

Nguyen, Ha Trong et al.

**Working Paper**

## Causal impact of physical activity on child health and development

GLO Discussion Paper, No. 1081

**Provided in Cooperation with:**

Global Labor Organization (GLO)

*Suggested Citation:* Nguyen, Ha Trong et al. (2022) : Causal impact of physical activity on child health and development, GLO Discussion Paper, No. 1081, Global Labor Organization (GLO), Essen

This Version is available at:

<https://hdl.handle.net/10419/251855>

**Standard-Nutzungsbedingungen:**

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

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

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

**Terms of use:**

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

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

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

# Causal impact of physical activity on child health and development

Ha Trong Nguyen<sup>\*,†</sup>

Hayley Christian<sup>†</sup>

Huong Thu Le<sup>†</sup>

Luke Connelly<sup>‡</sup>

Stephen R. Zubrick<sup>†</sup>

Francis Mitrou<sup>†</sup>

The relationship between physical activity and child health and development is well-documented, yet the extant literature provides limited causal insight into the amount of physical activity considered optimal for improving any given health or developmental outcome. This paper exploits exogenous variations in local weather conditions observed across random time use diary dates for the same individuals over time to investigate the causal impact of physical activity on a comprehensive set of health, non-cognitive development, and academic outcomes of children and adolescents. Applying an individual fixed-effects instrumental variables model to a nationally representative panel dataset from Australia, we find that physical activity leads to widespread benefits in child development. These include improved health, social and emotional development, and lower health expenditure. The results further indicate that physical activity offers greater developmental benefits for females. However, we find no evidence that physical activity improves academic performance. Our study highlights that the “optimal” amount of time that children and adolescents should spend physically active each day varies by the health or non-cognitive development outcome of interest. The results are robust to a series of specification and sensitivity tests, including an over-identification test and controlling for weather conditions recorded on the day when development outcomes were assessed.

**Keywords:** Time Allocation; Physical Activity; Time Use Diary; Health; Child Development; Instrumental variable; Panel data

**JEL classifications:** C36; I10; I12; I14; J13; J22; J24

---

\* **Corresponding author:** Telethon Kids Institute | Tel: +61 8 6319 1019 | Postal: GPO Box 855, Perth WA 6872, Australia | Email: [ha.nguyen@telethonkids.org.au](mailto:ha.nguyen@telethonkids.org.au).

† Telethon Kids Institute & The University of Western Australia.

‡ The University of Queensland & The University of Bologna.

**Acknowledgements:** This research was partly funded by the Australian Research Council Centre of Excellence for Children and Families over the Life Course (#CE140100027: Zubrick, #CE200100025: Christian) and Australian National Heart Foundation Future Leader Fellowship (#102549: Christian). This paper uses unit record data from Growing Up in Australia, the Longitudinal Study of Australian Children. The study is conducted in partnership between the Department of Social Services (DSS), the Australian Institute of Family Studies (AIFS) and the Australian Bureau of Statistics (ABS). The findings and views reported in this paper are those of the authors and should not be attributed to the DSS, the AIFS or the ABS.

## 1. Introduction

Time is a scarce resource that is extremely valuable (Becker 1965). This is particularly true for children as their time allocation among different activities may positively or negatively affect their health and developmental outcomes, which in turn can have long-lasting impacts on their later-life outcomes (Fiorini & Keane 2014; Del Boca *et al.* 2017). Because of this, estimating the magnitude of causal effects of children's time allocation on their development is clearly of strong interest to public policy makers, schools and parents. Establishing causal impacts of children's time allocation on development is challenging (Wooldridge 2010; Caetano *et al.* 2019). This is in part due to problems of unobservable individual heterogeneity correlated with both the child's time allocation and their development (such as the child's time preferences or genetic factors) and because of reverse causality (whether the child's time allocation affects development or *vice versa*).

In this paper we examine the causal effects of children and adolescents' time allocation towards physical activity on their health, non-cognitive development and academic outcomes. Our paper contributes to a large literature studying the relationship between physical activity and health (Janssen & LeBlanc 2010) and a growing literature examining the association between physical activity and academic performance (Donnelly *et al.* 2016; Erickson *et al.* 2019).<sup>1</sup> This study advances the evidence base in three important ways.

First, different from most observational studies which cannot successfully address the methodological challenges described above (i.e., the unobservable individual heterogeneity and reverse causality), this study employs a novel empirical model to estimate the causal impact

---

<sup>1</sup> Evidence from this vast physical activity literature (See review articles such as Brown *et al.* (2012); Erickson *et al.* (2019)) shows effects of physical activity vary significantly, depending on research methods (i.e., cross-sectional, prospective/panel, or randomized controlled trial), measures of physical activity (e.g., the type, amount and intensity), outcomes (e.g., health, psychosocial or cognitive outcomes) and the population considered (such as children or adults). This mixed evidence suggests a need for additional, well-designed studies. This paper also contributes to a rich literature examining the impact of the time allocated to other activities, including sleep (Gibson & Shrader 2018; Costa-Font & Fleche 2020; Bessone *et al.* 2021) or media (Gentzkow & Shapiro 2008; Nieto & Suhrcke 2021).

of physical activity.<sup>2</sup> Specifically, we exploit the exogenous variations in local weather conditions observed during random time use diary dates as a source of identification for the physical activity equation for each child and control for individual time invariant characteristics. Our identification strategy is further strengthened by the fact that, in our data, diary dates are mostly different from the dates when development outcomes are measured, allowing us to control for any potential direct impact of local weather conditions on developmental outcomes.

Second, this study presents causal evidence of the impact of physical activity on a comprehensive list of developmental outcomes. Existing studies are limited to a relatively small range of outcomes of interest (Janssen & LeBlanc 2010; Donnelly *et al.* 2016; Erickson *et al.* 2019). Focusing on a limited set of outcomes provides a partial, if not incomplete, picture of the potential impact of physical activity because individuals only have 24 hours per day and they have to choose to spend the time between physical activity and other alternative activities such as sleeping, eating, working, studying or socialising. Studies have shown that different activities may have differential effects on specific outcomes of interest (Fiorini & Keane 2014; Nguyen *et al.* 2022). A particular uncertainty in the evidence remains, however, regarding the “dose” of physical activity required to influence different child health and development outcomes.

This paper will address this evidence gap by concurrently examining the impact of physical activity on an extended set of health and developmental outcomes. Specifically, we investigate the impact of physical activity on health and non-cognitive developmental outcomes, including

---

<sup>2</sup> This paper while providing causal effects of physical activity like a small number of randomized control trials in this literature do (Erickson *et al.* 2019; Chen *et al.* 2020), offers some new insights to the literature. Randomized control trials often do not generalise very well due to the use of small samples with specific characteristics (such as children with diabetes or adults with dementia), the use of well controlled and monitored experimental settings and short temporal measurements. This paper uses “real-life” data on physical activity so evidence from this paper can be generalizable to physical activity in everyday life. This improved evidence is particularly important given that physical activity guidelines are intended to be undertaken by “normal people” in “real life”.

general health, mental health, anthropometric measures, health expenditures and general development indicators. We also consider children’s cognitive development as measured by test scores from various domains, including reading, writing, spelling, grammar, and numeracy. Many of these outcomes are of high quality because they were objectively measured, or available via linked administrative data sources and hence less prone to measurement errors. By providing evidence of the impact of physical activity on an extensive list of outcomes in one unified framework, this paper offers a more complete picture on the potential effects of physical activity than was previously possible.

Third, and specifically with reference to the “dose” effect of physical activity, this paper is the first to investigate causal evidence on the possible non-linear impact of physical activity on outcomes (i.e., due to diminishing returns). There may exist some thresholds in time spent per day physically active over which some of the identified benefits disappear or even become negative. The current literature has not been able to test this possibility (Brown *et al.* 2012; Erickson *et al.* 2019). This evidence gap is reflected in current physical activity guidelines which usually recommend the minimum time that children should spend physically active each day, essentially implying that “more is better”.<sup>3</sup> Our novel empirical method, coupled with high-quality data allow us to carefully explore the possible non-linear impact of physical activity on various child health and development outcomes, providing timely and important insights for the design of physical activity guidelines in Australia and elsewhere.

Employing a fixed-effects instrumental variables model to 16-year data from the Longitudinal Study of Australian Children (LSAC) survey, we find physical activity improves most general development and behavioural outcomes. Our results also demonstrate statistically significant

---

<sup>3</sup> Australia and most other countries have guidelines on the recommended amount, frequency and types of physical activity children and adults should do each day. For example, according to the current guidelines for individuals aged 5 to 17 years old (Department of Health 2022): “We recommend children and young people do at least 60 minutes each day of moderate to vigorous physical activity that makes the heart beat faster. More is better.”

and widespread health benefits of physical activity among children and adolescents. The results further indicate that physical activity offers greater developmental benefits to females. However, we find little evidence suggesting that physical activity fosters improvement in academic development per se. We also find that physical activity has a non-linear impact on most health and non-cognitive developmental outcomes. Based on this evidence, we calculate “optimal” hours that children and adolescents should spend each day physically active, assuming (somewhat arbitrarily) the objective is to obtain the maximum health and non-cognitive developmental benefit from physical activity. Finally, we find our results robust to a series of specification and sensitivity tests, including an over-identification test and controlling for weather conditions recorded on the same day or over the period when development outcomes were assessed.

The remainder of this paper proceeds as follows: Section 2 describes the data and Section 3 presents empirical models. Section 4 discusses the main empirical results while Section 5 reports results from numerous robustness checks. Section 6 presents additional results and Section 7 concludes.

## **2. Data**

Our main data source is the Longitudinal Study of Australian Children (LSAC) survey. The LSAC is a biennial nationally representative survey with a sampling frame of all children born between March 2003 and February 2004 (Birth or B-Cohort, 5,107 infants aged 0–1 year in 2004) and between March 1999 and February 2000 (Kindergarten or K-Cohort, 4,983 children aged 4–5 years in 2004). The LSAC was initiated in 2004 and the most recent wave 9 was surveyed in 2020 (Mohal *et al.* 2021).

### ***2.1. Measure of physical activity***

Time-use diaries (TUDs) of LSAC children and adolescents are used to construct daily time spent on physical activities. TUDs are thought to be highly accurate in quantifying time

allocation (Hamermesh 2016; Bauman *et al.* 2019). Physical activities include walking, active free play, cycling and organised sport, among other activities listed (Fiorini & Keane 2014; Nguyen *et al.* 2020). Appendix Table B1 and Appendix Table B2 provide detailed classifications of physical activities by LSAC wave and cohort.<sup>4</sup>

Like other time allocation measures derived from TUDs (Aguiar *et al.* 2013; Fiorini & Keane 2014), our physical activity measure is likely to capture physical activities regularly undertaken by children and adolescents over a longer time, despite it being only captured for a single day. This is due to several reasons. First, approximately two thirds of TUDs in LSAC were recorded on an “ordinary” day.<sup>5</sup> Second, this daily measure is correlated intertemporally (correlation of 0.15 and statistically significant at 1%, as can be seen from Appendix Table A3). Third, our TUD-based physical activity measure is also statistically significantly correlated with some other longer term measures capturing preferences toward, frequency, or intensity of physical activities available in the same dataset (see Appendix Table A3). Fourth, validation studies show that measures of physical activity constructed from TUDs are highly reliable because they are correlated with other objective measures of physical activity, including those obtained from accelerometers (Clifford *et al.* 2019).

There are four noteworthy changes to TUDs during the 16-year study period. First, from wave 1 to wave 3, families were given two TUDs to complete so each child had up to two TUDs (one on a weekday and one on a weekend day) each wave. However, from wave 4 to wave 8, each child was given only one TUD to complete, on either a weekday or a weekend day, each wave. Second, while children’s activities are reported according to the 96 15-minute periods

---

<sup>4</sup> Following prior studies (Fiorini & Keane 2014; Nguyen *et al.* 2020; Nguyen *et al.* 2022), we do not differentiate between the main and any concurrent activity. Likewise, we do not distinguish the child’s activities by who is present during each activity due to the ambiguity of the actual participation intensity of the nearby person(s) (if any) with the child. See Appendix Table A1 for variable description and summary statistics. Appendix Table A2 reports correlation structure of key variables. For examples of TUD, see Corey *et al.* (2014).

<sup>5</sup> This information is only available in the first three waves of LSAC.

of each 24-hour block in the first three waves of data, children's activities are listed in the form of an "activity episode" diary in the remaining waves (Corey *et al.* 2014). Third, while the parent completed a hard copy TUD in the first three waves, the study child was supposed to complete the TUD via computer assisted interview from wave 4 onwards. Fourth, K cohort children were asked to complete the TUD in the first six waves while B cohort children were not asked to do so in waves 4, 5 and 9. The available TUDs thus allow us to examine physical activities over a 16-year period for individuals aged from birth (for B cohort) or 4/5 years old (for K cohort) up to 15/16 years old (for both cohorts).

## **2.2. Meteorological data**

To measure local weather, as suggested by Hanigan *et al.* (2006), we assign daily weather elements from the three spatially closest weather stations to the individual's residential postcode centroid. Historical weather data from all monitoring stations throughout Australia were obtained from the Australian Bureau of Meteorology (BoM). Our spatial matching results indicate that, at the median, the first-, second-, and third-closest weather stations are within 5, 9, and 12 miles (8, 13, and 19 km), respectively, of the respondent's residential location. This close spatial distance matching ensures that the individual's physical activity is accurately linked to concurrent local weather conditions in all analyses presented.

## **2.3. Health and developmental outcomes**

We consider four main sets of developmental outcomes. The first outcome set captures the child's general development, constructed from the parent-report version of the Pediatric Quality of Life Inventory (PedsQL), which is designed for use with children and adolescents



aged 2 to 18 years (Varni *et al.* 2001). This set includes mean summary scores on Social, Emotional, and Physical development sub-scales along with a total scale score.<sup>6</sup>

The second outcome set measures child behavioural and socio-emotional development, derived from responses to the Strengths and Difficulties Questionnaire (SDQ). The SDQ contains five sub-scales: pro-social behaviour (hereafter called Pro-sociality), hyperactivity and inattention (Hyperactivity), emotional symptoms (Emotional), conduct problems (Conduct), and peer-relationship problems (Peer). Each SDQ sub-scale is scored as the summation of the item scores (from 0 (Not true), 1 (Somewhat true) to 2 (Certainly true)) on each of the five sub-items, so each has values from zero to 10. For ease of interpretation, as has been done with PedsQL measures, we rescale the SDQ measures so that higher SDQ scores indicate more desirable outcomes. We also construct an overall non-cognitive measure which is the average sum of all five sub-scales described above. To maximize the sample size and produce more reliable estimates, we use SDQ measures reported by the corresponding parent.

The third development set consists of four interviewer-administered anthropometric measures. Specifically, we use information on the child height and weight, and ages in months at the interview time, to calculate standardized gender- and age-adjusted BMI scores (Vidmar *et al.* 2013). We apply the World Health Organization (WHO) growth reference chart for children and adolescents in our data (WHO 2007). To capture the potential differential impact of physical activity on children at two tails of the BMI spectrum, we additionally use two BMI

---

<sup>6</sup> Specifically, in LSAC, the corresponding parent, also called Parent 1 and mostly the mother of the study child, was asked a series of questions, asking “In the past one month how often would you say this child has had a problem with...”. The “Social development” is the mean of responses to problems socialising with other kids, with other children not playing with study child, getting teased, unable to do what other children can, or problems keeping up with other children. The “Emotional development” sub-scale is constructed by averaging responses to problems feeling afraid or scared, feeling sad, feeling angry, sleeping, and with worrying. The “Physical development” index is calculated by taking mean of responses to problems with walking, running, sports or exercise, heavy lifting, bathing, helping to pick up toys, hurts or aches, or low energy levels. Responses are recorded as 1 Never; 2 Almost never; 3 Sometimes; 4 Often; 5 Almost always. As has been done previously (Varni *et al.* 2001), for all sub-scales, responses are rescaled so that 1=100, 2=75, 3=50, 4=25, and 5=0. We do not use two items describing “School functioning” because they were not asked in all waves. Appendix Table A4 describes LSAC contents by wave and cohort. Appendix Table B3 describes SDQ measures in more detail.

statuses: underweight and overweight, leaving normal weight as the omitted group. We further employ children's waist circumference as the fourth anthropometric measure.

The fourth outcome set includes variables describing the child's general health conditions. Specifically, we consider three separate indicators describing if the child (i) has "excellent health",<sup>7</sup> (ii) has "any ongoing condition",<sup>8</sup> or (iii) currently uses "prescribed medicine".<sup>9</sup> Moreover, we exploit a key strength of the LSAC data resource, being its linkage with administrative data from the Medicare Benefit Scheme (MBS) and the Pharmaceutical Benefit Scheme (PBS), which record all Australian Government subsidies and out-of-pocket payments for medical services and pharmaceuticals under Australia's universal and compulsory Medicare scheme, to investigate impact of physical activity on health expenditures for children. Almost all (97%) of LSAC children have information about health expenditures through MBS and PBS data observed from birth until March 2019 (Mohal *et al.* 2021). We measure annual health expenditures from MBS and PBS separately along with a combined indicator of these two expenditure types.

#### **2.4. Sample**

Our level of analysis is child-wave-diary. We exclude all diaries with obviously incorrect entries, incomplete information, incorrect diary time entries, or duplicated diary dates. These restrictions lead to a sample of around 55,000 TUDs. We further exclude a small number

---

<sup>7</sup> Particularly, this variable is constructed using responses to a question asking the corresponding parent: "In general, how would you say child current's health is: 1 Excellent; 2 Very good; 3 Good; 4 Fair; 5 Poor". We dichotomise the responses to mitigate potential reporting error issues, following prior literature (Le & Nguyen 2018). Combining "excellent" and "very good" responses produces similar results.

<sup>8</sup> Specifically, we use the corresponding parent's responses to the question "Does study child have any of these ongoing conditions? ('Ongoing conditions' exist for some period of time (weeks, months or years) or re-occur regularly. They do not have to be diagnosed by a doctor.)". The list of ongoing conditions varies by waves and includes hearing, sight or seeing properly, developmental delay, eczema, diarrhoea or colitis, anaemia, ear infections, food or digestive allergies, constipation, frequent headaches, tonsillitis, and others.

<sup>9</sup> This variable is derived from Parent 1's responses to the question "Does child currently need or use medicine prescribed by a doctor, other than vitamins?".

(around 200) of TUDs with calculated daily active time exceeding 12 hours.<sup>10</sup> We also exclude TUDs with missing information on weather (from linked BoM data) and basic explanatory variables that we control for in the regressions (see Section 3). Moreover, because our primary identification strategy relies on within-individual variation in physical activity and developmental outcomes, we focus on a subset of TUDs from individuals who have at least two TUDs during the study period. These sample restrictions lead to varying final sample sizes, which depend on which developmental outcomes are investigated. For instance, the final sample for the Social development outcome includes 45,321 complete TUDs for 8,323 unique individuals.

## **2.5. Descriptive analyses**

Children and adolescents in our data, at the median, spent roughly 2.2 hours per day physically active (See Appendix Figure A1 for distribution of physically active time and Appendix Table A1 for other relevant statistics). Summary statistics for the main characteristics and outcomes by physical activity sub-groups are presented in Table 1. Table 1 indicates that more physically active individuals (i.e., individuals with physically active time  $\geq$  median) were more likely to be younger, male, born to mothers who have higher education, or who themselves were born in Australia or born overseas in an English-Speaking-Background (ESB) country, or to live in two-parent families.

Table 1 also reveals that more physically active individuals do better in some general development or behavioural outcomes as measured by PedsQL (overall and two sub-scales: Social and Physical development), or the SDQ (two sub-scales: Emotional and Peer domains). By contrast, more physically active individuals tend to have lower scores for other behavioural outcomes such as Pro-sociality, Hyperactivity or Conduct. Table 1 additionally shows that

---

<sup>10</sup> Including these TUDs in the analysis or using other cut-offs, such as 10 or 15 hours, does not change the main findings.

more physically active individuals have lower BMI or waist circumference, or have a lower probability of being underweight or overweight. They are also more likely to have excellent (self-rated) health, and are less likely to have any ongoing health conditions or to use prescribed medicine. More physically active individuals are also found to have lower health related expenditures, as measured by receiving less MBS or PBS benefits. As might be expected, we also observe in Table 1 that individuals tend to be more physically active on warmer or drier days. However, it is important to note that the summary statistics in Table 1 are not adjusted for known observable or unobservable characteristics, and reverse causality. We address these issues directly in the following sections.

### 3. Empirical model

We employ the following model to estimate the impact of physical activity  $PA_{i\tau}$  recorded on diary date  $\tau$  on development outcome  $Y_{it}$  of individual  $i$  at time  $t$ :

$$Y_{it} = \alpha + \beta PA_{i\tau} + X_{it}\gamma + \delta_i + \mu_{it} \quad (1)$$

where  $X_{i,t}$  includes individual, household and neighbourhood characteristics.  $\delta_i$  relates to time invariant individual unobserved characteristics. We follow previous studies (Fiorini & Keane 2014; Nghiem *et al.* 2015; Le & Nguyen 2017) to include in  $X_{i,t}$  a rich set of factors contributing to the individual's development such as the individual's characteristics (e.g., age, gender, Indigenous status, low birthweight), the household's characteristics (e.g., maternal migration status, maternal education, number of siblings and living with both parents), and the characteristics of the neighbourhood.<sup>11</sup> In addition, we control for temporal or spatial

---

<sup>11</sup> These include percentages of individuals having an Aboriginal/Torres Strait Islander origin, completing year 12, speaking English, or being born in Australia in linked areas, percentages of households with household income less than AU\$1,000/week in linked areas, a metropolitan dummy. We do not include some commonly included variables such as mother's age or family income because they entail many missing values. All time invariant variables such as gender or low birthweight are dropped from FE regressions.

differences in developmental outcomes by including dummy variables for years and months of survey time and state/territory in all regressions.

In equation (1),  $\beta$  is the parameter of interest, which determines the effect of the amount of time that the individuals allocate to physical activities on their development. While the above fixed effects (FE) regression model (1) controls for time-invariant individual unobserved characteristics, it cannot deal with issues associated with reverse causality and measurement errors. As such, the estimate of  $\beta$  from equation (1) is likely to be biased. We employ the following auxiliary equation in an instrumental variables (IV) approach to investigate the amount of time allocated to PA by the individual  $i$ :

$$PA_{it} = \rho + X_{it}\sigma + Z_{it}\pi + \delta_i + \omega_{it} \quad (2)$$

where  $Z_{it}$  is a  $1 * H$  vector of instruments ( $H \geq 1$ ),  $\omega_{it}$  is an error term, and  $\rho, \sigma$  and  $\pi$  are vectors of parameters to be estimated.  $X_{it}$  and  $\delta_i$  are defined as in equation (1). To capture possible changes in physically active time throughout the week, we also include in  $X_{it}$  a series of day-of-week dummies.

We propose to use local daily maximum temperature and precipitation recorded on the diary date as two instruments in equation (2). These two instruments plausibly satisfy three conditions of good instruments (Wooldridge 2010), namely (i) they must be sufficiently correlated with  $PA_{it}$ ; (ii) they must be uncorrelated with  $Y_{it}$  except through  $PA_{it}$ ; and (iii) they cannot be correlated with time-variant, individual specific error term  $\mu_{it}$  in the development equation.

Specifically, prior studies have shown that the allocation of children's time to physical activities is particularly sensitive to both the local daily maximum temperature and

precipitation (Nguyen *et al.* 2021a; Nguyen *et al.* 2021b).<sup>12</sup> These instruments are also theoretically sound as plausibly exogenous local weather conditions directly affect individual's physical activity, but only indirectly affect their development outcomes through the physical activity channel. In particular, TUD dates in the LSAC survey are likely to be random because these dates were pre-selected by the interviewers to ensure a random distribution of weekdays and a random distribution of weekend days (Mohal *et al.* 2021). The realization of weather on these pre-selected TUD dates is also plausibly random.<sup>13</sup>

We will empirically test the strength of these instruments against criterion (iii) by controlling for some variables that are potentially associated with our instruments, such as the time spent on media activities, in Section 5. Importantly, this model is over-identified, because it has two instruments and one endogenous right-hand-side variable, thus enabling us to use a Sargan–Hansen test to formally test the external validity of instruments.

There is a concern raised in the literature regarding the use of weather conditions as instruments (see, for example, Mellon (2021) for a detailed discussion) that weather conditions may affect outcomes directly (i.e., criterion (ii)).<sup>14</sup> Our empirical model addresses this concern on four main fronts. First and foremost, as discussed above, the plausible randomness of both the diary date and weather conditions alleviate the concern that weather conditions on the diary date may directly affect developmental outcomes. Second, it is unlikely that the weather conditions recorded on one day (i.e., the diary date) exert an influence over outcomes that are measured over a much longer time horizon, such as over six months for SDQ outcomes, or one year for MBS or PBS expenditures. Even if weather conditions recorded on one day could affect

---

<sup>12</sup> We do not employ other weather elements such as wind speed or humidity as instruments because they do not statistically significantly influence children's time allocation.

<sup>13</sup> In line with a finding by Nguyen *et al.* (2021b), unreported results show that weather conditions are unrelated to the probability of a TUD being completed. These results suggest that weather conditions on diary dates are indeed exogenous.

<sup>14</sup> This source of “natural experiments” has been increasingly used to draw causal effects in other contexts (Dell *et al.* 2014; Fujiwara *et al.* 2016; Harris & Kessler 2019; Shenoy *et al.* 2021; Cantor *et al.* 2022).

developmental outcomes,<sup>15</sup> the fact that the diary dates are almost always different from the date that development outcomes were measured (hereafter called “survey dates”, which were also pre-selected), would allay this concern. In particular, and consistent with the design of LSAC, only 240 (less than 1%) out of over 45,000 TUDs were filled in on the same date as the main questionnaires, when most development outcomes were collected. Moreover, 60% of TUDs were completed before the survey date, with an average difference between the survey date and diary date of 6.6 days. The remaining 39% of TUDs were filled in roughly 1.4 days after the survey date. The difference between the dates of measured physical activity and those of measured outcomes, when viewed with evidence that physical activity is only affected by weather conditions recorded on (not before or after) the date of TUD (Nguyen *et al.* 2021b), further alleviates any concern that our instruments may affect outcomes directly.

Third, our empirical models include dummy variables for survey years, months, days of week, and states/territories, effectively controlling for temporal or spatial variations in various factors, including local weather conditions, that may directly influence developmental outcomes. Fourth, in a robustness check in Section 5, we also bolster our empirical results by controlling for weather conditions measured on the survey date or over the period that development outcomes refer to.

Because our instruments vary over time for the same individuals, we can apply an IV model to panel data in an FE-IV model, thus effectively controlling for both time-invariant and time-variant unobserved individual heterogeneity at the same time. This individual FE-IV model thus lessens the concern discussed above. Moreover, our ability to control for time-invariant unobserved individual factors is potentially important for our identification strategy as some of

---

<sup>15</sup> For instance, Graff Zivin *et al.* (2018) found that high temperature on the test day reduces test scores in math but not in reading of children in the US. However, they found little evidence of a longer-term impact of temperature on scores of both test subjects. Similarly, a more recent meta-analysis found no evidence of a significant relationship between temperature and work performance (Porrás-Salazar *et al.* 2021).

them may be simultaneously correlated with both the instruments and the outcomes. One such time-invariant unobservable factor is the individual's residential location preferences, as individuals may self-select into different climates based on their (or their parents', potentially correlated) responsiveness to weather conditions (Sinha *et al.* 2018).

We measure the time spent on physical activity on a daily basis. To have a sufficiently large sample to provide reliable estimates, we do not distinguish TUDs by whether they are recorded on weekends or weekdays.<sup>16</sup> Furthermore, we apply an Ordinary Least Squares (OLS) method to estimate equation (1) and a Two-Stage Least-Squares (2SLS) method to estimate the system of equations (1) and (2). In all regressions, robust standard errors are clustered at the individual level to account for serial correlation.

#### **4. Empirical results**

Table 2 and Table 3 report main results estimated from FE and FE-IV regressions. FE results<sup>17</sup> (reported in odd columns) show positive and statistically significant (at least at 5% level) associations between physical activity and some general development and behavioural outcomes, including Physical development, Conduct, SDQ overall, and excellent health status. Conversely, FE estimates indicate that physical activity is negatively and statistically

---

<sup>16</sup> An alternative approach is to measure the physically active time on a weekly basis (e.g., by multiplying the time from a weekday by 5 and a weekend day by 2). This approach requires that each child has at least two TUDs per wave to be included in the final sample. As previously discussed, this is only possible in the first three waves of LSAC (See Appendix Table A4 for timeline of TUDs and developmental outcomes in the LSAC).

<sup>17</sup> For comparison purposes, we also report results from regression (1) which does not control for individual fixed effects. The results, called "pooled OLS" results and reported in odd columns of Appendix Table A5 and Appendix Table A6, show a very different relationship between physical activity and developmental outcomes. For example, as compared to the FE results, pooled OLS results indicate a much more pronounced (in terms of the statistical significance and magnitude) positive relationship between physical activity and all general developmental and behavioural outcomes. Similarly, the pooled OLS results typically over-report health benefits of being physically active, including reducing BMI or health expenditures, than the FE results do. As discussed above, the pooled OLS regression does not control for individual unobservable characteristics, including genetic factors, so this may explain why the pooled OLS estimates often over-state the benefits of being physically active. Unreported test statistics from a Hausman style test confirm that FE models are always preferred to pooled OLS models, suggesting a need to control for individual unobservable characteristics when modelling the relationship between physical activity and child development. In all non-FE regressions, we do not restrict the sample to individuals with at least two TUDs as have been necessarily done with FE regressions. Nevertheless, the results are largely unchanged when we do.



correlated, at least at the 5% level, with BMI<sup>18</sup>, waist circumference, and prescribed medicine and MBS expenditures. However, FE estimates show no statistically significant correlations between physical activity and other outcomes, including Social development, Emotional development, PedsQL Overall, Hyperactivity, Peer problems, weight statuses or PBS expenditure.

The results of FE-IV estimation (reported in the even columns of Table 2 and Table 3) have three noteworthy findings.<sup>19</sup> First, the statistics from two tests reported at the bottom of the FE-IV results suggest that our instruments are empirically strong. Specifically, the lowest first-stage F statistic is 23, which is well above the rule-of-thumb value of 10 for a strong instrument (Stock & Yogo 2005).<sup>20</sup> Moreover, the statistics from a Sargan-Hansen test for over-identification restrictions from the FE-IV regressions indicate that our instruments are exogenous in 19 out of 20 cases.

---

<sup>18</sup> It is well-documented that anthropometric measures are determined by genetic and environmental factors (Polderman *et al.* 2015). Potential biological mechanisms underlying the impact of physical activity on anthropometric measures include that physical activity may cause changes in energy expenditure (Beer *et al.* 2017), growth hormone (Wideman *et al.* 2002) or bone health (Weeks *et al.* 2008). It is beyond the scope of this current study to examine potential mechanisms behind the estimated impact of physical activity on anthropometric and other developmental outcomes.

<sup>19</sup> We also report results from IV regressions in which we do not control for individual fixed effects in even columns of Appendix Table A5 and Appendix Table A6. There are three noticeable differences between IV and FE-IV results. First, F statistics for the excluded instruments in the first stage regression are always greater in IV regressions than in FE-IV regressions. Second, IV estimates of physical activity are slightly more pronounced than FE-IV estimates in terms of the statistical significance and magnitude. Third, results from a Sargan-Hansen test indicate that we can reject the null hypothesis that the instruments are exogenous in the regressions of five outcomes (i.e., Hyperactivity, Emotional problem, Peer problem, PBS expenditure, and MBS and PBS expenditure). As mentioned above, IV regressions may provide inaccurate estimates because they don't control for time invariant unobservable factors which may be associated with the instruments and outcomes at the same time.

<sup>20</sup> First-stage regression results from IV and FE-IV estimators are reported in column 1 and 2, respectively, of Appendix Table A7. The results are largely in line with those documented in other studies (Nguyen *et al.* 2021b; Nguyen *et al.* 2022). For instance, lower temperature or precipitation statistically significantly decreases the time spent on physical activity. Moreover, physically active time decreases as children age. We also observe that physically active time varies substantially across days of week. Appendix Table A8 reports estimation results of remaining variables from second-stage regressions. The results are largely as expected and akin to prior findings (Nghiem *et al.* 2015; Le & Nguyen 2017). For example, waist increases with children's ages. In addition, individuals living in two-parent households typically have better developmental outcomes while there is no clear relationship between maternal education and children's developmental outcomes. However, (diary) days of week do not explain a large proportion of the variation in most developmental outcomes.

Second, applying the FE-IV estimator substantially changes the results for Social development, PedsQL Overall and, SDQ Conduct. Specifically, the FE-IV estimator turns the estimates of physical activity on Social development and PedsQL Overall from statistically insignificant to statistically significant at the 5% and 10% levels, respectively. Thus, FE-IV results indicate a much more pronounced (in terms of the statistical significance and magnitude) benefit of being physically active on these two general development outcomes than previously observed with the FE results. By contrast, the FE-IV estimator turns the physical activity estimate on SDQ Conduct from positive and statistically significant at 1% level to negative and statistically significant at 5% level, suggesting physical activity worsens this behavioural outcome in children and adolescents.<sup>21</sup> Table 2 and Table 3 also show that changes in the magnitude and statistical significance level in the estimates of physical activity on the above-mentioned outcomes are consistent with results from a Hausman test which suggest physical activity is endogenous when modelling these outcomes. The results thus demonstrate that failing to adequately account for the endogeneity of physical activity would result in an inaccurate picture of the actual impact of physical activity on these outcomes.

Third, FE-IV estimates of physical activity on outcomes other than Social development, PedsQL Overall and SDQ Conduct are not statistically significant at any conventional level. These statistically insignificant estimates are consistent with the results from a Hausman test which indicate that we can model physical activity and these outcomes independently. Thus, the results from two Hausman-styled tests<sup>22</sup> support the use of a FE estimator to model the impact of physical activity on these outcomes.

---

<sup>21</sup> Unreported FE-IV regression results for five items of the Conduct sub-scale (see Appendix Table B3 for description of these items) show physical activity has a negative and statistically significant (at the 5% level) impact on Temper only, indicating that the negative impact of physical activity on Conduct is mainly driven by its impact on Temper.

<sup>22</sup> Specifically, the Hausman-styled test that supports the use of a FE model (over a pooled OLS model) and the one that rejects the endogeneity of physical activity in the FE-IV model.

As discussed above, the preferred FE results show that physical activity improves some general developmental and behavioural outcomes, including Physical development (the estimate is statistically significant at 5% level), Pro-sociality (10% level), Emotional symptoms (10% level), and SDQ Overall (5% level). The FE results also indicate statistically significant health benefits of being physically active, as represented by lower BMI and reduced waist circumference, increasing the probability of having excellent health, and decreasing the likelihood of having any ongoing condition or using prescribed medicine. Moreover, consistent with the finding of the statistically significant health benefits, the FE results also indicate that physical activity statistically significantly (at 1% level) decreases the individual's health expenditures. The latter result is driven mainly by lower MBS expenditures (essentially, expenditure on doctor visits). The estimates on health expenditures, while being modest in monetary terms, are quite considerable in relative terms as an increase of physical activity by one hour per day reduces annual MBS and PBS expenditures by AU\$ 2.81 (or 1.03 % of the sample mean, column 9 of Table 3). To our best knowledge, these statistically significant financial benefits of being physically active have not previously been documented in the extant literature and may represent a large fiscal saving when applied to a population. Of note, these financial benefits may be an under-estimate because MBS and PBS expenditures only include benefits covered by MBS and PBS and their associated co-payments.

## **5. Robustness checks**

This section checks whether our main findings are sensitive to: (i) a different functional form of the instruments, (ii) the exclusion or inclusion of some important time-variant variables, and (iii) the inclusion of weather conditions recorded on the dates when development outcomes were assessed.

We first test the robustness of the results to a different functional form of the temperature instrumental variable. In particular, we include daily maximum temperature and its square,

together with the daily precipitation, as instrumental variables in equation (2) to capture the potential non-linear impact of temperature on physically active time, as found in the literature (Nguyen *et al.* 2021a; Nguyen *et al.* 2021b).<sup>23</sup> The results, reported in Panel B1 of Appendix Table A9, show that using this modified specification, while not changing most of the baseline findings, which are reproduced in Panel A, tends to increase the statistical significance of FE-IV estimates for some outcomes. For example, it turns the estimates of physical activity on Emotional development (column 2) and Emotional symptoms (column 7) from statistically insignificant to statistically significant at 10% and 5% level, respectively. These changes are consistent with the result of the Hausman test which now rejects the exogeneity of physical activity in both regressions, indicating a more beneficial impact of being physically active on these emotional outcomes than previously observed. Similarly, using this new specification turns the negative estimate of physical activity on Hyperactivity (column 6) from statistically insignificant to marginally statistically significant at 10% level, suggesting an “unusual” impact of physical activity on children’s SDQ Hyperactivity score.

Our next set of sensitivity tests involves excluding or including some important time-variant variables, as suggested by Dell *et al.* (2014). We begin by excluding the list of the individual and household level explanatory variables, including number of siblings or whether the child lived with both parents at the survey time, which are potentially correlated with the instruments from FE-IV regressions. In addition, we test whether unobserved time-variant factors may account for the impact of physical activity on development by controlling for some important time-variant variables which are potentially associated with our instruments and developmental outcomes at the same time. Specifically, in light of previous findings, including a statistically

---

<sup>23</sup> We use a linear form of daily maximum temperature in the main regressions for two reasons. First, the linear specification appears to fit the data better as it produces higher F statistics. Second, in the following section, to investigate the non-linear impact of physical activity on developmental outcomes, for identification purposes, we need to introduce temperature in a quadratic form.

significant impact of local weather on children’s daily time spent on media and travel activities (Nguyen *et al.* 2021b) and some findings of beneficial impacts of media time on child development (Gentzkow & Shapiro 2008; Nieto & Suhrcke 2021), we separately include the individual’s daily time spent on media and travel as additional explanatory variables. Estimation results reported in Appendix Table A9 (Panels C1 to C3) show that either excluding or controlling for all of the above-mentioned time-varying variables does not change our results in any significant way. Results from these sensitivity checks, when viewed with results of a Sargan-Hansen test about the exogeneity of our instruments reported at the bottom of each panel in Appendix Table A9, suggest that our instruments are not correlated with time-variant unobservable characteristics in the outcome equations.

Finally, we experiment with directly including local weather conditions recorded on the survey date or over a period when development outcomes were assessed as additional explanatory variables. We consider the potential direct effects of two weather elements that we use as instruments: temperature and precipitation. Moreover, we follow Graff Zivin *et al.* (2018) to capture both contemporary and cumulative local weather impacts on developmental outcomes. To measure contemporary weather exposure, as has been done with the instruments, we use daily maximum temperature and daily total precipitation measured on the survey date when most development outcomes were assessed.<sup>24</sup> To capture potential cumulative local weather exposure, adopting previous studies (Deschênes & Greenstone 2012; Dell *et al.* 2014; Graff Zivin *et al.* 2018; Park *et al.* 2020), for each of the two weather elements, we construct two primary measures. In particular, for temperature, we calculate the average daily maximum

---

<sup>24</sup> In our sample, the correlation of daily maximum temperature measured on the diary date and that on the survey date is 0.84 and statistically significant at 1% level. Likewise, the correlation between daily total precipitation recorded on the diary date and that on the survey date is 0.31 and statistically significant at 1% level. These substantial pairwise correlations are as expected given local weather conditions, especially temperature, are highly correlated intertemporally and the relatively short time differences (e.g., less than one week as shown in Section 3) between the diary and survey dates in our data. The fact that the correlations are not unity is consistent with the previously documented pattern that survey dates and diary dates are mostly different in the data.

temperature and the number of days that temperature exceeded a given multiple of 10-degree Fahrenheit (°F) in the 365 days prior to the survey date. Likewise, we measure cumulative local precipitation exposure by the average of daily precipitation and the number of rainy days (with rainy days defined as days with precipitation  $\geq 0.1$  inch) in the 365 days preceding the survey date.

FE-IV estimates, reported in Panel E1 of Appendix Table A9,<sup>25</sup> show that additionally controlling for weather conditions on the survey date does not change our findings in terms of either statistical significance or magnitude. There are two noticeable exceptions, though. First, the FE-IV estimate of physical activity on Emotional development (column 2) becomes statistically significant at 5% level, suggesting a more pronounced beneficial impact of being physically active than previously observed in the baseline regressions. Second, by contrast, the FE-IV estimate of physical activity on Conduct (column 8) becomes statistically insignificant, indicating no negative impact of physical activity on this SDQ sub-scale. Panels E2 and E3 in Appendix Table A9 also show that the baseline findings are not sensitive to the inclusion of longer-term weather conditions which are recorded over 365 days before the survey date and represented in two alternative forms.

## **6. Additional results**

### ***6.1. Impact of physical activity on academic outcomes***

Having checked the robustness of our results, we next explore the impact of physical activity on cognitive skills. To measure cognitive skills, we employ scores from Peabody Picture

---

<sup>25</sup> Other results from these experiments suggest that the two instruments remain empirically strong as F statistics, which are slightly lower than that in the baseline analysis, from the first staged regressions are always greater than 16. Similarly, the statistics from a Sargan-Hansen test for over identification restrictions indicate that our instruments are exogenous in almost all cases. It is worth noting that we have a smaller sample size in some regressions, mainly due to missing information on weather variables measured over an extended period (i.e., 365 days). Unreported results show little evidence of a contemporary or cumulative weather impact on developmental outcomes. We also experiment with measuring weather conditions over a period when development outcomes were referred to (e.g., weather conditions in the 6 months prior to the survey time for SDQ outcomes or one month for general development outcomes) and find similar results (see Panel E4).

Vocabulary Test (PPVT), Matrix Reasoning (MR) and the National Assessment Program – Literacy and Numeracy (NAPLAN) tests. The PPVT is an interviewer-administered test to measure a child's knowledge of the meaning of spoken words for standard English (Mohal *et al.* 2021). The MR is a subtest of the Weschler Intelligence Scale for Children, which is also conducted by an interviewer, to measure a child's non-verbal visuospatial ability. PPVT was evaluated when children were 4 to 9 years of age while MR was administered when they were 6 to 11 years. The NAPLAN test is administered to all Australian students in grades 3, 5, 7 and 9 in the five domains of reading, writing, spelling, grammar and numeracy. The test scores range from 0 to 1000 and are comparable across students and over time. The NAPLAN test results were made available via data linkage with the LSAC data (Daraganova *et al.* 2013). The linked data have NAPLAN test scores from all test grades for both cohorts.

The results reported in Table 4 show that physical activity does not improve cognitive skills, and this is the case for almost all measured skills. In particular, FE results, which are preferred to FE-IV results according to the results of the Hausman test for endogeneity, show that physical activity is not statistically associated with test scores of six out of seven measured test subjects. By contrast, the preferred FE estimate shows that physical activity is negatively associated with NAPLAN spelling test scores. Likewise, the preferred FE-IV estimate for grammar indicates a negative impact of physical activity on this cognitive outcome. However, the estimates, when statistically significant, appear quantitatively small in terms of either statistical significance (i.e., the estimate is statistically significant at 10% level for spelling or 5% for grammar) or magnitude (e.g., an increase of one hour of physical activity per day is associated with a decrease of 0.22 (12.20) points, representing 0.04 (2.28) % of sample mean, in spelling (grammar) test score). Our finding of a null or relatively small negative impact of physical activity on cognitive skills is in line with that from prior studies which found that educational activities are the most productive input for academic development (Fiorini &

Keane 2014; Nguyen *et al.* 2020). These findings are consistent with the view that, given the limit of 24 hours per day, in order to increase physical activity time, individuals would have to reduce the time allocated to other activities (Becker 1965). A time allocation pattern that reduces time spent in other activities, especially education-based activities, could lead to poorer spelling or grammar test scores.<sup>26</sup>

## **6.2. *Heterogeneous impact of physical activity by gender***

To further our understanding of the physical activity effects on development, we implement a heterogeneity analysis by running separate regressions for males and females.<sup>27</sup> For this sub-population investigation, we initially use a FE-IV model for all outcomes and report the results from this model if the exogeneity of physical activity is rejected (i.e., when the  $p$  value of the Hausman test for exogeneity is greater than 0.1). When the exogeneity of physical activity is not rejected, we report results from the FE specification.

Sub-population results, reported in Table 5 and Table 6, suggest that physical activity appears to have some differential effects by gender. For example, the effects of physical activity on some general developmental and behavioural outcomes, including Physical development, PedsQL Overall, Emotional symptoms and SDQ Overall, are more pronounced for females because the estimates are typically greater or more statistically significant for them. We also

---

<sup>26</sup> Some findings on cognitive skills should be interpreted with caution since, for cognitive outcomes other than PPVT, the instruments are relatively weak, with F test statistics are slightly lower than 10, probably due to the small sample sizes. We reported the results using temperature and precipitation as joint instruments for all cognitive outcomes. IV results, whereby the explaining power of the two instruments increases substantially as F test statistics from the first staged regression range between 11 (for all NAPLAN test domains) and 23 (for PPVT), show no evidence suggesting that physical activity fosters cognitive skills (see POLS and IV results reported in Appendix Table A10). Introducing daily maximum temperature as a sole instrument improves the explaining power of the instrument in the FE-IV regressions, with F statistic of 11 for all NAPLAN test domains but does not change the findings. Similarly, using precipitation as the only instrument in the FE-IV regression of Matrix reasoning increases F statistic to 9 but does not change the findings. Because the instruments are statistically under-powered in the regressions of cognitive outcomes, we focus on non-cognitive outcomes in the remainder of this paper.

<sup>27</sup> We refrain from running separate regressions by ages mainly because activities, including physical activities, that the individuals undertook during the time diary day are listed slightly differently across waves and cohorts to reflect age-specific activities (see Appendix Table B1 and Appendix Table B2), making a comparison of sub-group results less meaningful.



observe that the impact of physical activity on Conduct is negative and statistically significant at the 5% level for females only, suggesting that the previously observed negative impact of physical activity on this behavioural outcome from the pooled sample is mainly driven by females.

The results additionally show that physical activity affects anthropometric measures of males and females very differently. Particularly, physical activity reduces the BMI of females more than that of males because the estimates are greater (in an absolute value) for females and statistically significant for females only.<sup>28</sup> Similarly, the preferred FE-IV estimate, which is positive and statistically significant at 5% level, indicates that physical activity affects the probability of being underweight for females only. The finding that physical activity increases the likelihood of being underweight for females helps explain why physical activity is more effective in reducing BMI of females than that of males. By contrast, physical activity decreases waist circumference among males more than that of females.

Sub-group results also suggest that physical activity offers more health benefits, as measured by reducing the probability of using prescribed medicine and lowering health expenditures, for females. Overall, the above sub-group analysis suggests that, except for Conduct and waist circumference, physical activity offers more health and non-cognitive developmental benefits to females. This finding when observed alongside the fact that females are less physically active than males (Farooq *et al.* 2020; Nguyen *et al.* 2022) supports the notion of delivering targeted interventions to increase physical activity among young females (Pearson *et al.* 2015).

---

<sup>28</sup> Our finding of a more pronounced impact of physical activity on BMI of young females is consistent with evidence of a greater role of environmental factors in contributing to body height of adult females (Silventoinen *et al.* 2001).

### 6.3. *Non-linear impact of physical activity*

We investigate if there is a non-linear/curvilinear relationship between physically active time and developmental outcomes. To do so, we introduce the daily physically active time variable in a quadratic form in equation (1). We initially employ a FE-IV model for all developmental outcomes. To identify this modified FE-IV model, as suggested by Wooldridge (2010), we also include both instruments (i.e., daily maximum temperature and precipitation) in a quadratic form. As the exogeneity of physical activity (and its square) is rejected (i.e., when the  $p$  value of the Hausman test for exogeneity is greater than 0.1) in almost all cases, we report results from the FE specification.<sup>29</sup>

The statistically significant estimates of the quadratic term of physically active time, reported in Appendix Table A12, suggests that physical activity has a non-linear impact on most general developmental and behavioural outcomes. Specifically, the results show that, except for Hyperactivity, the relationship between physically active time and all general developmental and behavioural outcomes follows an inverted U-shaped pattern (see Figure 1). Numerically, children's Social, Emotional, Physical and Overall PedsQL development outcomes first increase with physically active time, before starting to fall after 3.7, 3.5, 5.5, and 4.5 hours per day, respectively. Likewise, children's Pro-sociality, Emotional symptoms, Conduct,<sup>30</sup> Peer, and overall SDQ outcomes increase with physically active time up to 5.1, 4.5, 5.9, 4.4, and 4.8 hours per day, correspondingly, and starts to decrease afterwards. We also observe that children's health, as measured by having an excellent health status, improves with daily time spent on physical activity up to 6.4 hours, before starting to decline.

---

<sup>29</sup> FE-IV results from this modified model are reported in Appendix Table A11. Other test statistics suggest that the instruments are empirically strong because (i) the first-stage F-statistics are typically greater than 10, and (ii) the statistics from a Sargan-Hansen test suggest that the exogeneity of the instruments cannot be rejected in all cases.

<sup>30</sup> This new specification suggests a different relationship between physical activity and Conduct from the one estimated from the FE-IV estimator. Specifically, physical activity now improves Conduct up to the breakpoint of 5.9 hours.

By contrast, Figure 1 shows the link between daily physically active time and children's BMI, overweight status, waist circumference and MBS expenditure exhibits a U-shaped pattern. Specifically, children's BMI, overweight status, waist circumference and MBS expenditure arrive at their corresponding minimum value when time allocated to physical activity reaches 6.2, 4.8, 6.8, and 5.3 hours per day, respectively. We also observe some weak evidence (i.e., the estimate of the quadratic term of physically active time is statistically significant at 10% level) of a U-shaped relationship between physically active time and two other health outcomes: ongoing condition and prescribed medicine. Moreover, consistent with the baseline results, in these modified regressions, we do not find any statistically significant impact of physical activity on Hyperactivity, underweight status and PBS expenditure.

The above analysis, while confirming the health and non-cognitive developmental benefits of being physically active, suggests that diminishing marginal returns also set in at high levels of physical activity (see Appendix Figure A2 which represents marginal effects of physical activity). The turning points, which vary by outcomes, range between 3.5 (for Emotional development) and 6.8 hours (for waist circumference), indicate the "optimal" number of hours that children and adolescents should spend each day on physical activities to reap the maximum health and non-cognitive developmental benefits. These newly identified turning points, when observed with the distribution of physically active time that children and adolescents in our sample typically allocate each day (see Appendix Figure A1), suggest that most (up to 85%)<sup>31</sup> of them may not be physically active enough to obtain the full potential health and non-cognitive developmental benefits. This finding is consistent with that from a recent Physical Activity Report Card which finds only a small portion of Australian children and young people

---

<sup>31</sup> This figure is calculated using the mean of "optimal" hours of all outcomes with a statistically significant estimate of the quadratic term of physically active time.

meet the Australian Physical Activity Guidelines (AHKA 2018).<sup>32</sup> It is also in line with a pattern that has been referred to as a “global pandemic of physical inactivity” (Reis *et al.* 2016; WHO 2019). Unexpectedly, our results also indicate that 15% of children in our sample are probably more physically active than “optimally” desired which could be at the expense of time spent supporting other important developmental outcomes such as academic skills.

## 7. Conclusion

Applying a novel empirical model to a high-quality panel dataset, we have presented robust evidence on the causal effects of physical activity on a comprehensive list of health, non-cognitive development and academic outcomes of Australian children and adolescents. We find physical activity improves most general developmental and behavioural outcomes, including Social development, Physical development, Health related quality of life (PedsQL Overall), Pro-social behaviour, Emotional symptoms and behavioural and emotional difficulties generally (SDQ Overall). Our results also indicate statistically significant and widespread health benefits of being physically active. Identified health benefits include reduced BMI and waist circumference, increased probability of having excellent health, decreased likelihood of having any ongoing condition or using prescribed medicine, and reduced health expenditures. Finally, the results show statistically insignificant or a relatively small negative impact of physical activity on cognitive development.

Our findings have potentially important implications for both study methodologies and health policy. For instance, the results emphasize the importance of addressing potential endogeneity and consideration of the non-linear impact of physical activity on developmental outcomes. Moreover, our newly identified “optimal” hours of daily physical activity (ranging between 3.5 to 6.8 hours, depending on the health/development outcome of interest) indicate that most of

---

<sup>32</sup> TUDs do not allow us to accurately capture intensity of physical activity, making it hard to directly compare our identified “optimal” hours and those suggested in Guidelines.

children and adolescents are not physically active enough to reap the full potential health and non-cognitive developmental benefits. This finding provides higher quality evidence to support physical activity recommendations for children and adolescents and for physical activity interventions to increase physical activity, particularly among young females.

While we also found turning points for time spent in daily physical activity over which a further increase in physically active time produces no additional developmental gain and may come at the expense of other health or non-cognitive developmental outcomes, overall the findings highlight that the recommended minimum daily physical activity time for children should be reviewed, and that different thresholds exist for the optimal amount of daily physical activity children need to support various health and development outcomes.

This current paper has revealed novel information on the causal impacts of physical activity on various developmental outcomes in children and adolescents, which may be important considerations for future physical activity policy. It is beyond the scope of this paper to determine the precise mechanisms behind these identified effects. Future research should investigate hypotheses about the mechanisms through which physical activity causes changes in different child health and development outcomes. Furthermore, our measures of physical activity may not be the same as those employed in current physical activity guidelines (Department of Health 2022), limiting our ability to make a more direct comparison between our identified “optimal” daily physical activity hours and those suggested in the guidelines. To this end, more studies applying our proposed methods to data with more comparable measures of physical activity intensity, e.g., those obtained from accelerometers, may also be worthwhile.

## References

- Aguiar, M., Hurst, E., Karabarbounis, L., 2013. Time Use during the Great Recession. *American Economic Review* 103, 1664-96
- AHKA, 2018. Muscular Fitness: It's Time for a Jump Start. In: *The 2018 Active Healthy Kids Australia Report Card on Physical Activity for Children and Young People*. Active Healthy Kids Australia, Adelaide, South Australia
- Bauman, A., Bittman, M., Gershuny, J., 2019. A short history of time use research; implications for public health. *BMC Public Health* 19, 607
- Becker, G.S., 1965. A Theory of the Allocation of Time. *The Economic Journal* 75, 493-517
- Beer, N.J., Dimmock, J.A., Jackson, B.E.N., Guelfi, K.J., 2017. Providing Choice in Exercise Influences Food Intake at the Subsequent Meal. *Medicine & Science in Sports & Exercise* 49
- Bessone, P., Rao, G., Schilbach, F., Schofield, H., Toma, M., 2021. The Economic Consequences of Increasing Sleep among the Urban Poor. *The Quarterly Journal of Economics* 136, 1887–1941
- Brown, W.J., Bauman, A.E., Bull, F.C., Burton, N.W., 2012. Development of Evidence-based Physical Activity Recommendations for Adults (18-64 years). Report prepared for the Australian Government Department of Health
- Caetano, G., Kinsler, J., Teng, H., 2019. Towards causal estimates of children's time allocation on skill development. *Journal of Applied Econometrics* 34, 588-605
- Cantor, J., Sood, N., Bravata, D.M., Pera, M., Whaley, C., 2022. The impact of the COVID-19 pandemic and policy response on health care utilization: Evidence from county-level medical claims and cellphone data. *Journal of Health Economics*, 102581
- Chen, Z., Lan, W., Yang, G., Li, Y., Ji, X., Chen, L., Zhou, Y., Li, S., 2020. Exercise Intervention in Treatment of Neuropsychological Diseases: A Review. *Frontiers in Psychology* 11
- Clifford, S.A., Davies, S., Wake, M., 2019. Child Health CheckPoint: cohort summary and methodology of a physical health and biospecimen module for the Longitudinal Study of Australian Children. *BMJ Open* 9, 3
- Corey, J., Gallagher, J., Davis, E., Marquardt, M., 2014. The Times of Their Lives: Collecting time use data from children in the Longitudinal Study of Australian Children (LSAC). In: *LSAC Technical Paper No. 13*
- Costa-Font, J., Fleche, S., 2020. Child Sleep and Mother Labour Market Outcomes. *Journal of Health Economics* 69, 102258
- Daraganova, G., Edwards, B., Siphthorp, M., 2013. Using National Assessment Program—Literacy and Numeracy (NAPLAN) data in the Longitudinal Study of Australian Children (LSAC). In: *LSAC Technical Paper No. 8*, Australian Institute of Family Studies

- Del Boca, D., Monfardini, C., Nicoletti, C., 2017. Parental and child time investments and the cognitive development of adolescents. *Journal of Labor Economics* 35, 565-608
- Dell, M., Jones, B.F., Olken, B.A., 2014. What Do We Learn from the Weather? The New Climate–Economy Literature. *Journal of Economic Literature* 52, 740-798
- Department of Health, 2022. Australia's Physical Activity and Sedentary Behaviour Guidelines. URL <https://www.health.gov.au/health-topics/physical-activity-and-exercise/physical-activity-and-exercise-guidelines-for-all-australians/for-children-and-young-people-5-to-17-years>, accessed date: 17/02/2022
- Deschênes, O., Greenstone, M., 2012. The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather: reply. *The American Economic Review* 102, 3761-3773
- Donnelly, J.E., Hillman, C.H., Castelli, D., Etnier, J.L., Lee, S., Tomporowski, P., Lambourne, K., Szabo-Reed, A.N., 2016. Physical activity, fitness, cognitive function, and academic achievement in children: a systematic review. *Medicine & Science in Sports & Exercise* 48, 1197-1222
- Erickson, K.I., Hillman, C., Stillman, C.M., Ballard, R.M., Bloodgood, B., Conroy, D.E., Macko, R., Marquez, D.X., Petruzzello, S.J., Powell, K.E., 2019. Physical Activity, Cognition, and Brain Outcomes: A Review of the 2018 Physical Activity Guidelines. *Medicine & Science in Sports & Exercise* 51, 1242-1251
- Farooq, A., Martin, A., Janssen, X., Wilson, M.G., Gibson, A.-M., Hughes, A., Reilly, J.J., 2020. Longitudinal changes in moderate-to-vigorous-intensity physical activity in children and adolescents: A systematic review and meta-analysis. *Obesity Reviews* 21, e12953
- Fiorini, M., Keane, M.P., 2014. How the Allocation of Children's Time Affects Cognitive and Non-Cognitive Development. *Journal of Labor Economics* 3, 787-836
- Fujiwara, T., Meng, K., Vogl, T., 2016. Habit Formation in Voting: Evidence from Rainy Elections. *American Economic Journal: Applied Economics* 8, 160-88
- Gentzkow, M., Shapiro, J.M., 2008. Preschool Television Viewing and Adolescent Test Scores: Historical Evidence from the Coleman Study. *The Quarterly Journal of Economics* 123, 279-323
- Gibson, M., Shrader, J., 2018. Time Use and Labor Productivity: The Returns to Sleep. *The Review of Economics and Statistics* 100, 783-798
- Graff Zivin, J.S., Hsiang, S.M., Neidell, M.J., 2018. Temperature and Human Capital in the Short and Long Run. *Journal of the Association of Environmental and Resource Economists* 5, 77-105
- Hamermesh, D.S., 2016. What's to know about time use? *Journal of Economic Surveys* 30, 198–203

- Hanigan, I., Hall, G., Dear, K.B., 2006. A comparison of methods for calculating population exposure estimates of daily weather for health research. *International Journal of Health Geographics* 5, 1-16
- Harris, M.C., Kessler, L.M., 2019. Habit formation and activity persistence: Evidence from gym equipment. *Journal of Economic Behavior & Organization* 166, 688-708
- Janssen, I., LeBlanc, A.G., 2010. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *International journal of behavioral nutrition and physical activity* 7, 40
- Le, H.T., Nguyen, H.T., 2017. Parental health and children's cognitive and noncognitive development: New evidence from the longitudinal survey of Australian children. *Health Economics* 26, 1767–1788
- Le, H.T., Nguyen, H.T., 2018. The Impact of Maternal Mental Health Shocks on Child Health: Estimates from Fixed Effects Instrumental Variables Models for two Cohorts of Australian Children. *American Journal of Health Economics* 4, 185-225
- Mellon, J., 2021. Rain, Rain, Go Away: 176 potential exclusion-restriction violations for studies using weather as an instrumental variable. Available at SSRN: <http://dx.doi.org/10.2139/ssrn.3715610>
- Mohal, J., Lansangan, C., Gasser, C., Taylor, T., Renda, J., Jessup, K., Daraganova, G., 2021. Growing Up in Australia: The Longitudinal Study of Australian Children – Data User Guide, Release 9C1, June 2021. Australian Institute of Family Studies (Ed.), Melbourne
- Nghiem, H.S., Nguyen, H.T., Khanam, R., Connelly, L.B., 2015. Does School Type Affect Cognitive and Non-Cognitive Development in Children? Evidence from Australian Primary Schools. *Labour Economics* 33, 55-65
- Nguyen, H.T., Brinkman, S., Le, H.T., Zubrick, S.R., Mitrou, F., 2022. Gender differences in time allocation contribute to differences in developmental outcomes in children and adolescents. Global Labor Organization (GLO) Discussion Paper No. 1029
- Nguyen, H.T., Christian, H., Le, H.T., Connelly, L., Zubrick, S.R., Mitrou, F., 2021a. The impact of weather on time allocation to physical activity and sleep of child-parent dyads. Global Labor Organization (GLO) Discussion Paper No. 886
- Nguyen, H.T., Connelly, L., Le, H.T., Mitrou, F., Taylor, C., Zubrick, S., 2020. Ethnicity differentials in academic achievements: The role of time investments. *Journal of Population Economics* 33, 1381-1418
- Nguyen, H.T., Le, H.T., Connelly, L.B., 2021b. Weather and children's time allocation. *Health Economics* 30, 1559-1579
- Nieto, A., Suhrcke, M., 2021. The effect of TV viewing on children's obesity risk and mental well-being: Evidence from the UK digital switchover. *Journal of Health Economics* 80, 102543
- Park, R.J., Goodman, J., Hurwitz, M., Smith, J., 2020. Heat and learning. *American Economic Journal: Economic Policy* 12, 306-39



- Pearson, N., Braithwaite, R., Biddle, S.J., 2015. The effectiveness of interventions to increase physical activity among adolescent girls: a meta-analysis. *Academic pediatrics* 15, 9-18
- Polderman, T.J.C., Benyamin, B., de Leeuw, C.A., Sullivan, P.F., van Bochoven, A., Visscher, P.M., Posthuma, D., 2015. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics* 47, 702-709
- Porras-Salazar, J.A., Schiavon, S., Wargocki, P., Cheung, T., Tham, K.W., 2021. Meta-analysis of 35 studies examining the effect of indoor temperature on office work performance. *Building and Environment* 203, 108037
- Reis, R.S., Salvo, D., Ogilvie, D., Lambert, E.V., Goenka, S., Brownson, R.C., 2016. Scaling up physical activity interventions worldwide: stepping up to larger and smarter approaches to get people moving. *The Lancet* 388, 1337-1348
- Shenoy, A., Sharma, B., Xu, G., Kapoor, R., Aiden Rho, H., Sangha, K., 2021. God is in the Rain: The Impact of Rainfall-Induced Early Social Distancing on COVID-19 Outbreaks. *Journal of Health Economics*, 102575
- Silventoinen, K., Kaprio, J., Lahelma, E., Viken, R.J., Rose, R.J., 2001. Sex differences in genetic and environmental factors contributing to body-height. *Twin Research and Human Genetics* 4, 25-29
- Sinha, P., Caulkins, M.L., Cropper, M.L., 2018. Household location decisions and the value of climate amenities. *Journal of Environmental Economics and Management* 92, 608-637
- Stock, J.H., Yogo, M., 2005. Testing for Weak Instruments in Linear IV Regression. In: Andrews DWK (ed.) *Identification and Inference for Econometric Models. Identification and Inference for Econometric Models*. Cambridge University Press, New York, pp. 80-108.
- Varni, J.W., Seid, M., Kurtin, P.S., 2001. PedsQL™ 4.0: Reliability and Validity of the Pediatric Quality of Life Inventory™ Version 4.0 Generic Core Scales in Healthy and Patient Populations. *Medical Care* 39, 800-812
- Vidmar, S.I., Cole, T.J., Pan, H., 2013. Standardizing anthropometric measures in children and adolescents with functions for egen: Update. *Stata J* 13, 366-378
- Weeks, B.K., Young, C.M., Beck, B.R., 2008. Eight months of regular in-school jumping improves indices of bone strength in adolescent boys and girls: the POWER PE study. *Journal of bone and mineral research* 23, 1002-1011
- WHO, 2007. WHO child growth standards: head circumference-for-age, arm circumference-for-age, triceps skinfold-for-age and subscapular skinfold-for-age: methods and development. World Health Organization, Geneva.
- WHO, 2019. Global action plan on physical activity 2018-2030: more active people for a healthier world. World Health Organization (WHO).
- Wideman, L., Weltman, J.Y., Hartman, M.L., Veldhuis, J.D., Weltman, A., 2002. Growth hormone release during acute and chronic aerobic and resistance exercise: recent findings. *Sports Med* 32, 987-1004

Wooldridge, J.M., 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, Mass.

Table 1: Sample means of outcomes and key covariates by physically active sub-groups

Variable	More active group	Less active group	More active group - Less active group
	(1)	(2)	(3)
Child age (months)	88.480	111.144	-22.664***
Male	0.531	0.485	0.046***
Indigenous	0.022	0.020	0.001
Low birth weight	0.063	0.061	0.002
Mother has a certificate or diploma	0.380	0.407	-0.028***
Mother has a graduate degree	0.384	0.365	0.02***
Mother ESB migrant	0.104	0.091	0.012***
Mother NESB migrant	0.124	0.194	-0.07***
Number of siblings	1.493	1.479	0.015
Lived with both parents	0.863	0.815	0.048***
Social development	83.742	82.271	1.471***
Emotional development	74.568	74.420	0.148
Physical development	84.135	82.890	1.245***
PedsQL Overall	81.433	80.435	0.998***
Pro-sociality	8.154	8.231	-0.076***
Hyperactivity	6.944	7.019	-0.074***
Emotional	8.466	8.277	0.188***
Conduct	8.497	8.685	-0.187***
Peer	8.666	8.530	0.136***
SDQ Overall	8.145	8.148	-0.003
BMI	0.439	0.489	-0.05***
Underweight	0.054	0.059	-0.005**
Overweight	0.210	0.236	-0.026***
Waist (cm)	59.117	63.318	-4.201***
Excellent health	0.566	0.507	0.059***
Any ongoing condition	0.366	0.428	-0.062***
Prescribed medicine	0.127	0.152	-0.025***
MBS (\$100)	2.070	2.508	-0.438***
PBS (\$100)	0.208	0.382	-0.174*
MBS and PBS (\$100)	2.277	2.890	-0.612***
Maximum temperature ( <sup>0</sup> F)	67.363	66.829	0.534***
Precipitation (inches)	0.650	0.777	-0.127***
Number of observations	22734	22468	

Notes: Figures are sample means. Statistics are calculated using an estimated sample from the regression of “Social development” as an outcome. Tests are performed on the significance of the difference between the sample mean for “Less active” individuals (identified as those with active time < median of active time among individuals included in the final sample) and “More active” individuals (active time  $\geq$  median). The symbol \*denotes significance at the 10% level, \*\* at the 5% level, and \*\*\* at the 1% level.

Table 2: Impact of physical activity on general development and behavioural outcomes - results from FE and FE-IV models

	FE	FE-IV	FE	FE-IV	FE	FE-IV	FE	FE-IV	FE	FE-IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Social development		Emotional development		Physical development		PedsQL Overall		Pro-sociality	
Active time (hours)	-0.58	159.19**	-1.23	92.13	6.64**	75.84	2.60	106.85*	0.59*	10.53
	[3.06]	[73.38]	[2.76]	[66.22]	[2.83]	[68.94]	[2.23]	[55.57]	[0.33]	[11.66]
Observations	45,321	45,321	46,335	46,335	45,332	45,332	43,720	43,720	40,559	40,559
Individuals	8,232	8,232	8,274	8,274	8,223	8,223	8,122	8,122	7,979	7,979
Mean of dep. variable	83.01	83.01	74.30	74.30	83.34	83.34	80.95	80.95	8.16	8.16
F-statistic of IV		36.33		37.05		33.89		32.30		23.25
Hausman test (p value)		0.02		0.14		0.31		0.05		0.11
Sargan-Hansen test (p value)		0.77		0.64		0.93		0.98		0.17
	Hyperactivity		Emotional symptoms		Conduct		Peer problem		SDQ Overall	
Active time (hours)	-0.05	-17.12	0.58*	13.81	0.87***	-16.74**	0.31	4.83	0.46**	-0.61
	[0.38]	[11.08]	[0.35]	[9.64]	[0.30]	[8.23]	[0.31]	[8.68]	[0.20]	[5.64]
Observations	40,553	40,553	40,556	40,556	40,557	40,557	40,559	40,559	40,546	40,546
Individuals	7,977	7,977	7,978	7,978	7,979	7,979	7,979	7,979	7,976	7,976
Mean of dep. variable	6.95	6.95	8.35	8.35	8.56	8.56	8.57	8.57	8.12	8.12
F-statistic of IV		23.25		23.37		23.25		23.35		23.33
Hausman test (p value)		0.12		0.22		0.03		0.61		0.81
Sargan-Hansen test (p value)		0.93		0.18		0.51		0.86		0.30

Notes: FE results are from the regression (1) while FE-IV results from models (1) and (2). F-statistic of IV denotes the F statistic for the excluded instruments in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the physically active time variable in equation (1). Sargan-Hansen test indicates p value from a Sargan-Hansen over-identification test. Instruments: daily maximum temperature and daily precipitation. Other explanatory variables include age, maternal education, household size, two-parent household, local socio-economic background variables, state/territory dummies, survey year dummies, survey month dummies, and TUD day-of-week dummies. Robust standard errors clustered at the individual level in parentheses. Results (coefficient estimates and standard errors) are multiplied by 100 for aesthetic purposes. The symbol \*denotes significance at the 10% level, \*\*at the 5% level, and \*\*\*at the 1% level.

Table 3: Impact of physical activity on anthropometric and health outcomes - results from FE and FE-IV models

	FE	FE-IV	FE	FE-IV	FE	FE-IV	FE	FE-IV	FE	FE-IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	BMI		Underweight		Overweight		Waist (cm)		Excellent health	
Active time (hours)	-0.40***	-0.52	0.02	1.87	-0.11	1.10	-5.83***	-3.91	0.40***	0.46
	[0.15]	[3.76]	[0.04]	[1.22]	[0.07]	[1.76]	[1.19]	[28.92]	[0.09]	[2.34]
Observations	46,800	46,800	46,838	46,838	46,838	46,838	46,863	46,863	53,915	53,915
Individuals	8,334	8,334	8,337	8,337	8,337	8,337	8,331	8,331	8,710	8,710
Mean of dep. variable	0.46	0.46	0.06	0.06	0.22	0.22	61.07	61.07	0.55	0.55
F-statistic of IV		35.89		35.63		35.63		35.60		40.84
Hausman test (p value)		0.83		0.09		0.38		0.90		1.00
Sargan-Hansen test (p value)		0.03		0.24		0.24		0.59		0.50
	Any ongoing condition		Prescribed medicine		MBS (\$100)		PBS (\$100)		MBS and PBS (\$100)	
Active time (hours)	-0.18*	-1.99	-0.13**	-1.13	-2.07***	21.05	-0.75	-32.66	-2.81***	-11.66
	[0.11]	[2.44]	[0.06]	[1.56]	[0.73]	[18.85]	[0.73]	[33.85]	[1.06]	[40.39]
Observations	41,472	41,472	53,910	53,910	53,215	53,215	53,216	53,216	53,215	53,215
Individuals	8,115	8,115	8,710	8,710	8,555	8,555	8,555	8,555	8,555	8,555
Mean of dep. variable	0.40	0.40	0.14	0.14	2.43	2.43	0.29	0.29	2.72	2.72
F-statistic of IV		35.58		40.89		41.21		41.15		41.21
Hausman test (p value)		0.51		0.42		0.19		0.89		0.58
Sargan-Hansen test (p value)		0.16		0.30		0.87		0.34		0.42

Notes: FE results are from the regression (1) while FE-IV results from models (1) and (2). F-statistic of IV denotes the F statistic for the excluded instruments in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the physically active time variable in equation (1). Sargan-Hansen test indicates p value from a Sargan-Hansen over-identification test. Instruments: daily maximum temperature and daily precipitation. Other explanatory variables include age, maternal education, household size, two-parent household, local socio-economic background variables, state/territory dummies, survey year dummies, survey month dummies, and TUD day-of-week dummies. Robust standard errors clustered at the individual level in parentheses. Results (coefficient estimates and standard errors) are multiplied by 100 for aesthetic purposes. The symbol \*denotes significance at the 10% level, \*\*at the 5% level, and \*\*\*at the 1% level.

Table 4: Impact of physical activity on cognitive outcomes - results from FE and FE-IV models

	FE	FE-IV	FE	FE-IV	FE	FE-IV	FE	FE-IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	PPVT		Matrix reasoning		Reading		Writing	
Active time (hours)	-1.57	-24.51	0.04	28.30	-15.91	244.13	-0.53	-403.76
	[1.18]	[28.39]	[0.82]	[31.11]	[15.64]	[524.29]	[20.54]	[669.92]
Observations	21,900	21,900	14,346	14,346	18,817	18,817	18,814	18,814
Individuals	6,454	6,454	3,525	3,525	5,501	5,501	5,504	5,504
Mean of dep. variable	71.07	71.07	10.82	10.82	533.47	533.47	513.09	513.09
F-statistic of IV		11.43		4.99		8.02		7.36
Hausman test (p value)		0.69		0.31		0.65		0.50
Sargan-Hansen test (p value)		0.11		0.60		0.31		0.63
	Spelling		Grammar		Numeracy			
Active time (hours)	-21.81*	-599.76	-1.99	-1,219.62**	2.21	-688.69		
	[11.35]	[383.61]	[18.76]	[621.18]	[13.70]	[435.27]		
Observations	18,845	18,845	18,840	18,840	18,705	18,705		
Individuals	5,507	5,507	5,506	5,506	5,469	5,469		
Mean of dep. variable	522.16	522.16	535.60	535.60	532.31	532.31		
F-statistic of IV		8.58		8.58		8.86		
Hausman test (p value)		0.12		0.03		0.10		
Sargan-Hansen test (p value)		0.45		0.32		0.75		

Notes: FE results are from the regression (1) while FE-IV results from models (1) and (2). F-statistic of IV denotes the F statistic for the excluded instruments in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the physically active time variable in equation (1). Sargan-Hansen test indicates p value from a Sargan-Hansen over-identification test. Instruments: daily maximum temperature and daily precipitation. Other explanatory variables include age, maternal education, household size, two-parent household, local socio-economic background variables, state/territory dummies, survey year dummies, survey month dummies, and TUD day-of-week dummies. Robust standard errors clustered at the individual level in parentheses. Results (coefficient estimates and standard errors) are multiplied by 100 for aesthetic purposes. The symbol \*denotes significance at the 10% level, \*\*at the 5% level, and \*\*\*at the 1% level.

Table 5: Heterogeneous impact of physical activity on general development and behavioural outcomes by gender

	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Social development		Emotional development		Physical development		PedsQL Overall		Pro-sociality	
Estimator	FE	FE-IV	FE	FE	FE	FE	FE-IV	FE	FE	FE
Active time (hours)	4.24 [4.33]	144.67 [90.46]	4.51 [3.97]	-6.71* [3.82]	9.84** [4.11]	3.26 [3.90]	150.64* [86.22]	-2.88 [3.11]	0.60 [0.46]	0.53 [0.47]
Observations	22,307	23,014	22,686	23,649	22,163	23,169	21,460	22,260	19,855	20,704
Individuals	4,034	4,198	4,050	4,224	4,023	4,200	3,978	4,144	3,907	4,072
Mean of dep. variable	83.14	82.89	73.97	74.63	83.11	83.56	80.77	81.12	8.48	7.86
F-statistic of IV		23.44					15.20			
Hausman test (p value)		0.08					0.08			
Sargan-Hansen test (p value)		0.74					0.79			
	Hyperactivity		Emotional symptoms		Conduct		Peer problem		SDQ Overall	
Estimator	FE	FE	FE-IV	FE	FE-IV	FE	FE	FE	FE	FE
Active time (hours)	0.72 [0.52]	-0.74 [0.55]	43.01*** [16.62]	0.08 [0.47]	-29.22** [13.03]	0.47 [0.42]	0.20 [0.43]	0.43 [0.45]	0.81*** [0.28]	0.15 [0.28]
Observations	19,857	20,696	19,857	20,699	19,855	20,702	19,859	20,700	19,855	20,691
Individuals	3,907	4,070	3,907	4,071	3,907	4,072	3,907	4,072	3,907	4,069
Mean of dep. variable	7.45	6.47	8.24	8.45	8.70	8.43	8.69	8.46	8.31	7.93
F-statistic of IV			11.57		11.53					
Hausman test (p value)			0.01		0.01					
Sargan-Hansen test (p value)			0.32		0.37					

Notes: Results for different sub-populations are obtained from separate FE-IV or FE regressions. FE results are from the regression (1) while FE-IV results from models (1) and (2). F-statistic of IV denotes the F statistic for the excluded instruments in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the physically active time variable in equation (1). Sargan-Hansen test indicates p value from a Sargan-Hansen over-identification test. Instruments: daily maximum temperature and daily precipitation. Other explanatory variables include age, maternal education, household size, two-parent household, local socio-economic background variables, state/territory dummies, survey year dummies, survey month dummies, and TUD day-of-week dummies. Robust standard errors clustered at the individual level in parentheses. Results (coefficient estimates and standard errors) are multiplied by 100 for aesthetic purposes. The symbol \*denotes significance at the 10% level, \*\*at the 5% level, and \*\*\*at the 1% level.

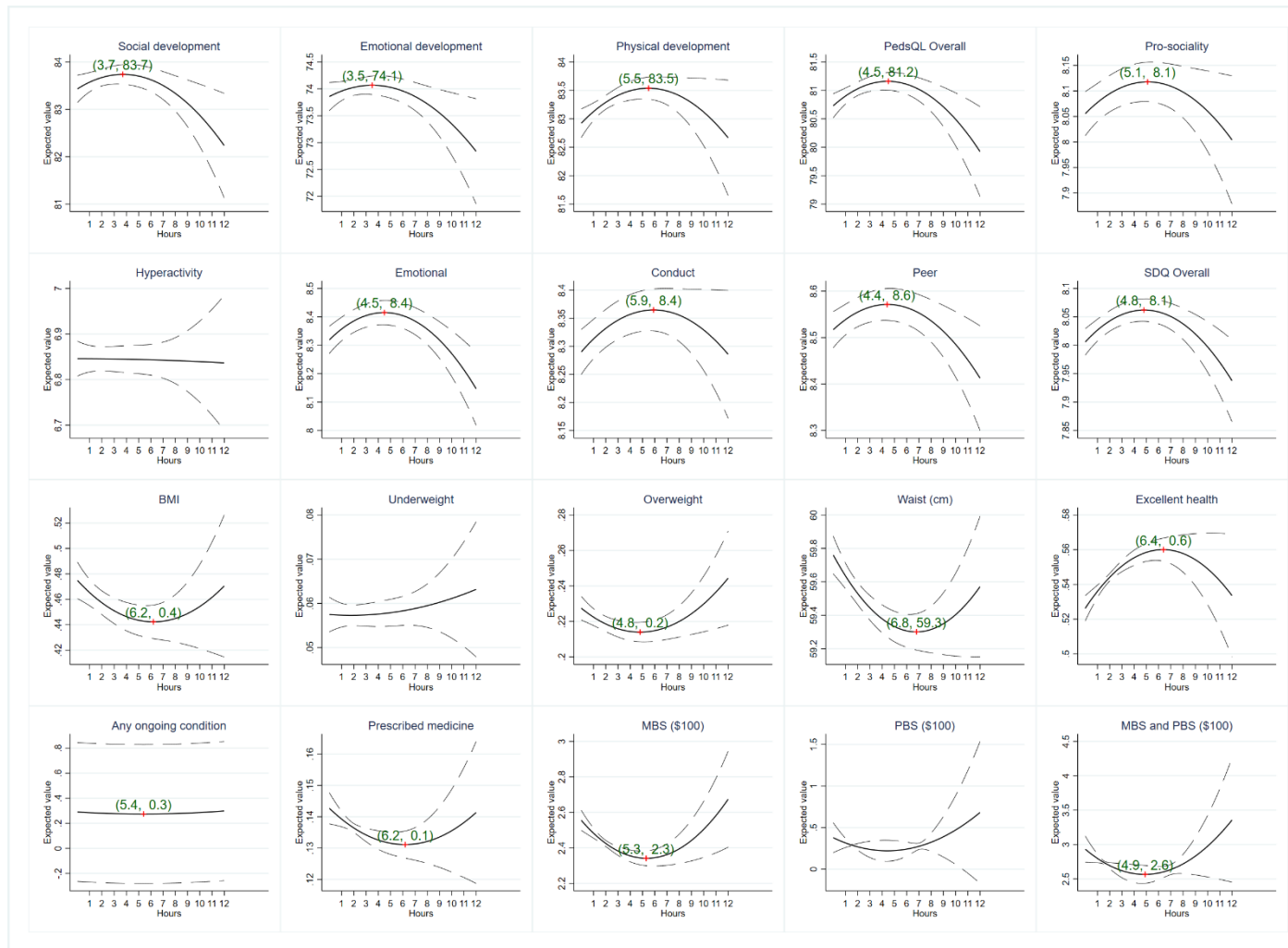
Table 6: Heterogeneous impact of physical activity on anthropometric and health outcomes by gender

	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	BMI		Underweight		Overweight		Waist (cm)		Excellent health	
Estimator	FE	FE	FE-IV	FE	FE	FE	FE	FE	FE	FE
Active time (hours)	-0.76*** [0.21]	-0.05 [0.21]	3.57** [1.81]	-0.03 [0.06]	-0.15 [0.10]	-0.07 [0.10]	-3.58** [1.62]	-7.30*** [1.72]	0.40*** [0.13]	0.40*** [0.13]
Observations	22,881	23,919	22,897	23,941	22,897	23,941	22,948	23,915	26,408	27,507
Individuals	4,076	4,258	4,076	4,261	4,076	4,261	4,078	4,253	4,267	4,443
Mean of dep. variable	0.41	0.52	0.06	0.06	0.23	0.21	60.66	61.45	0.56	0.53
F-statistic of IV			17.22							
Hausman test (p value)			0.03							
Sargan-Hansen test (p value)			0.68							
	Any ongoing condition		Prescribed medicine		MBS (\$100)		PBS (\$100)		MBS and PBS (\$100)	
Estimator	FE	FE	FE	FE	FE	FE	FE	FE	FE	FE
Active time (hours)	-0.12 [0.15]	-0.25* [0.15]	-0.27*** [0.09]	-0.01 [0.09]	-3.48*** [0.97]	-0.68 [1.08]	-0.76 [0.59]	-0.84 [1.42]	-4.24*** [1.22]	-1.52 [1.79]
Observations	20,329	21,143	26,409	27,501	26,001	27,214	26,002	27,214	26,001	27,214
Individuals	3,964	4,151	4,267	4,443	4,181	4,374	4,181	4,374	4,181	4,374
Mean of dep. variable	0.40	0.40	0.12	0.15	2.38	2.48	0.21	0.36	2.60	2.84
F-statistic of IV										
Hausman test (p value)										
Sargan-Hansen test (p value)										

Notes: Results for different sub-populations are obtained from separate FE-IV or FE regressions. FE results are from the regression (1) while FE-IV results from models (1) and (2). F-statistic of IV denotes the F statistic for the excluded instruments in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the physically active time variable in equation (1). Sargan-Hansen test indicates p value from a Sargan-Hansen over-identification test. Instruments: daily maximum temperature and daily precipitation. Other explanatory variables include age, maternal education, household size, two-parent household, local socio-economic background variables, state/territory dummies, survey year dummies, survey month dummies, and TUD day-of-week dummies. Robust standard errors clustered at the individual level in parentheses. Results (coefficient estimates and standard errors) are multiplied by 100 for aesthetic purposes. The symbol \*denotes significance at the 10% level, \*\*at the 5% level, and \*\*\*at the 1% level.



Figure 1: Non-linear impact of physical activity on non-cognitive development and health – Expected values



Notes: Results (in expected values and their 95% confidence intervals) are obtained from FE regressions. Green figures show the physically active time (the first figure in brackets) at which the expected value of the corresponding outcome reaches its maximum/minimum value (second figure). Full regression results are reported in Appendix Table A12.

## **Online Appendices**

for refereeing purposes and to be published online

Appendix Table A1: Variable description and summary statistics

Variable	Description	Mean	Min	Max	Standard deviations			N
					Overall	Between	Within	
Child age (months)	SC age at the survey time (months)	99.75	24.00	192.00	47.84	26.62	43.18	45202
Male	Dummy = 1 if SC is a male, = 0 if female	0.51	0.00	1.00	0.50	0.50	0.00	45202
Indigenous	Dummy: = 1 if SC has Aboriginal/Torres Strait Islander origin, = 0 otherwise	0.02	0.00	1.00	0.14	0.16	0.00	45202
Low birth weight	Dummy: = 1 if SC's birth weight is 2500 grams or less, = 0 otherwise	0.06	0.00	1.00	0.24	0.24	0.00	45202
Mother has a certificate	Dummy: = 1 if SC's mother has advanced diploma/diploma, = 0 otherwise	0.39	0.00	1.00	0.49	0.46	0.17	45202
Mother has a graduate degree	Dummy: = 1 if SC's mother has a bachelor degree or higher, = 0 otherwise	0.37	0.00	1.00	0.48	0.46	0.12	45202
Mother ESB migrant	Dummy: = 1 if SC's mother was born overseas in an English-Speaking Background (ESB) country, = 0 otherwise	0.10	0.00	1.00	0.30	0.29	0.02	45202
Mother NESB migrant	Dummy: = 1 if SC's mother was born overseas in a Non-ESB (NESB) country, = 0 otherwise	0.16	0.00	1.00	0.37	0.32	0.20	45202
Number of siblings	Number of siblings	1.49	0.00	11.00	0.98	0.94	0.38	45202
Lived with both parents	Dummy: = 1 if SC lived with both parents at the survey time, = 0 otherwise	0.84	0.00	1.00	0.37	0.35	0.17	45202
Social development	PedsQL social development sub-scale - P1	83.01	0.00	100.00	16.12	13.01	10.58	45202
Emotional development	PedsQL emotional development sub-scale - P1	74.49	0.00	100.00	15.40	12.53	9.67	44838
Physical development	PedsQL physical development sub-scale - P1	83.52	0.00	100.00	14.17	11.27	9.78	43966
PedsQL Overall	Mean of above three PedsQL sub-scales - P1	80.94	5.56	100.00	12.21	10.29	7.58	43672
Pro-sociality	SDQ Pro-social behaviour scale - P1	8.19	0.00	10.00	1.73	1.45	1.04	39160
Hyperactivity	SDQ Hyperactivity and inattention scale (reversed) - P1	6.98	0.00	10.00	2.27	1.99	1.19	39157
Emotional	SDQ Emotional symptoms scale (reversed) - P1	8.37	0.00	10.00	1.75	1.47	1.10	39155
Conduct	SDQ Conduct problems scale (reversed) - P1	8.59	0.00	10.00	1.61	1.34	1.00	39158
Peer	SDQ Peer-relationship problems scale (reversed) - P1	8.60	0.00	10.00	1.57	1.32	0.96	39160
SDQ Overall	Mean of above five SDQ sub-scales - P1	8.15	1.60	10.00	1.20	1.09	0.62	39152
BMI	SC's Body Mass Index (gender- and age-standardized z-scores) - ITV	0.46	-4.97	4.85	1.11	1.01	0.53	44585
Underweight	SC's gender- and age-standardized BMI is categorized as underweight, = 0 otherwise - ITV	0.06	0.00	1.00	0.23	0.18	0.15	44618
Overweight	SC's gender- and age-standardized BMI is categorized as overweight or obese, = 0 otherwise - ITV	0.22	0.00	1.00	0.42	0.35	0.25	44618
Waist (cm)	SC's waist circumference at the time of survey (cm) - ITV	61.20	30.00	180.00	11.04	7.99	8.56	44660
Excellent health	Dummy: = 1 if SC's health is in excellent condition, - 0 otherwise - P1	0.54	0.00	1.00	0.50	0.37	0.35	45181
Any ongoing condition	Dummy: = 1 if SC has any ongoing medical condition, - 0 otherwise - P1	0.40	0.00	1.00	0.49	0.36	0.35	39840
Prescribed medicine	Dummy: = 1 if SC currently uses prescribed medicine, - 0 otherwise - P1	0.14	0.00	1.00	0.35	0.26	0.24	45175
MBS (\$100)	Medicare Benefit Scheme amount during the survey year (AU\$100)	2.29	0.00	147	3.65	2.46	2.83	44612
PBS (\$100)	Pharmaceutical Benefit Scheme amount during the survey year (AU\$100)	0.29	0.00	2095	10.39	6.31	8.79	44613
MBS and PBS (\$100)	Medicare and Pharmaceutical Benefit Scheme amount during the survey year (AU\$100)	2.58	0.00	2127	11.30	7.03	9.41	44612
Maximum temperature (°F)	Daily maximum temperature on the TUD date (°F)	67.10	31.33	115.34	10.05	7.88	6.54	45202
Precipitation (inches)	Daily total precipitation on the TUD date (inches)	0.71	0.00	128.14	2.66	1.30	2.35	45202
Active time (hours)	Total time spent on physical activities per TUD day (hours)	2.66	0.00	12.00	2.30	1.14	2.05	45202

Notes: Statistics are calculated using an estimated sample from the regression of “Social development” as an outcome. English-Speaking Background (ESB) countries include UK, Ireland, Canada, New Zealand, South Africa and USA. SC refers to the Study Child. “P1” indicates Parent 1’s reported measures while “ITV” refers to the Interviewer’s.

Appendix Table A2: Correlations among physically active time and main outcomes

	Active time	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional	Conduct	Peer	SDQ Overall	BMI	Underweight	Overweight	Waist (cm)	Excellent health	Any ongoing condition	Prescribed medicine	MBS (\$100)	PBS (\$100)	MBS and PBS (\$100)	
Active time	1.00																					
Social development	0.05	1.00																				
Emotional development		0.51	1.00																			
Physical development	0.05	0.52	0.43	1.00																		
PedsQL Overall	0.05	0.81	0.76	0.85	1.00																	
Pro-sociality	-0.03	0.21	0.20	0.17	0.23	1.00																
Hyperactivity	-0.03	0.29	0.29	0.20	0.31	0.33	1.00															
Emotional	0.06	0.37	0.57	0.30	0.49	0.13	0.25	1.00														
Conduct	-0.08	0.26	0.35	0.19	0.32	0.42	0.48	0.27	1.00													
Peer	0.05	0.51	0.32	0.25	0.43	0.26	0.29	0.40	0.29	1.00												
SDQ Overall		0.48	0.51	0.33	0.52	0.63	0.75	0.60	0.72	0.64	1.00											
BMI	-0.02	-0.06	-0.02	-0.07	-0.07	0.02	-0.02	-0.03	-0.03	-0.07	-0.03	1.00										
Underweight	-0.02	-0.02		-0.02	-0.01			-0.02		-0.01		-0.50	1.00									
Overweight	-0.03	-0.07	-0.03	-0.08	-0.08	0.04	-0.01	-0.05	-0.02	-0.08	-0.03	0.72	-0.13	1.00								
Waist (cm)	-0.22	-0.13		-0.10	-0.10	0.09	0.04	-0.10	0.16	-0.08	0.03	0.52	-0.14	0.44	1.00							
Excellent health	0.07	0.19	0.19	0.24	0.26	0.11	0.13	0.21	0.11	0.18	0.22	-0.05	-0.03	-0.08	-0.11	1.00						
Any ongoing condition	-0.08	-0.15	-0.17	-0.14	-0.19	-0.03	-0.12	-0.22	-0.07	-0.13	-0.17	0.03		0.04	0.16	-0.19	1.00					
Prescribed medicine	-0.04	-0.09	-0.10	-0.10	-0.12	-0.05	-0.10	-0.13	-0.06	-0.10	-0.13	0.03		0.04	0.06	-0.20	0.27	1.00				
MBS (\$100)	-0.07	-0.13	-0.15	-0.14	-0.18	-0.04	-0.08	-0.20	-0.04	-0.13	-0.14	0.04	0.01	0.05	0.13	-0.16	0.22	0.23	1.00			
PBS (\$100)	-0.01	-0.03	-0.02	-0.02	-0.02		-0.01	-0.01	-0.02	-0.03	-0.02				0.01	-0.02	0.03	0.06	0.08	1.00		
MBS and PBS (\$100)	-0.03	-0.07	-0.07	-0.06	-0.08	-0.02	-0.04	-0.07	-0.03	-0.07	-0.06			0.01	0.05	-0.07	0.10	0.13	0.40	0.95	1.00	

Notes: Statistics are calculated using an estimated sample from the regression of “Social development” as an outcome. Only correlations which are statistically significant at 5% or lower are listed.

Appendix Table A3: Correlations among physically active measures in LSAC

Variable	Current physical activity time	Physical activity time in the following wave	Days per week exercise (30 mins)	Days per week exercise (60 mins)	Enjoys being physically active
Current physical activity time	1.00				
Physical activity time in the following wave	0.15	1.00			
Days per week exercise (30 mins)	0.19	0.15	1.00		
Days per week exercise (60 mins)	0.18	0.16	0.74	1.00	
Enjoys being physically active	0.18	0.14	0.43	0.40	1.00

Notes: All correlations are statistically significant at 0.1% level.

“Current physically active time” refers to our main indicator of physically active time and is measured at “current” survey wave while “Physically active time in the following wave” refers to the same indicator which is recorded one survey wave ahead.

“Days per week exercise (30 mins)” is derived from responses to a question, asking the study child about “How many days each week do you do at least 30 minutes of moderate or vigorous physical activity”. This question is asked in waves 7 and 8 for both cohorts.

“Days per week exercise (60 mins)” is derived from responses to a question, asking the study child about “How many days each week do you do at least 60 minutes of moderate or vigorous physical activity”. This question is asked in waves 7 and 8 for both cohorts.

“Enjoys being physically active” is derived from responses to a question asking the study child about “How much do you enjoy being physically active (doing things like sports, active games, walking or running, swimming)?”. Responses are coded as: 1 “Not at all”; 2 “Not very much”; 3 “Quite a lot”; and 4 “A lot”. This question is asked in waves 4, 5, and 6 for K cohort and waves 6, 7 and 8 for B cohort.

Appendix Table A4: LSAC contents by wave and cohort

LSAC wave	1	2	3	4	5	6	7	8	9
LSAC survey year	2004	2006	2008	2010	2012	2014	2016	2018	2020
Age									
B cohort	0/1	2/3	4/5	6/7	8/9	10/11	12/13	14/15	16/17
K cohort	4/5	6/7	8/9	10/11	12/13	14/15	16/17	18/19	20/21
TUD - P1 (wave 1 to 3) or SC (from wave 4)									
B cohort	Y	Y	Y			Y	Y	Y	
K cohort	Y	Y	Y	Y	Y	Y			
PedsQL measures - P1									
B cohort		Y	Y	Y	Y	Y	Y	Y	
K cohort	Y	Y	Y	Y	Y	Y	Y		
SDQ - P1									
B cohort			Y	Y	Y	Y	Y	Y	
K cohort	Y	Y	Y	Y	Y	Y	Y		
Weight - ITV									
B cohort	Y	Y	Y	Y	Y	Y	Y	Y	
K cohort	Y	Y	Y	Y	Y	Y	Y	Y	
Height - ITV									
B cohort		Y	Y	Y	Y	Y	Y	Y	
K cohort	Y	Y	Y	Y	Y	Y	Y	Y	
Waist circumference - ITV									
B cohort		Y	Y	Y	Y	Y	Y	Y	
K cohort	Y	Y	Y	Y	Y	Y	Y	Y	

Notes: “Y” indicates information is available in respective survey wave. PedsQL = Pediatric Quality of Life Inventory; SDQ = Strengths and Difficulties Questionnaire; P1 - reported by Parent 1; SC - reported by Study Child; ITV – assessed by Interviewer.

Appendix Table A4: LSAC contents by wave and cohort (continued)

LSAC wave	1	2	3	4	5	6	7	8	9
Excellent health - P1									
B cohort	Y	Y	Y	Y	Y	Y	Y	Y	
K cohort	Y	Y	Y	Y	Y	Y	Y	Y	
Any ongoing condition - P1									
B cohort		Y	Y	Y	Y	Y	Y	Y	
K cohort		Y	Y	Y	Y	Y	Y	Y	
Use prescribed medicine - P1									
B cohort	Y	Y	Y	Y	Y	Y	Y	Y	
K cohort	Y	Y	Y	Y	Y	Y	Y		
MBS and PBS									
B cohort	Y	Y	Y	Y	Y	Y	Y	Y	
K cohort	Y	Y	Y	Y	Y	Y	Y	Y	
PPVT - ITV									
B cohort			Y	Y	Y				
K cohort	Y	Y	Y						
MR - ITV									
B cohort				Y	Y	Y			
K cohort		Y	Y	Y					
NAPLAN test grade assigned									
B cohort				3	5	7	9		
K cohort		3	5	7	9				

Notes: “Y” indicates information is available in respective survey wave. P1 - reported by Parent 1; SC - reported by Study Child; ITV – assessed by Interviewer. PPVT = Peabody Picture Vocabulary Test; MR = Matrix Reasoning; NAPLAN = National Assessment Program – Literacy and Numeracy test score.

Appendix Table A5: Impact of physical activity on general development and behavioural outcomes - results from POLS and IV models

	POLS	IV	POLS	IV	POLS	IV	POLS	IV	POLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Social development		Emotional development		Physical development		PedsQL Overall		Pro-sociality	
Active time (hours)	19.50*** [3.88]	185.15* [104.77]	12.05*** [3.69]	229.16** [105.02]	26.20*** [3.36]	113.07 [91.11]	19.14*** [2.99]	173.37** [83.67]	1.72*** [0.45]	-0.01 [13.91]
Observations	45,703	45,700	46,704	46,701	45,735	45,733	44,145	44,143	41,055	41,052
Mean of dep. variable	82.94	82.94	74.28	74.28	83.28	83.28	80.88	80.88	8.16	8.16
F-statistic of IV		39.75		40.21		37.38		35.98		29.72
Hausman test (p value)		0.14		0.04		0.41		0.11		0.70
Sargan-Hansen test (p value)		0.54		0.59		0.58		0.34		0.19
	Hyperactivity		Emotional symptoms		Conduct		Peer problem		SDQ Overall	
Active time (hours)	1.59*** [0.57]	-7.91 [18.70]	2.37*** [0.44]	1.06 [13.65]	1.65*** [0.42]	-27.73** [12.48]	3.88*** [0.41]	2.19 [12.27]	2.23*** [0.31]	-6.37 [9.58]
Observations	41,050	41,047	41,052	41,049	41,053	41,050	41,055	41,052	41,043	41,040
Mean of dep. variable	6.94	6.94	8.34	8.34	8.56	8.56	8.56	8.56	8.11	8.11
F-statistic of IV		29.75		29.91		29.72		29.73		29.88
Hausman test (p value)		0.80		0.59		0.01		0.45		0.20
Sargan-Hansen test (p value)		0.10		0.10		0.85		0.07		0.47

Notes: POLS results are from the pooled OLS regression (1) while IV results from pooled regression models (1) and (2), without controlling for individual fixed effects. F-statistic of IV denotes the F statistic for the excluded instruments in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the physically active time variable in equation (1). Sargan-Hansen test indicates p value from a Sargan-Hansen over-identification test. Instruments: daily maximum temperature and daily precipitation. Other explanatory variables include age, gender, low birth weight, Indigenous status, maternal education, maternal migration status, household size, two-parent household, cohort dummy, local socio-economic background variables, state/territory dummies, survey year dummies, survey month dummies, and TUD day-of-week dummies. Robust standard errors clustered at the individual level in parentheses. Results (coefficient estimates and standard errors) are multiplied by 100 for aesthetic purposes. The symbol \*denotes significance at the 10% level, \*\*at the 5% level, and \*\*\*at the 1% level.



Appendix Table A6: Impact of physical activity on anthropometric and health outcomes - results from POLS and IV models

	POLS	IV	POLS	IV	POLS	IV	POLS	IV	POLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	BMI		Underweight		Overweight		Waist (cm)		Excellent health	
Active time (hours)	-0.52*	-10.89	-0.07	2.28	-0.24**	-1.63	-7.98***	-44.07	0.83***	-3.42
	[0.10]	[2.85]	[1.64]	[49.83]	[0.11]	[3.37]	[0.12]	[3.41]	[0.07]	[2.41]
Observations	47,156	47,151	47,193	47,188	47,193	47,188	47,223	47,218	54,214	54,207
Mean of dep. variable	0.46	0.46	0.06	0.06	0.22	0.22	61.12	61.12	0.54	0.54
F-statistic of IV		40.05		39.55		39.55		39.31		38.90
Hausman test (p value)		0.34		0.14		0.82		0.44		0.16
Sargan-Hansen test (p value)		0.20		0.73		0.37		0.98		0.47
	Any ongoing condition		Prescribed medicine		MBS (\$100)		PBS (\$100)		MBS and PBS (\$100)	
Active time (hours)	-0.53***	-6.44*	-0.39***	-3.69	-3.65***	-16.53	-1.81	-43.37	-5.45***	-59.94
	[0.12]	[3.41]	[0.07]	[2.41]	[1.00]	[22.97]	[1.49]	[26.82]	[1.86]	[37.97]
Observations	41,896	41,891	54,209	54,202	53,478	53,471	53,479	53,472	53,478	53,471
Mean of dep. variable	0.40	0.40	0.14	0.14	2.44	2.44	0.29	0.29	2.72	2.72
F-statistic of IV		37.88		38.75		37.23		37.14		37.23
Hausman test (p value)		0.12		0.16		0.57		0.73		0.53
Sargan-Hansen test (p value)		0.21		1.00		0.67		0.03		0.07

Notes: POLS results are from the pooled OLS regression (1) while IV results from pooled regression models (1) and (2), without controlling for individual fixed effects. F-statistic of IV denotes the F statistic for the excluded instruments in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the physically active time variable in equation (1). Sargan-Hansen test indicates p value from a Sargan-Hansen over-identification test. Instruments: daily maximum temperature and daily precipitation. Other explanatory variables include age, gender, low birth weight, Indigenous status, maternal education, maternal migration status, household size, two-parent household, cohort dummy, local socio-economic background variables, state/territory dummies, survey year dummies, survey month dummies, and TUD day-of-week dummies. Robust standard errors clustered at the individual level in parentheses. Results (coefficient estimates and standard errors) are multiplied by 100 for aesthetic purposes. The symbol \*denotes significance at the 10% level, \*\*at the 5% level, and \*\*\*at the 1% level.

Appendix Table A7: First-stage regression results

	POLS	FE
	(1)	(2)
Maximum temperature ( <sup>0</sup> F)	0.91*** [0.15]	1.24*** [0.20]
Precipitation (inches)	-1.87*** [0.33]	-1.83*** [0.37]
Child age (months)	-1.08*** [0.26]	-1.81*** [0.51]
Male	21.26*** [2.21]	
Indigenous	-1.55 [7.53]	
Low birthweight	-2.11 [4.33]	
Mother education: Certificate <sup>(a)</sup>	1.45 [2.92]	-1.99 [6.97]
Mother education: Graduate <sup>(a)</sup>	12.89*** [3.07]	-12.18 [9.12]
Mother ESB migrant <sup>(b)</sup>	8.51** [3.81]	
Mother NESB migrant <sup>(b)</sup>	-32.46*** [3.85]	
Number of siblings	7.02*** [1.13]	9.93*** [2.51]
Living with both parents	18.42*** [2.90]	-4.09 [5.90]
Tuesday <sup>(c)</sup>	-131.92*** [3.37]	-137.15*** [3.62]
Wednesday <sup>(c)</sup>	-141.03*** [3.28]	-148.01*** [3.54]
Thursday <sup>(c)</sup>	-147.57*** [3.21]	-155.27*** [3.47]
Friday <sup>(c)</sup>	-141.47*** [3.36]	-150.08*** [3.58]
Saturday <sup>(c)</sup>	-123.62*** [3.49]	-126.82*** [3.72]
Sunday <sup>(c)</sup>	24.04*** [3.69]	26.02*** [3.80]
K cohort	40.86*** [12.47]	
Observations	45,700	45,321
Number of unique individuals		8,232

Notes: POLS results are from the first stage of pooled IV regression of “Social development” as an outcome while FE results from the FE-IV regression. Coefficient estimates and standard errors are multiplied by 100 for aesthetic purposes. <sup>(a)</sup>, <sup>(b)</sup> and <sup>(c)</sup> denotes having year 12 or below qualification, Australian born mother, and Monday as the base group, respectively. Other included variables include local socio-economic background variables, state/territory dummies, survey year dummies and survey month dummies. Robust standard errors clustered at the individual level in parentheses. The symbol \*denotes significance at the 10% level, \*\*at the 5% level, and \*\*\*at the 1% level.

Appendix Table A8: Second-stage regression results for remaining variables

	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional symptoms	Conduct	Peer problem	SDQ Overall
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Child age (months)	-3.34 [2.82]	4.69* [2.43]	1.34 [1.92]	0.78 [1.96]	0.53 [0.33]	0.15 [0.26]	0.15 [0.34]	0.81*** [0.24]	0.24 [0.24]	0.39*** [0.15]
Mother education: Certificate <sup>(a)</sup>	-194.12*** [58.81]	-93.45* [54.75]	-119.99** [52.11]	-135.13*** [43.23]	-6.12 [6.61]	2.31 [7.72]	-12.09* [6.91]	-2.08 [6.49]	-14.97*** [5.70]	-6.60* [3.95]
Mother education: Graduate <sup>(a)</sup>	-241.93*** [78.42]	-192.52** [75.31]	-19.32 [71.04]	-145.84** [60.43]	-7.37 [8.96]	6.00 [10.54]	-15.43 [9.73]	-9.74 [8.16]	-13.14 [8.16]	-7.90 [5.48]
Number of siblings	-78.50*** [22.82]	-27.32 [21.64]	58.96*** [20.52]	-3.69 [17.57]	-7.24*** [2.60]	8.17*** [3.02]	4.25 [2.65]	-2.61 [2.43]	-2.18 [2.44]	0.05 [1.58]
Living with both biological parents	247.58*** [51.23]	364.28*** [48.46]	168.73*** [46.14]	249.60*** [38.01]	22.81*** [5.63]	16.64** [6.67]	31.80*** [6.20]	3.83 [5.29]	14.73*** [5.01]	17.98*** [3.44]
Tuesday <sup>(b)</sup>	243.52** [103.37]	131.41 [92.88]	136.61 [95.48]	165.54** [77.23]	15.56 [16.42]	-22.81 [15.63]	21.54 [13.69]	-21.66* [11.68]	9.43 [12.30]	0.89 [7.96]
Wednesday <sup>(b)</sup>	234.28** [110.30]	120.22 [100.12]	112.74 [104.27]	149.36* [83.28]	15.70 [17.86]	-24.25 [16.96]	22.11 [14.74]	-24.29* [12.63]	6.16 [13.33]	-0.40 [8.61]
Thursday <sup>(b)</sup>	256.37** [115.87]	109.43 [104.44]	100.54 [108.73]	147.49* [87.05]	18.82 [18.66]	-25.06 [17.67]	22.52 [15.42]	-26.81** [13.11]	7.07 [13.94]	-0.16 [9.01]
Friday <sup>(b)</sup>	236.72** [112.18]	107.54 [101.91]	123.02 [105.94]	152.19* [85.06]	17.56 [18.39]	-25.25 [17.55]	21.51 [15.27]	-26.11** [13.07]	9.71 [13.77]	0.00 [8.94]
Saturday <sup>(b)</sup>	230.93** [95.94]	136.61 [86.62]	117.36 [89.19]	153.21** [72.18]	16.51 [15.61]	-19.39 [14.76]	18.67 [12.90]	-16.37 [11.05]	5.63 [11.75]	1.43 [7.54]
Sunday <sup>(b)</sup>	-28.05 [27.01]	-20.57 [23.37]	-1.00 [23.68]	-22.34 [20.16]	-1.02 [4.26]	7.46* [4.39]	-3.62 [3.85]	6.22* [3.29]	-1.18 [3.36]	1.48 [2.20]
Observations	45,321	46,335	45,332	43,720	40,559	40,553	40,556	40,557	40,559	40,546
Individuals	8,232	8,274	8,223	8,122	7,979	7,977	7,978	7,979	7,979	7,976

Notes: Results are from the second stage of FE-IV regression. Coefficient estimates and standard errors are multiplied by 100 for aesthetic purposes. <sup>(a)</sup> and <sup>(b)</sup> denotes having year 12 or below qualification and Monday as the base group, respectively. Other included variables include local socio-economic background variables, state/territory dummies, survey year dummies, survey month dummies. Robust standard errors clustered at the individual level in parentheses. The symbol \*denotes significance at the 10% level, \*\*at the 5% level, and \*\*\*at the 1% level.

Appendix Table A8: Second-stage regression results (continued)

	BMI	Underweight	Overweight	Waist (cm)	Excellent health	Any ongoing condition	Prescribed medicine	MBS (\$100)	PBS (\$100)	MBS and PBS (\$100)
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Child age (months)	0.09 [0.14]	-0.02 [0.04]	0.07 [0.07]	13.99*** [0.97]	-0.12 [0.08]	0.32 [0.71]	-0.01 [0.04]	-1.10* [0.57]	-0.99 [0.97]	-2.09* [1.16]
Mother education: Certificate <sup>(a)</sup>	9.11*** [3.38]	-0.30 [0.78]	2.79* [1.44]	96.88*** [27.95]	-1.98 [1.74]	2.75 [1.96]	1.38 [1.21]	23.59* [12.13]	-5.37 [25.25]	18.22 [28.09]
Mother education: Graduate <sup>(a)</sup>	4.50 [4.48]	0.59 [1.10]	1.04 [2.01]	67.74* [37.61]	-3.24 [2.33]	-0.18 [2.63]	2.12 [1.62]	50.83*** [18.20]	2.60 [28.14]	53.42 [33.96]
Number of siblings	-5.42*** [1.28]	-0.05 [0.32]	-1.88*** [0.56]	-48.11*** [10.17]	0.19 [0.60]	-0.21 [0.76]	0.03 [0.42]	-17.56*** [4.33]	22.59 [16.43]	5.03 [17.25]
Living with both biological parents	-7.27** [2.83]	0.05 [0.73]	-4.31*** [1.25]	-78.58*** [22.28]	4.84*** [1.45]	-2.87* [1.63]	-0.50 [0.95]	-26.39** [12.21]	21.94 [20.76]	-4.45 [24.29]
Tuesday <sup>(b)</sup>	-2.25 [5.29]	3.12* [1.71]	1.28 [2.47]	-14.18 [40.70]	0.72 [2.84]	-2.64 [3.61]	-1.40 [1.89]	27.52 [23.07]	-34.19 [36.86]	-6.73 [45.44]
Wednesday <sup>(b)</sup>	0.18 [5.74]	2.56 [1.84]	0.74 [2.67]	-5.74 [44.09]	1.09 [3.28]	-1.72 [3.56]	-2.09 [2.18]	32.55 [25.60]	-18.16 [22.42]	14.33 [35.79]
Thursday <sup>(b)</sup>	-1.48 [5.97]	3.19* [1.93]	1.00 [2.78]	-9.16 [45.43]	1.30 [3.37]	-2.70 [3.68]	-2.17 [2.25]	35.41 [27.45]	-49.28 [51.70]	-13.94 [60.97]
Friday <sup>(b)</sup>	-3.32 [5.79]	3.32* [1.87]	0.84 [2.70]	-13.05 [44.37]	0.56 [3.28]	-2.40 [3.47]	-1.53 [2.20]	23.74 [26.90]	-36.01 [41.90]	-12.33 [52.09]
Saturday <sup>(b)</sup>	-2.26 [4.90]	2.85* [1.58]	0.62 [2.29]	-20.90 [37.58]	1.30 [2.86]	-0.89 [3.00]	-1.27 [1.91]	29.32 [23.23]	-37.73 [36.87]	-8.47 [45.53]
Sunday <sup>(b)</sup>	-0.74 [1.27]	-0.40 [0.38]	-0.84 [0.60]	-14.24 [9.70]	-0.31 [0.70]	1.50 [1.03]	-0.29 [0.47]	-4.03 [4.97]	10.90 [10.50]	6.87 [11.81]
Observations	46,800	46,838	46,838	46,863	53,915	41,472	53,910	53,215	53,216	53,215
Number of unique individuals	8,334	8,337	8,337	8,331	8,710	8,115	8,710	8,555	8,555	8,555

Appendix Table A9: Robustness checks

	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional symptoms	Conduct	Peer problem	SDQ Overall	BMI	Underweight	Overweight	Waist (cm)	Excellent health	Any ongoing condition	Prescribed medicine	MBS (\$100)	PBS (\$100)	MBS and PBS (\$100)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Panel A: Baseline																				
Active time (hours)	159.19**	92.13	75.84	106.85*	10.53	-17.12	13.81	-16.74**	4.83	-0.61	-0.52	1.87	1.10	-3.91	0.46	-1.99	-1.13	21.05	-32.66	-11.66
	[73.38]	[66.22]	[68.94]	[55.57]	[11.66]	[11.08]	[9.64]	[8.23]	[8.68]	[5.64]	[3.76]	[1.22]	[1.76]	[28.92]	[2.34]	[2.44]	[1.56]	[18.85]	[33.85]	[40.39]
Observations	45,321	46,335	45,332	43,720	40,559	40,553	40,556	40,557	40,559	40,546	46,800	46,838	46,838	46,863	53,915	41,472	53,910	53,215	53,216	53,215
Individuals	8,232	8,274	8,223	8,122	7,979	7,977	7,978	7,979	7,979	7,976	8,334	8,337	8,337	8,331	8,710	8,115	8,710	8,555	8,555	8,555
Mean of dep. variable	82.94	74.28	83.28	80.88	8.16	6.94	8.34	8.56	8.56	8.11	0.46	0.06	0.22	61.12	0.54	0.40	0.14	2.44	0.29	2.72
F-statistic of IV	36.33	37.05	33.89	32.30	23.25	23.25	23.37	23.25	23.35	23.33	35.89	35.63	35.63	35.60	40.84	35.58	40.89	41.21	41.15	41.21
Hausman test (p value)	0.02	0.14	0.31	0.05	0.11	0.12	0.22	0.03	0.61	0.81	0.83	0.09	0.38	0.90	1.00	0.51	0.42	0.19	0.89	0.58
S-H test (p value)	0.77	0.64	0.93	0.98	0.17	0.93	0.18	0.51	0.86	0.30	0.03	0.24	0.24	0.59	0.50	0.16	0.30	0.87	0.34	0.42
Panel B1: Using different instruments - maximum daily temperature (and its square) and daily precipitation																				
Active time (hours)	147.91**	121.30*	60.26	103.01**	7.57	-17.69*	18.94**	-13.92*	5.43	0.38	1.39	1.73	2.07	-3.18	-1.01	-3.90*	-0.80	22.66	-17.85	4.73
	[68.66]	[62.55]	[63.40]	[50.07]	[10.62]	[10.51]	[9.29]	[7.77]	[8.27]	[5.26]	[3.62]	[1.17]	[1.67]	[27.35]	[2.24]	[2.28]	[1.52]	[17.63]	[19.49]	[27.43]
Observations	45,321	46,335	45,332	43,720	40,559	40,553	40,556	40,557	40,559	40,546	46,800	46,838	46,838	46,863	53,915	41,472	53,910	53,215	53,216	53,215
Individuals	8,232	8,274	8,223	8,122	7,979	7,977	7,978	7,979	7,979	7,976	8,334	8,337	8,337	8,331	8,710	8,115	8,710	8,555	8,555	8,555
Mean of dep. variable	82.94	74.28	83.28	80.88	8.16	6.94	8.34	8.56	8.56	8.11	0.46	0.06	0.22	61.12	0.54	0.40	0.14	2.44	0.29	2.72
F-statistic of IV	27.41	28.26	26.58	26.39	16.99	16.99	17.08	17.00	17.07	17.06	26.84	26.68	26.68	26.52	31.17	29.18	31.24	31.39	31.37	31.39
Hausman test (p value)	0.02	0.05	0.38	0.04	0.23	0.08	0.06	0.05	0.55	0.98	0.53	0.10	0.15	0.88	0.58	0.13	0.50	0.14	0.89	0.35
S-H test (p value)	0.88	0.50	0.86	0.99	0.16	0.98	0.10	0.48	0.96	0.50	0.04	0.44	0.17	0.86	0.18	0.08	0.49	0.96	0.63	0.67
Panel C1: Excluding individual and household level variables																				
Active time (hours)	161.33**	91.85	88.57	113.05**	10.90	-17.07	14.31	-16.42**	4.64	-0.40	-0.97	1.92	0.99	-8.93	0.42	-1.98	-1.20	21.04	-32.91	-11.83
	[73.07]	[66.06]	[68.96]	[55.52]	[11.58]	[11.06]	[9.64]	[8.22]	[8.66]	[5.62]	[3.77]	[1.22]	[1.76]	[29.00]	[2.34]	[2.43]	[1.56]	[18.90]	[34.00]	[40.51]
Observations	45,440	46,454	45,448	43,836	40,668	40,662	40,665	40,666	40,668	40,655	46,941	46,983	46,983	47,006	54,053	41,574	54,048	53,371	53,372	53,371
Individuals	8,256	8,298	8,247	8,146	8,002	8,000	8,001	8,002	8,002	7,999	8,360	8,363	8,363	8,357	8,738	8,138	8,738	8,582	8,582	8,582
Mean of dep. variable	82.93	74.27	83.26	80.87	8.16	6.94	8.34	8.56	8.56	8.11	0.46	0.06	0.22	61.13	0.54	0.40	0.14	2.44	0.29	2.73
F-statistic of IV	36.72	37.36	34.33	32.81	23.31	23.31	23.43	23.31	23.41	23.39	36.02	35.59	35.59	35.55	40.92	36.03	40.97	41.08	40.97	41.08
Hausman test (p value)	0.02	0.14	0.23	0.04	0.10	0.12	0.20	0.03	0.62	0.83	0.91	0.08	0.41	0.98	1.00	0.51	0.40	0.19	0.87	0.60
S-H test (p value)	0.77	0.63	0.97	0.96	0.17	0.94	0.17	0.52	0.90	0.30	0.02	0.21	0.23	0.50	0.53	0.14	0.33	0.85	0.34	0.43

Notes: See Table 2. S-H test refers to Sargan-Hansen test.

Appendix Table A9: Robustness checks (continued)

	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional symptoms	Conduct	Peer problem	SDQ Overall	BMI	Underweight	Overweight	Waist (cm)	Excellent health	Any ongoing condition	Prescribed medicine	MBS (\$100)	PBS (\$100)	MBS and PBS (\$100)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Panel D1: Adding more variables - Media time																				
Active time (hours)	164.52**	90.37	71.74	106.72*	10.68	-18.90	14.07	-18.01**	4.47	-1.20	-0.46	1.98	1.22	-1.86	0.33	-2.04	-1.09	22.19	-35.07	-12.94
	[76.71]	[69.18]	[72.31]	[58.20]	[12.23]	[11.77]	[10.20]	[8.78]	[9.18]	[5.96]	[3.91]	[1.27]	[1.83]	[30.18]	[2.41]	[2.54]	[1.61]	[19.38]	[36.39]	[42.94]
Observations	45,321	46,335	45,332	43,720	40,559	40,553	40,556	40,557	40,559	40,546	46,800	46,838	46,838	46,863	53,915	41,472	53,910	53,215	53,216	53,215
Individuals	8,232	8,274	8,223	8,122	7,979	7,977	7,978	7,979	7,979	7,976	8,334	8,337	8,337	8,331	8,710	8,115	8,710	8,555	8,555	8,555
Mean of dep. variable	82.94	74.28	83.28	80.88	8.16	6.94	8.34	8.56	8.56	8.11	0.46	0.06	0.22	61.12	0.54	0.40	0.14	2.44	0.29	2.72
F-statistic of IV	33.77	34.31	31.08	29.75	21.10	21.10	21.21	21.10	21.16	21.16	33.46	33.20	33.20	33.04	38.77	33.27	38.80	39.08	39.03	39.08
Hausman test (p value)	0.02	0.16	0.35	0.06	0.11	0.10	0.23	0.03	0.65	0.76	0.83	0.08	0.37	0.86	0.97	0.51	0.45	0.18	0.90	0.58
S-H test (p value)	0.78	0.64	0.94	0.98	0.17	0.96	0.18	0.53	0.85	0.29	0.03	0.25	0.23	0.58	0.48	0.16	0.30	0.89	0.34	0.42
Panel D2: Adding more variables - Travel time																				
Active time (hours)	159.57**	91.13	75.23	106.63*	11.13	-17.07	14.12	-16.84**	4.86	-0.43	-0.71	1.90	1.04	-4.18	0.48	-2.08	-1.10	21.09	-33.43	-12.38
	[73.39]	[66.31]	[68.71]	[55.59]	[11.45]	[11.05]	[9.65]	[8.25]	[8.68]	[5.63]	[3.74]	[1.20]	[1.75]	[28.79]	[2.34]	[2.44]	[1.56]	[18.75]	[34.62]	[41.04]
Observations	45,321	46,335	45,332	43,720	40,559	40,553	40,556	40,557	40,559	40,546	46,800	46,838	46,838	46,863	53,915	41,472	53,910	53,215	53,216	53,215
Individuals	8,232	8,274	8,223	8,122	7,979	7,977	7,978	7,979	7,979	7,976	8,334	8,337	8,337	8,331	8,710	8,115	8,710	8,555	8,555	8,555
Mean of dep. variable	82.94	74.28	83.28	80.88	8.16	6.94	8.34	8.56	8.56	8.11	0.46	0.06	0.22	61.12	0.54	0.40	0.14	2.44	0.29	2.72
F-statistic of IV	36.69	37.43	34.50	32.61	23.53	23.54	23.65	23.53	23.63	23.62	36.56	36.27	36.27	36.35	40.96	36.23	41.02	41.43	41.37	41.43
Hausman test (p value)	0.02	0.14	0.31	0.05	0.10	0.12	0.20	0.03	0.61	0.83	0.87	0.08	0.40	0.91	0.99	0.48	0.43	0.19	0.89	0.59
S-H test (p value)	0.82	0.61	0.91	0.93	0.18	0.97	0.20	0.54	0.87	0.30	0.03	0.26	0.24	0.60	0.50	0.17	0.29	0.88	0.34	0.42
Panel E1: Controlling for weather conditions on survey date																				
Active time (hours)	168.15*	174.94**	82.83	133.90**	13.47	-0.51	18.84	-9.71	5.06	5.89	1.40	0.91	2.22	-1.39	0.34	-4.40	-2.77	12.25	-35.01	-22.79
	[86.32]	[78.62]	[81.34]	[66.34]	[14.66]	[12.56]	[11.59]	[9.39]	[10.08]	[6.87]	[4.11]	[1.33]	[1.98]	[31.75]	[2.67]	[2.74]	[1.77]	[20.89]	[37.14]	[44.70]
Observations	45,321	46,335	45,332	43,720	40,559	40,553	40,556	40,557	40,559	40,546	46,800	46,838	46,838	46,863	53,915	41,472	53,910	53,215	53,216	53,215
Individuals	8,232	8,274	8,223	8,122	7,979	7,977	7,978	7,979	7,979	7,976	8,334	8,337	8,337	8,331	8,710	8,115	8,710	8,555	8,555	8,555
Mean of dep. variable	82.94	74.28	83.28	80.88	8.16	6.94	8.34	8.56	8.56	8.11	0.46	0.06	0.22	61.12	0.54	0.40	0.14	2.44	0.29	2.72
F-statistic of IV	25.78	26.71	24.22	22.47	16.22	16.21	16.32	16.19	16.27	16.26	27.47	27.18	27.18	26.53	30.19	27.28	30.28	30.93	30.90	30.93
Hausman test (p value)	0.04	0.02	0.33	0.04	0.12	0.97	0.13	0.25	0.63	0.42	0.47	0.42	0.17	0.81	0.98	0.12	0.09	0.48	0.87	0.74
S-H test (p value)	0.98	0.81	0.74	0.93	0.32	0.97	0.21	0.96	0.96	0.36	0.03	0.26	0.22	0.30	0.78	0.33	0.28	0.93	0.30	0.33

Notes: See Table 2. S-H test refers to Sargan-Hansen test.

Appendix Table A9: Robustness checks (continued)

	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional symptoms	Conduct	Peer problem	SDQ Overall	BMI	Underweight	Overweight	Waist (cm)	Excellent health	Any ongoing condition	Prescribed medicine	MBS (\$100)	PBS (\$100)	MBS and PBS (\$100)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Panel E2: Controlling for cumulative weather conditions in the 365 days before the survey date - Average daily maximum temperature and average daily precipitation																				
Active time (hours)	148.30**	78.51	95.78	108.74*	16.56*	-15.72	12.71	-17.54**	4.50	0.45	1.34	2.00	1.96	5.53	1.46	-1.74	-0.45	17.64	-92.19	-74.50
	[73.94]	[67.08]	[69.30]	[56.63]	[9.67]	[11.00]	[9.59]	[8.04]	[8.61]	[5.35]	[3.82]	[1.25]	[1.78]	[29.18]	[2.34]	[2.36]	[1.52]	[21.11]	[90.55]	[93.74]
Observations	41,701	42,547	41,664	40,282	36,617	36,611	36,614	36,615	36,617	36,604	42,891	42,919	42,919	42,967	45,259	41,472	45,253	44,717	44,718	44,717
Individuals	8,075	8,120	8,067	7,968	7,816	7,814	7,816	7,816	7,816	7,813	8,177	8,178	8,178	8,170	8,299	8,115	8,298	8,168	8,168	8,168
Mean of dep. variable	82.82	74.47	83.29	80.90	8.20	6.99	8.34	8.66	8.58	8.15	0.47	0.06	0.23	61.69	0.53	0.40	0.14	2.40	0.30	2.69
F-statistic of IV	35.89	38.08	35.74	33.02	22.86	22.84	22.94	22.85	22.95	22.94	36.06	35.79	35.79	35.89	40.40	37.31	40.60	41.21	41.09	41.21
Hausman test (p value)	0.04	0.22	0.21	0.06	0.11	0.15	0.27	0.01	0.65	0.84	0.49	0.07	0.18	0.63	0.67	0.60	0.61	0.19	0.49	0.33
S-H test (p value)	0.91	0.59	0.76	1.00	0.23	0.92	0.22	0.80	0.91	0.35	0.02	0.30	0.28	0.43	0.81	0.14	0.20	0.34	0.47	0.90
Panel E3: Controlling for cumulative weather conditions in the 365 days before the survey date - Number of days with daily maximum temperature exceeding a given thresholds and number of rainy days																				
Active time (hours)	160.31**	82.61	100.81	114.09**	17.41*	-17.32	13.72	-19.00**	4.14	0.16	0.08	2.14*	1.55	1.32	1.25	-1.97	-0.48	19.33	-75.15	-55.86
	[76.08]	[68.09]	[71.05]	[57.96]	[9.87]	[11.21]	[9.84]	[8.29]	[8.74]	[5.46]	[3.90]	[1.29]	[1.83]	[29.90]	[2.40]	[2.44]	[1.55]	[21.24]	[72.21]	[76.13]
Observations	41,701	42,547	41,664	40,282	36,617	36,611	36,614	36,615	36,617	36,604	42,891	42,919	42,919	42,967	45,259	41,472	45,253	44,717	44,718	44,717
Individuals	8,075	8,120	8,067	7,968	7,816	7,814	7,816	7,816	7,816	7,813	8,177	8,178	8,178	8,170	8,299	8,115	8,298	8,168	8,168	8,168
Mean of dep. variable	82.82	74.47	83.29	80.90	8.20	6.99	8.34	8.66	8.58	8.15	0.47	0.06	0.23	61.69	0.53	0.40	0.14	2.40	0.30	2.69
F-statistic of IV	34.44	36.53	34.03	31.48	22.03	22.02	22.11	22.03	22.12	22.12	34.33	34.10	34.10	34.18	38.61	35.61	38.79	39.50	39.38	39.50
Hausman test (p value)	0.03	0.21	0.19	0.05	0.09	0.12	0.23	0.01	0.67	0.82	0.72	0.06	0.28	0.77	0.74	0.51	0.58	0.15	0.62	0.54
S-H test (p value)	0.83	0.70	0.80	0.92	0.29	0.98	0.23	0.83	0.94	0.43	0.03	0.36	0.29	0.61	0.84	0.17	0.22	0.26	0.34	0.66
Panel E4: Controlling for cumulative weather conditions over an outcome-dependent period before the survey date - Average daily maximum temperature and average daily precipitation																				
Active time (hours)	170.96**	126.75	73.83	119.45*	16.87*	-16.94	13.33	-18.04**	4.64	0.33	1.40	0.91	2.22	-1.39	0.34	-4.40	-2.77	17.64	-92.19	-74.50
	[84.75]	[77.35]	[78.11]	[64.27]	[9.70]	[11.03]	[9.63]	[8.07]	[8.64]	[5.36]	[4.11]	[1.33]	[1.98]	[31.75]	[2.67]	[2.74]	[1.77]	[21.11]	[90.55]	[93.74]
Observations	45,321	46,335	45,332	43,720	36,646	36,640	36,643	36,644	36,646	36,633	46,800	46,838	46,838	46,863	53,915	41,472	53,910	44,717	44,718	44,717
Individuals	8,232	8,274	8,223	8,122	7,817	7,815	7,817	7,817	7,817	7,814	8,334	8,337	8,337	8,331	8,710	8,115	8,710	8,168	8,168	8,168
Mean of dep. variable	82.94	74.28	83.28	80.88	8.20	6.99	8.34	8.66	8.58	8.15	0.46	0.06	0.22	61.12	0.54	0.40	0.14	2.40	0.30	2.69
F-statistic of IV	26.49	26.64	24.25	23.14	22.71	22.70	22.79	22.70	22.80	22.79	27.47	27.18	27.18	26.53	30.19	27.28	30.28	41.21	41.09	41.21
Hausman test (p value)	0.03	0.09	0.39	0.06	0.10	0.12	0.24	0.01	0.64	0.84	0.47	0.42	0.17	0.81	0.98	0.12	0.09	0.19	0.49	0.33
S-H test (p value)	0.65	0.93	0.95	0.84	0.25	0.99	0.19	0.71	0.90	0.39	0.03	0.26	0.22	0.30	0.78	0.33	0.28	0.34	0.47	0.90

Notes: See Table 2. S-H test refers to Sargan-Hansen test.

Appendix Table A10: Impact of physical activity on cognitive outcomes - results from POLS and IV models

	FE	IV	FE	IV	FE	IV	FE	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	PPVT		Matrix reasoning		Reading		Writing	
Active time (hours)	3.02*	-92.90**	-2.99***	18.99	-123.46***	-699.56	-42.63*	-1,179.35
	[1.66]	[41.87]	[1.04]	[30.98]	[27.00]	[964.86]	[25.48]	[896.00]
Observations	22,276	22,274	18,200	18,199	20,090	20,090	20,089	20,089
Mean of dep. variable	70.98	70.98	10.68	10.68	534.61	534.61	513.20	513.20
F-statistic of IV		23.23		15.87		11.48		11.92
Hausman test (p value)		0.10		0.51		0.62		0.24
Sargan-Hansen test (p value)		0.00		0.65		0.10		0.35
	Spelling		Grammar		Numeracy			
Active time (hours)	-111.27***	-1,033.00	-133.01***	-1,837.59*	-68.74**	-593.74		
	[24.86]	[901.77]	[28.29]	[1,017.61]	[26.95]	[942.12]		
Observations	20,118	20,118	20,114	20,114	20,005	20,005		
Mean of dep. variable	523.58	523.58	536.05	536.05	533.80	533.80		
F-statistic of IV		12.25		12.17		10.95		
Hausman test (p value)		0.42		0.13		0.72		
Sargan-Hansen test (p value)		0.25		0.13		0.17		

Notes: POLS results are from the pooled OLS regression (1) while IV results from pooled regression models (1) and (2), without controlling for individual fixed effects. F-statistic of IV denotes the F statistic for the excluded instruments in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the physically active time variable in equation (1). Sargan-Hansen test indicates p value from a Sargan-Hansen over-identification test. Instruments: daily maximum temperature and daily precipitation. Other explanatory variables include age, gender, low birth weight, Indigenous status, maternal education, maternal migration status, household size, two-parent household, cohort dummy, local socio-economic background variables, state/territory dummies, survey year dummies, survey month dummies, and TUD day-of-week dummies. Robust standard errors clustered at the individual level in parentheses. Results (coefficient estimates and standard errors) are multiplied by 100 for aesthetic purposes. The symbol \*denotes significance at the 10% level, \*\*at the 5% level, and \*\*\*at the 1% level.



Appendix Table A11: Non-linear impact of physical activity - Results from FE-IV regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional symptoms	Conduct	Peer problem	SDQ Overall
Active time (hours)	737.55 [3,188.41]	1,199.74 [3,336.13]	21.84 [3,532.09]	229.40 [1,831.35]	-116.77 [310.47]	64.70 [252.01]	-286.99 [600.47]	82.76 [235.66]	-14.79 [149.80]	-56.15 [136.67]
Active time squared	-110.40 [581.37]	-193.17 [591.95]	6.55 [629.46]	-23.97 [334.64]	23.35 [58.28]	-15.64 [47.72]	57.82 [113.88]	-18.39 [44.81]	3.62 [28.44]	10.58 [25.89]
Observations	45,321	46,335	45,332	43,720	40,559	40,553	40,556	40,557	40,559	40,546
Individuals	8,232	8,274	8,223	8,122	7,979	7,977	7,978	7,979	7,979	7,976
F-statistic of IV in active time	24.04	24.34	22.12	22.45	13.88	13.87	13.92	13.88	13.94	13.91
F-statistic of IV in active time squared	12.36	12.88	11.57	11.30	6.14	6.13	6.15	6.14	6.17	6.15
Hausman test (p value)	0.15	0.21	0.58	0.17	0.25	0.21	0.03	0.12	0.90	0.94
Sargan-Hansen test (p value)	0.38	0.50	0.13	0.64	0.24	0.73	0.94	0.62	0.99	0.41
	BMI	Underweight	Overweight	Waist (cm)	Excellent health	Any ongoing condition	Prescribed medicine	MBS (\$100)	PBS (\$100)	MBS and PBS (\$100)
Active time (hours)	126.10 [398.03]	28.38 [94.23]	65.15 [206.85]	7.64 [846.10]	-36.42 [36.33]	76.62 [107.63]	-14.37 [18.52]	-40.31 [237.91]	640.29 [869.55]	603.47 [882.01]
Active time squared	-22.77 [72.39]	-4.87 [17.14]	-11.49 [37.60]	-2.01 [154.89]	6.16 [6.30]	-14.07 [18.85]	2.35 [3.23]	10.48 [40.31]	-112.99 [151.54]	-103.15 [153.70]
Observations	46,800	46,838	46,838	46,863	53,915	41,472	53,910	53,215	53,216	53,215
Individuals	8,334	8,337	8,337	8,331	8,710	8,115	8,710	8,555	8,555	8,555
F-statistic of IV in active time	23.03	22.94	22.94	23.00	24.41	23.08	24.44	24.55	24.54	24.55
F-statistic of IV in active time squared	11.80	11.80	11.80	11.82	11.28	11.76	11.28	11.52	11.52	11.52
Hausman test (p value)	0.76	0.15	0.25	0.97	0.15	0.04	0.85	0.22	0.93	0.28
Sargan-Hansen test (p value)	0.65	0.73	0.77	0.87	0.21	0.91	0.11	0.38	0.80	0.88

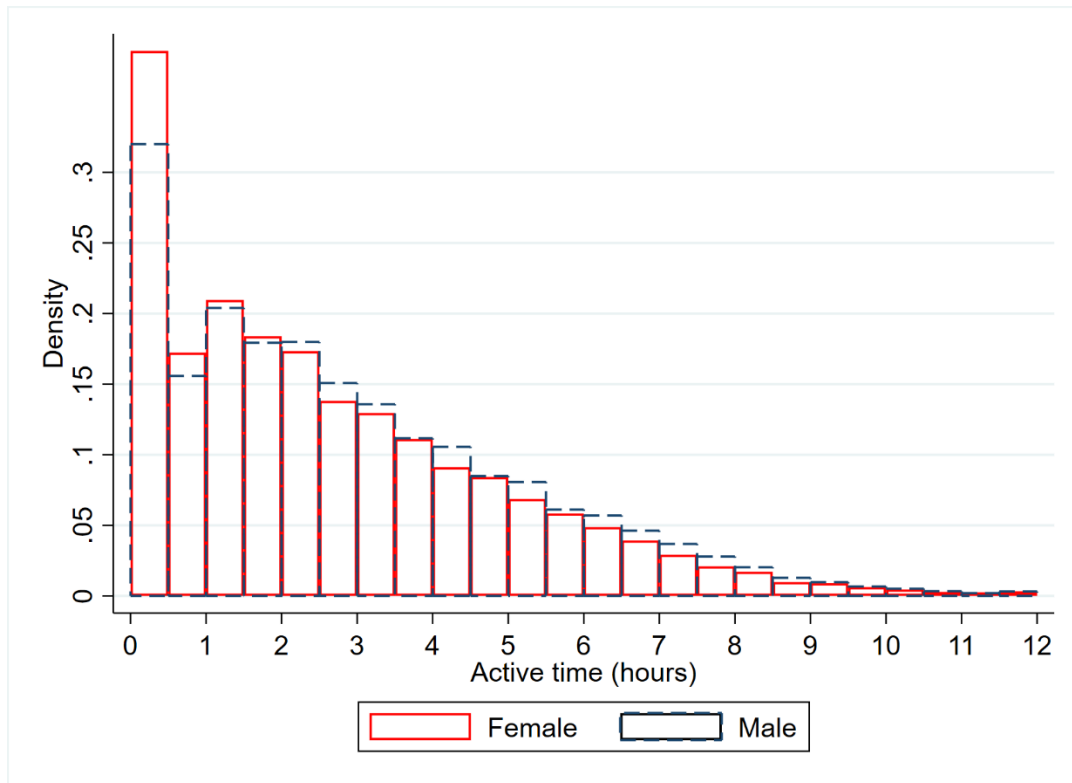
Notes: Results are obtained from FE-IV regressions. Instruments: daily maximum temperature (and its square) and daily precipitation (and its square). Other explanatory variables include age, maternal education, household size, two-parent household, local socio-economic background variables, state/territory dummies, survey year dummies, survey month dummies, and TUD day-of-week dummies. Each of all continuous explanatory variables is also included in a quadratic form. Robust standard errors clustered at the individual level in parentheses. Results (coefficient estimates and standard errors) are multiplied by 100 for aesthetic purposes. The symbol \*denotes significance at the 10% level, \*\*at the 5% level, and \*\*\*at the 1% level. Other notes: see Table 2.

Appendix Table A12: Non-linear impact of physical activity - Results from FE regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional symptoms	Conduct	Peer problem	SDQ Overall
Active time (hours)	16.42** [7.72]	11.97* [7.04]	22.51*** [7.42]	19.34*** [5.81]	2.45*** [0.82]	-0.01 [0.95]	4.28*** [0.89]	2.53*** [0.72]	2.46*** [0.77]	2.34*** [0.49]
Active time squared	-2.20** [0.87]	-1.71** [0.78]	-2.05** [0.83]	-2.17*** [0.64]	-0.24** [0.09]	-0.01 [0.11]	-0.48*** [0.10]	-0.21** [0.08]	-0.28*** [0.09]	-0.24*** [0.06]
Observations	45,321	46,335	45,332	43,720	40,559	40,553	40,556	40,557	40,559	40,546
Individuals	8,232	8,274	8,223	8,122	7,979	7,977	7,978	7,979	7,979	7,976
Optimal hours	3.70	3.50	5.50	4.50	5.10	5.10	4.50	5.90	4.40	4.80
	BMI	Underweight	Overweight	Waist (cm)	Excellent health	Any ongoing condition	Prescribed medicine	MBS (\$100)	PBS (\$100)	MBS and PBS (\$100)
Active time (hours)	-1.05*** [0.39]	-0.02 [0.11]	-0.56*** [0.18]	-13.55*** [3.14]	1.06*** [0.24]	-0.59** [0.26]	-0.37** [0.16]	-8.01*** [1.86]	-7.11 [7.24]	-15.12** [7.55]
Active time squared	0.08* [0.04]	0.01 [0.01]	0.06*** [0.02]	1.00*** [0.35]	-0.08*** [0.03]	0.05* [0.03]	0.03* [0.02]	0.75*** [0.21]	0.80 [0.84]	1.55* [0.87]
Observations	46,800	46,838	46,838	46,863	53,915	41,472	53,910	53,215	53,216	53,215
Individuals	8,334	8,337	8,337	8,331	8,710	8,115	8,710	8,555	8,555	8,555
Optimal hours	6.20	1.80	4.80	6.80	6.40	5.40	6.20	5.30	4.40	4.90

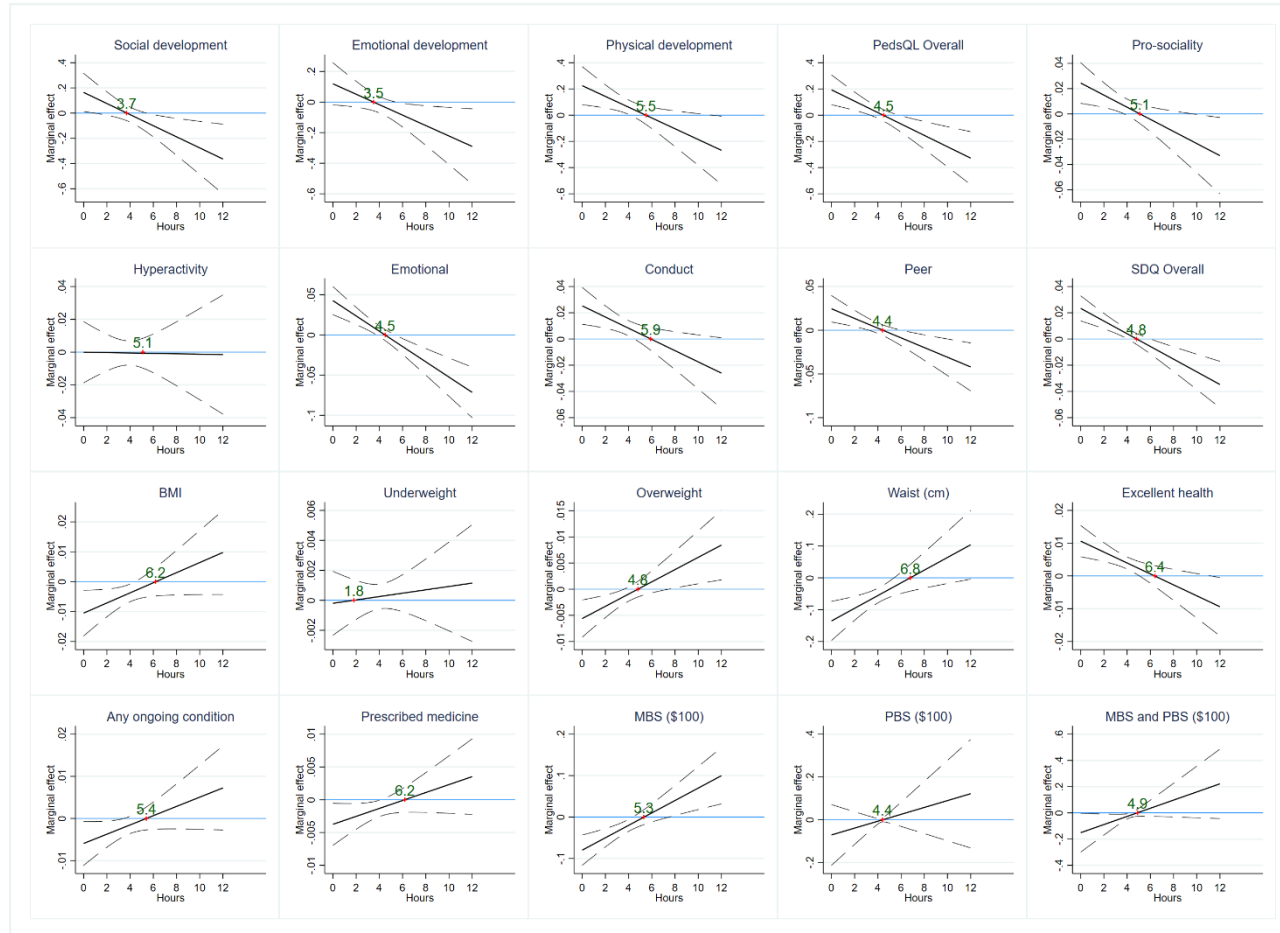
Notes: Results are obtained from FE regressions. “Optimal hours” is daily physical activity time at which the expected value of the corresponding outcome reaches its maximum/minimum. Other explanatory variables include age, maternal education, household size, two-parent household, local socio-economic background variables, state/territory dummies, survey year dummies, survey month dummies, and TUD day-of-week dummies. Robust standard errors clustered at the individual level in parentheses. Results (coefficient estimates and standard errors) are multiplied by 100 for aesthetic purposes. The symbol \*denotes significance at the 10% level, \*\*at the 5% level, and \*\*\*at the 1% level.

Appendix Figure A1: Distribution of physically active time by gender



Notes: Statistics are calculated using an estimated sample from the regression of “Social development” as an outcome.

Appendix Figure A2: Non-linear impact of physical activity on non-cognitive development and health– Marginal effects



Notes: Results (in marginal effects and their 95% confidence intervals) are obtained from FE regressions. Green figures, reported on the blue horizontal axis, show the physically activity time at which the expected outcome reaches its maximum/minimum value. Full regression results are reported in Appendix Table A12.

Appendix Table B1: Coding rules for activities by K cohort children

Grouping	Wave 1	Waves 2 and 3	Wave 4	Wave 5	Wave 6
Active	Walk for travel or for fun; ride bicycle, trike etc. (travel or fun); other exercise - swim / dance/ run about; visiting people, special event, party; other play, other activities	Walk for travel or for fun; Ride bicycle, trike etc. (travel for fun); Visiting people, special event, party; Organised sport/physical activity; Other organised lessons / activities	Organised team sports and training i.e.; Organised individual sport i.e. swimming; Ball games, riding a bike, scooter, ska; Taking Pet for a walk; Scouts, girl guides, etc.; Shopping; Going out to museums, cultural events,; Cinema; Live Sporting Events	Organised team sports and training; Organised individual sport and training; Unstructured active play; Walking pets / playing with pets; Active club activities; Shopping; Going out to a concert, play, museum, art gallery, community or school event , an amusement park etc.; Religious activities / ritual ceremonies; Attending live sporting events; Active activities nec.	Archery / Shooting sports; Athletics / Gymnastics; Fitness / Gym / Exercise; Ball Sports; Martial arts / Dancing; Motor Sports / Roller Sports / Cycling; Water/Ice/Snow Sports; Organised team sports and training other; Archery / Shooting sports (individual); Athletics / Gymnastics (individual); Fitness / Gym / Exercise (individual); Martial arts / Dancing (individual); Motor Sports / Roller Sports / Cycling (individual); Ball Sports (individual); Water/Ice/Snow Sports (individual); Organised individual sport and training other; Archery / Shooting sports (unstructured); Athletics / Gymnastics (unstructured); Fitness / Gym / Exercise (unstructured); Ball Sports (unstructured); Martial arts / Dancing (unstructured); Motor Sports / Roller Sports / Cycling (unstructured); Water/Ice/Snow Sports (unstructured); Unstructured active play Other; Walking pets/playing with pets; Active club activities; Shopping; Shopping; Purchasing consumer goods; Purchasing durable goods; Window shopping; Purchasing repair services; Purchasing administrative services; Purchasing personal care services; Purchasing other services; Attendance at movies / cinema; Attendance at concert/theatre; Attendance at museum / exhibition / art gallery; Attendance at zoo / animal park / botanic garden; Attendance at other mass events; Going out nec; Religious practice; Weddings, funerals, rites of passage; Religious activities / ritual ceremonies nec; Attending live sporting events; Active activities nec.
Media	Watching TV, video, DVD, movie; Listening to	Watching TV, video, DVD, movie; Listening to tapes, CD's,	Electronic media, games, computer use; Computer games - internet; Computer	Playing games; Watching TV programs or movies/videos; Spending time on social networking sites; Downloading/posting media	Playing games (electronic device); Playing games (Electronic device) nfd.; Watching TV programs or movies/videos; Spending time on social networking sites; Downloading/posting media;

Grouping	Wave 1	Waves 2 and 3	Wave 4	Wave 5	Wave 6
	tapes, CD's, radio, music; Use computer/computer games	radio, music; Use computer/computer games (if this activity done NOT for or NOT as part of homework)	games - not internet; Xbox, Playstation, Nintendo, Wii etc.; Internet not covered elsewhere; TV/DVD; Talking on a landline phone; Talking on a mobile phone; Texting, email, social networking - facebook/twitter; Skype or Webcam	(e.g. music, videos, applications); Internet shopping (excluding downloading/posting media); General Internet browsing (excluding homework); Creating/maintaining websites (excluding social networking profile); General application use (e.g. Microsoft Office; excluding homework); Electronic device use nec.; Talking on a landline phone (not video chat); Talking on a mobile phone (not video chat); Video chatting (e.g. Skype); Texting/emailing; Online chatting / Instant messaging	Internet shopping; General Internet browsing; Creating/maintaining websites; General application use; Electronic device use nec; Talking on a mobile phone; Video chatting; Texting/emailing; Online chatting / Instant messaging
Travel	Travel in pusher or on bicycle seat; travel in car / other household vehicle; travel on public transport, ferry, plane; taken places with adult (e.g. shopping)	Travel in car; Travel on public transport; Taken places with adult (e.g. Shopping)	Travel by foot; by bike, scooter, skateboard etc.; by private car; Travel by public transport such as bus	Travel by foot; by bike, scooter, skateboard etc.; by private motor vehicle/bike; by public/chartered transport such as bus, taxi or aeroplane; Travel nec.	Travel by foot; by bike, scooter, skateboard etc.; by private motor vehicle/bike; by public/chartered transport; Travel nec.

Appendix Table B2: Coding rules for activities by B cohort children

Grouping	Wave 1	Wave 2	Wave 3	Waves 6, 7 and 8
Active	Crawl, climb, swing arms or legs; Other play, other activities; Visiting people, special event, party	Active free play; Visiting people, special event, party; Walking; Ride bicycle/trike	Active free play; visiting people, special event, outing; walking; travel in pusher/bicycle seat; ride bicycle, trike, etc.	Archery / Shooting sports; Athletics / Gymnastics; Fitness / Gym / Exercise; Ball Sports; Martial arts / Dancing; Motor Sports / Roller Sports / Cycling; Water/Ice/Snow Sports; Organised team sports and training other; Archery / Shooting sports (individual); Athletics / Gymnastics (individual); Fitness / Gym / Exercise (individual); Martial arts / Dancing (individual); Motor Sports / Roller Sports / Cycling (individual); Ball Sports (individual); Water/Ice/Snow Sports (individual); Organised individual sport and training other; Archery / Shooting sports (unstructured); Athletics / Gymnastics (unstructured); Fitness / Gym / Exercise (unstructured); Ball Sports (unstructured); Martial arts / Dancing (unstructured); Motor Sports / Roller Sports / Cycling (unstructured); Water/Ice/Snow Sports (unstructured); Unstructured active play Other; Walking pets/playing with pets; Active club activities; Shopping; Shopping; Purchasing consumer goods; Purchasing durable goods; Window shopping; Purchasing repair services; Purchasing administrative services; Purchasing personal care services; Purchasing other services; Attendance at movies / cinema; Attendance at concert/theatre; Attendance at museum / exhibition / art gallery; Attendance at zoo / animal park / botanic garden; Attendance at other mass events; Going out nec; Religious practice; Weddings, funerals, rites of passage; Religious activities / ritual ceremonies nec; Attending live sporting events; Active activities nec
Media	Watching TV, video or DVD; Listening to tapes, CD's, radio, music	Watching TV, video, DVD, movie; Listening to tapes, CDs, radio, music; Using computer, computer game	Watching TV, video, DVD, movie; listening to tapes, CDs, radio, music; using computer, computer game	Playing games (electronic device); Playing games (Electronic device) nfd; Watching TV programs or movies/videos; Spending time on social networking sites; Downloading/posting media; Internet shopping; General Internet browsing; Creating/maintaining websites; General application use; Electronic device use nec.; Talking on a mobile phone; Video chatting; Texting/emailing; Online chatting / Instant messaging
Travel	Taken places with adult (e.g. shopping); Taken out in pram or bicycle seat; Travel in car / other household vehicle; Travel on public transport, ferry, plane	Travel in car; Travel in a pusher/bicycle seat; Travel on public transport; Taken places with adult (e.g. Shopping)	Travel in car; travel on public transport; taken places with adult	Travel by foot; by bike, scooter, skateboard etc.; by private motor vehicle/bike; by public/chartered transport; Travel nec.

Appendix Table B3: SDQ measures

Scale	Components
Pro-sociality	Considerate of other people's feelings Readily shared with children Helpful if someone is hurt etc. Kind to younger children Often volunteered to help
Hyperactivity <sup>(a)</sup>	Not been able to stay still Constantly fidgeting etc Easily distracted Stopped to think before acting <sup>(a)</sup> Has a good attention span <sup>(a)</sup>
Emotional <sup>(a)</sup>	Complained of headaches etc. Often seemed worried Often been unhappy or tearful Nervous or easily lose confidence Had many fears
Conduct <sup>(a)</sup>	Temper Obeys requests <sup>(a)</sup> Often fights/bullies children Often lies or cheats Steals
Peer <sup>(a)</sup>	Has been solitary Has at least one good friend <sup>(a)</sup> Liked by other children <sup>(a)</sup> Picked on/bullied by children Gets on better with adults

Notes: This table describes components of non-cognitive skill measures used in the paper. We use responses (1 Not true; 2 Somewhat true; 3 Certainly true) to the question "Please tick one box for each of the following statements to best describe the study child's behaviour over the past six months:". <sup>(a)</sup> indicates that reversed responses are used to calculate the corresponding sub-scale or the overall non-cognitive scale.