The First Five Years: What makes a difference?

5.2 Appendix

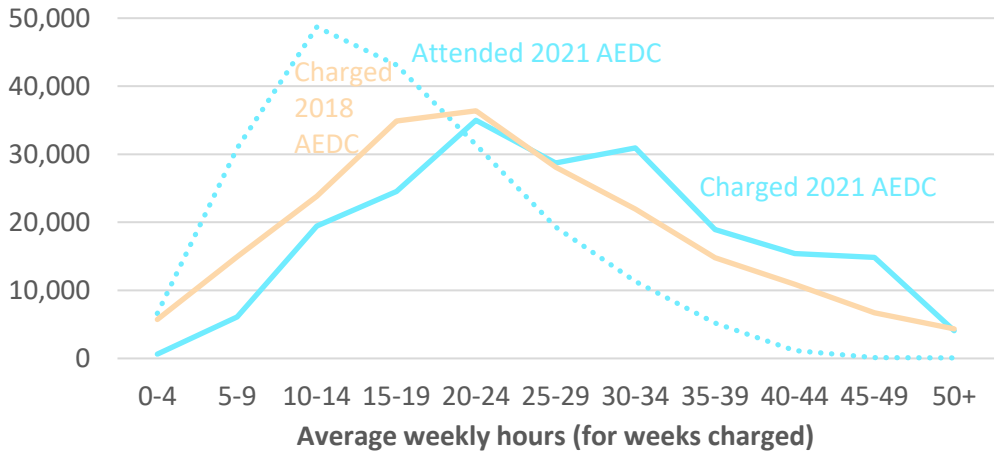
Appendix A: Note on child care charged and attended hours

It is well known that the hours children attend child care are generally less than charged hours. Child care centres typically charge for the whole time the centre is open to cover the cost of staffing and running the centre, while families rarely use the full session. Other reasons such as holidays and illness may also reduce the hours a child attends. Administrative data suggests that children attend for roughly 70 per cent of the hours that families are charged.

For the First Five Years project, the 2018 Australian Early Development Census (AEDC) is linked to the Child Care Management System (CCMS) data, where quarterly charged hours are recorded. Since July 2018, the new Child Care Subsidy (CCS) was introduced, and data on actual attendance at child care became available from 2019. We linked the CCS data with the 2021 AEDC to test the relationship between hours charged and hours attended, and the association with rates of children's developmental vulnerabilities.

The main goal of this subsection is to test the relationship between charged hours and attended hours, and then test whether the reported relationship between charged child care hours and children's development still holds for attended hours.

Figure 1: Number of people by average weekly charged hours and attended hours, 2018 and 2021 AEDC cohort.



Source: Custom Multi-Agency Data Integration Project (MADIP) extract Education, Skills and Employment National Data Asset (ESENDA), including AEDC data linked to CCS data, accessed 2023.

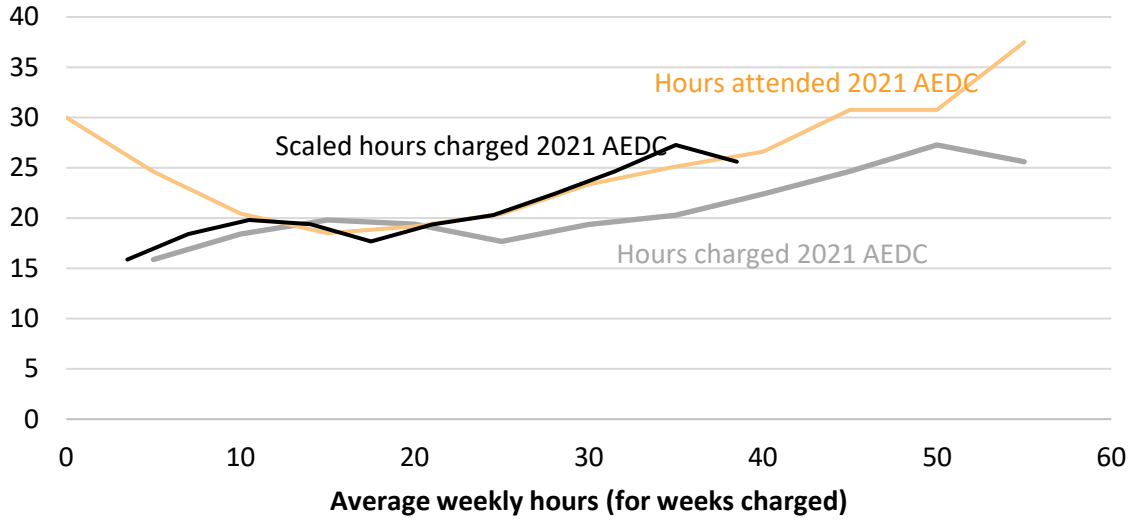
As shown in Figure 1, the distribution of hours families are charged is consistent between 2018 and 2021, with a slight shift towards higher amounts of charged hours in 2021. Most families are charged between 15 to 35 hours per week, equivalent to around 2 to 3 days per week. As expected, the actual hours attended are lower, with the peak at 10-14 hours of attendance per week.

Note that the data only capture attendance at a service eligible for the Child Care Subsidy, and as such does not include fully government-funded preschools.

CCS data contains both charged hours and attended hours, which allows the comparison of child developmental vulnerability for charged and attended hours. In Figure 2, the blue and grey lines show the rate of developmental vulnerability on one or more domains (DV1) based on charged hours and attended hours respectively. Low DV1 rates for the children with low charged hours are observed. However, the DV1 rates for the children attending short or no hours (while paying for them) are high. This low attendance could be due to family and/or health issues. A more detailed study in the future on the comparison of charged hours and attended hours will be useful to identify these extra factors.

We conducted a simple test of the validity of using hours charged as a proxy for hours attended, by comparing the distribution of rates of vulnerability for both measures. We scaled charged hours by 0.7 (in effect, assuming that hours attended represent 70 per cent of hours charged), and compared the resulting distribution of rates of developmental vulnerability to the true distribution based on actual hours attended (blue and black lines in Figure 2). Between 10 and 40 hours, the two curves are very similar, suggesting the charged hours is a reasonable proxy measure of the child care attendance (and developmental vulnerability for those attendance rates) for normal ranges of attendance. However, for the lowest and highest end of the attended hours, the results based on charged hours deviate and care should be taken when dealing with those cases.

Figure 2: Proportions (%) of children who were developmentally vulnerable on one or more domains (DV1), by average weekly charged hours and attended hours, 2021 AEDC cohort.



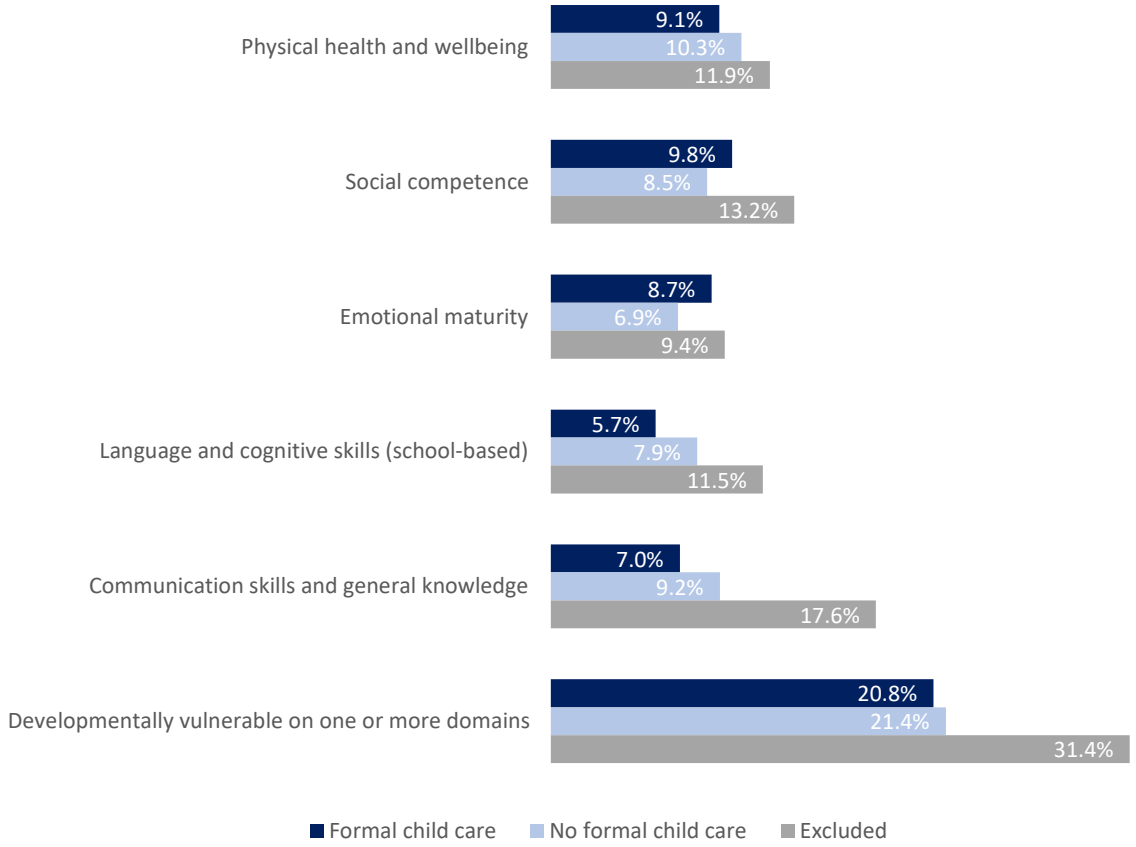
Source: Custom Multi-Agency Data Integration Project (MADIP) extract Education, Skills and Employment National Data Asset (ESENDA), including AEDC data linked to CCS data, accessed 2023.

Appendix B: Linkage bias

As shown in the *Methodology* paper (Table 2), the analytical cohort is smaller than the AEDC cohort due to linkage rates and exclusion of individuals without complete covariates.

For the descriptive analysis of formal child care attendance, we defined child care attendance as unknown if (1) the child could not be linked to the MADIP data, and (2) if the child did not have information of a primary carer. As can be seen from Figure 3, this group of children were more likely to be developmentally vulnerable in all domains compared to the analytical cohort, including both those who did and did not attend child care. Thus, the analytical cohort is biased towards those who are less likely to be developmentally vulnerable.

Figure 3. Proportion (%) of children who are developmentally vulnerable on different domains, for the analytical cohort attending child care and not attending child care and those excluded from the analytical cohort, 2018 AEDC cohort.



Source: Custom Multi-Agency Data Integration Project (MADIP) extract Education, Skills and Employment National Data Asset (ESENDA), including AEDC data linked to CCS data, accessed 2023.

Notes: Analytical cohort with valid DV1 score (including child care and no child care): N=272,616; Rest of cohort N=20,286

For the predictive modelling and statistical inference work, any child with one or more missing values was removed from the analytical cohort, resulting in a further reduced cohort of 211,885. A bias can be estimated for different variables as the difference between the data with no observations removed and the data with all observations with missing data removed. The study of the bias shows that the individuals with missing data were more likely to be developmentally vulnerable and disadvantaged in other ways, including being more likely to live in a low socio-economic area, less likely to have employed parents or parents with a bachelor degree. The results of our predictive modelling and inference work are based on a cohort that is biased towards being representative of less disadvantaged individuals.

Appendix C: Detailed statistical inference results for chapter 3.4

Detailed analysis results such as confidence intervals are available upon request.

Effect of child care duration analysis:

Sample sizes for statistical inference analysis on duration in child care							
	Average weekly hours						
Total hours	0 (control)	<10	10-19	20-29	30-39	40-49	50
0 (control)	55661						
1-2000		13456	25517	5933	877	101	
2001-4000		56	19839	21416	3124	571	1569
4001-6000			2012	25841	7631	1364	2137
6001-8000				6447	15736	2845	80
8001-10000				56	7117	6284	168
10001-12000					397	4780	269
12000						846	563

Appendix D: Individual income rules

We calculated individual incomes each financial year closely following an ABS methodology (ABS 2020; see Methodology for Household income rules). Income data sourced from Personal Income Tax (PIT), Pay As You Go summaries (PAYG) and Data Over Multiple Individual Occurrences (DOMINO).

Employee/Exertion:

- Salary or wages (PIT) +
- Allowances earnings tips directors' fees etc (PIT) +
- Employment termination payments taxable component (PIT) +
- Attributed personal services income (PIT) +
- Total reportable fringe benefits amount (PIT) +
- Reportable employer superannuation contributions (PIT) +
- Other net foreign employment income (PIT) +
- Foreign source income exempt foreign employment income (PIT) +
- Lump sum A (PAYG) +
- Lump sum B (PAYG) +
- Lump sum D (PAYG) +
- Lump sum E (PAYG)

When PIT data is unavailable or zero, use PAYG only definition:

- Lump sum A (PAYG) +
- Lump sum B (PAYG) +
- Lump sum D (PAYG) +
- Lump sum E (PAYG) +
- Total allowances (PAYG) +
- Gross Payment Amount (PAYG) +
- Exempt foreign employment income (PAYG)

Business:

- Net income or loss from business - primary production (PIT) +
- Net income or loss from business - non-primary production (PIT) +
- Distribution from trusts primary production (PIT) +
- Net PSI (PIT) +
- Distribution from partnerships less foreign income non primary production (PIT) +
- Distribution from partnerships primary production (PIT)

Investment:

- Gross interest (PIT) +
- Dividends unfranked (PIT) +
- Dividends franked (PIT) +
- Dividends franking credit (PIT) +
- Share of net income from trusts less net capital gains and foreign income non primary production (PIT) +
- Franked distributions from trusts non primary production (PIT) +
- Foreign source income Australian franking credits from a NZ company (PIT) +
- Foreign source income net foreign rent (PIT) +
- Rent net rent (PIT)

Superannuation:

- Australian annuities and superannuation income streams taxable component taxed element (PIT) +
- Australian annuities and superannuation income streams taxable component untaxed element (PIT) +
- Australian annuities and superannuation income streams lump sum in arrears taxable component taxed element (PIT) +

- Australian annuities and superannuation income streams lump sum in arrears taxable component untaxed element (PIT) +
- Australian superannuation lump sum payments taxed element (PIT) +
- Australian superannuation lump sum payments untaxed element (PIT) +
- Bonuses from life insurance companies and friendly societies (PIT)

Other:

- Foreign source income net foreign pension or annuity without UPP (PIT) +
- Foreign source income net foreign pension or annuity with UPP (PIT) +
- Foreign entities controlled foreign company income (PIT) +
- Foreign entities Transferor trust income (PIT) +
- Foreign source income other net foreign source income (PIT) +
- Other income category 1 (PIT) +
- Other income category 2 (PIT)

Welfare:

- Australian Government pensions and allowances (PIT)

When PIT data is unavailable or zero, use DOMINO based derivation for total welfare payments:

- Component amount paid (daily), multiplied by days received per financial year. All payment types summed. (DOMINO)

Tax:

- Net tax (PIT) +
- Medicare levy (PIT) +
- Medicare levy surcharge (PIT) -
- Medical expenses offset available

When PIT data is unavailable or zero, use PAYG only definition:

- Tax Withheld Amount (PAYG)

Gross:

- Employee/Exertion (derived) +
- Business (derived) +
- Investment (derived) +
- Superannuation (derived) +
- Welfare (derived)

When PIT data is unavailable or zero, use PAYG and DOMINO based definitions:

- Employee/Exertion (derived) +
- Reportable fringe benefits amount (PAYG) +
- Tax-free component (PAYG) +
- Taxable component (PAYG) +
- Welfare (derived)

Disposable:

- Gross (derived) -
- Total deductions (PIT) -
- Tax (derived)

When PIT data is unavailable or zero, use PAYG and DOMINO based definitions:

- Gross (derived) -
- Tax (derived)

Spousal gross:

- Your spouse's taxable income (PIT) +
- Amount of Australian Government pensions and allowances that your spouse received (PIT) +
- Amount of exempt pension income your spouse received (PIT) +
- Amount of your spouse's reportable superannuation contributions (PIT) +
- Your spouse's target foreign income (PIT) +
- Your spouse's total net investment loss (PIT) +
- Your spouse's total reportable fringe benefits amounts

References

ABS (Australian Bureau of Statistics) (2020) *Personal Income in Australia methodology, 2011—12 to 2017—18*, ABS website, accessed 23 February 2021.