecycle income. We are also able to look at a wider set of long term outcomes in the same dataset. Further, our study fills a gap between papers based on the 1958 NCDS and 1970 BCS cohorts, such as Goodman & Sianesi (2005), that consider long-term outcomes based on these rather old datasets, and more recent studies, such as the EPPE (Effective Provision of Preschool Education) study, that use recent data but only look at the impact of pre-compulsory education on primary school results. Thirdly, we explore ways to reduce the impact of endogenous selection into pre-school.

Following the most recent literature on the effect of pre-school, we analyse the different impact of this early form of education on children across various markers of

deprivation. In particular, we look at disadvantaged families, where disadvantage is defined in a variety of ways, and we analyse the effect of attending pre-school for them, compared to their advantaged peers. The results broadly support the idea that pre-school prior to compulsory education is particularly beneficial for the cognitive outcomes for children who come from disadvantaged socio-economic backgrounds while the effect on advantaged children, over the various specifications of our model, is largely statistically insignificant. Importantly, we find that non-cognitive outcomes, that seem likely to be the ones that we would normally associate with social benefits, are particularly insensitive to pre-school.

Our estimates imply that the statistically significant benefits are largely confined to private ones associated with cognitive outcomes. Since we are not able to find significant social benefits, the case for subsidising pre-school rests largely on the greater effects on cognitive outcomes for the disadvantaged children. Such a policy would be likely to generate less lifecycle inequality in the future but a cost benefit analysis is beyond the scope of this paper. The case for universal subsidies seems weak.

2. Policy Background

The expansion of pre-school provision in the UK started in the 1970's with the idea of promoting school readiness for children who were considered economically and socially disadvantaged. The number of children attending pre-school nurseries (managed by local government or privately), child-minders, and playgroups increased steadily from the late 1970's. The Children Act 1989 introduced mandatory inspections by the social services departments of local government for all childcare centres and imposed new requirements for the registration and review of private and voluntary day care centres. It prompted increased attention to the contents of the educational curriculum for children under 5 and to the implications for the training for staff working in pre-school settings. The introduction of a system of National Vocational Qualifications (NVQs) in 1991 brought nationally agreed standards across various professions including childcare. Before that, there were many different qualifications in the childcare sector, with very different levels of study ranging from short to 4-year degrees.

The government introduced income tax relief for childcare provided by the employer at the workplace for the first time in the 1990 Budget and, since October 1994, the claimants of Working Tax Credit (an in-work welfare programme) have been entitled to deduct childcare expenditure, up to a limit, from their income which provided an implicit subsidy for

such expenditure. Furthermore, the Conservative government piloted a pre-school education voucher scheme, which offered parents of all 4-year-olds vouchers worth £1,100 to purchase pre-school provision of their choice. The underlying idea of this intervention was to give parents the choice to send their children to maintained, private or voluntary institutions (Audit Commission, 1996 and Brewer, Crawford, & Dearden, 2005). The subsequent Labour government adopted the voucher scheme and subsequently turned these vouchers into direct subsidies for time-limited care. The result was that the proportion of children attending pre-school grew, at the expense of children with child-minders who generally did not receive such subsidies. Thus, by the mid 1990's the shape of early education in the UK had been determined and the changes since have been largely confined to expanding provision rather than changing the nature of what is provided. The proportion of children in childminder arrangements has fallen considerably so childcare arrangements for 3 and 4 year olds now are similar to the pre-school and playgroup provision in the mid 1990's when the cohort used in our analysis were aged 4. Free part-time pre-school provision has been available since 2004 to every child in England and Wales from three years of age, and 95% of eligible children take up this offer (see Department for Education, 2011). In addition the SureStart programme available in deprived areas, while providing a similar pre-school experience to that available elsewhere, also provided parental support and this has had important impacts on the mother and the home environment (NESS, 2010). Most recently, the UK government has advanced proposals for (almost) universal tax deductibility of childcare expenses.[1]

3. Overview of the existing literature

The effect of parental time and home inputs on child development has been widely analysed by psychologists and sociologists (see, for example, McCartney, 1984; Lamb, 1996; and Bianchi, 2000). A substantial body of literature has focused on maternal time *vs*. alternative care time and, to a lesser extent, on the effect of household expenditures and income on children's outcomes. Economists are interested in these topics, especially because some recent literature has showed that long term labour market outcomes, such as wages and employment, that determine lifecycle incomes, largely depend on factors and skills that are already in place by adolescence (see, for example, Cunha, Heckman, Lochner, & Masterov, 2006; and Keane & Wolpin, 2001 and 2010). Extensive research has showed that early cognitive achievements are strong predictors of later educational and labour market

[1] See McLean (2012) for further details on the policy background in the UK and its comparison to the US.

outcomes. For example, Bernal & Keane (2011) show that test scores at ages 4 and 6 are strongly correlated with completed education for the children of US single mothers. Indeed, Heckman & Masterov (2007) make a compelling case for government subsidies for intervention at an early age, for example with high quality childcare, because of the social benefits associated with long term outcomes, such as lower crime.

Although there have been recent important contributions from elsewhere, most of the economics literature on the effect of childcare is based on UK and US cohort data, much of it quite dated. Some recent UK research has relied on the EPPE study of over three thousand children who attended childcare institutions in the late 1990's. EPPE research is available for outcomes up to the age of 14 (see Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2012) but is largely concerned with childcare quality differentials rather than childcare *per se*. Our research is for a cohort only a few years earlier than the EPPE children and so has almost the same relevance for current policy. Moreover, we adopt a matching methodology and we extend the EPPE work to include a wide variety of later outcomes. These include: educational outcomes up to 16 and beyond; a variety of behavioural and attitudinal outcomes up to age 16; and, for some of these outcomes up to age 21

Melhuish (2003) provides an excellent survey of a wide variety of literature for the UK and elsewhere. He focusses, in particular, on experimental and quasi-experimental work. He finds that the existing evidence on childcare (aged 0-2) is equivocal, while the evidence on pre-school education (age 3-4) points to a beneficial effect across the population with notable effects of both months of use and the quality of provision, with larger effects for disadvantaged groups.

Our literature review will mostly focus on prior studies of the effect of pre-school in the UK, as this is the context in which our paper is placed. The majority of this literature has relied on two cohort studies - the 1958 National Child Development Study (NCDS) and the 1970 British Cohort Study (BCS). Most of the recent research based on these analyse the effect of pre-school in the short run (up to attainments in primary school at age 10), although some address long-term outcomes too.[2]

The 1970 BCS was analysed by Osborn & Milbank (1987) who performed an analysis of variance and controlled for a wide set of characteristics, such as social class, family size,

[2] NCDS at age 7 asked for information on the type of pre-school attended and identifies five types of pre-school provision. BCS asked about time in pre-school (hours) for nine types of pre-school. Feinstein et al. (1998) reports that 72% of BCS children attended some form of pre-school provision (48% excluding playgroups).

neighbourhood, gender, mother's age, mental state and employment, type of family, ethnic origin and the presence of handicaps. They found that children with pre-school education have better results in cognitive tests at age 5 and 10 than their peers who did not go to pre-school, with a slightly greater effect for children from disadvantaged backgrounds.

Feinstein, Robertson, & Symons (1998) conducted an analysis of the 1970 BCS and the 1958. They constructed a model of pre-school choice and used the price of pre-school as an instrument for the amount of hours of pre-school. Pre-school was found to have positive effects on cognitive tests up to age 11 (particularly on mathematics skills), which then fade away by age 16 for the 1958 cohort. However, for the 1970 BCS70 children, hours of pre-school were associated with worse social adjustment and reduced vocabulary at 5, worse reading skills at 11 and no effects on maths skills. The authors concluded that "over about a decade (1962–1973), the pre-school experience appears to have ceased to improve test scores in children as they enter secondary school". However, the validity of the instrument seems dubious – the price of childcare is correlated with quality and this is likely to have an effect an independent effect on outcomes.

A subsequent paper by Goodman & Sianesi (2005) analyses the 1958 NCDS looking at the effect of any early education (specifically, early entry into primary school, as well as attendance of private or LEA pre-school schools and playgroups) on a wide range of outcomes, both in the short and in the long run, including cognitive achievements at age 7 through to 16, socialisation, and later outcomes such as wage and employment at age 33. This paper aimed at estimating the total policy effect of early education, using Ordinary Least Squares, OLS with a broad set of interacted variables, and Propensity Score Matching. The authors show a positive effect of pre-school education on test scores, diminishing in size as the children grow, while the effect on socialisation was more mixed. In adulthood, pre-compulsory education was found to increase the probabilities of obtaining qualifications and of being employed at age 33.

The major limitation of all these studies is that they rely on relatively old data, and have only a relatively small percentage of children attending pre-school schools or other formal childcare centres. The early childhood industry has rapidly evolved since the 1960's and many changes have taken place, in terms of pedagogy, teachers' qualifications, and focus on cognitive development. Our own work updates and extends the earlier analyses.

A separate strand of literature has used a rich data-set specifically collected on pre-school-age children known as the Effective Provision of Pre-School Education (EPPE) study. This was launched in 1997 and included a sample of 3,000 children from various socio-economic backgrounds, who attended a range of different pre-schools. The study included only approximately 300 'home' children with no pre-school experience at all. Children were followed until age 7 and pre-school was found to have had a positive impact on cognitive and social development with a particularly positive effect on children from disadvantaged backgrounds. They found a stronger effect for pre-schools with a strong educational focus (see Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2004). At age 11 EPPE children who had been in high quality childcare (as opposed to low) were found to have higher literacy and numeracy levels (around 0.2 and 0.4 of a standard deviation respectively). These results did not seem to vary by parental social background. The EPPE team are continuing to track the children but the effect on long-term outcomes is not yet available. Nonetheless, one influential outcome of the EPPE research has been to show that the estimated effects of high quality childcare centres are larger than more informal forms of childcare and this motivated the UK Department for Education to concentrate resources on high quality centres.

Ruhm & Waldfogel (2013) provides an excellent wider review of the existing literature on the effects of parental and childcare, including studies outside the UK. Identification is clearly an issue and few studies use IV estimation because of the difficulties in finding valid exogenous variation in childcare use. Bernal & Keane (2011) uses the US NLSY data and exploits welfare reforms around the mid 1990's to provide instruments for childcare use to analyse the effect of childcare on cognitive development. This work (see also Bernal & Keane, 2010) shows that formal centre-based early education has positive effects on the children of US single mothers.

Recently, there have been several further studies that have used excellent data and convincing identification strategies. For example, Baker, Gruber & Milligan (2008), looks at the impact of the introduction of childcare subsidies in Quebec and finds significant *adverse* outcomes on a range of outcomes including illness, aggression and social skills. Datta Gupta & Simonsen (2010), uses local variation in waiting lists in Denmark for identification and finds no significant effects on a commonly used non-cognitive index at age 7, but their data contains no test scores. In contrast, Havnes & Mogstad (2011) uses spatial variation in the timing of childcare subsidies in Norway in the mid 1970's and finds strong beneficial effects

on adult outcomes. Bingley & Westergaard-Nielsen (2012) show that pre-school attendance is positively associated with completed schooling (especially for disadvantaged children) and earnings at age 22-30 in Denmark. Dumas & Lefranc (2012) analyze a large-scale expansion of pre-school enrollment in France and show that pre-school attendance is particularly beneficial for children from low socio-economic background.

Finally, there is a literature on the effect of specific programs that target children from disadvantaged socio-economic backgrounds, such as the Perry Pre-School program in the US (see Currie, 2001; Carneiro & Heckman, 2003; Blau & Currie, 2006; and Heckman, Moon, Pinto, Savelyev, & Yavitz, 2010). A few recent papers examine the effect of such pre-school programs on school readiness and behavioural problems (see Loeb, Bridges, Bassok, Fuller & Rumberger, 2007; and Magnuson, Ruhm, & Waldfogel, 2007). Perry combined pre-school for deprived 3 year olds, at about the same intensity as is now freely provided in the UK, with intensive home support. The Heckman et al. (2010) re-evaluation of the Perry data suggests more modest, but still significantly positive, effects than was claimed in earlier work and crime effects are a major component of the return.

The heterogeneity in findings is, perhaps, not surprising because of the variety of treatments and the variation in the populations they are applied to. Thus, it is too early to say that there is a consensus – although most of the few studies that do investigate long-term non-cognitive outcomes do seem to find beneficial effects, at least for disadvantaged children.

4. Data

Our work uses data from the Longitudinal Study of Young People in England (LSYPE), which is a large scale panel survey of English adolescents, interviewed for the first time when they were in school year 9 in 2004 at the age of 14-15[3]. The questionnaires cover a variety of topics, including academic achievements, family relationships, attitudes toward school, family and labour market, and some sensitive or challenging issues, such as risky health behaviours (smoking, alcohol drinking, drug taking), personal relationships, etc.

In the first wave, 15,500 young people were interviewed who were selected in a two stage sampling to provide sufficient representation of young people in England by oversampling deprived schools, and then over-sampling the major ethnic minorities. In the first four waves, parents/guardians were also interviewed. The wave 1 response rate was 74%

[3] Note that grade repetition is very rare in the British school system.

and subsequent response rates were much higher. In addition, LSYPE can be linked to the National Pupil Database (NPD), a pupil-level administrative database that matches pupil and school characteristics data to pupil attainment and contains detailed information on test scores for all pupils in England and Wales including the LSYPE chldren. Retrospective information about the LSYPE child and the family was asked of the parent at waves 2 and 3 of the survey and we use this to determine whether the child attended pre-school, and to provide information on the child's and family's situation at birth. Our final sample includes around 11,000 observations of children with non-missing information on test scores, early education and other essential information on the child's birth and family background.[4]

4.1 Outcomes

We are interested in analysing the impact of pre-school education on a variety of outcomes. The LSYPE data includes information on various outcomes, measured from adolescence to early adulthood. This allows us to attain a more complete picture of the skills and behaviours that are affected by early education than previous studies. Table 1 lists the outcomes we use in our analysis and the age of the child for each outcome. We focus on five groups of outcomes:

Cognitive development and test scores. We analyse data from the NPD on children's results in tests and particularly: Key Stage 2 (age 11) of the National Curriculum, Key Stage 3 (age 14) and Key Stage 4 (also known as the General Certificate of Secondary Education, GCSE) at age 16. Key Stage 2 consists of national curriculum tests in English (reading, and writing, including handwriting and spelling) and mathematics, together with teacher assessments in English, mathematics and science. Key Stage 3 consists of teacher assessments only, in all National Curriculum subjects. Schools have to submit scores for English, maths and science. At the end of Key Stage 4, pupils generally take the national public examinations known as GCSE in most subjects studied – often in as many as 10 subjects. GCSE grades range from A* to G. The dependent variables in our analysis are: Key Stage 2 and Key Stage 3 scores (in aggregate and in single subjects); the number of subjects with "pass" grades from A* to C in GCSE exams; a binary variable indicating having 5 GCSE passes including maths and English which is usually required for students following an academic track for progression beyond age 16; and two binary variables equal to 1 if the child attained A* to C in English and Mathematics. A-levels are taken at age 18 and are the main admission criterion for

[4] The selected observations were not significantly different from the original data in terms of their observable characteristics.

university admission. The A-level points score measures attainment at age 18 – admission to elite institutions, or to high demand courses, usually requires 360 points or more, while the less prestigious institutions and less popular subjects will typically demand around 240 points.[5]

Attitudinal variables and socialisation. Recent literature has showed that personality traits may have an important role in determining the future success of individuals over and above test scores (see Heckman et al., 2010). Therefore, we extended our analysis to take into consideration some additional outcomes such as: the intensity of effort at school; whether the child likes her/his school and teachers; the number of close friends the child has; and the psychological well-being of the child (measured by General Health Questionnaire score[6]).

Economic activity and education at age 20-21: intentions to attend university, attending university, being employed, and NEET (not being employed, or in education, or in training).

Health risky behaviours such as smoking, drinking, use of cannabis, and early pregnancy.

Problematic behaviours such as being involved in fighting, being suspended from school, vandalism, shoplifting, and being in contact with the police.

4.2 Pre-school

We are interested in investigating the effect of pre-school on various children outcomes, including cognitive development. In order to address these research questions, we use the information recorded in the LSYPE history data file where the parents are asked whether the child went to pre-school school. Unfortunately, we don't have any information on how many days/hours the child spent in pre-school school. Nor do we have any information about its quality. We are clearly estimating an effect of the average exposure. Around 80% of the children in the estimation sample record having attended pre-school.

[5] We do not consider the complex array of vocational training courses that less academic students can take from 16 to 18.

[6] The GHQ Caseness score is constructed from the responses to 12 questions covering feelings of strain, depression, inability to cope, anxiety-based insomnia and lack of confidence. The twelve answers are combined into a total GHQ score that indicates the level of mental distress, giving a scale running from 0 (the least distressed) to 12 (the most distressed)

Table 1 Outcomes

Cognitive development	***Test scores***
Cognitive development at 11	KS2 Score (points, divided by standard deviation)
Cognitive development at 14	KS3 Score (points, divided by standard deviation)
Cognitive development at 16	GCSE – N subjects pupil achieved A*-C
5+ GCSE incl English/Maths	=1 if at least 5 GCSE A*-C incl English and Maths
Language skills at 11	KS2 Score in English (points)
Maths skills at 11	KS2 Score in Maths (points)
Science skills at 11	KS2 Score in Science (points)
Language skills at 14	KS3 Score in English (points)
Maths skills at 14	KS3 Score in Maths (points)
Science skills at 14	KS3 Score in Science (points)
Language skills at 16	=1 if pupil has level A*-C in GCSE English
Maths skills at 16	=1 if pupil has level A*-C in GCSE Maths
Has any post-GCSE qualifications	= 1 if pupil has further qualifications after GCSE
Has A levels	= 1 if pupil has A levels
Attitudinal variable, free time and socialisation	
Effort in school work	=1 if child agrees with the following statement: At school I work as hard as I can (wave 1)
Likes school	=1 if child agrees with the following statement: On the whole, I like being at school (wave 1)
Likes teachers	=1 if the child says: She/he likes all/most teachers (wave 1)
N. of friends	of close friends the child has (only wave 6 or 7)
Psychological well-being	General Health Questionnaire (from 0=not distressed to 12 = most distressed) recorded (wave 2)
Free time with friends	=1 if child declares he mainly spends his free time with friends (rather than with family or alone)
Reading often	=1 if child reads for pleasure at least once a week (wave 1)
Sports often	=1 if child plays sports at least once a week (at wave 1)
TV	=1 if child watches TV 4+ hours during a school day
Bullied	=1 if the child has ever been bullied (up to wave 3)
Economic activity and education	
Tertiary education intentions	=1 if the child has applied to university in wave 4 (age 17)
Education at 20-21	=1 if child attending university or other course (at wave 6/7)
Working at 20-21	=1 if the child is working (wave 6/7)
NEET	=1 if child not in employment, education, training (wave 6/7)
Health risky behaviours	
Smoking	=1 if child ever smokes cigarettes
Smoking cannabis	=1 if child ever tried cannabis
Pregnant	=1 if child has ever been pregnant
Early first sexual intercourse	=1 if younger than 16 at first sexual intercourse
Frequent drinking	=1 if drinks at least once a week (wave 3)
Problematic behaviours	
Suspended from school	=1 if child has ever been suspended from school (wave 1)
Vandalism, Shoplifting	=1 if child has ever shop lifted, vandalised, or graffitied
Police	=1 if police ever in contact because of the child's behaviour

There is no a clear definition of pre-school in the LSYPE data, but official data from the British government (see Department of Education and Employment, 1999a and 1999b) show that day-care facilities in England in the 1990's included day nurseries (independent or run by Local Education Authorities), playgroups, child-minders, after school clubs and holiday schemes and that a vast majority of children were in some form of early years education provision (around 95% of four years old population, according to Department of Education and Employment, 1999a and 1999b). We believe that the self-reported "pre-school" is likely to include various forms of care in centre-based institutions, such as day care centres, local authority and independent nurseries, and play-groups. Some of these will provide a formal curriculum, but most will focus on play. Some will provide full time care, most will be for just for three hours most weekdays and most, but not all, will be run by professionally trained staff with little reliance on parental help. Local Authority and voluntary day nurseries were targeted on disadvantaged or 'at risk' groups whereas private day nurseries were more likely to be used by relatively advantaged families, usually with two incomes. That is, our treatment covers a wide variety of arrangements but excludes child-minder arrangements that have become less and less common over time as centre-based care became increasingly subsidised.

Children in LSYPE were born in 1989-1990, so they are likely to have attended pre-school schools between 1992 and 1995. Before 1997, there were no requirements for Local Authority in terms of educational provision for children under compulsory school age, and the decision of whether to provide free pre-school places (and if so how many to provide) was left to each individual Local Authority, leading to substantial variation in provision. According to Brewer et al. (2005) provision across the country ranged from zero free pre-compulsory education places provided, to a high of 27.5 places per 100 children and this variation in access to pre-school education persisted into the 1990's (see Dickson, 2008 for a discussion of changes to the provision of pre-school places in the late 1990s).

4.3 Other explanatory variables

There is extensive information available in LSYPE. As we will discuss in greater detail in section 4, we try to capture all factors that determine early education attendance and child outcomes. All of the variables we control for are, arguably, pre-determined variables – that is, not themselves influenced by pre-school education. Inputs in children's outcomes

include individual mental and physical endowments, parental and family inputs (such as income, time, size of the family and number of siblings), and local area characteristics. Our first, most parsimonious, model only includes at-birth characteristics such as: birth-weight; whether the child was premature; ethnic background; sex of the child; month of birth; and family characteristics such as marital status and age of the mother at birth. In the second model we include other family's characteristics (measured at wave 1, but which are unlikely to have changed since the child's birth) such as: main language of the family, maternal education; child's and mother's disability; grandparents' education and older siblings. Table 2 lists the explanatory variables used in the empirical model.

We explore the possibility of heterogeneity in the effect of pre-school on children by socio-economic background and by various indicators of socio-economic disadvantage. We follow Ruhm (2008) and construct a multivariate indicator of socioeconomic status by regressing total family income on mother's age at birth, education, and marital status. Youths are classified as "advantaged" ("disadvantaged") if they live in households that are above (below) the median prediction. This SES index simultaneously accounts for a larger number of determinants than simple income and possibly reduces the endogeneity problem. Secondly, we also divide youths by maternal marital status at birth, maternal employment when the child was 5, and single parenthood at age 5.

Table 2 Control variables

Model 1	
Child	
Pre-school	=1 if main parent reports young person attended pre-school
Birth-weight	In kg
Month of birth	Omitted: September
Premature birth	=1 if the child at least 3 weeks early
Sex of the child	Boy=1
Ethnic background:	White (omitted), black, Asian, mixed
Mother	
Young mother	=1 if mother was <21 year old at child's birth
Single mother	=1 if mother was not married at child's birth
Model 2 – above, plus	
Child	
Child's disability	=1 if child has disability/long standing illness
Mother	
Maternal education:	Degree (omitted); Higher education (not degree); Junior high school graduate (GCSE A*-C); No qualifications
Family	
English	=1 if English main language spoken in household

Older siblings	Number of older siblings
Grandparents' education	=1 if main parent's parent went to university
Main parent's disability	=1 if main parent disability/long standing illness

4.4 Descriptive statistics

Table 3 presents the distribution of outcomes, split by pre-school school attendance. 82% of children in the sample recorded attending pre-school – this was far higher than the maternal labour force participation rate at the time. Official data on early year childcare arrangements did not begin until 1999 and the earliest available Early Years Census for 1999 reports that 98% of all 4 year old children were "in early years provision".

Table 3 Distribution of outcome variables

	Whole sample	**Attended pre-school**	**Not attended pre-school**	**Test of difference in means (p-value)**
Attended pre-school %	82	1	0	
KS2 points - Average	27.2 (3.9)	27.3 (3.8)	26.6 (4.1)	0.000
KS3 points - Average	34.4 (6.5)	34.6 (6.4)	33.2 (6.8)	0.000
KS2 English	26.7 (4.2)	26.9 (4.2)	26.0 (4.5)	0.000
KS2 Maths	26.7 (4.8)	26.9 (4.7)	26.2 (4.9)	0.000
KS2 Science	28.4 (3.6)	28.5 (3.6)	27.9 (3.7)	0.000
KS3 English	33.6 (6.0)	33.82 (6.0)	32.7 (6.2)	0.000
KS3 Maths	36.1 (7.8)	36.36 (7.7)	35.0 (8.3)	0.000
KS3 Science	33.5 (6.5)	33.73 (6.5)	32.5 (6.9)	0.000
GCSE A*-C in English(%)	62.3	63.6	54.9	0.000
# GCSE subject grade A*-C	6.1 (4.2)	6.2 (4.2)	5.5 (4.3)	0.000
5+ GCSE incl. English and Maths	50	51	43	0.000
GCSE A*-C in Maths (%)	57.2	58.2	51.6	0.000
Effort in school work (%)	81.5	81.8	80.2	0.091
Has any post GCSE Qualification	41	42	37	0.000
Has A level	33	34	29	0.000
Like school (%)	84.8	85.2	82.8	0.001
Like teachers (%)	42.6	43.1	39.8	0.001
Psychological wellbeing GHQ	1.67 (2.51)	1.68 (2.51)	1.63 (2.55)	0.090
Applied to university age 17 (%)	63.1	63.8	59	0.527
University/Education age 19-20 (%)	56.8	51.9	51.3	0.000
Working at age 19-20	39.2	36	34.2	0.002
NEET at age 19-20	8.8	8.4	11	0.000
Ever Smoking (%)	9.6	9.2	11.8	0.000
Ever tried cannabis (%)	8.6	8.5	9.4	0.000
Frequent drinking (%)	26	26	25	0.490
<16 at first sexual intercourse (%)	31	31	31	0.830
Ever been pregnant (%)	4.5	4.4	5.3	0.000
Suspended (%)	9.4	9.1	11.0	0.000
Problematic behaviour	22.2	22.0	23.6	0.170

Note: Test scores are recorded at the appropriate age. All the other variables are measured at wave 1, unless differently specified. Problem behaviour examples are vandalism, shoplifting, graffiti, and having been in

contact with the police. NEET= not in employment, education or training. Standard deviations are in brackets. KS scores are translated from an achievement "level" and the conversion to KS points is explained in http://www.education.gov.uk/schools/performance/secondary_11/PointsScoreAllocation2011.pdf

This will be an overestimate relative to LSYPE for several reasons: the EYC figure is for 4 year olds only, 56% of the EYC figure is in infant classes in schools (which may be perceived from LSYPE parents as "going to school" rather than "pre-school"), and there was a rise in childcare use between 1994 and 1999.

Children who went to pre-school seem more likely to be happy with their secondary school and teachers and likely to put more effort into school work. It seems plausible that this would translate into higher test scores. On average, children who went to pre-school perform better in test-scores up to age 16 than did those who did not receive early education. These differences do not seem to fade in later scores, at least up to age 16. The differences in GCSE results is similar in size to the difference in KS2 and KS3 scores and are particularly pronounced in performance in English (rather than Maths or Science). However, we are not able to find any significant effects on A-level performance at 18. Children who went to pre-school seem less likely to be out of education or employment when they reach the age of 20-21 and less likely to engage in health-risky behaviours, such as smoking, trying cannabis, early pregnancy.

5. Estimation

We begin our analysis by looking at cognitive achievements and test scores at age 11, 14 and 16 and we further exploit the richness of our dataset to investigate test results in various subjects. We then examine children's intentions to apply for university at age 17, subsequent attendance at university, and economic activity at age 19-20. We look into school outcomes in more detail, by exploiting information available on children's attitudes and efforts in school work, psychological well-being, and general happiness of the child at school. We also look at some health related behaviours between 14 and 20, such as smoking, use of cannabis, and teenage pregnancy.

Estimating the causal effect of pre-school on children's outcomes raises the missing counterfactual problem. The evaluation problem is to provide unbiased estimates of the average counterfactual using appropriate methods and assumptions. Like many other studies, we provide an upper bound by estimating a linear regression to examine the correlation

between the children's' outcomes and pre-school education. The linear model can be written as:

$$C_i = \alpha + \beta_1 N_i + \beta_2 X_i + \varepsilon_i$$

where C_i represents a particular outcome, N_i is a binary variable equal to 1 if the child attended pre-school school and X_i is a vector of characteristics. We present estimates from linear probability models rather than nonlinear probit which produce similar marginal effects.

The major challenge for such analyses is that of establishing causal connections between pre-school education and child outcomes, given that children who went to pre-school may have unobserved characteristics that also affect their cognitive development and other outcomes. Furthermore, mothers who decided to send their child to pre-school may be systematically different from those who did not, and their child's cognitive ability can itself influence mothers' decisions. The possibility of self-selection suggests that we should think of OLS as providing an upper bound to the causal effects.

Nonetheless, the effect of pre-school on children's outcomes may be estimated correctly through OLS if several assumptions hold. First, the "selection on observables" assumption must be satisfied (see Heckman, 1979). This means that all variables that predict both pre-school attendance and children's outcomes should be included in our model. These variables are sometimes called "confounding variables" because if they are not appropriately controlled for, their effect on the outcomes is confounded with the effect of the causing variable of interest (see Angrist & Krueger, 1999). Second, the model must be correctly specified: this is problematic, because functional form assumptions are difficult to test when we include many independent variables. Further, if there is a lack of overlap in covariate distributions across children who went and did not go to pre-school then linear regression models extrapolate results over unsupported portions of the covariate distribution. That is, OLS attempts to compare incomparable children.

We cannot, in this data, address the selection on unobservables problem. There is simply no quasi-experimental variation across our sample to exploit. However, we can go some way towards addressing the other problems. Firstly, we try to lower the upper bound provided by OLS estimation, through the inclusion of a more detailed set of independent variables. Second, we exploit propensity score matching. This does not rely on functional form assumptions and restricts inference to samples where we can find overlap in the

distribution of covariates across the treatment. In this context, we compare a group of children who did not go to pre-school school that looks as similar as possible to our treatment group (children who did attend pre-school school). Caliendo & Kopeinig (2008) is a recent clear exposition of the issues. The validity depends on the assumption of "unconfoundedness" (Rosenbaum & Rubin, 1983) – also called "selection on observables" (Heckman & Robb, 1985) or the "conditional independence assumption" (CIA) (Lechner, 1999). A major development in the matching literature was Rosenbaun & Rubin (1983) who introduced the concept of propensity score matching to solve the problem of the possible high dimensionality in the vector of observed characteristics *X*. Rosenbaum & Rubin (1983) suggest using the balancing scores. They show that if the outcomes are independent of treatment conditional on covariates *X*, they are also independent of treatment conditional on a balancing score *b(X)*. The propensity score $P(D_i = 1|\mathbf{X}_i, u_i) = P(\mathbf{X}_i, u_i)$, i.e. the probability for an individual to participate in a treatment given his observed covariates **X**, is one possible balancing score. Hence, if the unconfoundness assumption holds, all biases due to observable components can be removed by conditioning on the propensity score (Imbens, 2004).

Specifically, the conditional probability of going to pre-school for each child, given our covariates, is estimated. We follow the overwhelming majority of the literature on propensity score matching and use a discrete choice model (probit) to estimate this propensity score (see for example Caliendo & Kopeinig, 2008). Recent literature has suggested the possibility of using semi-parametric techniques to estimate the propensity score but these methods are computationally very demanding and their advantages are debatable – for example, Zhao (2008) shows that poorly estimated propensity scores have little influence on the estimates of the treatment effect and propensity; and Leher & Kordas (2013) suggest that semiparametric estimation of the propensity score yields large benefits only when the error distribution is highly asymmetric and the treatment effects do not vary in a monotonic manner with the true propensity score (see also Heckman, Ichimura, & Todd, 1998 for a discussion).

The estimated propensity scores can then be used to create a matched control group in various ways. Here, we match to the nearest neighbour by finding the comparison untreated sample member with the closest propensity score for each treated sample member. Unmatched individuals are dropped from the analysis. This can be conducted using various parametric and nonparametric techniques. Figure 1 shows a histogram of the propensity scores while Figure 2 shows a kernel density estimate of propensity scores for treatment and

control group. Both graphs are based on Model 2 estimates and show that there is extensive overlap between the treatment and control groups.

Figure 1 - Histogram of propensity scores of treatment vs. control group

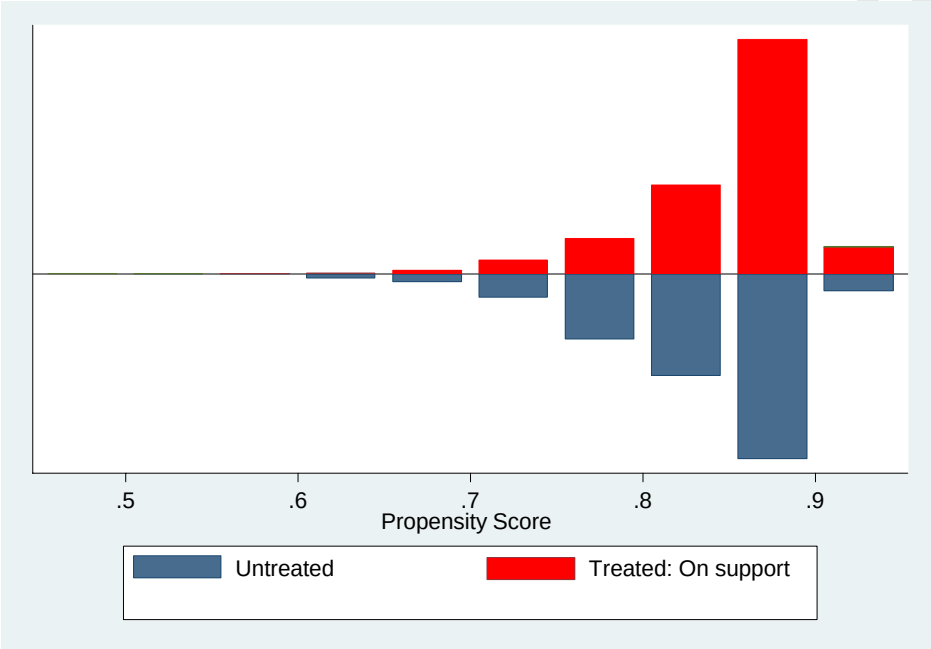
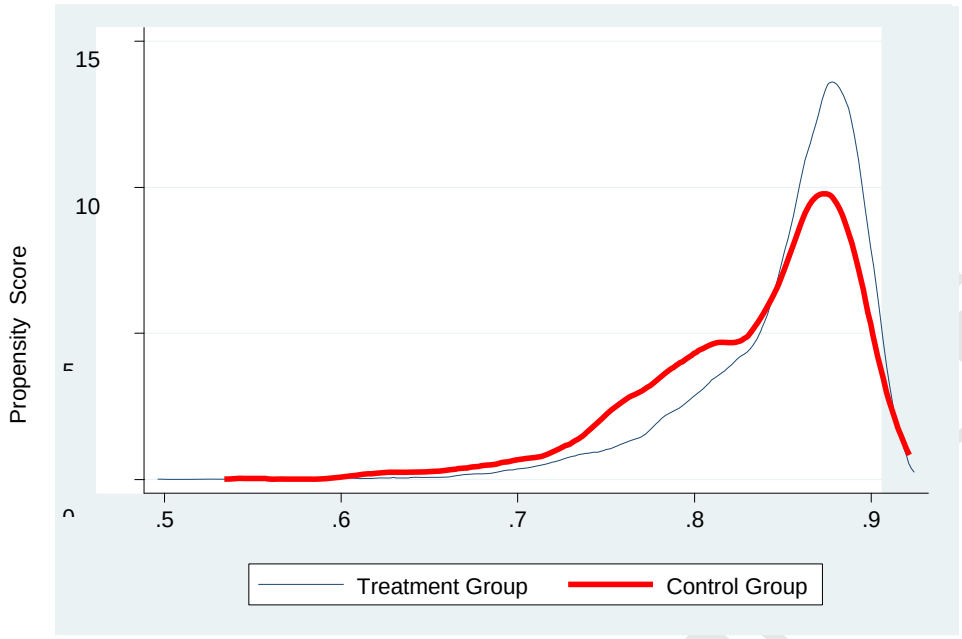


Figure 2 - Kernel graphs of propensity score for treated and control group



Matching is more robust than OLS because it does not restrict the way in which pre-school school affects child outcomes to be linear, and inference is limited to samples that are effectively comparable, based on the covariates distribution. Matching attaches appropriate weights to the observations in the control group, so that the distribution of their observable characteristics is realigned to the treatment group. Our analysis is performed using the STATA routine *psmatch2*, performing propensity score matching with the nearest neighbour method with replacement.[7] Off-support observations (only around 0.01% of the sample) have been dropped from the analysis. We follow Leuven & Sianesi (2003) in reporting approximate standard errors on the treatment effects and assuming independent observations, fixed weights, homoskedasticity of the outcome variable within the treated and within the control groups, and that the variance of the outcome does not depend on the propensity score.

However, in order to test the stability of our results with respect to the matching methodology, we run a series of sensitivity tests, using different matching techniques (Heckman et al., 1998). In particular: we provide estimates of our model bootstrapping the standard errors from *psmatch2*; we use nearest neighbor (Mahalanobis-metric) matching on the **X** with analytical standard errors as in Abadie and Imbens (2006); we use Radius matching (caliper equal 0.001); we use Kernel nonparametric matching (see Heckman et al., 1998); and we use Inverse-propensity score weighting as in Imbens and Wooldridge (2009). The results from these estimations are reported in Appendix Table A2. In all cases the results are very similar to those presented in the body of the paper. Finally, appropriate tests have been run, in order to compare covariate distribution across our matched groups to ensure that adequate balance has been obtained and to verify the sensitivity of matching estimates to unobserved heterogeneity. The results from the balancing tests are not presented for reasons of parsimony but are available on request.

We estimate the causal effect of pre-school education on cognitive and non-cognitive outcomes for those children who attended pre-school school, the so-called Average Effect of the Treatment on the Treated (ATT). We also estimate the Average Effect of the Treatment on the Untreated (ATU); and we think of the latter as being useful for policymakers considering extending the treatment to untreated individuals.

[7] The propensity score matching with replacement has been shown to reduce bias relative to matching without replacement (see Dehejia and Wahba, 2002 and Goodman and Sianesi, 2005).

6. Results

The results from the estimation of the effect of pre-school on child outcomes are presented in Tables 4 to 7. Table 4 presents results on the effect of pre-school education on cognitive development and test scores for the whole sample. Pre-school generally has a moderately positive effect on test scores at age 11, 14 and 16. We analyse the effect on pre-school on KS2 and KS3 scores (age 11 and 14) and on the probability of getting at least 5 GCSE with a grade between A* and C, including English and Maths (age 16). Therefore the results at age 11 and 14 are not directly comparable with those at age 16, but this GCSE indicator is commonly used in the UK to evaluate school performance, because receiving five or more A*–C grades at age 16, including English and Maths, is often a requirement for taking A-levels in post-compulsory schooling. KS2 and KS3 results are directly interpretable as effects relative to a standard deviation.

The effects are found both for average test scores and separately for maths and language skills and results from OLS and PSM estimation are very similar. Interestingly, the magnitude of the effect is higher for language skills (between 8 and 10 per cent of a standard deviation in Model 2) rather than for mathematics or science skills in KS2 and KS3 scores (around 4-5 per cent of a standard deviation). The size of these effects is comparable with that of some other important characteristics, such as birth-weight, sex of the child, number of older siblings or parental disability. The effects on A-level and other post-GCSE qualifications are similar in model 1 but smaller and less significant in model 2. The effect of pre-school school for children who received early education (ATT) and the effect for those who did not go to pre-school (ATU) are generally quite similar in size and significance, showing that the moderate beneficial effect of going to pre-school school is comparable to the gains that could be had from expanding provision to those not attending pre-school. Predictably, including more covariates, as in Model 2, results in smaller estimates and we think of this specification as driving out some selection so as to tighten the upper bound.

The positive effect of pre-school education can be mediated through a variety of factors, including early exposure to literacy and numeracy, and socialisation. Moreover, the children may build up independence and self-confidence and get used to school routines, and this is an advantage that enhances their learning when they are in formal education.

Appendix Table A1 presents ATT results for the effect of other independent variables in the PSM Model 2 on KS2, KS3 and Number of GCSE subjects with grade A*-C – and so

completes Table 4. There is a strong education gradient so that children with more educated mothers (or with grandparents who went to university) are more likely to have higher test scores. Birth-weight has a positive effect; disability a negative one; being the child of a teen mother has a negative effect similar in magnitude to the effect of living in a lone parent household; there are very important month of birth effects that show that summer born do less well; and being black is associated with lower attainment, while being Asian is associated with higher attainment relative to white. All of these results are consistent with findings elsewhere and, while important, are not the focus of this paper.

Appendix Table A2 reports results from estimation performed using alternative matching techniques for Model 1 and Model 2. The results are generally stable across the various specifications and because we think of Model 2 as providing a tighter upper bound, hereafter we focus only on Model 2.

The Model 2 results are then split by sex of the child in Table 5. Pre-school attendance seems to be beneficial for girls, while the effects on boys' test scores are generally not significantly different from zero. Girls who went to pre-school school significantly score better in all tests at different ages and in all subjects. This result is consistent with Feinstein et al. (1998) who show that gains from pre-school are less for boys than girls and negative effect on social adjustments are negligible for girls, and with Goodman & Sianesi (2005), who show that long-term positive effects of pre-school (i.e. chances of getting a high-degree or wages at 33) are more pronounced for girls. This result is also quite common in the international literature (see for example Anderson, 2008; Berlinski, Galiani, & Gertler, 2009; Havnes & Mogstad, 2011) and may be due to a stronger effect from improved language skills (usually higher in girls), combined with the lower impact of negative behaviours (such as aggressiveness, and anti-social behaviours) which are more common in boys.

In Table 6, we analyse the effect of pre-school education by socio-economic status and gender, using various definitions of socio-economic disadvantage. Pre-school education is more beneficial for children coming from disadvantaged socio-economic backgrounds and positive effects are more pronounced for girls than boys. Key Stage 2 and Key Stage 3 average scores increase by more than 20 per cent of a standard deviation for disadvantaged girls who went to pre-school school and girls from single mothers' households. Children with an unemployed mother also benefit from pre-school attendance (the chance of getting 5 or more GCSE with grade A*-C improve by around 5 percentage points). These results are stable across subjects and estimation technique. The chances of getting at least 5 GCSE A*-C

increases by around 6 and 9 percentage points for disadvantaged boys and girls. The size and significance of the effects decrease when we look at post-GCSE qualification and A levels results. However, the chances of applying and actually attending university are also positively affected by pre-school attendance for girls who come from a disadvantaged socio-economic background.

Table 4 Effect of pre-school on cognitive development (Whole sample)

	Model 1			N. observ.	Model 2			N. observ.
	OLS	**PSM – ATT**	**PSM – ATU**	**Treated/ Control**	**OLS**	**PSM- ATT**	**PSM – ATU**	**Treated/ Control**
Age 11								
KS2 Overall	0.148	0.153	0.138	9,462	0.070	0.071	0.074	9,381
	(0.025)**	(0.038)**	(0.040)**	1,672	(0.024)*	(0.036)+	(0.038)+	1,652
KS2 Maths	0.110	0.113	0.089	9,372	0.046	0.051	0.032	9,608
	(0.025)**	(0.037)**	(0.040)*	1,656	(0.025)+	(0.037)	(0.038)	1,637
KS2 English	0.166	0.130	0.170	9,387	0.086	0.099	0.069	9,310
	(0.025)**	(0.038)**	(0.040)*	1,661	(0.024)**	(0.037)**	(0.039)+	1,642
KS2 Science	0.119	0.131	0.079	9,377	0.050	0.032	0.072	9,299
	(0.025)**	(0.037)**	(0.040)*	1,650	(0.024)*	(0.036)	(0.038)+	1,631
Age 14								
KS3 Overall	0.166	0.160	0.150	9,456	0.081	0.052	0.084	9,366
	(0.025)**	(0.037)**	(0.039)**	1,696	(0.023)**	(0.037)	(0.038)*	1,675
KS3 Maths	0.036	0.094	0.102	9,341	0.060	0.049	0.050	9,254
	(0.026)**	(0.038)**	(0.040)*	1,666	(0.024)*	(0.037)	(0.038)	1,645
KS3 English	0.152	0.076	0.114	9,213	0.071	0.074	0.052	9,154
	(0.002)**	(0.037)**	(0.040)**	1,635	(0.024)**	(0.037)*	(0.039)	1,614
KS3 Science	0.141	0.150	0.085	9,263	0.061	0.032	0.078	9,158
	(0.026)**	(0.037)**	(0.040)*	1,644	(0.024)*	(0.037)	(0.039)*	1,624
Treated/Control								
Age 16								
# GCSE at A*-C	0.627	0.678	0.746	9,813	0.298	0.346	0.189	9,730
	(0.110)**	(0.150)**	(0.16)**	1,758	(0.091)**	(0.150)**	(0.150)	1,737
5+ GCSE A*-C incl	0.0706	0.075	0.074	9,813	0.036	0.062	0.0402	9,730
English and Maths	(0.013)**	(0.017)**	(0.018)**	1,758	(0.012)**	(0.017)**	(0.018)**	1,737
A*-C in English	0.083	0.093	0.096	9,813	0.047	0.052	0.043	9,730
	(0.010)**	(0.017)**	(0.018)**	1,758	(0.011)**	(0.020)**	(0.018)**	1,737
A*-C in Maths	0.060	0.046	0.073	9,813	0.018	0.041	0.031	9,730
	(0.010)**	(0.017)**	(0.018)**	1,758	(0.010)**	(0.021)**	(0.018)**	1,737
Age 18+								
Has any Post-GCSE	0.046	0.047	0.0507	10,050	0.018	0.0324	0.028	9,956
qualification	(0.012)**	(0.016)**	(0.018)**	1,813	(0.012)	(0.016)**	(0.017)	1,792
Has A levels	0.024	0.022	0.033	10,050	0.0009	0.017	0.019	9,956
	(0.011)*	(0.015)	(0.016)*	1,813	(0.010)	(0.014)	(0.015)	1,792
Has applied to	0.056	0.047	0.064	7,558	0.028	0.058	0.045	7,301
university	(0.014)**	(0.019)**	(0.021)**	1,286	(0.013)*	(0.019)**	(0.021)**	1,217
Is at university	0.027	0.037	0.032	6,564	-0.001	-0.009	0.013	6,504
	(0.015)+	(0.029)+	(0.023)**	1,108	(0.015)	(0.021)	(0.022)	1,093

Notes: Standard errors are in brackets + indicates that the underlying coefficient is significant at 10% level, * at 5% and **at 1%.
KS2 and KS3 results are interpretable as % of a standard deviation.

Table 5 Effect of pre-school on cognitive development by gender: Model 2

	BOYS			N. observations	GIRLS			N. observations
	OLS	**PSM – ATT**	**PSM – ATU**	**Treated/ Control**	**OLS**	**PSM- ATT**	**PSM – ATU**	**Treated/ Control**
Age 11								
KS2 Overall	0.016	0.065	0.006	4,732	0.130	0.100	0.124	4,589
	(0.034)	(0.052)	(0.055)	848	(0.033)*	(0.049)*	(0.052)*	804
KS2 Maths	0.008	-0.015	0.061	4,711	0.089	0.118	0.085	4,567
	(0.035)	(0.052)	(0.056)	837	(0.034)**	(0.051)*	(0.053)*	800
KS2 English	0.031	0.068	0.038	4,710	0.148	0.091	0.021	4,564
	(0.035)	(0.053)	(0.057)	843	(0.032)**	(0.051)+	(0.052)**	799
KS2 Science	-0.003	0.000	0.0005	4,717	0.109	0.090	0.125	4,556
	(0.035)	(0.053)	(0.056)	837	(0.035)*	(0.051)*	(0.052)**	794
Age 14								
KS3 Overall	0.021	0.017	0.003	4,756	0.14	0.115	0.128	4,623
	(0.034)	(0.052)	(0.055)	854	(0.032)*	(0.050)**	(0.053)	821
KS3 Maths	0.009	-0.006	0.061	4,700	0.113	0.089	0.150	4,516
	(0.035)	(0.053)	(0.056)	837	(0.034)*	(0.051)+	(0.053)	808
KS3 English	-0.005	0.021	-0.038	4,635	0.148	0.155	0.151	4,497
	(0.035)	(0.052)	(0.057)	819	(0.033)**	(0.050)**	(0.051)	795
KS3 Science	0.021	0.084	-0.0005	4,631	0.106	0.093	0.084	4,477
	(0.035)	(0.053)	(0.056)	826	(0.034)*	(0.051)+	(0.053)	798
Age 16								
# GCSE at A*-C	0.092	0.133	0.158	4,949	0.524	0.415	-0.739	4,740
	(0.140)	(0.201)	(0.225)	893	(0.14)**	(0.211)**	(0.220)**	844
A*-C in English	0.032	0.026	0.040	4,949	0.066	0.046	0.085	4,740
	(0.016)*	(0.023)	(0.025)	893	(0.015)**	(0.042)*	(0.024)**	844
A*-C in Maths	0.003	0.004	0.004	4,949	0.057	0.062	0.071	4,740
	(0.017)	(0.002)	(0.025)	893	(0.017)**	(0.025)**	(0.026)**	844
5+ GCSE A*-C incl	0.012	0.016	0.021	4,949	0.062	0.055	0.077	4,740
English and Maths	(0.017)	(0.023)	(0.025)	893	(0.017)**	(0.025)**	(0.026)**	844
Age 18+								
Has any Post GCSE	0.005	.0017	0.0109	5,087	0.032	0.0093	0.037	4,828
qualification	(0.016)	(0.022)	(0.024)	916	(0.017)+	(0.023)	(0.025)	876
Has A levels	-0.0128	-.0074	0.0207	5,087	.0152	0.001	0.042	4,828
	(0.014)	(0.019)	(0.0207)	916	(0.016)	(0.020)	(0.022)+	876
Has applied to	0.022	0.021	0.004	3,693	0.037	0.046	0.045	3,599
university	(0.019)	(0.027)	(0.03)	647	(0.019)+	(0.029)	(0.029)	570
Is at university	0.0002	0.026	0.026	3,215	-0.0044	0.004	0.028	3,285
	(0.021)	(0.030)	(0.03)	560	(0.021)	(0.031)	(0.033)	533

Notes: Standard errors are in brackets + indicates that the underlying coefficient is significant at 10% level, * at 5% and **at 1%.
KS2 and KS3 results are interpretable as % of a standard deviation.

Table 6 Effect of pre-school on cognitive development (by socio-economic status)

Model 2 **PSM**	**Disadvantaged youths**		**Advantaged youths**		**Single mother at birth**		**Partnered mother at birth**		**Mother not employed when child 5**		**Mother employed when child 5**	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Age 11												
KS2 Overall	0.017	0.170	-0.047	-0.031	0.171	0.280	0.018	0.016	0.069	0.062	-0.041	0.107
	(0.06)	(0.07)**	(0.073)	(0.075)	(0.111)	(0.11)**	(0.057)	(0.055)	(0.076)	(0.069)	(0.065)	(0.069)
Age 14												
KS3 Overall	0.050	0.220	-0.0061	0.120	0.113	0.286	0.010	0.092	0.069	0.212	-0.010	0.111
	(0.060)	(0.06)**	(0.078)	(0.085)	(0.101)	(0.10)**	(0.060)	(0.057)	(0.073)	(0.06)**	(0.070)	(0.069)
Age 16												
# GCSE at A*-C	0.410	0.880	0.270	0.430	0.94	0.840	-0.04	0.22	0.360	0.430	0.07	0.60
	(0.25)*	(0.25)**	(0.30)	(0.32)	(0.39)**	(0.44)+	(0.24)	(0.23)	(0.271)	(0.250)*	(0.28)	(0.29)+
5+ GCSE A*-C incl	0.060	0.090	0.027	0.040	0.06	0.110	0.003	0.05	0.020	0.051	0.010	0.071
English and Maths	(0.03)**	(0.03)**	(0.04)	(0.04)	(0.04)	(0.04)**	(0.02)	(0.03)+	(0.030)	(0.030)*	(0.030)	(0.03)*
Age 18+												
Has any Post GCSE	0.026	0.027	-0.027	0.0494	0.027	0.016	-0.017	0.043	0.0177	0.0034	-0.033	0.0756
qualification	(0.027)	(0.027)	(0.039)	(0.042)	(0.040)	(0.041)	(0.027)	(0.028)	(0.032)	(0.033)	(0.0338)	(0.037)
Has A levels	-0.010	0.013	-0.026	0.062	-0.045	0.003	-0.031	0.0204	0.019	0.033	-0.030	-0.009
	(0.022)	(0.022)	(0.043)	(0.041)	(0.046)	(0.063)	(0.031)	(0.031)	(0.033)	(0.034)	(0.037)	(0.040)
Has applied to university	0.01	0.07	0.02	0.02	0.046	0.13	0.030	0.003	0.02	0.024	-0.005	0.06
	(0.04)	(0.041)+	(0.04)	(0.04)	(0.058)	(0.075)+	(0.033)	(0.03)	(0.03)	(0.034)	(0.04)	(0.04)
Is attending university	-0.01	0.02	0.04	-0.01	0.05	-0.033	0.030	-0.005	0.02	-0.02	0.09	0.01
	(0.039)	(0.038)	(0.04)	(0.05)	(0.061)	(0.065)	(0.033)	(0.035)	(0.04)	(0.04)	(0.04)*	(0.04)

Notes: Standard errors are in brackets + indicates that the underlying coefficient is significant at 10% level, * at 5% and **at 1%.
KS2 and KS3 results are interpretable as % of a standard deviation.

Table 7 presents results from the estimation of the effect of pre-school education on non-cognitive and economic activity outcomes. The results are estimated using PSM and the independent variables from Model 2. The evidence on non-cognitive and economic activity outcomes is less precise than we found for test scores. The effects are generally insignificant for girls with the exception that they are significantly more likely to practice sports regularly. The effects for boys are all insignificant in Table 7 apart from those relating to liking school and teachers and this is confined to disadvantaged boys only.

Our findings are consistent with previous literature looking at short-run effects of pre-school education in the UK and especially with Goodman & Sianesi (2005), who find that obtaining education before age 5 is associated with an increase of 7 per cent of a standard deviation in average test scores at age 11 and an increase of 5 per cent of a standard deviation at age 16. It is also consistent with a positive but moderate long run effect on higher education, even if we do not find any evidence on actual university attendance (but we do find an effect on the probability of applying to university). Our results are also consistent with Osborn & Milbank (1987) who show a stronger positive effect of pre-school attendance on vocabulary expansion (indeed, we find stronger effects on language rather than mathematical skills). Moreover, our insignificant results on higher education would be consistent with Feinstein et al. (1998) who show that the marginally positive effects of pre-school fade when the children grow up.

There are various key issues in terms of comparability of our findings with previous British studies. First of all, there is a difference in the measurement of the key independent variable, i.e. pre-school attendance. We use a binary self-report and we have no information on the type (so we cannot separately analyse private and Local Authority nurseries from playgroups, as it is done by Osborn & Milbank, 1987 and Goodman & Sianesi, 2005), or on the hours of attendance (as in Feinstein et al., 1998). It is possible that our moderately positive results reflect this definition, as we may capture the positive effect of infrequent or low levels of exposure to various types of formal pre-school. However, Osborn & Milbank (1987) underlines that "The actual type of pre-school experience matters very little" as the majority of different pre-school institutions shows common elements, such as interesting and various activities for the children, proper care, and social interaction. This suggests that variation in the definition of childcare across datasets may not play an important role in the results.

Table 7 Effect of pre-school education on other outcomes: (Model 2 PSM)

	All Boys	All Girls	Advantaged Boys	Advantaged Girls	Disadvantaged Boys	Disadvantaged Girls
Attitudes/socialisation						
Number of friends	-0.005	0.051	0.176	0.161	-0.052	0.118
at wave 6 or 7	(0.074)	(0.060)	(0.095)+	(0.095)+	(0.106)	(0.092
Often spends free time	0.030	0.039	-0.0108	0.069	0.036	0.000
with friends (wave 1)	(0.022)	(0.023)+	(0.038)	(0.040)+	(0.028)	(0.01)
Reads for pleasure at	0.052	0.014	-0.030	-0.013	0.013	0.008
least weekly (wave 1)	(0.022)*	(0.019)	(0.034)	(0.030)	(0.028)	(0.025)
Plays sports at least	0.015	0.058	0.004	0.003	0.028	0.079
weekly (wave 1)	(0.016)	(0.022)*	(0.024)	(0.036)	(0.021)	(0.028)**
TV for 4+ hours in	0.005	0.031	0.006	-0.028	0.049	.0415
school day (wave 1)	(0.018)	(0.019)	(0.028)	(0.032)	(0.033)	(0.033)
Ever been bullied	0.011	-0.032	-0.046	-0.021	0.038	0.003
(up to wave 3)	(0.025)	(0.026)	(0.041)	(0.043)	(0.034)	(0.034)
Like her/his school	0.033	0.020	0.007	-0.003	0.0621	0.023
(wave 1)	(0.018)+	(0.017)	(0.028)	(0.028)	(0.026)**	(0.022)
Like her/his teachers at	0.061	0.007	0.075	0.075	0.089	.0.023
wave 1	(0.023)*	(0.023)	(0.038)+	(0.041)+	(0.030)**	(0.021)
Lots of effort in	0.021	0.019	0.029	0.031	0.008	0.0140
schoolwork (wave 1)	(0.019)	(0.018)	(0.033)	(0.032)	(0.026)	(0.022)
Economic activity						
NEET	0.0003	-0.017	-0.007	0.009	-0.005	0.0318
at wave 6 or 7	(0.018)	(0.018)	(0.019)	(0.018)	(0.031)	(0.031)
Is working	-0.011	-0.021	0.004	-0.064	-0.041	0.066
at wave 6 or 7	(0.03)	(0.031)	(0.040)	(0.045)	(0.046)	(0.043)
Health Behaviours						
Ever Smoked	0.015	-0.039	0.044	-0.011	0.031	0.037
	(0.019)	(0.023)+	(0.032)	(0.038)	(0.026)	(0.028)
Ever tried cannabis	-0.002	0.033	-0.017	0.044	-0.024	0.002
	(0.023)	(0.022)	(0.041)	(0.039)	(0.028)	(0.004)
Alcohol at least weekly	0.011	0.015	-0.024	-0.004	0.028	0.065
(wave 3)	(0.027)	(0.027)	(0.041)	(0.045)	(0.040)	(0.04)
Ever been pregnant	NA	-0.004	NA	-0.004	NA	0.039
		(0.023)		(0.030)		(0.032)
Early first sexual	-0.023	0.050	0.034	0.044	-0.010	0.121
intercourse	(0.036)	(0.034)	(0.049)	(0.048)	(0.054)	(0.071)
Mental health	0.015	-0.14	-0.227	-0.346	-0.126	0.189
(ghq score wave 2)	(0.15)	(0.11)	(0.167)	(0.257)	(0.162)	(0.188)
Problem behaviours						
Ever shoplifted	-0.007	0.013	-0.015	0.027	-0.001	0.016
	(0.009)	(0.010)	(0.013)	(0.017)	(0.013)	(0.013)
Ever vandalised	0.002	0.004	0.003	0.000	0.014	0.008
	(0.010)	(0.007)	(0.014)	(0.001)	(0.014)	(0.009)
Ever graffitied	0.008	0.004	0.016	0.008	0.009	-0.004
	(0.008)	(0.005)	(0.013)	(0.009)	(0.011)	(0.007)

Notes: Standard errors are in brackets + indicates that the underlying coefficient is significant at 10% level, * at 5% and **at 1%.

Secondly, some existing studies analyse the overall effects of early education, including both pre-school and early entry into primary school (Goodman & Sianesi, 2005) in contrast to the clearer institutional context in which the LSYPE cohort received pre-school (see Section 2 for details). Further, we show that disadvantaged children are those who benefit more from pre-school attendance in terms of cognitive outcomes. This result is also reflected in the previous literature.

Further, we show that disadvantaged children are those who benefit more from pre-school attendance. These children were more likely to attend pre-school in the early Nineties than in the Seventies (mostly because of the very little tax and financial incentives for working parents in that early period). Therefore, another factor to keep into consideration when comparing results is the difference in the characteristics of the population of children attending pre-school in LSYPE data with respect to older data sets, such as 1970 BCS and 1958 NCDS. Lastly, our study shows mildly positive effects of pre-school received by the LSYPE children in 1993-94. This was the beginning of a period characterised by an increasing focus on the provision of early education as well as care, as we highlighted in Section 2. Therefore, it is possible that our results are driven by the increased quality in the provision of childcare, with respect to the previous decades. However, our results are generally consistent with previous studies and we do not find big discrepancies that are hard to justify.

In comparison with US studies, we are not able to demonstrate significant effects on crime or on health related behaviour. The difference in the effects on crime, relative to Perry pre-school work, may arise because of the substantially lower rates of crime and incarceration in the UK. The Perry evaluation data does not contain health information although Carneiro & Ginja (2012) investigate the effects of the US Head Start programme on health and find important effects on obesity in adolescence, as well as crime in early adulthood. The health differences may arise because the UK's National Health Service provides better access to healthcare for children from disadvantaged backgrounds. The effects of pre-school on sports participation by disadvantaged girls we find here would be consistent with the HeadStart obesity effects. While we are not able to demonstrate important non-cognitive effects even for disadvantaged children, we are able to demonstrate important impacts on cognitive achievement for disadvantaged children that provides some support the case for early interventions.

7. Conclusion

In this paper, we have investigated the effect of pre-school on various children's outcomes, including cognitive development and non-cognitive outcomes in adolescence and early adulthood. We have used the Longitudinal Study of Young People in England, which is a rich source of information on English teen-agers and can be linked to the National Pupil Database, in order to get detailed information on school outcomes. We find that pre-school moderately increases test scores at age 11, 14 and 16 and is more beneficial for girls and children from disadvantaged socio-economic backgrounds. The size of the effect is notable and is comparable to other important variables, such as birth-weight, sex of the child, or parental disability.

Our analysis is performed using Ordinary Least Squares and Propensity Score Matching. We make extensive use of the extensiveness of the information contained in LSYPE and increase our set of independent variables, in order to control for wider factors affecting both school outcomes and pre-school attendance. Our results are stable over the two different specifications of our model. Moreover, Propensity Score Matching allows us to better compare a group of children who did not go to pre-school school with observationally similar children who did go pre-school school, given our independent variables.

The paper provides further evidence of the positive effects of pre-school on cognitive outcomes that are well known to have important effects on lifecycle living standards. We also find positive effects on attitudes towards schooling and this might be one transmission mechanism through which cognitive scores are increased. However, we find less significant effects on post-GCSE qualifications, intentions to attend higher education or on actual attendance. The fact that pre-school raises attainment at age 16 but has no significant effect on higher education is a puzzle that deserves further research. It is unclear why the effects fade at this stage of the education process.

Moreover, the results on non-cognitive outcomes are generally insignificant relative to those on cognitive achievements. While there is a significant effect on sports participation for disadvantaged girls, we do not find any significant effect on mental well-being or on problematic behaviours even for disadvantaged children. We find no effects on petty crime, nor do we find effects on risky health behaviours. These are precisely the outcomes that have been emphasised as social benefits in the recent literature. It would appear that our positive results are largely confined to cognitive outcomes and mostly for disadvantaged youths.

These are important since they are strongly and causally associated with lifecycle living standards. The fact that they are strong for those from disadvantaged backgrounds suggests that subsidies should be means-tested rather than universal. However, the results do not lend strong support for the idea that there are important *external* effects in later life associated with pre-school participation. The benefits are largely private ones: social benefits associated with health risky and criminal behaviours do not seem to be among the advantages that UK pre-school confers.

References

Abadie, A. and Imbens, G.W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1), 235-267.

Anderson, M.L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training projects. *Journal of the American Statistical Association* 103 (484), 1481-1495.

Angrist, J.D. & Krueger, A.B. (1999). Empirical strategies in labor economics, in O. Ashenfelter, & D. Card (eds.), *Handbook of Labor Economics*, edition 1, volume 3, (pp.1277-1366). Amsterdam: Elsevier.

Audit Commission (1996). *Counting to Five: Education of Children under Five*, London, HMSO.

Berlinski S., Galiani S., & Gertler, P.(2009). The effect of pre-primary education on primary school performance. *Journal of Public Economics* 93(1-2), 219-234.

Baker M., Gruber J., & Milligan K. (2008). Universal childcare, maternal labor supply and family well-being. *Journal of Political Economy*, 116(4), 709-745.

Bernal, R., & Keane, M.P. (2010). Quasi-structural estimation of a model of childcare choices and child cognitive ability production. *Journal of Econometrics*, 156(1), 164-189.

Bernal, R., & Keane, M.P. (2011). Child care choices and children's cognitive achievement: the case of single mothers. *Journal of Labor Economics*, 29(3), 459-512.

Berger, L.M., Hill, J., & Waldfogel, J. (2005). Maternity leave, early maternal employment and child health and development in the US. *The Economic Journal*, 115(501), F29-F47.

Bianchi, S.M. (2000). Maternal employment and time with children: dramatic change or surprising continuity? *Demography*, 37(4), 401–414.

Bingley, P., & Westergaard-Nielsen, N. (2012). Intergenerational transmission and day care in Denmark. In J. Ermisch, M. Jantti, & T. Smeeding (eds.) *Inequality from Childhood to Adulthood: A Cross-National Perspective on the Transmission of Advantage*. New York: Russell Sage Foundation.

Blau, D., & Currie, J. (2006). Pre-school, day care, and after school care: who's minding the kids? In E. Hanushek, & F. Welch (eds.), *Handbook of the Economics of Education*, Vol. 2(2) (pp. 1163–1278). Amsterdam: Elsevier.

Brewer, M., Crawford, C., & Dearden, L. (2005). Reforms to childcare policy. In Chote, R., Emmerson, C., Miles, D., Olderfield, Z (eds.) *The IFS Green Budget 2005*, London: The Institute of Fiscal Studies.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31-72.

Carneiro, P. & Ginja, R. (2012). Long term impacts of compensatory preschool on health and behaviour: evidence from Head Start, *Centre for Microdata Methods and Practice, Institute for Fiscal Studies*, CWP01/12.

Carneiro, P., & Heckman, J.J. (2003). Human capital policy. In: J.J. Heckman, A.B. Krueger & B.M. Friedman. (Eds.), *Inequality in America: What role for human capital policies?* Cambridge Mass: MIT Press.

Cunha, F., Heckman, J.J., Lochner, L., & Masterov, D.V. (2006). Interpreting the evidence on life cycle skill formation. In: Hanushek, E., and Welch, F. (eds.), *Handbook of the Economics of Education.* (pp. 697- 812) Amsterdam: North-Holland.

Currie, J. (2001). Early childhood education programs. *Journal of Economic Perspectives*, 15(2), 213–238.

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

Dehejia, R.H. & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics*, 84(1), 151–61.

Department for Education and Employment (1999a). Early education provision for four year old in England, *Statistical First Release* 20/1999.

Department for Education and Employment (1999b), Children day care facilities at 31/3/1999: England, *Statistical First Release* 28/1999.

Department for Education (2011). Provision for children under five years of age in England, *Statistical First Release* 2011.

Dickson, M. (2008), The effect of free pre-school education on children's subsequent academic performance: Empirical evidence from England, *University of Warwick Working paper.*

Dumas, C. & Lefranc, A. (2012). Early schooling and later outcomes: Evidence from pre-school extension in France. In J. Ermisch, M. Jantti, and T. Smeeding, (eds). *Inequality from Childhood to Adulthood: A Cross-National Perspective on the Transmission of Advantage*. New York: Russell Sage Foundation.

Feinstein L., Robertson, D., & Symons, J. (1998). Pre-school education and attainment in the NCDS and BCS, Centre for Economic Performance, London School of Economics and Political Sciences, Paper no. 382.

Goodman, A., & Sianesi, B. (2005). Early education and children's outcomes: how long do the impacts last? *Fiscal Studies* 26(4), 513-548.

Havnes, T., & Mogstad, M. (2011). No child left behind. Universal child care and children's long-run outcomes, *American Economic Journal: Economic Policy* 3(2), 97-129.

Heckman, J.J. (1979). Sample selection bias as a specification error. *Econometrica* 47(1), 153-161.

Heckman, J.J., & Robb, R. (1985). Alternative models for evaluating the impact of interventions. *Journal of Econometrics*, 30(1-2), 239-267.

Heckman, J.J., Ichimura, H., & Todd, P.E. (1998). Matching as an Econometric Evaluation Estimator. *Review of Economic Studies*, 65(2), 261–294.

Heckman, J.J., and Masterov, D.V. (2007). The Productivity Argument for Investing in Young Children, *Review of Agricultural Economics*, 29(3), 446-493.

Heckman, J.J., Moon, S.H., Pinto, R., Savelyev, P.A.& Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program, *Journal of Public Economics*, 94(1-2), 114-128.

Imbens, G.W. (2004). Nonparametric estimation of average treatment effects under exogeneity: a review. *Review of Economics and Statistics*, 86(1), 4–29.

Imbens, G.W., & Wooldridge, J.M. (2009). Recent developments in the econometrics of program evaluation, *Journal of Economic Literature*, 47(1), 5-86.

Keane, M.P., & Wolpin, K.I. (2001). The effect of parental transfers and borrowing constraints on educational attainment, *International Economic Review* 42(4), 1051-1103.

Keane, M.P., & Wolpin, K.I. (2010). The role of labor and marriage markets, preference heterogeneity and the welfare system in the life cycle decisions of black, hispanic and white women, *International Economic Review* 51(3), 851-892.

Lamb, M.E. (1996). Effects of nonparental child care on child development: an update. *Canadian Journal of Psychiatry*, 41(6), 330-342.

Lechner, M. (1999). Earnings and employment effects of continuous off-the-job training in East Germany after unification. *Journal of Business & Economic Statistics*, 17, 74–90.

Leher, S.F., & Kordas, G. (2013). Matching using semi-parametric propensity scores. *Empirical Economics*, 44(1), 13-45.

Leuven, E., & Sianesi, B. (2003). PSMATCH2: Stata module to perform full Mahalanobis propensity score matching, common support graphing, and covariate imbalance testing.

Loeb, S., Bridges, M., Bassok, D., Fuller, B. & Rumberger, R.W. (2007). How much is too much? The influence of preschool centers on children's social and cognitive development. *Economics of Education Review* 26(1), 52-66.

Magnuson, K.A., Ruhm, C., and Waldfogel, J. (2007). Does prekindergarten improve school preparation and performance? *Economics of Education Re*view 26(1), 33-51.

McCartney, K. (1984). The effect of quality of day care environment on children's language development. *Developmental Psychology* 20(2), 244-260.

McLean, C. (2012), Governing the Market: Childcare Policy and Provision in the US and UK, mimeo.

Melhuish, E.C. (2003). *A literature review of the impact of early years provision on young children, with emphasis given to children from disadvantaged backgrounds*, Institute for the Study of Children, Families & Social Issues Birkbeck, University of London.

National Evaluation Sure Start Team (NESS) (2010). The Impact of Sure Start Local Programmes on Child Development and Family Functioning: Report of the Longitudinal Study of 5-year-old Children and their Families. London: Department for Education

Osborn, A. F., & Milbank, J.E. (1987). *The effects of early education: a report from the child health and education study*. Oxford, Clarendon Press.

Rosenbaum, P.R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.

Ruhm, C.J. (2008). Maternal Employment and Adolescent Development. *Labour Economics* 15(5), 958-983.

Ruhm, C.J., & Waldfogel, J. (2013 forthcoming), Long term effect of early childcare and education, *Nordic Economic Policy Review*.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2004). The final report: effective pre-school education. The Effective Provision of Pre-School Education (EPPE) Project, Technical Paper 12. London: Department for Education and Skills / Institute of Education, University of London.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2012). Effective pre-school, primary and secondary education 3-14 project (EPPSE 3-14) - Final report from the Key Stage 3 phase: influences on students' development from Age 11-14, London: Department for Education / Institute of Education, University of London.

Zhao, Z. (2008). Sensitivity of propensity score methods to the specifications. *Economics Letters*, 98(3), 309–319.

Appendix

Tables A1 and A2 are explained in the text.

Table A1 Effects of other independent variables on cognitive development

	KS 2 Overall Score	KS3 Overall Score	# GCSEs at A*-C
Premature birth	-0.001 (0.126)	0.022 (0.204)	0.098 (0.133)
Main parent has a disability	-0.135 (0.087)	-0.400 (0.14)**	-0.445 (0.091)**
Grandparents went to university	0.847 (0.128)**	1.718 (0.206)**	0.692 (0.132)**
English as main language	0.568 (0.192)**	0.815 (0.292)**	-0.029 (0.192)
Child has a disability	-1.251 (0.101)**	-1.996 (0.164)**	-1.198 (0.106)**
N. older siblings	-0.372 (0.033)**	-0.743 (0.053)**	-0.492 (0.034)**
Mother senior high school graduate	-1.624 (0.118)**	-3.485 (0.192)**	-1.824 (0.122)**
Mother junior high school graduate	-3.067 (0.150)**	-5.968 (0.243)**	-3.427 (0.156)**
Mother no qualification	-3.524 (0.140)**	-6.822 (0.225)**	-3.746 (0.145)**
Birth-weight in kg	0.468 (0.068)**	0.672 (0.109)**	0.305 (0.071)**
Sex - Male	-0.341 (0.070)**	-0.813 (0.113)**	-0.940 (0.074)**
Single parent household at birth	-0.964 (0.086)**	-2.070 (0.139)**	-1.496 (0.090)**
Mother<20 at birth	-1.303 (0.146)**	-2.489 (0.233)**	-1.796 (0.152)**
Month of birth			
October	-0.314 (0.173)+	-0.640 (0.28)**	-0.375 (0.182)*
November	-0.287 (0.175)	-0.509 (0.282)+	-0.377 (0.183)*
December	-0.658 (0.173)*	-0.987 (0.278)**	-0.501 (0.182)**
January	-0.562 (0.172)*	-0.788 (0.276)**	-0.382 (0.18)*
February	-0.900 (0.171)*	-1.152 (0.277)**	-0.698 (0.18)**
March	-0.673 (0.171)*	-0.772 (0.275)**	-0.417 (0.179)*
April	-1.323 (0.172)*	-1.527 (0.277)**	-0.864 (0.181)**
May	-1.290 (0.168)*	-1.537 (0.272)**	-0.719 (0.177)**
June	-1.275 (0.168)*	-1.497 (0.271)**	-0.695 (0.176)**
July	-1.362 (0.168)*	-1.500 (0.271)**	-0.577 (0.176)**
August	-1.629 (0.167)**	-2.092 (0.27)**	-0.992 (0.176)**
Ethnic background			
Black	-1.028 (0.153)**	-1.834 (0.235)**	-0.168 (0.155)
Asian	0.210 (0.13)	1.077 (0.207)**	1.621 (0.136)**
Mixed	0.119 (0.135)	0.257 (0.216)	0.511 (0.141)**

Notes: Standard errors are in brackets + indicates that the underlying coefficient is significant at 10% level, * at 5% and **at 1%.

Table A2 – Results from estimation with alternative matching techniques

	Model 1						Model 2					
Age 11	PSM NNeighbour	PSM Bootstrap se's	Radius Matching	Kernel Matching	Mahalanob. Matching	IPW	PSM NNeighbour	PSM Bootstrap se's	Radius Matching	Kernel Matching	Mahalanob. Matching	IPW
KS2 Overall	0.153	0.153	0.184	0.184	0.142	0.150	0.071	0.071	0.076	0.093	0.0095	0.062
	(0.038)**	(0.040)**	(0.026)**	(0.026)**	(0.034)**	(0.025)**	(0.036)+	(0.038)+	(0.029)**	(0.028)*	(0.029)	(0.024)*
KS2 Maths	0.113	0.113	0.116	0.123	0.101	0.113	0.051	0.051	0.045	0.072	-0.006	0.046
	(0.037)**	(0.036)**	(0.028)**	(0.027)**	(0.034)**	(0.026)**	(0.037)	(0.027)+	(0.029)	(0.028)*	(0.031)	(0.025)+
KS2 English	0.130	0.130	0.170	0.185	0.188	0.170	0.099	0.099	0.081	0.106	0.041	0.086
	(0.038)**	(0.037)**	(0.028)**	(0.028)**	(0.036)**	(0.026)**	(0.037)**	(0.034)**	(0.029)**	(0.029)**	(0.031)	(0.024)**
KS2 Science	0.131	0.131	0.123	0.129	0.106	0.121	0.032	0.032	0.050	0.076	-0.009	0.045
	(0.037)**	(0.033)**	(0.028)**	(0.028)**	(0.035)**	(0.026)**	(0.036)	(0.038)	(0.029)+	(0.029)*	(0.027)	(0.020)+
Age 14												
KS3 Overall	0.160	0.158	0.160	0.181	0.198	0.165	0.052	0.053	0.063	0.103	0.035	0.065
	(0.037)**	(0.041)**	(0.028)**	(0.028)**	(0.037)**	(0.26)**	(0.037)	(0.038)	(0.029)*	(0.029)	(0.032)	(0.024)**
KS3 Maths	0.094	0.093	0.128	0.170	0.168	0.134	0.049	0.050	0.055	0.079	0.012	0.048
	(0.038)**	(0.037)*	(0.029)**	(0.026)**	(0.036)**	(0.026)**	(0.037)	(0.039)	(0.030)+	(0.029)*	(0.033)	(0.024)+
KS3 English	0.076	0.076	0.141	0.168	0.176	0.151	0.074	0.076	0.063	0.094	0.046	0.059
	(0.037)*	(0.036)+	(0.028)**	(0.028)**	(0.038)**	(0.026)**	(0.037)*	(0.037)*	(0.029)*	(0.028)**	(0.032)	(0.024)*
KS3 Science	0.150	0.150	0.180	0.153	0.178	0.138	0.032	0.032	0.058	0.084	0.027	0.047
	(0.037)**	(0.041)**	(0.026)**	(0.028)**	(0.038)**	(0.026)**	(0.037)	(0.041)	(0.030)+	(0.029)**	(0.033)	(0.024)+
Age 16												
# GCSE at A*-C	0.678	0.678	0.665	0.678	0.643	0.645	0.346	0.346	0.347	0.43	0.270	0.306
	(0.150)**	(0.14)**	(0.11)**	(0.11)**	(0.14)**	(0.11)**	(0.15)**	(0.13)**	(0.12)*	(0.12)**	(0.12)+	(0.12)*
A*-C in English	0.093	0.093	0.088	0.089	0.087	0.084	0.052	0.052	0.049	0.059	0.054	0.045
	(0.02)**	(0.015)*	(0.01)**	(0.01)**	(0.016)**	(0.01)**	(0.02)**	(0.01)**	(0.01)*	(0.01)*	(0.01)**	(0.01)*
A*-C in Maths	0.046	0.046	0.061	0.065	0.066	0.060	0.041	0.041	0.037	0.0415	0.0238	0.028
	(0.02)**	(0.01)**	(0.01)**	(0.01)**	(0.016)**	(0.01)**	(0.02)**	(0.01)**	(0.01)*	(0.01)*	(0.01)+	(0.01)*
5+ A*-C incl	0.075	0.075	0.074	0.076	0.074	0.072	0.062	0.062	0.046	0.052	0.035	0375
Math + English	(0.02)**	(0.02)**	(0.01)**	(0.01)**	(0.02)**	(0.01)**	(0.017)**	(0.01)**	(0.01**	(0.01)**	(0.01)*	(0.01)*

Notes: Standard errors are in brackets + indicates that the underlying coefficient is significant at 10% level, * at 5% and **at 1%.
KS2 and KS3 results are interpretable as % of a standard deviation.

JEL Classification

JEL codes: J13 I21

Highlights

The relationship between pre-school attendance and children's outcomes is analysed

We find evidence that pre-school improves results in cognitive tests

Positive effects are found for children from disadvantaged socioeconomic background