Childcare Use and Its Role in Indigenous Child Development: Evidence from the Longitudinal Study of Indigenous Children in Australia*

Francisco Azpitarte[†,‡], Abraham Chigavazira[†], Guyonne Kalb[†,§], Brad M. Farrant[‖], Francisco Perales[#] and Stephen R. Zubrick[‖,¶]

[†] Melbourne Institute of Applied Economic and Social Research,
The University of Melbourne
[‡] Brotherhood of St Laurence
[§] Institute of Labor Economics (IZA)
[‖] Telethon Kids Institute, The University of Western Australia
[#] Institute for Social Science Research, The University of Queensland
[¶] Centre for Child Health Research, The University of Western Australia

March 2017

* For correspondence, contact Francisco Azpitarte; e-mail: fraz@unimelb.edu.au.

Abstract

We investigate patterns of childcare use and its influence on the cognitive development of Indigenous children. The influence of childcare on Indigenous children's cognitive outcomes is less well understood than for non-Indigenous children due to a lack of appropriate data. We focus on a cohort of Indigenous children in Australia who have been followed from infancy and for whom rich information on childcare use and cognitive outcomes is observed. Compared to Indigenous children who never participated in childcare, Indigenous children who participated in childcare performed better on several early cognitive outcomes. Using regression and propensity score matching, we show that this difference is driven by selection into childcare, with children from more advantaged families being more likely to attend formal childcare. However, matching analysis results suggest that relatively disadvantaged children might benefit more from attending childcare, as indicated by the positive estimated effects found for those who never attended childcare.

1. Introduction

Disadvantage is widespread amongst Indigenous populations across the world. As the recent study for 28 populations in 23 countries by Anderson *et al.* (2016) documents, Indigenous disadvantage starts from birth through higher infant mortality rates and continues throughout life, leading to poorer health outcomes and shorter life expectancy in Indigenous communities. Importantly, the disadvantage experienced by these communities has consequences for children's development. Indigenous children in developed countries tend to perform more poorly than their non-Indigenous counterparts on a range of cognitive scores. In their study of Australian children Leigh and Gong (2009) found that Indigenous children at age five had similar cognitive scores as non-Indigenous children at age four. In New Zealand, Māori children achieved significantly lower levels on conventional school literacy measures than non-Māori children at the time of school entry (Phillips *et al.*, 2004; and references therein). Similarly, Indian-American children in the U.S. and Aboriginal children in Canada have lower levels of reading and numeracy skills in early school years than non-minority groups (Clotfelter *et al.,* 2009; Fischer and Stoddard, 2013; Richards *et al.*, 2010). In light of these disparities and given the importance of these skills for outcomes later in life, closing the gap in outcomes between Indigenous and non-Indigenous children in the early years has become a policy goal in countries with Indigenous populations (see for Australia: Department of the Prime Minister and Cabinet, 2015; and for New Zealand: Gould, 2000).

One of the most prominent policy interventions aimed at addressing the gap in outcomes between Indigenous and non-Indigenous children is improving the access to early childhood education.[1] Policy initiatives that promote access to learning opportunities in the early years are particularly important in supporting disadvantaged children. Although the literature on childcare and preschool programmes has not converged toward a unanimous evaluation regarding their effect on cognitive development,[2] several studies for the US show that such programmes may be effective in improving school readiness, especially for children from disadvantaged backgrounds (Fitzpatrick *et al.*, 2011; Magnuson *et al.*, 2007; and Ludwig and Miller, 2007). This literature provides valuable insights into the impact of childcare on

---

[1] In Australia, for example, ensuring access to early childhood education for all Indigenous children in remote communities was one of the policy targets set by the Council of Australian Governments in 2008. A recent Government report (Department of the Prime Minister and Cabinet, 2015: page 5) mentions that ensuring access to childcare is one of the key goals of the closing the gap agenda.

[2] For a review of this literature see Duncan and Magnuson (2013) and Kalb *et al.* (2014).

cognitive outcomes; however, whether these results can be extrapolated to Indigenous populations is not clear. Evaluations of childcare and preschool programmes are usually based on general population surveys where Indigenous communities are typically underrepresented, compromising the capacity to draw conclusions on the impact of preschool activities in these communities. For example, Bradbury *et al.* (2011) compare inequality in child development at age 4-5 in Australia, Canada, the US and the UK. Although they control for ethnicity, no specific attention is paid to Indigenous populations due to the small numbers observed in national surveys.[3]

Understanding the influence of early childhood education and care activities on Indigenous children is crucially important to prevent the large disparities in educational performance observed between Indigenous and non-Indigenous children later in school (Bradley *et al.,* 2007; Ford, 2013). Unfortunately, however, research on the potential impact of childcare in Indigenous populations is still very limited and mostly based on small samples. The current paper contributes to the existing literature by investigating childcare use and its influence on children's cognitive development in Indigenous communities using data from the first five waves of the Longitudinal Study of Indigenous Children (LSIC). The younger of two cohorts of children followed in LSIC has detailed information about childcare participation from infancy to school entry age, children's cognitive outcomes, and a broad range of family background characteristics and other socioeconomic characteristics. Making use of this unique dataset, the paper contributes to the literature in two ways. First, we investigate the prevalence and determinants of Indigenous children's childcare attendance during the preschool years. The literature puts considerable emphasis on the first three to five years in life (especially in relation to school readiness). Understanding differences in the probability of using childcare across families with different socioeconomic characteristics is therefore important. Second, we estimate the influence of childcare attendance on cognitive outcomes at different stages of childhood using multivariate regression and propensity-score matching techniques. Although we cannot estimate causal relationships with the data we currently have, we control for as much of the endogeneity of childcare use as is feasible by using information on a comprehensive set of potential confounders of preschool attendance and children's outcomes. We use year-on-year data rather than the five waves of data in one panel regression, because the outcome

---

[3] The sample includes 219 Indigenous Australian children and 129 Indigenous Canadian children, who represent about 5 and 2% of the sample, respectively. No observations from Indigenous populations are included in the sample for the U.S.

variable of interest is slightly different every year. Although some researchers have used cognitive skill measures from subsequent waves in a multi-period panel regression (e.g. see Fiorini and Keane, 2014), the different measures in LSIC are not strictly comparable across waves and are likely to measure slightly different skills or abilities each wave.

Our paper builds on the recent work by Arcos Holzinger and Biddle (2015) — AHB (2015) henceforth — who used LSIC data to study the relationship between early childhood education and children's outcomes. However, our paper differs from theirs in the following important aspects. AHB (2015) study the influence of preschool and childcare attendance on children's outcomes by the time they are between five and seven years of age. Given our interest in understanding how skills develop in early years, we expand their analysis by examining the impact of childcare attendance on children's outcomes at different ages between two and five before they enter school. Even if limited impacts at school age are found (as in AHB), there could be a larger effect earlier on, which has diminished over time (or has not been sustained through the formal education system), and it is important to investigate this possibility. As in AHB (2015), we study the relationship between childcare use and outcomes using linear regression techniques. However, differently to AHB (2015) we also look at this relationship using propensity score matching techniques, and we pay particular attention to the definition of variables and the composition of the sample of analysis as explained in the data section. Furthermore, our analysis considers a more comprehensive set of covariates when evaluating the relationship between childcare attendance and cognitive outcomes to ensure we compare children that are as similar as is feasible, except for their childcare use. In particular, following the literature on child development, our models account for differences in home environments and family investments in children which turn out to be important.

We find significant differences in cognitive outcomes between childcare users and those who never participate in childcare during the first five years of life. Children who attend childcare perform better on a range of cognitive scores measured at different ages in the preschool period. Selection into childcare contributes to explain these results. Indeed, our analysis of childcare participation shows that the probability of using childcare is not uniform across different socioeconomic groups. Children from more advantaged families are more likely to attend formal childcare than other groups. Thus, children whose primary carer has a University degree, children with employed carers, and children who are read to regularly at home have a higher probability of ever attending childcare than other children. With the data at hand we find no significant association between childcare use and cognitive outcomes after controlling for a

broad range of individual and household characteristics that are correlated with formal childcare use and child development. We find that for most cognitive measures, the statistical relationship between outcomes and childcare attendance ceases to be significant once we control for differences in family environments. However, the results suggest that relatively disadvantaged children might benefit from attending as indicated by some modestly significant estimated effects for the children currently not attending childcare.

The remainder of the paper is structured as follows. The small literature on childcare use and child development in Indigenous communities is reviewed in Section 2. Section 3 describes the features of the LSIC data including a description of the variables that are constructed and used, and presents a range of summary statistics and descriptive graphs. This is followed by the methodology in Section 4, and results in Section 5. We discuss the results for both childcare use and cognitive outcomes in detail in two separate subsections. Section 6 concludes.

2. Indigenous Childcare Use and Child Development – The Literature

Mounting evidence on the importance of early childhood experiences for individuals' subsequent development and life outcomes has led to a growing interest in understanding the potential impact of childcare participation on developmental outcomes. Thus, in recent years there has been a substantial amount of research aimed at quantifying the influence of childcare and other preschool activities on children's cognitive and non-cognitive outcomes (i.e. the returns to childcare). Most of this research is focused on developed countries, including those of interest to us (the US, Australia and Canada) where the effect of childcare is evaluated using data from general population surveys. Estimates of the returns to childcare for Indigenous children based on data from these surveys should be interpreted with caution for several reasons. First, Indigenous populations in these countries are generally small, and so their numbers in representative surveys tend to be very small. Second, these populations are typically underrepresented in general population surveys, with higher than average non-participation rates. Third, general surveys often fail to recognise the specific circumstances and characteristics of Indigenous communities which are likely to be relevant for the study of childcare use and its impact on the outcomes of children in these communities. Fourth, Perales *et al.* (2015) show that there are higher rates of survey problems (e.g. poor understanding of survey questions, lack of cooperation, having responses influenced by others, and being suspicious of the interviewers) amongst Indigenous people in surveys. This may lead to poorer interview data quality for the Indigenous population in general population surveys. Taken

together, these factors severely hamper researchers in making robust statistical inferences about Indigenous populations based on these surveys.

2.1. *Childcare usage amongst Indigenous children in Australia*

There are no quantitative studies on formal childcare use amongst Indigenous children. Therefore, we examine two papers on preschool use by Indigenous children, noting that preschool is used by children from 3-4 years old and amongst the non-Indigenous population, preschool use is more prevalent than formal childcare use. Biddle (2007) uses data from the 2011 Australian Census of Population and Housing to compare the levels and predictors of preschool use between Indigenous and non-Indigenous populations in Australia. His results show that Indigenous children are less likely to attend preschool than their non-Indigenous counterparts. Equivalised household income and parental education positively influence preschool attendance in both Indigenous and non-Indigenous populations, although their importance is greater amongst Indigenous children. The presence of Indigenous preschool workers increases the probability of preschool use amongst Indigenous children, which may highlight the importance of cultural factors and/or support in enabling the child to access childcare in these populations. Hewitt and Walter (2014) analyse preschool attendance of Australian Indigenous children using data from the first two waves of LSIC, and find that housing stability and the number of books in the household are factors positively associated with the probability of attending preschool.

2.2. *Impacts of childcare usage on Indigenous children's cognitive development*

Research on how childcare participation influences cognitive outcomes in Indigenous populations is scarce, partly due to the limited availability of suitable longitudinal data for these populations. Thus, to date much of the research for Indigenous communities is based on small samples. Gormley *et al.* (2005) analyse the impact of a prekindergarten programme on the outcomes of Native Americans in the US using data from 240 Native American children. Similar to the other populations, Native American children benefitted from the programme in terms of school readiness. Their improvements were smaller than those found for Hispanic children, but bigger than those found for Black and White populations.

Benzies *et al.* (2011, 2014) evaluate the impact of a Canadian preschool programme using data collected before and after implementation for 112 children, including 40 Indigenous children. It was found to have a positive impact on receptive language skills, and for Indigenous children the time in the programme also had a positive impact on the test results obtained immediately

after programme completion. Mughal *et al.* (2015) used follow-up data on the same children to compare subsequent development across Indigenous Canadian children, other Canadian-born children and migrant children. Although the total sample is further reduced to 78 (21 Indigenous children) and the results should be interpreted with caution, the effects of the programme on Indigenous Canadian children were maintained up to age 10.

The work by AHB (2015) constitutes the only attempt to date to investigate the relationship between childcare and preschool attendance, and child development in Indigenous communities in Australia. Using LSIC data, they evaluate the influence of preschool activities on children's cognitive and developmental outcomes two years, and three to five years after participation in early childhood education. Using linear regression models, they find a positive relationship between attendance at preschool and childcare in early years, and children's cognitive and developmental outcomes after school entry,[4] but this disappears once they control for family characteristics.

3. Data

3.1. *The Longitudinal Study of Indigenous Children*

To study patterns of childcare use and their impact on Australian Indigenous children's cognitive outcomes, we use data from the first five waves of the Longitudinal Study of Indigenous Children (LSIC). LSIC is a panel survey which collects annual information on Australian Indigenous children and their families. The survey includes a Baby cohort (B-cohort) comprised mostly of children who were between 6 months and 1.5 years of age at the time of the first interview, and a Child cohort (K-cohort) where the majority of children were between 3.5 and 4.5 years old in Wave 1. This is a unique and unprecedented dataset capturing rich information that is particularly suitable for the analysis of the developmental process of Indigenous children. Indigenous children are the sampling unit in LSIC. Most families in the study were contacted and recruited using administrative information regarding the presence of an Indigenous child in one of the target age groups as provided by Centrelink or Medicare Australia (Department of Families, Housing, Community Services and Indigenous Affairs,

---

[4] Although requirements vary slightly across states, most children in Australia start school at the age of five. From age six, enrolment at school is compulsory across all states.

2011).[5] However, other informal methods of recruitment such as word of mouth, local knowledge and study promotion were also used to supplement the sample. In practice, the Child cohort consists of children born in 2003, 2004 and 2005 and the Baby cohort consists of children born in 2006, 2007 and 2008.

Taking into account the diversity of Indigenous-specific circumstances, LSIC collects data on, for example: children's cognitive measures at different stages, prenatal health, Indigenous culture integration, parental characteristics, family environments and other demographic information. The survey started in 2008 with a sample of 1,671 Indigenous children who have subsequently been interviewed every year. By Wave 5, 1,267 children remained in the study.

Due to the high dispersion and the relatively low number of families with Indigenous children in Australia, LSIC used non-random purposive sampling clustered across 11 geographical sites.[6] These sampling sites were specifically chosen such that the sample, even if not nationally representative, represents the broad distribution of the Indigenous population in Australia.[7]

3.2. Analytical Sample

Given our interest on the influence of childcare on cognitive development in early years, we focus exclusively on children in LSIC's B-cohort. This is because in this group we observe a wider range of measures of cognitive development, and children's exposure to early childcare was observed (for the older K-cohort early exposure to childcare is not captured). We have defined our sample of analysis, based on the intended age range of the B-cohort in the first wave, i.e. between 0.5 and 1.5 years. To ensure that cognitive development is measured for children of comparable ages, children who are older or younger than this age range in wave 1 (n=283; about 29% of the sample) are excluded from the analyses.

The derived sample includes 672 children, 671 from the original B-cohort and one child from the K-cohort whose age at the time of the first interview was between 0.5 and 1.5 years. Besides these 672 children, there are 28 children who joined the survey in Wave 2 and satisfy the age

---

[5] Families who had at least one child who was identified by Centrelink or Medicare Australia as being Aboriginal or a Torres Strait Islander and was born between December 2003 and November 2004 (Child cohort) or between December 2006 and November 2007 (Baby cohort) were in the sampling frame.

[6] We do not have sufficient information in the LSIC data to control for the clustered design of the LSIC study as recommended by Hewitt (2012). The only measure of geographic area available with the general release of LSIC is the Level of Relative Isolation classification. This indicator is based on geocoding of the home address of participants. Five categories of isolation have been defined, ranging from 1= "None" to 5 = "Extreme".

[7] For a detailed discussion of the properties of the LSIC data see Dodson *et al.* (2012).

criterion to be included in our sample.[8] In subsequent waves, the age range of those in the derived Baby cohort broadens again due to different timings of the survey in these later waves, but it never extends beyond two years. The age range for children participating in the child development measurements in waves 3 to 5 is two years as shown in Appendix I Table A.1. There are 24 children in the derived sample who were enrolled in Year 1 at school in wave 5 and one child was enrolled in wave 4. These children are excluded from the analyses of childcare attendance and its impact on children's outcomes. Furthermore, given the importance of age in the developmental process, all models that estimate the influence of childcare on cognitive development control for age differences across children.

This paper pays particular attention to the composition of the sample of analysis to ensure the different analyses are comparable (i.e. the same children are included in each model variation). That is, when comparing the models only including childcare use as an explanatory variable with the models including additional sets of explanatory variables, we ensure that these models are based on the exact same samples so that any differences are driven by the additional variables rather than different children being included in our sample.

3.3. Outcome measures

Formal childcare use

We have classified formal care to include any form of day care centre care, family day care, care by a nanny[9] or other formal care; while informal care includes care by brother/sister, grandparents, other relative, friend/neighbour, or another person. Table 1 presents the proportion of children who use formal childcare in each wave. To put these proportions in context we also present the proportion of children who use informal childcare in each wave separately. This shows that a relatively small proportion of families use formal childcare for their children (less than half of all children have ever used formal childcare by around age 5), while the use of informal childcare is widespread (over 95% of all children have ever used informal childcare by around age 5). Amongst formal care, day care centre use is most common with 35.4% of all children having attended a day care centre by around age 5.

---

[8] For those who entered LSIC in Wave 2, we used the intended age range plus one year so we can categorise them according to their age in Wave 2.

[9] Only one family in our sample of analysis used a nanny to care for their child.

Table 1 Formal and informal childcare, and preschool use by wave

Wave [Age in years]	Wave 1 [0.5-1.5]	Wave 2 [1.5-2.5]	Wave 3 [2.5-3.5]	Wave 4 [3.5-4.5]	Wave 5 [4.5-5.5]	Ever
Formal childcare (%)	19.2	26.7	32.4	24.0	14.8	41.6
Day care centre (%)	13.4	20.0	27.1	19.3	9.5	35.4
Family day care (%)	2.8	3.4	2.7	2.3	1.4	5.9
Other formal care (%)	3.3	3.6	3.1	2.3	2.2	8.7
Preschool (%)	0.0	0.0	8.4	49.7	47.4	56.7
Informal childcare (%)	72.6	74.6	75.9	79.2	76.7	95.1
Mixed formal/informal (%)	14.6	19.7	24.1	19.3	11.3	35.6
Ever used formal care (%)	19.2	31.6	40.1	42.7	46.6	41.6
Average number of waves in formal care	0.19	0.45	0.77	0.99	1.16	0.98
Total (N)	672	614	584	529	497	700

Source: Authors' calculations based on LSIC.

Observing this high usage of informal childcare, it is important to note that traditional Indigenous contexts feature a more communal approach to the care, mentoring and education of children. There are substantial cultural differences between Indigenous and non-Indigenous Australians concerning child rearing (e.g. see Lohoar *et al.*, 2014). For many Indigenous people –especially those living in Indigenous communities– caring for and educating young children is not a job that is done (or expected to be done) solely by parents, or formal institutions such as long day care centres. Instead, informal childcare, child mentoring and support to parents are routinely provided by other family members (chiefly grandparents, but also aunts, uncles, cousins and older siblings), neighbours, and the community as a whole. This distinctive Indigenous approach to informal childcare might affect Indigenous parents' decisions about participation in formal childcare, and about the relative outcomes of childcare usage on Indigenous children. In particular, it could be argued that Indigenous communities may provide better alternatives to formal childcare, in the form of collective arrangements and informal childcare practices.

Formal childcare use starts at a low level and reaches a peak in Wave 3, when children were between 2.5 and 3.5 years old. As expected, the rate drops off when children start to attend school (typically from age 5). Table 2 presents the percentage of children who use formal and informal childcare and preschool by their age.[10] The results reported in the table show that formal childcare is mostly used when the children are aged between 2 and 4 years. At age 3-4,

---

[10] Note that the number of children in the age range 1 to 2 exceeds the actual sample size because some children interviewed in wave 1 were also in that age group when they were interviewed in wave 2.

preschool starts to become important, and at age 4-5 preschool use reaches its peak with nearly 60% of all children attending. In our analyses, we control for preschool attendance by including a dummy variable for preschool attendance as an explanatory variable.

Table 2 Formal and informal childcare and preschool use by age of study child

Age range	0-1	1-2	2-3	3-4	4-5	5-6
Formal childcare (%)	15.1	23.5	29.7	29.8	20.4	9.8
Informal childcare (%)	69.3	74.6	74.7	81.1	76.9	77.6
Preschool (%)	0	0	2.2	26.5	59.2	31.8
Mixed formal/informal (%)	10.6	18.2	22.6	22.7	15.2	7.5
Mixed formal/preschool (%)	0	0	0.2	6.5	13.7	5.6
Total (N)	218	765	585	554	540	214

Note: The sample size for preschool information is slightly smaller than for the childcare information: 552 for age 3-4 and 539 for age 4-5.

Source: Authors' calculations based on LSIC.

For our analyses we also construct a variable that measures the cumulative extent to which formal childcare is used by counting the number of waves in which formal childcare attendance was reported. A relative version of this measure is also used, which divides the number of waves in which childcare was used by the number of waves in which the child was observed.[11]

Children's cognitive development
To investigate the influence of childcare attendance on children's cognitive development, we consider three measures at different stages of childhood. These include:

- the *MacArthur Bates Communicative Development Index* in Wave 3 (2.5-3.5 years)[12], which is parent-reported and measures expressive vocabulary and early grammar skills[13];

---

[11] An alternative relative version assumes no formal childcare was attended in the non-responding waves. The two relative variables are very similar (see Appendix I Tables A.2a and A.2b). Comparing the two alternative definitions in Wave 5 shows that 88.5% of all cases would have the same value for both variables. In the cases where they are different, they are usually the next-closest possibility. In our analysis we use the version based on observed waves only.

[12] The indicated age ranges are the intended ages of children in each wave. Appendix I Table A.1 shows the actual (slightly wider) age distribution of the children when doing these tests.

[13] Farrant *et al.* (2014) express some concern regarding whether the MacArthur Bates measure was culturally appropriate enough in terms of the words assessed and/or sensitive enough as a measure of Indigenous vocabulary development.

- the *Who Am I* measure in Wave 5 (4.5-5.5 years), a developmental test of school readiness (which is not language dependent and suitable for children with limited English) that looks at children's ability to write their name, letters and copy shapes; and

- the *Renfrew Word Finding* vocabulary test in Wave 4 (3.5-4.5 years) and Wave 5 (4.5-5.5 years), which tests children's expressive vocabulary by examining their capacity to name pictures of objects arranged in order of difficulty. Children can name pictures in their preferred language.

Buckley *et al.* (2009) report on the selection of the two last measures for the LSIC data collection. The Steering Committee which oversees LSIC has mandated that the design and conduct of the survey must have the acceptance and support of Aboriginal and Torres Strait Islander communities and of participating families. The *Who Am I* measure was tested in 2007 for its usefulness in a population of Aboriginal and Torres Strait Islander children. It was found to be satisfactory, but a few amendments were made. The children in the LSIC sample scored lower on average than the norm group of comparable age, but the LSIC children's pattern of development was similar to that of the norm groups (Buckley *et al.*, 2009).

Table 3 presents the mean scores on these tests by childcare use. In the statistical analyses, all scores are standardised to have a mean value of zero and a standard deviation equal to one. Inspection of average scores reveal significant differences in cognitive development, with children who ever attended formal childcare outperforming children who never attended formal childcare. Figures 1 to 3 display the differences in performance on each of these measures by children who did or did not attend formal childcare.

Table 3 Mean scores on cognitive tests by formal childcare use and type of test

| Cognitive test | Target Age | Score range | Mean score | | Difference (%) | Number of observations (never/ever) |
			Never in formal childcare	Ever in formal childcare		
MacArthur Bates scale	2.5 - 3.5	0-100	59.1	66.0	11.7	344/226
Renfrew Word test	3.5 - 4.5	0-46	16.0	21.5	34.4	278/213
Renfrew Word test	4.5 - 5.5	3-47	22.0	26.7	21.4	253/220
Who Am I?	4.5 -5.5	0-42	22.0	23.6	7.3	231/210

Note: All differences are significant at the 1%-level at least.
Source: Authors' calculations based on LSIC.

Figure 1 shows the unconditional distribution of the MacArthur Bates scores in Wave 3 by the number of waves in which childcare was attended (up to Wave 3). As evidenced by the graph, differences in cognitive scores between these two groups are not restricted to the mean. Indeed, the distribution of scores amongst those who attended childcare in more waves has shifted to the right. That is, among those children who attended childcare in more waves, a higher proportion have a high score and a lower proportion have a low score in the cognitive development measures, compared to those children who attended childcare in fewer waves (or not at all). As Figures 2 and 3 show, similar differences in the distributions were found for the Renfrew test administered to children at an older age (Waves 4 and 5) and also for the Who Am I measure observed in Wave 5. That is, the distributions of scores amongst children who attended childcare in more waves are located more to the right than those of the group who attended childcare in fewer waves. These differences in the distributions suggest that either formal childcare attendance improves the performance of children on each of these scores, or it reflects selection into childcare and differences in the characteristics and circumstances of children who attended childcare and those who did not, which are correlated with the scores.

Figure 1: MacArthur Bates level 3 by number of waves in formal childcare (Wave 3)
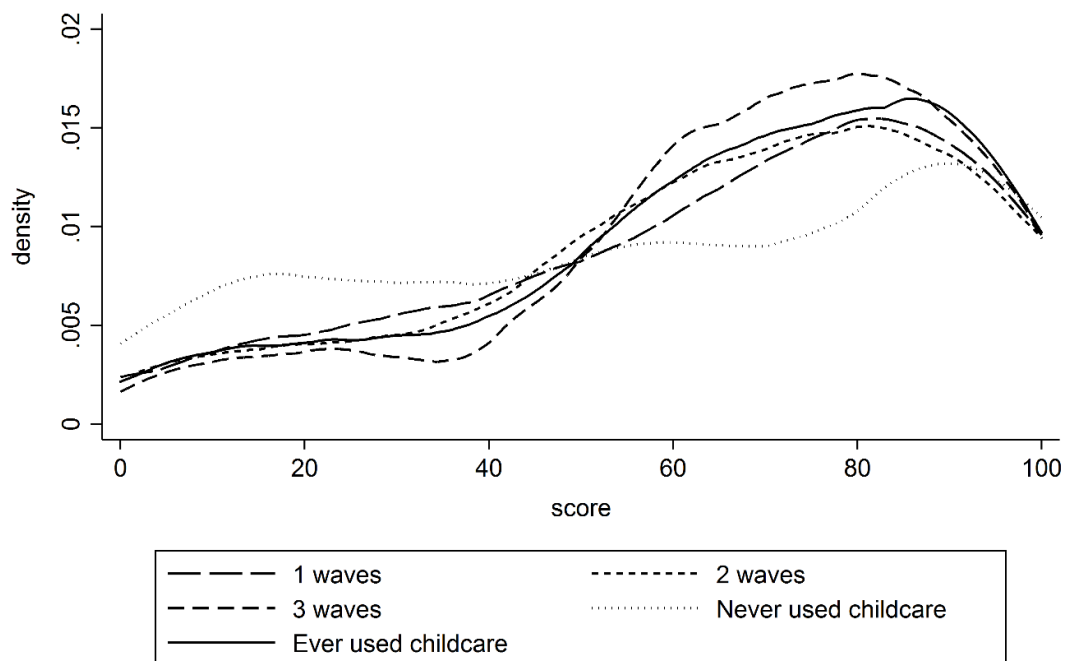
Overall, the graphs suggest that formal childcare attendance is positively associated with children's cognitive outcomes. However, these are unconditional estimates. In the sections that follow, we control for factors that may confound the association between childcare use and

children's cognitive outcomes. The control variables capturing these factors are discussed in the next subsection.

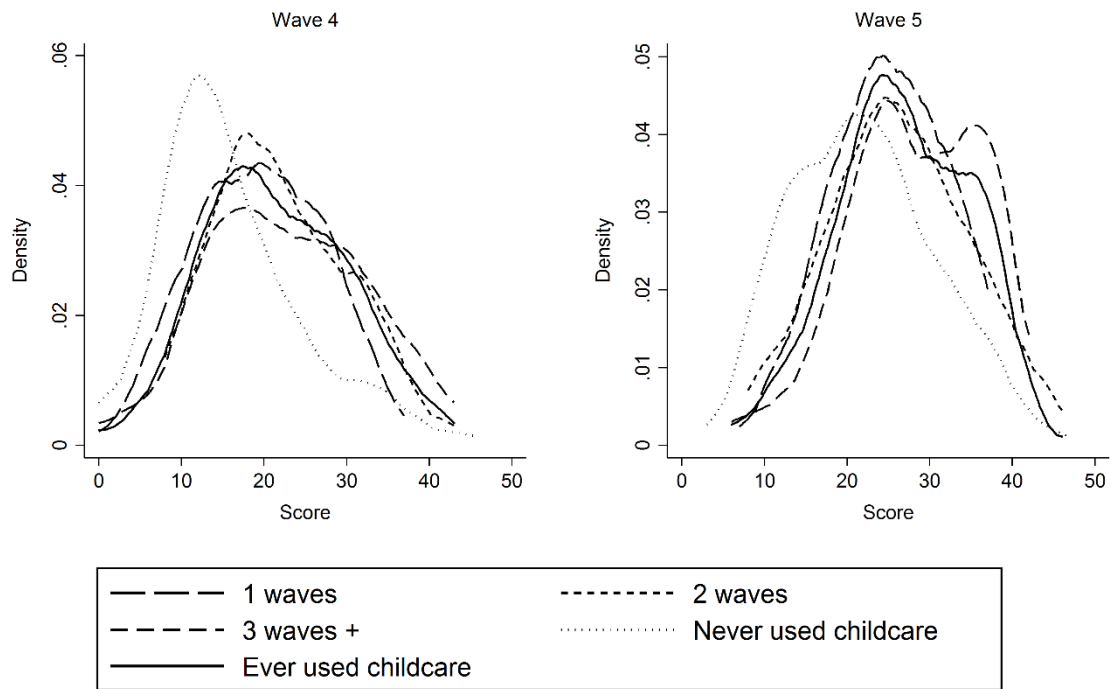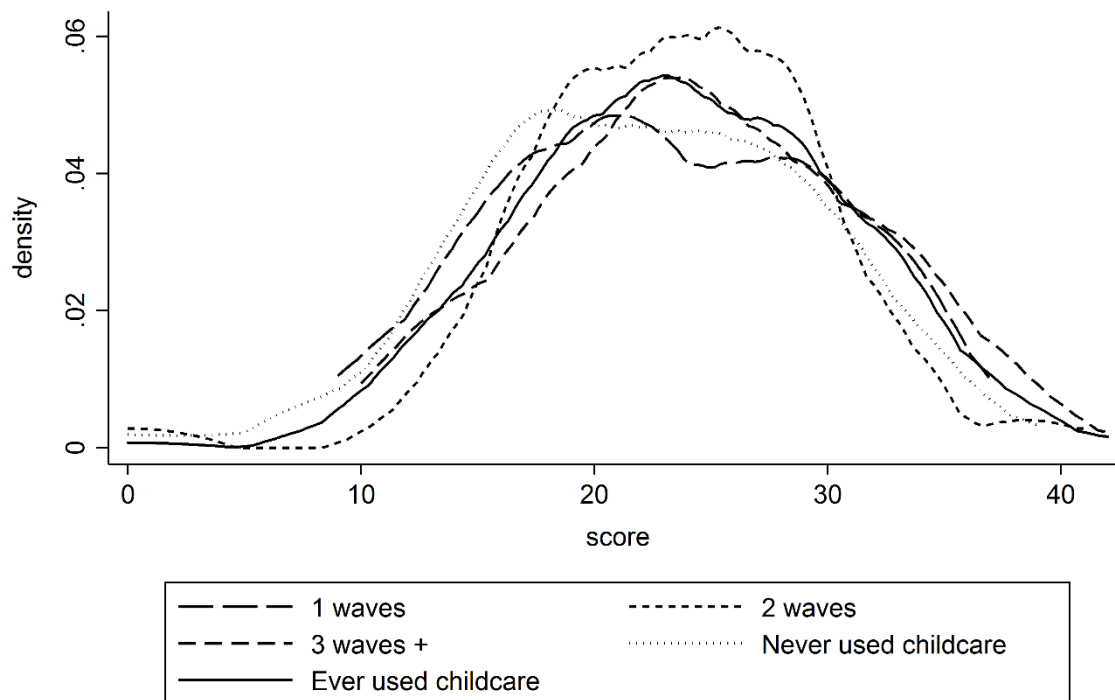**Figure 2: Renfrew test by number of waves in formal childcare (waves 4 and 5)**

Figure 3: Who Am I measure by number of waves in formal childcare (wave 5)

3.4. Explanatory variables

In our multivariate analyses we include a set of variables expected to affect child development, as well as a set of variables expected to be correlated with childcare use. These include child characteristics, such as age and gender; parental characteristics, such as education and employment; family environment characteristics, such as number of siblings and household income; and community characteristics, such as socio-economic ranking and geographic remoteness (as measured by the Level of Relative Isolation variable).

Our empirical analysis exploits the information available in LSIC to the maximum extent possible, paying particular attention to the robustness of our findings to the definitions of some variables that were not collected consistently across waves. This includes parental education which was collected in different ways across waves 2 to 5. In waves 2 to 4, there was one question on the highest level of education achieved which combined years of school completed and post-school/tertiary qualifications. In Wave 5, years of school completed was asked separately from post-school qualifications. Comparing the answers in waves 2 to 4 to the answers in Wave 5, it appears that several parents did not report post-school qualifications in waves 2 to 4, only highest year of school completed. We therefore used the information in Wave 5 to update the information in waves 2 to 4. However, in some cases the information in waves 2 to 4 could not be reconciled with the information in Wave 5 (e.g. education in Wave 5 was lower than in one of the earlier waves). In these cases, we create two alternative versions of the education variable: one representing a pessimistic view on the education level, reporting the lowest level observed, and the other representing an optimistic view on the education level, reporting the highest level observed. As a result, we have a lower bound on the education level (the variable *imputed low values education*) and an upper bound on the education level (the variable *imputed high values education*) for waves 2 to 4. Both variables are reported in Table 4 with the summary statistics. These two versions can be used to investigate the sensitivity of the results to the pessimistic and optimistic assumption. Given the likely relevance of parental education in childcare use and child development this is important. However, regressions using the higher bound turn out to give similar results to those using the lower bounds on education so only the results based on the lower bound are reported.

Information on income is available as a categorical variable in all waves except for Wave 3, where the income question was not asked. Compared to earlier waves, the categorisation is more detailed in Wave 5. We aggregate income groups in Wave 5 to match the categorisation

used in the earlier waves. In the analyses we use the midpoints of the income category, treating income as a continuous variable to limit the number of parameters to be estimated.

We include a broad range of family and individual characteristics based on the findings of previous research. Using Programme for International Student Assessment (PISA) data, De Bortoli and Thomson (2010) demonstrated the importance of home resources (including number of books at home) for children's reading performance and performance in science, and specifically for children with an Indigenous background where higher correlations are found than for other Australian children. For this reason, the number of books at home, reading and storytelling to children are included in the analyses, as well as other family activities.

Biddle (2007) examined preschool use by Indigenous children and found that low income, low parental education levels, and the remoteness of the family's residence explain a large component of the lower participation by Indigenous children relative to their non-Indigenous counterparts. The presence of an Indigenous preschool worker in the local area increased participation, but relatively few households benefited from this. A small qualitative study by Trudgett and Grace (2011) provided a number of potential reasons for the lower attendance of Indigenous children at early childhood education institutions while emphasising that Indigenous families are not a homogenous group. However, one key factor they identified is trust and a good carer-child connection which for some families is connected to indigeneity.

We do not observe the characteristics of local childcare workers. Instead, we control for the degree to which the child's primary carer at home identifies with Indigenous culture and participates in Indigenous-specific activities. The assumption is that the presence of Indigenous childcare workers may be more important to parents/carers identifying strongly with Indigenous culture, thus potentially discouraging childcare use.

LSIC includes multiple indicators about the financial stress, family environments, Indigenous cultural practices, and life events that children were exposed to during childhood. We use data from these indicators to derive a composite index for each of these aspects of children's lives to provide a more reliable and succinct measure of these constructs to use in regressions, at the same time avoiding potential multicollinearity of the separate indicators. Results from exploratory factor analysis for the financial stress, family activities and Indigenous cultural practices variables suggest that one underlying factor for each of these sets of variables exists. It is not clear, however, that such an underlying factor exists for the life events variables. A brief discussion of the derivation of the composite measures is provided in Appendix II.

A number of relevant control variables that are collected in other survey waves are not collected in Wave 3 of the study. These include income (as mentioned earlier), the number of books at home and several of the family activities. Therefore, fewer variables are included in the analyses of Wave 3.

3.5. Summary Statistics

Table 4 shows summary statistics for the variables that we use in our analyses for each wave, based on the samples used in the childcare use regressions. These are organised in the following categories: childcare use; cognitive development measures; child characteristics; primary carer characteristics; family environment; Indigenous culture and community; and birth and early childhood characteristics.

Table 4 Summary statistics for each wave (percentages unless otherwise indicated)

Variable	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
Ever in formal childcare	19.6	30.7	41.3	44.3	46.3
Cognitive development measures					
MacArthur Bates score[a]			62.0		
Renfrew test score[b]				18.5	24.6
Who Am I test score[c]					22.8
Study child (SC) characteristics					
SC: Female	51.0	49.4	49.6	47.9	49.3
SC: Oldest child	1.5	4.1	12.2	16.6	19.5
SC: Youngest child	66.2	60.8	50.6	44.7	33.6
SC: Age in months[*]	13.2	23.1	35.4	47.2	58.8
SC: Speaks Indigenous language	16.3		19.8	19.8	19.5
Prep Year 1 student	0.0	0.0	0.0	2.0	34.3
Primary carer (PC) characteristics					
PC: non-Indigenous	17.7	17.2	18.8	19.6	21.0
PC: Partnered	41.8	39.8	40.4	38.0	34.0
PC: Employed	27.2	27.6	29.8	30.1	32.8
PC: Number of hours worked[*]	6.7	7.6	8.5	8.4	9.4
PC: Highest educational qualification, low imputation:					
Year 9 or less	13.5	13.1	13.2	13.0	7.8
Year 10/11	30.3	32.2	29.8	30.2	25.5
Year 12	10.4	10.2	9.9	10.3	12.0
Certificate (other)	17.6	17.4	16.5	16.0	19.0
Advanced Diploma or Certif. III/IV	21.6	20.5	23.6	23.8	25.6
University degree	6.7	6.6	7.0	7.0	10.0

Variable	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
PC: Highest educational qualification, high imputation[d]:					
Year 9 or less	9.8	9.3	9.5	9.9	8.9
Year 10/11	30.1	31.3	29.1	28.5	25.5
Year 12	11.8	12.4	12.0	11.7	11.8
Certificate (other)	16.5	16.2	15.5	15.1	18.5
Advanced Diploma or Certif. III/IV	23.5	22.4	24.2	25.2	25.7
University degree	8.3	8.3	9.5	9.2	9.8
PC: Speaks Indigenous language	22.6	25.1	23.0	21.6	23.0
Family environment					
Number of other siblings[*]	1.5	1.6	1.8	1.9	1.8
Household size[*]	5.1	5.2	5.2	5.3	5.3
Grandparents present	6.3	7.5	8.7	12.4	16.0
Weekly family income ($)	490.4	527.3		620.7	669.5
Missing income dummy	4.4	7.3		6.3	5.0
Income support main source of income	68.2	73.9	78.5	85.6	81.3
Ever homeless	7.9	8.3	8.7	8.5	9.3
Housing tenure:					
Home being paid off	17.7	17.0	19.2	20.0	21.8
Private rental	23.8	21.6	23.0	24.5	26.3
Public housing	40.1	38.4	35.7	33.5	29.8
Social housing	16.5	21.0	17.8	14.8	12.5
Other housing tenure	1.8	2.1	4.3	7.2	9.8
Moved house between waves	36.2	24.7	27.1	26.1	26.5
Number of homes since SC's birth[*]	1.4	1.4			
Home needs major repairs	38.4	38.0			
Number of people per bedroom[*]	1.6	1.7	1.6	1.7	1.7
Number of books at home: 0-10	37.2	35.5		25.0	21.3
11-20 books	13.5	9.8		12.1	13.0
31-30 books	8.9	17.6		9.9	9.8
30+ books	40.5	37.1		53.0	56.0
Number of waves SC was read books[*]	0.7	1.5	2.0	2.7	3.3
Number of waves SC was told stories[*]	0.6	1.3	2.4	3.2	3.9
Family activities index[e]	0.003	-0.029		0.032	0.03
Index of life events[e]	0.017	-0.002	-0.012	0.002	0.009
Number of major life events	4.0	3.8	4.0	4.1	4.0
Financial stress index[e]			0.019	0.022	0.013
Indigenous culture, and community					
Participation in Indigenous culture index[e]	-0.036		-0.065	-0.031	
Community bad for kids (mean value: 1=very good to 5=really bad)	2.2	2.2	2.1	2.0	1.9
No places to play in community (mean value: 1=lots of parks/playgrounds to 5=no, none)	2.5	2.5	2.4	2.4	2.3

Variable	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
Dangerous community (mean value: 1= very safe to 5=dangerous)	2.4	2.4	2.4	2.3	2.2
Level or relative isolation:					
None	32.5	31.1	32.5	34.0	33.3
Low	48.6	47.1	50.1	48.5	49.4
Moderate	12.0	14.7	12.0	12.1	11.5
Extreme	6.8	7.1	5.6	5.4	5.8
Index of Relative Indigenous and Socioeconomic Outcomes (IRISEO):					
Quintile 1	8.1	13.1	11.8	11.2	9.8
Quintile 2	17.6	17.8	18.4	18.2	18.5
Quintile 3	41.2	36.5	38.2	37.1	35.0
Quintile 4	15.3	13.9	14.9	15.7	17.3
Quintile 5	17.7	18.7	16.7	17.8	19.4
Sample size for the above variables (N)	541	482	484	445	400

Birth and early years characteristics[f]

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
Birth timing					
On time	29.5	31.7	29.5	29.6	30.0
Late	25.2	24.9	25.3	25.1	23.0
Early	33.6	33.0	33.6	34.6	36.8
Very early	11.7	10.5	11.6	10.7	10.2
Sample size (N)	515	458	455	419	383
Birth weight					
Less than 2.5kg	6.7	5.8	6.8	6.1	6.3
2.5-4.5kg	74.7	73.7	72.9	74.6	76.0
More than 4.5kg	18.7	20.5	20.3	19.3	17.8
Sample size (N)	541	482	484	455	400
Smoked whilst pregnant	49.1	49.4	47.5	45.6	47.6
Sample size (N)	501	443	440	406	368
Drank alcohol whilst pregnant	21.6	22.4	21.9	21.0	21.5
Sample size (N)	500	442	439	405	367
Ever breast fed	80.3	81.0	79.8	79.5	81.7
Sample size (N)	538	480	481	444	399
Child had sleeping problems	30.9	27.5	27.4	28.3	19.3
Sample size (N)	534	476	483	445	400

Notes: a) sample size is 470, b) sample size is 415 and 385 respectively, c) sample size is 360; d) in Wave 5 the low and high imputation variables for education are the same; e) a higher value on these indexes indicate more family activities, more adverse life events, more financial stress, and more participation in indigenous cultural events; f) including birth and early years characteristics would reduce the sample size for regressions considerably, so these are not included in the main regressions; the relevant sample size combining the variables in the other groups with each of the variables in this group separately. *numbers are not percentages but levels.
Source: Authors' calculations based on LSIC.

Many variables are quite similar across the waves with the obvious exceptions of child's age, being the oldest or youngest child (as children move out of home, and new children are born), employment and hours worked (both increasing with the child's age), number of books at home (also increasing with child's age), cumulative number of waves of book reading and storytelling, presence of grandparents in the child's home (possibly increasing with the grandparents' age) and income. The proportion of families whose main income source is a Government pension, benefit or allowance is also increasing, which is somewhat surprising given the observed increases in employment and hours worked.

3.6. Attrition

Attrition is higher for LSIC data than for comparable general population surveys, such as the Longitudinal Study of Australian Children (LSAC). Starting from the full sample size as presented in Table 5, just over 56% of children are observed in all five waves, while Table 6 shows that attrition rates between consecutive waves are over 10% per wave.

Table 5 Sample sizes and number of waves in the survey by time of entry

Number of waves	Started in W1	%	Started in W2	%	All	%
Total number (Starting sample)	672	100.0	28	100.0	700	100.0
1	34	5.1	2	7.1	36	5.1
2	43	6.3	2	7.1	45	6.3
3	65	9.7	9	32.1	74	10.8
4	137	20.4	15	53.7	152	21.7
5	393	58.5			393	56.1

Table 6 Attrition in the LSIC sample of analysis

	Attrited observations	Rate (%)
Wave 1-Wave 2	86 (out of 672)	12.8
Wave 2-Wave 3	68 (out of 613)	11.1
Wave 3-Wave 4	94 (out of 584)	16.1
Wave 4-Wave 5	75 (out of 530)	14.2

We checked whether the probability of dropping out of the sample between two consecutive waves is systematically related to children's socioeconomic (and other) characteristics using a probit model. We ran two sets of regressions: one for the probability of attrition between consecutive waves, and another for the probability of staying in the panel for the whole sample period. For these regressions we considered a comprehensive set of covariates capturing information on the characteristics of the child (age, gender, weight at birth, timing of birth,

number of siblings, order of birth, health), the characteristics of the family (income, housing tenure, parent's health and employment, family type, number of books) and the use of formal and informal childcare. Full results are presented in Appendix I Table A.3.

The probability of dropping out of the sample between two consecutive waves does not appear to be systematically related to most of the covariates included in the model. However, we find that the study child's age and the family's type of housing tenure predicts the probability of attrition between waves and also the probability of being in all waves of the panel. In particular, families with older children are more likely to drop out of the sample in most waves and therefore also less likely to remain in the sample for all waves than families with younger children. Families who own their main residence and have outstanding mortgage debt are more likely to stay in the panel in all waves. Caution is needed in interpreting this result as the statistical relationship between housing tenure and attrition could be driven by other unobserved variables not included in the model. Children whose principal carer has a higher education level also seem more likely to participate in all waves. We find no significant relationship between the probability of attrition and childcare use. We assessed the robustness of our empirical findings by applying re-weighting techniques to control for non-random attrition; the main conclusions remained unaltered.

4. Methodology

4.1. Predicting childcare attendance

In this paper we analyse childcare use decisions and their association with cognitive development for a sample of Indigenous Australian children using parametric techniques. Attendance at formal childcare facilities is examined using a probit model that expresses the probability of ever attending childcare during early childhood as a function of child and family characteristics as follows:

$$P(D_t = d_t | x_t) = \Lambda(x_t'\beta) \tag{1}$$

where D_t is an indicator variable that takes the value one if the child has ever attended childcare by time t, x_t is a vector of explanatory variables, and Λ is the normal distribution function. We estimate the model at different stages during childhood using LSIC data (waves 2 to 5). At each stage, the vector of covariates x_t includes a set of time-variant and time-invariant child and family-environment characteristics. These include the variables described in Sections 3.4 and 3.5, grouped into six categories: birth and early childhood conditions; child

characteristics; parent characteristics; family/home environment; Indigenous culture, and community; and financial and life events experienced by families.

4.2. *The impact of childcare attendance on cognitive outcomes*

We investigate the strength of the association between childcare attendance and children's cognitive outcomes using regression models and propensity score matching techniques.

4.2.1 Regression analysis

We propose to estimate a linear model of the form:

$$Y_t = Childcare_t \beta + x_t'\delta + \varepsilon_t, \tag{2}$$

where Y_t is a given cognitive outcome measure, $childcare_t$ is a variable that summarises formal childcare use up to time t, x_t is the same vector of explanatory variables as in Equation (1), and ε_t is a residual term assumed to be identically, independently and normally distributed. Two childcare use variables are considered. First, we use a prevalence measure containing information on whether or not the child has ever attended formal childcare by time t. Although informative about childcare use, this variable does not take into account the intensity of the exposure, as it provides no information about the length or persistence of childcare attendance. To overcome this limitation to some extent we also estimate equation (2) using a count variable capturing the number of waves in which the child attended formal childcare by time t.[14] As we show in the results section, the main conclusions regarding the impact of formal childcare on children's cognitive outcomes are similar using either construct.

4.2.2 Propensity score matching

We also evaluate the impact of childcare on children's cognitive outcomes using propensity score matching techniques.[15] Under certain assumptions, these techniques allow the identification and estimation of the average treatment effects on treated (ATET) and untreated (ATEU) individuals, defined as:

---

[14] Using more detailed information by also measuring the intensity of childcare use through usual weekly hours of childcare would give a sense of accuracy that is unwarranted as childcare usage may have varied over the year, and the hours of childcare are only available for one week. However, the results when using childcare counts weighted by usual hours of care to estimate the outcome regressions are similar to the results from the specifications used in this paper.

[15] For an overview of propensity-score matching and its application to estimate treatment effects see Caliendo and Kopeinig (2008).

$$ATET = E(Y_{D=1} - Y_{D=0} \mid D = 1) = E(Y_{D=1} \mid D = 1) - E(Y_{D=0} \mid D = 1) \tag{3}$$

$$ATEU = E(Y_{D=1} - Y_{D=0} \mid D = 0) = E(Y_{D=1} \mid D = 0) - E(Y_{D=0} \mid D = 0), \tag{4}$$

where D is the treatment variable that takes the value 1 if the child ever attended childcare and the value 0 otherwise, and $Y_{D=1}$ and $Y_{D=0}$ are the potential cognitive outcomes when the child ever attends childcare and never attends childcare, respectively. The main problem in estimating the ATET and the ATEU is that $E(Y_{D=0} \mid D = 1)$ and $E(Y_{D=1} \mid D = 0)$ cannot be directly estimated from the data, as we do not have information to estimate $Y_{D=0}$ for those who ever attended, or $Y_{D=1}$ for those who never attended childcare. However, under the *conditional independence assumption* (CIA) these two expectations can be identified from observational data. The CIA assumption requires orthogonality between the potential outcomes and the treatment, conditional on a propensity score (or the probability of treatment) $p(x) = P(D=1 \mid x)$. As defined by equation (1), the propensity score reflects the probability of treatment assignment given a set of observable characteristics x. Under the CIA, the ATET and the ATEU can be identified by matching children from the treated and untreated groups. In particular, the ATET is estimated by matching every child j in the treated group to a group $C(j)$ of comparable children from the non-treated group with a similar $p(x)$ value. A weighted average of the outcomes of children in $C(j)$ is used to derive a counterfactual of $Y_{D=0}$ for j. Similarly, the ATEU is estimated by matching every child in the non-treated group to a group of comparable children (i.e. with similar $p(x)$ values) who attended childcare and whose outcomes can be used to estimate the potential outcome $Y_{D=1}$ of those who never attended childcare.[16]

In addition to the binary treatment case, we also derive matching estimates of the treatment effects using a multiple treatment framework that allows us to control for the intensity of childcare use to some extent. For this we consider a treatment variable D that reflects the number of waves in which children attended childcare. In principle, this variable ranges between 0 and 5, however, because of the small number of children who went to childcare in

---

[16] We use the *Stata* command *psmatch2* to match children from the treated and non-treated groups. In particular, we use a kernel matching algorithm with a bandwidth of 0.06. When computing the potential outcomes of child *j*, this technique assigns more weight to those individuals in the control group whose scores are closer to that of child *j*. We checked the sensitivity of the results using 0.1, 0.2 and 0.3 bandwidths and other matching techniques including the radius, the k-nearest neighbours, and local linear regressions and the results are robust to these changes.

more than 3 waves, we consider the set of treatments $D=\{0,1,2,3\}$, where category 3 pools children observed in childcare in at least three waves. We denote the potential outcomes associated to each treatment category by $\{Y_{D=0}, Y_{D=1}, Y_{D=2}$ and $Y_{D=3}\}$. Treatment evaluations in this framework are based on the pairwise comparison of treatments. Given our interest in the impact of using childcare, we take $D=0$ as the treatment of reference for the evaluation of the other treatments. As shown by Imbens (2000) and Lechner (2002), these effects can be identified assuming a generalised version of the CIA which requires the orthogonality between potential outcomes and the treatment conditional on the propensity score that reflects the probability of treatment assignment: $p_d(x) = P(D=d \mid D=\{0,d\}, x)$. Following Lechner (2002), the average treatment effects of treatment d can then be expressed as

$$ATET(D=d) = E(Y_{D=d} \mid D=d) - E[E(Y_{D=0} \mid p_d(x), D=0) \mid D=d] \qquad (5)$$

$$ATEU(D=d) = E[E(Y_{D=d} \mid p_d(x), D=d) \mid D=0] - E(Y_{D=0} \mid D=0), \qquad (6)$$

where the counterfactuals $E(Y_{D=0} \mid p_d(x), D=0)$ and $E(Y_{D=d} \mid p_d(x), D=d)$ are estimated matching subjects in treatments d and 0 according to $p_d(x)$, the probability of being assigned to treatment d conditional on being in either d or 0. This selection probability is estimated with a logit model using the restricted sample with observations from treatments d and 0.[17]

A key element of any matching analysis is the evaluation of the matching procedure's success in balancing the distribution of the relevant covariates in the treatment and matched control groups. We assess the quality of the matching between childcare users and non-users using two indicators of covariate balancing widely used in the literature. We report the median standardised bias of covariates before and after matching computed over the set of covariates included in the propensity score. Following Rosenbaum and Rubin (1985), the standardised bias for any covariate X before and after the matching is given by:

$$Bias_{before} = 100 \frac{\overline{X}_T - \overline{X}_U}{\sqrt{\dfrac{V(X_T) + V(X_U)}{2}}}$$

$$Bias_{after} = 100 \frac{\overline{X}_{MT} - \overline{X}_{MU}}{\sqrt{\dfrac{V(X_{MT}) + V(X_{MU})}{2}}} \, ,$$

where $\overline{X}_T$ and $\overline{X}_U$ are the sample means of the full treated and untreated subsamples, and $V(X_T)$ and $V(X_U)$ their respective variances; $\overline{X}_{MT}$ and $\overline{X}_{MU}$ are the sample means of the matched treated and untreated samples, and $V(X_{MT})$ and $V(X_{MU})$ are the corresponding variances. As discussed in Caliendo and Kopeinig (2008), there is no clear criterion to determine what constitutes a good matching. In practice, however, many evaluations consider a standardised bias below 5% as sufficient. Following Sianesi (2004) we also report the p values of the likelihood ratio test on the joint significance of the regressors in a logit model of the treatment before and after the matching. The rationale for this test is that the set of covariates should explain little of the variability in the probability of treatment within the matched sample if the CIA holds. Thus, if the matching works well (and is necessary), the hypothesis of joint significance should not be rejected before matching but should be rejected after matching. Finally, the quality of the matching was also evaluated using the R and B statistics proposed by Rubin (2001). These measures use information on the distribution of the propensity scores in the control and treatment groups after the matching to evaluate its quality.

5. Results

5.1. Childcare use

We estimated probit regression models of the probability of attending formal childcare.[18] In particular, we modelled the probability of using formal childcare in Wave 1 and the probability of having ever attended formal childcare in waves 2 to 5. Wave-on-wave results, expressed as marginal effects, are reported in Table 7. We only show the marginal effects for those characteristics that significantly influence the outcome variable in at least one wave.

We considered six sets of explanatory variables: *birth and early childhood characteristics*; *child's characteristics*; *characteristics of the primary carer*; *characteristics of the family/home*

---

[18] Results for ever attending a day care centre were very similar to the results presented here. These results are available online (Table OA1). Other types of childcare were used by too few children to analyse separately. The data do not allow us to look at other characteristics of the childcare attended either, since these are only available through the childcarer/teacher surveys, usage of which would have considerably reduced the sample further.

environment; *Indigenous culture, and community*; and *financial situation and life events*. A detailed list of the variables included in each block is reported in Table 4.

Table 7 Probit analyses of ever using formal childcare up to the relevant wave (marginal effects)

Wave [Target age in years]	Wave 1 [0.5-1.5]	Wave 2[a] [1.5-2.5]	Wave 3 [2.5-3.5]	Wave 4 [3.5-4.5]	Wave 5 [4.5-5.5]
Study child (SC) characteristics					
Female	0.019	-0.0621[+]	-0.00522	-0.0707[+]	-0.045
	[0.0304]	[0.0361]	[0.0384]	[0.0406]	[0.0441]
Oldest child	0.0345	-0.0989	0.0401	-0.0221	0.059
	[0.124]	[0.109]	[0.0648]	[0.0607]	[0.0610]
Youngest child	-0.0726	-0.0514	0.0139	0.0647	0.147[**]
	[0.0470]	[0.0466]	[0.0456]	[0.0454]	[0.0486]
Age (in months)	0.0082	-0.0022	0.003	0.0059	0.0095
	[0.0055]	[0.0049]	[0.0043]	[0.0047]	[0.0058]
Primary carer (PC) characteristics					
No partner	0.0294	0.0848[*]	0.106[*]	0.0837	-0.0364
	[0.0399]	[0.0428]	[0.0484]	[0.0513]	[0.0553]
Employed	0.192[**]	0.206[**]	-0.122	0.112	0.0607
	[0.0575]	[0.0778]	[0.105]	[0.102]	[0.0991]
Hours of work	-0.0005	0.0027	0.0120[**]	0.0023	0.0013
	[0.0019]	[0.00253]	[0.0034]	[0.0034]	[0.0031]
Highest Educational qualification[b] (reference is Year 10/11)					
Year 9 or less	0.0438	0.0284	0.0585	0.0879	-0.029
	[0.0566]	[0.0593]	[0.0641]	[0.0669]	[0.0924]
Year 12	0.0118	-0.0611	0.0469	0.0755	0.0539
	[0.0550]	[0.0650]	[0.0718]	[0.0733]	[0.0780]
Certificate (other)	-0.0711	-0.172[**]	-0.0707	0.00698	-0.0129
	[0.0524]	[0.0580]	[0.0596]	[0.0623]	[0.0665]
Advanced Diploma or Certificate III/IV	-0.012	-0.0527	-0.042	0.0237	0.0092
	[0.0446]	[0.0512]	[0.0547]	[0.0564]	[0.0638]
University Degree	0.0525	0.0476	0.398[**]	0.388[**]	0.251[**]
	[0.0649]	[0.0865]	[0.124]	[0.132]	[0.0963]
Speaks Indigenous language	0.0089	-0.0615	-0.144[*]	0.117[+]	0.0615
	[0.0545]	[0.0541]	[0.0688]	[0.0679]	[0.0695]
Family environment					
Number of other siblings	0.0047	-0.0195	-0.0496[*]	-0.0884[**]	0.000878
	[0.0220]	[0.0206]	[0.0218]	[0.0252]	[0.0294]
Household size	-0.0133	0.00572	0.00269	0.0014	-0.0399
	[0.0154]	[0.0155]	[0.0138]	[0.0204]	[0.0243]

Wave [Target age in years]	Wave 1 [0.5-1.5]	Wave 2[a] [1.5-2.5]	Wave 3 [2.5-3.5]	Wave 4 [3.5-4.5]	Wave 5 [4.5-5.5]
Episode of Homeless: Yes	-0.124+	-0.152+	-0.0816	-0.0207	0.088
	[0.0684]	[0.0809]	[0.0730]	[0.0745]	[0.0767]
Housing Tenure (reference is public housing):					
Owned/mortgage	0.019	0.029	0.0552	0.0695	0.0894
	[0.0559]	[0.0672]	[0.0680]	[0.0713]	[0.0764]
Private rental	0.0233	0.0447	0.0898	0.0735	0.111+
	[0.0473]	[0.0559]	[0.0579]	[0.0597]	[0.0620]
Social	-0.0299	0.00357	0.0117	-0.026	-0.0475
	[0.0551]	[0.0568]	[0.0630]	[0.0717]	[0.0782]
Other	0.201+	-0.00601	0.1	-0.0604	-0.0642
	[0.108]	[0.136]	[0.101]	[0.0911]	[0.0882]
	[0.0454]	[0.0520]		[0.0521]	[0.0494]
Number of books at home (reference is 0-10):					
Books: 11-20[c]	-0.0278	-0.0417		0.160*	0.165*
	[0.0527]	[0.0730]		[0.0746]	[0.0820]
Books: 21-30	0.0178	0.0689		0.223**	0.0296
	[0.0615]	[0.0565]		[0.0779]	[0.0917]
Books: 30+	0.0272	0.0033		0.200**	0.0888
	[0.0443]	[0.0568]		[0.0650]	[0.0710]
Number of waves SC was read books	0.0184	0.0926**	0.0470+	-0.0053	-0.0069
	[0.0407]	[0.0328]	[0.0257]	[0.0242]	[0.0233]
Family activities index	-0.0188	-0.0751*		-0.0229	-0.00413
	[0.0235]	[0.0317]		[0.0314]	[0.0364]
Indigenous culture, and community					
Indigenous background index	0.0487*		-0.0332	-0.0153	
	[0.0221]		[0.0267]	[0.0298]	
Community bad for kids (1=very good to 5=really bad)	-0.0089	-0.0005	0.0620*	0.0345	0.0054
	[0.0209]	[0.0245]	[0.0255]	[0.0272]	[0.0316]
No places to play in community (1=lots of parks/playgrounds to 5=no, none)	0.0123	-0.0211	-0.0484*	0.0018	0.0367
	[0.0158]	[0.0192]	[0.0203]	[0.0215]	[0.0235]
Level of relative isolation (reference is low):					
None	0.0455	0.0202	0.148**	0.0236	0.0737
	[0.0457]	[0.0523]	[0.0536]	[0.0549]	[0.0565]
Moderate	0.0174	-0.0348	-0.0573	-0.196*	-0.203*
	[0.0627]	[0.0781]	[0.0782]	[0.0861]	[0.0944]
Extreme	0.134+	0.194*	0.229*	-0.00245	0.0618
	[0.0814]	[0.0959]	[0.0994]	[0.116]	[0.115]

Wave [Target age in years]	Wave 1 [0.5-1.5]	Wave 2[a] [1.5-2.5]	Wave 3 [2.5-3.5]	Wave 4 [3.5-4.5]	Wave 5 [4.5-5.5]
IRISEO (reference is quintile 3)					
Quintile 1	-0.0557	-0.0681	0.073	0.0123	-0.113
	[0.0776]	[0.0870]	[0.0845]	[0.0940]	[0.111]
Quintile 2	-0.07	-0.0502	-0.0665	-0.0133	-0.156[*]
	[0.0509]	[0.0553]	[0.0569]	[0.0601]	[0.0647]
Quintile 4	0.0315	0.0392	0.0161	0.0441	-0.095
	[0.0489]	[0.0597]	[0.0662]	[0.0634]	[0.0651]
Quintile 5	-0.013	0.0321	-0.049	0.0744	-0.0668
	[0.0545]	[0.0633]	[0.0715]	[0.0679]	[0.0680]
Observations	541	482	484	445	400
Pseudo R^2	0.25	0.293	0.267	0.289	0.264

Notes: a) The actual age of the child may vary in waves 2 to 5, as discussed in Section 3.2.
b) The lower bound education variable is used; results based on the higher bound education variable are available online (Table OA2)).
c) In Wave 3 families were not asked about the number of books in the home.
Standard errors in brackets; +, *, and ** denote significance at the 10, 5, and 1%-level, respectively.
Other variables included in the regression, but not reported here: study child speaks indigenous language, study child enrolled in pre-primary schooling, principal carer is non-Indigenous, grandparent present, moved house in last 12 months, number of homes since birth, home needs major repairs, number of people per bedroom, number of waves child was told stories, community is safe, weekly family income, missing income dummy, number of major life events, index of life events, index of financial stress, and income support being the main source of income (full results available online in Table OA2a).
Source: Authors' calculations based on LSIC.

Results in Table 7 suggest that the probability of using childcare is not uniform across groups and there are some children who are more likely to attend formal childcare than others. The strongest patterns we found indicate that:

- Children whose primary carer went to University have a significantly higher probability of ever using childcare after reaching the age of two years. The marginal effects are substantial: around 40 percentage points higher at age three and four, and just over 26 percentage points higher at age five.

- The employment status and hours of work of primary carers are correlated with ever using childcare especially during the first three years of life. Children with employed primary carers and carers who work for more hours are more likely to attend formal childcare. Again, the impacts can be substantial: around 20 percentage points higher if the primary carer is employed, and over 1 percentage point higher per hour worked. However, this effect is not statistically significant at all ages of the child.

- Single parents are more likely to use formal childcare for their child (an increase of around 10 percentage points for children aged between 1.5 and 4.5 years). This is to be expected, as there is no second parent who can also provide childcare for the child.[19]

- Children with more siblings are less likely to attend formal childcare. The cost of childcare for multiple children may be prohibitively high, and older children may look after younger siblings.

- Children who are read to at home and children who have more than 10 books at home have a higher probability of ever attending childcare, with 'being read to' being more important in the earlier years, and the number of books being more important after three years of age.

Weaker and less consistent patterns suggest the following:

- Compared to children living in low remoteness areas, young children in very remote areas are more likely to use childcare, and children older than 3.5 in moderately remote areas are less likely to have ever used childcare.[20]

- Children in families that have experienced homelessness are less likely to have ever used childcare up to age 2.5; although this association is only statistically significant at the 10%-level.

- Children living in a community that is considered to be bad for children and in which there are places for children to play are more likely to have ever attended childcare in Wave 3.

- Children with a higher score on the Participation in Indigenous culture index are more likely to use formal childcare in Wave 1.

- Children in families with a broader array of parental investments as measured by our family activities index are less likely to have ever attended childcare in Wave 2.

There is substantial cross-wave variation and sometimes a lack of consistency in the effect of the covariates on the probability of using childcare. Thus, there are variables that are significant only in one wave and variables for which the direction of the effect changes across waves. Some variation is to be expected as the reasons for using childcare change over time, and the raw data show considerable differences in childcare use by age of the child. However, only in

---

[19] For many families, information on the partner is missing, so no further partner characteristics could be included in the analysis without reducing the sample size considerably.

[20] Hewitt and Walter (2014) report that in 2011 Indigenous children living in remote areas were more likely to be enrolled in preschool than Indigenous children in other areas. We included an indicator for being from Torres Strait Islander descent to check whether this is driving the results, and found that this was not the case.

one instance does a significant negative effect turn into a significant positive effect from one wave to the next. This occurs for the variable indicating that the parent speaks an Indigenous language from Wave 3 to Wave 4.

As a sensitivity analysis, we re-estimated all models using a balanced panel of just under 250 observations to check whether this variability was (partly) due to changes in the sample composition across waves. The results (available online in Table OA3) are similar to those from the original regressions, although some coefficients are larger for this more selective subsample. Examples are the primary carer having a university degree and their labour supply, as well as some other characteristics (such as income or community characteristics).

To further understand changes in childcare use from wave to wave, we also estimated models for the probability of using childcare in waves 2 to 5 conditioning on childcare attendance in the previous wave. Estimates from these models, available online (Table OA4), indicate strong persistence in childcare use especially at waves 3 and 4 (for ages between 2.5 and 4.5). Results from these regressions are mostly consistent with results from the original regressions. Thus, for instance, children from more advantaged families in terms of education and employment are still more likely to use childcare, even after controlling for childcare use in the past. Also, children who are read to and who have more books at home have a higher probability of attending childcare than other children.

Our results are similar to those reported in Houng *et al.* (2011) for the general population of Australian children (at ages 2 to 3). Particularly, parental employment and education are important predictors of childcare usage for both groups. The marginal effect on the probability of using formal childcare estimated in Houng *et al.* is slightly larger at just over an additional 30 percentage points for a full-time working main carer, and 22 percentage points for a part-time working main carer. In the general population, it is the lowest educated group that is 6 percentage points less likely to use childcare for their children. The difference in childcare use by parental education level is much larger for Indigenous children than children in the general population, but the impact of the main carer having no partner is about half the size for Indigenous children. The marginal effects of the number of siblings are similar in size when comparing Indigenous children with the general population.

In addition, several other characteristics that are not (always) statistically significant in our sample of Indigenous children have the same sign as comparable characteristics in the general

population by Houng *et al.* (2011). These include speaking a language other than English at home and residing in a more or less disadvantaged area.

Despite these similarities, there is an important difference between the two samples: the use of formal childcare is much lower for Indigenous children than for children in the general Australian population. No more than 30% of Indigenous children attend childcare at any given age, while Houng *et al.* (2011) report that 55% of the general population of children aged 2 to 3 use formal childcare. Kalb *et al.* (2014) examined the same population, and report that children in families on low and middle incomes (under $73,000) and children in two-parent families where only one of the parents is employed are the least likely to attend any formal childcare. Compared to non-Indigenous parents, Indigenous parents are on average on lower incomes and they are also less likely to be employed (Kalb *et al.*, 2014). This may partially explain some of the differences in childcare usage between Indigenous and non-Indigenous children. Furthermore, such differences could also be due in part to the different locations in which these groups reside, with the Indigenous population being substantially more likely to live in remote areas where less formal childcare is available.[21]

5.2. Formal childcare usage and child cognitive development

5.2.1 <u>Regression results</u>

Tables 8 and 9 report the estimation results for different specifications of the model in equation (2) when childcare use is measured using the indicator and the count variables, respectively.[22] For ease of presentation and given the focus of this paper, we only report the estimated coefficients on the childcare variables.[23] Standardised values of the cognitive scores are used in all regressions so that coefficients are expressed in standard deviations.

---

[21] Knowing whether this is the case would be useful to interpret these different levels of usage. Average usage conditional on the local availability of formal childcare locally would allow us to understand whether this is purely a low demand for childcare issue or whether families are constrained in their choice due to where they live. However, information on the availability of formal childcare is not observed by us.

[22] Results using attendance at a day care centre only were very similar to the results presented here. As were results based on the higher bound education variable. These results are available online (Table OA5 and Table OA6).

[23] Full results are available online (Table OA7), and are consistent with expectations.

Table 8 Cognitive outcomes and formal childcare use - regression results for the childcare variable "ever in formal childcare" (in standard deviations)

Specification	MacArthur Bates Wave 3	Renfrew Wave 4	Renfrew Wave 5	Who am I Wave 5
Ever in formal childcare	0.252**	0.560**	0.509**	0.186*
No other control variables	[0.0928]	[0.0940]	[0.0983]	[0.0920]
	0.016	0.079	0.065	0.011
Ever in formal childcare	0.161	0.484**	0.419**	0.138
Controls: Birth and early years	[0.0987]	[0.101]	[0.105]	[0.0958]
	0.044	0.151	0.113	0.082
Ever in formal childcare	0.180*	0.514**	0.404**	0.190+
Controls: Study child characteristics	[0.0873]	[0.0923]	[0.0968]	[0.0808]
	0.153	0.148	0.172	0.306
Ever in formal childcare	0.162	0.333**	0.303**	0.0847
Controls: Parent characteristics	[0.103]	[0.0990]	[0.102]	[0.0995]
	0.083	0.193	0.182	0.071
Ever in formal childcare	0.137	0.202*	0.126	0.0319
Controls: Family environment	[0.0979]	[0.0989]	[0.099]	[0.103]
	0.083	0.241	0.276	0.078
Ever in formal childcare	0.141	0.333**	0.263**	0.0934
Controls: Indigenous culture and community	[0.0954]	[0.0923]	[0.0949]	[0.0962]
	0.105	0.243	0.25	0.085
Ever in formal childcare	0.252**	0.499**	0.391**	0.106
Controls: Financial and life events	[0.0956]	[0.0953]	[0.101]	[0.0945]
	0.022	0.105	0.112	0.048
Ever in formal childcare	0.0774	0.108	0.0639	-0.0115
Controls: All variables except birth and early years	[0.100]	[0.101]	[0.0982]	[0.0920]
	0.29	0.381	0.432	0.431
Sample size with all variables except birth and early years characteristics (used for all analyses, except the one including birth and early years)	470	415	385	360
Sample size including birth and early years characteristics (used only for analysis including birth and early years)	422	377	350	330

Notes: We report three estimates for each specification: the coefficient of the childcare use variable, its standard error in brackets, and the value of the R^2 below that. +, *, and ** denote significance at the 10, 5, and 1%-level, respectively.
Source: Authors' calculations based on LSIC.

Table 9 Cognitive outcomes and formal childcare use - regression results for the childcare variable "waves in formal childcare count" (in standard deviations)

Specification	MacArthur Bates Wave 3	Renfrew Wave 4	Renfrew Wave 5	Who am I Wave 5
Formal childcare count	0.113**	0.199**	0.157**	0.0569*
No other control variables	[0.0418]	[0.0335]	[0.0311]	[0.0287]
	0.015	0.079	0.062	0.011
Formal childcare count	0.0858+	0.166**	0.127**	0.0437
Controls: Birth and early years	[0.0449]	[0.0363]	[0.0333]	[0.0299]
	0.047	0.147	0.11	0.082
Formal childcare count	0.0834*	0.191**	0.131**	0.0688*
Controls: Study child characteristics	[0.0393]	[0.0327]	[0.0301]	[0.0248]
	0.154	0.155	0.175	0.31
Formal childcare count	0.0654	0.0968*	0.0815*	0.0173
Controls: Parent characteristics	[0.0486]	[0.0375]	[0.0355]	[0.0337]
	0.082	0.184	0.174	0.07
Formal childcare count	0.0575	0.0542	0.0089	-0.0113
Controls: Family environment	[0.0451]	[0.0364]	[0.0326]	[0.0333]
	0.082	0.237	0.273	0.078
Formal childcare count	0.0707	0.110**	0.0689*	0.0175
Controls: Indigenous culture and community	[0.0432]	[0.0334]	[0.0303]	[0.0301]
	0.106	0.239	0.245	0.083
Formal childcare count	0.117**	0.174**	0.110**	0.0252
Controls: Financial and life events	[0.0439]	[0.0350]	[0.0328]	[0.0302]
	0.023	0.100	0.103	0.047
Formal childcare count	0.0458	0.0137	0.0005	-0.0119
Controls: All variables except birth and early years	[0.0477]	[0.0386]	[0.0338]	[0.0308]
	0.29	0.379	0.431	0.431
Sample size with all variables except birth and early years characteristics (used for all analyses, except the one including birth and early years)	470	415	385	360
Sample size including birth and early years characteristics (used only for analysis including birth and early years)	422	377	350	330

Notes: We report three estimates for each specification: the coefficient of the childcare use variable, its standard error in brackets, and the value of the R^2 below that. +, *, and ** denote significance at the 10, 5, and 1%-level, respectively.
Source: Authors' calculations based on LSIC.

Results from the basic specification without additional covariates indicate that Indigenous Australian children who attended childcare have on average higher scores on all cognitive measures than Indigenous children who never attended childcare. All results are statistically significant and, despite the smaller sample used in Table 8, they are similar to the results reported in Table 3. In the case of the MacArthur Bates scale, childcare users score on average over 0.25 standard deviations more than non-users, which implies a difference of more than 7.52 points between the two groups. Differences in the Renfrew tests scores show that children aged 3.5 to 5.5 years who attended childcare in at least one wave scored on average 0.5 to 0.6 standard deviations higher than non-users of the same age. For the "Who Am I?" test, the difference in the average score between formal childcare users and non-users is 0.19 standard deviations These results are of a similar order of magnitude as those found for the general Australian population in Houng *et al.* (2011) and Kalb *et al.* (2014) using the LSAC data. These papers find that formal childcare users score 0.1 to 0.5 standard deviations better on a learning index measured at age 4 to 5, depending on household income and gender of the child.

Analogous estimates using the childcare count variable also show a positive relationship between formal childcare use and children's cognitive development. Indeed, results from the basic specification suggest that an additional year of childcare is associated with an increase in the average score of between 0.06 and 0.20 of a standard deviation per wave attending childcare, depending on the cognitive measure considered. [24]

Interestingly, our results suggest that differences in cognitive scores between Indigenous Australian children who attend and do not attend childcare are *entirely* driven by differences in observable characteristics between these two groups. In our preferred specification including all covariates except the characteristics at birth and in the early years,[25] the coefficients on the different measures of childcare use are not statistically significant for any of the cognitive outcomes. Results from the regression models that control for one set of covariates at a time show that differences in birth and early years characteristics; child and parental characteristics; Indigenous culture, and community; home environments; and family and life events all

---

[24] An alternative specification using dummy variables for childcare use in 1 wave, 2 waves, and 3 or more waves. leads to similar results, implying an increase in score with more waves in childcare. Another specification, incorporating weekly hours of childcare use in the intensity measure, again leads to similar results. These results are available online (Table OA8a and Table OA8b).

[25] The inclusion of the birth and early year variables would eliminate about 20-23% of the sample due to the large number of missing values for these variables. The estimation of the preferred specification including these covariates led to the same conclusions. These results are available from the authors online (Table OA7).

contribute to explaining differences in cognitive outcomes between formal childcare users and non-users. That is, in all specifications, the estimated coefficient on the childcare use variable becomes smaller (and less significant) after controlling for each set of characteristics. The quality of the family environment appears to have the most pronounced role in explaining differences in cognitive scores between childcare users and non-users.[26] Indeed, controlling for these differences alone eliminates the gap in cognitive scores, regardless of the childcare variable used in the regressions. The exception to this is the Renfrew test score in Wave 4 for which the strongest impacts are observed and for which a significant difference still persists after controlling for the family environment, although this is only when childcare attendance is measured by the "ever use" variable.

The pattern of these results is again similar to what is found in Houng *et al.* (2011) and Kalb *et al.* (2014), but in these two studies the coefficient on childcare remains significant after adding a broad range of household and parental characteristics. However, this could be due to the much larger sample size available when using the LSAC (compared to using LSIC). Gormley *et al.* (2005) estimated effects of around 0.6 to 0.9 of a standard deviation for Native Americans, depending on the cognitive outcome measure, which is more substantial than our estimate, possibly because the intervention is a pre-kindergarten programme rather than just childcare. It is even more difficult to compare the results in this paper to those for the Indigenous populations in Benzies *et al.* (2011, 2014) and Mughal *et al.* (2015) as a very different intervention (a parent and child programme) and outcome measure (within-child improvement in an age-standardised measure) was used in these papers.

We also estimated the models in Table 8 adding interactions between the "ever in childcare" dummy and indicators for being from advantaged and disadvantaged backgrounds as measured by high and low education level of the parents respectively. There is no evidence from our preferred specification (including all covariates) that childcare has a different impact on children from the most advantaged and disadvantaged families, except for the Renfrew test outcomes in wave 5 where children from highly educated parents appear to benefit more from attending childcare. Similar results were found when community disadvantage as measured by the IRISEO index was interacted with the "ever in childcare" dummy, except that now children from more advantaged communities appear to benefit less from attending childcare in terms of

---

[26] Amongst the family environment variables, housing tenure and reading or telling stories as well as the number of books are most relevant in terms of reducing the coefficient on childcare use.

Renfrew test outcomes in wave 5. When the impact of childcare was allowed to vary by the level of remoteness, the impact of childcare is the same for all outcomes, except for the "Who Am I?" score, where children living in non-isolated locations benefit more from attending childcare (these additional results are available online in Table OA9).

5.2.2 Propensity score matching results

Results from the propensity-score matching analyses yield similar conclusions to those from the regression analyses. In particular, we find no statistical evidence of the impact of formal childcare use on the cognitive outcomes of Australian Indigenous children once we control for differences in the characteristics of children and their families.

As the results reported in Table 10 show, our estimates of the average treatment effect on the treated (ATET) for the four cognitive measures are small and sometimes even negative. In only one case, the ATET is significant at the 10-% level. This lack of significance holds irrespective of whether we consider a binary or a multinomial treatment variable for the use of childcare.

Interestingly, our results suggest that Indigenous Australian children who never used childcare might have benefited from attending childcare, had they done so. Estimates of the average treatment effect on the untreated (ATEU) for the binary treatment case are always positive although not statistically significant, except for the Renfrew test in Wave 4 which is significant at the 5%-level. Further investigation of this result, using larger samples and/or better information on the availability and affordability of formal childcare, would be useful.

The measures of covariate balancing for the binary treatment suggest a good balance of covariates in the matched samples used to derive the ATET and the ATEU. That is, the treatment and control groups have a similar distribution of characteristics. In fact, for all cognitive measures, the median standardised bias is below 5% most of the time –the threshold usually employed in the literature to identify good matching quality. Furthermore, the null hypothesis of joint significance is not rejected before the matching, but it is rejected after the matching. This indicates that before matching the characteristics of the control and treatment groups were statistically different, while after the matching that was no longer the case. Finally, tests based on the R and B statistics also indicate a good quality match between the treatment and control groups (not reported here but available online in Table OA10).

Table 10. Formal childcare use and cognitive outcomes (in standard deviations) - propensity-score matching estimates

Treatment	Number treated	ATET (Std.)	Median Bias in % Before	After	p>chi2 Before	After	ATEU (Std.)	Median Bias in % Before	After	p>chi2 Before	After
			MacArthur Bates (Wave 3, sample=470)								
Binary											
Ever in childcare	192	-0.03 (0.15)	21.70	7.70	0.00	1.00	0.02 (0.13)	25.70	5.70	0.00	0.87
Multiple											
1 wave	71	0.02 (0.16)	17.80	4.50	0.00	1.00	0.01 (0.18)	17.80	8.20	0.00	0.07
2 waves	61	0.06 (0.18)	23.90	4.60	0.00	1.00	0.05 (0.20)	23.90	7.90	0.00	0.89
3 or more waves	60	0.11 (0.29)	37.40	11.60	0.00	1.00	0.42+ (0.25)	37.40	13.50	0.00	0.80
			Renfrew test (Wave 4, sample=415)								
Binary											
Ever in childcare	187	0.11 (0.13)	17.80	2.20	0.00	1.00	0.24* (0.14)	17.80	4.00	0.00	1.00
Multiple											
1 wave	53	0.06 (0.18)	16.40	5.40	0.00	1.00	0.19 (0.19)	16.40	8.40	0.00	0.98
2 waves	48	0.34+ (0.21)	19.60	8.30	0.00	1.00	0.67** (0.23)	19.60	10.20	0.00	0.58
3 or more waves	86	-0.13 (0.25)	30.30	10.80	0.00	1.00	-0.15 (0.26)	30.30	7.30	0.00	0.70
			Renfrew test (Wave 5, sample=385)								
Binary											
Ever in childcare	179	0.08 (0.13)	20.70	3.70	0.00	1.00	0.07 (0.14)	20.70	4.50	0.00	1.00
Multiple											
1 wave	46	0.08 (0.16)	9.90	4.20	0.06	1.00	0.03 (0.16)	9.90	6.10	0.26	1.00
2 waves	46	0.07 (0.20)	17.30	5.40	0.02	1.00	0.05 (0.20)	17.30	3.20	0.02	1.00
3 or more waves	87	0.22 (0.22)	22.10	9.70	0.00	1.00	0.06 (0.25)	22.10	13.20	0.00	0.92
			Who Am I (Wave 5, sample=360)								
Binary											
Ever in childcare	171	-0.11 (0.13)	23.00	3.70	0.00	1.00	0.02 (0.13)	23.00	4.10	0.00	1.00
Multiple											
1 wave	41	-0.08 (0.19)	11.40	5.50	0.13	1.00	0.03 (0.22)	11.40	8.90	0.13	0.91
2 waves	45	-0.09 (0.18)	20.80	5.40	0.01	1.00	0.02 (0.17)	20.80	4.90	0.01	1.00
3 or more waves	85	0.23 (0.23)	24.30	8.30	0.00	1.00	-0.11 (0.21)	24.30	8.40	0.00	1.00

Notes: +, * and ** denote significance at the 10, 5, and 1%-level, respectively. For an explanation of the p value of the likelihood ratio test and the median bias indicators see Section 3.

Source: Authors' calculations based on LSIC.

For the multiple treatment case, the quality of the matching declines as we change our definition of treatment from ever attending childcare to attending childcare for one wave, for two waves, or for three or more waves. As the number of waves of attending childcare increases, the quality of the matching decreases. The median standardised bias measures for attending childcare for two and for three waves are in general well above 5%, which highlights the difficulty of finding good matches between children who never attended formal childcare and those who attended regularly during their childhood.

6. Conclusions

The literature indicates that the disadvantage experienced by Indigenous children is evident on a range of cognitive and developmental measures from an early age. Trying to improve access to early childhood education has been a common policy response to the Indigenous gap in countries like Australia and New Zealand. But, to what extent does formal childcare influence the cognitive development of Indigenous children? Existing evidence on the relationship between childcare and children's cognitive outcomes is mostly based on general surveys that fail to recognise the specificities of Indigenous populations and are not representative of these populations due to small population numbers. As a consequence, the influence of childcare attendance on Indigenous children's cognitive outcomes is not well understood. This paper has contributed to filling this gap in knowledge by investigating the influence of attendance at formal childcare on the cognitive development of Indigenous Australian children. This was achieved by using unique data from LSIC, an unprecedented longitudinal survey that tracks a cohort of Indigenous children from birth and includes detailed information on their childcare use and cognitive outcomes.

Our results indicate that childcare users perform better on a range of cognitive tests measured at different ages in the preschool period. However, selection into childcare explains these results in their entirety. Hence, our results indicate that it is not that Indigenous Australian children improve their cognitive outcomes by attending formal childcare, but that children who have characteristics (particularly family environment characteristics) associated with better cognitive outcomes tend to be enrolled into childcare. This is consistent with our analyses of the predictors of childcare participation, which show that the probability of using childcare is far from uniform across different socioeconomic groups. Children from more advantaged families (e.g. those with highly educated or employed primary carers) are substantially more likely to attend formal childcare than children from less advantaged families. These results are similar to those for Australian children more generally (Houng *et al.* 2011; Kalb *et al.* 2014).

Using the current data and taking particular care in defining our variables (especially the very important education variable) and in selecting the same sample across all specifications, we find no significant association between formal childcare use and Indigenous-Australian children's cognitive outcomes after controlling for a broad range of individual and household characteristics that are correlated with formal childcare use and child development. This is consistent with the results reported in AHB (2015), where no statistical relationship between childcare attendance and school-age cognitive outcomes was found after taking into account differences in the socioeconomic characteristics of families. Studies based on the general Australian population of children usually find reduced relationships as well, although these are often still significant (possibly due to their larger sample size). Some of the overseas studies have found significant effects for Indigenous children, but these are often difficult to compare directly as somewhat different (and more substantial) interventions have been studied in these studies.

However, our results suggest that the children who are currently not attending childcare (i.e. the relatively disadvantaged children) might benefit from attending, as indicated by a modest result on the Renfrew Wave 4 test which is just short of being significant at the 5%-level. This is consistent with findings from other studies demonstrating positive results from high quality early childhood education programmes, especially for children from more disadvantaged backgrounds (Felve and Lalive, 2012). Indeed, future research should investigate whether the quality of childcare provided remains associated with positive cognitive outcomes for Australian Indigenous children from more disadvantaged backgrounds even after controlling for other characteristics.

Our results further suggest that there may be benefit for cognitive development in improving the home environment (e.g., by educating current and future parents on the importance of the home environment). Indeed, our results and the findings of other studies (e.g., Farrant el al. 2014) indicate that interventions that focus on educating current and future parents on the importance of parenting practices like reading or telling stories to their children, as well as providing children's books to increase the number of these in the home, are likely to have significant benefits. Comparing the impact of childcare attendance to this kind of parenting programme could be another fruitful avenue for future research.

The LSIC data is a major improvement compared to what was available before this data collection started, both in terms of collecting information on a larger number of Indigenous children than ever before and in terms of the type of information that is collected. However,

the sample size is still relatively small, information on local formal childcare availability is missing and specific characteristics of formal childcare are not available for a sufficient proportion of children attending formal childcare to include in the analysis. In addition, the LSIC data are not collected as a probabilistic sample which means no general inferences can be drawn for the population of Indigenous children based on our analysis. If these limitations could be overcome, this question could be investigated further, so that stronger evidence for or against recommending childcare as a way to improve Indigenous children's education outcomes could be provided. That said, the LSIC data is a resource that has enabled a much more detailed analysis than would have been feasible otherwise, and that is not currently feasible in other countries with large Indigenous populations, such as Canada or the US.

References

Anderson, I. *et al.* (2016) Indigenous and tribal peoples' health (*The Lancet*-Lowitja Institute Global Collaboration): a population study, *The Lancet*, published online on 20 April 2016: DOI: 10.1016/S0140-6736(16)00345-7.

Arcos Holzinger, L. and Biddle, N. (2015) The relationship between Early Childhood Education and Care (ECEC) and the outcomes of Indigenous children: Evidence from the Longitudinal Study of Indigenous Children (LSIC), CAEPR paper, Research School of Social Sciences, Australian National University, Canberra.

Benzies, K., Edwards, N., Tough, S., Nagan, K., Mychasiuk, R., Keown, L.-A. and Donnelly, C. (2011) Effects of a two-generation preschool programme on receptive language skill in low-income Canadian children, *Early Child Development and Care*, 181(3), 397-412, DOI: 10.1080/03004430903424579.

Benzies, K., Mychasiuk, R., Kurilova, J., Tough, S., Edwards, N. and Donnelly, C. (2014) Two-generation preschool programme: immediate and 7-year-old outcomes for low-income children and their parents, *Child and Family Social Work*, 19, 203–214.

Biddle, N. (2007) Indigenous Australians and preschool education: Who is attending? *Australian Journal of Early Childhood*, *32* (3), 9-16.

Bradbury, B., Corak, M., Waldfogel, J., Washbrook, E. (2011) Inequality during the early years: Child outcomes and readiness to learn in Australia, Canada, United Kingdom, and United States, IZA Discussion Paper series No. 6120, Forschungsinstitut zur Zukunft der Arbeit, Bonn, Germany.

Bradley, S., Draca, M., Green, C., and Leeves, G. (2007) The magnitude of educational disadvantage of indigenous minority groups in Australia, *Journal of Population Economics*, 20, 547–569.

Buckley, S., Underwood, C. and Purdie, N. (2009) Report on Who Am I? and the Renfrew Word Finding Vocabulary Test: Footprints in Time, Wave 2, Australian Council for Educational Research for Footprints in Time: Longitudinal Study of Indigenous Children.

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

Clotfelter, C.T., Helen F. L., and Jacob L. (2009) The Academic Achievement Gap in Grades 3 to 8, *Review of Economics and Statistics*, 91 (2): 398–419.

De Bortoli, L.J. and Thomson, S. (2010) Contextual factors that influence the achievement of Australia's Indigenous students: Results from PISA 2000 – 2006, Australian Council for Educational Research (ACER), Camberwell: ACER press.

Department of Families, Housing, Community Services and Indigenous Affairs (2011) Footprints in Time: The Longitudinal Study of Indigenous Children—Key Summary Report from Wave 2, FaHCSIA, Canberra.

Department of the Prime Minister and Cabinet (2015) Closing the gap: Prime Minister's report 2015. Canberra: Department of the Prime Minister and Cabinet, Australia; available at https://www.dpmc.gov.au/sites/default/files/publications/Closing_the_Gap_2015_Report.pdf.

Dodson, M., Hunter, B., McKay, M. (2012) Footprints in Time: The Longitudinal Study of Indigenous Children: A guide for the uninitiated. *Family Matters* 91, 68–82.

Duncan, G.J. and Magnuson, K. (2013) Investing in Preschool Programs. *Journal of Economic Perspectives*, 27(2), 109–132.

Edwards, M.C. (2009) An introduction to item response theory using the Need for Cognition Scale. Social and Personality Psychology Compass. 3/4:507–529.

Farrant, B.M., Shepherd, C.C.J., Walker, R.D. and Pearson, G.C. (2014) Early Vocabulary Development of Australian Indigenous Children: Identifying Strengths. *Child Development Research*, Article ID 942817.

Felfe, C. and Lalive, R., (2012) Early Child Care and Child Development: For Whom it Works and Why. IZA Discussion Paper No. 7100, Bonn, Germany.

Fiorini, M. and Keane, M.P. (2014) How the Allocation of Children's Time Affects Cognitive and Noncognitive Development. *Journal of Labor Economics*, 32(4), 787–836.

Fischer, S., and Stoddard, C. (2013) The academic achievement of American Indians. *Economics of Education Review*, 36, 135–152.

Fitzpatrick, M.D., Grissmer, D. and Hasted, S. (2011) What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment. *Economics of Education Review*, 30, 269–279.

Ford, M. (2013) Achievement gaps in Australia: What NAPLAN reveals about education inequality in Australia, *Race, Ethnicity and Education*, 16 (1), 1–23.

Gormley, W.T., Gayer, T., Phillips, D. and Dawson, B. (2005) The Effects of Universal Pre-K on Cognitive Development, *Developmental Psychology*, 41(6), 872 – 884.

Gould, J. (2000) Closing the gaps? *Political* Science, 52: 116-122.

Hewitt, B. (2012) The Longitudinal Study of Indigenous Children; Implications of the study design for analysis and results, Institute for Social Science Research report, University of Queensland.

Hewitt, B. and Walter, M. (2014) Preschool participation among Indigenous children in Australia, *Family Matters*, no. 95, 41 – 50.

Houng, B., Jeon, S.-H., and Kalb, G. (2011) The effects of childcare on child development, Report for the Department of Education, Employment and Workplace Relations, Melbourne Institute of Applied Economic and Social Research, University of Melbourne; viewed March 2016 <http://melbourneinstitute.com/downloads/labour/1-10_FINAL_REPORT.pdf>

Imbens, G. (2000) The Role of the Propensity Score in Estimating Dose-Response Functions, *Biometrika*, 87(3), 706-710.

Kalb, G., Tabasso, D. and Zakirova, R. (2014) Children's participation in early childhood education and care, and their developmental outcomes by Year 5: A comparison between disadvantaged and advantaged children, Report for the Department of Education, Employment and Workplace Relations, Melbourne Institute of Applied Economic and Social Research, University of Melbourne; viewed March 2016 <http://melbourneinstitute.com/downloads/labour/4-13_FINAL_REPORT.pdf>

Lechner, M, (2002), Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies, *Review Economics and Statistics*, 84(2): 205-220, May.

Leigh, A. and Gong, X. (2009) Estimating cognitive gaps between Indigenous and non-Indigenous Australians, *Education Economics,* 17(2), 239 – 261.

Lohoar, S., Butera, N. and Kennedy, E. (2014) Strengths of Australian Aboriginal cultural practices in family life and child rearing. Child Family Community Australia Paper No. 25, Australian Institute of Family Studies.

Ludwig, J. and Miller, D.L., (2007) Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design. *Quarterly Journal of Economics*, 122(1), 159 – 208.

Magnuson, K.A., Ruhm, C., and Waldfogel, J. (2007). Does Prekindergarten Improve School Preparation and Performance? *Economics of Education Review*, 26(1). 33 – 51.

Mughal, M.K., Ginn, C.S., Perry, R.L. and Benzies, K.M. (2015) Longitudinal effects of a two-generation preschool programme on receptive language skill in low-income Canadian children to age 10 years, *Early Child Development and Care*, DOI: 10.1080/03004430.2015.1092141.

Perales F., Baffour B. and Mitrou F. (2015) Ethnic Differences in the Quality of the Interview Process and Implications for Survey Analysis: The Case of Indigenous Australians. *PLoS ONE* 10(6): e0130994. doi:10.1371/journal.pone.0130994

Phillips, G., McNaughton, S., and MacDonald, S. (2004). Managing the mismatch: Enhancing early literacy progress for children with diverse language and cultural identities in mainstream urban schools in New Zealand. *Journal of Educational Psychology*, 96(2), 309–323.

Richards, J., Vining, A., and Weimer, D. (2010) Aboriginal performance on standardized tests: Evidence and analysis from provincial schools in British Columbia, *The Policy Studies Journal*, 38, 47–67.

Rosenbaum P.R. and Rubin D.B. (1985) Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score, *The American Statistician*, 39(1), 33-39.

Rubin, D.B. (2001). Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation, *Health Services & Outcomes Research Methodology* 2, 169-188.

Sianesi, B. (2004) An Evaluation of the Active Labour Market Programmes in Sweden, *The Review of Economics and Statistics*, 86(1), 133–155.

Trudgett, M. and Grace, R. (2011) Engaging with early childhood education and care services: The perspectives of Indigenous Australian mothers and their young children, *Kulumun, Indigenous Online Journal*. 1, 15 – 36.

Appendix I

Table A.1 Proportion of children in targeted age category by type of test (in column percentages)

Cognitive test	MacArthur Bates scale (Wave 3)	Renfrew Word test (Wave 4)	Renfrew Word test (Wave 5)	Who Am I? (Wave 5)
Younger than targeted age range (up to 5-7 months)	9.79	9.90	10.20	10.35
In targeted age range	80.63	83.34	83.58	83.55
Older than targeted age range (up to 5-6 months)	9.58	6.76	6.23	6.11
Number of observations	470	414	402	377

Table A.2a Percentage of all observations in formal childcare as observed in Wave 5 when using alternative definitions (cell percentages)

Using definition 1[a]	Using definition 2[b]						
	0%	20%	40%	60%	80%	100%	Total
0%	53.55	0	0	0	0	0	53.55
20%	0	7.87	0	0	0	0	7.87
25%	0	4.41	0	0	0	0	4.41
33.33%	0	1.34	0	0	0	0	1.34
40%	0	0	8.83	0	0	0	8.83
50%	0	0	1.73	0	0	0	1.73
60%	0	0	0	8.06	0	0	8.06
66.66%	0	0	1.15	0	0	0	1.15
75%	0	0	0	1.92	0	0	1.92
80%	0	0	0	0	6.72	0	6.72
100%	0	0	0.19	0	0.77	3.45	4.41
Total	53.55	13.63	11.9	9.98	7.49	3.45	100

Notes: a) Definition 1: number of waves in childcare divided by number of waves in which the child is observed;
b) Definition 2: number of waves in childcare divided by number of study waves, including the current wave.

Table A.2b Percentage of observations in childcare for all observations across all waves when using alternative definitions

	Definition 1[a]	Definition 2[b]
0%	64.88	64.88
20%	1.4	2.43
25%	2.57	2.36
33.33%	3.39	3.08
40%	1.57	2.12
50%	6.3	5.72
60%	1.44	1.78
66.66%	2.74	2.57
75%	2.16	2.02
80%	1.2	1.34
100%	12.36	11.71
Number of observations	2,921	2,921

Notes: a) Definition 1: number of waves in childcare divided by number of waves in which the child is observed;
b) Definition 2: number of waves in childcare divided by number of study waves, including the current wave.

Table A.3 Probit model for the probability of remaining in the sample between consecutive waves, and across all waves (marginal effects)

	Waves 1-2	Waves 2-3	Waves 3-4	Waves 4-5	All waves[a]
Study child (SC) characteristics					
Female	-0.00324	-0.0153	-0.0565[+]	0.0528[+]	-0.0647
	[0.0202]	[0.0274]	[0.0308]	[0.0305]	[0.0398]
Youngest child	0.0292	-0.00396	-0.035	-0.0131	-0.100[+]
	[0.0300]	[0.0337]	[0.0361]	[0.0335]	[0.0602]
Oldest child[b]		0.0577	0.0178	0.0900[+]	
		[0.0707]	[0.0510]	[0.0519]	
age in months	-0.00395	-0.00834[*]	-0.00855[*]	-0.00800[*]	-0.0179[**]
	[0.00350]	[0.00375]	[0.00341]	[0.00362]	[0.00682]
Primary carer (PC) characteristics					
Non-Indigenous	-0.0115	0.0652	0.104[*]	-0.0253	0.124[*]
	[0.0299]	[0.0460]	[0.0501]	[0.0428]	[0.0598]
No partner	0.00271	-0.0520[+]	-0.0218	-0.0710[+]	-0.05
	[0.0228]	[0.0308]	[0.0355]	[0.0367]	[0.0469]
Employed	-0.0162	-0.0711	-0.0795	-0.0588	0.0493
	[0.0444]	[0.0615]	[0.0808]	[0.0803]	[0.0925]
hours of work	0.000504	0.000873	0.00201	0.00313	-0.00279
	[0.00153]	[0.00192]	[0.00277]	[0.00291]	[0.00301]
Highest educational qualification (reference is year 10/11):					
Year 9 or less	0.0312	0.0287	0.0508	-0.0252	0.062
	[0.0416]	[0.0471]	[0.0510]	[0.0426]	[0.0661]
Year 12	0.00254	-0.0579	-0.102[+]	0.00792	-0.059
	[0.0443]	[0.0405]	[0.0528]	[0.0488]	[0.0735]
Certificate other	-0.0172	0.0296	0.0136	0.102[*]	0.162[**]
	[0.0301]	[0.0405]	[0.0456]	[0.0500]	[0.0599]
Advanced diploma or	-0.0501[+]	0.0843[+]	0.052	0.116[*]	0.128[*]
Certificate III/IV	[0.0277]	[0.0459]	[0.0473]	[0.0476]	[0.0587]
University degree	-0.0102	0.0218	0.0126	0.123	0.137
	[0.0540]	[0.0621]	[0.0725]	[0.0935]	[0.101]
Speaks Indigenous language	-0.00025	0.00503	-0.071	-0.0725	-0.116
	[0.0374]	[0.0370]	[0.0468]	[0.0475]	[0.0705]
Family environment					
Number of other siblings	-0.0087	-0.00673	0.0287[+]	0.00615	0.0581[*]
	[0.0121]	[0.0151]	[0.0157]	[0.0190]	[0.0243]
Grandparent present[c]		0.14		0.0986	
		[0.0954]		[0.0694]	

	Waves 1-2	Waves 2-3	Waves 3-4	Waves 4-5	All waves[a]
Episode of Homelessness: Yes			0.0225	-0.0281	
			[0.0530]	[0.0570]	
Housing tenure (reference is public housing):					
House being paid off	0.0638	-0.013	0.0538	-0.00334	0.131[+]
	[0.0489]	[0.0552]	[0.0533]	[0.0574]	[0.0745]
Private rental	0.00563	-0.0429	0.143**	-0.0478	0.0729
	[0.0306]	[0.0415]	[0.0515]	[0.0442]	[0.0608]
Social housing	-0.0198	0.0834[+]	-0.00206	0.0693	0.0595
	[0.0284]	[0.0459]	[0.0450]	[0.0537]	[0.0634]
Other housing tenure	-0.0343	0.0486	-0.0734	0.119	0.00544
	[0.0606]	[0.103]	[0.0674]	[0.0765]	[0.148]
Number of books at home (reference is 0-10):					
Books at home: 11-20c	0.0418	-0.0338		0.0077	0.0615
	[0.0413]	[0.0521]		[0.0511]	[0.0637]
Books at home: 21-30		-0.0596		0.0382	
		[0.0462]		[0.0549]	
Books at home: 30+	-0.0326	-0.0966*		0.0517	-0.0482
	[0.0263]	[0.0425]		[0.0472]	[0.0527]
Number of waves SC was read books	0.00748	0.0444[+]	-0.00986	0.0119	0.0364
	[0.0235]	[0.0228]	[0.0196]	[0.0160]	[0.0491]
Number of waves SC was told a story	0.0385[+]	-0.00197	0.0407*	0.009	0.0606
	[0.0228]	[0.0195]	[0.0170]	[0.0136]	[0.0448]
Family activities index	0.0258[+]	0.026		0.0104	0.0532[+]
	[0.0150]	[0.0253]		[0.0230]	[0.0298]
Indigenous culture, and community					
Participation in Indigenous culture index	-0.0102		0.0501*	-0.0104	0.00933
	[0.0133]		[0.0204]	[0.0225]	[0.0285]
Community bad for kids (1=very good to 5=really bad)	-0.0103	-0.0112	-0.0242	0.0265	-0.0475[+]
	[0.0126]	[0.0185]	[0.0204]	[0.0210]	[0.0260]
No places to play in community (mean value: 1=lots of parks/playgrounds to 5=no, none)	-0.0123	0.0131	0.00701	-0.00179	0.000527
	[0.00977]	[0.0142]	[0.0163]	[0.0159]	[0.0197]
Level of Relative Isolation (reference is low):					
None	0.0432	0.0577	0.042	0.0234	0.0963
	[0.0333]	[0.0441]	[0.0472]	[0.0425]	[0.0602]
Moderate	-0.0225	-0.0709	0.0659	0.0618	-0.0212
	[0.0399]	[0.0496]	[0.0555]	[0.0601]	[0.0755]
Extreme	-0.0541	-0.1	0.0764	0.125	0.00554
	[0.0452]	[0.0688]	[0.0763]	[0.0800]	[0.0997]

	Waves 1-2	Waves 2-3	Waves 3-4	Waves 4-5	All waves[a]
IRISEO (reference is quintile 3):					
Quintile 1	0.0479	-0.0723	0.0764	-0.0815	0.0197
	[0.0429]	[0.0604]	[0.0763]	[0.0636]	[0.0895]
Quintile 2	0.0198	-0.0741+	-0.0409	0.00339	-0.0878
	[0.0331]	[0.0397]	[0.0431]	[0.0439]	[0.0594]
Quintile 4	-0.0376	-0.0113	0.0188	0.0223	-0.0763
	[0.0339]	[0.0522]	[0.0524]	[0.0471]	[0.0678]
Quintile 5	-0.0268	-0.108*	-0.0222	0.0927	-0.0381
	[0.0385]	[0.0501]	[0.0587]	[0.0624]	[0.0735]
Missing income information	0.0485	-0.0561		-0.0542	0.219+
	[0.0550]	[0.0662]		[0.0732]	[0.114]
Income support is the main source	0.0198	-0.0423	-0.0122	0.123*	0.071
of income	[0.0296]	[0.0448]	[0.0504]	[0.0514]	[0.0591]
Number of major life events	-0.00263	0.00175	-0.0250*	0.00336	0.00272
	[0.00748]	[0.0114]	[0.0118]	[0.0111]	[0.0154]
Ever in formal childcare	-0.0119	0.00203	-0.0202	-0.00813	-0.00909
	[0.0281]	[0.0332]	[0.0369]	[0.0341]	[0.0464]
Observations	541	482	483	444	541
Pseudo R^2	0.229	0.213	0.195	0.227	0.11

Notes: (a) The model for the probability of remaining in the sample in all waves was estimated using data from the first wave.

(b) The indicator variable for whether the study child is the oldest child was excluded from the estimation of the models for the probability of attrition between waves 1 and 2, and the probability of remaining in the sample in all waves due to perfect collinearity with the respective endogenous variables of: leaving the sample between Waves 1 and 2, and remaining in the sample for all five waves.

(c) The indicator variables for whether there is a grandparent present in the household and whether the number of books at home is between 21 and 30 were excluded from the estimation in the models for waves 1 and 2, and the model for being in all waves because of perfect collinearity with the respective endogenous variables of: leaving the sample between Waves 1 and 2, and remaining in the sample for all five waves. The dummy variable for the presence of grandparents was excluded from the analysis of attrition between waves 3 and 4 for the same reason. In Wave 3, families were not asked about the number of books in the home.

Standard errors in brackets; +, *, and ** denote significance at the 10, 5, and 1%-level, respectively.

Other variables included in the regression but not reported here as they were not significant in any of the waves are: number of adults in the household, moved house in last 12 months, number of homes since birth, home needs major repairs, number of people per bedroom, community is safe, weekly family income, index of life events, and index of financial stress.

Source: Authors' calculations based on LSIC.

Appendix II

Composite measures

We are interested in analysing how children's outcomes are influenced by the type of family environment children live in, the family's participation in Indigenous culture, the level of financial stress experienced by family members, and the prevalence of life events. We quantify these latent traits using Item Response Theory models.[27] In particular, we estimate a two-parameter logistic model using Stata's *irt2pl* routine. Table A.4 below shows the list of indicators used to derive each of the four composite measures. Results from exploratory factor analyses suggest the existence of an underlying factor for the sets of indicators capturing family activities, participation in Indigenous culture, and financial stress. It is not clear, however, that such an underlying factor exists for the life events indicators.

Playing indoors, playing outdoors, reading a book, and drawing, art or craft are in general the most discriminating items in the family activities composite measure. For the Indigenous culture measure, the most discriminating items appear to be the study child's identification with a tribe and language, the child's connection to the country, and whether the child spends time with Indigenous family members or other Indigenous children. In the case of financial stress, whether families have problems to pay bills and heat their homes are found to be the most discriminating items in most waves. Having been humbugged, mugged robbed or assaulted, and arrested/in jail are the most discriminating items for the life events measure.

Table A.4 Indicators used for the derivation of the composite measures

Composite measure	Wave				
Family activities	*1*	*2*	*3*	*4*	*5*
Music and dance	x	x		x	x
Read a book	x	x	x	x	x
Tell an oral story	x	x	x	x	x
Take shopping	x	x		x	
Housework or cooking	x	x		x	x
Drawing, art or craft	x	x	x	x	x
Play indoors	x	x		x	x
Play outdoors	x	x		x	x
Play computer, Xbox	x	x		x	x

---

[27] For a discussion of IRT models and their application to the estimation of cognitive constructs see Edwards (2009).

Composite measure	Wave				
Have dinner together as a family					x
Have gone to a playground	x				
Have gone swimming	x				
Participation in Indigenous culture	*1*	*2*	*3*	*4*	*5*
How often study child goes to Indigenous cultural event	x			x	
How often you teach study child traditional practices, fishing, hunting	x			x	
How often you teach study child traditional arts, painting, dance	x			x	
Study child: identified with a tribe, language	x			x	
Study child: connection to country	x			x	
Study child: homelands near here	x			x	
Study child: ever visited homelands	x			x	
Parent 1: identifies with mob			x		
Parent 1: identifies with non-Indigenous groups			x		
Study child spends time with Indigenous family member			x		
Study child spends time with Indigenous children			x		
Family experiences racism, discrimination			x		x
Financial stress	*1*	*2*	*3*	*4*	*5*
Could not pay bills on time			x	x	x
Could not pay housing payments on time			x	x	x
Went with meals			x	x	x
Unable to heat or cool home			x	x	x
Pawned/sold something			x	x	x
Assistance from welfare organisation			x	x	x
Unable to send child to preschool/childcare/school				x	x
Child could not do school activities					x
Life events	*1*	*2*	*3*	*4*	*5*
Child hurt/sick	x	x	x	x	x
Family member passed away	x	x	x	x	x
Passed away, someone else who lived there	x	x	x	x	x
Passed away, someone else who doesn't live there	x	x	x	x	x
Worries about money	x	x	x	x	x
Humbugged	x	x	x	x	x
Mugged robbed or assaulted	x	x	x	x	x
Arrested jail, police	x	x	x	x	x
Child upset by family arguments	x	x	x	x	x
Child scared by other people	x	x	x	x	x
Family split up	x	x	x	x	x
Child cared for by someone else	x	x	x	x	x

Author/s:
Azpitarte, F; Chigavazira, A; Kalb, G; Farrant, BM; Perales, F; Zubrick, SR

Title:
Childcare Use and Its Role in Indigenous Child Development: Evidence from the Longitudinal Study of Indigenous Children in Australia

Date:
2019-03-01

Persistent Link:
http://hdl.handle.net/11343/219502

File Description:
Accepted version