BANKWEST CURTIN ECONOMICS CENTRE

WORKING PAPER SERIES

20/02: HOUSEHOLD DENSITY AND CHILDREN'S WELLBEING IN AUSTRALIA: ARE CHILDREN'S HOMES TOO EMPTY?

Michael Dockery

Bankwest Curtin Economics Centre Working Paper Series
ISSN: 2202-2791
ISBN: 978-1-925757-09-5

Suggested Citation
Michael Dockery "Household density and children's wellbeing in Australia: Are children's homes too empty " Bankwest Curtin Economics Centre Working Paper 20/02, Perth: Curtin University.

**Household density and children's wellbeing in Australia:
Are children's homes too empty?**

AM. Dockery
Principal Research Fellow
Bankwest Curtin Economics Centre

Abstract: Living in crowded housing has been associated with poorer physical and mental health outcomes, with a greater emphasis on the psychological effects when the underlying quality of housing is high. Hence 'overcrowding' features as a housing condition extensively studied in housing economics. This paper looks at physical health, mental health and cognitive outcomes for Australian children and adolescents conditional upon measures of occupant density of their family homes using three longitudinal datasets: the Household, Income and Labour Dynamics in Australia Survey; the Longitudinal Study of Australian Children; and the Longitudinal Study of Indigenous Children. Overall, measures of occupant density – including some commonly used as measures of 'crowding' – are found to be positively associated with children's wellbeing. It is argued that Australian children face a greater problem of homes being too empty, rather than too crowded, for optimal physical health and social and emotional development. Surprisingly, the lack of evidence of negative effects of occupant density on health and socio-emotional wellbeing extends to Indigenous children who, on average, live in housing with markedly higher household density levels. There is, however, some evidence of higher density detracting from educational outcomes for Indigenous children. The results should not be taken to dismiss problems of extreme household density levels experienced by some Australian children, and notably by Indigenous children, but do point to a need to reconsider what we call 'crowding' and how it is measured.

Acknowledgements:

I would like to acknowledge the contribution of Ha Nguyen to earlier research that ultimately led to this paper, and two referees who provided valuable comments on an earlier draft.

This report uses:

- Unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA project was initiated and is funded by the Australian Government Department of Social Services (DSS) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute).
- Unit record data from *Growing Up in Australia:* The Longitudinal Study of Australian Children (LSAC). LSAC is conducted in partnership between the Australian Government Department of Social Services, the Australian Institute of Family Studies (AIFS) and the Australian Bureau of Statistics (ABS), with advice provided by a consortium of leading researchers from research institutions and universities throughout Australia.
- Unit record data from Footprints in Time—the Longitudinal Study of Indigenous Children (LSIC). LSIC was initiated and is funded by the Australian Government Department of Social Services.

The findings and views reported in this paper, however, are those of the author and should not be attributed to DSS, the Melbourne Institute, AIFS or the ABS.

1.0 Introduction

In the literature relating housing circumstances to child development, there has been a long-standing concern with the negative effects of 'density' or 'crowding' upon outcomes spanning the domains of physical health, psychological wellbeing and cognitive development. Living in crowded conditions has been associated with the spread of communicable diseases such as meningococcal meningitis in New Zealand (Baker et al. 2000); a higher incidence of respiratory illness in UK children (Mann et al. 1992); lower parental assessed general health for US children (Solari and Mare 2012) and higher mortality rates in later life, particularly associated with heart disease, for children from Chesterfield, England (Coggon et al. 1993). Crowding at home has been associated with poorer psychological health, less persistence in problem solving and poorer school behaviours (Evans, Lercher and Koffler 2002; Evans, Saegert and Harris 2001, Solari and Mare 2012), plus poorer cognitive development and schooling outcomes such as IQ, memory, language development and reading proficiency.[1] The effects of crowding on parenting styles, parent-child conflict and access to quiet spaces for homework have been identified as potential causal links (Bartlett 1998; Evans, Maxwell and Hart 1999; Evans et al. 2001, Wachs & Camli 1991).

Despite the international body of empirical studies identifying negative effects of crowding on children's outcomes, Solari and Mare argue the literature remains ambiguous regarding whether or not crowding has an effect on child wellbeing net of other socio-economic factors. They attribute the ambiguity to inconsistency in crowding measures used in the literature, and inadequate controls for family socio-economic background (2012: 465–466). The evidence is even less clear for Australian children, who have been the subject of relatively few empirical studies of the impact of living in crowded housing. The commonly used term in Australia of 'overcrowding' carries with it the connotation of negative effects of household density beyond some threshold but, as noted by Memmott et al. (2012), is inherently tautological. Much of the Australian overcrowding literature has focussed on Aboriginal and Torres Strait Islander persons, for whom inferior housing has been identified as contributing to poorer relative health and schooling outcomes relative to non-Indigenous Australians (AIHW 2014; Biddle 2008; Dockery et al. 2013).

In the empirical literature on crowding and children's outcomes, measures of crowding are typically constructed on the basis of occupant density, such as persons per dwelling or persons per bedroom, including the commonly used Canadian National Occupancy Standard (CNOS). The CNOS sets out the number of bedrooms required in a household based on the number, age, sex and relationships between occupants and assumptions on who can 'reasonably' be expected to share a bedroom (see Appendix 1). As far back as 1976, Amos Rapoport noted the conceptual misalignment between density and crowding: "Density can be seen as a measure of people per unit area, and crowding as a negative perception of excessive density – a subjective experience of sensory and social overload." (1976: 8). Recognising that subjective experience is integral to any notion of crowding implies the need to take into account factors that may lead to differences in the way density may affect occupants' experiences of density, such as cultural norms and the relationships between the occupants. It also invokes the possibility that "… [T]he negative subjective experience may also be of excessively low interaction, i.e., of too low densities which we may call isolation for want of a better term." (Rapoport 1976: 8). Issues of cultural norms and kinship relationships may be particularly important in creating differences in how measures of occupancy density translate into subjective feelings of crowding for Indigenous and non-Indigenous Australians (Memott et al. 2012).

[1] For a more detailed review and further references, see Dockery et al. 2010, pp. 13-15.

In the context of these debates, this paper presents new empirical estimates of the associations between measures of household density and the physical health, social and emotional wellbeing and cognitive outcomes of Australian children using three major longitudinal datasets. If anything, and contrary to expectations, measures of occupant density are found to be positively associated with children's wellbeing. I argue the results challenge the appropriateness of existing measures of crowding and suggest that, in contemporary Australia, children often face a problem of social isolation within their own homes.

Before proceeding, a note on terminology is warranted. The empirical analyses in this paper uses measures based on the ratio of the number of occupants to bedrooms in children's homes, and the number of extra bedrooms required under varying assumptions of who might share a bedroom. These will be referred to as measures of household density or occupant density (or just density), rather than crowding. This is because of the misalignment, noted above, between measures of density and 'crowding', and because it is almost nonsensical to speak of 'crowding' having positive effects when the term simultaneously implies a negative effect. However, I do use the term 'crowding' or 'overcrowding' in reviewing literature that also uses those terms, and hence purports to relate to issues of excessive household density levels. Further, note that density measures based on the number of additional bedrooms required, such as the CNOS, are quite explicitly intended as measures of crowding, albeit with a somewhat arbitrary standard imposed as to what constitutes an adequate number of bedrooms given the household composition.

2.0 Existing Australian research

There have been very few Australian empirical studies of the effects of household density for either adult or child occupants, and the available evidence is quite ambiguous as to whether any effect is detrimental or beneficial. Dockery et al. (2013) looked at children's physical health outcomes, social and emotional wellbeing and cognitive outcomes using data from the initial three waves (2004-2008) of the Longitudinal Study of Australian Children (LSAC) and two waves (2008-2009) of the Longitudinal Study of Indigenous Children (LSIC). Multivariate, random-effects panel models were estimated which included density as an independent variable along with a wide range of other controls relating to housing and socio-economic background. Density was measured as the ratio of the number of occupants to bedrooms in the child's dwelling, assuming parents in a couple household share one bedroom. In the models based on LSAC data, no significant association between density and physical health or social and emotional wellbeing was observed. However, a significant and negative association was observed for learning outcomes.

For Indigenous children, evidence of occupant density effects in models based on LSIC data was ambiguous, with coefficients implying a higher ratio of persons per bedroom to be associated with inferior parental assessment of the child's general health, but also fewer reported concerns with social and emotional behaviour (with weak significance, $p \approx 0.10$ in both cases). Density was not significant in a model for learning concerns (Dockery et al. 2013: 45-50). Analyses of the Western Australian Aboriginal Child Health Survey also found that Aboriginal children living in houses with a high occupancy level displayed fewer emotional and behavioural difficulties compared to those living in low occupancy households (Zubrick et al. 2005: 129).

Booth and Carroll (2005) investigate the effect of 'overcrowding' on adult health using data from the 2001 Australian Bureau of Statistics' (ABS) National Health Survey, which included an expanded Indigenous sample. They measure crowding as the ratio $(N_{persons})/(N_{bedrooms} + 1)$, with the '+1' in the denominator included to account for bedsits with zero bedrooms, and look at both self-assessed general health and whether the respondent had been hospitalised in the past 12 months.

Defining households as ‘overcrowded’ if the above ratio is greater than 2, their simple bivariate statistical analysis found that persons living in overcrowded houses in fact report better health than those living in non-crowded housing. This association held in a multivariate model with controls for demographics and socio-economic status, but the estimate was not significantly different from zero. In an alternative specification, Booth and Carroll included separate control variables for the number of persons and number of bedrooms. They found evidence of positive effects on health and fewer hospitalisations associated with more bedrooms for a given number of persons, although one reported model also found a significant, positive association between adult health and the number of children in the house for a given number of bedrooms. On this basis, Booth and Carroll argue the more flexible specification of crowding, allowing the number of rooms, adults and children to enter separately, is preferable to a dichotomous classification of ‘overcrowded’ (2005: 3).

Booth and Carroll (2005) inferred crowding to be an issue for Indigenous households on the basis of relatively minor changes in the estimated coefficient for the Indigenous household dummy between models estimated with and without the ‘crowding’ variables. However, this assumes any density effect is the same for the two samples: models were not estimated separately for Indigenous and non-Indigenous households. Along with their datasets’ topcoding of variables for the number of occupants and number of bedrooms, a major limitation of the study was the lack of other controls relating to housing. Notably, it is well known that tenure is important, with homeowners displaying superior outcomes relative to renters, and particularly relative to renters in public or community housing (Dockery and Bawa 2019, Dockery et al. 2010, 2013). Measures of density are higher for private renters compared to homeowners, and higher again for those in public and community housing (Dockery et al. 2013, Waters 2001b).

Using data from the earlier 1995 ABS National Health Survey, Waters (2001a, b) defined households as being crowded if they did not meet the CNOS. Multivariate analyses found that adults living in crowded households reported a significantly lower number of health conditions and fewer doctor consultations than those living in non-crowded households. Those living in crowded households were estimated to have significantly lower self-assessed health, but that result was not robust to controlling for sample clustering by area. In summarising her findings relating to overcrowding, Waters warns “Despite the fact that the current study found virtually no evidence of an adverse effect of overcrowding on health in Australia, it cannot be assumed that such effect does not exist.” (2001a: 23). However, as with the analysis by Booth and Carroll (2005), controls for housing tenure were not included in Waters’ models of the effect of crowding on health. Given the bivariate statistics presented, it seems clear their inclusion would have strengthened the evidence of a positive association between ‘crowding’ and health.

Despite this, at best, flimsy evidence-base, acceptance of a substantial problem of household crowding (or overcrowding) appears to be widespread in Australian housing policy discourse. In the 2019 issue of the Australian Institute of Health and Welfare’s (AIHW’s) *Housing Assistance in Australia* report, the incidence of overcrowding is estimated at 4 per cent of households in public housing, 24 per cent in state owned and managed Indigenous housing (SOMIH) and 4 per cent in community housing based on the CNOS. While limitations of the CNOS are acknowledged, the report states: ‘Regardless of the appropriateness of the measure, overcrowding based on CNOS has been found to adversely affect the physical and mental health of residents (AIHW 2014, Booth & Carroll 2005, SCRGSP 2016)’. This is the same Booth and Carroll (2005) discussed above, and which did not in fact use the CNOS. The other two references provided in support of the claim of adverse effects relate only to Indigenous households. Based on 2011 Census data, AIHW (2014) notes that Indigenous households are 3 times more likely to be overcrowded than non-Indigenous households, and reports how the incidence of crowding increases with remoteness. However, only descriptive data on the incidence and trends in crowding are presented, not empirical evidence relating 'crowding' to outcomes.

The referenced 'SCRGSP 2016' is the most recent of the regular *Overcoming Indigenous Disadvantage* reports produced by the Steering Committee for the Review of Government Service Provision. The Overcoming Indigenous Disadvantage reporting framework then included seven 'strategic areas for action', and under each of these a number of 'strategic change indicators' are identified. The home environment is one of the strategic areas for action, and 'overcrowding in housing' the first of its three indicators, stating 'reducing overcrowding is associated with positive effects on health, family relationships and children's education' (2016: 10.1). SCRSGP also notes that 'reducing overcrowding, particularly in remote areas and discrete communities' was a key outcome agreed to by the Council of Australian Governments (COAG) for the 2009 National Affordable Housing Agreement (2016: 10.3). A number of review papers are cited, but little by way of empirical Australian evidence of a negative effect of crowding on outcomes. Dockery et al. (2013), discussed above, is noted as providing evidence that 'learning and cognitive development of all Australian children has been shown to be negatively affected by crowding'. Studies are also noted which suggest there may be benefits from additional household occupants by way of greater connectivity (SCRGSP 2016: 10.4).

The background 'Research Agenda' document for the Australian Housing and Urban Research Institute's (AHURI's) 2020 National Housing Research Program states: 'Addressing overcrowding is a central objective of housing policy in Australia, yet, defining and measuring overcrowding is problematic. Overcrowding has been identified as one of the key indicators in need of improvement under the NHHA's data improvement plan.' (AHURI 2019: 19)[2]. All seven references listed in relation to that overcrowding topic are studies focussing specifically on Indigenous Australians. The Western Australian Commissioner for Children and Young People's (2014) report on the wellbeing of children aged 0 to 17 in Western Australia included 'Overcrowded households' as one of the key indicators of 'material wellbeing'. The report indicates the measure is included because:

> "Overcrowded housing has been associated with potential negative impacts on children's wellbeing and development which can have significant and far reaching effects. While more Australian research into the effects on children is required, possible impacts of overcrowding cited in the literature here and abroad include an increased risk of infection-based illness, irregular sleep, poor school performance, increased parental stress resulting in punitive parenting and parent-child conflict, greater vulnerability to abuse and poor mental health." (CCYP 2014: 211)

No original empirical Australian evidence is cited in support of this statement, in line with the noted need for more research.

In developing a small area index of housing disadvantage for Australian children, McNamara et al. (2010) undertake a principal component analysis of variables relating to overcrowding, public housing, dwelling structure and housing affordability stress. While they cite international literature linking crowding to poorer outcomes for children, they also note uncertainty as to whether such findings are applicable in the Australian setting, including reference to Waters' (2001a, b) null findings discussed above.[3] The resulting index classifies 8.8 per cent of Australian children as living in overcrowding, including 31.3 per cent of children in the Northern Territory. This is attributed to the high proportion of Indigenous Australians in the Northern Territory population, drawing a caution

[2] NHHA is the National Housing and Homelessness Agreement.
[3] To be precise, they refer to Waters (2002), which provides a policy summary of the 2001a and 2001b reports.

that the NCOS-based overcrowding measure may have limited cultural applicability. However, they observe that '… overcrowding is a very real problem for Indigenous Australians, noting that it intensifies problems such as sexual abuse and domestic violence (Cooper & Morris, 2005, pp. ii–iii).' (McNamara et al. 2010: 639). The cited Cooper and Morris (2005) is a large qualitative study of the supports needed for Indigenous tenancies, based on interviews with over 100 Indigenous women. McNamara et al. are appropriately citing 'key findings' given in Cooper and Morris' executive summary:

> "Overcrowding is a huge problem, exacerbating domestic and family violence, sexual abuse and property damage, often leading to women and/or children leaving 'home', or eviction." (Cooper and Morris 2005: p. ii.)

However, in the section on overcrowding in the body of the report (pp. 25-27) there are numerous reports by participants of having experienced crowding, but none of any actual negative impacts attributed directly to crowding. The subsection entitled 'Impact of overcrowding' in fact states:

> "There were very few comments about the impact of overcrowding on health, especially communicable diseases, children's education, personal safety, or the increased likelihood of sexual abuse. This may suggest a lack of awareness of the range of consequences. In the case of sexual abuse, there is possibly shame and stigma associated with disclosure." (Cooper and Morris 2005: 26)

This apparent misalignment between Cooper and Morris' strongly worded conclusion and the reported evidence from their own study seems symptomatic of a wider phenomenon in Australian housing policy and research, whereby the spectre of a major problem of overcrowding has taken on a life of its own, unfettered by any need to be sustained by evidence.

3.0 Contemporary evidence for Australian children

Data from three separate surveys are utilised to investigate the impact of crowding on children's outcomes in Australia: the Longitudinal Study of Australian Children (LSAC), Longitudinal Study of Indigenous Children (LSIC) and the Household, Income and Labour Dynamics in Australia survey (HILDA). The differing sampling frames for these datasets permit a focus on young Australian children (LSAC), young Indigenous children (LSIC) and adolescent/young adult Australians (HILDA). Each survey collects relatively rich data on housing circumstances, plus a variety of measures of physical health, social and emotional wellbeing and cognitive or educational outcomes. Using each dataset, the associations between a range of measures of household density and outcomes in each of those domains are estimated. This section first provides an overview of the three survey datasets and the constructs of household density to be used, and then reports estimates from multivariate models of the effect of household density on outcomes.

3.1 The data

The initial samples for both LSAC and LSIC comprise of children from a Baby cohort (aged 0 to 18 months) and a Kindergarten cohort (aged 4-5 years), which are tracked with follow-up surveys every year in the case of LSIC, and every second year for LSAC. HILDA is a household panel survey, in which all household members aged 15 and above are interviewed. With 18 annual waves of HILDA data now available, it is possible to analyse self-reported outcomes for youth aged 15 and over conditional upon their housing circumstances in childhood.

Also referred to as Growing up in Australia, LSAC is managed jointly by the Department of Social Services (DSS), the Australian Institute of Family Studies (AIFS), and the Australian Bureau of

Statistics (ABS). Initiated in 2004, it is Australia's first nationally representative longitudinal study of child development, with the aim of facilitating '…a comprehensive understanding of [children's] development and life-course trajectories within Australia's current social, economic and cultural environment' (AIFS 2017: 1). Seven waves of data were available at the time of this analyses. Table 1 provides a snapshot of the responding sample by cohort and wave.

Table 1: LSAC responding sample by cohort, wave and children's age

Year	**2004**	**2006**	**2008**	**2010**	**2012**	**2014**	**2016**
Wave	1	2	3	4	5	6	7
Baby 'B' cohort							
Age range (years)	0-1	2-3	4-5	6-8	8-9	10-11	12-14
Responding sample	5,107	4,606	4,386	4,242	4,085	3,764	3,381
Kindergarten 'K' cohort							
Age range (years)	4-5	5-7	8-9	10-11	12-13	14-15	15-18
Responding sample	4,983	4,464	4,331	4,169	3,956	3,537	3,089

Notes: Responding sample numbers based on frequency for the cohort variable.

The LSIC, also known as "Footprints in Time", was initiated by the Australian Government with the aim of providing insights into how a child's early years affect their development and on strategies to help close the gap in life circumstances between Indigenous and non-Indigenous Australians (Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) 2009; DSS 2017).

The sample was drawn from 11 different sites around Australia designed to 'cover the range of socio-economic and community environments where Aboriginal and Torres Strait Islander children live' and provide roughly equal representation of urban, regional and remote areas, among other criteria. Wave 1 interviews were conducted from April 2008 to February 2009 for around 150 children from each site, or 1,650 in total. FaHCSIA estimated that this represented around 6 per cent of the total Aboriginal and Torres Strait Islander population in each cohort (2009: 12). The availability of such data is important for researchers because Aboriginal and Torres Strait Islander people represent only around three per cent of the Australian population, meaning that population-representative surveys typically contain too few observations for many statistical purposes. Table 2 reports the sample numbers for the 9 waves of LSIC available at the time of writing.

Table 2: LSIC responding sample by cohort, wave and children's age

Year	2008	2009	2010	2011	2012	2013	2014	2015	2016
Wave	1	2	3	4	5	6	7	8	9
Baby cohort Age (years)	$^1/_2$–2	1$^1/_2$–3	2$^1/_2$–4	3$^1/_2$–5	4$^1/_2$–6	5$^1/_2$–7	6$^1/_2$–8	7$^1/_2$–9	8$^1/_2$–10
Resp. sample	954	868	813	749	728	737	734	756	754
Child cohort Age (years)	3$^1/_2$–5	4$^1/_2$–6	5$^1/_2$–7	6$^1/_2$–8	7$^1/_2$–9	8$^1/_2$–10	9$^1/_2$–11	10$^1/_2$–12	11$^1/_2$–13
Resp. sample	717	655	591	534	530	502	519	499	514

Note: responding sample numbers based on frequencies for the 'cohort' variable.

Initiated in 2001, HILDA is an annual panel survey of individuals from a representative sample of private households (Watson and Wooden 2010). Around 13,000 individuals from over 7,000 households have responded in each year, with year-on-year attrition rates averaging below 10 per cent. In 2011 an additional top-up sample of 2,153 households encompassing 4,009 responding individuals was recruited to the survey sample (Wilkins 2014).[4] There are three main survey instruments used in collecting the data: a household questionnaire that is completed for the whole household by one responding adult; an individual 'person questionnaire' that is completed by interview and collects detailed information relating to the responding individual; and a self-completion questionnaire that the responding individual fills out independently and returns to the researchers.

Although individuals within a HILDA household are only interviewed if they are aged 15 years or older, the household questionnaire collects some information on all household occupants, including basic demographics such as their gender, date of birth and relationship to other household members. In the terminology of the HILDA survey, people who are recorded as household occupants but not interviewed directly are referred to as 'enumerated persons', and assigned a unique identifying number to allow them to be tracked over time and included in analyses. So for children aged less than 15, information collected on their household is known (such as housing tenure and geographical variables), and the relationship grid enables them to be matched to information given by their parents as responding persons within the household.

3.1.1 Household density measures

Using each dataset, multivariate models were estimated with selected outcomes in the domains of physical health, social and emotional wellbeing and cognitive development/educational achievement as the dependent variables. Four measures of household density were explored. These were:

- Occupants per bedroom (*Density1*) – the straight ratio of the number of usual household occupants to the number of bedrooms. The number of bedrooms was set to 1 where no bedrooms were recorded, effectively classifying bedsits as having one bedroom.
- Occupants per bedroom, couples share (*Density2*) – as with *Density1*, but reducing the number of occupants by 1 for couple households, on the assumption that the parents will share a bedroom. Hence, *Density1* and *Density2* are identical for sole parent households.
- Extra bedrooms required (*Bedreq1*) – the number of bedrooms required is calculated as number of household occupants minus 1 if the parent/carer is partnered. If this exceeds the actual number of bedrooms, then *Bedreq1* is set at required bedrooms minus actual bedrooms, and 0 otherwise.
- Extra bedrooms required, kids share (*Bedreq2*) – As with *Bedreq1*, the parents in a couple household are assumed to share a bedroom. Two children aged 15 and under are also assumed to be able to share a bedroom, and all other household occupants are assumed to require their own bedroom. If the resulting 'required' number of bedrooms exceeds the actual number of bedrooms, then *Bedreq2* is set at required bedrooms minus actual bedrooms, and 0 otherwise. This cannot be calculated with LSAC, as only the total number of occupants is recorded, not the number of adults and children. *Bedreq2* more closely matches the CNOS, but the data do not distinguish the number of children by gender. Hence, unlike the CNOS, the measure does not impose the standard that children aged 5 to 18 of the opposite gender should not share a bedroom.

[4] See http://melbourneinstitute.unimelb.edu.au/hilda for further details on the HILDA survey.

In the first wave of LSAC, the number of bedrooms is topcoded to 5 for households with 5 or more bedrooms, while in subsequent years the actual number was recorded. Given the small number of dwellings with more than 5 bedrooms (less than 3 per cent), the *Density1*, *Density2* and *Bedreq1* variables are calculated using the topcoded value to enable inclusion of the Wave 1 data in the analyses. In Wave 4, data on the number of bedrooms is collected only if the family had moved since the last interview. For those who had not moved, the number of bedrooms reported in Wave 3 is assumed to apply.

For LSIC, the number of bedrooms is collected for all respondents in Waves 1 and Waves 8, and in other waves asked only if respondents had moved house. The previous recorded number of bedrooms is assumed for all respondents who had not moved house. The number of bedrooms is collected in HILDA in all waves. LSIC and HILDA also record the total number of people and the total number of children aged 15 and under in each wave.

Table 3 presents descriptive statistics for these measures of household density from the pooled samples for each dataset. By definition, all observations for LSAC and LSIC relate to households with children, since at least the study child must be present. For HILDA, the figures are calculated with children aged under 15 as the unit of analysis, rather than households. So a household with, for example, two children aged under 15 in a particular wave will contribute two observations.

Table 3: Descriptive statistics for household density measures: LSAC, LSIC and HILDA

	Mean	Median	Mode	Min	Max	Pooled obs.
Density1						
LSAC	1.30	1.25	1.00	0.40	8.00	51,396
LSIC	1.64	1.50	1.00	0.33	12.00	11,781
HILDA	1.29	1.25	1.00	0.25	6.00	77,168
Density2						
LSAC	1.06	1.00	1.00	0.40	7.00	51,396
LSIC	1.46	1.33	1.00	0.33	11.00	11,781
HILDA	1.03	1.00	1.00	0.17	6.00	75,051
Bedreq1						
LSAC	0.47	0.00	0.00	0.00	9.00	51,396
LSIC	1.53	1.00	0.00	0.00	16.00	11,781
HILDA	0.40	0.00	0.00	0.00	9.00	75,051
Bedreq2						
LSAC	—	—	—	—	—	—
LSIC	0.59	0.00	0.00	0.00	16.00	11,781
HILDA	0.08	0.00	0.00	0.00	5.00	75,051

Notes: estimates calculated for LSAC and HILDA using relevant person weights. Weights are not provided for LSIC. The sample for HILDA excludes group or multi-family households.

From Table 3, it can be seen that most Australian children live in housing with a sufficient number of bedrooms, even allowing for every child to have their own room (median for *Bedreq1*=0 for LSAC and HILDA). On this rather generous measure, the LSAC data indicate that 70 per cent of children lived in housing requiring no additional bedrooms, with a similar figure of 73 per cent derived from HILDA. Results from LSIC indicate that 36 per cent of Indigenous children live in houses with enough bedrooms for all children to have their own room.

Allowing for two children under the age of 15 to share a bedroom (*Bedreq2*), the results from HILDA indicate that 94 per cent of children live in housing with an adequate number of rooms, and from LSIC that 69 per cent of Indigenous children do so. In effect, few Australian children live in housing that could be considered 'overcrowded'. Even among Indigenous households, the vast bulk of children do not live in crowded housing, although clearly residential density levels are higher among this population and there are incidences of quite extreme crowding by any standard (see the maximum values reported in Table 3).

3.2 Multivariate estimates - LSAC

Panel models were estimated for selected indicators of the child's physical health, social and emotional development and educational outcomes as dependent variables. Physical health is measured using the responding parent's (or carer's) assessment of the child's general health on a scale of poor, fair, good, very good or excellent. These ratings were collected in all waves of the survey.

Social and emotional development is measured using scores from the Strengths and Difficulties Questionnaire (SDQ). The SDQ is a brief instrument to assess psychological adjustment of children and youths aged 3 to 16 and designed so that near identical versions can be completed by parents and teachers (Goodman 2001). The SDQ instruments have been incorporated into the LSAC surveys. Parents completed the SDQ for all children from age 4-5 years (i.e. from Wave 1 for the Kindergarten cohort and Wave 3 for the Baby cohort). Teachers completed the SDQ for the study children from age 4-5 years to age 14-15 (Waves 1-6 for the Kindergarten cohort and, to date, Waves 3-7 for the Baby cohort).

The SDQ instrument contains 25 items relating to positive and negative aspects of psychological adjustment grouped into five sub-scales: emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and pro-social behaviour. The LSAC data contain a derived 'total difficulties score', which is derived by summing the values on 20 of the 25 items (the pro-social behaviour scale is excluded), with the responses assigned numerical values of 0, 1, or 2, respectively, where the respondent indicates 'no', 'sometimes', or 'yes' in respect to whether the child displayed particular behaviours in the past 6 months. The scores based on the parent-assessed SDQ were categorised into three ranges of normal (0-13 on the total difficulties score), borderline (14-16) and abnormal (17-40); and the teacher-assessed SDQ as normal (0-11), borderline (12-15) and abnormal (16-40), consistent with cut-points recommended by Goodman (1997).[5]

Educational outcomes are modelled using parent and teacher assessments of the child's 'overall school achievement' provided on a 5-point scale, ranging from well below average, below average, average, above average and excellent. This is collected from parents for children aged from 6-7 years to 14-15 years. Teacher assessments were collected for the Kindergarten cohort at ages 8-9 years and 10-11 years, and for the Baby cohort at ages 6-7 years, 8-9 years and 10-11 years.

With each of the outcome measures represented by an ordered, discrete dependent variable, ordered probit models were estimated to model progression from poor assessed health to better assessed health, from the abnormal range to borderline to normal range for the total difficulties scores, and from well below average school performance to excellent. Hence, in each model positive coefficients for a covariate imply 'better' outcomes for the child as that variable increases. More specifically, Stata's XTOPROBIT command was used to estimate random-effects panel models to take

[5] Teachers typically identify fewer behavioural problems than parents.

account of the fact that we have repeat observations on children. While fixed-effects models would control more rigorously for unobserved child-specific effects, no fixed-effects version of the ordered probit model is available.

Controls are included for the child's age, cohort, gender, non-English speaking background and Indigeneity, and for sole parent and home-ownership status. Variables included to capture background socio-economic status of the family are parental-assessed financial prosperity, a derived 'socio-economic position index', the ABS socio-economic index of disadvantage for the neighbourhood in which the family lives, and a parent-assessed neighbourhood 'liveability' scale (see Appendix Table A2.1 for variable definitions). Full results for models using the *Density2* variable are presented in Appendix Table A1. *Density2*, the ratio of occupants to bedrooms allowing for couples to share one bedroom, is chosen for reporting as the estimates for *Density2* were the most robust. Table 4 reports results for the various specifications of the household density variables from all 15 models (5 outcome measures and three different specifications of density).

For parent-assessed general health of the child and parent-assessed school achievement, there are negative associations with the household density variables after controlling for basic child demographic characteristics and family socio-economic background. However, none of these estimates are close to accepted levels of statistical significance, meaning we cannot reject the hypothesis of no associations between density and the outcomes observed. For the SDQ measures and teacher-assessed overall school achievement, the coefficients are all positive, indicating better outcomes for children living in housing with higher density. Moreover, these density estimates are significant in a number of models for teacher assessment at the five per cent level, and at the one per cent level in the case of the estimated effect of *Bedreq1* for teacher-assessed school achievement.

Table 4: Summary of results for density variables, LSAC models

	Density1			Density2			Bedreq1		
	β		P>|z|	β		P>|z|	β		P>|z|
General health [1-5]	-0.019		0.321	-0.014		0.528	-0.002		0.828
SDQ (parent)	0.004		0.921	0.013		0.762	0.003		0.868
SDQ (Teacher)	0.091	**	0.013	0.102	**	0.012	0.023		0.204
School achieve. (parent)	-0.011		0.715	-0.015		0.647	-0.012		0.409
School achieve. (teacher)	0.099	**	0.033	0.123	**	0.017	0.077	***	0.001

Notes: ***, **, and * indicate the coefficient is significantly different from zero at the 1, 5 and 10 per cent level, respectively. Refer to Appendix Table A3.1 for details of control variables included.

If models are estimated with only the controls for basic demographics (child gender, age, cohort, Indigenous status and non-English speaking background), the coefficients for measures of density are negative in all the models with parent-assessed outcomes as the dependent variables, and these are often statistically significant. On face value, this would indicate deleterious impacts of household density on children's outcomes. However, the addition of the various controls for family socio-economic status and for neighbourhood characteristics reduce the magnitude of the estimates, demonstrating that these are confounding variables. The estimates for density are most sensitive to the inclusion of variables for the family's relative socio-economic position followed by home-ownership, indicating the importance of controlling for these effects in any empirical estimates of household density effects on children.

3.2.1 LSAC robustness checks

To allow estimation of the panel models using fixed effects, and therefore to control more rigorously for unobserved individual effects, the models were re-estimated using simple linear regression rather than the ordered probit model. Those fixed-effects estimates resulted in a combination of positive and negative estimated coefficients for the variables for density, but none were significantly different from zero at accepted levels. Estimation of the preferred models using the one period lag of *Density2* does not greatly change the results. The estimate for the teacher-assessed SDQ score remains positive but is insignificant, while the estimate for teacher-assessed school achievement is marginally larger and more significant (β=0.131, p=0.013). The other estimates for *Density2* remain insignificant when the variable is lagged. Recall that with LSAC, a one period lag represents a time period of two years.

The inclusion of the continuous density measures as independent variables imposes a monotonic, linear relationship between the density measure and the outcome variables. However, it is likely that any effect of household density on children's wellbeing varies over the range of density levels. Two alternative specifications were tested to allow for more flexible relationships: the inclusion of both the density measure and its quadratic; and replacing the continuous density variable with a series of dummy variables relating to separate intervals along the density range.

The inclusion of the quadratic term allows for a second order effect of density: *a priori* we might anticipate that increases in density are initially associated with positive effects on children's wellbeing but at a diminishing rate, such that a turning point is reached where further increases in density have a detrimental effect on children. Table A4.1 in the Appendix reports the corresponding results to Table 4, but with each of the 15 models estimated with the addition of the square of the household density measure. For parent-assessed general health, SDQ and school achievement, it remains the case that there is no significant association between child outcomes and the density measures with this more flexible specification. For teacher-assessed school achievement, the quadratic terms are insignificant, while the linear terms remain positive and significant, albeit at a lower level of significance. Hence, these models continue to show a positive association between household density and children's school achievement.

Only the models for the teacher-assessed SDQ total difficulties score show significant estimates for both the linear (positive) and quadratic (negative) coefficients, consistent with a concave relationship between density and outcomes. Using those estimated coefficients it is possible to calculate the 'turning point' at which further increases in density cease to be associated with improved outcomes, and begin to be associated with negative impacts on children (reported in bottom row, Table A4.1). For *Density1*, the straight ratio of occupants to bedrooms, this turning point is 2.99, which is above the 99^{th} percentile on the distribution. For *Density2* and *Bedreq1*, the turning points correspond to around the 99^{th} and the 97^{th} percentiles, respectively, of their distributions. Hence, this more flexible specification implies density levels do reach a point where they become 'crowding' in the sense of having a negative association with children's outcomes, but only one to three percent of children live in homes with such density levels.

Specifying the density measures through a series of mutually exclusive dummy variables provides a more flexible form again, freely allowing levels of density to have different effects across ranges of the variable. The *Density1* and *Density2* variables were coded into four categories of 'low', 'medium', 'high' and 'very high' where, as close as possible, these corresponded to children living in houses with the lowest 40 per cent of values of the density variables, the next 40 per cent (medium), and then next 10 per cent (high) and the top 10 per cent of values (very high). Exact alignment with these ranges was not possible due to the 'lumpy' nature of the distribution (see notes to Table A4.2 for actual ranges). This was chosen, partly, due to the skewed nature of the distribution, with many living in household with low values of the density variable, and, partly to ensure the models would provide estimates of the effects for children living in more extreme levels of the density distribution: the 9^{th} and 10^{th} deciles. For the number of extra bedrooms required, *Bedreq1*, the dummy variables relate to no bedrooms required (low, 72.7 per cent of children), 1 bedroom required (modest, 18.3 per cent of children), 2 bedrooms required (high, 5.9 per cent of children) and 3 or more bedrooms required (very high).

Appendix Table A4.2 reports the results for the coefficients for the dummy variables from models corresponding to those reported in Table 4. In all models the 'low' dummy variable is the omitted category, so the coefficients on the modest, high and very high dummies can be interpreted as their estimated effect relative to living in a low density household. All estimated coefficients that are statistically significant are positive, providing no evidence of negative density effects. As before, there are no significant coefficients for the density variables for the parent-assessed outcomes variables (general health, SDQ and school achievement). Turning to teacher-assessed SDQ scores, for the dummies based on ranges of the *Density1* and *Density2* measures, it is living in high density, equating to around the 9^{th} decile, that is estimated to be associated with the highest scores and these estimates are statistically significant (p=0.003 for *Density1*, p=0.035 for *Desnity2*). In results for the dummies based on ranges of the *Bedreq1* variable, it is living in homes that require 1 extra bedroom that has the highest estimated coefficient in the model for teacher-assessed SDQ, and that estimate is highly significant (p=0.002). Recall that the base category of requiring no extra bedrooms applies to 73 per cent of observations.

The results for teacher-assessed school achievement display the strongest evidence of a positive association between household density and children's outcomes. Generally, the magnitude of the coefficients and their level of significance are increasing with density, consistent with the results from the linear specification (Table 4). The dummy for living in very high density, as opposed to low density housing, is statistically significant for *Density1* (p=0.021), *Density2* (p=0.044) and *Bedreq1* (p=0.008). The latter, highly significant association applies to the 3.2 per cent of children living in a household assessed as requiring 3 or more extra bedrooms.

3.3 Multivariate estimates – LSIC

Models for parent-assessed general health of the child, parent- and teacher-assessed SDQ scores, and outcomes on standardised tests for reading and numeracy were estimated for the sample of Indigenous children from the LSIC. In each wave, the study child's general health was rated by the responding parent using that same 5-point scale described above for the LSAC. The SDQ was included in the parent surveys in waves 3, 4, 6 and 8 for the Kindergarten cohort and Waves 3, 6 and 8 for the Baby cohort; and in teacher surveys in Wave 2 for the Kindergarten cohort, and for all study children in Waves 3 to 6 and again in Wave 9. The abnormal, borderline, and normal ranges are used as recommended by Goodman (1997).

Cognitive outcomes were modelled using the reading and mathematics scores from the 'progressive assessment tasks' (PAT) undertaken by the interviewers directly with the children. Scaled scores for the assessment tasks for reading comprehension (PAT-R) are available for the Child cohort for Waves 4, 5, 6 and 8 and for the Baby cohort for Waves 7 and 8. Scaled scores for the mathematics tests (PAT-M) are available for the Kindergarten cohort in Waves 6, 7 and 9, and in Wave 9 only for the Baby cohort. The original test scores achieved by the student are not contained in the data, rather a scaling of the scores produced by the Australian Council for Educational Research (ACER), and designed to be comparable within and across waves (DSS 2017: 27–30). These are continuous variables ranging from 17 to 150, with mean 88 for the reading scale and from 57 to 150, with mean 105 for the maths scale. These are used as indicators of cognitive development. While the mathematics tests were only intended for completion by those from the Child cohort, there are a small number of scores in the Wave 6 and Wave 7 data for children from the Baby cohort, and these are retained in the analysis. As a robustness check, the children's matched National Assessment Program – Literacy and Numeracy (NAPLAN) scores for reading and mathematics are also tested as dependent variables. Consent to link the study children's NAPLAN data to LSIC data was requested from Wave 4 for both cohorts.

Ordered probit models were again used to model parental assessment of the child's general health, and the parent- and teacher-assessed SDQ ranges; and linear regression for the scaled PAT-R and PAT-M test scores. These were all initially estimated as panel models with random effects to account for unobserved child-individual effects. In all models, controls are included for study child's cohort, age, gender, and family type (lives with two parents, lives with sole parent, lives with carer(s)), housing tenure and remoteness. The decile of the ABS SEIFA indices for the child's address have been included in the data. These were tested as measures of the socio-economic status of the neighbourhood, with only the most significant of the four different indices retained. Controls relating to the responding parent include self-assessed health and level of education in Wave 1; employment status, parenting style ('warmth'), whether the family experienced serious money troubles, and assessment of the liveability of the neighbourhood (see Appendix Table A2.2 for variable definitions). As some of these latter variables are not collected in all waves and/or have significant incidence of missing values, they were dropped from the models when insignificant to maximise the sample available for estimation.

Table 5 reports the results from the preferred models for the four household density constructs. Full results for the models with the *Bedreq1* variable are reported in Appendix Table A3.2, as this is the specification that returns the most statistically significant estimates for the density variable. For children's general health, as assessed by the responding parent, and for both parent- and teacher-assessed outcomes on the Strengths and Difficulties Questionnaire, the results for the density measures are small and generally insignificant. However, as with the LSAC sample, the results suggest that, if anything, additional occupants are associated with better child outcomes even for this sample of Indigenous children with significantly higher household density levels than the wider population of Australian children. The estimated coefficients on all density measures are positive, and in two models they attain weak significance (10 per cent level). These are for the effect of *Bedreq1* on the child's general health and parent-assessed SDQ, recalling *Bedreq1* represents the number of extra bedrooms required to ensure one couple and all other household occupants have their own bedrooms.

There is, however, evidence that greater household density is associated with poorer cognitive outcomes. Each of the constructs are associated with lower maths scores, with the effect highly significant. Negative associations between density are also observed for the children's reading scores, significant at the 5 per cent or 10 per cent levels. The magnitude of the estimated effects is relatively modest. Recalling these estimates are from a linear regression, the estimated effect of the household requiring an additional bedroom is to reduce the score by one point or less. Differences by child gender and age are greater, with girls and older children performing better (see Appendix Table A3.2). As another comparison, estimated differences in scores between children living in very remote and non-remote areas amount to around 8 points on both the reading and maths scales. However, the effect implied for extreme cases of household density –the *Bedreq1* variable ranges up to a value of 16 extra bedrooms required – is substantial.

Table 5: Summary of results for density variables, LSIC models

	Density1		Density2		Bedreq1		Bedreq2	
	β	P>|z|	β	P>|z|	β	P>|z|	β	P>|z|
General health [1-5][a]	0.03	0.303	0.04	0.255	0.02*	0.065	0.03	0.171
SDQ (Parent) [a]	0.04	0.457	0.04	0.472	0.04*	0.094	0.04	0.182
SDQ (Teacher) [a]	0.05	0.609	0.05	0.609	0.06	0.116	0.05	0.470
PAT Reading[b]	-1.36**	0.021	-1.46**	0.017	-0.69**	0.011	-0.69*	0.096
PAT Maths[b]	-1.72***	0.001	-1.72***	0.001	-0.73***	0.000	-1.06***	0.000

Notes: a. Estimated by ordered probit model; b. Estimated by linear regression; ***, **, and * indicate the coefficient is significantly different from zero at the 1, 5 and 10 per cent level, respectively.

Commencing with basic models including only child and family demographics, and then successively adding groups of variables provides an indication of confounding factors between density and child outcomes. In the child health models, parental education, parental work status, and remoteness act as confounders. That is to say, lower parental education and absence of work is associated with both greater household density and lower assessed child health. Similarly, living in a more remote area is associated with great density and lower assessed child health. Housing tenure acts as a confounder for the parent-assessed SDQ, and parental education and work status for the teacher-assessed SDQ. Remoteness is the main confounding factor when it comes to the effect of density on the PAT scores – children in more remote areas live in higher density households and achieve lower scores. It is difficult to say how well the available control variables control for these confounding effects in the full models. The other variables, where significant, have the expected signs with the exception of the parental warmth scale, which is highly significant but negatively associated with parental assessments of the child's physical health and the parent-assessed SDQ score. Possibly this reflects that parents with more affectionate and engaged parenting styles are also more likely to be familiar with, or to recognise any health or behavioural problems.

In each wave, the questionnaires included a 'major life events' section in which the responding parent or carer was asked whether, in the past 12 months, they had felt too crowded where they lived, moved house or had housing problems. From Waves 4 to 9, respondents were asked to specifically indicate whether they had felt too crowded, allowing some comparison to be made between the derived density measures and parents' subjective sense of crowding. As would be expected, the density and crowding measures were, on average, higher for parents who reported feeling too crowded – by around one-half an additional person per bedroom or one-half an additional required bedroom. However, this subjective measure was not significant in the models of child outcomes. It is also the case that some people lived in houses that, by conventional measures, would be considered extremely overcrowded, but did not report feelings of being too crowded – many requiring 5 or more additional bedrooms by the measure allowing for 2 children to share the same bedroom (*Bedreq2*). However, it should be noted that the subjective assessment relates to life events over the previous year, and that experience may not relate to the composition of the household at the time of the survey, limiting the inferences that can be made through comparing the subjective reports and density measures.

3.3.1 LSIC robustness checks

The results for models for the PAT-R and PAT-M scales have been chosen for reporting as there are substantially more observations on those tests than for NAPLAN. However, comparable models estimated using the students' NAPLAN reading and maths test results tell a very similar story. Due to the intermittent collection of the outcomes variables, estimation with the *Bedreq1* variable lagged was tested only for the model for general health. The coefficient for the 1-year lag of *Bedreq1* in that model was similar in magnitude, but not quite significant at the 10 per cent level (β=0.019, p=0.123).

As with the LSAC analysis, models were re-estimated with the inclusion of quadratic terms of the density variables, and using a series of mutually exclusive dummy variables to capture density effects (see discussion in Section 3.2.1, Appendix Tables A4.3 and A4.4 for LSIC results). With the addition of the quadratic terms, significant associations are no longer observed between *Bedreq1* and general health or the parent-assessed SDQ score (these were significant only at the 10 per cent level in the linear model). However, in the case of the parent-assessed SDQ score, the estimates now imply concave relationships with *Density1*, *Density2* and *Bedreq2* are now implied, with outcomes improving as density increases but at a declining rate. For *Density1* the coefficients imply a turning point of 3.82, and for *Density2* a turning point of 3.57. In both cases these turning points are above the 98th percentile in the distribution, suggesting negative associations of increasing density only at the very upper extremes of the distribution of households in which Indigenous children live. It should be noted, however, that the estimates on the linear terms are only weakly significant for both these measures.

In contrast to the simple linear specification, the quadratic specifications generate minimal evidence of a negative association between density and the PAT-R reading score, with the exception of *Bedreq1* for which the quadratic term is negative and highly significant. The negative association between household density and the PAT-M scores is confirmed, with the results implying scores declining as household density increases, but at a declining rate. The turning points implied by the estimated coefficients are at the very upper end of the distribution, indicating a negative association with additional density applies across almost the full range of household densities. In all, the quadratic specifications support the finding of a general absence of negative associations between household density and Indigenous children's general health or social and emotional adjustment, and offering a hint of positive associations. Evidence of negative associations between household density and cognitive development is reinforced in the case of mathematics, but weakened in the case of reading.

These findings are further reinforced by the specifications using dummy variables (see Appendix Table A4.4). For the child's general health and the two SDQ total difficulties scores, there is no evidence that medium, high or very high density levels area associated with poorer outcomes compared to children in low density households. Only three of 36 coefficients are negative, each small in magnitude and insignificant. For the *Bedreq2* measure, which allows for children aged 15 and under to share a bedroom, the household requiring one extra bedroom as opposed to no extra bedrooms, is associated with higher parent- and teacher-assessed SDQ scores, both significant at the 5 per cent level. The specifications with dummy variables show strong evidence that increased density is associated with lower PAT-M test scores and some further evidence of density being associated with lower PAT-R scores.

3.4 Multivariate estimates – HILDA

Individuals within a HILDA household are interviewed as 'responding persons' once they turn 15 years of age. To assess the impact of children's housing circumstances on their outcomes as young adults, outcome variables observed for responding persons between the ages of 15 years and 19 years are modelled conditional upon housing variables averaged over two stages of childhood: over the ages of 5 to 9 years and 10 to 14 years. With data from 18 annual waves of the HILDA survey available at the time of writing, a person who was 5 years of age in the first (2001) HILDA survey would have turned 15 in Wave 11, and 19 in Wave 15. For such an individual, there are potentially 5 observations on outcomes from age 15 to 19 which can be related to their circumstances in young childhood (ages 5-9). This also holds for persons who were aged 5 in Waves 2, 3 and 4. For younger persons, the number of repeated observations on outcomes declines – someone aged 5 in Wave 8 will have turned 15 in Wave 18, and contribute observations on outcomes for just that one year. Hence, conditioning on information in young childhood generates an unbalanced panel. Conditioning on information while the person was aged 10-14 years similarly results in an unbalanced panel, but with substantially more observations, with outcomes observed for 15-year-olds from Wave 6 onward.

Outcomes are modelled in the domains of physical health; mental health and wellbeing; and educational attainment. Two measures of physical health are analysed. The first is self-assessed health reported on a five-point scale ranging from poor, fair, good, very good to excellent. The second is the physical health component score based on the 'SF-36'. HILDA contains the set of questions that make up the 36 item 'short form survey' developed in the Medical Outcomes Study (the SF-36). These items are used to generate eight 'scales' capturing physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional and mental health (Ware et al. 2000[6]). In turn, a physical health component summary score (PCS) and mental health component summary score (MCS) are generated through factor analysis of those eight scales.

Earlier releases of HILDA contained the PCS and MCS as derived variables, based on scoring coefficients for the eight scales reported in the Australian Bureau of Statistics' National Health Survey. These have not been included in more recent releases. For this study, a factor analysis was conducted on the pooled sample of 15–19-year-olds across the 18 waves of HILDA to generate physical and mental health component summary scores with scoring coefficients specific to the age group being analysed. The PCS correlates most strongly with the physical functioning, role-physical and bodily pain scales (see Appendix Table A2.3 for details). The resulting summary score is a continuous variable standardised to have a mean of 50 and standard deviation of 10 across the Wave 1-to-18 pooled sample of 15 to 19-year-olds.

The Mental Health Component summary score derived from the SF-36 items is the variable used to measure mental health. The MCS correlates most strongly with the mental health, vitality, role-emotional and social functioning scales (see Appendix Table A2.3). As with the PCS, it is standardised to have a mean of 50 and standard deviation of 10. Subjective wellbeing is measured using the question on life satisfaction, in which respondents indicate how satisfied they are with their life, 'all things considered', on a scale from 0 (totally dissatisfied) to 10 (totally satisfied).

Indicators of educational outcomes are whether or not the respondent had completed Year 12, assessed at age 18, and whether they entered university by age 19. This later indicator is based on whether or not they had completed a university degree or were enrolled at university at age 19. Educational outcomes are modelled only using data on housing circumstances when the person was aged 10-14 years, due to the limited sample size when data for earlier childhood are included.

Variables capturing families' housing circumstances and socio-economic background were defined for each child for each year from age 5 years to age 14 years, including the four density measures: *Density1*, *Density2*, *Bedreq1* and *Bedreq2*, as defined above. These variables were then averaged over two five-year childhood periods: 5–9 years and 10–14 years of age. Housing variables included

[6] See also: http://www.sf-36.org/tools/sf36.shtml

tenure (ownership status), neighbourhood socio-economic status (ABS SEIFA deciles) and remoteness. Variables designed to capture the socio-economic background of the family were also averaged for the two periods of childhood. These were annual household income (in logs), whether the child lived in a sole parent household, the highest level of education attained by a parent, whether a parent had a professional occupation, and whether the child lived in a household with no parent in work. Contemporaneous control variables measured at the time the youth was aged 15 to 19 were included in the models where appropriate. These included the youth's gender, age, whether English was their first language, Indigenous status, living status and main current activity (see Appendix Table A2.3 for variable definitions).

Models for the continuous, linear outcome variables (the SF-36 PCS and MCS) were estimated by panel ordinary least squares, and self-assessed health and life satisfaction using the panel ordered probit model. Models for completion of Year 12 and entering university are estimated as cross-sectional binary logit models, since there is only one observation per individual. All panel models are estimated using random-effects. Because the historical housing variables are fixed for each individual, their effects cannot be estimated using fixed-effects.

Table 6 provides a summary of the estimates for the various density measures from models which include independent variables defined when the person was aged 5-9 years, and Table 7 the estimates for the density variables from models estimated with independent variables relating to when the youth was aged 10-14 years. As the estimates for the *Bedreq2* specification are the most consistently significant, full results from the corresponding models using that variable are reported in Appendix Tables A3.3 to A3.5. The substantially lower number of observations available when variables defined at age 5-9 years are included (approx. 4,000) compared to when variables defined at age 10-14 (approx. 9,000) can be seen by comparing Appendix Tables A3.3 and A3.4.

The results appear quite striking. People who lived in more densely occupied homes as children are both physically and mentally healthier as young adults, based on their own reports. Starting with the measures defined when the respondents were aged 5-9 years (Table 6), the estimates for self-assessed general health are positive and highly significant for all four specifications of the density variables, with weak confirmation of this from the effect of *Bedreq2* on the SF-36 physical health component summary score. The estimates for all four constructs are also positive and highly significant in the models of the SF-36 mental health summary score, supported by modest evidence that a more densely occupied home in childhood is associated with higher life satisfaction as a young adult.

Table 6: Summary of results for crowding variables measured at age 5-9 years: HILDA

	Density1		Density2		Bedreq1		Bedreq2	
	β	P>|z|	β	P>|z|	β	P>|z|	β	P>|z|
General health [1-5][a]	0.40***	0.001	0.42***	0.001	0.17***	0.001	0.38***	0.003
SF-36 Physical health [b]	0.17	0.786	0.09	0.900	0.00	0.989	0.70*	0.092
Life satisfaction [0-10] [a]	0.15	0.155	0.19*	0.080	0.09**	0.026	0.22**	0.021
SF-36 Mental health [b]	2.14***	0.002	2.37***	0.002	0.86***	0.002	2.15***	0.001

Notes: a. Estimated by ordered probit model; b. Estimated by linear regression; ***, **, and * indicate the coefficient is significantly different from zero at the 1, 5 and 10 per cent level, respectively.

The same pattern of a beneficial association with general health and the SF-36 mental health score applies for the effects of household density when the person was aged 10-14, although the effects are not as robust for self-assessed general health. Interestingly, the estimated effects of density on mental health and life satisfaction are larger and generally more significant for the variables defined in younger childhood than for those defined in later childhood. This seems surprising given the greater time lag and the markedly fewer number of observations available for estimating those effects.

No significant associations are observed between household density during the ages of 10-14 and either the probability of completing Year 12 or of having commenced at university by age 19. The results for other variables do concord with factors known to affect educational outcomes and access to higher education, including inferior outcomes for males, youth from sole parent families and who lived in regional and remote Australia, and better outcomes for children of more educated parents and who own their own home.

Table 7: Summary of results for crowding variables measured at age 10-14 years: HILDA

	Density1		Density2		Bedreq1		Bedreq2	
	β	P>|z|	β	P>|z|	β	P>|z|	β	P>|z|
General health [1-5][a]	0.21**	0.025	0.23**	0.020	0.10**	0.016	0.21**	0.020
SF-36 Physical health [a]	-0.17	0.715	-0.29	0.567	-0.30	0.166	-0.45	0.285
Life satisfctn [0-10][a]	0.11	0.163	0.14	0.107	0.03	0.389	0.18***	0.010
SF-36 Mental health [a]	1.52***	0.005	1.62***	0.005	0.64***	0.005	1.80***	0.000
Completed Yr 12[c]	-0.19	0.366	-0.23	0.313	-0.09	0.303	-0.15	0.409
Entered university [c]	0.07	0.723	0.02	0.940	-0.01	0.889	0.02	0.914

Notes: a. Estimated by ordered probit model; b. Estimated by linear regression; c. estimated by cross-section logit model. ***, **, and * indicate the coefficient is significantly different from zero at the 1, 5 and 10 per cent level, respectively.

3.4.1 HILDA robustness checks

When the models including *Bedreq2* were estimated with the historical variables for age 5-9 years and age 10-14 years included simultaneously, the *Bedreq2* variable averaged for age 5-9 years remains positive and weakly significant for self-assessed general health (β=0.407, p=0.093) and the SF-36 mental health summary score (β=2.214, p=0.082), but insignificant for life satisfaction and the SF-36 physical health score. The *Bedreq2* variable averaged over the ages of 10-14 years is not significant in any of the models once the corresponding variable in younger childhood is included.

Results from corresponding models from Table 6 with the square of the density measures added are reported in Appendix Table A4.5 (averaged over ages 5-9 years). It seems clear the simple linear specification is preferred when it comes to self-assessed general health. For *Density1*, *Density2* and *Bedreq1* the coefficients are insignificant and, unusually, the linear and second order effects are both positive. For Bedreq2, the significant linear relationship with self-assessed health is confirmed, with a larger coefficient offset by a small and insignificant estimate for the quadratic term. For the SF-36 physical health summary score, the coefficients indicate a convex relationship, with the score initially falling and then rising with density. This is counterintuitive, and the negative coefficients for the linear measure are only weakly significant. The implied turning points at which the estimated associations with density become positive are at around the 80th percentile for *Density1* and *Density2*, and 98th percentile for *Bedreq2*. In the case of *Bedreq2*, the positive coefficient for the linear variable is weakly significant, while the positive coefficient on the quadratic term is highly significant. None of the density estimates are significant in the life satisfaction or mental health models with the exception of *Bedreq2*, for which both the linear (β=5.65, p=0.000) and quadratic terms (β=-1.09, p=0.005) are highly significant. This generates the expected concave relationship with mental health positively associated with density up to the point of around 2.6 bedrooms required, which is above the 99th percentile on the distribution.

Results obtained when the density measures are specified as a series of dummy variables provide strong evidence of positive associations between household density at age 5-9 years and later physical and mental health (Appendix Table A4.6). Living in the very highest levels of household density – equating roughly to the top decile of the distribution at age 5-9 years - is associated with significantly better self-assessed general health and higher SF-36 mental health summary scores relative to those living in low density households, and this is consistent across the four density measures. Only two weakly significant negative associations suggest children from higher density households may have inferior outcomes relative to those who lived in low density households: for the high *Density1* dummy in the model for the SF-36 PCS, and the medium *Bedreq1* dummy in the model for life-satisfaction. Coefficients for the other density dummies in those models are insignificant, but not consistent with a pattern of a detrimental association with increasing density.

The inclusion of quadratic terms for the density variables measured at age 10-14 years provides no evidence of any negative association between household density at that age and physical health, life satisfaction or mental health as a young adult (see Appendix Table A4.7). There is weak evidence of higher household density being associated with better SF-36 mental health scores but with a lower likelihood of entering university. The results from modelling household density at age 10-14 years as a series of dummy variables (Appendix Table A4.8) suggest those who lived in houses with medium density report the highest life satisfaction as young adults, and those who lived in very high density homes had the highest SF-36 mental health scores. In contrast there is some evidence that living in a home requiring more bedrooms is associated with lower SF36 physical health scores and a lower probability of completing Year 12.

4.0 Discussion and conclusions

Crowding is often raised as a key characteristic of housing circumstances that can impact upon children's developmental outcomes. In Australia, there appears to be a general acceptance in housing policy discourses of the existence of a significant problem of 'overcrowding.' However, a deeper investigation suggests this view has formed around a very scant evidence base. In this paper, I have sought to provide new empirical evidence on the association between household density and children's physical, emotional and cognitive development in Australia, and the implications relating to the existence and extent of a problem of 'crowding' for children in contemporary Australia.

The key finding from the analyses is that, on average, Australian children live in homes with household density levels not too high, but too low for their optimal health and social and emotional development. Evidence from LSAC and LSIC suggests that household density—measured on the basis of occupants per bedroom or the need for additional bedrooms—is not significantly associated with child physical health. Analysis of data from those surveys provides weak evidence that higher density at home is associated with better child social and emotional development. So while this evidence of positive associations between density and children's general health and social and emotional development is not strong, it is clearly at odds with the hypothesis of a problem of crowding contributing to inferior outcomes within the range of those variables that apply to most

Australian children. More compelling are the results from HILDA relating to children's housing circumstances when they were aged 5–9 years or 10–14 years. Children who lived in homes assessed as requiring additional bedrooms for the number of occupants subsequently had better self-assessed general health as young adults, as well as greater life satisfaction and better mental health.

Findings relating to the effect of household density levels on children's cognitive and educational outcomes were mixed. For the LSAC sample of the general population of Australian children, higher occupancy levels were found to be associated with better teacher ratings of school achievement. However, there is quite robust evidence of lower performance on standardised tests of reading and mathematics for Indigenous children who live in homes with higher household density. There is also weak evidence from HILDA that higher household density in later childhood is associated with a reduced probability of school completion or progression to university.

The absence of evidence of negative effects of household density on Indigenous children's health and social and emotional wellbeing seems the most remarkable finding, given the emphasis placed on the problems of overcrowding for the Indigenous population (SCRGSP 2016: Chapter 10). The mean number of occupants per bedroom for families in the LSIC sample is 1.64, but ranging as high as 12, and some children were living in homes assessed as requiring as many as 16 additional bedrooms. Certainly, the underlying variables extend into ranges that would be considered overcrowding by any norms, and yet significant effects on health or social and emotional adjustment are not found. Careful consideration is needed in trying to understand this finding and its implications. As noted, however, a previous study of Indigenous children in Western Australia also measured fewer emotional and behavioural difficulties for children living in higher occupancy dwellings (Zubrick et al. 2005), so it is not without precedence.

It must be stressed again that these results do not imply that 'crowding' is a good thing. It makes little sense to refer to levels of household density or of bedroom requirements as 'crowding' if the effect of an increase in those measures is positive, given that the word 'crowding' carries with it connotations of negative psychological impacts. An analogy might be useful here. Malnourishment is detrimental to health and development. However, I imagine a study of the health of the Australian population would find that a modest reduction in food intakes would have positive health benefits. Of course, this fact should not be used to dismiss the need to address malnourishment for those living in food poverty, or to suggest malnourishment is not a problem. Equally, however, the fact that some Australians are malnourished should not blind us to the fact that the Australian population, in general, would benefit from shedding a few pounds.

In a similar vein, crowding – or 'overcrowding' is a problem for those experiencing it. At the same time, I believe the results presented above do indicate that Australian children, in general, live in homes that are too empty. There is no contradiction, but ambiguity arises because of the measures used and the labels given to them. The density ratios and criteria for required bedrooms used here are typical of measures used in the existing literature. The need for extra bedrooms has not been defined here using the CNOS, the most commonly used standard. Potentially, the CNOS could be replicated using the HILDA data, but not with LSAC or LSIC. However, the *Bedreq2* variable will closely mirror indicators of crowding based on the CNOS. The key difference between the two is that while *Bedreq2* does assume children under 15 years of age can share a bedroom, it does not impose the standard that children aged 5 and over should not share a bedroom if they are of opposite genders. However, with over 70 per cent of Australian children living in houses where there is a bedroom for everybody, the added requirement based on gender would mean the two measures would deviate in only a small proportion of cases. Hence, the measures used above do correspond to those that are commonly labelled 'crowding'.

This comes back to the important distinction between household density and household crowding measures. Density is a ratio of two objectively determined quantities – typically the number of people and the number of bedrooms - with no reference point to what is or is not considered crowding. Implicitly, a crowding measure imposes a subjectively determined standard regarding the point beyond which density becomes detrimental to occupants' wellbeing. The evidence presented here suggests, at least in respect to the wellbeing of children, those standards for the Australian population may need to be reconsidered. The cusp implicit in common standards for the number of bedrooms required appears instead to fall within the range, going back to Rapoport, 'of too low densities which we may call isolation'. Experimentation with more flexible model specifications using quadratic terms and dummy variables to capture non-monotonic effects over the range of the density measures suggests density may have adverse associations with children's physical health and social and emotional development only at the very extremes of the distribution experienced by Australian children. Even here, positive associations between household density when children were aged 5-9 and their physical and mental health as young adults appear to persist into even the highest (tenth) decile of the household density distribution.

The findings point to the need for further research in a number of areas. One is the question of the definition and measurement of crowding, and this is currently the subject of an ongoing project. The results for Indigenous children suggest the need for more culturally nuanced definitions and indicators of crowding for Indigenous households, as proposed by Memmott et al. (2012), and the need to incorporate the diversity of cultural norms more generally into the measurement of crowding.

There is also a need for more detailed exploration and understanding of the way density actually impacts upon physical health and psychological wellbeing, both with respect to children and adults. My working hypothesis for the finding of a positive association between household density measures and the physical health and social and emotional wellbeing of children in the general population is that children now grow up too isolated within their own homes. Too often, they have separate bedrooms and living spaces when they would instead benefit from more interaction with other siblings and adults. The evidence from HILDA indicates that young adult's health and wellbeing are more sensitive to their housing circumstances in young childhood (aged 5-9 years) than the circumstances in later childhood. Following Bronfenbrenner's (1979a,b) bio-ecological theory of the development of the child, this may be explained by the immediate microsystem of the home and family being most important for child development in the early years, with external social settings (mesosystems) of the neighbourhood and school becoming of greater relative importance as children age. Hence, the level of interaction they experience within the family home when they are young may have the greatest impact on later outcomes.

Finally, this paper's focus on the effects of density, and its implications for what we call 'crowding', has come at the neglect of comment on findings relating to the effect of other housing and non-housing factors on children's wellbeing. Those results provide evidence that higher family and neighbourhood socio-economic background, and home-ownership are associated with better outcomes for children. Moreover, it is critical to control for these as confounding factors in any empirical analysis of the effects of household density on children's outcomes, particularly if analysts are aiming to make inferences about crowding.

References

Australian Bureau of Statistics (ABS) (2011), Position Paper - ABS Review of Counting the Homeless Methodology, Catalogue No. 2050.0.55.002, ABS, August.

Australian Bureau of Statistics (2013), *Census of Population and Housing: Socio-economic Indexes for Areas (SEIFA), Australia, 2011*, ABS cat. no. 2033.0.55.001, Canberra.

Australian Housing and Urban Research Institute (2019), National Housing Research Program 2020 Research Agenda. https://www.ahuri.edu.au/__data/assets/pdf_file/0027/39546/2020-NHRP-Research-Agenda.pdf

Australian Institute of Family Studies (2017), *The Longitudinal Study of Australian Children Annual Statistical Report 2016*, AIFS, Melbourne.

Australian Institute of Health and Welfare (AIHW) (2014), Housing circumstances of Indigenous households: tenure and overcrowding. Cat no. IHW 132. Canberra: AIHW.

Australian Institute of Health and Welfare (AIHW) (2019), Housing assistance in Australia 2019, AIHW. https://www.aihw.gov.au/reports/housing-assistance/housing-assistance-in-australia-2019/report-editions

Baker, M., McNicholas, A., Garrett, N., Jones, N., Stewart, J., Koberstein, V. and Lennon, D. (2000), 'Household crowding a major risk factor for epidemic meningococcal disease in Auckland children', *Pediatric Infectious Disease Journal*, 19(10), 983-90.

Baker, K., Sipthorp, M. and Edwards, B. (2017), 'A longitudinal measure of socioeconomic position in LSAC', LSAC Technical Paper No. 18, Australian Institute of Family Studies, January.

Bartlett, S. (1998), 'Does Inadequate Housing Perpetuate Children's Poverty?', *Childhood*, 5(4), 403–420.

Biddle, N. (2008), 'The scale and composition of Indigenous housing need, 2001-2006, CAEPR Working Paper No. 47/2008, Centre of Aboriginal Economic Policy Research, ANU.

Booth, A. and Carroll, N. (2005), 'Overcrowding and Indigenous health in Australia', Centre for Economic Policy Research Discussion Paper No. 498, ANU.

Bronfenbrenner, U. (1979a), 'Contexts of Child Rearing: Problems and Prospects', *American Psychologist*, 34(10): 844–850.

Bronfenbrenner, U. (1979b), *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge, Massachusetts: Harvard University Press.

Coggon, D., Barker, D., Inskip, H. and Wield, G. (1993), 'Housing in Early Life and Later Mortality', *Journal of Epidemiology and Community Health*, 47, 345–348.

Commissioner for Children and Young People (CCYP) (2014), *The state of Western Australia's children and young people – Edition 2*, CCYP, Western Australia.

Cooper, L. and Morris, M. (2005), 'Sustainable tenancy for Indigenous families: what services and policy supports are needed?', AHURI Final Report No. 81, Australian Housing and Urban Research Institute, June.

Department of Families, Housing, Community Services and Indigenous Affairs (FaCHSIA) (2009), *Footprints in Time: The Longitudinal Study of Indigenous Children – Key Summary Report from Wave 1*, FaHCSIA, Canberra.

Department of Social Services (DSS) (2017), *Footprints in Time: the Longitudinal Study of Indigenous children – Data User Guide*, release 8.1, DSS, Canberra.

Dockery, AM. and Bawa, S. (2019), 'Why do home-owners do better?', State of Australian Cities Conference, 3-5 December, University of Western Australia, Perth (proceedings forthcoming).

Dockery, AM., Kendall, G., Li, J., Mahendran, A., Ong, R. and Strazdins, L. (2010), *Housing and children's development and wellbeing: a scoping study*, AHURI Final Report No. 149, Australian Housing and Urban Research Institute, July.

Dockery, AM., Ong, R., Colquhoun, S., Li, J. and Kendall, G. (2013), *Housing and children's development and wellbeing: Evidence from Australian Data*, AHURI Final Report No. 201, Australian Housing and Urban Research Institute, Melbourne.

Evans, GW., Lercher, P. and Kofler, WW. (2002), 'Crowding and Children's Mental Health: The Role of House Type', *Journal of Environmental Psychology*, 22(3), 221–231.

Evans, GW., Maxwell, L. and Hart, B. (1999), 'Parental Language and Verbal Responsiveness to Children in Crowded Homes', *Developmental Psychology*, 35, 1020–1023.

Evans, GW., Saegert, S. and Harris, R. (2001), 'Residential Density and Psychological Health among Children in Low-Income Families', *Environment and Behavior*, 33(2), 165–180.

Goodman, R. (1997), 'The Strengths and Difficulties Questionnaire: A research note', *Journal of Child Psychology & Psychiatry,* 38(5), 581-586.

Goodman, R. (2001), 'Psychometric properties of the Strengths and Difficulties Questionnaire', *Journal of the American Academy of Child & Adolescent Psychiatry,* 40(11), 1337-45.

MacNamara, J., Cassells, R., Wicks, P. and Vidyattama, Y. (2010), 'Children in housing disadvantage in Australia: development of a summary small area index', *Housing Studies*, 25(5), 625-646.

Mann, S., Wadsworth, M. and Colley, J. (1992), 'Accumulation of Factors Influencing Respiratory Illness in Members of National Birth Cohort and Their Offspring', *Journal of Epidemiology and Community Health*, 46, 286–292.

Memmott, P., Greenop, K., Clarke, A., Go-Sam, C., Birdsall-Jones, C., Harvey-Jones, W., Corunna, V. and Western, M. (2012), 'NATSISS crowding data: What does it assume and how can we challenge the orthodoxy?', in *Survey Analysis for Indigenous Policy in Australia: Social Science Perspectives*, Research Monographs, ANU E Press, pp. 241-279.

Rapoport, A. (1976), 'Toward a redefinition of density', in S. Saegert (ed.), *Crowding in Real Environments*, Sage Publications, California, pp. 7-32.

Solari, CD. and Mare, RD. (2012), 'Housing crowding effects on children's wellbeing', *Social Science Research*, 41, 464-476.

Steering Committee for the Review of Government Service Provision (SCRGSP) (2016), *Overcoming Indigenous disadvantage: Key indicators 2016 report*, Productivity Commission.

Wachs, TD. and Camli, O. (1991), 'Do Ecological or Individual Characteristics Mediate the Influence of the Physical Environment on Maternal Behavior?', *Journal of Environmental Psychology* 11(3), 249–264.

Ware, JE., Snow, KK., Kosinski, M. and Gandek, B. (2000), *SF-36 Health Survey: Manual and Interpretation Guide*, QualityMetric Inc., Lincoln, RI.

Waters, A-M. (2001a), 'Do housing conditions impact on health inequalities between Australia's rich and poor?, AHURI Final Report No. 4, Australian Housing and Urban Research Institute.

Waters, A-M. (2001b), 'Do housing conditions impact on health inequalities between Australia's rich and poor?, AHURI Positioning Paper No. 2, Australian Housing and Urban Research Institute.

Waters, A-M. (2002), 'Do Housing Conditions Make a Difference to Our Health?', Australian Housing and Urban Research Institute, Research and Policy Bulletin, Issue 6, April.

Watson, N. and Wooden, M. (2010), 'The HILDA survey: progress and future developments', *Australian Economic Review*, 43, 326–36.

Wilkins, R. (ed) (2014), *Families, Incomes and Jobs, Volume 9: A statistical report on waves 1 to 11 of the Household, Income and Labour Dynamics in Australia Survey*, Melbourne Institute of Applied Economic and Social Research.

Zubrick, SR., Silburn, SR., Lawrence, DM., Mitrou, FG., Dalby, RB., Blair, EM., Griffin, J., Milroy, H., De Maio, JA., Cox, A. and Li, J. (2005), *The Western Australian Aboriginal Child Health Survey: The social and emotional wellbeing of Aboriginal children and young people,* Curtin University and Telethon Institute for Child Health Research, Perth.

Appendix 1: The Canadian National Occupancy Standard (CNOS)

The CNOS assesses the bedroom requirements of a household based on the following criteria (https://meteor.aihw.gov.au/content/index.phtml/itemId/386254; ABS 2011):

- There should be no more than 2 persons per bedroom;
- Children less than 5 years of age of different sexes may reasonably share a bedroom;
- Children 5 years of age or older of opposite sex should have separate bedrooms;
- Children less than 18 years of age and of the same sex may reasonably share a bedroom; and
- Single household members 18 years or older should have a separate bedroom, as should parents or couples.
- A lone person household may reasonably occupy a bed sitter (i.e. a residence with no separate bedroom)

The CNOS is commonly used as the basis for household-specific measures of crowding through either:

- A binary variable designating households as either crowded or not crowded depending upon whether the dwelling has a sufficient number of bedrooms, or
- A continuous variable equal to the required number of bedrooms minus the actual number of bedrooms where the required number exceeds the actual number, and equal to zero where the actual number is equal to or greater than the required number of bedrooms.

Appendix 2: Additional variable definitions

Table A2.1: LSAC Variable definitions

Variable	Definition and range
Non-English speaking background	Dummy variable = 1 if the language spoken at home is other than English, = 0 if the language spoken at home is English.
Financial prosperity	Parental assessment of how family is getting on financially, coded 1='very poor', 2='poor', 3='just getting along', 4='reasonably comfortable', 5='very comfortable', 6='prosperous'.
Socio-economic position	A derived variable in LSAC based on the parent's or both parents' (as applicable) education, occupational status and income. A change in the methodology was necessitated due to the change in occupational coding used. While the versions are not strictly comparable, we use the original measure from ages 0 to 6 and the revised version from age 8 onwards. The variable is calculated as an index with a mean of zero and standard deviation of 1 for each wave, with higher values indicating higher family socio-economic position (see Baker et al. 2017 for details on methodology).
Sole parent	Dummy variable = 0 if the child has both parents living in the home, = 1 otherwise.
Homeowner	Dummy variable = 1 if the current dwelling is either owned outright or being paid off; = 0 otherwise.
SEIFA disadvantage	Based on the ABS socioeconomic indices for areas (SEIFA) for disadvantage derived from the Census and linked to the family's address, rounded to the nearest 10. The original index has a mean of around 1,000. This is scaled down by a factor of 10. The resulting variable has a mean of 10.1 and standard deviation of 0.64 for the pooled sample, with higher values indicating a more advantaged neighbourhood.
Liveability index	Mean of parental ratings to three questions on whether neighbourhood is safe, is clean and has good play areas. Coded 1='strongly disagree', 2='disagree', 3='agree' and 4='strongly agree'.

Table A2.2: LSIC Variable definitions

Variable	Definition and range
Family type: Lives with both parents Lives with one parent Lives with carer(s)	Series of three mutually exclusive dummy variables based on who the study child (SC) lives with. SC lives with parent and partner (including adoptive or step-parents). SC lives with lone parent. SC lives with carer (either lone or partnered).
Parent 1 health	Parent 1 self-assessment of their health in Wave 1. Coded 1='poor', 2='fair' 3='good', 4='very good', 5='excellent'.
Parenting warmth scale	Based on four questions about how often parent/carer hugs SC, ensures SC does what is told, praises SC for doing well, enjoys doing things with SC. Each rating is from 1 (always) to 5 (never). Scale is reversed and the mean calculated, so that a higher value implies more frequent affectionate parenting behaviours. In waves in which these questions were not asked, the parental warmth value from the most recent prior wave is used.

Variable	Definition and range
Parent 1 works	Dummy variable = 1 if responding parent/carer or their partner is in work, = 0 otherwise.
Parent level of education: Less than year 10 Completed Year 10 Completed Year 12 Post-school quals	Four mutually exclusive and time-invariant dummy variables, based on parent/carer's reported level of qualification in Wave 2. Less than Year 10 = 1 if highest year of school completed is Year 9 or below, or if parent never went to school. Completed Year 10 = 1 if parent competed Year 10 or 11. Completed Year 12 = 1 if highest qualification is Certification of completion, Year 12 or equivalent, Certificate I/II or other non-school certificate. Post-school qualifications = 1 if highest qualification is Certificate III/IV, trade qualification, diploma/associate diploma, degree or post-graduate qualification. Completed Year 10 is used as the default category.
Family had money trouble	Dummy variable = 1 if answer to 'In the last year has your family had serious worries about money?" is 'yes', = 0 if 'no'.
Housing tenure: Community housing Public housing Private rental Homeowner Other	Set of mutually exclusive dummy variables. Community housing = 1 if home is rented from community or co-operative housing group or Indigenous Community Housing Organisation. Public housing = 1 if home is rented from government housing authority. Private rental = 1 if home is rented from private landlord or employer. Homeowner = 1 if home is being purchased or owned outright by parent and/or their partner. Other = 1 if home is occupied rent free or none of the above apply. Public housing is used as the default category.
SEIFA deciles: Socio-econ. disadv. Socio-econ adv. & disadv. Economic resources	The decile of the area in which the study child's family lives. These are derived variables available in LSIC. The indices are based on ABS 2011 census data by area (see ABS 2013 for details). Each variable ranges from 1 to 10 and is coded such that a higher number represents an area of higher socio-economic status.
Liveability scale	Average of responses on 1–5 scales to 3 questions asking if this is a good community for little kids, whether there are good places to play and whether it is a safe community. In each case 5 indicates a better/safer community for kids. In waves in which these questions were not asked, the value from the preceding wave is used provided the family has not moved.
Remoteness level None Low Moderate High	The Level of Relative Isolation is included in the LSIC data from geocoding of the home address in which the study child lives. The underlying variable has values of 1 (none), 2 (low), 3 (moderate) and 4 (high/extreme). This is transformed into four mutually exclusive dummy variables. 'Low' is used as the default category.

Table A2.3: HILDA Variable definitions

Variable	Definition and range
Self-assessed health	Five-point scale: 1='poor', 2='fair', 3='good', 4='very good', 5='excellent'.
SF-36 physical health	Weighted index of eight SF-36 subscales. Weights determined by factor analysis with 2-factor solution for pooled 15–19-year-old sample: Loadings for raw PCS are 0.542*physical-functioning + 0.480*role-physical + 0.371*bodily-pain + 0.081*social-functioning + 0.058*general-health – 0.091*role-emotional – 0.132*vitality – 0.176*mental-health. The final PCS variable is standardised with mean = 50 and standard deviation = 10.
SF-36 mental health	As for PCS. Loadings for raw MCS are 0.345*mental-health + 0.315*vitality + 0.209*social-functioning + 0.262*role-emotional + 0.183*general-health – 0.018*bodily-pain – 0.093*role-physical – 0.196*physical-functioning. The final MCS variable is standardised with mean = 50 and standard deviation = 10.
English is 2nd language	Dummy variable = 1 if English was not first language learned as a child; = 0 otherwise.
Indigenous	Dummy variable = 1 if respondent identifies as Aboriginal, Torres Strait Islander, or both; = 0 otherwise
SEIFA decile of neighbourhood	The decile of the area in which the respondent currently lives. These are derived variables available in HILDA. The indices are based on ABS 2011 census data by area (see ABS 2013 for details). Each variable ranges from 1 to 10 and is coded such that a higher number represents an area of higher socio-economic status.
Living status: Couple with children Couple, no children Sole parent Lives both parents Lives sole parent Lives with others	Series of 7 mutually exclusive dummy variables based on respondent's current living status and based on HILDA derived variables for HH type and respondent's relationships within family. Respondent is deemed dependent child if living with their parent(s), irrespective of personal financial means. 'Lives with others' defined as being unrelated to all household members or living with family members who are not their partner, child or parent(s). Living with both parents is used as the default category.
Activity status: Still at school Work Post-school study Work and study Inactive	Series of 5 mutually exclusive dummy variables based on respondent's activity at time of survey. Still at school is used as the default category. For persons who had left school: Work = 1 if working full-time or working part-time and not studying; Post-school study = 1 if studying full-time or studying part-time but not working; Inactive = 1 if neither working nor studying; Work and study = 1 if working part-time and studying part-time. People were coded as working in the handful of observations where the respondent indicated they had left school and were working full-time and studying full-time.
Long term health condition	Dummy variable = 1 if respondent was recorded on HH form as having a long-term health condition; = 0 otherwise.
Log of household income	Log of household income indexed by consumer price index to 2018 dollars and averaged over the five years. If the five-year average household income was less than $1 (including negative incomes) the log of real household income was set at 0.
Sole parent family	Dummy variable = 1 if household type was lone parent; = 0 otherwise.
Parental education	Based on parent with highest level of education in two-parent households. 1 = Year 11 or below; 2 = Year 12; 3 = Certificate Level

Variable	Definition and range
	III/IV; 4 = diploma/advanced diploma; 5 = bachelor degree; 6 = postgraduate masters or doctorate.
No working parent	Dummy variable = 1 if no parent in the household works; = 0 otherwise.
Parent professional	Dummy variable = 1 if a parent's occupation is professional or managerial; = 0 otherwise.
Housing tenure of family Homeowner Rents privately Public housing Other	Series of 4 mutually exclusive dummy variables based on tenure of child's household. Homeowner includes owning or currently paying off mortgage; Public housing includes households who rent from a government housing authority or community/cooperative housing group; Rents privately includes all other renters. Public housing and private renters include rent-buy schemes. Other includes rent-free/life tenure. Homeowner is used as the default category.
Remoteness: Major city Inner regional Outer regional Remote/very remote	Series of 4 mutually exclusive dummy variables based on the Accessibility/Remoteness Index of Australia (ARIA), available as a derived variable in HILDA. Migratory, offshore, shipping and 'no usual address' excluded. Major city is used as the default category. Remote and very remote areas collapsed into one category due to the small proportion (< 2 per cent) of households in very remote areas.

Appendix 3: Full regression results for selected models

Table A3.1: Multivariate (probit) models of children's outcomes: LSAC

			SDQ total difficulties range				Overall school achievement			
	General health		Parent-assessed		Teacher-assessed		Parent-assessed		Teacher-assessed	
	β	P>z	β	P>z	β	P>z	β	P>z	β	P>z
Baby cohort	-0.126	0.000	0.092	0.018	0.026	0.422	-0.089	0.002	0.088	0.053
Gender - male	-0.088	0.000	-0.434	0.000	-0.721	0.000	-0.389	0.000	0.312	0.000
Indigenous	-0.142	0.008	-0.381	0.000	-0.407	0.000	-0.179	0.029	0.675	0.000
Non-English sp background	-0.341	0.000	-0.211	0.000	0.169	0.003	0.262	0.000	-0.151	0.040
Age (years)	-0.026	0.000	0.066	0.003	-0.107	0.000	-0.170	0.000	0.106	0.300
Age squared	0.000	0.296	-0.003	0.022	0.006	0.000	0.007	0.000	-0.008	0.180
Financial prosperity [1-5]	0.109	0.000	0.166	0.000	0.088	0.000	0.107	0.000	-0.053	0.016
Socio-economic position	0.065	0.000	0.179	0.000	0.113	0.000	0.135	0.000	-0.440	0.000
Sole parent	-0.066	0.006	-0.131	0.003	-0.219	0.000	-0.093	0.006	0.115	0.039
Homeowner	0.060	0.002	0.192	0.000	0.215	0.000	0.046	0.089	-0.162	0.000
SEIFA disadvantage	0.008	0.561	0.161	0.000	0.114	0.000	0.015	0.452	-0.091	0.006
Liveability index	0.164	0.000	0.197	0.000	0.072	0.001	0.073	0.000	-0.010	0.716
Density2	-0.014	0.528	0.013	0.762	0.102	0.012	-0.015	0.647	0.123	0.017
Observation	48,444		38,722		31,476		30,285		13,808	
Individuals	10,065		9,427		9,077		8,834		7,847	
Obs/indiv										
Minimum	1		1		1		1		1	
Average	5.0		4.1		3.5		3.4		1.8	
Maximum	7		6		6		5		3	
Wald Chi-sq	1003.8	0.000	809.5	0.000	901.2	0.000	702.6	0.000	700.5	0.000

Notes: probit model cut-points not reported. Reported β's are the estimated coefficients, not marginal effects.

Table A3.2: Multivariate regression models of children's outcomes: LSIC

	General Health		SDQ total difficulties range				Progressive assessment task scales			
			Parent-assessed		Teacher-assessed		Reading score		Maths score	
	β	P>z	β	P>z	β	P>z	β	P>z	β	P>z
Constant							-21.661	0.074	30.425	0.117
Kindergarten cohort	-0.060	0.272	0.011	0.893	0.152	0.231	-5.774	0.000	-2.932	0.000
SC is male	-0.125	0.012	-0.334	0.000	-0.813	0.000	-5.172	0.000	-1.453	0.033
SC age in years	0.010	0.765	0.046	0.403	-0.211	0.088	18.150	0.000	10.966	0.003
SC age-squared	0.000	0.927	-0.001	0.877	0.010	0.217	-0.531	0.000	-0.315	0.071
Family type										
Lives with both parents	—		—		—		—		—	
Lives with one parent	0.093	0.045	0.049	0.533	-0.337	0.008	-1.073	0.255	-0.431	0.541
Lives with carer(s)	-0.152	0.109	-0.040	0.780	-0.617	0.009	-5.154	0.002	-1.691	0.163
Parent 1 Health [1-5]	0.365	0.000	0.083	0.012	0.076	0.128				
Parenting warmth scale [1-4]	-0.245	0.000	-0.428	0.000						
Parent 1 works	0.012	0.797	0.127	0.099	-0.133	0.282	2.302	0.012	1.326	0.068
Parent 1 level of education:										
Less than Year 10	0.091	0.208	-0.155	0.131	-0.084	0.624	-3.528	0.007	-2.521	0.013
Completed Year 10	—		—		—		—		—	
Completed Year 12	-0.048	0.448	0.028	0.792	0.478	0.003	3.059	0.018	2.320	0.012
Post-school qualification	0.116	0.144	0.195	0.098	0.351	0.043	3.737	0.007	3.185	0.002
Family had money trouble			-0.290	0.000						
Housing tenure:										
Community housing	0.039	0.495	-0.148	0.142	0.168	0.349	0.318	0.787	-0.192	0.855
Public housing	—		—		—		—		—	
Private rental	0.038	0.504	0.075	0.411	0.063	0.671	2.983	0.014	0.997	0.265
Homeowner	0.192	0.012	0.375	0.002	0.191	0.299	4.274	0.004	1.345	0.202
Other	-0.063	0.463	0.028	0.863	-0.170	0.462	5.584	0.001	3.083	0.027
SEIFA deciles:										
Socio-econ. disadvantage	-0.011	0.290	0.021	0.183			0.176	0.369		

	General Health		SDQ total difficulties range				Progressive assessment task scales			
			Parent-assessed		Teacher-assessed		Reading score		Maths score	
	β	P>z	β	P>z	β	P>z	β	P>z	β	P>z
Socio-econ adv. & disadv.					0.038	0.194				
Economic resources									0.201	0.114
Liveability scale					0.182	0.006				
Remoteness level:										
None	-0.005	0.931	-0.072	0.411	-0.053	0.679	4.356	0.000	2.315	0.006
Low	—		—		—		—		—	
Moderate	-0.483	0.000	-0.105	0.405	0.033	0.869	-7.322	0.000	-4.334	0.000
High	-0.334	0.000	0.184	0.223	0.368	0.339	-4.428	0.033	-6.253	0.000
Bedreq1	0.024	0.065	0.036	0.094	0.065	0.116	-0.691	0.011	-0.732	0.000
Observations	6,852		3,442		1,768		2,749		1,835	
Individuals	1,383		1,354		946		1,257		1,156	
Obs/indiv										
Minimum	1.0		1.0		1.0		1.0		1.0	
Average	5.0		2.5		1.9		2.2		1.6	
Maximum	7.0		4.0		5.0		4.0		3.0	
R-squared							0.33		0.22	
Wald Chi-sq	406.1	0.000	147.7	0.000	89.1	0.000	1794.2	0.000	685.4	0.000

Notes: probit model cut-points not reported. Reported β's are the estimated coefficients, not marginal effects.

Table A3.3: Multivariate regression models of young adults' physical health and socio-emotional wellbeing: conditional upon circumstances when aged 5-9 years, HILDA

	Self-assessed health		SF-36 physical health		Life satisfaction		SF-36 mental health	
	β	P>z	β	P>z	β	P>z	β	P>z
Constant			44.248	0.000			72.342	0.000
Contemporary variables								
Male	0.379	0.000	0.439	0.340	0.279	0.000	4.565	0.000
Age (years)	-0.096	0.000	0.468	0.000	-0.160	0.000	-1.016	0.000
English is 2nd language	-0.721	0.296	2.062	0.303	0.141	0.844	-1.482	0.701
Indigenous	0.039	0.822	-3.008	0.024	0.394	0.037	1.562	0.172
SEIFA decile of neighbourhood								
Advantage/disadvantage [1-10]			0.004	0.966				
Education & occupation [1-10]	0.021	0.199			0.009	0.481	-0.005	0.956
Living status								
Part of couple with children	0.221	0.558	-6.582	0.089	0.489	0.169	-0.135	0.965
Part of couple, no children	-0.289	0.114	-1.590	0.209	0.371	0.037	1.058	0.351
Sole parent	0.468	0.249	5.880	0.003	0.286	0.601	4.143	0.121
Lives with both parents	—		—		—		—	
Lives with sole parent	-0.209	0.030	-0.668	0.259	-0.156	0.055	-0.514	0.405
Lone person	-0.174	0.201	-0.999	0.213	-0.149	0.158	-1.281	0.071
Lives with others	-0.400	0.018	-1.213	0.345	-0.118	0.572	-3.814	0.006
Activity status								
Still at school					—		—	
Work					0.053	0.482	2.252	0.000
Post-school study					0.062	0.392	1.904	0.000
Work and study					-0.157	0.353	0.321	0.772
Inactive					-0.136	0.249	0.001	0.999
History variables (averaged, age 5-9)								
Long term health condition	-0.867	0.000	-6.392	0.000	0.018	0.916	-3.932	0.007
Log of household income	-0.093	0.152	-0.337	0.115	-0.112	0.050	-0.733	0.003

	Self-assessed health		SF-36 physical health		Life satisfaction		SF-36 mental health	
	β	P>z	β	P>z	β	P>z	β	P>z
Sole parent family	-0.207	0.166	1.354	0.148	-0.326	0.013	-1.696	0.076
Parental education [1-6]	0.019	0.583	0.176	0.354	-0.003	0.918	-0.270	0.216
No working parent	-0.102	0.642	-1.328	0.351	-0.010	0.960	0.052	0.970
Parent professional	0.000	0.997	0.409	0.538	-0.106	0.280	0.148	0.848
Housing tenure of family								
Homeowner	—		—		—		—	
Rents privately	-0.193	0.125	-0.305	0.696	-0.205	0.038	-0.430	0.561
Public housing	-0.129	0.583	-1.520	0.283	-0.003	0.989	-0.339	0.832
Other	0.119	0.826	2.725	0.086	0.285	0.436	0.200	0.951
SEIFA decile of neighbourhood								
Advantage/disadvantage [1-10]	0.060	0.005	0.303	0.027				
Economic resources [1-10]					0.014	0.359	0.184	0.108
Remoteness								
Major city	—		—		—		—	
Inner regional	0.052	0.628	0.276	0.652	0.083	0.327	0.667	0.297
Outer regional	0.073	0.606	0.873	0.264	0.163	0.169	0.619	0.452
Remote & very remote	0.154	0.580	-0.168	0.917	-0.366	0.153	2.531	0.166
Bedreq2	0.380	0.003	0.701	0.092	0.217	0.021	2.150	0.001
Observations	3,996		3,964		4,433		3,964	
Individuals	1,173		1,168		1,215		1,168	
Obs/indiv								
Minimum	1		1		1		1	
Average	3.4		3.4		3.6		3.4	
Maximum	5		5		5		5	
R-squared			0.04				0.08	
Wald Chi-sq	166.0	0.000	95.1	0.000	189.3	0.000	213.3	0.000

Notes: probit model cut-points not reported. Reported β's are the estimated coefficients, not marginal effects.

Table A3.4: Multivariate regression models of young adults' physical health and socio-emotional wellbeing: conditional upon circumstances when aged 10-14 years, HILDA

	Self-assessed health		SF-36 physical health		Life satisfaction		SF-36 mental health	
	β	P>z	β	P>z	β	P>z	β	P>z
Constant			40.733	0.000			72.747	0.000
Contemporary variables								
Male	0.403	0.000	-0.107	0.706	0.274	0.000	4.110	0.000
Age (years)	-0.086	0.000	0.469	0.000	-0.154	0.000	-0.885	0.000
English is 2nd language	-0.511	0.027	-0.587	0.713	-0.262	0.230	-0.205	0.882
Indigenous	-0.216	0.123	-2.066	0.016	0.210	0.124	0.577	0.474
SEIFA decile of neighbourhood								
Advantage/disadvantage [1-10]			0.045	0.527				
Education & occupation [1-10]	0.018	0.203			0.027	0.003	0.134	0.075
Living status								
Part of couple with children	0.332	0.160	-2.250	0.308	0.953	0.000	1.836	0.273
Part of couple, no children	-0.082	0.483	-2.871	0.000	0.228	0.011	1.058	0.092
Sole parent	-0.104	0.676	-1.466	0.564	0.188	0.412	-4.566	0.128
Lives with both parents	—		—		—		—	
Lives with sole parent	-0.082	0.269	-0.460	0.299	-0.128	0.050	-0.388	0.386
Lone person	-0.176	0.067	-2.060	0.001	-0.152	0.035	-1.291	0.025
Lives with others	-0.275	0.013	-1.209	0.150	-0.199	0.067	-1.938	0.021
Activity status								
Still at school					—		—	
Work					0.083	0.088	1.752	0.000
Post-school study					0.059	0.216	1.724	0.000
Work and study					-0.077	0.513	1.162	0.131
Inactive					-0.123	0.094	0.270	0.603
History variables (averaged, age 5-9)								
Long term health condition	-1.097	0.000	-5.800	0.000	-0.214	0.123	-3.508	0.000
Log of household income	-0.011	0.855	0.148	0.554	0.016	0.768	-0.855	0.003

	Self-assessed health		SF-36 physical health		Life satisfaction		SF-36 mental health	
	β	P>z	β	P>z	β	P>z	β	P>z
Sole parent family	-0.188	0.073	0.117	0.847	-0.274	0.003	-0.864	0.167
Parental education [1-6]	0.012	0.605	0.015	0.907	-0.051	0.013	-0.299	0.033
No working parent	0.000	0.999	-1.465	0.066	0.123	0.323	-0.055	0.948
Parent professional	0.120	0.155	0.017	0.970	0.100	0.156	0.294	0.553
Housing tenure of family								
Homeowner	—		—		—		—	
Rents privately	-0.416	0.000	-0.405	0.431	-0.132	0.082	-2.408	0.000
Public housing	0.104	0.516	-1.146	0.225	0.098	0.520	0.149	0.876
Other	-0.107	0.747	-0.688	0.684	0.009	0.979	-0.149	0.935
SEIFA decile of neighbourhood								
Advantage/disadvantage [1-10]							0.061	0.507
Economic resources [1-10]					0.002	0.838		
Education & occupation [1-10]	0.051	0.004	0.146	0.080				
Remoteness								
Major city	—		—		—		—	
Inner regional	0.042	0.555	0.343	0.342	0.028	0.647	0.493	0.240
Outer regional	0.100	0.308	0.847	0.077	0.071	0.382	0.605	0.269
Remote & very remote	-0.005	0.979	0.339	0.765	-0.135	0.470	1.814	0.103
Bedreq2	0.213	0.020	-0.449	0.285	0.182	0.010	1.798	0.000
Observations	8,800		8,729		9,919		8,729	
Individuals	2,437		2,428		2,511		2,428	
Obs/indiv								
Minimum	1		1		1		1	
Average	3.6		3.6		4		3.6	
Maximum	5		5		5		5	
R-squared			0.04				0.07	
Wald Chi-sq	354.3	0.000	158.1	0.000	367.6	0	368.2	0

Notes: probit model cut-points not reported. Reported β's are the estimated coefficients, not marginal effects.

Table A3.5: Results of logit models of young adults' education outcomes: conditional upon circumstances when aged 10-14 years, HILDA

	Completed Year 12		Entered University	
	β	P>z	β	P>z
Constant	-0.933	0.575	-0.708	0.615
Contemporary variables				
Wave (time trend)	0.084	0.000	0.020	0.373
Male	-0.647	0.000	-0.462	0.000
English is 2nd language	1.136	0.153	1.950	0.001
Indigenous	0.101	0.731	-0.362	0.371
History variables (averaged, age 5-9)				
Long term health condition	-0.138	0.653	-1.820	0.000
Log of household income	0.068	0.648	-0.108	0.386
Sole parent family	-0.360	0.050	-0.689	0.000
Parental education [1-6]	0.265	0.000	0.365	0.000
No working parent	-0.370	0.154	0.246	0.457
Parent professional	0.787	0.000	0.369	0.035
Housing tenure of family				
Homeowner	—		—	
Rents privately	-0.713	0.000	-0.703	0.001
Public housing	-0.773	0.016	-0.443	0.297
Other	-0.006	0.992	-0.079	0.917
SEIFA decile of neighbourhood				
Education & occupation [1-10]	0.042	0.173	0.090	0.001
Remoteness				
Major city	—		—	
Inner regional	-0.333	0.041	-0.246	0.096
Outer regional	-0.587	0.004	-0.346	0.106
Remote & very remote	-0.474	0.342	-0.482	0.286
Bedreq2	-0.151	0.409	0.019	0.914
Observations	1,745		1,538	
LR chi-square	291.4	0.000	356.6	0.000
Pseudo R-squared	0.17		0.17	

Appendix 4: Robustness checks

Table A4.1: Summary of results for density variables, LSAC models with quadratic terms

	Density1			Density2			Bedreq1		
	β		P>|z|	β		P>|z|	β		P>|z|
General health [1-5]									
Linear term	-0.063		0.203	-0.050		0.342	0.002		0.900
Quadratic term	0.012		0.317	0.012		0.426	-0.001		0.777
SDQ (parent)									
Linear term	-0.005		0.958	-0.022		0.843	-0.020		0.530
Quadratic term	0.002		0.925	0.011		0.739	0.007		0.376
SDQ (Teacher)									
Linear term	0.230	**	0.011	0.316	***	0.002	0.086	***	0.010
Quadratic term	-0.039	*	0.078	-0.071	**	0.019	-0.020	**	0.032
School achieve. (parent)									
Linear term	0.002		0.977	-0.003		0.974	0.005		0.858
Quadratic term	-0.004		0.833	-0.004		0.856	-0.005		0.463
School achieve. (teacher)									
Linear term	0.238	**	0.036	0.280	**	0.030	0.082	*	0.078
Quadratic term	-0.038		0.168	-0.051		0.182	-0.002		0.902
Implied turning point									
SDQ (Teacher)	2.99			2.22			2.16		

Notes: all models estimated by ordered probit; ***, **, and * indicate the coefficient is significantly different from zero at the 1, 5 and 10 per cent level, respectively. Refer to Appendix Table A3.1 for details of control variables included. Turning points calculated only for models in which both the linear and quadratic terms are significant at the 10 per cent level or higher.

Table A4.2: Summary of results for density variables specified as a series of binary dummies, LSAC

Dependent variable and	Density1[a]			Density2[b]			Bedreq1[c]		
Range of density variable	β		P>|z|	β		P>|z|	β		P>|z|
General health [1-5]									
Low	—			—			—		
Medium	-0.001		0.960	0.005		0.759	0.005		0.808
High	0.004		0.895	0.011		0.692	0.008		0.789
Very high	-0.034		0.226	-0.035		0.195	-0.028		0.522
SDQ (parent)									
Low	—			—			—		
Medium	0.006		0.858	-0.003		0.923	0.042		0.230
High	0.042		0.418	0.029		0.568	0.019		0.731
Very high	-0.041		0.444	-0.035		0.495	-0.114		0.151
SDQ (Teacher)									
Low	—			—			—		
Medium	0.053	*	0.083	0.045		0.156	0.103	***	0.002
High	0.145	***	0.003	0.100	**	0.035	0.022		0.691
Very high	0.081		0.115	0.072		0.146	0.075		0.340
School achieve. (parent)									
Low	—			—			—		
Medium	-0.011		0.647	-0.019		0.429	0.006		0.819
High	-0.011		0.779	0.020		0.605	-0.012		0.790
Very high	0.000		0.993	-0.021		0.611	-0.066		0.278
School achieve. (teacher)									
Low	—			—			—		
Medium	0.053		0.189	0.051		0.226	0.081	*	0.066
High	0.121	*	0.059	0.131	**	0.037	0.163	**	0.028
Very high	0.153	**	0.021	0.129	**	0.044	0.267	***	0.008

Notes: all models estimated by ordered probit; ***, **, and * indicate the coefficient is significantly different from zero at the 1, 5 and 10 per cent level, respectively. Refer to Appendix Table A3.1 for details of control variables included. a. for *Density1*, Low equates to lowest 39.6 per cent of observations, Medium to next 41.5 per cent, High to next 9.4 per cent and Very high to the highest 9.6 per cent. b. Corresponding ranges for *Density2* are 33.8 per cent, 44.6 per cent, 10.4 per cent and 11.1 per cent c. for *Bedreq1* Low corresponds to no extra rooms required (72.7 per cent), Medium to 1 room required (18.3 per cent), High to 2 rooms required (5.9 per cent) and Very high to 3 or more rooms required (3.2 per cent).

Table A4.3: Summary of results for density variables, LSIC models with quadratic terms

	Density1		Density2		Bedreq1		Bedreq2	
	β	P>|z|	β	P>|z|	β	P>|z|	β	P>|z|
General health [1-5]								
Linear term	0.09	0.151	0.08	0.182	0.03	0.268	-0.02	0.634
Quadratic term	-0.01	0.270	-0.01	0.371	0.00	0.890	0.01	0.126
SDQ (Parent)								
Linear term	0.15*	0.080	0.16*	0.061	0.06	0.171	0.10*	0.056
Quadratic term	-0.02**	0.027	-0.02***	0.009	0.00	0.541	-0.01	0.129
SDQ (Teacher)								
Linear term	0.25	0.186	0.30	0.125	0.07	0.325	0.10	0.414
Quadratic term	-0.04	0.114	-0.05*	0.080	0.00	0.881	-0.02	0.592
PAT Reading								
Linear term	-1.88	0.109	-1.86	0.103	-0.31	0.537	0.24	0.682
Quadratic term	0.09	0.508	0.08	0.563	-0.06	0.379	-0.17***	0.002
PAT Maths								
Linear term	-3.79***	0.003	-4.40***	0.000	-1.06***	0.002	-1.84***	0.000
Quadratic term	0.44*	0.056	0.60***	0.003	0.05	0.192	0.13**	0.029
Implied turning point								
SDQ (Parent)	3.82		3.57					
PAT Maths	4.31		3.68				7.01	

Notes: ***, **, and * indicate the coefficient is significantly different from zero at the 1, 5 and 10 per cent level, respectively. Refer to Appendix Table A3.2 for details of control variables included in models containing *Density2*. Turning points calculated only for models in which both the linear and quadratic terms are significant at the 10 per cent level or higher.

Table A4.4: Summary of results for density variables specified as a series of dummy variables, LSIC models

	Density1[a]		Density2[b]		Bedreq1[c]		Bedreq2[d]	
	β	P>|z|	β	P>|z|	β	P>|z|	β	P>|z|
General health [1-5]								
Low	—		—		—		—	
Medium	0.03	0.449	0.02	0.722	0.06	0.181	0.03	0.534
High	0.08	0.266	0.07	0.253	0.06	0.193	0.01	0.843
Very high	0.01	0.903	0.08	0.282	0.09	0.185	0.07	0.386
SDQ (Parent)								
Low	—		—		—		—	
Medium	0.07	0.337	0.07	0.337	0.06	0.484	0.16**	0.042
High	0.01	0.932	0.05	0.657	0.11	0.190	0.01	0.959
Very high	0.25*	0.076	0.14	0.247	0.13	0.252	0.23	0.125
SDQ (Teacher)								
Low	—		—		—		—	
Medium	0.11	0.316	0.07	0.524	0.04	0.768	0.31**	0.042
High	0.12	0.581	0.18	0.364	0.21	0.123	-0.12	0.588
Very high	-0.01	0.974	-0.02	0.915	0.19	0.395	0.12	0.685
PAT Reading								
Low	—		—		—		—	
Medium	-1.71*	0.056	-2.06**	0.020	-0.94	0.345	-0.49	0.641
High	-0.34	0.833	-1.23	0.388	-2.25**	0.038	-0.69	0.648
Very high	-2.48	0.150	-2.20	0.154	-2.17	0.161	-1.28	0.481
PAT Maths								
Low	—		—		—		—	
Medium	-1.21*	0.082	-1.21*	0.078	-0.09	0.907	-1.14	0.168
High	-3.50***	0.010	-2.38**	0.047	-1.57**	0.045	-3.48***	0.007
Very high	-5.07***	0.000	-4.55***	0.000	-3.78***	0.001	-4.79***	0.001

Notes: ***, **, and * indicate the coefficient is significantly different from zero at the 1, 5 and 10 per cent level, respectively. Refer to Appendix Table A3.2 for details of control variables included in models containing *Density2*. a. for *Density1*, Low equates to lowest 45.5 per cent of observations, Medium to next 38.4 per cent, High to next 7.5 per cent and Very high to the highest 8.6 per cent. b. Corresponding ranges for *Density2* are 42.7 per cent, 36.9 per cent, 9.8 per cent and 10.6 per cent c. For *Bedreq1* Low corresponds to no extra rooms required (35.9 per cent), Medium to 1 room required (25.1 per cent), High to 2-3 rooms required (26.9 per cent) and Very high to 4 or more rooms required (12.2 per cent). d. for Bedreq2 Low corresponds to no extra rooms required (69.5 per cent), Medium to 1 room required (16.6 per cent), High to 2 rooms required (7.2 per cent) and Very high to 3 or more rooms required (6.8 per cent).

Table A4.5: Summary of results for density variables measured at age 5-9 years, HILDA models with quadratic terms

	Density1		Density2		Bedreq1		Bedreq2	
	β	P>|z|	β	P>|z|	β	P>|z|	β	P>|z|
General health [1-5]								
Linear	0.24	0.607	0.03	0.947	0.11	0.246	0.64**	0.031
Quadratic	0.05	0.700	0.14	0.363	0.01	0.486	-0.08	0.344
SF36 Physical Health								
Linear	-4.20*	0.060	-4.78**	0.036	-0.93*	0.075	-1.32	0.319
Quadratic	1.38**	0.023	1.81**	0.012	0.21***	0.005	0.63**	0.048
Life satisfaction								
Linear	-0.10	0.804	-0.13	0.729	0.04	0.640	0.37	0.113
Quadratic	0.08	0.529	0.12	0.373	0.01	0.422	-0.05	0.491
SF36 Mental Health								
Linear	-0.87	0.786	-0.27	0.926	0.80	0.166	5.65***	0.000
Quadratic	0.95	0.337	0.98	0.326	0.01	0.888	-1.09***	0.005
Implied turning point								
SF36 Physical Health	1.53		1.32		2.27			
SF36 Mental Health							2.59	

Notes: ***, **, and * indicate the coefficient is significantly different from zero at the 1, 5 and 10 per cent level, respectively. Refer to Appendix Table A3.3 for details of control variables included in models containing *Bedreq2*. Turning points calculated only for models in which both the linear and quadratic terms are significant at the 10 per cent level or higher.

Table A4.6: Summary of results for density variables measured at age 5-9 years and specified as a series of dummy variables, HILDA models

	Density1[a]		Density2[b]		Bedreq1[c]		Bedreq2[d]	
	β	P>|z|	β	P>|z|	β	P>|z|	β	P>|z|
General health [1-5]								
Low	—		—		—		—	
Medium	0.18**	0.050	0.16*	0.073	0.02	0.832	0.24	0.306
High	0.19	0.239	-0.06	0.726	-0.04	0.803	0.41**	0.023
Very high	0.45***	0.003	0.51***	0.001	0.38**	0.017		
SF36 Physical Health								
Low	—		—		—		—	
Medium	0.38	0.452	-0.07	0.888	-0.89	0.136	-2.07	0.124
High	-1.72*	0.070	-1.44	0.141	-0.54	0.529	0.34	0.677
Very high	0.18	0.823	-0.15	0.860	-1.34	0.151		
Life satisfaction								
Low	—		—		—		—	
Medium	0.04	0.647	0.00	0.968	-0.15*	0.063	-0.06	0.692
High	0.14	0.293	0.12	0.413	0.09	0.483	0.18	0.218
Very high	0.12	0.341	0.16	0.181	0.14	0.262		
SF36 Mental Health								
Low	—		—		—		—	
Medium	0.25	0.670	0.52	0.343	-0.97	0.161	1.40	0.206
High	0.25	0.800	0.58	0.600	0.60	0.533	3.04***	0.002
Very high	2.72***	0.003	2.98***	0.001	1.98**	0.033		

Notes: ***, **, and * indicate the coefficient is significantly different from zero at the 1, 5 and 10 per cent level, respectively. Refer to Appendix Table A3.3 for details of control variables included in models containing *Bedreq2*. a. for *Density1*, Low equates to lowest 39.8 per cent of observations, Medium to next 40.3 per cent, High to next 10.1 per cent and Very high to the highest 9.8 per cent. b. Corresponding ranges for *Density2* are 40.3 per cent, 40.0 per cent, 9.5 per cent and 10.2 per cent c. For *Bedreq1* 59.7 per cent (no extra rooms required), 20.6 per cent, 10.3 per cent and 9.4 per cent. d. for *Bedreq2* Low 89.74 per cent (no extra rooms required), Medium 4.2 per cent and High 6.0 per cent.

Table A4.7: Summary of results for density variables measured at age 10-14 years, HILDA models with quadratic terms

	Density1		Density2		Bedreq1		Bedreq2	
	β	P>|z|	β	P>|z|	β	P>|z|	β	P>|z|
General health [1-5]								
Linear	-0.20	0.562	-0.03	0.917	0.03	0.716	0.16	0.374
Quadratic	0.14	0.212	0.09	0.397	0.02	0.150	0.02	0.689
SF36 Physical Health								
Linear	0.66	0.703	-0.80	0.628	-0.57	0.120	-1.49	0.115
Quadratic	-0.29	0.590	0.16	0.727	0.07	0.211	0.39	0.121
Life satisfaction								
Linear	-0.27	0.370	0.01	0.969	-0.06	0.305	0.22	0.155
Quadratic	0.13	0.170	0.04	0.647	0.02**	0.029	-0.01	0.734
SF36 Mental Health								
Linear	-1.81	0.368	-1.26	0.525	0.10	0.809	1.86*	0.059
Quadratic	1.14*	0.073	0.93	0.116	0.14**	0.029	-0.02	0.926
Completed Year 12								
Linear	-0.61	0.483	-0.81	0.317	-0.15	0.317	-0.08	0.826
Quadratic	0.14	0.621	0.18	0.456	0.02	0.615	-0.02	0.835
Entered university								
Linear	-0.91	0.247	-1.52**	0.046	-0.26*	0.087	-0.52	0.197
Quadratic	0.32	0.198	0.48**	0.037	0.06*	0.061	0.18	0.145
Implied turning point								
Entered university			1.58		2.34			

Notes: ***, **, and * indicate the coefficient is significantly different from zero at the 1, 5 and 10 per cent level, respectively. Refer to Appendix Table A3.4 for details of control variables included in models containing *Bedreq2*. Turning points calculated only for models in which both the linear and quadratic terms are significant at the 10 per cent level or higher.

Table A4.8: Summary of results for density variables measured at age 10-14 years and specified as a series of dummy variables, HILDA models

	Density1[a]		Density2[b]		Bedreq1[c]		Bedreq2[d]	
	β	P>|z|	β	P>|z|	β	P>|z|	β	P>|z|
General health [1-5]								
Low	—		—		—		—	
Medium	0.09	0.161	0.04	0.516	-0.11	0.187	-0.11	0.327
High	0.06	0.536	0.08	0.452	0.00	0.968	0.21	0.127
Very high	0.14	0.195	0.21*	0.051	0.16	0.131		
SF36 Physical Health								
Low	—		—		—		—	
Medium	0.48	0.139	-0.04	0.911	-1.27***	0.004	-0.35	0.521
High	-0.19	0.740	0.27	0.654	-0.59	0.286	-1.36*	0.065
Very high	0.23	0.685	-0.13	0.816	-0.56	0.314		
Life satisfaction								
Low	—		—		—		—	
Medium	0.09*	0.089	0.15***	0.005	0.04	0.614	0.16*	0.094
High	0.01	0.874	0.04	0.706	-0.04	0.656	0.13	0.267
Very high	0.00	0.987	0.05	0.564	-0.01	0.897		
SF36 Mental Health								
Low	—		—		—		—	
Medium	0.35	0.356	0.63*	0.084	-0.51	0.324	0.03	0.957
High	0.06	0.926	0.00	0.998	-0.05	0.933	2.05***	0.006
Very high	1.05	0.105	1.66***	0.006	0.83	0.190		
Completed Year 12								
Low	—		—		—		—	
Medium	-0.17	0.292	-0.03	0.860	-0.26	0.155	-0.42*	0.055
High	-0.33	0.168	-0.27	0.256	-0.23	0.266	-0.06	0.827
Very high	-0.35	0.128	-0.21	0.345	-0.39*	0.073		
Entered university								
Low	—		—		—		—	
Medium	0.10	0.445	-0.05	0.701	-0.13	0.428	-0.17	0.468
High	-0.32	0.160	-0.09	0.698	-0.07	0.728	-0.40	0.177
Very high	-0.07	0.781	-0.17	0.473	-0.32	0.175		

Notes: ***, **, and * indicate the coefficient is significantly different from zero at the 1, 5 and 10 per cent level, respectively. Refer to Appendix Table A3.4 for details of control variables included in models containing *Bedreq2*. a. for *Density1*, Low equates to lowest 40.8 per cent of observations, Medium to next 39.4 per cent, High to next 9.4 per cent and Very high to the highest 10.4 per cent. b. Corresponding ranges for *Density2* are 40.5 per cent, 40.4 per cent, 9.2 per cent and 9.9 per cent c. For *Bedreq1* 65.7 per cent (no extra rooms required), 14.4 per cent, 10.6 per cent and 9.4 per cent. d. for *Bedreq2* 87.0 per cent (no extra rooms required), Medium 7.8 per cent and High 5.2 per cent.

The Bankwest Curtin Economics Centre is an independent economic and social research organisation located within the Curtin Business School at Curtin University. The Centre was established in 2012 through the generous support from Bankwest (a division of the Commonwealth Bank of Australia), with a core mission to undertake high quality, objective research on the key economic and social issues of relevance to Western Australia.

The Centre's research and engagement activities are designed to influence economic and social policy debates in state and Federal Parliament, regional and national media, and the wider Australian community. Through high quality, evidence-based research and analysis, our research outcomes inform policy makers and commentators of the economic challenges to achieving sustainable and equitable growth and prosperity both in Western Australia and nationally.

The Centre capitalises on Curtin University's reputation for excellence in economic modelling, forecasting, public policy research, trade and industrial economics and spatial sciences. Centre researchers have specific expertise in economic forecasting, quantitative modelling, microdata analysis and economic and social policy evaluation.

A suite of tailored and national economic models and methods are maintained within the Centre to facilitate advanced economic policy analysis: these include macroeconomic and time series models, micro(simulation) models, computable general equilibrium (CGE) models, spatial modelling methods, economic index analysis, and behavioural modelling methods.

CONTACT

Bankwest Curtin Economics Centre
Curtin University
Kent Street Bentley WA 6102
GPO Box U1987 Perth WA 6845

Tel: +61 8 9266 2873

bankwesteconomicscentre@curtin.edu.au
business.curtin.edu.au/bcec