

Peter, Frauke H.; Schober, Pia S.; Spiess, C. Katharina

Working Paper

Early birds in day care: The social gradient in starting day care and children's non-cognitive skills

DIW Discussion Papers, No. 1438

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Peter, Frauke H.; Schober, Pia S.; Spiess, C. Katharina (2014) : Early birds in day care: The social gradient in starting day care and children's non-cognitive skills, DIW Discussion Papers, No. 1438, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/106201>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

1438

Discussion Papers

Deutsches Institut für Wirtschaftsforschung

2014

Early Birds in Day Care: The Social Gradient in Starting Day Care and Children's Non-cognitive Skills

Frauke H. Peter, Pia S. Schober and C. Katharina Spiess

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

IMPRESSUM

© DIW Berlin, 2014

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
<http://www.diw.de>

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
<http://www.diw.de/discussionpapers>

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
<http://ideas.repec.org/s/diw/diwwpp.html>
<http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html>

Early Birds in Day Care: The Social Gradient in Starting Day Care and Children's Non-cognitive Skills

Frauke H. Peter, Pia S. Schober and C. Katharina Spiess

DIW Berlin, Germany and Freie Universität Berlin, Germany

December 2014

* Corresponding author: C. Katharina Spiess, Email: kspiess@diw.de, phone: +49(0)30-89789-254, fax: +49(0)30-89789-109, DIW Berlin, Mohrenstraße 58, 10117 Berlin, Germany.

Acknowledgements: Valuable comments by Jan Marcus and Ludovica Gambaro on an earlier version of this paper are gratefully acknowledged. We thank Chrysanthi Tsiasioti and Sebastian Streibel for helpful research assistance.

Abstract

In recent years, almost all children below school age in Western industrialized countries have some experience of attending day care institutions. However, the age at which children enter day care and therefore the overall time spent in day care varies substantially. We investigate the potential impact of later day care entry on the social and emotional behaviour of children, one important aspect of non-cognitive skills. Based on the English sample of the Millennium Cohort Study, we analyse the effects on children's development at the age of five and seven, using propensity score techniques. We find clear evidence of effects on children's development at the age of seven: Later day care entry increases children's peer-problems and reduces prosocial behaviour. We find that boys with low educated mothers and from families with a household income below the poverty line are most strongly affected.

Keywords: Day care entrance, early start, socio-emotional behaviour, propensity score matching.

JEL classifications: J13, I21.

1 Introduction

With a growing labour force participation of mothers with young children in many industrialized countries and increasing acceptance of day care centres as educational institutions, the percentage of children attending day care has gone up. In 2010, at the EU27-level almost 83 percent of all children 3 to 5 years of age attended some sort of day care institution (or preschool) before they enter school¹. By contrast, only every third child below the age of three years attended some kind of day care in 2010 (33%), with huge differences between the countries: These range from 7 to 66 percent of all under three years olds in day care. The UK ranges with 93 percent of children from age 3 and 42 percent of children under three years slightly above the EU average.²

Overall day care has become an important phase in children's education biography and questions regarding the effects of day care attendance on children's development have gained increasing attention also in the economic literature (see chapter 2). The results of economic studies have been diverse, among other reasons due to significant cross-national variation in how and to which groups day care services are provided. Furthermore, most of these studies mainly focus on the effects of attending versus not attending day care. Given that in the European context nowadays nearly all children attend day care institutions at some point in their lives, the more relevant and challenging question concerns the optimal starting age³. We therefore investigate the question 'is it better to start early or later?' based on a rich data set for England as one European day care context.

The theory on skill formation posits that children develop cognitive and non-cognitive skills by childhood stages (see Cunha/Heckman 2007, 2008). It emphasizes that children's development benefits from high quality care and education contexts. In particular children who were exposed to adverse (family) environments are assumed to benefit most from the

¹ The term day care is used in various ways in the literature, covering various types of early childhood education and care services or preschool provision. We use this term for all kinds of non-compulsory formal child care arrangements taking place outside the home of a child.

² All relevant figures in this chapter on attendance rates in EU countries are taken from the OECD Family Data base (OECD 2014).

³ For an overview of this discussion, see Baker (2011).

experience of extra-familial care. Thus one might argue that an earlier starting age in high-quality day care and longer experience in terms of years of attendance are more beneficial for the development of these children than a later entry and fewer years of experience. This can mainly be explained by the complementarity of skill formation as higher levels of skills at an earlier level beget further development of skills at a later stage (e.g. Cunha et al. 2010). Emphasizing this later mechanism, one might also argue that an earlier entry may be beneficial for children from high quality family environments as their skills developed in the family context would be complemented by positive learning experiences in a day care context. However, it should be noted that for very young children, e.g. under one year of age, or in combination with long hours of care, some studies point to adverse developmental consequences (e.g. Belsky et al. 2007). Therefore further evidence is needed on how the consequences of day care attendance vary by the age when children start day care.

We focus on the effects of the time of day care entry on non-cognitive skills, as an increasing number of economic studies suggest that non-cognitive skills are important predictors of later educational achievements, health outcomes, and labour market success during adulthood (e.g., Blanden et al 2007, Carneiro et al 2007, Currie/Stabile 2006, Heckman et al. 2013, and Prevoort/Weel 2013). Cunha and Heckman (2006) even find that non-cognitive skills promote the formation of cognitive skills but not vice versa. Heckman et al. (2006) also show that non-cognitive skills are as important for school enrolment decisions as cognitive skills. Nevertheless, few economic studies so far have focussed on the effects of day care on non-cognitive skills. There is one prominent exception, a reanalysis of Heckman and co-authors of the Perry Preschool Project, an early intervention study targeted at disadvantaged children. Heckman et al. (2013) show, that non-cognitive skills explain more of the variance in later outcomes than cognitive skills, in particular for boys. Nevertheless, as pointed out in many studies (see e.g. Baker et al. 2008, Havnes/Mogstad 2011), these results cannot be transferred to more universal programs.

We are interested in whether starting day care in a more universal system after age 2.5 has adverse effects on English children's non-cognitive skills at primary school age. Rather than

using simply an overall index of socio-emotional behaviour as most previous economic studies, we differentiate also between individual dimensions, such as social versus emotional development. This allows a better understanding of the mechanisms how day care environments affect some skill dimensions more than others and to disentangle which dimensions drive the overall effects. Furthermore, we focus on the heterogeneity of effects by gender and parental socio-economic background. This enables us to explore the assumption that earlier day care experience is more beneficial for children from disadvantaged family backgrounds. If disadvantaged children would benefit in particular, policies to promote their earlier entry might be one means to reduce the large developmental gap that exists between children from different socio-economic backgrounds already at school entry age. To contribute to the understanding of effects of starting age on socio-emotional behaviour, we aim at identifying a causal relationship by using propensity score matching. An advantage of propensity score methods is that estimates are less dependent on the functional form in the model, i.e. assumption of a linear relationship compared to OLS. Propensity Score Matching has also been used in several recent papers that investigate the determinants of child outcomes in the economic literature (e.g., Berger et al 2005, Goodman/Sianesi 2005, Ruhm 2008, and Apps et al. 2013).

2 Related literature

Many economic studies have investigated effects of universal day care services on children. Given our research focus, in this review we first of all focus only on studies investigating the effects on non-cognitive skills (for studies analysing day care effects on cognitive outcomes or health outcomes see e.g. Brilli et al 2011, Fitzpatrick 2008, or Magnuson et al. 2007). Secondly, we concentrate on studies considering age at day care entry or the duration of the day care experience rather than studies on the effects of any day care attendance. Thirdly, as we observe short- and medium-term outcomes of children, we review mainly studies with a similar time frame, excluding those which analyse long-run effects (for studies on effects in adulthood, see e.g. Bingley/Westergaard-Nielsen 2010, Dumas/Lefran 2010 or Havnes/Mogstad 2011, for

effects on adolescents see e.g., Apps et al. 2013 and above)⁴. Moreover, only a few studies explicitly analyse the heterogeneity of these effects for different groups.

Among the few studies that investigate specifically the duration of day care and focus on socio-emotional outcomes, the study by Datta-Gupta and Simonsen (2010) is the most comparable. Based on Danish survey data, they investigate the effects of a Danish policy reform on children's non-cognitive skills at age 7 using the same measure as we do. They use a reform regarding the differences in the take up of preschool⁵ or family day care, respectively, between municipalities after they introduced guaranteed access. Although they do not explicitly focus on age of entry, their results can roughly speaking be interpreted as the effects of enrolling in preschool one year earlier. They find no overall effects of attending preschool earlier on the socio-emotional behaviour of children, not even for subgroups. However, they find significant adverse effects of attending family day care on socio-emotional behaviour of boys with less educated mothers. The authors only use one measure to capture the socio-emotional behaviour and do not differentiate between subscales to investigate which dimensions of a child's socio-emotional behaviour are most strongly affected by preschool attendance. In a later study by the same authors (2012) they use the same reform as an instrument. This study focuses on a set of various non-cognitive and cognitive measures at age eleven. Again they find no significant effects of attending preschool at age three on socio-emotional behaviour. They only find one significant effect: that children attending day care liked school more.

Another study by Baker et al. (2008) uses a reform in Canada's largest province Quebec to identify the effects of a large extension of access to a full-day place at a day care institution first for 3-year olds, then for 2-year olds and finally for all children aged less than 2 on various outcomes. With respect to non-cognitive skills their estimates indicate that behaviour deteriorated in the treated province. However, they do not examine the age of entry. This is done in the study by Kottelenberg and Lehrer (2014), which extends the study by Baker et al. (2008). Kottelenberg and Lehrer (2014) explicitly focus on differences in the age children got

⁴ For a very good summary on various studies see Ruhm and Waldfogel (2011).

⁵ The preschool term they use is very similar to what we consider day care - at least for the majority of care settings we analyse.

access to subsidized child care. Their estimates show that children who got earlier access experience significantly larger negative impacts on motor-social development scores and behavioural outcomes. The only behavioural outcomes for which the significant negative relationship steepens with age are hyperactivity and inattention scores. However, their further subgroup analysis shows that access to day care at 3 years of age may benefit the most disadvantaged.

A study based on US data from the Early Childhood Longitudinal Study Kindergarten-Class (ECLS-K) (Loeb et al. 2007) explicitly takes into account differences in day care starting age and in intensity. They identify effects using OLS, matching, and instrumental variables for five-year-old children. The authors find that on average, attending centre care is associated with positive gains in cognitive skills, but negative for social behaviour. Across economic levels, the worst behavioural effects occur for those who enter the earliest. Starting earlier than age 2 is related to more pronounced negative social effects. However, for some disadvantaged children, the socio-behavioural effects are neutral.

Studies on the effect of (early) day care attendance on non-cognitive skills using German data report different results. Based on data of the German Socio-Economic Panels Study (SOEP) and focusing on the same outcome measure we use, Müller et al. (2013) show that children attending day care before the age of three have less socio-emotional problems at primary school age (5-10 year old children). Using an instrumental variable approach and based on the 'DJI children panel', Schlotter and Wößmann (2010) estimate a positive effect of an earlier day care entry on children's assertiveness and ability to form friendships. In a recent study, Felfe and Lalive (2014) analyse the effect of early day care entrance based on administrative data for one German state. Using a marginal treatment effects framework, they show that early day care is more beneficial for developing socio-emotional maturity based on paediatrician and parent reports for children of less educated mothers than for children of highly educated mothers.⁶

⁶ Apart from the economic literature there are many other studies, mainly by educational scientists, on the effectiveness of day care programs on child outcomes (for recent meta-analysis see e.g. Burger 2010 or

3 Institutional context

In this paper we focus on the UK, and, more specifically, on England. While all UK countries share similar early childhood education and care systems, countries within the UK have considerable autonomy in the way they fund, organise and regulate services. In particular, compulsory schooling starts one year earlier in Northern Ireland, while in Scotland primary schools do not offer full-day provision the year before compulsory schooling as England and Wales do. Moreover, quality regulation and curricula differ across the countries (for summaries for Wales, Scotland, and Northern Ireland, see EACEA 2009, OECD 2000).

We restrict our analysis to the effects on children attending day care (or preschool institutions) in England in order to reduce the contextual variation in early childhood education and care institutions. The following brief description will discuss day care policies for children from birth until they reach compulsory school age, which is age five in England. We focus on the years between 2000 and 2005, which are the ones observed in the empirical analysis, and also coincide with several reforms and initiatives by the Labour government in power at the time. In the UK, mothers of children born in 2000 or 2001, the cohort we focus on, were entitled to 14 weeks, 18 weeks, or 29 weeks, depending on the length of their employment tenure with the same employer (Moss/O'Brien 2006, Ringen 1997).

Historically, day care provision in England was a matter for local decision and dominated by a market economy of child care. In England providers of early education and care have included organisations from the public, private and voluntary sectors, such as nursery schools, nursery classes, infant classes, and reception classes based in primary schools, day nurseries, preschool/playgroups, and child minders (West 2006, OECD 2000). Generally, day care providers offered services throughout the year on a full-time basis, while services based in primary schools were open only during term time and are often part-time. Another important

Camilli et al. 2010). Some using these use the same outcome measure as we do. The study by Jensen et al. (2013), for instance, shows that children who attended high-quality day care developed fewer emotional symptoms and conduct problems, became less hyperactive and were more attentive.

difference concerns child to staff ratios and staff educational qualification requirements. Non-school settings, such as private day nurseries, required lower child-teacher-ratios than reception classes, nursery classes, and nursery schools, where pedagogic staff had higher education levels.

Common across all types of setting was however the Foundation Stage – a standardized curriculum covering all provision for children from age three and a system of regular inspections (West 2006). Many children attend more than one type of day care during their preschool years, and the transition ages and combinations of providers vary. Therefore, we focus on the age at which children started to attend any of these types of day care.

A local system of planning was introduced in 1998 to try and ensure that there was adequate provision to meet demand. Furthermore, a free, part-time entitlement to an early years education place for all 4-year olds was introduced in 1999. All 3-year olds were guaranteed the same free, part-time entitlement to an early years education place by 2004. Parents could top up the free place with additional fee-paying hours in the same or another setting. As part of the Sure Start Programme, additional funding was allocated for extending the provision in disadvantaged neighbourhoods to counteract the large differences in day care attendance rates of different social groups (Bell and Finch 2004). Besides funding going directly to service providers, parents also received financial support. In particular the Childcare Tax Credit (CTC) was paid to the low-paid, working more than 16 hours per week and receiving the Working Family Tax Credit. The CTC covered up to 70 percent of formal childcare costs with a registered provider up to a certain maximum.

Day care attendance rates of children under the age of three were low in the early 2000s. About 20 percent of this age group had access to licensed services, of which most were two-year olds. Younger children were mostly cared for by informal carers or child minders and only few attended private day-nurseries (OECD 2000, OECD 2006). The percentage of three-year olds attending day care rose from 86 percent to 99 percent between 2000 and 2003 (Department for Education and Skills 2003). Full enrolment was reached for all four year olds already in 2000 (Lewis 2003).

Furthermore, attendance rates varied significantly between socio-economic groups. Whereas 7 percent of three- and four-year old children with parents earning less than £ 10,000 did not attend any day care services in 2002, this was the case for only two percent of children whose parents earned £ 30,000 or more (Bell/Finch 2004). For children who had just turned three years, this difference was even larger (23 percent versus 5 percent). Previous studies and parent surveys suggested that the high costs of private provision and limited availability of affordable public provision limited low income families' choices in terms of early years education for their children (West 2006). In 2008, the free-entitlement beginning the term after children's third birthday was extended to 38 weeks of the year, and in 2010 it rose to 15 hours per week. Since 2013, the entitlement has been extended to 2-year-olds looked after by the local authority or from low-income families who would be eligible for free school meals (for details on eligibility, see Brewer et al. 2014).

4 Empirical strategy

Our empirical strategy aims at identifying the causal effect of later day care entry on non-cognitive skills of children at age 5 and 7. In general, estimating the causal effect of later day care entry on children's outcomes raises the missing counterfactual problem.

Equation 1 summarizes the linear relationship of later day care entry (our treatment) on children's non-cognitive skills:

$$Y_i = \beta_1 + \beta_2 \text{TREATMENT}_i + \beta_3 X_i + \nu_i \quad (1)$$

Y_i denotes *child i's non-cognitive skills* and X_i is the vector of conditioning variables.

The *Treatment* variable is defined as a binary measure, which equals 1 if a child enters day care at the age of 31 months or older and 0 if day care is started prior or equal to the age of 2.5 years (i.e. 30 months). This cut-off is driven by recent day care policy developments in England and by the actual distribution in starting age (see Section 5). There has been a continuing debate whether universal access to a day care place granted to children from age three years should be

extended to all, and not just to disadvantaged, two-year-olds (as done since 2013). If such a guarantee for a place were introduced, due to fixed day care starting dates mainly in September and April children would in practice start sometime between age 2 and 2.5 depending on the month of their birth.

The effect of later day care entry may be estimated correctly through OLS if the “selection on observables” assumption is satisfied (see Heckman 1979). This means that all variables which predict both later day care entry and children’s non-cognitive skills must be included in our model. Furthermore, for OLS to consistently estimate the effect of later day care entry, the underlying model must be correctly specified: the assumption that the relationship between later day care entry and a child’s non-cognitive skills is linear or additive is really hard to verify when including several explanatory variables.

In this study we apply propensity score matching methods to solve the missing counterfactual problem. We need to exclude the possibility that a child who starts day care later may have unobserved characteristics that also affect a child’s non-cognitive development. Furthermore parents who decided to send their child to day care later may be systematically different from those who did so earlier. A prevalent method to increase similarity between two groups is propensity score matching (e.g. Blundell/Costa Dias 2000, Rosenbaum/Rubin 1983). Thus, compared to the existing literature on day care, we do not restrict our analysis to compliers of a reform (as done e.g. by Datta Gupta/Simonsen 2010, 2012, Baker et al. 2008 or Kottelenberg/Lehrer 2014).

Propensity score matching assumes that conditional on observable characteristics assignment to treatment and control group is random. To obtain a “random sample”, we estimate a child’s propensity to start day care later than at the age of 2.5 years. Based on a set of relevant characteristics (X_i) we obtain a comprehensive measure of all covariates for each child. Propensity score matching methods enable us to reweight individuals of the control group based on the propensity score to match treatment observations for the average treatment effect of the treated (ATT). Non-matched individuals are dropped from the analysis. We use kernel

matching⁷ to obtain the weight that reweights the control group observations to match the treatment group and to compare nearly identical children either starting day care early or later.

We use propensity score methods to estimate two parameters of interest: the average treatment effect of the treated (ATT) and the average treatment effect of the untreated (ATU). The ATT measures the difference between the average outcome for children who enter later and the average outcome assuming they had not entered later: $ATT = E(Y_1 | D=1) - E(Y_0 | D=1)$, whereas the ATU in our paper compares the difference in outcomes assuming the children had entered day care earlier: $ATU = E(Y_1 | D=0) - E(Y_0 | D=0)$.

Equation 2 depicts the average treatment effect of the treated (ATT) using a regression-adjusted matching approach⁸:

$$ATT = \sum_{i \in T} W_i [(Y_{1i} - x_i \hat{\beta}) - \sum_{j \in C} W_{ij} (Y_{0j} - x_j \hat{\beta})] \quad (2)$$

In Equation 2, W_{ij} is the weight placed on *individual j* (of the control group) to be comparable to *individual i* (of the treatment group). The weight W_{ij} includes values obtained from kernel matching.

However, in order to causally interpret our estimations, the method assumes that there are no unobserved variables that simultaneously influence children's socio-emotional behaviour and the probability of starting day care later. Meaning that in the absence of variation in the age of entry in day care, the non-cognitive skills of treated children and matched controls would be identical. If this unconfoundedness assumption is violated, i.e. if children who are treated differ systematically from children who are not treated in terms of unobservable characteristics, our model suffers from endogeneity. To make the unconfoundedness assumption more plausible,

⁷ Matching is implemented in Stata 13 using the program `psmatch2` provided by Leuven and Sianesi (2003). We use kernel matching with an Epanechnikov kernel function and a bandwidth parameter of 0.06.

⁸ This regression-adjustment decreases the standard errors of the treatment effect estimates, since the weighting leads to mean-independent treatment (see Stuart 2010).

we include measures⁹ of children's cognitive and non-cognitive skills at the age of 9 months, which might influence both a child's socio-emotional behaviour and starting day care later.

For the implementation of matching, we predict a child's propensity to enter day care later than age 2.5 years. Since our day care definition covers all care arrangements outside the home of the child, there is less risk of bias due to the potential correlation between different types of day care, age at entry and child's non-cognitive skills than in other studies which, for instance, differentiate between family day care and preschool with different average starting ages. Moreover, the estimations include child-related characteristics, such as child's temperament and child's skill development at the age of 9 months, as covariates to ensure that our measure of age at entry in day care is unaffected by child behaviour.

5 Data

In order to analyse the effect of our treatment (later entry into day care), we use data from the Millennium Cohort Study (MCS). The Millennium Cohort Study is a representative sample of births covering children born in the UK between September 2000 and January 2002 (Hansen 2012). The sample is clustered geographically and disproportionately stratified to over-represent: (1) the three smaller countries of the UK – Wales, Scotland and Northern Ireland; (2) areas of high child poverty; (3) within England only, areas with high concentration of ethnic minorities (Plewis 2007). The first wave of MCS occurred when children were aged about nine months, gathering information from the parents of 18,818 children, about themselves and the babies (e.g. problems during pregnancy, birth weight). Since then, families have been interviewed again five times. Here we use data from the first four waves, which took place when the children were aged 9 months, three years, five years (thus entering compulsory schooling) and seven years.

We restrict our sample to children living in England, so that the childcare supply is similar for all children observed (see section 3). Furthermore, we keep in our sample only

⁹ Figure 2 shows that after propensity score matching all conditioning variables have a standardized bias within the -5 percent and 5 percent interval. Children's non-cognitive skills have a standardized bias close to 0 percent suggesting that they are nearly identical in the treatment and control group after matching.

children who are not twins or triplets, whose mothers responded to the survey and who have valid information on non-cognitive skills at wave 3 and 4 and on their age at day care entry. This reduces the sample to 6,460 children. Furthermore we restrict our sample to children who have non-missing information on children's early skill measures 'temperament of the child': 5,448 children have non-missing information on this conditioning variable. The drop in sample size is driven by item non-response¹⁰. Predicting the propensity of later entry based on this skill measure and other relevant conditioning variables (see description below) leaves us with 4,825 observations in the matched sample.

Day care entry measures. To investigate the effect of later day care entry, we use information on the year of first day care attendance provided by mothers. In the MCS, the day care history is surveyed in wave 2 and, more comprehensively, in wave 3 (when children are 3 and 5 years old, respectively). Based on the provider type mothers mentioned first, the majority of children, 51 percent, attended nursery schools or nursery classes, 15 percent a play group, 16 percent a preschool, 3 percent a child minder, and the remaining 15 percent were cared for by day nurseries or some other than the before mentioned institutions. We consider all these types of care as day care, including preschools, as in England preschool education and day care in a more narrow sense are intertwined. Moreover the information collected from mothers about the type of day care is insufficient to distinguish systematically between all relevant settings with varying quality dimensions. Therefore our treatment covers a wide variety of extra-familial care arrangements, which are subject to some quality regulations (see section 3) (for the same approach in the UK context, see e.g. Goodman/Sianesi 2005 and Apps et al. 2013). From a methodological point of view this definition allows us a large enough sample of children in day care and a clear distinction between treatment and control group. Unfortunately we do not have any further information about day care quality. Thus we are clearly estimating an effect of average exposure. Moreover, we focus on age at entry and do not take into account different intensity of day care attendance, i.e. full-time or part-time care.

¹⁰ In our sensitivity analysis section we report results for "relaxed" sample restrictions and remove 'temperament of the child' as conditioning variable from our estimations. By doing this, we infer if item non-response influences our estimations.

The treatment group comprises children who enter day care later, namely later than the age of 2.5 years. This treatment cut-off reflects the distribution of entry age in our sample with 42 percent of children entering day care prior or equal to the age of 2.5 years and 58 percent of children entering later (see Figure 1).

[Figure 1 about here]

The control group consists of children who attend day care early, namely prior or equal to the age of 2.5 years. Similar to the treatment group children have to have valid non-cognitive skill information and their mothers have to have participated in the survey in wave 1 to wave 4. The control group does not include children who do not participate in day care, as all children in England are attending day care at some point in time; at the latest before they enter school (see section 4).

Measures of non-cognitive skills. To measure the outcome of interest, the socio-emotional behaviour of children, the Millennium Cohort Study (MCS) includes the Strength and Difficulties Questionnaire (SDQ), proposed by Goodman (1997). The SDQ is used to collect information on the socio-emotional behaviour of children at age three, at age five, and at age seven. Research suggests that the SDQ and the well-known Rutter questionnaires (Rutter 1970) correlate highly and do equally well in terms of classifying behaviour (see Goodman 1997). The SDQ is reported to have high test-retest reliability and good validity, and has been used in other large epidemiological, educational, psychological and, as indicated above, in a few economic studies.

We use measures of the SDQ at two points in time: at age five and seven. The SDQ is part of a self-completion module filled out mainly by mothers. It consists of statements to which the responses are: ‘not true’, ‘somewhat true’, and ‘certainly true’. The socio-emotional behaviour is gathered through 25 items over five separate dimensions respectively subscales: Conduct Problems, Hyperactivity/Inattention, Emotional Symptoms, Peer Relationship Problems, and Prosocial behaviour. Higher scores on the first four of these subscales and lower

scores on the pro-social subscale indicate greater problems. A ‘total difficulties’ score is generated by summing the first four scales and excluding the prosocial scale, which can be used as a positive counter measure to the overall SDQ score. The resultant score ranges from 0 to 40, the subscales from 0 to 10. For our analysis we use the overall SDQ score as well as its dimensions to cover various aspects of a child’s socio-emotional behaviour. Moreover, the overall SDQ score can also be divided in two scales of externalising and internalising behaviour¹¹. In the estimations, the SDQ and all subscales are standardized to have a mean of zero and a standard deviation of one. In addition, we apply the SDQ instructions¹² regarding bandings for categorization as ‘normal’, ‘borderline’ or ‘abnormal’ in order to identify children who are likely ‘cases’ with mental health disorders¹³. We collapse the categories ‘borderline’ and ‘abnormal’ behaviour into one. In our sample 11 percent of children are classified ‘borderline/abnormal’ at age 5 and 14 percent of children at age 7.

Mothers’ assessments of children’s behaviour include a risk of measurement error. However, as pointed out by Datta Gupta and Simonsen (2010), even if mothers’ responses are biased, as long as this is unrelated to early or later day care entry, it will not cause problems for our identification strategy. Since mothers are not asked to evaluate child behaviour directly in the context of day care entry but instead are likely to use other children at primary school age as the comparison group for their assessments¹⁴, it is less likely that biases in mothers’ responses will vary systematically with day care entrance age.

Our measure of socio-emotional behaviour has been applied in many other studies, among them a few by economists (see, e.g. Andersen et al. 2007, Ermisch 2008, Datta-Gupta/Simonsen 2010 and 2012, Ermish et al. 2012, Dearden et al. 2010, and Müller et al. 2013) as well as by researchers from other disciplines using the UK Millennium Cohort Study

¹¹ The externalising score ranges from 0 to 20 and is the sum of the conduct and hyperactivity dimensions. The internalising score ranges from 0 to 20 and sums the emotional and peer problems scales.

¹² See www.sdqinfo.com (download: October 2014) for further details.

¹³ Children whose overall SDQ score ranges from 0 to 13 are categorized ‘normal’, children whose overall SDQ score is greater or equal to 14 and less than 17 are categorized ‘borderline’, and children whose overall SDQ score is greater or equal to 17 fall into the category ‘abnormal’.

¹⁴ For a detailed overview of the dimensions and questions, see Table A1 in Appendix.

(e.g. Griffiths et al. 2011, Heikkilä et al. 2011, Kelly et al. 2011, McMunn et al. 2012). However, most studies mainly focus on the total SDQ without differentiating which dimension(s) might be driving the overall effect.

Conditioning variables. The empirical strategy relies on the assumption that the variables to predict day care entry and children's non-cognitive skills are observed. Therefore the set of variables used to condition on children's propensity to enter day care later is crucial for the identification strategy. The choice of conditioning variables is based on other empirical studies¹⁵ investigating day care (entry) effects on child outcomes. These variables are among others children's skills and other characteristics around birth (e.g. birth weight, temperament and a measure for skill development at 9 months), maternal characteristics that determine maternal preference for day care (e.g. employment status, education, ethnic background, age at birth, mental health and educational attitudes) and household characteristics (e.g. household type, number of siblings and poverty status). The majority of our conditioning variables are measured at wave 1. Table A2 in the Appendix summarises the variables.

Descriptive Statistics. A first descriptive comparison between treatment and control group shows that a child's socio-emotional behaviour differs by age at entry (see Table 1). Overall 58 percent of the children attend day care later than at the age of 2.5 years. The mean of children's overall SDQ score at age 5 and at age 7 is 0.6¹⁶ score points lower for children who enter day care early. This difference is statistically significant and hinges towards a potential negative effect of later day care entry on children's socio-emotional behaviour: the higher the SDQ, the more socio-emotional problems a child has. In general, Table 1 shows that children who enter later have a higher score on all SDQ dimensions, i.e. have on average relatively more conduct problems, peer problems, emotional problems or hyperactivity. This negative relationship between later day care entry and children's non-cognitive skills is found at age 5 and descriptively persists until age 7. These mean differences between treatment and control group show up for all sub-dimensions of the socio-emotional behaviour. The differences are

¹⁵ For other studies see Section 2 of this paper.

¹⁶ This translates into a difference between treatment and control group of 11% of a standard deviation.

higher for peer problems and relatively small for emotional problems. For the mean differences between the treatment and the control group based on the matched sample, see Table A3 in the Appendix. In addition, we have run appropriate tests to contrast covariate mean differences across our matched groups to ensure that adequate matches have been obtained. Results from the balancing tests are shown in Figure 2 and in Table A4 in the Appendix. They show that balancing children of the control group to match children of the treatment group was successful. We obtain a large reduction in the standardised percent bias, which is the percent difference of the sample means in the treatment and the matched control sample as a percentage of the square root of the average of the sample variances in both groups. Figure 2 depicts that the means of nearly every conditioning variable lie on the zero percentage line, and definitely in the interval of -5 and +5 percent of standardized bias. In addition, Figure 3 shows the kernel density estimate of the obtained propensity score for treatment and control group. Both curves show an extensive overlap between the treatment and control groups.

[Table 1, Figures 2 and 3 about here]

Heterogeneous effects. To examine heterogeneous effects of starting day care late, we differentiate our analysis by gender of the child, by maternal education, by a poverty threshold and by combining gender and education as well as gender and the poverty threshold. We use information on maternal education based on the national vocational qualification framework, and distinguish between “low education” and “high education”.¹⁷ The poverty subgroup utilizes a measure taken directly from the Millennium Cohort Study. This measure is based on the equivalised household income according to the OECD equivalence scale. Households are defined as poor if their equivalised household income is less than 60 percent of median household income. All indicators for the subgroups are measured at wave 1 to ensure pre-

¹⁷ In our analysis we define “low education” if mothers have an educational level below A-levels, i.e. “none of these qualifications”, “GCSE grades A-G”, or “O-levels”. Mothers are categorised “high educated” if their educational level is equal to or above A-levels, i.e. “A-levels”, “diplomas in higher education”, “first or higher degree”.

treatment information, as we perform propensity score matching for each heterogeneous effect separately. Matching within the subgroups ensures that we compare each treatment effect “subgroup-specific”, e.g., girls who enter late are matched to nearly identical girls who enter early based on the set of conditioning variables¹⁸.

6 Results

This section presents the effect of later day care entry using OLS and propensity score methods; we mainly interpret the results of our preferred specifications using propensity score matching.

Table 2 presents the results of later day care entry (entry > 2.5 years of age) affecting children’s non-cognitive skills, namely the socio-emotional behaviour at age 5 (columns 2-4) and at age 7 (columns 6-8). First of all a comparison between the mean difference estimates (OLS) and the results using propensity score matching (PSM) shows that the mean differences are larger in respect to the size of the effects than PSM estimates controlling for the set of conditioning variables. This suggests that not taking into account the selection effect of later entry age may overestimate the actual impact of age at entry on children’s non-cognitive skills.

Later day care entry significantly increases children’s socio-emotional problems (SDQ) in the medium run (at age 7) but not in the short run (at age 5). In the short run (at age 5) entering day care later only impacts children’s internalising score, mainly through greater peer problems (increase by 8 percent of a standard deviation), and their probability of having borderline or abnormal behaviour. A child’s SDQ at age 7 increases by 6.7 percent of a standard deviation if she enters day care later than at the age of 2.5 years. This ‘lifts’ the average child closer to ‘borderline or abnormal’ behaviour. A later entry in day care results in a larger externalising score, mainly due to a rise in hyperactivity. A later entry also increases the internalising score by 6.6 percent of a standard deviation, reflecting in particular the increase of peer problems. Children who enter later also behave less prosocial at age 7.

¹⁸ In each subgroup, we compare the covariate distribution across our matched groups to ensure that adequate matches have been obtained. Results from the balancing tests for each subgroup are available from authors upon request.

As shown by the ATU specifications, the effect of *late entry* on socio-emotional behaviour of children in the control group at age 5 and at age 7 is at most slightly smaller with respect to the size of the effects and their statistical significance but shows the expected relationships.

[Table 2 about here]

Given different developmental patterns between girls and boys (see e.g., Heckman et al. 2013), in particular for their socio-emotional behaviour, we split the *sample by gender* (see Table 3). At age 5, late day care entry increases peer problems and the likelihood of borderline or abnormal behaviour for both boys and girls. Some of these effects however are only significant at the 10 percent level. Moreover, the prosocial behaviour of boys decreases by 9 percent of a standard deviation if they enter day care later.

At age 7, we find large and significant effects for boys of a later entry age on their socio-emotional behaviour, whereas later entry does not seem to affect girls in the medium-term. The overall SDQ score of boys increases by almost 11 percent of a standard deviation if they enter later. The results for the different subscales show increases in externalising behaviour with significant effects on conduct problems and hyperactivity. The increase in the internalising score is mainly driven by the large effect of late day care entry on greater peer problems for boys. Moreover, boys who enter later are less prosocial and have a higher probability of being classified as borderline or abnormal. The average treatment effects of the untreated render almost the same picture in the short-run as the ATT. The children who enter earlier would not benefit from a later entry in respect to their socio-emotional behaviour.¹⁹

[Table 3 about here]

¹⁹ Only for girls, there is one additional marginally significant effect on their internalising behaviour, which is less pronounced if they enter later.

In a next step, we differentiate our analysis by two *additional subgroups*, by the education of the mother and by poverty status. Other studies have shown that in particular children from disadvantaged families benefit from earlier entries (see above, for a summary of empirical studies, see Ruhm/Waldfogel 2011). Thus, if treatment effects are heterogeneous, we will not expect the parameters in Table 2 to be representative for all groups²⁰. Table 4 comprises the estimates of the ATT at age 5 differentiated by maternal education and by the 60 percent threshold of median household income. In addition, Table 4 also depicts the subgroup analysis for boys and girls separately.

In the short run at age 5, the results show that children in poor households are more likely to be hyperactive than their peers who entered earlier. The difference amounts to 19 percent of a standard deviation. This effect is found for boys and girls. The peer problem effect, however, is triggered by children of high educated mothers - interestingly by boys and girls of such mothers. For girls of poor households, however, the results show that once they enter later they have more peer problems as well. Their overall SDQ score increases once they enter later.

Furthermore, it is remarkable that the prosocial behaviour effect is mainly driven by boys with low educated mothers. This provides support that a social gradient in the effect of day care entry age exists, however the direction differs by the dimension of the socio-emotional behaviour.

[Table 4 about here]

In the medium-term (see Table 5), we find more consistently adverse effects of later day care entry on most SDQ dimensions for disadvantaged groups than in the short-term. Overall, the effects found at age 5 for children from low educated mothers and for children living below the poverty threshold mostly persist and become larger. For children from higher educated mothers and households above the poverty line, the effects mostly disappear. Their socio-

²⁰ In these tables we restrict the various measures for the socio-emotional behavior to the overall score and the five subscales. Moreover we only present the average treatment effects on the treated, as the ATT is our preferred specification.

emotional behaviour at the age of 7 seems to be unaffected by the age of day care entry, with one exception²¹. All other effects only apply to the group of disadvantaged children: Their overall SDQ score is much higher when they enter day care later. This effect is more pronounced for children from poor households than for children from low educated mothers (22 respectively 12 percent of a standard deviation). The negative effect of day care for boys with low educated mothers has also been found in other studies. For example, Datta Gupta and Simonsen (2010) observe that boys born to mothers with a high school degree or below experience an increase in the SDQ by 2.1 points if they are not taken care of at home. Furthermore, Loeb et al. (2007) find that the negative effects for children who enter earlier fade out for some groups of disadvantaged children.

In contrast to the results at age 5, at the age of 7 disadvantaged children who enter day care later have significantly more peer problems. Remarkable is that once we restrict the sample to disadvantaged boys the effects of later day care entry appear to impact all dimensions of children's socio-emotional behaviour. Boys with low educated mothers have greater conduct problems, are more likely to be hyperactive, and display more emotional and peer problems when they enter day care later. By contrast, the earlier effects of disadvantaged girls fade out in the medium-term.

[Table 5 about here]

7 Sensitivity Analysis

First we test if the results are sensitive to our age of entry cut-off age. We test an alternative cut-off which is solely data driven, the median age of entry, which is 2.7 years of age. Table 6 shows that the results remain the same.²² In addition we test if our results are robust regarding outliers in the age of entry distribution – at the lower and the upper bound. To

²¹ When these children enter day care later, they are less prosocial. This effect is driven by boys of this group, whose prosocial behaviour decreases by almost 15 percent of a standard deviation.

²² Only for the prosocial behaviour at age 5 and for hyperactivity at age 7 the statistical significance decreases from the 5-percent to 10-percent-level.

test whether children who enter day care very early, i.e. during the first year of their life, or relatively late, i.e. between age 49 months and 60 months might bias our analysis, we re-estimate the treatment effect excluding these two small groups of children in separate steps. The results of these tests show that limiting the control group to those entering after 12 months or restricting the treatment group to entries between the age of 31 months and 48 months respectively, does not affect our estimates (see Table A5).

[Table 6 about here]

Furthermore we test if our sample restriction to valid observations for children's temperament at the age of 9 months or poverty threshold information – both confounding variables have particularly high non-response – might bias our results. Excluding and not matching on these variables increases our sample to 5,568 and 5,934, respectively, and renders similar results. As shown in Table A6 in the Appendix, the only substantive change is that later entry is less influential for prosocial behaviour at age 5 and the significance level of its effect decreases.

Since the Millennium Cohort Study is clustered geographically and disproportionately stratified, we further check the robustness of our results by estimating the propensity to enter day care later using the sampling weights provided by the MCS. To consider these sampling weights, we use propensity score weighting (Hirano/Imbens 2001)²³. Compared to the treatment effect obtained from propensity score matching in Table 2, the size of the estimates of the average treatment effect on the treated (ATT) is slightly larger but remains significant at the same significance level. We also re-estimate the probit regression of entering day care after 2.5 years of age using the sampling weights of the MCS. The results are robust and similar in size to whether we match conditioning on stratification criteria or whether we use the country specific weights using propensity score weighting.

²³ A critical aspect of using the estimated propensity score as weight is its sensitivity to large estimated propensity scores which, however, decreases with sample size. By restricting the post-estimations to the common support area, the problem of “large propensity score values” should have only a minor impact.

8 Conclusion

In this paper, we investigate the causal effects of entering day care later than age 2.5 years compared to earlier on the socio-emotional behaviour of children as one particular measure of non-cognitive skills. A growing number of studies provide evidence of the importance of non-cognitive skills for educational, health, and labour market outcomes in the longer run. We contribute to the literature on how children's early learning and care experiences in day care centres affect their non-cognitive skills. We extend previous studies by not simply focussing on the effects of day care in general, but on the age of entry. Furthermore, whereas the few previous studies using the same skill measures mostly used one composite score, we analyse several dimensions of socio-emotional behaviour to gain a more in-depth understanding of the underlying mechanisms. We also extend the existing literature by performing detailed analysis of social gradients in the effects of later day care entry on children's development considering moderating factors of child gender, maternal education, and household income. Furthermore we focus on short- and mid-term effects, while most other studies only focus on a measure at one point in time. We use Ordinary Least Squares and Propensity Score Matching - given the richness of the Millennium Cohort Study; we match on a rich set of confounding variables, covering for instance children's earlier skills and birth weight, mothers' parenting attitudes, health, and educational level, and household characteristics.

Across the whole sample, we find that a day care entry past age 2.5 years adversely affects the socio-emotional behaviour of children at age 7. The average treatment effect on the untreated confirms that at the age of 7 children benefited from an earlier entry. Looking at the SDQ sub-dimensions reveals that the effects are mainly driven by relatively large effects of later day care entry on prosocial behaviour and peer problems. Although we find no overall effect at age 5, a detailed analysis of the different SDQ dimensions shows that children who enter day care later are less prosocial and have more peer problems already at age 5. In general, these results for the full sample are in line with the results found by studies mainly based on German data. Our effects sizes at age 7 are very similar to what Felve and Lalive (2014) find for the

socio-emotional maturity of children of less educated mothers in Germany. The estimates differ to the results of Datta-Gupta and Simonsen (2010 and 2012), who do not find any significant effects for their overall sample in Denmark. Baker et al. (2008) find negative effects; yet the latter use a different measure of non-cognitive skills and focus on a reform in Quebec, which included incentives for full-day use of care.

In respect to social gradients in the effect of day care entry age, late entry mainly seems to have adverse medium-term effects on boys from disadvantaged families in terms of low maternal education levels or household income below the poverty threshold. In the short-term at age 5, we find more diverse effects on different groups. By the age of 7, however, the effects found at age 5 for boys from low educated mothers and income-poor households mostly persist or become larger. By contrast, they are mostly non-significant for girls across all groups and for boys from higher educated mothers and households above the poverty line. These findings are in line with some previous studies and indicate that boys from lower socio-economic groups may benefit most from entering day care before age 2.5 years (see section 2). Thus our study also provides further evidence of the importance to focus on the heterogeneity of the effects on different groups of children.

In general our results provide support for the hypothesis that day care institutions provide important environments where children can learn to interact with each other. Entering day care relatively late increases the probability of them developing more peer-problems and being less prosocial. Our results indicate, that day care entry before or after age 2.5 years is less strongly related to developing emotional or conduct problems. By analysing several developmental dimensions, our study is able to pinpoint more in detail than previous studies how the day care entry age impacts children's formation of non-cognitive skills.

It should be noted that our measure of early entry mainly covers children entering day care after their first birthday and thus does not focus on children entering very early for whom more negative effects of long hours of day care have been found (e.g. Loeb et al. 2007 or for a non-economic study Belsky et al. 2007). Another limitation of our study is that it only follows children until the age of 7. Further research is needed to see how long the effects last. Thus we

agree with Kottelenberg and Lehrer (2014: 360) who suggest that 'more work is needed to explore differences in the timing at which treatments are received and investments are made, and how they interact with child characteristics'.

As most other economic studies (with the exception of Bauchmüller et al. 2014 or Müller et al. 2013), we cannot control for the quality of the day care institutions which the children attended, which according to educational studies is most important for the skill formation of children (e.g., Vandell et al. 2010, Anders et al. 2011, Sammons et al. 2008). We also did not consider the daily intensity of children's attendance at day care institutions, which has an influence as well (e.g. Loeb et al. 2007 or DeCicca 2007). Future research should therefore aim to consider more detailed day care histories covering aspects of entry age, quality, and daily intensity over the children's preschool years.

Provided that future research confirms our findings also for children's longer-term development, this would emphasize the importance of public investments in earlier day care in England. This seems justified as improved social skills of children have been shown to benefit not only the individual child but also society as a whole in the long-term (Heckman et al. 2010). Given our findings of stronger benefits for children from disadvantaged families, the recent extension of the free day care entitlement to 2-year-olds in low income families (for details on eligibility see Brewer et al. 2014) seems like a promising investment.

References

- Anders, Y., P. Sammons, B. Taggart, K. Sylva, E. Melhuish, I. Siraj-Blatchford (2011), “The influence of child, family home factors and pre-school education on the identification of special educational needs at age 10”. *British Educational Research Journal*, **37**, 421-441.
- Andersen, A., M. Deding and M. Lautsen (2007), “How Much Does Parental Divorce Affect Children’s Well-being?”, SFI Working Paper.
- Apps, P., S. Mendolia and I. Walker (2013), “The Impact of pre-school on adolescents' outcomes: Evidence from a recent English cohort”, *Economics of Education Review*, **37**, 183-199.
- Baker, M. (2011), Innis lecture: “Universal Early Childhood Interventions: What is the Evidence Base?”, *Canadian Journal of Economics*, **44**, 1069-1105.
- Baker, M., J. Gruber and K. Milligan (2008), “Universal child care, maternal labour supply, and family well-being”, *Journal of Political Economy*, **116**(4), 709-745.
- Bauchmüller, R., Gørtz, M. and A. Würtz Rasmussen (2014), “Long-Run Benefits from Universal High Quality Pre-Schooling”, *Early Childhood Research Quarterly*, **29**, 457-470
- Bell, A. and S. Finch (2004), “Sixth survey of parents of three and four year old children and their use of early years services”, *DfES Research Report*, **525**, London, DfES,
- Belsky, J., D. Vandell, M. Burchinal, K. A. Clarke-Stewart, K. McCartney, M. Owen, and the NICHD Early Child Care Research Network (2007), “Are there long-term effects of early child care?”, *Child Development*, **78**, 681-701.
- Berger, L. M., Hill, J. and J. Waldfogel (2005). “Maternity leave, early maternal employment and child health and development in the US”, *Economic Journal* **115**(501), F29-F47.
- Bingley, P. and N. Westergaard-Nielsen (2012), “Intergenerational Transmission and Day Care”, in J. Ermisch, M. Jäntti and T. Smeeding, eds., *From Parents to Children: The Intergenerational Transmission of Advantage*, Russell Sage Foundation, New York, 190-203.
- Blanden, J., P. Gregg and L. Macmillan (2007), “Accounting for intergenerational income persistence: non-cognitive skills, ability and education.” *Economic Journal*, **117**, C43-C60.
- Blundell, R. and M. Costa Dias (2000), “Evaluation methods for non-experimental data”, *Fiscal Studies* **21**(4), 427-468.

- Brewer, M., S. Cattan and C. Crawford (2014), “State support for early childhood education and care in England”, in C. Emmerson, P. Johnson and H. Miller eds., *The IFS Green Budget*, 172-197, London, Institute for Fiscal Studies.
- Brilli, Y., Del Boca, D. and C. Pronzato (2011). “Exploring the impacts of public childcare on mothers and children in Italy: Does rationing play a role?”, Bonn: *IZA Discussion Paper* **5918**.
- Burger, K. (2010), “How does early childhood care and education affect cognitive development? An international review of the effects of early intervention for children from different social backgrounds”, *Early Childhood Research Quarterly* **25**(2), 140-165.
- Camilli, G., S. Vargas, S. Ryan and W.S. Barnett (2010), “Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development”, *Teachers College Record* **112**(3), 579-620.
- Carneiro, P., C. Crawford and A. Goodman (2007), “The Impact of Early Cognitive and Non-cognitive Skills on Later Outcomes”. London: *Centre for the Economics of Education Discussion Paper* **92**.
- Cunha, F., J. Heckman and S. Schennach (2010), “Estimating the Technology of Cognitive and Noncognitive Skill Formation”, Cambridge MA: *NBER Working Paper* **15664**.
- Cunha, F. and J. J. Heckman (2007), “The technology of skill formation.” *American Economic Review Papers and Proceedings*, **97**, 31-47.
- Cunha, F., J. J. Heckman, L. Lochner and D. V. Masterov (2006), “Interpreting the Evidence on Life Cycle Skill Formation”, Hanushek, Eric A. and F. Welch, eds., *Handbook of the Economics of Education*, Volume 1, Chapter 12.
- Currie, J. and M. Stabile (2006), “Child mental health and human capital accumulation: The case of ADHD”, *Journal of Health Economics* **25**(6), 1094-1118.
- Datta Gupta, N. and M. Simonsen (2010), “Non-cognitive child outcomes and universal high quality child care”, *Journal of Public Economics* **94**(1-2), 30-43
- Datta Gupta, N. and M. Simonsen (2012), “The effects of type of non-parental child care on pre-teen skills and risky behaviour”, *Economic Letters* **116**(3), 662-625.
- Dearden, L. et al. (2010), “The socio-economic gradient in early child-outcomes : evidence from the Millennium Cohort Study”, *IFW Working Paper*, **W11/03**.
- DeCicca, P. (2007): “Does full-day kindergarten matter? Evidence from the first two years of schooling”, *Economics of Education Review* **26**, 67-82.
- DeCicca, P. and J. Smith (2013), “The long-run impacts of early childhood education: Evidence from a failed policy experiment”, *Economics of Education Review* **36**, 41-59.
- Department for Education and Skills (2003), “Provision for children under five years of age in England-January 2003”, London, DfES.

- Dumas, C. and A. Lefranc (2010), “Early schooling and later outcomes: Evidence from pre-school extension in France”, Université Cergy-Pontoise: *THEMA Working Papers* **2010-07**.
- EACEA (2009): “Early Childhood Education and Care in Europe: Tackling Social and Cultural Inequalities”, 197-181, http://eacea.ec.europa.eu/education/eurydice/documents/thematic_reports/098EN.pdf: Education, Audiovisual and Culture Executive Agency.
- Ermisch, J. (2008), “Origins of Social Immobility and Inequality: Parenting and Early Child Development”. *National Institute Economic Review*, **205**, 62-71.
- Ermisch, J., F. H. Peter and C. K. Spieß (2012), “Early Childhood Outcomes and Family Structure”, in J. Ermisch, M. Jäntti and T. Smeeding, eds., *From Parents to Children: The Intergenerational Transmission of Advantage*, Russell Sage Foundation, New York, 120-139.
- Felfe, C. and R. Lalive (2014), “Does Early Child Care Help or Hurt Children’s Development?”, *IZA Discussion Paper* **8484**.
- Fitzpatrick, M. (2008), “Starting school at four: The effect of universal pre-kindergarten on children’s academic achievement.” *The B.E. Journal of Economic Analysis and Policy*, **8**(1), 1-40.
- Griffiths, L. J. et al. (2011), “Is obesity associated with emotional and behavioural problems in children? Findings from the Millennium Cohort Study”, *International Journal of Pediatric Obesity*, **6**, e423-e432.
- Goodman, R. (1997), “The strengths and difficulties questionnaire: A research note”, *Journal of Child Psychology and Psychiatry*, **38**(5), 581–586.
- Goodman, A. and B. Sianesi (2005), “Early Education and Children’s Outcomes: How Long Do the Impacts Last?”, *Fiscal Studies* **25**(4), 513-548.
- Hansen, K. (2012), “First, Second, Third and Fourth Surveys: A Guide to the Datasets (Seventh Edition)”, London: *Centre for Longitudinal Studies*.
- Hansen, K. and D. Hawkes (2009), “Early childcare and child development”, *Journal of Social Policy*, **38**, 211–239.
- Havnes, T. and M. Mogstad (2011), “No Child Left Behind: Subsidized Child Care and Children’s Long-Run Outcomes”, *American Economic Journal: Economic Policy*, **3**(2), 97-129.
- Heckman, J. J., R. Pinto and P. Savelyev (2013), “Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes”, *American Economic Review*, **103**(6), 2052-2086.
- Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev and A. Yavitz (2010), "The rate of return to the HighScope Perry Preschool Program", *Journal of Public Economics*, **94**, 128.
- Heckman, J. J. (2008), “Schools, Skills and Synapses”, *Economic Inquiry*, **46**(3), 289-324.

- Heckman, J. J., Stixrud, J. and S. Urzua (2006), “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior”, *Journal of Labor Economics*, **24**(3), 411-482.
- Heckman, J. J. (1979), “Sample Selection Bias as Specification Error”, *Econometrica*, **47**(1), 153-161.
- Heikkilä, K. et al. (2011), “Breast feeding and child behaviour in the Millennium Cohort Study”, *Archives of Disease in Childhood*, **96**, 635-647.
- Hirano, K. and G. Imbens, (2001), “Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization”, *Health Services and Outcomes Research Methodology* **2**(3-4), 259-298.
- Jensen, B. et al. (2013), “Effectiveness of a Danish early year preschool program: A randomized trial”, *International Journal of Educational Research*, **62**, 115-128.
- Kelly, Y. et al. (2011), “What role for the home learning environment and parenting in reducing the socioeconomic gradient in child development? Findings from the Millennium Cohort Study”, *Archives of Disease in Childhood*. **96**(9), 832-837.
- Kottelenberg, M. J. and S. F. Lehrer (2014), “Do the Perils of Universal Childcare Depend on the Child’s Age?”, *CESifo Economic Studies*, **60**(2), 338-365.
- Land, H. (2002), "Meeting the child poverty challenge: Why universal childcare is key to ending child poverty", *Facing the Future Policy Papers* **3**, London, Daycare Trust.
- Leuven, E. and B. Sianesi (2003), “PSMATCH2: Stata module to perform full mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing”. <http://ideas.repec.org/c/boc/bocode/s432001.html>. This version 4.0.10 10feb2014.
- Lewis, J. (2003), "Developing early years childcare in England, 1997-2002: the choices of (working) mothers", *Social Policy and Administration*, **37**, 219-238.
- Loeb, S., M. Bridges, D. Bassok, B. Fuller and R. W. Rumberger (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., C. Ruhm, and J. Waldfogel (2007), “Does prekindergarten improve school preparation and performance?”, *Economics of Education Review* **26**(1), 33–51.
- McMunn, A. et al. (2012), “Maternal employment and child socio-emotional behaviour in the UK: longitudinal evidence from the Millennium Cohort Study”, *Journal of Epidemiology and Community Health*, **66**(7), 19-25.
- Moss, P. and M. O’Brien (2006), “International review of leave policies and related research 2006”, London, *DTI Employment Relations Research Series*.
- Müller, K. U., C. K. Spieß, C. Tsiasioti and K. Wrohlich et al. (2013), “Evaluationsmodul: Förderung und Wohlergehen von Kindern”, DIW Berlin: Politikberatung kompakt Nr. **73**.

- OECD (2014), OECD Family Data Base, <http://www.oecd.org/social/family/database.htm>
- OECD (2006), "United Kingdom (England). Starting strong II: Early childhood education and care", pp. 415-425, Paris, OECD.
- OECD (2000), "OECD Country Note: Early Childhood Education and Care Policy in the United Kingdom", Paris, OECD.
- Plewis, I. (2007), "The Millennium Cohort Study: Technical report on sampling", London: Centre for Longitudinal Studies.
- Prevoo, T. and B. ter Weel (2013), "The Importance of Early Conscientiousness for Socio-Economic Outcomes: Evidence from the British Cohort Study", Bonn: *IZA Discussion Paper 7537*.
- Ringen, S. (1997), "Great Britain", in S. B. Kamerman and A. J. Kahn eds., *Family Change and Family Policies in Great Britain, Canada, New Zealand, and the United States*, pp. 29-102, Oxford: Clarendon Press.
- Rosenbaum, P. and D. Rubin, (1983), "The central role of the propensity in observational studies for causal effects. *Biometrika* **70**(1), 41-55.
- Ruhm, C. and J. Waldfogel (2011), "Long-Term Effects of Early Childhood Care and Education", *IZA DP 6149*.
- Ruhm, C. (2008), "Maternal employment and adolescent development", *Labour Economics* **15**(5), 958-983.
- Rutter, M., Tizard, J. and Whitmore, K. (1970), *Education, Health and Behaviour*. London: Longmans
- Sammons, P., Y. Anders, K. Sylva, E. Melhuish, I. Siraj-Blatchford, B. Taggart, and S. Barreau (2008), "Children's cognitive attainment and progress in English primary schools during Key Stage 2. Investigating the potential continuing influences of pre-school education", *Zeitschrift für Erziehungswissenschaft*, **10**, 179-198.
- Schlotter, M. and L. Wößmann (2010), "Frühkindliche Bildung und spätere kognitive und nicht-kognitive Fähigkeiten: Deutsche und internationale Evidenz", *Vierteljahrshefte für Wirtschaftsforschung* **79**(3), 99-120.
- Stuart, E. (2010), "Matching methods for causal inference: A review and a look forward", *Statistical Science* **25**(1), 1-21.
- Vandell, D., J. Belsky, M. Burchinal, L. Steinberg, N. Vandergrift and the NICHD Early Child Care Research Network (2010), "Do effects of early child care extend to age 15 years? Results from the NICHD Study of Early Child Care and Youth Development", *Child Development*, **81**, 737-756.
- West, A. (2006), "The pre-school education market in England from 1997: quality, availability, affordability and equity", *Oxford Review of Education*, **32**, 283-301..

TABLES

Table 1: Differences in child's non-cognitive skills by age at entry (score points¹)

	Full sample		Early entry (≤ 2.5 yrs)		Later entry (> 2.5 yrs)		Mean Difference
	Mean	SD	Mean	SD	Mean	SD	
Age 5							
<i>Overall SDQ score</i>	7.21	4.89	6.87	4.61	7.40	5.04	-0.528***
Externalising score	4.72	3.36	4.55	3.22	4.81	3.46	-0.254***
Conduct problems	1.46	1.47	1.40	1.42	1.49	1.50	-0.0919**
Hyperactivity	3.26	2.34	3.16	2.26	3.32	2.40	-0.162***
Internalising score	2.49	2.49	2.32	2.37	2.59	2.55	-0.274***
Emotional problems	1.36	1.57	1.30	1.49	1.39	1.60	-0.0966**
Peer problems	1.13	1.42	1.02	1.34	1.20	1.47	-0.177***
<i>Prosocial behaviour</i>	8.39	1.65	8.46	1.60	8.35	1.68	0.104**
'Borderline/abnormal' behaviour	0.11	0.31	0.08	0.28	0.13	0.33	-0.045***
Age 7							
<i>Overall SDQ score</i>	7.42	5.34	7.06	5.03	7.65	5.53	-0.590***
Externalising score	4.68	3.54	4.49	3.36	4.80	3.64	-0.312***
Conduct problems	1.35	1.51	1.28	1.44	1.40	1.55	-0.116***
Hyperactivity	3.23	2.48	3.20	2.37	3.40	2.54	-0.196***
Internalising score	2.75	2.76	2.58	2.64	2.86	2.83	-0.278***
Emotional problems	1.53	1.75	1.46	1.70	1.57	1.78	-0.107**
Peer problems	1.22	1.54	1.12	1.45	1.29	1.60	-0.172***
<i>Prosocial behaviour</i>	8.61	1.62	8.70	1.55	8.55	1.66	0.144***
'Borderline/abnormal' behaviour	0.14	0.34	0.11	0.32	0.15	0.36	-0.033***
N	6,678		2,706		3,754		6,460

Source: UK Millennium Cohort Study. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. ¹ For 'borderline/abnormal behaviour' the mean depicts percent and not score points.

Table 2: Estimation of the effect of later day care entry on child's socio-emotional behaviour at age 5 and at age 7

	Age 5			Age 7		
	OLS	ATT	ATU	OLS	ATT	ATU
<i>Overall SDQ score</i>	0.108*** (0.025)	0.039 (0.028)	0.033 (0.026)	0.110*** (0.025)	0.067** (0.028)	0.044* (0.026)
Externalising score	0.076*** (0.025)	0.018 (0.028)	0.005 (0.027)	0.088*** (0.025)	0.050* (0.028)	0.033 (0.027)
Conduct problems	0.062** (0.025)	-0.001 (0.028)	-0.011 (0.026)	0.077*** (0.025)	0.036 (0.029)	0.012 (0.027)
Hyperactivity	0.069*** (0.025)	0.026 (0.029)	0.014 (0.028)	0.079*** (0.025)	0.050* (0.029)	0.040 (0.028)
Internalising score	0.110*** (0.025)	0.053* (0.028)	0.058** (0.026)	0.101*** (0.025)	0.066** (0.028)	0.043 (0.027)
Emotional problems	0.062** (0.025)	0.012 (0.028)	0.016 (0.027)	0.061** (0.025)	0.032 (0.029)	0.012 (0.028)
Peer problems	0.124*** (0.025)	0.079*** (0.028)	0.084*** (0.027)	0.112*** (0.025)	0.081*** (0.028)	0.063** (0.028)
<i>Prosocial behaviour</i>	-0.063** (0.025)	-0.063** (0.029)	-0.050* (0.028)	-0.089*** (0.025)	-0.089*** (0.028)	-0.066** (0.028)
'Borderline/abnormal' behaviour (marginal eff.)	0.047*** (0.008)	0.026*** (0.009)	0.028*** (0.008)	0.034*** (0.009)	0.019* (0.010)	0.014 (0.009)
Conditioning variables	No	Yes	Yes	No	Yes	Yes
N	6,460	4,825	4,825	6,460	4,825	4,825

Standard errors in parentheses, all dependent variables in terms of one standard deviation (z-score), * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note: Each cell represents estimations from a separate regression analysis. ATT and ATU models include as control variables all conditioning variables as outlined in Section 5 of this paper. The number of observations in the ATU-analysis consists of 4823 observations. Source: UK Millennium Cohort Study.

Table 3: Estimation of the effect of later day care entry on socio-emotional behaviour at age 5 and at age 7 by gender (ATT)

	Boys		Girls	
	Age 5	Age 7	Age 5	Age 7
<i>Overall SDQ score</i>	0.041 (0.042)	0.105** (0.042)	0.031 (0.037)	0.019 (0.038)
Externalising score	0.017 (0.042)	0.087** (0.042)	0.014 (0.038)	0.009 (0.039)
Conduct problems	-0.005 (0.042)	0.085** (0.043)	-0.011 (0.039)	-0.018 (0.039)
Hyperactivity	0.028 (0.042)	0.073* (0.042)	0.027 (0.038)	0.024 (0.039)
Internalising score	0.057 (0.041)	0.092** (0.042)	0.043 (0.039)	0.025 (0.039)
Emotional problems	0.009 (0.040)	0.053 (0.042)	0.005 (0.042)	-0.001 (0.042)
Peer problems	0.089** (0.042)	0.105** (0.042)	0.069* (0.038)	0.046 (0.038)
<i>Prosocial behaviour</i>	-0.094** (0.044)	-0.152*** (0.045)	-0.038 (0.038)	-0.011 (0.036)
‘Borderline/abnormal’ behaviour (marginal eff.)	0.024* (0.014)	0.039*** (0.015)	0.024** (0.012)	-0.005 (0.013)
N	2,444		2,379	

Standard errors in parentheses, all dependent variables in terms of one standard deviation (z-score), * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note: Each cell represents estimations from a separate regression analysis. All models include as control variables all conditioning variables as outlined in Section 5 of this paper. Source: UK Millennium Cohort Study.

Table 4: Other heterogeneous effects of later day care entry on socio-emotional behaviour at age 5 (ATT)

	Overall SDQ score	Conduct problems	Hyperactivity	Emotional problems	Peer problems	Prosocial behaviour	N
Low education	0.026 (0.045)	-0.023 (0.047)	0.058 (0.044)	-0.033 (0.046)	0.055 (0.045)	-0.099** (0.044)	2,307
High education	0.052 (0.033)	0.017 (0.034)	-0.005 (0.036)	0.050 (0.035)	0.113*** (0.034)	-0.029 (0.037)	2,513
Below 60% median income	0.130* (0.071)	-0.028 (0.075)	0.193*** (0.066)	0.054 (0.069)	0.098 (0.070)	-0.088 (0.070)	1,091
Above 60% median income	0.005 (0.029)	-0.008 (0.030)	-0.025 (0.032)	-0.003 (0.031)	0.072** (0.030)	-0.054* (0.031)	3,732
Boys with low educated mothers	0.003 (0.070)	-0.075 (0.071)	0.037 (0.067)	-0.048 (0.064)	0.080 (0.068)	-0.164** (0.067)	1,151
Boys with high educated mothers	0.060 (0.049)	0.036 (0.051)	0.001 (0.054)	0.053 (0.049)	0.107** (0.050)	-0.008 (0.057)	1,280
Girls with low educated mothers	0.009 (0.059)	-0.017 (0.064)	0.068 (0.058)	-0.066 (0.066)	0.010 (0.059)	-0.030 (0.059)	1,156
Girls with high educated mothers	0.047 (0.045)	-0.010 (0.047)	-0.007 (0.052)	0.058 (0.051)	0.120** (0.046)	-0.041 (0.050)	1,220
Boys living in households with income below 60% of median	0.081 (0.113)	-0.054 (0.114)	0.209** (0.099)	-0.030 (0.104)	0.025 (0.113)	-0.114 (0.108)	535
Boys living in households with income above 60% of median	0.019 (0.043)	-0.004 (0.043)	-0.033 (0.047)	0.023 (0.042)	0.098** (0.043)	-0.076 (0.046)	1,895
Girls living in households with income below 60% of median	0.178** (0.088)	-0.031 (0.095)	0.208** (0.087)	0.109 (0.095)	0.180** (0.084)	-0.071 (0.092)	539
Girls living in households with income above 60% of median	-0.018 (0.041)	-0.016 (0.043)	-0.032 (0.043)	-0.032 (0.046)	0.040 (0.041)	-0.018 (0.042)	1,834

Standard errors in parentheses, all dependent variables in terms of one standard deviation (z-score), * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note: All cells represent estimations from separate regressions controlling for all conditioning variables used for propensity score matching. Columns display the ATT of each outcome. Source: UK Millennium Cohort Study.

Table 5: Other heterogeneous effects of later day care entry on socio-emotional behaviour at age 7 (ATT)

	Overall SDQ score	Conduct problems	Hyperactivity	Emotional problems	Peer problems	Prosocial behaviour	N
Low education	0.117*** (0.044)	0.060 (0.045)	0.064 (0.044)	0.087* (0.046)	0.147*** (0.042)	-0.142*** (0.043)	2,307
High education	0.019 (0.034)	0.022 (0.035)	0.033 (0.037)	-0.025 (0.036)	0.018 (0.037)	-0.041 (0.037)	2,513
Below 60% median income	0.195*** (0.069)	0.125* (0.074)	0.168** (0.066)	0.093 (0.070)	0.178*** (0.067)	-0.072 (0.068)	1,090
Above 60% median income	0.019 (0.030)	-0.003 (0.030)	0.009 (0.032)	0.009 (0.032)	0.045 (0.031)	-0.082*** (0.031)	3,732
Boys with low educated mothers	0.219*** (0.064)	0.138** (0.068)	0.149** (0.064)	0.166** (0.064)	0.197*** (0.062)	-0.281*** (0.066)	1,151
Boys with high educated mothers	-0.007 (0.052)	0.039 (0.052)	0.000 (0.055)	-0.078 (0.052)	0.024 (0.054)	-0.047 (0.058)	1,280
Girls with low educated mothers	-0.026 (0.059)	-0.045 (0.062)	-0.023 (0.059)	-0.056 (0.067)	0.053 (0.059)	0.027 (0.054)	1,156
Girls with high educated mothers	0.039 (0.047)	-0.008 (0.048)	0.050 (0.052)	0.039 (0.051)	0.021 (0.051)	-0.044 (0.048)	1,220
Boys living in households with income below 60% of median	0.313*** (0.109)	0.261** (0.117)	0.268** (0.098)	0.194* (0.105)	0.175 (0.109)	-0.134 (0.104)	535
Boys living in households with income above 60% of median	0.033 (0.044)	0.023 (0.044)	0.017 (0.046)	-0.005 (0.044)	0.070 (0.044)	-0.146*** (0.049)	1,895
Girls living in households with income below 60% of median	0.031 (0.092)	-0.025 (0.097)	0.077 (0.092)	-0.091 (0.101)	0.113 (0.088)	0.010 (0.082)	539
Girls living in households with income above 60% of median	-0.004 (0.041)	-0.040 (0.043)	-0.015 (0.044)	0.023 (0.046)	0.024 (0.042)	-0.007 (0.041)	1,834

Standard errors in parentheses, all dependent variables in terms of one standard deviation (z-score), * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note: All cells represent estimations from separate regressions controlling for all conditioning variables used for propensity score matching. Columns display the ATT of each outcome. Source: UK Millennium Cohort Study.

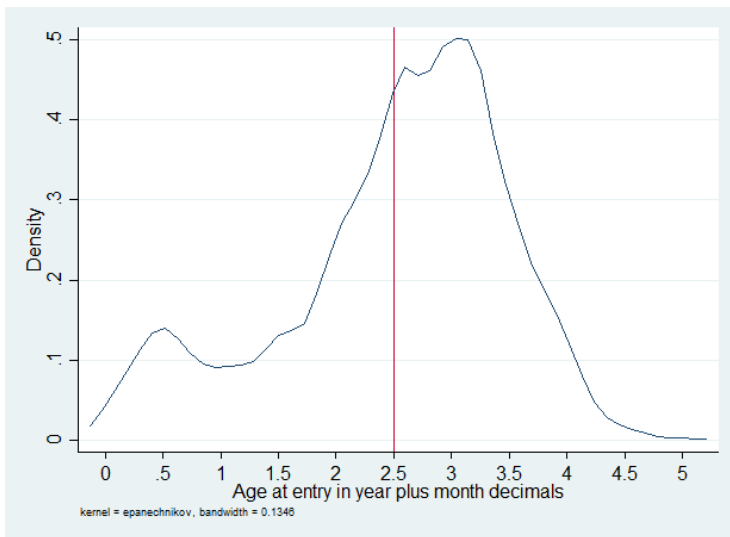
Table 6: Sensitivity analysis using median of age at entry as cut-off (median=2.7 years of age) (ATT)

	Age 5	Age 7
<i>Overall SDQ score</i>	0.036 (0.028)	0.064** (0.028)
Externalising score	0.018 (0.028)	0.043 (0.029)
Conduct problems	-0.012 (0.029)	0.030 (0.029)
Hyperactivity	0.033 (0.029)	0.043 (0.029)
Internalising score	0.047* (0.029)	0.069** (0.029)
Emotional problems	0.002 (0.029)	0.028 (0.029)
Peer problems	0.081*** (0.029)	0.093*** (0.029)
<i>Prosocial behaviour</i>	-0.052* (0.029)	-0.099*** (0.029)
‘Borderline/abnormal’ behaviour (marginal eff.)	0.029*** (0.009)	0.017* (0.010)

Standard errors in parentheses, all dependent variables in terms of one standard deviation (z-score), * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note: N=4,825. Each cell represents estimations from a separate regression analysis. All models include as control variables all conditioning variables as outlined in Section 5 of this paper. Source: UK Millennium Cohort Study.

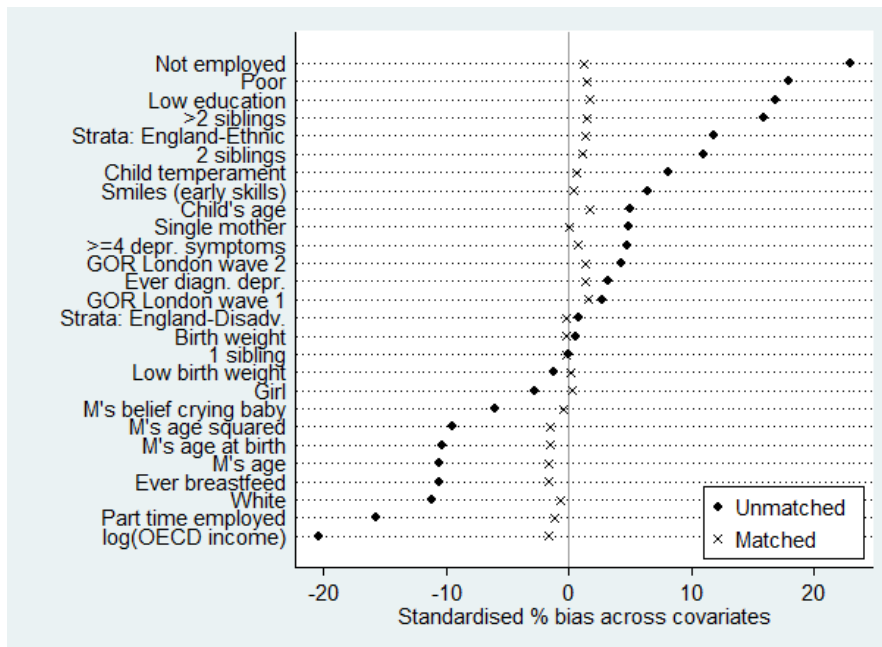
Figures

Figure 1: Distribution of age at entry in full sample



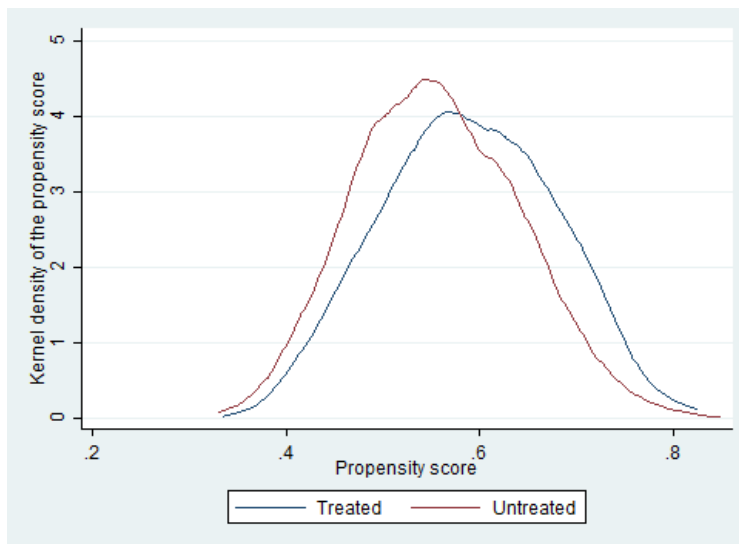
Source: UK Millennium Cohort Study. Note: This figure displays the distribution of starting age in the overall sample. The red line marks the cut-off point used in our empirical strategy.

Figure 2: “Goodness of fit” – Bias reduction of propensity score matching in full sample



Source: UK Millennium Cohort Study. Own calculations. Note: All conditioning variables are measured at wave one except for GOR London, age of child and employment status, which are measured at wave 2. Figure 2 presents the standardised percent bias after matching. It is the percent difference of the sample means in the treatment and the matched control sample as a percentage of the square root of the average of the sample variances in both groups. It shows that after propensity score matching all conditioning variables have a standardized bias within the -5 percent and 5 percent interval. Prior matching (black dot) a positive standardised percent bias indicates higher mean of treatment group and a negative standardised percent bias higher mean of control group. Source: UK Millennium Cohort Study. Own calculations using *pstest* provided by Leuven and Sianesi (2003) as part of *psmatch2*.

Figure 3: Kernel density of propensity score for treatment and control group in full sample



Source: UK Millennium Cohort Study. Note: This figure compares the overlap of the estimated propensity scores for treatment and control group in the full sample (N=4,825).

Appendix

Table A1: Items of the Strength and Difficulties Score (SDQ) and Subscales

The parents rank each item to be “not true”, “somewhat true” or “certainly true”

Subscales	Items
Conduct problems scale	Often has temper tantrums or hot tempers... Generally obedient... Often fights with other children... Often lies or cheats... Steals from home, school or elsewhere...
Hyperactivity scale	Restless, overactive... Constantly fidgeting or squirming... Easily distracted, concentration wanders... Thinks things out before acting... Sees tasks through to the end...
Emotional problems scale	Often complains of headaches... Many worries... Often unhappy, downhearted... Nervous or clingy in new situations... Many fears, easily scared...
Peer problems scale	Rather solitary, tends to play alone... Has at least one good friend... Generally liked by other children... Picked on or bullied... Gets on better with adults than with other children...
Prosocial scale	Considerate of other people's feelings ... Shares readily with other children... Helpful if someone is hurt... Kind to younger children... Often volunteers to help others...

Source: Scoring the Strengths/Difficulties Questionnaire for age 4-17, August 2014, www.sdqinfo.com (download: October 2014) for further details.

Table A2: Overview of conditioning variables

Child	
Age ¹	Measured in months
Girl	0=boy, 1=girl
Birth weight	Measured in kg
Low birth weight	0=birth weight is ≥ 2.5 kg, 1=birth weight is < 2.5 kg
Mother ever tried to breastfeed	0=no, 1=yes
Main Development Index: Child ...	
...smiles	
...sits up	
...stands up holding on	
...puts hands together	
...grabs objects	
...holds small objects	1=often
...passes a toy	2=once or twice
...walks a few steps	3=not yet
...gives toy	
...waves bye-bye	
...extends arms	
...nods for yes	
Child temperament measured in terms of	Index that ranges from 0=most uncomplicated child behaviour to 4=most complicated child behaviour.
Mother	
Age at birth	Measured in years
Age	Measured in years
Age squared	Measured in years
White	0=not ethnic group white, 1=ethnic group white
Low education	0=A-levels or higher qualification, 1=below A-levels qualification
Employment status ¹	Ref.: full time, part-time, not employed (binary variables)
Depressive symptoms	0=mother shows < 4 depressive symptoms, 1=mother shows ≥ 4 depressive symptoms
Depression diagnosed	0=mother never had a diagnosed depression, 1=mother had at least once a diagnosed depression
Mother's belief whether crying baby should be picked up immediately	Range from 1=highest degree of agreement to 5=highest degree of disagreement
Household	
Single parent	0=no single parent (mother), 1=single parent (mother)
Number of siblings	Ref. no sibling, one sibling, two siblings, more than two siblings (binary variables)
HH Income	Log(OECD HH Income) in MCS1
Poor	0=above 60percent median OECD HH income, 1=below 60percent median OECD HH income
Living in London ¹	0=not living in GOR London, 1=living in GOR London (measured at wave 1 and 2)
Region	Ref.: England – Advantaged, England – Disadvantaged, England – Ethnic (binary variables)

Source: UK Millennium Cohort Study. All variables are measured pre-treatment – at wave 1 if not indicated otherwise. ¹ These variables are measured in wave 2.

Table A3: Differences in child's non-cognitive skills by age at entry in matched sample (score points¹)

	Full sample		Early entry (≤2.5 years)		Later entry (> 2.5 years)		Mean Difference
	Mean	SD	Mean	SD	Mean	SD	
Age 5							
<i>Overall SDQ score</i>	6.95	4.76	6.66	4.50	7.16	4.93	-0.509***
Externalising score	4.57	3.31	4.43	3.18	4.68	3.41	-0.247**
Conduct problems	1.40	1.44	1.35	1.39	1.44	1.47	-0.0892**
Hyperactivity	3.17	2.32	3.08	2.24	3.24	2.38	-0.158**
Internalising score	2.37	2.39	2.22	2.28	2.49	2.47	-0.261***
Emotional problems	1.30	1.50	1.26	1.45	1.34	1.54	-0.0765*
Peer problems	1.07	1.38	0.96	1.30	1.15	1.43	-0.185***
<i>Prosocial behaviour</i>	8.43	1.61	8.51	1.56	8.37	1.64	0.143***
'Borderline/abnormal' behaviour	0.10	0.30	0.07	0.26	0.12	0.32	-0.044***
Age 7							
<i>Overall SDQ score</i>	7.22	5.24	6.87	4.92	7.48	5.45	-0.615***
Externalising score	4.57	3.49	4.37	3.32	4.72	3.60	-0.343***
Conduct problems	1.32	1.50	1.24	1.42	1.37	1.55	-0.124***
Hyperactivity	3.25	2.45	3.13	2.35	3.35	2.51	-0.220***
Internalising score	2.65	2.71	2.49	2.59	2.76	2.79	-0.271***
Emotional problems	1.49	1.72	1.44	1.67	1.53	1.75	-0.0953*
Peer problems	1.16	1.52	1.06	1.43	1.23	1.58	-0.176***
<i>Prosocial behaviour</i>	8.64	1.59	8.74	1.50	8.57	1.64	0.172***
'Borderline/abnormal' behaviour	0.13	0.33	0.11	0.31	0.14	0.35	- 0.0328***
N	4,825		2,055		2,770		

Source: UK Millennium Cohort Study. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. ¹ For 'borderline/abnormal behaviour' the mean depicts percent and not score points.

Table A4: Quality of matching

	Mean		Standardised bias	p > t
	Treatment	Matched Control		
Child				
Age ¹	37.445	37.405	1.7	0.531
Girl	0.484	0.483	0.2	0.931
Birth weight	3.389	3.391	-0.3	0.914
Low birth weight	0.057	0.056	0.2	0.951
Mother ever tried to breastfeed	0.743	0.749	-1.6	0.550
Main Development Index: Child ...				
...smiles	1.0051	1.0048	0.4	0.881
...sits up	1.041	1.039	0.8	0.760
...stands up holding on	1.429	1.424	0.7	0.796
...puts hands together	1.228	1.228	0.0	0.998
...grabs objects	1.004	1.005	-0.5	0.842
...holds small objects	1.135	1.131	0.9	0.731
...passes a toy	1.059	1.057	0.6	0.820
...walks a few steps	2.814	2.806	1.6	0.560
...gives toy	1.537	1.534	0.4	0.870
...waves bye-bye	1.964	1.956	0.9	0.725
...extends arms	1.234	1.229	0.8	0.748
...nods for yes	2.777	2.781	-0.7	0.787
Child temperament index	-1.491	-1.493	0.4	0.876
Mother				
Age	29.883	29.972	-1.6	0.566
Age squared	926.28	931.31	-1.5	0.585
White	0.881	0.883	-0.6	0.810
Low education	0.515	0.506	1.7	0.519
<i>Ref.: Full time employed</i>				
Part time employed ¹	0.440	0.446	-1.2	0.663
Not employed ¹	0.475	0.469	1.3	0.643
Depressive symptoms	0.135	0.132	0.7	0.803
Depression diagnosed	0.243	0.237	1.4	0.604
Mother's belief whether crying baby should be picked up immediately	3.048	3.051	-0.3	0.898
Household				
Single parent	0.116	0.115	0.0	0.991
<i>Ref.: No sibling</i>				
One sibling	0.378	0.378	-0.1	0.982
Two siblings	0.159	0.155	1.1	0.685
More siblings	0.069	0.066	1.3	0.636
HH Income	5.607	5.618	-1.6	0.553
Poor	0.254	0.248	1.3	0.623
Living in London	0.155	0.150	1.4	0.599
Living in London ¹	0.139	0.135	1.3	0.622
<i>Ref.: England - Advantaged</i>				
Region: England - Disadvantaged	0.3726	0.374	-0.3	0.910
Region: England - Ethnic	0.108	0.104	1.2	0.626

Note: All variables are measured pre-treatment – at wave 1 if not indicated otherwise. ¹ These variables are measured in wave 2. The third column displays the standardised percent bias after matching. It is the percent difference of the sample means in the treatment and the matched control sample as a percentage of the square root of the average of the sample variances in both groups. The fifth column shows the p-value of the likelihood ratio test of the joint insignificance of all relevant variables after matching. Source: UK Millennium Cohort Study.

Table A5: Sensitivity analysis using matching on a limited age at entry treatment (ATT)

	Early entry ≥ 12 months		Later entry < 48 months	
	Age 5	Age 7	Age 5	Age 7
<i>Overall SDQ score</i>	0.044 (0.029)	0.061** (0.030)	0.038 (0.028)	0.064** (0.028)
Externalising score	0.020 (0.030)	0.049 (0.030)	0.016 (0.028)	0.046 (0.028)
Conduct problems	0.003 (0.030)	0.026 (0.031)	-0.003 (0.029)	0.034 (0.029)
Hyperactivity	0.027 (0.030)	0.053* (0.031)	0.024 (0.029)	0.045 (0.029)
Internalising score	0.060** (0.030)	0.055* (0.031)	0.054* (0.028)	0.067** (0.029)
Emotional problems	0.013 (0.030)	0.027 (0.032)	0.012 (0.029)	0.033 (0.030)
Peer problems	0.090*** (0.030)	0.068** (0.030)	0.081*** (0.028)	0.082*** (0.028)
<i>Prosocial behaviour</i>	-0.068** (0.031)	-0.082*** (0.031)	-0.068** (0.029)	-0.091*** (0.029)
'Borderline/abnormal' behaviour (marginal eff.)	0.026*** (0.009)	0.021** (0.010)	0.027*** (0.009)	0.019* (0.0103)
N	4,292		4,766	

Standard errors in parentheses, all dependent variables in terms of one standard deviation (z-score), * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note: Each cell represents estimations from a separate regression analysis. All models include as control variables all conditioning variables as outlined in Section 5 of this paper. Source: UK Millennium Cohort Study.

Table A6: Sensitivity analysis – propensity score matching without “temperament of child” variable (models I) and without “temperament of child” variable and “poor” variable (models II) (ATT)

	Models I		Models II	
	Age 5	Age 7	Age 5	Age 7
<i>Overall SDQ score</i>	0.036 (0.026)	0.065** (0.026)	0.036 (0.025)	0.063** (0.025)
Externalising score	0.021 (0.026)	0.055** (0.026)	0.017 (0.026)	0.050* (0.025)
Conduct problems	0.005 (0.027)	0.044* (0.027)	-0.003 (0.026)	0.043* (0.026)
Hyperactivity	0.027 (0.027)	0.052* (0.027)	0.027 (0.026)	0.045* (0.026)
Internalising score	0.042 (0.027)	0.054** (0.027)	0.046* (0.026)	0.057** (0.026)
Emotional problems	0.001 (0.027)	0.031 (0.028)	0.005 (0.027)	0.032 (0.027)
Peer problems	0.073*** (0.027)	0.062** (0.027)	0.076*** (0.026)	0.067** (0.026)
<i>Prosocial behaviour</i>	-0.042 (0.027)	-0.085*** (0.027)	-0.042 (0.026)	-0.081*** (0.026)
‘Borderline/abnormal’ behaviour (marginal eff.)	0.026*** (0.009)	0.018** (0.009)	0.026*** (0.008)	0.019** (0.009)
N	5,568		5,934	

Standard errors in parentheses, all dependent variables in terms of one standard deviation (z-score), * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each cell represents estimations from a separate regression analysis. All models include as control variables all conditioning variables as outlined in Section 5 of this paper. Source: UK Millennium Cohort Study.