This is a repository copy of *Unlocking potential: Investigating the prolonged impact of formal childcare intensity on non-cognitive skills*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/208601/

Version: Published Version

**Monograph:**
Ward, L. (2024) Unlocking potential: Investigating the prolonged impact of formal childcare intensity on non-cognitive skills. Working Paper. Sheffield Economic Research Paper Series, 2024001 . Department of Economics, University of Sheffield ISSN 1749-8368

# The University Of Sheffield.

# Department Of Economics

# Unlocking Potential: Investigating the Prolonged Impact of Formal Childcare Intensity on Non-Cognitive Skills

Lucy Ward

# Unlocking Potential: Investigating the Prolonged Impact of Formal Childcare Intensity on Non-Cognitive Skills[*]

LucyWard[†1]

[1]Department of Economics, University of Sheffield, UK

January 30, 2024

## Abstract

To support the expansion of universal early childcare programs, policymakers often refer to the positive outcomes documented in the literature. In reality, the evidence is mixed. In addition, most evidence is based on the impacts of enrolment in childcare for children aged 3-to-5-year-old. This research focuses on the intensity of childcare, measured in hours per week, for children under the age of 3. This area has received very little attention in the literature despite recent policy focus, specifically in the UK, on the number of hours of childcare which should be subsidised. We use data from a large, nationally representative English birth cohort, the Millennium Cohort Study, and an instrumental variables strategy that leverages exogenous variation in both the probability that the mother works shift work and has uncertain working hours to estimate whether hours in formal childcare prior to the age of 3 have an impact on non-cognitive skills at ages 3-14, which are measured by the Strength and Difficulties Questionnaire (SDQ), proposed by Goodman (1997). Results indicate that increasing hours in formal childcare has an initial positive impact on non-cognitive skills which persists over time. Moreover, we estimate heterogeneous impacts across family background characteristics, suggesting that increasing access to more time in childcare for disadvantaged children may hold potential for decreasing early inequalities in child development. The results are robust to a number of sensitivity checks including weak instrument robust testing and discussion of potential omitted variables.

**Keywords**: Childcare, child outcomes, non-cognitive skills, instrumental variables
**JEL Classifications**: J13, I21, C26

# 1　Introduction

Are we falling short in providing adequate support for our children before they start school? Whilst 72% of non-disadvantaged pupils reach the expected level of development by the age of 5, only 55% of disadvantaged pupils reach the same level in the UK.[1] (Andrews et al., 2017). Disparities in skills during early childhood are likely to contribute to inequalities later in life, including health (Case et al., 2005), education (Almond and Currie, 2011), and employment (Black et al., 2007). The UK government's commitment to halve this gap by 2028, has created a conversation on how to achieve this goal. Early childhood has become a central focus in this discussion, particularly in relation to the development of human capital. There is a wide range of literature on the benefits of attending childcare, with specific focuses on the benefits for disadvantaged children. Less research has focused on the intensity of this treatment. This paper aims to shed light on the causal relationship between hours spent in childcare at age 3 and non-cognitive development for children aged 3-14. Using nationally representative data from the Millennium Cohort Study, we are able to show that there are not only short run gains but also long run impacts of early access to childcare for more hours.

Despite the recognised importance of early childhood in the human capital production function, it remains the only stage in childhood and adolescence with relatively limited public investment. In the UK, less than 0.1% of GDP is allocated to childcare, placing it among the lowest investments within the OECD. Notably, 84% of childcare in the UK is provided by for-profit providers, a sharp contrast to the 3% in Germany and 4% in France (Darlington, 2021). The UK childcare system is also much more heterogeneous than in other European countries (Sweden, Denmark, and France), where an affordable, fulltime public childcare system has been in place for several decades. The cost of childcare in the UK has been notoriously high with the government first introducing childcare subsidies in the early 2000s.[2]

Allocating resources to childcare could create numerous advantages, including the enhancement of children's development, the increase in family income, and a general improvement in productivity. This is because childcare is likely having knock-on implications for parents' ability and incentive to work. In the short term, it boosts productivity by expanding labour force participation and hours of work. Additionally, there are medium-term benefits, as childcare prevents the deterioration of parents' skills during periods of non-employment. Over the long term, investing in childcare is expected to contribute to sustained productivity by improving children's development.

This research aims to explore the relationship between hours in childcare prior to the age of 3 and non-cognitive skills, an area which has received limited attention in the literature. In this paper, we use the term non-cognitive skills to describe the personal attributes not thought to be measured by achievement tests.[3] Non-cognitive skills are often overlooked within the education assessment system where there is heavy reliance on written tests to screen and sort individuals, to evaluate students and schools, and to assess the performance of entire nations. This is despite the argument

---

[1] Disadvantaged is defined as being eligible for free school meals.

[2] All 4-year-olds since 2000 and all 3-year-olds since 2005 have been entitled to free, part-time nursery education. This was set at 12.5 hours per week for 33 weeks per year until 2010, then 15 hours per week for 38 weeks. In 2013, free part-time education was also extended to disadvantaged 2 year olds. In 2017, this was increased to 30 hours for working parents. In the 2023 budget, the government announced their plans to provide 30 hours of free childcare for children over the age of nine months with working parents.

[3] These attributes go by many names in the literature, including soft skills, personality traits, non-cognitive abilities, character skills, and socio-emotional skills.

that non-cognitive skills predict a wide range of outcomes, including educational achievement, labour market outcomes, health, and criminality (see Heckman and Kautz, 2014; Almlund et al., 2011, Borghans et al., 2008; and Roberts et al., 2007). Furthermore, the predictive power of non-cognitive skills has been shown to rival that of measures of cognitive ability. For example, conscientiousness predicts years of schooling with the same strength as measures of intelligence (Almlund et al., 2011).

Identifying a causal relationship is a challenge due to formal childcare being determined by the parents' choices.[4] Parents' parenting skills and preferences for childcare are likely correlated with the amount of time spent in childcare and the child's non-cognitive skills. These factors are difficult to observe, potentially creating a spurious correlation between hours in childcare and non-cognitive development. To correct for endogeneity of hours in childcare, this research creates and uses instrumental variables.

Making use of the information on mother's occupation, we calculate the probability that the mother would work shifts or have uncertain working hours. The assumptions are that due to the structured nature of formal childcare, mothers working in sectors with more shift work, place their child in formal care for fewer hours. On the other hand, if the mother has varying hours each week, the uncertainty causes them to place their child in formal care for more hours.

Utilising the Millennium Cohort Study (MCS) for the UK, which provides detailed information about childcare and child outcomes, and employing the instrumental variables, our results are summarised as follows: the amount of time spent in formal childcare has a positive and significant impact on non-cognitive skills at age 3 through to age 7. We find that the impact on emotional skills, conduct and peer relationships is the most persistent over time with larger effects for males, first-borns and disadvantaged children.

There is a large literature that has examined the impact of attending childcare on both cognitive and non-cognitive development. Previous literature has provided international evidence on the positive impacts of attending childcare prior to starting school. This research has demonstrated both short run and long run impacts on cognitive and non-cognitive development (see Elek et al., 2020 for an overview of the literature). Areas that have received less attention in the literature are on the effectiveness of attending childcare for children under 3 and the intensity of childcare. If we think of early childcare as an investment in children's human capital, it is natural to think that the more time spent in childcare the better.

Datta Gupta and Simonsen (2010) and Berger et al., (2021), who both examine the impact of attending childcare, extend their analysis to provide descriptive evidence on the impact of hours in childcare on cognitive and non-cognitive skills. Datta Gupta and Simonsen (2010), using data from Denmark on individuals born in 1995, finds that increases in the hours from 0-10 to 10-20 and 10-20 to 20-30 are benign, no matter the choice of care. Further increasing hours, however, seems to significantly worsen non-cognitive skills, and this is significant in the case of preschool. Berger et al. (2021), using French data on children born in 2011, provides evidence to suggest that additional hours or days per week in creche are associated with greater language and motor skills. They find the opposite results for behaviour, children spending three and five days in creche exhibit poorer behaviour than those experiencing other types of care. It should be noted that these estimates are taken from an OLS model and therefore are likely biased.

---

[4]This research focuses on 3 year olds in 2003/4, who were not eligible for the childcare policies of the time.

To examine whether there is causal relationship, there has been a larger focus on the age that childcare started. The most common method used is regression discontinuity. Blanden et al (2022) and Kuehnle and Oberfichtner (2020) both exploit variation in month of birth to examine the impact of starting childcare earlier. Children born earlier in the academic year become eligible first for subsided childcare. Blanden et al (2022) find that eligibility for an extra term of childcare leads to a small increase in the probability (just under 1 percentage point) that a child reaches the expected level of competencies after the first year in primary school, at age 5, compared to children not eligible for the extra term. Kuehnle and Oberfichtner (2020) find no short or long run effect of starting childcare earlier. Fort et al. (2020) also exploit a discontinuity; they use admission thresholds. They find that one additional day care month at age 0–2 reduces intelligence quotient by 0.5% (4.7% of a standard deviation) at age 8–14 in a relatively affluent population. For non-cognitive skills, an extra month decreases openness and agreeableness by 1.4% and 1.2% and increases neuroticism by 0.9%.

Overall, this area of the literature is limited, and the current studies provide mixed results. The literature has broadly found that childcare has positive effects on cognitive outcomes and children who start childcare at a younger age benefit more. The current research has however failed to reach an agreement on the impacts of childcare on non-cognitive skills. The impact of childcare on non-cognitive skills in comparison to cognitive skills seems to be more impacted by the country analysed, specific measure of skills, starting age and quality of childcare. This overview points towards important heterogeneity in the effects of early childhood intervention, specifically on non-cognitive skills, highlighting the need for further research.

Against this backdrop, we make three contributions to the literature. Firstly, we add to the knowledge base on the impacts of childcare on non-cognitive skills. Due to variation in findings, providing useful evidence for policy guidance is a challenge, and by increasing the evidence we may start to be able to draw some conclusions. Secondly, this research specifically focuses on the impact of hours in childcare. In the UK, for example, the current childcare policy focus is on the number of hours which should be subsidised though the evidence base on this question is limited. Finally, we focus on childcare taken at ages up to age 3, while the majority of the literature has examined the impact of childcare on children aged 3-5.

This research provides evidence in support of a causal relationship between time spent in childcare at younger ages and non-cognitive skills. We add to the evidence-based case for significant additional government investment in early childcare.

This paper is organised as follows. Sections 2 outline the data whilst section 3 discusses the methodology. Section 4 presents the estimated impact and tests the robustness, before concluding in Section 5.

## 2   Data

This research uses data from the Millennium Cohort Study (MCS). The MCS is a multidisciplinary cohort survey run by the Centre for Longitudinal Studies at the University College London.

The MCS is a valuable data source as it tracks the lives of a sample of about 19,000 babies born in

the UK between 1st September 2000 and 11th January 2002.[5] The sample was constructed to be representative of the total UK population. The data collectors selected electoral wards with the aim to recruit 100 per cent of the children born in the eligible period within them.[6] They also wanted to adequately represent disadvantaged and ethnic minority children. The population of wards was therefore stratified by ethnicity and the Child Poverty Index. This means that certain sub-groups of the population were intentionally over sampled.

The survey is conducted in several waves, the first occurred when children were aged nine months, gathering information from the parents of 18,818 children. Since then, families have been interviewed again six times at ages 3, 5, 7, 11, 14 and 17.[7] The survey was originally answered by just the parents. From wave 2 onwards, the child was also surveyed along with older siblings. Class teachers responded to a survey in waves 3-5. Topics covered include family composition, housing and local area, parental education, employment and income, parental and child health, parenting activities and attitudes, physical, social and cognitive development of the child, preschool experiences, schooling and leisure activities etc.

The main advantage of the MCS is the rich range of information regarding the experiences and outcomes of the MCS children and their families. This allows many individual and family characteristics to be controlled for. The main limitation of the survey is that the longitudinal pattern of response is complex, with attrition and re-entry. At each wave there has been non-response owing to refusal and non-contact.

## 2.1 Sample

In this research we use data from the first six waves, which took place when the children were aged 9 months, 3, 5, 7, 11 and 14 years old.[8] We are interested in the impact of childcare at age 3, therefore the majority of our focus is on the information provided in wave 2. We use the data provided in wave 1, at 9 months, to construct our control variables and the information in waves 3-6 to measure non-cognitive skills at older ages.

We restrict our sample to children living in England, so that the childcare supply is similar for all children observed. We also exclude twins and triplets since childcare arrangements and their effects may be different when more children need to be looked after.

The Strength and Difficulties questionnaire, the outcome of interest, as well as reported hours in childcare, suffer from non-responses, we therefore keep cohort members who have responded at least once to the SDQ and report hours in childcare at age 3.[9] The initial sample is therefore 6,567 cohort members.[10]

---

[5]1 September 2000 and 31 August 2001 (for England and Wales), and between 24 November 2000 and 11 January 2002 (for Scotland and Northern Ireland).

[6]They achieved a response rate of 72 per cent of all the families with eligible children living at nine months in the sampled wards.

[7]Wave 8 at age 22, is currently being undertaken.

[8]This research stops at age 14 as this is the last year parents completed the Strengths and Difficulties Questionnaire, the main outcome of interest.

[9]We examine the determinants of non-response as well as examine the impact of a change in the sample and find no impact on the estimated results.

[10]Focusing on England reduces the sample to 11,680. Removing twins reduces the sample by 308. A further 3,336 observations are dropped due to the childrencare questions not being responded to. The sample is reduced to 7,957 as

## 2.2 Childcare measure

This research is interested in the impact of time spent in formal childcare at the age of 3. We define formal childcare as any type of childcare provided by someone who is not family or friend. This includes centre-based care for example nursery but also childminders. We do not include nannies or au pairs as we are interested in the impact of being cared for in a group setting outside the home. Nurseries are likely to be more structured than childminders, however both are required to have some level of childcare qualifications. We also believe that the difference in structure is less important for non-cognitive skills than cognitive development.

In the first three waves of the survey, the mother was asked about the childcare choices made. In wave 1, when the child was 9 months old, working mothers were asked to state the types of care being used at the time of the survey when they were at work. In wave 2, when the child was 3 years old, all mothers were asked details about the type of childcare that had been used since the first survey, including starting dates, stop dates, and the number of hours per week.[11]

This research focuses on information from the second wave for two reasons. Firstly, it allows us to include working and non-working mothers, generating a more representative sample. Secondly, information on hours in childcare, the main focus of this research, is only surveyed in wave 2.

There is potentially measurement error in the number of hours of care reported by the mother. Whilst some parents will know exactly how many hours they enrol their child in childcare for, other may only know how many sessions a week they take them. This uncertainty could lead to mothers guessing how many hours this equates to. Whether the guessing or quick in head calculations would give under or over estimates is not clear.[12]

## 2.3 Measure of non-cognitive skill

The outcome of interest in this research is non-cognitive development. We focus on non-cognitive skills for two reasons. Firstly, in comparison to cognitive skills, the literature has found more varied results meaning there is a lack of a consensus on the impact of childcare on non-cognitive skills. Secondly, we are interested in both the short and long run effect of time in childcare and it has been argued that non-cognitive skills are more malleable as well as more predictive of long-term outcomes than are test scores (Heckman & Rubinstein, 2001; Lindquist & Vestman, 2011; Mueller & Plug, 2006).

As our measure of non-cognitive skills we use the Strength and Difficulties Questionnaire (SDQ), proposed by Goodman (1997). The questionnaire consists of 25 questions over five separate dimensions: Conduct Problems, Hyperactivity/Inattention, Emotional Symptoms, Peer Relationship Problems, and Prosocial behaviour. The questionnaire consists of statements to which the responses

there are 78 non-response to the Strength and Difficulties questionnaire. The final reduction in the sample is due to non-response to other variables included in the model.

[11]In the third wave, when the child is 5 years old, all mothers are asked about all childcare prior to starting school. This research focuses on children under the age of 3 and therefore, focuses on the first 2 waves of the survey.

[12]See section 3 for how we try to control for this measurement error.

are: 'not true', 'somewhat true', and 'certainly true'.[13][14] Higher scores on the first four dimensions 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 total difficulties score ranges from 0 to 40, the subscales from 0 to 10. For the total difficulties score, 0–15 is defined as having low needs, 16–19 some needs and 20–40, high needs. For our analysis we use the overall SDQ score as well as its dimensions to cover various aspects of a child's non-cognitive skills. In the estimations, the SDQ and all subscales are standardised to have a mean of zero and a standard deviation of one.

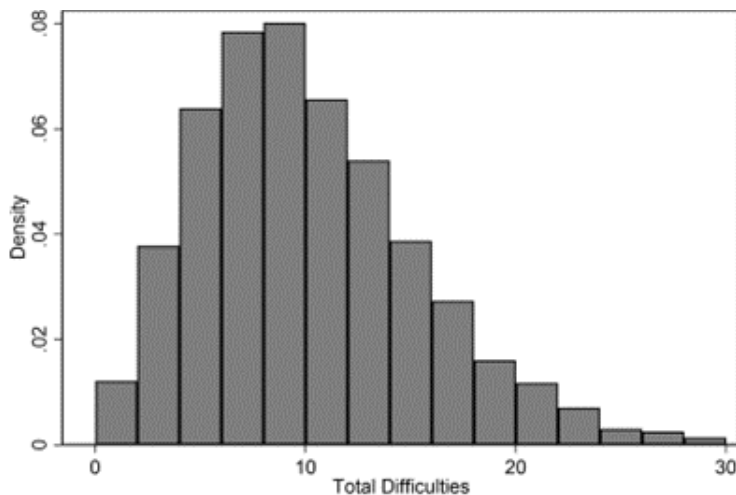Figure 1: Distribution of total difficutiles at age 3



Figure 1 illustrates the distribution of the total difficulties at age 3. The distribution is normal with a few children reported to have very low or very high difficulties.

## 2.4 Conditioning variables

Cohort members were first surveyed at 9 months old, before many of them started childcare, and we therefore use the information given at age 9 month to construct the control variables. Due to the large amount of information held on both the children and their families we are able to control for many variables (see Table 2 in Section 4). We consider the child's characteristics (gender, ethnicity, age, birth-weight, and three indicators of child development[15] ); household characteristics (number

---

[13]Statements include: Often has temper tantrums or hot tempers, generally obedient, often fights with other children, often lies or cheats, steals from home, school or elsewhere, restless and overactive, constantly fidgeting or squirming, easily distracted and concentration wanders, thinks things out before acting, sees tasks through to the end, often complains of headaches, many worries, often unhappy and downhearted, nervous or clingy in new situations, many fears and easily scared, rather solitary and tends to play alone, has at least one good friend, generally liked by other children, picked on or bullied by other children, gets on better with adults than with other children, 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.

[14]For negative statements 'Not true' is coded as zero, 'somewhat true' is given one point and 'Certainly true' two points. For positive statements, the opposite is true.

[15]We include three factors suggested by Del Boc et al., (2018). They are: "s/he waves bye-bye on her/his own when someone leaves", "s/he can pick up a small object using forefinger and thumb only", "s/he can sit up without

of siblings, real weekly equivalent income[16] and region of residence); and the mother's characteristics (age and level of education[17] and economic activity status.)

## 2.5    Instrumental variables

To identify the causal effect of hours in formal childcare on non-cognitive skills, we exploit exogenous variation in hours by using instrumental variables.

The job characteristics of the parents, specifically the mother, are likely to be key determinants of the amount of time children are enrolled in childcare. Using the Labour Force Survey, which is a nationally representative study of the employment circumstances of the UK population, covering approximately 40,000 responding UK households per quarter, we obtain information on a set of job characteristics. We focus on two characteristics, firstly shift work. We make the assumption that mothers who work shifts are likely to place their child in formal childcare for fewer hours per week. This is due to the structured opening hours in formal childcare not working for the unsocial working hours of shift workers. The second characteristic we focus on is having variation in working hours each week. We believe that the uncertainly of hours of working will drive parents to enrol their child in formal childcare for more hours per week, in order to cover all possible times when they might be working.[18]

To try to increase the validity of the instruments we use the probability of shift work and uncertain hours calculated at the three-digit SOC code level. Using the Labour Force Survey, we calculate the percentage of individuals within a three-digit SOC code that report working shift work/having varying hours which we link to the MCS using the three-digit SOC code in which the mother was employed when the child was 9 months old.[19] Further discussion on the quality of the instrument as provided in section 4.

The average percentage of individuals within a three-digit SOC code that reported working shifts was 22.1% which relates to SOC code 221, health professionals for humans and animals. The occupation that reported the highest percentage of shift work at 75.3% was transport associate professionals. Workers in this group command and navigate aircraft and vessels. The occupation reporting the lowest percentage of shift work at 1.1% was architects, town planners and surveyors.

For uncertain hours of work, the mean was 45.5% associated with SOC code 213, information technology and telecommunication professionals. The SOC code associated with the highest percentage of worker with uncertain hours (80.5%) was 121, managers and proprietors in agriculture related services who plan, organise, direct, and control the activities and resources of agricultural, forestry, fishing and similar establishments and services. The SOC code associated with the lowest percentage of workers with uncertain hours (22.1%) was 923, elementary cleaning occupations.

---

being supported"; answers are "often", "once or twice", and "not yet".

[16]The equivalent income is the income of the household taking into account the number of people in the family and assigning weights. The ones provided in the MCS follows the OECD equivalence scale, which assign a value of 1 to the first household member, of 0.7 to each additional adult, and of 0.5 to each child.

[17]The level of education is controlled for by including a range of dummy variables relating to level of qualification.

[18]These assumptions are tested in section 4.

[19]If the mother was unemployed we code the instruments as zero.

# 3    Method

The data allow this research to observe hours spent in childcare by the age of 3 as well as non-cognitive outcomes from age 3-14. This analysis aims to identify the causal effect of hours spent in childcare by the age of 3 on non-cognitive outcomes both in the short run and the medium run. As previously discussed, establishing causality between hours spent in childcare and non-cognitive outcomes faces methodological challenges. There are two primary threats to causality: (i) omitted variable bias, and (ii) measurement error. This section will outline how this research will try to correct for the bias.

The analysis starts by estimating an ordinary least squares (OLS) regression in which we regress total difficulties reported in each year on hours spent in childcare at age 3 and the covariates.[20] The models take the following form:

$$TotalDifficulties_{it} = \beta_0 + \beta_1 Hours_i + \beta_2 Child_i + \beta_3 Family_i + \beta_4 Mother_i + \epsilon_{it} \tag{1}$$

where $TotalDifficulties_{it}$ is the total number of difficulties reported by the mother in the strength and difficulties questionnaire for child i at age t. Hours is the number of reported hours spent in formal childcare at age 3. $Child$ is a vector of individual characteristics including gender, ethnicity, special educational need, birth-month, birth-weight, and three indicators of child development. $Family$ is a vector of household characteristics including the number of siblings, real weekly equivalent income and region of residence. $Mother$ is a vector of the mother's characteristics including age at birth, level of education and economic status.[21] $\epsilon_{it}$ captures the unobservable determinants of non-cognitive skills.[22]

The main interest of this research is the estimation of $\beta_1$, which is the effect of hours in formal childcare on non-cognitive skills. To interpret $\beta_1$ as the causal effect of an hour in formal childcare at age 3 on an individual's non-cognitive skills, we require independent variation in hours in formal childcare, meaning the zero conditional mean assumption must hold, $E(\epsilon_{it}|Child_i, Family_i, Mother_i) = 0$. Due to the endogeneity of hours in formal childcare, it can be argued that this assumption may not hold.

Parents who enrol their children in formal childcare for more hours may differ in both observable and unobservable ways from those who opt for a few hours or none at all. These differences could be driven by child and parent factors. Parents of children with higher natural ability may place their child in formal childcare for more hours because they believe they will be able to cope with the separation from their parents better. On the other hand, they might enrol their child in formal childcare for fewer hours, due to a belief their child does not require early intervention. Parents with more time to devote to their child may have chosen fewer hours of childcare though having the time does not always lead to effective use. More "education oriented" parents may choose to place their child in childcare for many hours at an early age. They may also believe that they can provide their child with a strong early education at home and therefore place them in formal childcare for fewer hours. Many of these factors may be not observable and may influence the child's non-cognitive

---

[20]We make the assumption that hours in formal childcare depends linearly on total reported difficulties. This assumption is discussed in section 4.

[21]All control variables are measured at age 9 months.

[22]A justification for these control variables is provided below.

development, biasing the estimates. The direction of the bias is not straightforward. Take a child's ability for example, whilst this is likely to increase non-cognitive skills, it is not clear whether this would increase or decrease hours in formal childcare. The same can be said for a number of potential omitted variables. The number of recorded hours in childcare may also be subject to measurement error. Much of the literature to date has suggested a negative bias, we therefore interpret the OLS results as a lower bound of the association between hours in childcare and non-cognitive skills.

To check whether this may be the case, we use an instrumental variables approach that leverages exogenous variation in childcare intensity—by the likelihood of the mother being required to work shifts or having varying hours week by week to estimate the unbiased local average treatment effect (LATE) of hours in formal childcare on non-cognitive skills. We make the assumption that the structured nature of formal childcare means that mothers who are more likely to work shift work enrol their child in formal care for fewer hours or make less use of childcare, whereas mothers who are more likely to experience uncertain hours have to enrol their child in formal childcare for more hours due to the uncertainty of knowing in advance when the childcare would be needed. We use a two-stage least-squares (2SLS) regression approach to first estimate the hours in formal childcare as a function of the probability of shift work and uncertain hours for the mother, net of child, household, and mother characteristics. The predicted hours in formal childcare is then forwarded to a second-stage regression to predict the unbiased LATE of hours in formal childcare on non-cognitive skills. The first stage equation takes the following form:

$$Hours_i = \beta_0 + \beta_1 Shift_i + \beta_2 Uncertain_i + \beta_3 Child_i + \beta_4 Family_i + \beta_5 Mother_i + \gamma_{it} \quad (2)$$

Where $Shift_i$ is the percentage of individuals, working in the same three-digit SOC code as the mother, who reported working shift work, and $Uncertain_i$ is the percentage of individuals working in the same three-digit SOC code as the mother, who reported having variation in their weekly hours.

The second-stage equation takes the following form:

$$TotalDifficulties_{it} = \beta_0 + \beta_1 H\hat{o}urs_i + \beta_2 Child_i + \beta_3 Family_i + \beta_4 Mother_i + \epsilon_{it} \quad (3)$$

where $H\hat{o}urs_i$ is the predicted hours of childcare based on the first stage.

The IV strategy requires that two assumptions be met. First, the instruments must be relevant, meaning that they are highly predictive of hours in formal childcare. Second, they must be valid, meaning uncorrelated with the error term in the explanatory (second-stage) equation, such that the instruments affect non-cognitive skills only through their effect on hours in formal childcare.

The first assumption is easily tested and, as shown in the first stage results (Appendix Table 10), holds true in all our models. The second may be violated if, for example, mothers select into professions based on children's needs.

Although we cannot fully rule out this possibility, we attempt to minimise it by adjusting for an extensive array of covariates. Adjusting for these factors should reduce the risk of bias. As a robustness checks we drop observations where mothers report leaving a job due to not having flexible

hours. We also drop observations where mothers report having a change in hours in their new job. Despite the change in sample, the results reported are very similar to the baseline results.

# 4    Results

## 4.1    Descriptive Statistics

This research focuses on hours in formal childcare for children under the age of 3. We focus on an initial sample of 6,567 individuals who record at least one year of non-cognitive skills as well as hours in childcare (see section 2 for further discussion on construction of the sample.) Within the sample, 13% of children have attended formal childcare by the age of 3. 87% have not yet attended formal childcare, of whom 83% have only been cared for by their parents, leaving 17% who have received informal childcare, mainly from the maternal grandmother.

The main variable of interest within this research is non-cognitive development, measured by the parent's response to the strength and difficulties questionnaire as discussed above. Table 1 presents the average number of difficulties reported for ages 3, 5, 7, 11 and 14.[23] Children who receive formal childcare by age 3 have statistically significantly lower reported difficulties at every age compared to children who receive no formal childcare at age 3.

Table 1: Average number of reported difficulties

|  | Childcare | No childcare |
|---|---|---|
| Total difficulties age 3 | 8.02 | 10.07 |
|  | (4.30) | (5.41) |
| Total difficulties age 5 | 6.40 | 7.50 |
|  | (4.32) | (5.02) |
| Total difficulties age 7 | 6.51 | 7.72 |
|  | (4.81) | (5.45) |
| Total difficulties age 11 | 6.91 | 7.92 |
|  | (5.52) | (5.84) |
| Total difficulties age 14 | 7.13 | 8.37 |
|  | (5.71) | (5.95) |

Descriptive statistics for the covariates, shown in Table 2[24], reinforce that there is likely systematic selection into childcare. Focusing on individual characteristics, females, children of white ethnicity, children born with a higher birth weight, having no learning difficulties and having fewer siblings are all associated with a greater likelihood of attending childcare by age 3. Differences are also evident for family and parental characteristics. For example, children of less educated, lower-income, and younger parents are less likely to attend formal childcare by the age of 3.[25] All covariates, apart

---

[23]Standard deviation is in parentheses.

[24]Standard deviation is in parentheses.

[25]Mother's education is a dummy variable equal to one if the mother has a degree or equivalent. In the main estimations we include mother's education as a range of dummy variables. 14.97% of mothers in the sample have no

from gender, are statistically significant. As discussed in section 2, we use three indicators of child development measured at age 9 months to proxy ability,but we do not find significant differences.

Table 2: Summary statistics for covariates

| Variable | Childcare | No Childcare |
|---|---|---|
| Female | 49.38% | 49.05% |
| | (0.50) | (0.50) |
| White | 85.39% | 76.37% |
| | (0.35) | (0.76) |
| SEN | 11.62% | 19.44% |
| | (0.32) | (0.40) |
| Birth Weight | 3.40 | 3.33 |
| | (0.52) | (0.59) |
| Development 1 | 1.13 | 1.14 |
| | (0.42) | (0.45) |
| Development 2 | 1.94 | 1.95 |
| | (0.83) | (0.84) |
| Development 3 | 1.02 | 1.06 |
| | (0.15) | (0.31) |
| Number of Siblings | 0.55 | 0.99 |
| | (0.77) | (1.10) |
| Year of mother's birth | 1969 | 1972 |
| | (5.44) | (5.93) |
| Mother degree | 44.91% | 16.46% |
| | (0.49) | (0.37) |
| Mother employed | 71.62% | 47.45% |
| | (0.45) | (0.49) |
| Equivalised household income | 522.19 | 307.35 |
| | (249.74) | (209.96) |

This research is predominantly interested in the impact of the intensity of childcare measured by hours. The average number of hours per week being cared for by someone other than their parents is 20. Children who attend formal childcare by age 3 on average, spent 27 hours per week in that setting whilst children in informal care settings spent 17 hours per week.

To conclude the descriptive analysis, Figure 2 documents the association between the average number of reported difficulties across all years surveyed and hours spent in formal childcare. As expected, the correlation is negative. Spending up to 8 hours per week in childcare is associated with an average of 8.6 reported difficulties, spending 40 hours a week in childcare is associated with an average of 7.9 reported difficulties.

Appendix figures 3-7 show the correlation by each surveyed year. Reported difficulties reduce over time and the correlation gets weaker.

---

qualifications, 47.63% receive school qualifications including GCSEs, A-levels or equivalent, 28.68% received a degree or higher or equivalent, 8.71% hold other qualifications.

Figure 2: Descriptive relationship between hours in childcare and non-cognitive skills between age 3-14
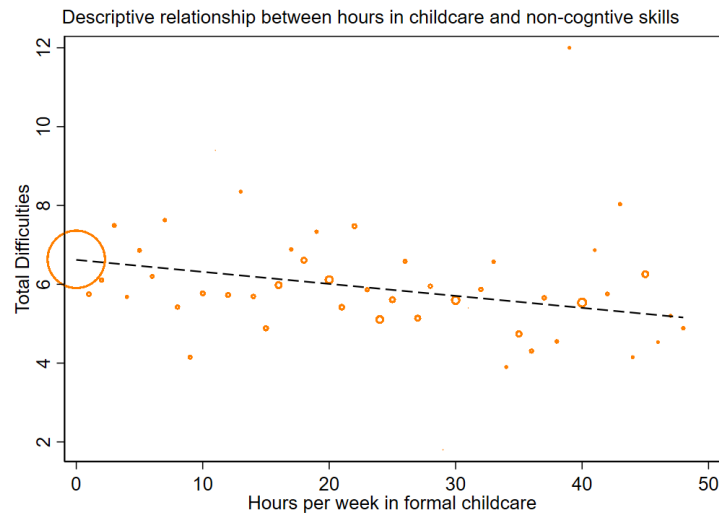

Descriptive relationship between hours in childcare and non-cogntive skills

Notes: Correlation between hours per week in formal childcare at age 3 and average reported difficulties from age 3-14. Reported difficulties are collapsed on the integer of hours in childcare. The size of the marker indicates the relative number of observations in the hours cell. The fitted line is taken from a simple linear regression of reported difficulties on weekly hours in childcare.

## 4.2 Regression Results

Table 3 contains the OLS estimates for non-cognitive skills from age 3-14 from the production function specified in equation (1). The OLS results are positive and small in magnitude. The estimated coefficients indicate negative impacts of hours in childcare on non-cognitive skills. As previously discussed, these estimates are likely biased. Previous childcare research has found their unbiased estimates larger than OLS, suggesting a downwards bias.

Table 3: OLS results

|  | (Non-cog 3) | (Non-cog 5) | (Non-cog 7) | (Non-cog 11) | (Non-cog 14) |
|---|---|---|---|---|---|
| Hours | 0.001 | 0.001 | 0.001 | 0.002** | 0.001 |
|  | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| N | 6387 | 6567 | 5926 | 5617 | 5118 |
| $R^2$ | 0.16 | 0.16 | 0.15 | 0.12 | 0.11 |

*Notes: The dependent variable is the total number of reported difficulties at each age surveyed. The dependent variable is standardised with a mean of zero and a standard deviation of one. Hours is the number of hours in formal childcare at age 3. Controls include gender, ethnicity, month of birth, SEN, birth weight, development at 9 months old, location, mother's age at birth, number of siblings, mother's education, mother's economic status and household income. All controls are measured at age 9 months. Standard errors are in parentheses. We assume iid errors.[a] The change in sample size is due to non-response and attrition. The impacts of the change in the sample size are examined in the robustness checks.*

*\* $p < 0.10$, \*\* $p < 0.05$, \*\*\* $p < 0.01$*

[a]We also estimated all models with non-iid errors and results remained similar.

Turning to the IV results, the instruments perform well. The first-stage F statistics (Cragg-Donald Wald F statistic) range from 43.5-39.3. Tests for both under and over identification are satisfied in each model. The weak instrument (Anderson-Rubin) robustness test of the joint significance of the instruments in the reduced form model is satisfied for non-cognitive skills at ages 3-7 but not age 11 and 14.

The first stage estimates (see Appendix Table 10) suggest that a 10 percentage points increase in the probability of the mother working shift work, reduces hours in formal childcare by 51.6 minutes per week. For a 10-percentage point increase in the probability of having uncertain hours, hours in formal care increase by 44.4 minutes per week.

The second-stage IV results, shown in Table 5[26], are larger in magnitude than the OLS estimates.[27] The IV results indicate that on average, increasing hours in formal childcare at age 3 by 1, reduces reported difficulties at age 3 by 3.8% of a standard deviation. This reduces to 1.9% of a standard deviation by age 5, followed by 3.4% at age 7[28] and 1.7% and 1.6% at age 11 and 14 respectively. When looking at the robust confidence intervals, we find significant impacts for children aged 3 to 7 however the impact declines over time and slips into insignificance at ages 11 and 14. We therefore argue that hours in childcare has a significant impact on initial non-cognitive skills which persists in the medium term. However, the findings for the long run are less clear.

Another way to interpret these findings is that increasing childcare at age 3 by half a day (4 hours) per week suggests that reported difficulties would reduce by slightly more than half a reported difficulty. Large findings are not uncommon within the literature. Felfe et al. (2015), Datta Gupta

---

[26]Full model second stage results are shown in table 11 in the appendix.

[27]We calculate low values of the coefficient of proportionality, as suggested by Ciacci (2021), which offer supportive evidence that IV estimates are not too large with respect to OLS.

[28]The reduction in the impact at age 5 followed by an increase in the estimates impact at age 7 is consistent with some of the literature. Peter et al (2016), who also use data from the Millennium Cohort Study, report similar findings. 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). As children enter school at age 5, there is a suggestion that such a big event in a child's life makes it difficult for the parent to correctly rank their difficulties.

and Simonsen (2016) and Berger et al. (2021) all estimate impacts of attending childcare from 15-34% of a standard deviation.

Table 4: Over-identified IV model

|  | (Non-cog 3) | (Non-cog 5) | (Non-cog 7) | (Non-cog 11) | (Non-cog 14) |
|---|---|---|---|---|---|
| Hours | -0.038*** | -0.019** | -0.034*** | -0.017* | -0.016* |
|  | (0.009) | (0.008) | (0.009) | (0.009) | (0.009) |
|  |  |  |  |  |  |
| First Stage F | 43.5 | 43.2 | 39.9 | 39.3 | 39.4 |
| A-R Wald (P Value) | 0.00 | 0.03 | 0.00 | 0.15 | 0.16 |
| Hansan J Stat (P Value) | 0.50 | 0.13 | 0.77 | 0.85 | 0.78 |
| A-R CI | [-0.063, -0.019] | [-0.038,-0.004] | [-0.061, -0.013] | [-0.041, 0.004] | [-0.041, 0.005] |
| N | 6387 | 6567 | 5926 | 5617 | 5118 |

*Notes: The dependent variable is the total number of reported difficulties at each age surveyed. The dependent variable is standardised with a mean of zero and a standard deviation of one. Hours is the number of hours in formal childcare at age 3. Cragg-Donald Wald F statistic and the p-values associated with the Hansan J overidentification test and the weak instrument (Anderson-Rubin) robustness test are shown as well as the 95% confidence intervals. Standard errors are in parentheses. Controls include gender, ethnicity, month of birth, SEN, birth weight, development at 9 months old, whether the child has older siblings, mother's age at birth, mother's education, mother's economic status and household income. All controls are measured at age 9 months. The change in sample size is due to non-response and attrition.[a] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$*

[a]We run a model where the sample size is reduced to 4148 individuals who responded to every wave and the results did not change.

The difference between the OLS and IV results may reflect that children who are exogenously induced into hours in childcare based on their mother's work patterns are more heavily impacted by formal childcare than children whose parents select them into hours in childcare regardless of their work patterns. We may also expect mothers on lower incomes to have more uncertain work patterns.

The overall pattern of results suggests that any bias induced by the endogeneity of hours in formal childcare likely results in underestimation of the positive effect of hours in formal childcare on non-cognitive skills.[29]

The simplifying assumption made in the baseline model is that there are neither diminishing nor increasing benefits to hours in childcare. Whilst we may expect increasing or decreasing returns, we are unable to examine the non-linear impact as we have been unsuccessful in finding an instrument for the quadratic of hours in childcare at age 3 that satisfies weak instrument testing.

[29]To test for the potential of omitted variables, we run the model on a smaller sample who reported preferences for childcare and the distance lived from grandparents. We find similar estimated coefficients.

### 4.3 Heterogeneity

On average, increasing the amount of time spent in formal childcare has a positive impact on non-cognitive development both in the short and medium term. The next stage of this analysis explores the extent to which the effect of hours in formal childcare differs across observable characteristics. Understanding heterogeneous impacts is important for policy development as it can identify groups which may benefit the most from the intervention.

The Strength and Difficulties questionnaire asks about 25 attributes, some positive and others negative. These 25 items are divided between 5 scales: emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems and prosocial behaviour. For the baseline estimates we use total difficulties as the dependent variable. We start examining the heterogeneous impact by estimating the impact of hours in childcare across the 5 scales.[30]

Table 5: Impact of hours in formal care on type of non-cognitive skill estimated by an over-identified IV model

|                  | (Emotion)   | (Conduct)   | (Hyper)   | (Peer)      | (Pro-social) |
|------------------|-------------|-------------|-----------|-------------|--------------|
| Impact at age 3  | -0.017**    | -0.029***   | -0.025**  | -0.026***   | 0.006        |
|                  | (0.008)     | (0.008)     | (0.008)   | (0.008)     | (0.006)      |
| Impact at age 5  | -0.010      | -0.020**    | -0.010    | -0.015*     | 0.004        |
|                  | (0.008)     | (0.008)     | (0.008)   | (0.008)     | (0.006)      |
| Impact at age 7  | -0.031***   | -0.029***   | -0.014    | -0.030***   | -0.007       |
|                  | (0.009)     | (0.009)     | (0.009)   | (0.009)     | (0.006)      |
| Impact at age 11 | -0.014      | -0.016*     | -0.007    | -0.017*     | -0.008       |
|                  | (0.009)     | (0.009)     | (0.009)   | (0.009)     | (0.009)      |
| Impact at age 14 | -0.012      | -0.021**    | -0.013    | -0.004      | 0.002        |
|                  | (0.009)     | (0.009)     | (0.009)   | (0.009)     | (0.007)      |

*Notes: The dependent variable is the total number of reported difficulties in each sub-section at each age. The dependent variables are standardised with a mean of zero and a standard deviation of one. Hours is the number of hours in formal childcare at age 3. Standard errors in parentheses. Controls include gender, ethnicity, month of birth, SEN, birth weight, development at 9 months old, whether the child has older siblings, mother's age at birth, mother's education and household income. All controls are measured at age 9 months. Standard errors are in parentheses. $* p < 0.10$, $** p < 0.05$, $*** p < 0.01$*

We find no significant impact on pro-social behaviour. Pro-social behaviour is not included in total difficulties and is a measure of positive non-cognitive skills. We therefore argue that time in childcare at age 3, improves (i.e. reduces) negative non-cognitive skills but has no effect on positive non-cognitive skills. We find significant improvements in all other measures of strength and difficulties for age 3 and all, expect hyper, for age 5 and 7. The impact on emotional skills, conduct and peer relationships could be argued to be the most persistent over time. Difficulties related to

---

[30]First stage F statistics range from 43.5-39.3 and over-identification checks are valid for all models.

emotional skills are worrying, being nervous and scared. For conduct, difficulties include getting into fights, getting angry and not doing as they have been told. In relation to peer relationship, difficulties include being solitary and getting on better with adults. Spending more time in childcare at a young age could be argued to increase independence, get children used to being out of their comfort zone, increasing interaction with other children and being in a more structured environment where people are trained to help children understand their emotions.

We also conduct a series of subgroup analyses based on individual and family characteristics; the results are shown below.[31]

We start by examining the impact across gender.

Table 6: The impact of hours in childcare at age 3 on total difficulties, by gender. (Estimated by an over-identified IV model)

|  | (Female) | (Male) |
|---|---|---|
| Impact at age 3 | -0.035*** | -0.043*** |
|  | (0.013) | (0.012) |
| Impact at age 5 | -0.022* | -0.019* |
|  | (0.012) | (0.011) |
| Impact at age 7 | -0.035*** | -0.033*** |
|  | (0.013) | (0.013) |
| Impact at age 11 | -0.010 | -0.025** |
|  | (0.013) | (0.012) |
| Impact at age 14 | -0.001 | -0.031** |
|  | (0.013) | (0.013) |

*Notes: See Table 5*

With respect to gender, we find positive impact of hours in childcare for both males and females. We estimate a larger impact for males which is persistent over time.

Blanden et al (2022), who also use English data, find no significant impact for females. Studies looking at other European countries for example, Havnes and Mogstad (2015) and Felfe et al., (2015), both find larger impacts for females.

We also examine whether there is a difference in impact depending on if the child has older siblings. We could assume that children who are first born may be most impacted by time in childcare as their parents have less experience. On the other hand, children with older siblings may be impacted more by increasing time in childcare as they have to share their parents' time with their siblings. Tables 7 shows that it is first born children which seem to benefit the most from spending more time in childcare, this effect is also persistent over time.

---

[31]We estimate all subgroup analysis using the over-identified IV model. First stage F statistics remain above 10 for all subgroups and overidentification tests are satisfied.

Table 7: The impact of hours in childcare at age 3 on total difficulties, by first born status. (Estimated by an over-identified IV model)

|  | (First Born) | (Older Siblings) |
|---|---|---|
| Impact at age 3 | -0.039*** | -0.032*** |
|  | (0.013) | (0.012) |
| Impact at age 5 | -0.024** | -0.011 |
|  | (0.012) | (0.011) |
| Impact at age 7 | -0.033** | -0.028** |
|  | (0.013) | (0.012) |
| Impact at age 11 | -0.017 | -0.015 |
|  | (0.013) | (0.012) |
| Impact at age 14 | -0.012 | -0.019 |
|  | (0.012) | (0.013) |

*Notes: See Table 5*

There is some degree of consensus in the literature that childcare is more beneficial for disadvantaged children than for their more advantaged counterparts. To examine this, we carry out subsample analyses using mother's education[32] and equivaled household income.[33] Table 8 present the results for mother's education whilst table 9 shows the estimated impact across the household income distribution.

Table 8: The impact of hours in childcare at age 3 on total difficulties, by mother's education. (Estimated by an over-identified IV model)

|  | (Degree or higher) | (No degree) |
|---|---|---|
| Impact at age 3 | -0.029** | -0.044*** |
|  | (0.011) | (0.011) |
| Impact at age 5 | -0.027** | -0.018* |
|  | (0.011) | (0.011) |
| Impact at age 7 | -0.036*** | -0.034*** |
|  | (0.013) | (0.012) |
| Impact at age 11 | -0.016 | -0.014 |
|  | (0.011) | (0.011) |
| Impact at age 14 | -0.019* | -0.011 |
|  | (0.011) | (0.011) |

*Notes: See Table 5*

Whilst we find that the initial impact of hours in childcare is larger for children with mothers who have lower levels of education, this impact declines over time and becomes insignificant. The impact for children with mothers of high levels of education remains borderline significant at age 14. We could therefore argue that children who have lower educated mothers do benefit greatly from time spent in childcare however, children with higher educated mothers are better are maintaining the skills learnt.

---

[32]Mother's education is a dummy variable equal to one if she has a degree or equivalent

[33]To examine the impact of childcare attendance across the income distribution we generate three dummy variables relating to the bottom 20%, top 20% and middle 60% of the distribution.

Table 9: The impact of hours in childcare at age 3 on total difficulties, by household income. (Estimated by an over-identified IV model)

|  | (Top 20%) | (Middle 60%) | (Bottom 20%) |
|---|---|---|---|
| Impact at age 3 | -0.023** | -0.038 *** | -0.094 |
|  | (0.009) | (0.009) | (0.109) |
| Impact at age 5 | -0.023** | -0.022** | -0.239 |
|  | (0.009) | (0.009) | (0.235) |
| Impact at age 7 | -0.023** | -0.030*** | -0.102 |
|  | (0.011) | (0.009) | (0.080) |
| Impact at age 11 | -0.005 | -0.019** | -0.027 |
|  | (0.009) | (0.009) | (0.071) |
| Impact at age 14 | -0.007 | -0.016* | -0.052 |
|  | (0.010) | (0.009) | (0.107) |

*Notes: See Table 5*

Regarding household income, as expected, we find the largest impact on children in households with income levels which fall into the lowest 20% of the distribution, followed by those in the middle and the smallest impact is found for those at the top of the distribution. It should be noted that due to children from households with higher income being more likely to enrol in childcare for more hours, the sample of children in the bottom 20% of the distribution which has hours in childcare larger than zero is small. The instruments are weak for the sub-sample analysis on the bottom 20% and should therefore be interpreted with caution. Since the bias will be in the direction of the OLS we can interpret these estimates as lower bounds.

# 5    Conclusion

This research provides the first attempt at identifying the causal relationship between hours in childcare prior to the age of 3 and non-cognitive skills reported at ages 3-14. We aim to shed light on the causal relationship by adopting an instrumental variables strategy that leverages exogenous variation in both the probability that the mother works shift work and has uncertain working hours. Our IV results suggest that the relationship between hours in childcare and non-cognitive skills is likely causal and that the more naive OLS estimates likely underestimate the causal effect.

Using the Millennium Cohort Study which focuses on children born in the UK in the year 2000, we find that increasing hours in formal childcare at age 3 by 1, reduces reported difficulties at age 3 by 3.8% of a standard deviation. This reduces to 1.9% of a standard deviation by age 5, followed by 3.4% at age 7. We also find smaller impacts of 1.7% and 1.6% at age 11 and 14 respectively. These findings show that time spent in childcare at age 3 has initial impacts on non-cognitive skills which persist into the medium run. In respect to the long run, we are cautious in our interpretation as the impacts slip into insignificance when performing weak instrument robust testing. The fact that these effects do persist underlines the importance of early childhood education.

Our findings, especially for the medium run impact, deliver very robust results across different samples of individuals and measures of non-cognitive skills. The robustness checks consistently

suggest that the existence of weak instruments or omitted variable bias is rather unlikely.

How do our findings fit within the literature? We estimate positive effects of increased time in childcare on non-cognitive skills whereas Datta Gupta and Simonsen (2010) and Berger et al. (2021) who provide descriptive evidence on the impact of time spent in childcare, both estimate negative effects.[34] Firstly, our estimates represent LATEs rather than average treatment effects. As such, they are less readily comparable to estimates from prior work. Secondly, the literature on the impact of hours spent in childcare is very small and for the most part descriptive, making it impossible to draw any general conclusions.

This research has useful policy implications. Firstly, consistent with the broader literature on childcare enrolment, we estimate positive effects of childcare. There is a current interest in the optimal number of hours in childcare in England due to the upcoming changes in childcare subsidies. The subsided hours in childcare for working parents are increasing from 30 hours for 3- and 4-year-olds to cover all under 5-year-olds by 2025. Whilst this policy was driven by the impact on female labour supply, our findings provide evidence to show that the policy change also has benefits for the child.

Secondly, results from our sub-group analysis suggest that there is heterogeneity in the magnitude and persistence of impacts across population subgroups. Most notably, less advantaged children—particularly those with low-educated mothers and those in lower-income households appear to benefit most from hours spent in childcare. These findings are consistent with a large proportion of the literature which focuses on childcare for disadvantaged children. These findings support the argument that facilitating access to more time in childcare for disadvantaged children may hold potential for decreasing early socioeconomic disparities in child development. Though we also find that these initial impacts do not persist, these findings require further investigation to understand whether disadvantaged children need more assistance to maintain skills learnt in early childcare.

These results should be interpreted in the context of a few limitations. Firstly, they are English specific and therefore should be generalised with caution. Secondly, the counterfactual condition to more hours in childcare in our IV analysis is heterogenous. Children spending fewer hours in childcare could be in many different childcare settings for different lengths of time.[35] Heterogeneity in the counterfactual condition is a common limitation of studies of childcare. Morris et al. (2018) show that the benefits of high-quality centre-based care are more pronounced when compared with parental care than with other types of formal childcare. Finally, our IV analyses rely on the assumption that mothers do not select into industries based on their child's non-cognitive skills. If this assumption is incorrect, our IV estimates will be biased. Although our sensitivity analyses help to reduce such concerns, we cannot be certain of the absence of such behaviours.

In future research, it would be worthwhile to explore the impact across different care settings, assess the functional form, extend the analyses to cognitive measures and investigate some of the mechanisms involved in the impact.

This research has contributed to the literature by investigating the causal relationship between hours in childcare and non-cognitive skills. We estimate positive findings which we argue suggests that an expansion of subsided hours in childcare will be beneficial for all, especially children from

---

[34]There is a larger literature which examines the impact of attending childcare which also find mixed evidence.
[35]Due to sample size, we cannot explore this further.

disadvantaged backgrounds, which may have potential to contribute to decreasing early gaps in child development.

# 6 Appendix

Figure 3: Descriptive relationship between hours in childcare and non-cognitive skills at age 3
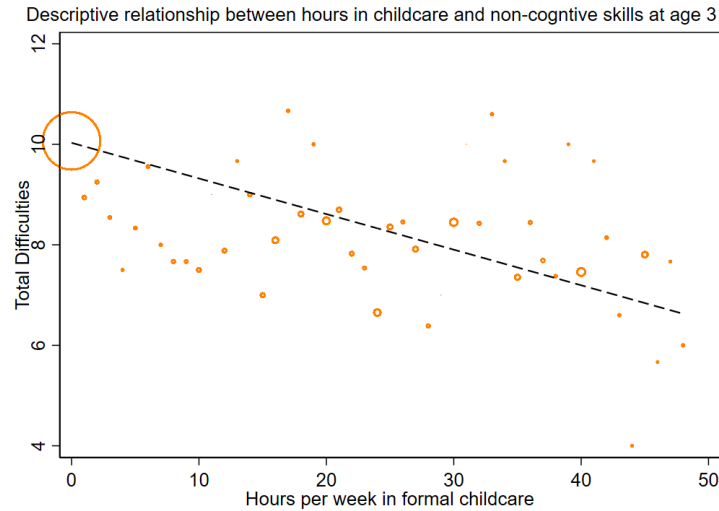


Notes: Correlation between hours per week in formal childcare at age 3 and average reported difficulties at age age 3. Reported difficulties are collapsed on the integer of hours in childcare. The size of the marker indicates the relative number of observations in the hours cell. The fitted line is taken from a simple linear regression of reported difficulties on weekly hours in childcare.

Figure 4: Descriptive relationship between hours in childcare and non-cognitive skills at age 5



Notes: Correlation between hours per week in formal childcare at age 3 and average reported difficulties at age 5. Reported difficulties are collapsed on the integer of hours in childcare. The size of the marker indicates the relative number of observations in the hours cell. The fitted line is taken from a simple linear regression of reported difficulties on weekly hours in childcare.

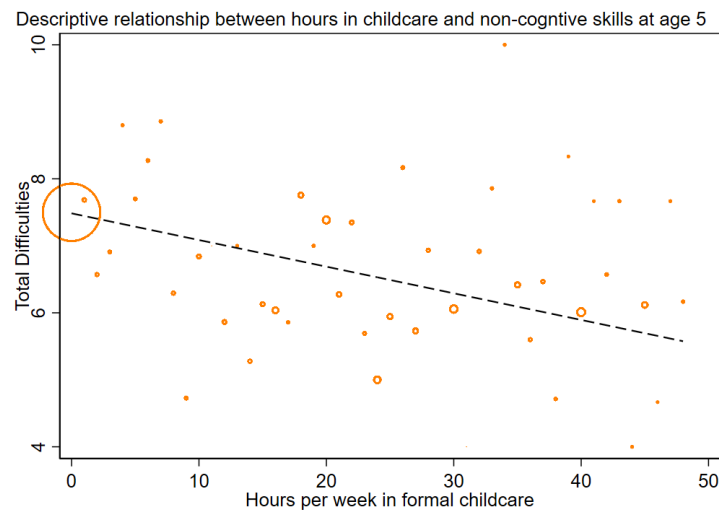Figure 5: Descriptive relationship between hours in childcare and non-cognitive skills at age 7



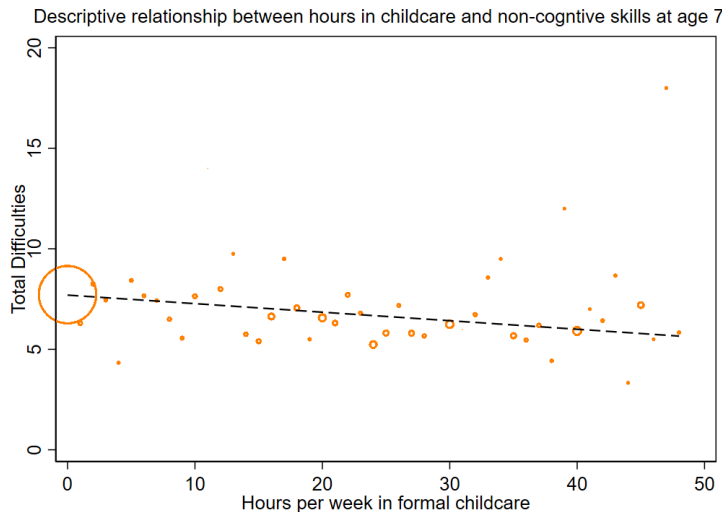Descriptive relationship between hours in childcare and non-cogntive skills at age 7

Notes: Correlation between hours per week in formal childcare at age 3 and average reported difficulties at age 7. Reported difficulties are collapsed on the integer of hours in childcare. The size of the marker indicates the relative number of observations in the hours cell. The fitted line is taken from a simple linear regression of reported difficulties on weekly hours in childcare.

Figure 6: Descriptive relationship between hours in childcare and non-cognitive skills at age 11



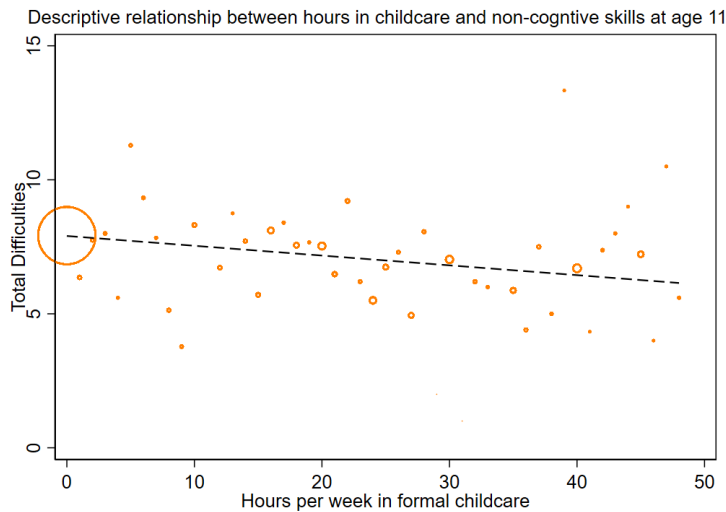Descriptive relationship between hours in childcare and non-cogntive skills at age 11

Notes: Correlation between hours per week in formal childcare at age 3 and average reported difficulties at age 11. Reported difficulties are collapsed on the integer of hours in childcare. The size of the marker indicates the relative number of observations in the hours cell. The fitted line is taken from a simple linear regression of reported difficulties on weekly hours in childcare.

Figure 7: Descriptive relationship between hours in childcare and non-cognitive skills at age 14



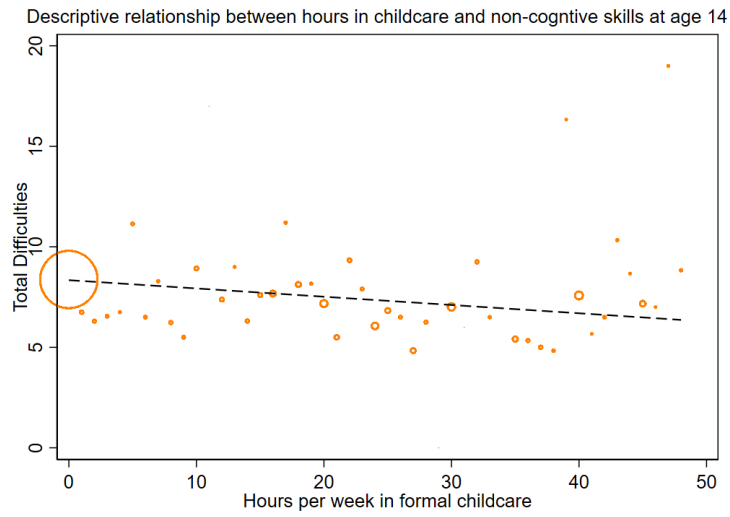Descriptive relationship between hours in childcare and non-cogntive skills at age 14

Notes: Correlation between hours per week in formal childcare at age 3 and average reported difficulties at age 14. Reported difficulties are collapsed on the integer of hours in childcare. The size of the marker indicates the relative number of observations in the hours cell. The fitted line is taken from a simple linear regression of reported difficulties on weekly hours in childcare.

Table 10: First Stage

|                        | (First Stage) |
|------------------------|---------------|
| Shift work (IV)        | -0.086***     |
|                        | (0.010)       |
| Uncertain hours (IV)   | 0.074***      |
|                        | (0.013)       |
| Female                 | -0.472        |
|                        | (0.316)       |
| Mixed ethnicity        | 4.431***      |
|                        | (1.500)       |
| Asian ethnicity        | 0.439         |
|                        | (0.543)       |
| Black ethnicity        | 6.949***      |
|                        | (0.831)       |
| Other ethnicity        | 1.658         |
|                        | (1.162)       |
| Month of birth         | -0.002        |
|                        | (0.044)       |
| Birth weight           | -0.221        |
|                        | (0.283)       |
| SEN                    | -0.076        |
|                        | (0.550)       |
| Siblings               | -1.349***     |
|                        | (0.178)       |
| Household income       | 0.013***      |
|                        | (0.001)       |
| Mother L1 education    | 0.117         |
|                        | (0.725)       |
| Mother L2 education    | 0.534         |
|                        | (0.570)       |
| Mother L3 education    | 1.292**       |
|                        | (0.656)       |
| Mother L4 education    | 0.534         |
|                        | (0.744)       |
| Mother L5 education    | 1.624**       |
|                        | (0.688)       |
| Mother L6 education    | 2.989***      |
|                        | (1.032)       |
| Mother other education | 1.019         |
|                        | (0.733)       |
| Mother age             | -0.029        |
|                        | (0.033)       |
| Mother employed        | 3.886***      |
|                        | (1.181)       |
| Mother self-employed   | -1.945        |
|                        | (1.364)       |
| Development 1          | 0.304         |
|                        | (0.362)       |
| Development 2          | -0.102        |
|                        | (0.192)       |
| Development 3          | -0.359        |
|                        | (0.588)       |
| N                      | 6930          |

Notes: First stage regression for hours in childcare at age 3. All controls are measured at age 9 months. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

Table 11: Second stage over-identified IV model of the impact of hours in formal childcare at age 3 on non-cogntive skills at age 3

|  | (Non-cognitive age 3) |
|---|---|
| Hours | -0.038*** |
|  | (0.009) |
| Female | -0.177*** |
|  | (0.027) |
| Mixed ethnicity | 0.150 |
|  | (0.132) |
| Asian ethnicity | 0.305*** |
|  | (0.045) |
| Black ethnicity | 0.247*** |
|  | (0.090) |
| Other ethnicity | 0.114 |
|  | (0.098) |
| Month of birth | 0.001 |
|  | (0.004) |
| Birth weight | -0.044* |
|  | (0.024) |
| SEN | 0.348*** |
|  | (0.046) |
| Siblings | -0.067*** |
|  | (0.020) |
| Household income | -0.000 |
|  | (0.000) |
| Mother age | 0.018*** |
|  | (0.003) |
| Mother L1 education | -0.135** |
|  | (0.060) |
| Mother L2 education | -0.222*** |
|  | (0.048) |
| Mother L3 education | -0.244*** |
|  | (0.048) |
| Mother L4 education | -0.334*** |
|  | (0.062) |
| Mother L5 education | -0.350*** |
|  | (0.061) |
| Mother L6 education | -0.275*** |
|  | (0.093) |
| Mother other education | -0.183*** |
|  | (0.063) |
| Mother employed | 0.143 |
|  | (0.104) |
| Mother self-employed | -0.146 |
|  | (0.112) |
| Development 1 | 0.106*** |
|  | (0.030) |
| Development 2 | 0.052*** |
|  | (0.016) |
| Development 3 | 0.188*** |
|  | (0.049) |
| N | 6930 |

*Notes: The dependent variable is the total number of reported difficulties at age 3. All controls are measured at age 9 months. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$*

# 7 References

Almond, D. and Currie, J., 2011. Killing me softly: The fetal origins hypothesis. *Journal of economic perspectives*, 25(3), pp.153-172.

Almlund, M., Duckworth, A.L., Heckman, J. and Kautz, T., 2011. Personality psychology and economics. *Handbook of the Economics of Education* (Vol. 4, pp. 1-181). Elsevier.

Andrews, J. (2017) *Closing the gap? Trends in Educational Attainment and Disadvantage.* London: Education Policy Institute.

Berger, L.M., Panico, L. and Solaz, A., 2021. The impact of center-based childcare attendance on early child development: evidence from the French Elfe cohort. *Demography*, 58(2), pp.419-450.

Black, S.E., Devereux, P.J. and Salvanes, K.G., 2007. From the cradle to the labor market? The effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, 122(1), pp.409-439.

Blanden, J., Del Bono, E., Hansen, K. and Rabe, B., 2022. Quantity and quality of childcare and children's educational outcomes. *Journal of Population Economics*, 35(2), pp.785-828.

Borghans, L., Duckworth, A.L., Heckman, J.J. and Ter Weel, B., 2008. The economics and psychology of personality traits. *Journal of human Resources*, 43(4), pp.972-1059.

Case, A., Fertig, A. and Paxson, C., 2005. The lasting impact of childhood health and circumstance. *Journal of health economics*, 24(2), pp.365-389.

Ciacci, R., 2021. A Matter of Size: Comparing IV and OLS estimates.(Unpublished)

Darlington, C., 2021. *The childcare market in England.* London: Children England.

Del Boca, D., Piazzalunga, D. and Pronzato, C., 2018. The role of grandparenting in early childcare and child outcomes. *Review of Economics of the Household*, 16, pp.477-512.

Elek, C., Gubhaju, L., Lloyd-Johnsen, C., Eades, S. and Goldfeld, S., 2020. Can early childhood education programs support positive outcomes for indigenous children? A systematic review of the international literature. *Educational Research Review*, 31, p.100363.

Felfe, C., Nollenberger, N. and Rodríguez-Planas, N., 2015. Can't buy mommy's love? Universal childcare and children's long-term cognitive development. *Journal of population economics*, 28, pp.393-422.

Fort, M., Ichino, A. and Zanella, G., 2020. Cognitive and noncognitive costs of day care at age 0–2 for children in advantaged families. *Journal of Political Economy*, 128(1), pp.158-205.

Goodman, R., 1997. The Strengths and Difficulties Questionnaire: a research note. *Journal of child psychology and psychiatry*, 38(5), pp.581-586.

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

Gupta, N.D. and Simonsen, M., 2016. Academic performance and type of early childhood care. *Economics of Education Review*, 53, pp.217-229.

Havnes, T. and Mogstad, M., 2015. Is universal child care leveling the playing field?. *Journal of public economics*, 127, pp.100-114.

Heckman, J.J. and Kautz, T., 2014. Achievement tests and the role of character in American life.

Heckman, J.J. and Rubinstein, Y., 2001. The importance of noncognitive skills: Lessons from the GED testing program. *American economic review*, 91(2), pp.145-149.

Kuehnle, D. and Oberfichtner, M., 2020. Does starting universal childcare earlier influence children's skill development?. *Demography*, 57(1), pp.61-98.

Roberts, B.W., Kuncel, N.R., Shiner, R., Caspi, A. and Goldberg, L.R., 2007. The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological science*, 2(4), pp.313-345.

Lindqvist, E. and Vestman, R., 2011. The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment. *American Economic Journal: Applied Economics*, 3(1), pp.101-128.

Mueller, G. and Plug, E., 2006. Estimating the effect of personality on male and female earnings. *Ilr Review*, 60(1), pp.3-22.

Peter, F.H., Schober, P.S. and Spiess, K.C., 2016. Early birds in day care: The social gradient in starting day care and children's non-cognitive skills. *CESifo Economic Studies*, 62(4), pp.725-751.