The effects of sleep duration on child health and development

Nguyen, Ha and Zubrick, Stephen and Mitrou, Francis

Telethon Kids Institute, The University of Western Australia

2024

The effects of sleep duration on child health and development

Ha Trong Nguyen [a, b, *]

Stephen R. Zubrick [a, b]

Francis Mitrou [a, b]

Children and adolescents spend more than one-third of their time sleeping. Yet, we know little about the causal impact of sleeping on their development. This paper is the first to exploit variation in local daily daylight duration measured on pre-determined diary dates across the same individuals through time as an instrument in an individual fixed effects regression model to draw causal estimates of sleep duration on a comprehensive set of child development indicators. Applying this model to about 50 thousand time use diaries from two cohorts of Australian children spanning over 16 years, we first document that children sleep substantially less on days with longer daylight duration. Our results show that sleeping longer improves selected general developmental, behavioural and health outcomes in children and adolescents. By contrast, sleeping more statistically significantly increases the BMI scores, mainly by increasing the risk of being overweight. Moreover, while the impact of sleep duration on general and behavioural outcomes is more pronounced for females or older individuals, the effect on BMI is largely driven by males. The results indicate a null or relatively small positive impact of sleeping longer on cognitive skills.

Keywords: Sleep; Time Allocation; Circadian Rhythms; Human Capital; Child Development.

JEL classifications: I00; I12; 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.

[a] Telethon Kids Institute, Australia

[b] Centre for Child Health Research, The University of Western Australia, Australia

Acknowledgements: We thank the Co-Editor Scott Adams, anonymous reviewers of Journal of Economic Behavior & Organization, and participants at 44th Australian Health Economics Society Conference for helpful comments and suggestions. This research was partly funded by the Australian Research Council Centre of Excellence for Children and Families over the Life Course (#CE200100025). 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 author and should not be attributed to the DSS, the AIFS or the ABS.

1. Introduction

Humans spend approximately one-third of their lives sleeping, with children sleeping more than adults (Ohayon *et al.* 2004; Hirshkowitz *et al.* 2015). Given time is a scarce resource, individuals make choices about how they allocate time to sleep (Biddle & Hamermesh 1990). Because optimal sleep is a biological necessity, an understanding of the consequences of choices made by individuals (or for them) about time spent sleeping is of value in recommending health advice specifically, but also more broadly to aspects of human development – particularly that of children and young people. A large scientific literature repeatedly attests to the association between a range of sleep qualities – including the amount of time – and aspects of child development (see Section 2 for a literature review). And yet, studies that directly estimate the causal effects of time sleeping on various health, cognitive and non-cognitive outcomes in children and adolescents are scant (see, for instance, recent reviews by Matricciani *et al.* (2019) or Jagnani (2022)). This paper examines the causal impact of sleep duration on the health and development of children and adolescents. It contributes to our understanding of the relationship both between sleep and child health (Chaput *et al.* 2016) and between sleep and academic performance in two important ways.

First, this paper moves beyond observational studies of association to more directly address unobservable individual heterogeneity and reverse causality issues (Wooldridge 2010) by employing a new empirical model to estimate the causal impact of sleep duration on child development. Particularly, we exploit variation in local daily daylight duration measured on pre-determined diary dates across the same individuals over time as an instrument in an individual fixed effects (FE) regression model to draw causal estimates of sleep duration on child development indicators. Motivated by medical research on circadian rhythm (Reppert & Weaver 2002; Roenneberg *et al.* 2007), previous studies have successfully employed solar cycle-based instruments in an instrumental variables (IV) approach to identify the causal impacts of adults' sleep duration (Giuntella *et al.* 2017; Gibson & Shrader 2018; Kajitani 2021). Our paper is the first to adopt this IV identification strategy to explore the impact of sleep duration in children and adolescents. We augment this IV approach by applying it to an individual FE regression model.

Second, this paper presents causal evidence of the impact of sleep duration on an extensive list of child developmental outcomes. Prior studies have focused on a limited range of child development outcomes, potentially due to data constraints (Taras & Potts-Datema 2005; Matricciani *et al.* 2019). This would provide an incomplete picture of the potential impacts of sleep duration which may have differential effects on specific outcomes of interest (Fiorini & Keane 2014; Nguyen *et al.* 2022a). To provide a more complete picture, we utilise high-quality longitudinal data with rich information on both child sleep and development outcomes. In particular, we quantify sleep duration using time-use diaries, which are considered one of the most accurate tools to record time allocation (Frazis & Stewart 2012), from two cohorts of children observed on multiple occasions over 16 years. During the same period, our data also contain a rich suite of child development outcome measures, including general development, health, anthropometric measures, health expenditures, and cognitive test scores. Many of these outcomes were objectively measured or available via linked administrative data sources and hence are less prone to measurement errors. By providing evidence of the impact of sleep duration on a comprehensive set of outcomes in one unified framework, this paper depicts a much broader picture of the effects of sleep duration than previously possible, providing important insights for the design of sleep recommendations for children and adolescents (Paruthi *et al.* 2016).

Employing 16 years of data from the Longitudinal Study of Australian Children (LSAC) survey, we first document that on days with longer daylight duration, children sleep statistically significantly less. Using a fixed-effects instrumental variables (FE-IV) approach, we then find that sleeping longer improves selected general development, behavioural and health outcomes in children and adolescents. By contrast, sleeping more increases their Body Mass Index (BMI), mainly by increasing the likelihood of their being overweight. Moreover, while the general and behavioural developmental benefits of sleeping longer are mainly found among females and older individuals, the potentially detrimental effects of sleeping longer on BMI are only observed for males. The results further suggest a null or at most a small positive impact of sleep duration on cognitive skills. Finally, we find our results are robust to a series of sensitivity tests, including employing alternative instruments or additionally controlling for numerous time-variant observable factors.

The rest of this paper is structured as follows. Section 2 provides a brief review of related studies and Section 3 describes the data and empirical model. We present our main empirical results in Section 4. Section 5 presents robustness checks and additional findings while Section 6 concludes.

2. Literature review

Our empirical work is theoretically motivated by a relatively small number of economists' contributions on sleep. For instance, building on the work of Becker (1965) on time allocation and Grossman (1972) on demand for health, Biddle and Hamermesh (1990) develop an optimal model of time allocation among work, leisure and sleep. Two main implications from the seminal work by Biddle and Hamermesh (1990) are: (i) sleep duration affects the amount of time allocated to other activities, and (ii) higher labour productivity increases the opportunity cost of sleep time. As Biddle and Hamermesh (1990)'s model is developed primarily to explain sleep choices in the working-age population, Jagnani (2022) extends their model to predict the sleep choices of children. It is clear from these theoretical frameworks that regardless of who makes the decision about how long to sleep, sleep is a choice variable, suggesting it is important to properly control for endogeneity of sleep when quantifying its causal impact on outcomes of interest. However, these theoretical frameworks provide ambiguous predictions about the direction, as well as the magnitude, of sleep effects on developmental outcomes in young individuals. As such, it remains an empirical issue to determine to what extent sleep affects developmental outcomes.

There is a rich literature exploring the effects of sleep on adults (Watson *et al.* 2015) and children (Chaput *et al.* 2017; Matricciani *et al.* 2019; Schlieber & Han 2021). Most of this literature is from non-economics fields, and concerns the effects of sleep on children's developmental outcomes, producing mixed results, reflecting differences in sleep measures, developmental outcomes and empirical methods employed by prior studies (see, for example, Matricciani *et al.* (2019) for a recent meta review). This literature has been criticised for relying on correlational cross-sectional designs (Matricciani *et al.* 2019). Thus, despite a large literature documenting the relationship between sleep and child development, we remain uncertain about the causal impact of sleep on health, cognitive and non-cognitive outcomes in children and adolescents.

Quantifying the causal impact of sleep is challenging due to problems related to unobserved heterogeneity and reverse causality. Specifically, there are unobservable individual characteristics (such as the individual's time preferences or genetic factors) which are correlated with both the child's sleep and their development. Reverse causality is a threat to estimate validity because it is not clear whether the child's sleep influences development or *vice versa*. To overcome these research challenges, previous studies have employed experimental research designs (Van Dongen *et al.* 2003; Lo *et al.* 2016; Beebe *et al.* 2017; Bessone *et al.* 2021) or instrumental variables methods (Giuntella *et al.* 2017; Gibson &

Shrader 2018; Groen & Pabilonia 2019; Costa-Font & Fleche 2020; Kajitani 2021; Costa-Font *et al.* 2024).[1]

Four studies which use solar cycles-based instruments share commonalities with our empirical approach.[2] Gibson and Shrader (2018) use daily sunset time recorded on the diary date as an instrument to explore the short-run impact of sleep duration on earnings in the US.[3] Likewise, Giuntella *et al.* (2017) employ yearly average sunset time at a local level as an instrument to study the effects of sleep duration on cognitive skills and depression symptoms of older workers in urban China. Kajitani (2021) also exploits the annual variation in the average daylight duration between cities as an instrument to examine the impact of sleep duration on labour market outcomes of Japanese men. More recently, Costa-Font *et al.* (2024) employ the sunset time recorded on the interview date as an instrument in an individual FE framework to examine the impact of sleep duration on labour market outcomes of German adults.

Previous IV studies have primarily focused on adults' sleep, making our study the first to adopt this IV identification strategy[4] to investigate the causal impact of sleep duration on child and adolescent development. We enhance this IV approach by applying it to an individual fixed effects (FE) model. To accurately estimate the causal effects of sleep on child development, it is crucial to control for unobservable individual factors, including the individual's time

[1] Experimental studies, particularly on students (Lo *et al.* 2016; Beebe *et al.* 2017), are not without criticism because their results may not be generalized well to real-world settings (Matricciani *et al.* 2019).
[2] Using UK data, Costa-Font and Fleche (2020) employ child sleep disruption as an instrument to examine the effect of maternal sleep duration on her labour market outcomes in an individual FE-IV model. Groen and Pabilonia (2019) examine the effects of school start time on test scores, time allocation and health of US students. In an additional analysis, they use school start time as an instrument for sleep to identify the causal impact of sleep on test scores of female students.
[3] They also use annual average sunset as a second instrument to examine the long-run impacts of sleep.
[4] Unlike our direct IV approach estimating the causal impact of sleep, Jagnani (2022) employs a reduced-form regression model with daily sunset time as a predictor, revealing only indirect sleep effects. This distinction in methodology highlights the unique contribution of our study in directly estimating the true magnitude of sleep's impact. Furthermore, our study differs from Jagnani's (2022) in three key aspects: 1) context (our study focuses on developed countries like Australia vs. developing countries in Jagnani's), 2) scope (we examine broader developmental outcomes beyond cognitive skills), and 3) data and methodology (we use panel data and individual fixed effects to account for time-invariant confounding factors, whereas Jagnani uses cross-sectional data).

preferences, genetic predispositions, sleep habits, and environmental conditions, as these factors can correlate with both the child's sleep and their development (Matricciani *et al.* 2013; Lo *et al.* 2016). Some of these time-invariant unobservable factors may also correlate with local solar cycles, potentially invalidating the use of solar cycle-based variables as an instrument. For instance, residential preference could be one such factor, as individuals may self-select into different locations based on their sensitivity to solar cycles (Gibson & Shrader 2018). Another potentially significant time-invariant factor could be persistent reporting bias in sleep duration (Wooldridge 2010; Frazis & Stewart 2012). Our FE-IV model effectively controls for time-invariant factors that may be associated with both the instrument and child development outcomes.

Our empirical approach is also relevant to those of studies which exploit exogenous sleep induced by sunset time or daylight-saving times (DST) transitions to examine the impact of sleep on adult health (Giuntella & Mazzonna 2019; Jin & Ziebarth 2020), adult economic performance (Giuntella & Mazzonna 2019), automobile accidents (Smith 2016), adults' voting behaviours (Holbein *et al.* 2019) or children's cognitive scores (Jagnani 2022).[5] To deal with the fact that sleep and outcomes of interest are not available in one common dataset, these studies have to employ a reduced-form regression approach (i.e., by including sunset time or DST transitions as an explanatory variable in the outcome equation). This reduced-form approach can only reveal the indirect impact of sleep and may lead to uncertainty in the magnitude of the actual impact of sleep on such outcomes.

[5] There are several related studies concerning the effects of school starting time on test scores (Carrell *et al.* 2011; Edwards 2012; Minges & Redeker 2016; Heissel & Norris 2018; Groen & Pabilonia 2019). These studies rarely observe students' sleep or other time uses. This paper is also related to studies examining impacts of time allocation to other activities, such as media (Gentzkow & Shapiro 2008; Nieto & Suhrcke 2021) or physical activities (Nguyen *et al.* 2022b) on development of young individuals.

3. Data and empirical strategy

3.1. Data

1.1.1. Sleep data

We use time-use diaries (TUD) from two cohorts of children surveyed in the Longitudinal Study of Australian Children (LSAC) to document the time children and adolescents allocate to sleep. 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 in 2004) and between March 1999 and February 2000 (Kindergarten or K-Cohort, 4,983 children aged 4–5 in 2004). The LSAC began in 2004 and the most recent wave 9 was surveyed in 2020/21 (Mohal *et al.* 2021).

TUDs embedded in the LSAC also were collected biennially (see Appendix Table A2 for LSAC contents by wave and cohort). There were four major changes to TUDs during the study period which are worth mentioning. First, in each of the first three waves of LSAC, the corresponding parent was given two TUDs (one on a weekday and one on a weekend day) to complete on the study child's activities. However, from wave 4 onwards, each family was given one TUD to complete each wave. Second, activities are recorded according to 96 15-minute slots in the first three waves, while activities are reported in the form of an "activity episode" diary from wave 4 onwards (See Corey *et al.* (2014) for examples of TUDs). Third, from Wave 4 onwards the study child was requested to fill in the TUD via computer assisted interview. Fourth, K cohort children were requested to complete TUDs in the first six waves while B cohort children were not asked to do so in waves 4, 5 and 9. The available TUDs enable us to study the topic over a 16-year period for study children and young people aged from birth (for B cohort) or 4/5 years old (for K cohort) up to 15/16 years old (for both cohorts).

Our primary sleep related variable is sleep duration which is calculated by summing all time slots or episodes recorded as sleep or napping during the diary date. Our sleep duration variable captures the “actual” time spent on sleeping/napping as it excludes time spent in bed awake. We measure sleep duration in hours per day.

Summary statistics, reported in Appendix Table A1 and Appendix Figure A1, show that on average, children and adolescents in our data spend about 10.5 hours per day sleeping. It should be noted that while being recorded on one day, sleep duration derived from LSAC TUDs is likely to capture sleep behaviours on a longer horizon for two main reasons. First, 67% of TUDs are explicitly stated as being recorded on an “ordinary” day. Second, the sleep duration variable used in this paper is statistically significantly correlated with other sleep-related variables which are measured over a longer period, such as “during the last month”, “regular times” or “usual” time, and are available in LSAC (See Appendix Table A3).[6] Appendix Table A3 additionally reports statistically significant correlations between sleep duration and some variables describing sleep adequacy, sleep routine or sleep quality. For instance, the significant correlations suggest individuals with a longer sleep duration are more likely to report that they have enough sleep or have a sleep routine. Moreover, individuals who sleep longer are more likely to report that they sleep well. Due to the high quality and well-established validity of TUDs (Hamermesh 2016), they have been used in various studies to investigate the connection between sleep duration and child development outcomes, with many of these development outcomes mirroring ours (Fiorini & Keane 2014; Nguyen *et al.* 2022a). Studies by Gibson and Shrader (2018) and Jagnani (2022) similarly employ time use diaries to examine the effects of sleep on labour market outcomes of US adults and academic performance of Indian children, respectively.

[6] We do not use these sleep-related variables in the main analysis because they are not available frequently enough for us to apply our empirical method (see Appendix Table A2 for descriptions and availability of main variables). For the same reason, we do not use other developmental outcomes available in LSAC in this paper.

1.1.2. Child health and development outcomes

We use the LSAC dataset and its linked datasets to construct five sets of child development outcomes. The first outcome set measures general development of children and adolescents aged 2 to 18 years and is derived from the parent-report version of the Pediatric Quality of Life Inventory (PedsQL) (Varni *et al.* 2001). This set includes three sub-scales describing Social, Emotional, and Physical development and an Overall PedsQL scale.[7] For ease of interpretation, we rescale all PedsQL measures so that a higher score indicates a more desirable trait. Moreover, for a similar reasoning, we standardize each of these outcomes to have a zero mean and a unit standard deviation.

The second development set describes child behavioural and socio-emotional development, constructed from the corresponding parent's responses to the Strengths and Difficulties Questionnaire (SDQ). This set includes an overall SDQ summary scale and 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). As has been done with the PedsQL measures, we rescale the SDQ measures so that higher SDQ scores indicate more desirable outcomes. We also standardize all SDQ - based measures.

The third outcome set includes four interviewer-administered anthropometric measures. The first measure is standardized gender- and age-adjusted Body Mass Index (BMI) scores, which are calculated using child height, weight and ages (in months) and the World Health

[7] Particularly, the corresponding parent 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" sub-scale is constructed from 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 calculated from responses to problems feeling afraid or scared, feeling sad, feeling angry, sleeping, and with worrying. The "Physical development" sub-scale is constructed from 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. See Appendix Table A2 for timeline of TUDs and developmental outcomes in the LSAC.

Organization (WHO) growth reference chart (Vidmar *et al.* 2013). To capture the potential differential impact of sleep duration on individuals at two ends of the standardized BMI scores, we additionally use two binary indicators describing if the individual is classified as being underweight or overweight. The last anthropometric measure is the waist-for-height ratio.

The fourth set consists of six measures describing the individual's health. The first three of these include indicators describing if the individual (i) has "excellent health",[8] (ii) has "any ongoing condition",[9] or (iii) currently uses "prescribed medicine".[10] For the remaining indicators we also consider the impact of sleep on three health expenditure measures derived from linked LSAC-administrative Medicare data. Medicare data record all Australian Government subsidies and out-of-pocket payments for medical services (from the Medicare Benefit Scheme (MBS)) and pharmaceuticals (from the Pharmaceutical Benefit Scheme (PBS)) under Australia's universal and compulsory Medicare scheme. About 97% of LSAC children are linked to Medicare data and, for them, we have information on health expenditures from birth until March 2019 (Mohal *et al.* 2021). We measure yearly health expenditures from MBS and PBS separately along with the sum of these two expenditure types.

The fifth outcome set captures child cognitive skills which are constructed using scores from Matrix Reasoning (MR) and the National Assessment Program – Literacy and Numeracy (NAPLAN) tests. The MR is a subtest of the Weschler Intelligence Scale to measure a child's non-verbal visuospatial ability. MR were administered by the interviewer when children were 6 to 11 years old (Mohal *et al.* 2021). The NAPLAN test is administered to all Australian

[8] This binary variable takes the value of one if the corresponding parent responses "Excellent" to a question asking: "In general, how would you say child current's health is: 1 Excellent; 2 Very good; 3 Good; 4 Fair; 5 Poor", and zero otherwise.

[9] This binary measure takes the value of one if the corresponding parent responses "Yes" to the question "Does study child have any of these ongoing conditions?", and zero otherwise. The list of ongoing conditions varies by wave, preventing us from using a particular condition as an outcome.

[10] This binary variable takes the value of one if the corresponding parent responses "Yes" to the question "Does child currently need or use medicine prescribed by a doctor, other than vitamins?", and zero otherwise.

students in grades 3, 5, 7 and 9 in the five domains of reading, writing, spelling, grammar and numeracy. The NAPLAN test results were made available via data linkage with the LSAC data (Daraganova *et al.* 2013). We also standardize each of these cognitive outcomes to facilitate interpretation of the results.

Appendix Table A4 shows that the developmental outcomes described above are not usually strongly correlated with each other. This suggests that each outcome may measure a different aspect of child development (Heckman & Kautz 2013). The fact that the developmental outcomes are not strongly correlated with each other also suggests a need to study the effects of sleep on as many different developmental outcomes as possible. Our data uniquely allow us to investigate the sleep impacts on a comprehensive set of child development outcomes. However, using too many outcome variables increases the risk of finding spurious correlations. We address this multiple inference issue in two ways. First, where possible, we reduce the number of outcome measures by generating summary indices. For instance, we use two summary indices of behavioural and socio-emotional development outcomes, namely the overall PedsQL scale and overall SDQ summary scale. Likewise, we combine both MBS and PBS expenditures as one. Second, we calculate p-values that are adjusted for the multiple inference issue using the method proposed by Simes (1986) and Benjamini and Hochberg (1995).

3.2. Empirical strategy

We empirically investigate the effect of sleep duration on developmental outcome Y of individual i at time t using the following equation:

$$Y_{it} = \alpha + \beta D_{it} + X_{it}\gamma + \delta_i + \varepsilon_{it} \quad (1)$$

where Y_{it} is a child development outcome and D_{it} is child sleep duration (measured in hours per day). X_{it} is a vector of explanatory variables described in detail below. δ_i is an individual fixed effect and ε_{it} is an error term. $\alpha,\ \beta$ and γ are parameters to be estimated. The coefficient of interest is β which gauges the effect of sleep duration on a child development outcome.

Equation (1) which controls for individual time-invariant factors (δ_i) and a rich list of time-variant factors (X_{it}) would produce a more accurate estimate of sleep duration than a pooled regression model which does not control for such factors. However, it cannot control for unobserved time-variant factors or address the reverse causality issue, leaving a causal interpretation of FE estimate uncertain. We tackle these issues by employing an instrumental variable for sleep duration variable in Equation (1). In particular, we employ an auxiliary sleep duration equation of the following form:

$$D_{it} = \lambda + S_{i(p)t}\rho + X_{it}\varphi + \delta_i + \mu_{it} \qquad (2)$$

In Equation (2), $S_{i(p)t}$ is an instrumental variable and μ_{it} is an error term. X_{it} and δ_i are defined as in Equation (1) and λ, ρ and φ are parameters to be estimated.

A successful application of an IV model relies on finding at least one valid instrument, which satisfies two conditions: (i) it is strongly correlated with sleep duration and (ii) it does not covary with other child development determinants (Wooldridge 2010). Following prior studies (Giuntella *et al.* 2017; Gibson & Shrader 2018; Kajitani 2021; Costa-Font *et al.* 2024) and building on the results in the study by Nguyen *et al.* (2022c) which shows sleep duration is most sensitive to daily daylight duration, we propose to use local daylight duration recorded on the diary date as an instrument. Specifically, our instrument $S_{i(p)t}$ measures daylight duration (in hours) recorded on diary date t in the child's residential postcode p. This instrument is calculated using the diary date, geographic coordinates (i.e., longitude and latitude) of the child's residential postcode centroid, daylight saving adjusted time zone offsets and astronomical algorithms developed by Meeus (1999).[11] We augment this IV approach with an individual fixed effects (FE) regression model.

Our identification strategy leverages two sources of variation in daily daylight duration. The primary source stems from seasonal variation in local daylight duration recorded on pre-determined TUD dates[12] for the same individual. A secondary source of variation comes from a small number (i.e., about 10% of children in final our sample) of movers and geographical variation in local daylight duration. Our data appear to support our empirical strategy. Particularly, Figure 1 represents distributions of sleep duration by local daily daylight duration, showing that longer daylight durations shift the distribution of sleep duration leftward. Moreover, Appendix Figure A2 and Appendix Table A1 show substantial variation in daily daylight duration between and within individuals for us to employ a FE model.

Our instrument is closest to that in a recent work by Costa-Font *et al.* (2024) which uses the sunset time recorded on the interview date as an instrument in an individual FE framework. Previously, Gibson and Shrader (2018) employ sunset time recorded on the diary date as an instrument in a cross-sectional setting and hence cannot control for individual heterogeneity. Our empirical model improves this approach by effectively controlling for individual time-invariant factors that may be simultaneously correlated with the daily solar cycle variable and sleep duration. As discussed by Gibson and Shrader (2018), one of such time-invariant unobservable factors would be residential sorting as individuals may self-select into different

[11] Similar astronomical algorithms have been employed in previous studies (Giuntella *et al.* 2017; Gibson & Shrader 2018). We use a STATA command written by Gibson and Shrader (2018) to perform this task.

[12] Specifically, TUD dates were pre-selected by the interviewers to obtain a random distribution of weekdays and a random distribution of weekend days (Corey *et al.* 2014). Moreover, an attempt has been made to keep the survey duration between two adjacent waves within a 24 month period (Mohal *et al.* 2021), easing a concern that survey dates and hence TUD dates were solely determined by the respondent. In line with this survey design, Appendix Figure A3 shows that the median duration between two adjacent survey waves is 24.67 months. Because our empirical strategy exploits variations in daily solar cycles recorded on the pre-determined diary dates across the same individuals over time, there is not sufficient variation in daily solar cycles in the data for us to control for a temporal level lower than a quarter level (e.g., by controlling for month dummies).

locations based on their responsiveness to solar cycles. Another potentially important time-invariant factor would be persistent reporting bias (Wooldridge 2010; Frazis & Stewart 2012).

Like the solar cycle based instruments employed by Costa-Font *et al.* (2024) and Gibson and Shrader (2018), our instrument relies on seasonal short-term variation in local areas. Potentially motivated by limitations in their data which have insufficient seasonal variation in local solar cycles, two other studies by Giuntella *et al.* (2017) and Kajitani (2021) employ a longer-term variation in yearly solar cycles across local areas as an instrument.[13] Unlike these yearly instruments, our season-based instrument can address a critical threat to studies that employ location-based instruments to infer the causal impact of sleep duration. There is compelling evidence demonstrating the significance of where individuals reside (Currie 2009; Chetty & Hendren 2018). This evidence raises doubts about the validity of yearly solar-based instruments, which remain constant over time for a given location. As a result, they are likely to be correlated with unobserved local factors that covary with outcomes of interest (Dell *et al.* 2014). Our combination of a season-based instrument and individual fixed-effects modelling effectively address this concern by enabling us to control for time-invariant unobservable factors, whether they are measured at the local or individual level.

To obtain more accurate estimates of sleep duration effects, echoing the work of Costa-Font *et al.* (2024) and Gibson and Shrader (2018), this paper utilizes seasonal variation in local solar cycles as the main identification strategy. In contrast to estimates derived from studies employing longer-term yearly variation in solar cycles, such as those by Giuntella *et al.* (2017) and Kajitani (2021), our estimates capture a shorter-term impact of sleep duration. In line with this approach, our methodology aligns with that in other sleep-related studies that exploit short-

[13] Indeed, Giuntella *et al.* (2017, page 1729) note that "Unfortunately, 87 % of the interviews took place in July and August, so we do not have sufficient seasonal variation to exploit". We have experimented with including yearly average daylight duration in the residential postcode as an additional explanatory variable in the sleep duration equation. Unreported results show that this variable does not statistically significantly explain children's sleep duration, and this is the case in both pooled and fixed effects regressions.

term exogenous variation in sleeping patterns induced by daylight-saving time transitions for identification purposes (Smith 2016; Giuntella & Mazzonna 2019; Holbein *et al.* 2019; Jin & Ziebarth 2020). Likewise, random experiments often rely on short-term interventions. For example, Bessone *et al.* (2021) conduct a randomized three-week sleep intervention in India.

The key threat to our identification strategy is seasonal factors which would correlate with both daylight duration and developmental outcomes within the same location (Gibson & Shrader 2018). To alleviate this concern, our FE-IV model controls for quarter dummies.[14] In Section 5, we further check the robustness of our results to the second condition for a strong instrument by additionally controlling for various time-variant observable factors which potentially covary with the daily daylight duration and child development outcomes.

We include in X_{it} a rich set of characteristics which have been shown to be associated with child development (Fiorini & Keane 2014) or time allocation (Nguyen *et al.* 2021; Nguyen *et al.* 2022a). These include the individual's characteristics (e.g., age and its square, gender, Aboriginal status, low birthweight), the household's characteristics (e.g., maternal migration status, maternal education, number of siblings and two-parent household), and neighbourhood characteristics.[15] We additionally control for seasonal or spatial differences in time allocation by including TUD quarter,[16] survey wave and state/territory dummies. The inclusion of

[14] LSAC was implemented mostly in non-summer months, which do not include school summer holidays or Christmas/New Year holidays, to maximize the response rate (Mohal *et al.* 2021). Consistent with this survey design, Appendix Figure A4 shows that about 87% of TUDs were completed between April and September. This survey period does not include summer months in Australia and exhibits shorter daylight duration than the rest of year. Appendix Figure A5 exhibits that daylight duration follows a yearly cycle pattern, suggesting a need to control for other seasonal factors potentially covarying with both daylight duration and child development.

[15] To provide a comprehensive overview of the variables considered, we describe the time-invariant variables, such as gender and migration status, that are included in the pooled cross-sectional regressions but excluded from the individual FE regressions. Additionally, for comparative purposes, these time-invariant variables are included in some pooled cross-sectional regressions that do not control for individual fixed effects (see Appendix Table A8, Appendix Table A9 and Appendix Table A10). Neighbourhood characteristics include percentages of individuals having an Aboriginal/Torres Strait Islander origin, speaking English, being born in Australia or completing year 12 in linked areas, percentages of households with household income less than AU$1,000/week in linked areas, a metropolitan dummy.

[16] We obtain very similar results using season dummies instead of quarter dummies. To mitigate potential multicollinearity concerns, we employ survey wave dummies instead of survey year dummies. This choice is

state/territory dummies additionally controls for different time zones across Australia. To capture likely variation in time use patterns throughout the week, we further include in X_{it} a series of day-of-week dummies[17] and an indicator describing whether the diary was completed on holidays (Nguyen *et al.* 2021; Nguyen *et al.* 2022a).

The unit of analysis in this section is diary level and we do not distinguish whether a diary is recorded on weekends or weekdays to have a sufficiently large sample to provide reliable estimates. We estimate the FE-IV model using a Two-Stage Least-Squares (2SLS) method. As documented above, we utilize daily daylight duration recorded on the diary date within each child's residential postcode. Given the variation in diary dates across individuals interviewed in the same wave (see Appendix Figure A4), employing these dates strengthens the identification assumption. This is because, due to differing diary dates and daily solar cycles, individuals within the same postcode may experience varying levels of sunlight exposure. As the treatment varies within individuals over time, standard errors are clustered at the individual level to address potential serial correlation (Cameron & Miller 2015). A robustness check using standard errors clustered at both individual and postcode levels yields largely similar results.

From an initial sample of about 55 thousand TUDs collected across Waves 1 through 8, we exclude TUDs with obviously incorrect entries or incomplete information. We also exclude TUDs with missing information on basic explanatory variables that we control for in the regressions. Final sample sizes aggregated across the 8 Waves vary by empirical models or developmental outcomes considered. For example, the sample size used to examine the effect of sleep duration on the "Social development" outcome includes 45,137 complete TUDs, from 8,222 unique children.

---

motivated by the fact that within a given survey wave, nearly all participating children were interviewed in the same calendar year.

[17] The absence of day-of-the-week weights in all LSAC waves precluded the use of weighted analysis (Mohal *et al.* 2021).

3.3. Sample representativeness

There is a concern in our study design is that the daylight duration could affect participants' propensity to complete a diary. We address this by estimating a probit model. The dependent variable indicates completion of a time use diary, determining inclusion in our final sample. Explanatory variables include basic demographics and daylight duration recorded on scheduled diary date (available for all participants, regardless of completion). Analysis is restricted to wave 4 onwards because scheduled diary dates are not available in prior waves (Mohal *et al.* 2021). Consistent with previous studies (Baxter 2007; Nguyen *et al.* 2021), Appendix Table A5 shows statistically significant differences in some demographics like age, gender, and socio-economic background between included and excluded samples. However, the small pseudo-R2 (0.05) suggests minimal quantitative impact. Crucially, the estimate of daylight duration variable is statistically insignificant ($p = 0.28$), alleviating the concern about daylight-driven selection bias.

4. Main empirical results

4.1. Descriptive analysis

Table 1 reports summary statistics for the main explanatory variables and outcomes by sleep duration sub-groups. It shows that individuals who sleep longer (i.e., individuals with sleep duration ≥ median) tend to be younger, female, born to mothers who have lower education or mothers who were born in Australia or born overseas in an English-Speaking-Background (ESB) country, to have fewer siblings or to live in two-parent families. Table 1 also indicates that individuals who sleep longer do better in some general development or behavioural outcomes as measured by PedsQL (all three sub-scales) or SDQ (two sub-scales: Emotional and Peer). By contrast, children who sleep longer tend to have lower scores for other behavioural outcomes such as Pro-sociality, Hyperactivity, Conduct or SDQ Overall. Moreover, individuals with a longer sleep duration have lower BMI, a lower probability of being overweight or higher waist-to-height ratios. They are more likely to have better self-reported health conditions or lower health-related expenditures. Contrarily, individuals who sleep longer have lower test scores in all cognitive domains. However, it is important to note that this simple comparison does not account for observable or unobservable characteristics, and reverse causality. We address these issues directly in the following sections.

4.2. Main results

FE and FE-IV results are reported in Tables 2, 3 and 4.[18][19] FE results for general development and behavioural outcomes, reported in odd columns of Table 2, show a statistically significant (at least at 5% level) and positive estimate of sleep duration on Emotional development, PedsQL Overall, Emotional symptoms, Conduct and SDQ Overall. These significant estimates suggest that sleeping longer benefits such developmental outcomes. Similarly, the statistically significant FE estimates of sleep duration on health-related outcomes, reported in odd columns of Table 3, suggest sleeping longer offers some health benefits. Specifically, health benefits include a reduction in BMI score, a reduced risk of being overweight or having any health condition, and a higher probability of having excellent health. However, apart from a

[18] Three observations support the validity of using a FE model despite potential concerns about insufficient variable variation. Firstly, substantial within-individual variation for instrumental variable and child development outcomes exists within our data, as evidenced by large standard deviations reported in Appendix Table A1. Secondly, standard errors for the sleep duration variable are consistently greater in pooled OLS regressions (Appendix Tables A8, A9 and A10) compared to FE models (Tables 2, 3 and 4), aligning with Wooldridge (2010)'s observation that insufficient variation wouldn't lead to such a pattern. Finally, unreported F-statistics from a Hausman-style test confirm that FE models are consistently preferred over OLS models, further validating our approach.

[19] Results from the first and second stage regressions are reported in Appendix Table A6 and Appendix Table A7, correspondingly. The results are largely as expected and in line with that in previous studies (Le & Nguyen 2017, 2018; Nguyen *et al.* 2023). For instance, child ages are strongly associated with various development outcomes. Moreover, children in two-parent families have better developmental outcomes. However, there is little evidence suggesting that child development outcomes vary by seasonal factors, as measured by quarter or day-of-week dummies. We also report results from pooled OLS (POLS) and IV regressions where we do not control for individual fixed effects in Appendix Table A8, Appendix Table A9 and Appendix Table A10. In these regressions, we additionally control for a list of time-invariant variables such as child gender, Aboriginal status, low birthweight status, cohort dummy, and maternal migration statuses. As compared to FE estimates, POLS estimates are more pronounced in terms of the statistical significance or magnitude. Moreover, while the POLS estimates suggest a highly statistically significant and negative association between sleep duration and all cognitive outcomes, the FE estimates indicate a statistically insignificant relationship. As compared to IV estimates, FE-IV estimates tend to be more statistically significant for some outcomes such as PedsQL Overall, BMI or overweight.

marginally statistically significant (at 10% level) and positive estimate of sleep duration on Grammar, FE estimates are not statistically significant for all considered cognitive outcomes (see odd columns in Table 4), suggesting that sleeping more may not statistically significantly improve cognitive skills in children and adolescents.

Before turning to the results from FE-IV regressions, we briefly discuss the results from the first stage regressions. The results, reported in Appendix Table A6, show that children and adolescents sleep statistically significantly less on days with longer daylight duration. Specifically, the highly statistically significant (at 1% level) FE estimate of daylight duration indicates that an increase of one hour in daylight duration is associated with a decrease of 3.51 minutes (or 0.06 hours) in sleep duration per day. This estimate is quite substantial in magnitude since an increase of 6 hours in daylight duration (i.e., the maximum variation of daylight duration observed in our data) can reduce sleep duration by 21 minutes per day (or 3.33 % of sample mean). Estimates of other variables are as expected and in line with previous studies (Nguyen *et al.* 2021; Nguyen *et al.* 2022a).

FE-IV estimates, reported in even columns of Tables 2 to 4, present four main findings. First, the weak identification tests from FE-IV regressions produce large Kleibergen-Paap statistics (the lowest F statistic is 18, as for Social development) that compare favourably to the statistics reported in Stock and Yogo (2005). These test statistics thus reject the hypothesis of a weak instrument for all regressions. Second, applying a FE-IV estimator substantially changes the results for some developmental outcomes. Specifically, the FE-IV estimator noticeably increases the size of sleep duration impact on PedsQL Overall while preserving its statistical significance at 5% level. Thus, FE-IV results indicate a much more pronounced benefit (in terms of the magnitude) of sleeping longer on this general development outcome than previously observed with the FE results.

The FE-IV estimator changes the sleep duration estimates on BMI and overweight from negative to positive but reduces the statistical significance to the 10% level. The FE-IV results therefore indicate that sleeping more increases BMI scores, mainly by increasing the probability of being overweight, in children and adolescents. Consistent with the positive impact of sleep on these BMI-related measures, the FE-IV estimate shows that sleeping longer also statistically (at 5% level) increases MBS expenditures. Lastly, the FE-IV estimate on Spelling is now statistically significant at 5% level, suggesting that sleeping longer improves Spelling scores.

Third, the changes in the magnitude and statistical significance level in the estimates of sleep duration on the above-mentioned development outcomes are consistent with results from a Hausman test which indicate that sleep duration is endogenous when modelling these outcomes (see Hausman test statistics reported in Tables 2 to 4). The results thus demonstrate that failing to adequately account for the endogeneity of sleep duration would lead to an inaccurate picture of the impact of sleep duration on these outcomes.

Fourth, FE-IV estimates of sleep duration on other development outcomes are not statistically significant at any conventional level. These statistically insignificant estimates are in line with the results from a Hausman test which suggest that we can model sleep duration and these outcomes independently. Therefore, the results from two Hausman-styled tests[20] suggest that a FE model would be suitable and hence preferred to identify the causal effects of sleep duration on these outcomes.

Fifth, adjusted *p*-values, which are calculated using raw *p*-values from the respective estimation model (i.e., FE or FE-IV) to address the multiple inference issue using the method proposed by Simes (1986) and Benjamini and Hochberg (1995) and reported in curly brackets in Tables

[20] Unreported statistics from a Hausman test suggest that the FE model is preferred to the pooled OLS model for all outcomes.

2 to 4, indicate that accounting for the multiple inference problem makes most estimates less statistically significant. For example, it turns FE estimates on PedsQL Overall, BMI, overweight, any ongoing health condition and grammar test scores to statistically insignificant. Likewise, when accounting for the multiple inference problem, none of the FE-IV estimates are statistically significant. However, the preferred adjusted *p*-values, which are calculated using the *p*-values from the preferred estimation results and reported in angle brackets, show that the preferred estimates remain statistically significant (at 10% level or higher) for some specific outcomes. These include Emotional development, Emotional symptoms, Conduct, SDQ Overall, and excellent health.

Overall, the preferred results from this section show that sleeping longer improves selected general developmental and behavioural outcomes, including Emotional development, Physical development, PedsQL Overall, Emotional symptoms, Conduct, and SDQ Overall. Sleeping more is also found to increase the probability of having excellent health or decrease the likelihood of having any ongoing health condition. By contrast, sleeping longer statistically significantly increases BMI scores, mainly by increasing the risk of being overweight, in children and adolescents. This causal evidence of sleep duration on BMI scores helps verify an unproven hypothesis that "sleep duration seems to influence weight gain in children" (Felső *et al.* 2017).

In line with the previous finding on children from the developing world (Jagnani 2022), our results also indicate some cognitive benefits of sleeping longer. However, the estimates, when statistically significant, appear quantitatively small.[21] Our finding of a null or relatively small positive impact of sleep duration on cognitive skills is consistent with that in the US study by Groen and Pabilonia (2019). Our finding is also in line with previous findings indicating that

[21] The discrepancy in findings between the two studies may be attributed to differences in study contexts (Australia vs. India) and modelling approaches.

educational activities are the most productive input for cognitive development (Fiorini & Keane 2014; Nguyen *et al.* 2020). These findings are in line with the premise that given the limit of 24 hours per day, to increase sleep duration individuals must reduce the time spent on other activities, especially educational activities.

5. Robustness checks and additional results

5.1. Robustness checks

This section checks whether our main findings are robust to: (i) different instruments, (ii) two-way clustering, (iii) the exclusion or inclusion of some potentially important time-variant variables, (iv) the inclusion of local weather conditions, and (v) a reduced-form regression approach. These checks address concerns about the validity of the instrument so they are applied to the FE-IV model only.

We first experiment with using daily sunrise time or daily sunset time in place of daily daylight duration as a separate instrument in the original FE-IV regression framework. We obtain largely similar results (reported in Panel B1 of Appendix Table A11) when employing daily sunrise time as an instrument. One notable difference is that the estimate of sleep duration on the waist-to-height ratio is (still) positive and statistically significant at the 5% level. Thus, the sleep duration estimates on BMI- and waist-based scores all indicate that sleeping longer increases the risk of being overweight. We also arrive at a broadly similar conclusion, although at a slightly lower precision level, when instrumenting sleep duration by daily sunset time (see Panel B2). This decrease in precision is consistent with the fact that children's sleep duration is least responsive to daily sunset time, resulting in the lowest F statistics (see F statistics reported at the bottom of each panel).

To address potential spatial correlation of daylight duration within postcodes, we conduct a robustness check by clustering standard errors at both individual and postcode levels in the FE-IV regressions. This explores the sensitivity of our findings to the previously employed individual-level clustering. Panel C demonstrates that two-way clustering (individual and postcode) does not present a definitively more conservative approach compared to individual-level clustering alone. While standard errors fluctuate in different cases, the overall findings remain unaffected, highlighting the robustness of our results. The third set of robustness checks consists of excluding or including some important time-variant variables. We start by excluding all individual and household level time-variant explanatory variables other than the child age-related variables and find our results (reported in Panel D1) are largely similar to the baseline results (reproduced in Panel A). Following exclusion of day-of-week fixed effects (Frazis & Stewart 2012), our findings remain largely unchanged (Panel D2).

As discussed in Section 4, the primary threat to the exclusion restriction would be that time-variant unobserved shocks are systematically associated with daily daylight duration measured on the diary date and development outcomes. Although it is challenging, if not impossible, to rule out the existence of confounding factors that would influence our estimates, we provide evidence that our estimates are insensitive to the inclusion of an extensive set of such time-variant variables. In particular, following Gibson and Shrader (2018), we additionally and separately control for those grouped activities which have been shown to be affected by daily daylight duration (Nguyen *et al.* 2022c) and, some of which may also influence the child development (Fiorini & Keane 2014; Nguyen *et al.* 2022a).[22] These include the daily time allocated to personal care, school, physical and media activities and results are reported in Panel E1, E2, E3 and E4 of Appendix Table A11, respectively. The results show that our

[22] The current literature, including studies using a laboratory setting or random experiments (Van Dongen *et al.* 2003; Bessone *et al.* 2021), faces an unresolved challenge to precisely pinpoint the impact of the time allocated to one activity (e.g., sleep as in our case) from that to other activities (Fiorini & Keane 2014; Nguyen *et al.* 2022a). This is mainly because individuals only have 24 hours per day to spend on all activities and hence increasing the time allocated to one activity must be met by decreasing the time devoted to other activities. Fortunately, we can identify the direct impact of the instrument (i.e., daylight duration) on the time allocated to all activities, allowing us to simultaneously control for other affected activities. Following Gibson and Shrader (2018), as we also lack exogenous variation in other activities, we employ a quasi-random design to estimate the causal impact of a change in sleep duration, allowing for endogenous adjustments in time allocated to other activities.

estimated sleep effects are not contaminated by the inclusion of these non-sleep activities, supporting that our estimates are likely to capture the true separate impacts of sleep.

Other time-variant variables would be sleep loss associated with longer daylight duration of related individuals, such as parents, teachers or classmates, in the same residential postcode. Quantifying the potential impact of these variables on our estimates is unfortunately not feasible due to the lack of reliable measures of sleep duration for related individuals in our data and the absence of prior evidence suggesting a spillover effect of their sleep on child development outcomes. Nevertheless, in light of the existing evidence on the impact of sleep on adult individuals' labour market outcomes (Gibson & Shrader 2018; Costa-Font & Fleche 2020; Costa-Font *et al.* 2024), we additionally control for the corresponding parent's employment status (results are reported in Panel E7) or household income (E8) in the regression. Moreover, for a similar reasoning (Giuntella *et al.* 2017; Mitrou *et al.* 2024), we separately control for the corresponding parent's general health (results are reported in Panel E5) and mental health (Panel E6). The stability of our estimates suggests that the potential influence of these time-varying factors is minimal, alleviating concerns about their impact on our findings.

Fourth, we additionally control for weather conditions recorded on the diary date (Panel E9) or cumulative weather conditions in the 365 days before the survey date (Panel E10).[23] The results show our findings are insensitive to the inclusion of these weather variables. Fifth, more evidence demonstrating the credibility of our findings is that the reduced form effects of daily daylight duration on child development outcomes (Panel F) display similar patterns as the 2SLS estimates.

[23] Daily weather conditions are measured by daily maximum temperature (and its square) and precipitation. To capture potential cumulative local weather exposure, following previous studies (Dell *et al.* 2014; Graff Zivin *et al.* 2018), we include the number of days with daily maximum temperature exceeding some thresholds and number of rainy days in the 365 days prior to the survey date.

5.2. Heterogeneity

We study heterogeneous effects of sleep with respect to: (i) child gender (male versus female) and (ii) child age (young versus old, identified relative to the median age of all individuals in the pooled sample).[24] To do this, we run separate regressions by subgroup distinguished by each of the above characteristics using a FE-IV model for all outcomes and report the results from this model if the exogeneity of sleep duration is rejected (i.e., when the p value of the Hausman test for exogeneity is equal to or smaller than 0.1). When the exogeneity of sleep duration is not rejected, we report results from the FE estimator.

Sub-population results by gender and age (reported in Table 5 and Table 6, respectively) suggest that sleep duration appears to have some differential effects by gender and age. For example, the impacts on some general developmental and behavioural outcomes, including Emotional development, Physical development, PedsQL Overall, Emotional symptoms, Conduct and SDQ Overall, are more pronounced for females or older individuals because the estimates of sleep duration are typically greater (i.e., more positive) or more likely to be statistically significant for them. By contrast, the sleep duration estimates on BMI and the probability of being overweight are positive and statistically significant (at 5% level) for males only, suggesting that the previously observed impacts of sleep duration on these BMI-based outcomes from the pooled sample are entirely driven by males.

6. Conclusion

This paper exploits variation in local daily daylight duration recorded on diary dates across the same individuals to assess the causal impacts of sleep duration on child development. Our

[24] We refrain from running separate regressions by other potentially important characteristics, such as maternal education level, mainly because we lack a statistical power, including a weak instrument issue, for some sub-groups or outcomes. Some findings in this section should be interpreted with caution because, for some sub-groups and outcomes, the instrument is relatively weak, probably because of the small sample sizes. While our empirical model is not ideal to explore the potentially non-linear causal effect of sleep duration on child development, we attempt to explore this possibility in Appendix B.

results show that longer daylight duration statistically significantly reduces sleep duration in children and adolescents. Employing a fixed effects instrumental variables approach, we find that sleeping longer improves selected general developmental and behavioural outcomes, such as Emotional development, Physical development, Health related quality of life (i.e., PedsQL Overall), Emotional symptoms, Conduct and behavioural and emotional difficulties generally (SDQ Overall). Our results also reveal that sleeping more increases the probability of having excellent health or decreases the likelihood of having any ongoing condition. By contrast, sleeping longer statistically significantly increases BMI scores, mainly by increasing the risk of being overweight. Moreover, while the beneficial effects of sleeping longer on general and behavioural outcomes are more pronounced for females or older individuals, the impact on BMI is only observed for males. Furthermore, the results show a null or relatively small positive impact of sleeping more on cognitive development. Finally, we find the results are robust to a range of sensitivity checks, including the utilization of alternative instrumental variables and the incorporation of additional time-varying observable control factors.

The findings presented in this paper highlight the importance of addressing potential endogeneity of sleep duration when quantifying its impact on child developmental outcomes. The findings of substantial health and development benefits of sleeping longer from this study reinforce the need to formulate policies to reduce sleep deprivation in young individuals, especially in females and adolescents who appear to benefit more. This paper also identifies undesirable effects of sleeping longer on some developmental outcomes, including increased BMI and a higher risk of being overweight for males, and these side effects should be considered when designing such policies.

The main objective of this paper is to provide evidence on the causal relationship between sleep duration and child development. With this said, it is beyond the scope of this paper to explore the precise mechanisms behind the estimated impacts. More and better research is needed to reveal potential underlying mechanisms. Moreover, our data and emperical model only capture the short-run impacts of sleep on child development, leaving the long-run effects for future research (Costa-Font *et al.* 2024). Similarly, our empirical model is not ideally suited to detect a non-linear causal impact of sleep duration on child development, causing some uncertainty around establishing an amount of sleep duration considered optimal for improving any given health or developmental outcome in young individuals. More studies, such as field experiments (Bessone *et al.* 2021), which have the power to find more definitive answers to this important question are necessary.

References

Baxter, J., 2007. Children’s time use in the Longitudinal Study of Australian Children: Data quality and analytical issues in the 4-year cohort. The Australian Institute of Family Studies Technical Paper No. 4

Becker, G.S., 1965. A Theory of the Allocation of Time. The Economic Journal 75, 493-517

Beebe, D.W., Field, J., Milller, M.M., Miller, L.E., LeBlond, E., 2017. Impact of Multi-Night Experimentally Induced Short Sleep on Adolescent Performance in a Simulated Classroom. Sleep 40, zsw035

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) 57, 289-300

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

Biddle, J.E., Hamermesh, D.S., 1990. Sleep and the Allocation of Time. Journal of Political Economy 98, 922-943

Cameron, A.C., Miller, D.L., 2015. A Practitioner’s Guide to Cluster-Robust Inference. Journal of Human Resources 50, 317-372

Cappuccio, F.P., D'Elia, L., Strazzullo, P., Miller, M.A., 2010. Sleep duration and all-cause mortality: a systematic review and meta-analysis of prospective studies. Sleep 33, 585

Cappuccio, F.P., Taggart, F.M., Kandala, N.-B., Currie, A., Peile, E., Stranges, S., et al., 2008. Meta-Analysis of Short Sleep Duration and Obesity in Children and Adults. Sleep 31, 619-626

Carrell, S.E., Maghakian, T., West, J.E., 2011. A's from Zzzz's? The Causal Effect of School Start Time on the Academic Achievement of Adolescents. American Economic Journal: Economic Policy 3, 62-81

Chaput, J.-P., Gray, C.E., Poitras, V.J., Carson, V., Gruber, R., Birken, C.S., et al., 2017. Systematic review of the relationships between sleep duration and health indicators in the early years (0–4 years). BMC Public Health 17, 855

Chaput, J.-P., Gray, C.E., Poitras, V.J., Carson, V., Gruber, R., Olds, T., et al., 2016. Systematic review of the relationships between sleep duration and health indicators in school-aged children and youth. Applied Physiology, Nutrition, and Metabolism 41, S266-S282

Chetty, R., Hendren, N., 2018. The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. The Quarterly Journal of Economics

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. Australian Government Department of Social Services, Canberra

Costa-Font, J., Fleche, S., 2020. Child Sleep and Mother Labour Market Outcomes. Journal of Health Economics 69, 102258

Costa-Font, J., Fleche, S., Pagan, R., 2024. The Labour Market Returns to Sleep. Journal of Health Economics 93, 102840

Currie, J., 2009. Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. Journal of Economic Literature 47, 87-122

Daraganova, G., Edwards, B., Sipthorp, 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

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

Edwards, F., 2012. Early to rise? The effect of daily start times on academic performance. Economics of Education Review 31, 970-983

Felső, R., Lohner, S., Hollódy, K., Erhardt, É., Molnár, D., 2017. Relationship between sleep duration and childhood obesity: Systematic review including the potential underlying mechanisms. Nutrition, Metabolism and Cardiovascular Diseases 27, 751-761

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

Frazis, H., Stewart, J., 2012. How to Think about Time-Use Data: What Inferences Can We Make about Long- and Short-Run Time Use from Time Diaries? Annals of Economics and Statistics, 231-245

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

Giuntella, O., Han, W., Mazzonna, F., 2017. Circadian Rhythms, Sleep, and Cognitive Skills: Evidence From an Unsleeping Giant. Demography 54, 1715-1742

Giuntella, O., Mazzonna, F., 2019. Sunset time and the economic effects of social jetlag: evidence from US time zone borders. Journal of Health Economics 65, 210-226

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

Groen, J.A., Pabilonia, S.W., 2019. Snooze or Lose: High School Start Times and Academic Achievement. Economics of Education Review 72, 204-218

Grossman, M., 1972. The Demand for Health: A Theoretical and Empirical Investigation. Columbia University Press for National Bureau of Economic Research Books, New York.

Hamermesh, D.S., 2016. What's to know about time use? Journal of Economic Surveys 30, 198-203

Heckman, J.J., Kautz, T., 2013. Fostering and measuring skills: Interventions that improve character and cognition. In: Heckman JJ, Humphries JE & Kautz T (eds.) The Myth of Achievement Tests: The GED and the Role of Character in American Life. The University of Chicago Press, Chicago, IL, pp. 341-430.

Heissel, J., Norris, S., 2018. Rise and Shine: The Effect of School Start Times on Academic Performance from Childhood through Puberty. Journal of Human Resources 53, 957-992

Hirshkowitz, M., Whiton, K., Albert, S.M., Alessi, C., Bruni, O., DonCarlos, L., et al., 2015. National Sleep Foundation's sleep time duration recommendations: methodology and results summary. Sleep Health 1, 40-43

Holbein, J.B., Schafer, J.P., Dickinson, D.L., 2019. Insufficient sleep reduces voting and other prosocial behaviours. Nature Human Behaviour 3, 492-500

Jagnani, M., 2022. Children's Sleep and Human Capital Production. The Review of Economics and Statistics forthcoming

Jin, L., Ziebarth, N.R., 2020. Sleep, health, and human capital: Evidence from daylight saving time. Journal of Economic Behavior & Organization 170, 174-192

Kajitani, S., 2021. The return of sleep. Economics & Human Biology 41, 100986

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

Lo, J.C., Ong, J.L., Leong, R.L.F., Gooley, J.J., Chee, M.W.L., 2016. Cognitive Performance, Sleepiness, and Mood in Partially Sleep Deprived Adolescents: The Need for Sleep Study. Sleep 39, 687-698

Matricciani, L., Blunden, S., Rigney, G., Williams, M.T., Olds, T.S., 2013. Children's sleep needs: is there sufficient evidence to recommend optimal sleep for children? Sleep 36, 527-534

Matricciani, L., Paquet, C., Galland, B., Short, M., Olds, T., 2019. Children's sleep and health: A meta-review. Sleep Medicine Reviews 46, 136-150

Meeus, J.H., 1999. Astronomical algorithms. Willmann-Bell, Incorporated, Richmond, VA.

Minges, K.E., Redeker, N.S., 2016. Delayed school start times and adolescent sleep: A systematic review of the experimental evidence. Sleep Medicine Reviews 28, 86-95

Mitrou, F., Nguyen, H.T., Le, H.T., Zubrick, S.R., 2024. The causal impact of mental health on tobacco and alcohol consumption: an instrumental variables approach. Empirical Economics 66, 1287-1310

Mohal, J., Lansangan, C., Gasser, C., Taylor, T., Renda, J., Jessup, K., et al., 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

Nguyen, H.T., Brinkman, S., Le, H.T., Zubrick, S.R., Mitrou, F., 2022a. Gender differences in time allocation contribute to differences in developmental outcomes in children and adolescents. Economics of Education Review 89, 102270

Nguyen, H.T., Christian, H., Le, H.T., Connelly, L., Zubrick, S.R., Mitrou, F., 2022b. Causal impact of physical activity on child health and development. Global Labor Organization (GLO) Discussion Paper No. 1081

Nguyen, H.T., Christian, H., Le, H.T., Connelly, L., Zubrick, S.R., Mitrou, F., 2023. The impact of weather on time allocation to physical activity and sleep of child-parent dyads. Science of The Total Environment 880, 163249

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., 2021. Weather and children's time allocation. Health Economics 30, 1559-1579

Nguyen, H.T., Zubrick, S., Mitrou, F., 2022c. The effects of sleep duration on child health and development. GLO Discussion Paper, No. 1150, Global Labor Organization (GLO), Essen

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

Ohayon, M.M., Carskadon, M.A., Guilleminault, C., Vitiello, M.V., 2004. Meta-Analysis of Quantitative Sleep Parameters From Childhood to Old Age in Healthy Individuals: Developing Normative Sleep Values Across the Human Lifespan. Sleep 27, 1255-1273

Paruthi, S., Brooks, L.J., D'Ambrosio, C., Hall, W.A., Kotagal, S., Lloyd, R.M., et al., 2016. Recommended Amount of Sleep for Pediatric Populations: A Consensus Statement of the American Academy of Sleep Medicine. Journal of Clinical Sleep Medicine 12, 785-786

Reppert, S.M., Weaver, D.R., 2002. Coordination of circadian timing in mammals. Nature 418, 935-941

Roenneberg, T., Kumar, C.J., Merrow, M., 2007. The human circadian clock entrains to sun time. Current Biology 17, R44-R45

Schlieber, M., Han, J., 2021. The Role of Sleep in Young Children's Development: A Review. The Journal of Genetic Psychology 182, 205-217

Simes, R.J., 1986. An improved Bonferroni procedure for multiple tests of significance. Biometrika 73, 751-754

Smith, A.C., 2016. Spring Forward at Your Own Risk: Daylight Saving Time and Fatal Vehicle Crashes. American Economic Journal: Applied Economics 8, 65-91

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.

Svensson, T., Saito, E., Svensson, A.K., Melander, O., Orho-Melander, M., Mimura, M., et al., 2021. Association of Sleep Duration With All- and Major-Cause Mortality Among Adults in Japan, China, Singapore, and Korea. JAMA Network Open 4, e2122837-e2122837

Taras, H., Potts-Datema, W., 2005. Sleep and Student Performance at School. Journal of School Health 75, 248-254

Van Dongen, H.P.A., Maislin, G., Mullington, J.M., Dinges, D.F., 2003. The Cumulative Cost of Additional Wakefulness: Dose-Response Effects on Neurobehavioral Functions and Sleep Physiology From Chronic Sleep Restriction and Total Sleep Deprivation. Sleep 26, 117-126

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 Journal 13, 366-378

Watson, N.F., Badr, M.S., Belenky, G., Bliwise, D.L., Buxton, O.M., Buysse, D., et al., 2015. Joint Consensus Statement of the American Academy of Sleep Medicine and Sleep Research Society on the Recommended Amount of Sleep for a Healthy Adult: Methodology and Discussion. Journal of Clinical Sleep Medicine 11, 931-952

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

Table 1: Summary statistics by sleep duration

Variable	Longer sleep duration group	Shorter sleep duration group	Longer sleep group - Shorter sleep group
	(1)	(2)	(3)
Child age	6.044	9.659	-3.614***
Male	0.501	0.515	-0.014***
Indigenous	0.021	0.021	0.000
Low birth weight	0.062	0.061	0.001
Mother has a certificate or diploma	0.384	0.404	-0.02***
Mother has a graduate degree	0.364	0.386	-0.022***
Mother ESB migrant	0.097	0.098	0.000
Mother NESB migrant	0.118	0.201	-0.082***
Number of siblings	1.453	1.520	-0.067***
Lived with both parents	0.860	0.816	0.044***
Social development	0.084	-0.010	0.094***
Emotional development	0.046	0.008	0.038***
Physical development	0.065	0.028	0.038***
PedsQL Overall	0.071	0.004	0.067***
Pro-sociality	-0.017	0.047	-0.065***
Hyperactivity	0.034	0.075	-0.041***
Emotional	0.112	0.016	0.096***
Conduct	-0.067	0.134	-0.201***
Peer	0.057	0.037	0.02**
SDQ Overall	0.037	0.091	-0.053***
BMI	0.418	0.512	-0.094***
Underweight	0.057	0.056	0.000
Overweight	0.207	0.239	-0.032***
Waist-for-height ratio	0.486	0.465	0.021***
Excellent health	0.556	0.518	0.038***
Any ongoing condition	0.352	0.442	-0.089***
Prescribed medicine	0.133	0.146	-0.013***
MBS ($1000)	0.204	0.253	-0.049***
PBS ($1000)	0.018	0.042	-0.024**
MBS and PBS ($1000)	0.222	0.295	-0.073***
Matrix reasoning	0.020	0.064	-0.045***
Reading	-0.064	0.365	-0.429***
Writing	-0.014	0.369	-0.384***
Spelling	-0.063	0.382	-0.445***
Grammar	-0.036	0.343	-0.379***
Numeracy	-0.058	0.433	-0.491***
Daylight duration (hour/day)	10.826	11.088	-0.262***
Number of observations	23,108	22,029	

Notes: Figures are sample means. Statistics are calculated using an estimated sample from the FE-IV regression for "Social development" as an outcome. Tests are performed on the significance of the difference between the sample mean for "Shorter sleep duration" individuals (identified as those with sleep duration < median of sleep duration among individuals included in the final sample) and "Longer sleep duration" individuals (sleep duration >=median). The symbol *denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Table 2: Impact of sleep duration 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	
Sleep duration (hour/day)	0.24	8.92	1.26***	15.94	-0.23	22.94	0.47**	26.34**	-0.16	-4.27
	[0.23]	[13.19]	[0.23]	[11.25]	[0.23]	[14.26]	[0.22]	[13.02]	[0.24]	[11.29]
Adjusted p-value	{0.45}	{0.82}	{0.00}	{0.46}	{0.47}	{0.46}	{0.14}	{0.38}	{0.53}	{0.92}
Preferred adjusted p-value	<0.40>		<0.00>			<0.23>		<0.14>	<0.53>	
Observations	45,137	45,137	46,141	46,141	45,132	45,132	43,539	43,539	40,421	40,421
Individuals	8,222	8,222	8,264	8,264	8,210	8,210	8,114	8,114	7,962	7,962
Mean of dep. variable	0.04	0.04	0.01	0.01	0.04	0.04	0.04	0.04	0.00	0.00
F-statistic of IV		18.13		24.30		19.20		20.10		21.98
Hausman test (p value)		0.51		0.18		0.08		0.03		0.71
	Hyperactivity		Emotional symptoms		Conduct		Peer problem		SDQ Overall	
Sleep duration (hour/day)	0.26	-0.90	0.79***	16.45	0.65***	8.98	0.20	-9.50	0.51**	3.16
	[0.21]	[9.89]	[0.25]	[12.00]	[0.23]	[10.62]	[0.25]	[12.00]	[0.20]	[9.55]
Adjusted p-value	{0.42}	{1.00}	{0.03}	{0.46}	{0.05}	{0.78}	{0.48}	{0.78}	{0.07}	{0.92}
Preferred adjusted p-value	<0.34>		<0.03>		<0.05>		<0.48>		<0.07>	
Observations	40,414	40,414	40,418	40,418	40,419	40,419	40,421	40,421	40,407	40,407
Individuals	7,960	7,960	7,961	7,961	7,962	7,962	7,962	7,962	7,959	7,959
Mean of dep. variable	0.04	0.04	0.05	0.05	0.02	0.02	0.03	0.03	0.04	0.04
F-statistic of IV		22.11		21.88		21.93		22.02		22.02
Hausman test (p value)		0.91		0.17		0.43		0.41		0.78

Notes: FE results are from the regression (1) while FE-IV results from models (1) and (2). Results (coefficient estimates and standard errors, which are reported in squared brackets) are multiplied by 100 for aesthetic purposes. "Adjusted *p*-values" to account for multiple inference issue calculated using the Simes-Benjamini-Hochberg method are in curly brackets. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration. Other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. The symbol *denotes significance at the 10% level, **at the 5% level, and ***at the 1% level.

Table 3: Impact of sleep duration 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-for-height ratio		Excellent health	
Sleep duration (hour/day)	-0.39**	22.96*	0.05	-0.14	-0.15*	8.38	-0.01	0.92	0.29**	1.24
	[0.20]	[11.76]	[0.05]	[3.13]	[0.09]	[5.18]	[0.01]	[0.59]	[0.12]	[6.26]
Adjusted p-value	{0.16}	{0.38}	{0.48}	{1.00}	{0.25}	{0.46}	{0.48}	{0.46}	{0.07}	{0.95}
Preferred adjusted p-value		<0.15>	<0.48>			<0.23>	<0.24>		<0.07>	
Observations	46,599	46,599	46,637	46,637	46,637	46,637	46,495	46,495	53,691	53,691
Individuals	8,321	8,321	8,324	8,324	8,324	8,324	8,311	8,311	8,699	8,699
Mean of dep. variable	0.46	0.46	0.06	0.06	0.22	0.22	0.48	0.48	0.55	0.55
F-statistic of IV		20.07		20.28		20.28		20.69		21.84
Hausman test (p value)		0.03		0.95		0.08		0.09		0.88
	Any ongoing condition		Prescribed medicine		MBS ($1000)		PBS ($1000)		MBS and PBS ($1000)	
Sleep duration (hour/day)	-0.25**	1.18	0.07	-3.73	-0.24	13.56**	-0.32	-7.48	-0.55	6.06
	[0.13]	[5.74]	[0.08]	[4.36]	[0.17]	[6.00]	[0.36]	[9.76]	[0.39]	[11.05]
Adjusted p-value	{0.16}	{0.95}	{0.48}	{0.78}	{0.34}	{0.38}	{0.48}	{0.78}	{0.34}	{0.90}
Preferred adjusted p-value	<0.14>		<0.48>			<0.13>	<0.48>		<0.27>	
Observations	41,362	41,362	53,686	53,686	53,000	53,000	53,001	53,001	53,000	53,000
Individuals	8,109	8,109	8,699	8,699	8,546	8,546	8,546	8,546	8,546	8,546
Mean of dep. variable	0.40	0.40	0.14	0.14	0.24	0.24	0.03	0.03	0.27	0.27
F-statistic of IV		26.65		21.92		22.00		22.04		22.00
Hausman test (p value)		0.80		0.37		0.01		0.44		0.54

Notes: FE results are from the regression (1) while FE-IV results from models (1) and (2). Results (coefficient estimates and standard errors, which are reported in squared brackets) are multiplied by 100 for aesthetic purposes. "Adjusted *p*-values" to account for multiple inference issue calculated using the Simes-Benjamini-Hochberg method are in curly brackets. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration. Other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. The symbol *denotes significance at the 10% level, **at the 5% level, and ***at the 1% level.

Table 4: Impact of sleep duration on cognitive outcomes - results from FE and FE-IV models

	FE	FE-IV	FE	FE-IV	FE	FE-IV
	(1)	(2)	(3)	(4)	(5)	(6)
	Matrix reasoning		Reading		Writing	
Sleep duration (hour/day)	-0.19	15.13	0.23	6.58	-0.40	0.06
	[0.36]	[10.35]	[0.21]	[6.46]	[0.29]	[9.42]
Adjusted p-value	{0.59}	{0.46}	{0.45}	{0.75}	{0.34}	{1.00}
Preferred adjusted p-value	<0.59>		<0.40>		<0.28>	
Observations	14,384	14,384	18,854	18,854	18,849	18,849
Individuals	3,519	3,519	5,503	5,503	5,506	5,506
Mean of dep. variable	0.08	0.08	0.17	0.17	0.20	0.20
F-statistic of IV		30.47		24.40		24.49
Hausman test (p value)		0.12		0.32		0.96
	Spelling		Grammar		Numeracy	
Sleep duration (hour/day)	0.18	11.55**	0.40*	3.59	0.14	-2.53
	[0.16]	[5.44]	[0.24]	[7.41]	[0.19]	[5.65]
Adjusted p-value	{0.43}	{0.38}	{0.25}	{0.90}	{0.49}	{0.90}
Preferred adjusted p-value		<0.14>	<0.23>		<0.49>	
Observations	18,881	18,881	18,876	18,876	18,742	18,742
Individuals	5,510	5,510	5,509	5,509	5,472	5,472
Mean of dep. variable	0.18	0.18	0.17	0.17	0.21	0.21
F-statistic of IV		23.96		23.95		23.49
Hausman test (p value)		0.02		0.67		0.64

Notes: FE results are from the regression (1) while FE-IV results from models (1) and (2). Results (coefficient estimates and standard errors, which are reported in squared brackets) are multiplied by 100 for aesthetic purposes. "Adjusted *p*-values" to account for multiple inference issue calculated using the Simes-Benjamini-Hochberg method are in curly brackets. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration. Other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. The symbol *denotes significance at the 10% level, **at the 5% level, and ***at the 1% level.

Table 5: Heterogenous impact of sleep duration by gender

	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Social development		Emotional development		Physical development		PedsQL Overall		Pro-sociality		Hyperactivity		Emotional symptoms	
Estimator	FE	FE	FE	FE	FE-IV	FE	FE-IV	FE	FE	FE	FE	FE	FE	FE
Sleep duration (hour/day)	0.39	-0.02	1.75***	0.62**	38.76*	-0.38	35.17*	0.03	-0.01	-0.34	0.41	0.11	1.50***	-0.07
	[0.33]	[0.33]	[0.32]	[0.31]	[22.59]	[0.33]	[19.77]	[0.31]	[0.32]	[0.36]	[0.29]	[0.32]	[0.36]	[0.36]
Observations	22,206	22,931	22,580	23,561	22,057	23,075	21,360	22,179	19,782	20,639	19,783	20,631	19,784	20,634
Individuals	4,028	4,194	4,046	4,218	4,019	4,191	3,973	4,141	3,902	4,060	3,902	4,058	3,902	4,059
Mean of dep. variable	0.04	0.03	-0.01	0.04	0.02	0.05	0.02	0.05	0.18	-0.17	0.26	-0.16	-0.01	0.10
F-statistic of IV					9.47		9.95							
Hausman test (p value)					0.04		0.04							
	Conduct		Peer problem		SDQ Overall		BMI		Underweight		Overweight		Waist-for-height ratio	
Estimator	FE	FE	FE-IV	FE	FE	FE	FE	FE-IV	FE	FE	FE	FE-IV	FE	FE
Sleep duration (hour/day)	0.87***	0.37	-27.22	-0.10	0.95***	0.00	-0.44	43.99**	0.07	0.01	-0.31**	21.06**	-0.01	-0.01
	[0.31]	[0.35]	[16.91]	[0.36]	[0.27]	[0.30]	[0.28]	[21.08]	[0.08]	[0.08]	[0.13]	[9.83]	[0.01]	[0.02]
Observations	19,782	20,637	19,786	20,635	19,781	20,626	22,771	23,828	22,787	23,850	22,787	23,850	22,760	23,735
Individuals	3,902	4,060	3,902	4,060	3,902	4,057	4,071	4,250	4,071	4,253	4,071	4,253	4,068	4,243
Mean of dep. variable	0.11	-0.06	0.10	-0.04	0.20	-0.10	0.41	0.52	0.06	0.06	0.23	0.21	0.48	0.48
F-statistic of IV			12.41					9.55				9.79		
Hausman test (p value)			0.06					0.00				0.00		

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 instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration. Other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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: Heterogenous impact of sleep duration by gender (continued)

	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Excellent health		Any ongoing condition		Prescribed medicine		MBS ($1000)		PBS ($1000)		MBS and PBS ($1000)	
Estimator	FE	FE	FE	FE	FE-IV	FE	FE-IV	FE	FE	FE	FE-IV	FE
Sleep duration (hour/day)	0.48***	0.10	-0.18	-0.30	-9.81	0.20	18.20*	-0.00	-0.07	-0.59	19.57*	-0.59
	[0.17]	[0.17]	[0.17]	[0.19]	[6.32]	[0.13]	[9.60]	[0.12]	[0.06]	[0.73]	[10.54]	[0.73]
Observations	26,278	27,413	20,271	21,091	26,279	27,407	25,874	27,126	25,875	27,126	25,874	27,126
Individuals	4,262	4,437	3,961	4,148	4,262	4,437	4,177	4,369	4,177	4,369	4,177	4,369
Mean of dep. variable	0.56	0.53	0.40	0.40	0.12	0.15	0.24	0.25	0.02	0.04	0.26	0.28
F-statistic of IV					11.63		12.27				12.27	
Hausman test (p value)					0.08		0.02				0.03	
	Matrix reasoning		Reading		Writing		Spelling		Grammar		Numeracy	
Estimator	FE	FE	FE-IV	FE	FE	FE-IV	FE-IV	FE-IV	FE	FE-IV	FE	FE
Sleep duration (hour/day)	-0.05	-0.30	13.69*	0.41	-0.40	32.99	10.23	34.67**	-0.00	24.13	0.12	0.16
	[0.52]	[0.48]	[8.02]	[0.30]	[0.41]	[21.36]	[6.31]	[16.65]	[0.34]	[16.47]	[0.26]	[0.27]
Observations	6,987	7,397	9,273	9,581	9,280	9,569	9,294	9,587	9,296	9,580	9,206	9,536
Individuals	1,700	1,819	2,710	2,793	2,719	2,787	2,722	2,788	2,722	2,787	2,689	2,783
Mean of dep. variable	0.15	0.02	0.26	0.08	0.37	0.04	0.26	0.10	0.29	0.06	0.16	0.26
F-statistic of IV			16.22			7.78	15.91	7.02		7.04		
Hausman test (p value)			0.06			0.06	0.07	0.00		0.10		

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 instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration. Other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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: Heterogenous impact of sleep duration by age

	Young	Old	Young	Old	Young	Old	Young	Old	Young	Old	Young	Old	Young	Old
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Social development		Emotional development		Physical development		PedsQL Overall		Pro-sociality		Hyperactivity		Emotional symptoms	
Estimator	FE	FE-IV	FE	FE-IV	FE	FE-IV	FE	FE-IV	FE	FE	FE	FE	FE	FE-IV
Sleep duration (hour/day)	-0.00	0.20	0.59**	38.64**	-0.15	32.48*	0.02	44.97**	0.08	0.06	0.26	0.07	0.39	30.10*
	[0.24]	[0.41]	[0.24]	[17.51]	[0.23]	[19.03]	[0.22]	[19.50]	[0.25]	[0.41]	[0.25]	[0.34]	[0.26]	[17.13]
Observations	22,909	22,228	23,083	22,291	22,694	21,664	21,851	20,895	20,262	20,159	20,258	20,156	20,260	19,380
Individuals	7,307	7,439	7,409	6,714	7,313	6,648	7,103	6,554	6,813	7,305	6,811	7,304	6,811	6,526
Mean of dep. variable	0.13	-0.05	-0.01	0.04	0.05	0.02	0.08	0.00	-0.12	0.12	-0.03	0.12	0.13	-0.03
F-statistic of IV				13.95		14.38		13.22						12.25
Hausman test (p value)				0.01		0.06		0.00						0.06
	Conduct		Peer problem		SDQ Overall		BMI		Underweight		Overweight		Waist-for-height ratio	
Estimator	FE	FE	FE	FE	FE	FE-IV	FE-IV	FE	FE	FE	FE-IV	FE	FE	FE-IV
Sleep duration (hour/day)	0.47	0.40	0.02	0.24	0.36	23.74*	44.07	-0.46*	0.06	0.08	22.98*	-0.16	-0.01	1.05*
	[0.29]	[0.34]	[0.27]	[0.44]	[0.23]	[13.13]	[28.01]	[0.26]	[0.07]	[0.09]	[13.61]	[0.14]	[0.01]	[0.63]
Observations	20,260	20,159	20,262	20,159	20,252	19,376	23,008	23,230	23,385	23,252	23,024	23,252	23,281	22,431
Individuals	6,812	7,305	6,812	7,305	6,810	6,524	7,124	7,510	7,488	7,512	7,127	7,512	7,471	6,716
Mean of dep. variable	-0.22	0.26	0.02	0.04	-0.06	0.15	0.40	0.52	0.05	0.06	0.20	0.25	0.50	0.45
F-statistic of IV						12.39	4.85				4.80			16.26
Hausman test (p value)						0.04	0.02				0.01			0.06

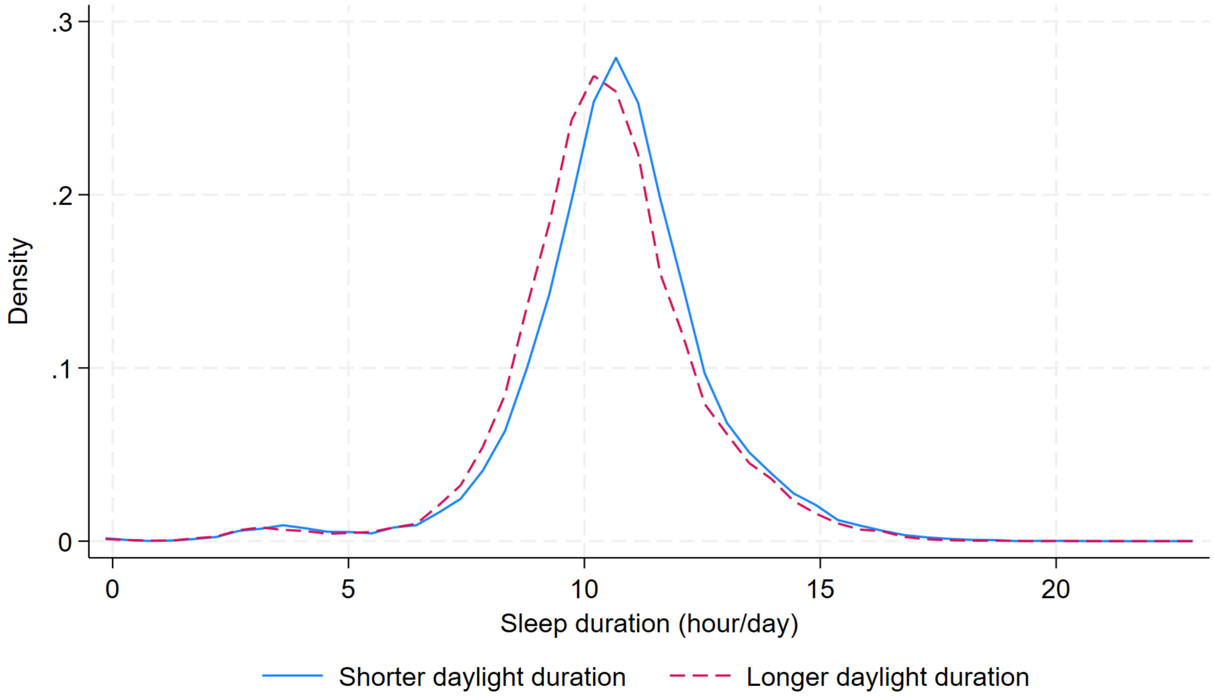
Notes: FE results are from the regression (1) while FE-IV results from models (1) and (2). F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration. Other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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: Heterogenous impact of sleep duration by age (continued)

	Young	Old	Young	Old	Young	Old	Young	Old	Young	Old	Young	Old
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Excellent health		Any ongoing condition		Prescribed medicine		MBS ($1000)		PBS ($1000)		MBS and PBS ($1000)	
Estimator	FE-IV	FE-IV	FE	FE	FE	FE	FE	FE	FE	FE	FE	FE
Sleep duration (hour/day)	-23.08	14.48**	-0.34**	-0.08	-0.03	0.26**	-0.22	9.74	0.01	0.09	-0.22	-0.19
	[18.29]	[6.67]	[0.14]	[0.24]	[0.11]	[0.13]	[0.29]	[6.44]	[0.03]	[0.11]	[0.29]	[0.22]
Observations	26,457	26,012	20,965	20,397	26,899	26,787	27,308	24,912	27,308	25,693	27,308	25,692
Individuals	7,230	6,931	7,083	7,405	7,675	7,709	7,628	6,823	7,628	7,603	7,628	7,603
Mean of dep. variable	0.58	0.51	0.30	0.50	0.12	0.15	0.23	0.26	0.02	0.04	0.24	0.30
F-statistic of IV	4.15	24.22						21.51				
Hausman test (p value)	0.10	0.02						0.10				
	Matrix reasoning		Reading		Writing		Spelling		Grammar		Numeracy	
Estimator	FE	FE	FE	FE	FE	FE	FE	FE	FE	FE	FE	FE
Sleep duration (hour/day)	-0.56*	0.18	0.23	-0.51	-0.29	-0.50	0.09	0.42	0.38	-0.50	0.02	-0.08
	[0.30]	[0.49]	[0.25]	[0.44]	[0.29]	[0.81]	[0.17]	[0.37]	[0.26]	[0.58]	[0.22]	[0.36]
Observations	7,569	6,815	9,544	9,310	9,534	9,315	9,550	9,331	9,546	9,330	9,520	9,222
Individuals	3,317	3,203	3,284	5,215	3,277	5,213	3,277	5,220	3,276	5,219	3,280	5,171
Mean of dep. variable	0.06	0.11	-0.33	0.68	-0.24	0.66	-0.37	0.74	-0.27	0.63	-0.34	0.78
F-statistic of IV												
Hausman test (p value)												

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 instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration. Other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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: Distributions of sleep duration by daylight duration



Notes: This figure reports univariate kernel density estimation of sleep duration (in hours per day) for a pooled sample of LSAC children with a valid TUD. "Longer daylight duration" indicates all TUDs recorded on dates with daylight duration at or above the median while "Shorter daylight duration" refers to those under the median.

Online Appendixes

of

Nguyen, H.T., Zubrick, S., Mitrou, F., 2024. The effects of sleep duration on child health and development. Journal of Economic Behavior & Organization 221, 35-51. https://doi.org/10.1016/j.jebo.2024.03.016

Appendix A reports additional results.

Appendix B reports results on non-linear impact of sleep duration.

Appendix Table A1: Variable description and summary statistics

Variable	Description	Mean	Min	Max	Standard deviations			N
					Overall	Between	Within	
Child age	SC age at the survey time (years)	6.84	0.00	16.00	4.51	2.67	3.94	53740
Male	Dummy = 1 if SC is a male, = 0 if female	0.51	0.00	1.00	0.50	0.50	0.00	53740
Indigenous	Dummy: = 1 if SC has Aboriginal/Torres Strait Islander origin, = 0 otherwise	0.02	0.00	1.00	0.15	0.17	0.00	53729
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	0.25	0.00	53673
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	53740
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	0.13	53740
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	53680
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	0.36	0.33	0.19	53680
Number of siblings	Number of siblings	1.41	0.00	11.00	1.00	0.96	0.44	53740
Lived with both parents	Dummy: = 1 if SC lived with both parents at the survey time, = 0 otherwise	0.85	0.00	1.00	0.36	0.34	0.18	53740
Sleep onset time	Time the SC went to sleep on the diary date (hour, 24-hour clock)	20.20	0.00	23.98	3.48	1.73	3.14	53713
Wakeup time	Time the SC woke up on the diary date (hour, 24-hour clock)	6.94	0.00	17.25	1.89	1.15	1.60	53739
Sleep duration	Total time spent on sleeping and napping per TUD day (hour/day)	10.55	0.00	22.75	2.03	1.16	1.75	53740
Personal care	Total time spent on personal care per TUD day (hour/day)	4.15	0.00	20.00	2.28	1.24	1.99	53740
School	Total time spent on school related activities per TUD day (hour/day)	1.86	0.00	19.75	2.81	1.23	2.58	53740
Educational activity	Total time spent on sleeping and napping per TUD day (hour/day)	1.00	0.00	14.42	1.33	0.74	1.14	53740
Physical activity	Total time spent on sleeping and napping per TUD day (hour/day)	2.56	0.00	23.75	2.29	1.12	2.06	53740
Chores	Total time spent on sleeping and napping per TUD day (hour/day)	0.34	0.00	11.50	0.77	0.38	0.69	53740
Media activity	Total time spent on sleeping and napping per TUD day (hour/day)	2.19	0.00	18.75	2.12	1.24	1.81	53740
Travel	Total time spent on sleeping and napping per TUD day (hour/day)	1.34	0.00	18.50	1.40	0.71	1.25	53740
Sunrise time	Sunrise time on the TUD date (hour, 24-hour clock)	6.73	4.69	7.81	0.54	0.40	0.39	53740
Sunset time	Sunset time on the TUD date (hour, 24-hour clock)	17.66	16.68	20.95	0.73	0.48	0.60	53740
Daylight duration	Daylight duration on the TUD date (hour/day)	10.93	8.98	15.01	1.07	0.66	0.90	53740

Notes: 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. "ADM" indicates linked administrative data sources.

Appendix Table A1: Variable description and summary statistics (continued)

Variable	Description	Mean	Min	Max	Standard deviations			N
					Overall	Between	Within	
Social development	PedsQL social development sub-scale - Standardized - P1	0.04	-4.94	1.06	0.97	0.79	0.63	45318
Emotional development	PedsQL emotional development sub-scale - Standardized - P1	0.01	-4.68	1.64	0.98	0.80	0.61	46298
Physical development	PedsQL physical development sub-scale - Standardized - P1	0.03	-5.46	1.13	0.94	0.75	0.65	45328
PedsQL Overall	Mean of above three PedsQL sub-scales - Standardized - P1	0.04	-5.85	1.53	0.95	0.82	0.59	43773
Pro-sociality	SDQ Pro-social behaviour scale - Standardized - P1	0.00	-4.63	1.04	0.99	0.83	0.60	40708
Hyperactivity	SDQ Hyperactivity and inattention scale (reversed) - Standardized - P1	0.04	-2.95	1.36	0.98	0.87	0.52	40702
Emotional	SDQ Emotional symptoms scale (reversed) - Standardized - P1	0.05	-4.51	0.95	0.96	0.81	0.60	40705
Conduct	SDQ Conduct problems scale (reversed) - Standardized - P1	0.02	-5.20	0.90	0.99	0.84	0.62	40706
Peer	SDQ Peer-relationship problems scale (reversed) - Standardized - P1	0.03	-5.29	0.92	0.98	0.83	0.60	40708
SDQ Overall	Mean of above five SDQ sub-scales - Standardized - P1	0.04	-5.18	1.55	0.98	0.88	0.50	40695
BMI	SC's Body Mass Index (gender- and age-standardized z-scores) - ITV	0.46	-4.97	4.85	1.11	1.00	0.54	46729
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	0.16	46765
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	46765
Waist-for-height ratio	SC's waist circumference at the time of survey (cm) - ITV	0.48	0.15	1.01	0.06	0.04	0.04	46630
Excellent health	Dummy: = 1 if SC's health is in excellent condition, - 0 otherwise - P1	0.55	0.00	1.00	0.50	0.36	0.37	53700
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	41567
Prescribed medicine	Dummy: = 1 if SC currently uses prescribed medicine, - 0 otherwise - P1	0.14	0.00	1.00	0.34	0.25	0.25	53695
MBS	Medicare Benefit Scheme amount during the survey year (AU$1000) - ADM	0.24	0.00	30.68	0.41	0.27	0.32	53002
PBS	Pharmaceutical Benefit Scheme amount during the survey year (AU$1000) - ADM	0.03	0.00	209.51	0.96	0.44	0.85	53003
MBS and PBS	Medicare and Pharmaceutical Benefit Scheme amount during the survey year (AU$1000) - ADM	0.27	0.00	212.74	1.07	0.55	0.92	53002
Matrix reasoning	Matrix reasoning test score - Standardized - ITV	0.04	-3.17	2.79	0.99	0.92	0.49	18110
Reading	NAPLAN Reading test score - Standardized - ADM	0.18	-5.20	3.94	0.96	0.79	0.57	20045
Writing	NAPLAN Writing test score - Standardized - ADM	0.20	-4.58	3.52	0.99	0.83	0.62	20036
Spelling	NAPLAN Spelling test score - Standardized - ADM	0.19	-3.33	3.44	0.97	0.81	0.57	20066
Grammar	NAPLAN Grammar test score - Standardized - ADM	0.18	-3.77	3.70	0.96	0.79	0.57	20062
Numeracy	NAPLAN Numeracy test score - Standardized - ADM	0.23	-5.10	4.11	0.97	0.80	0.58	19953

Notes: 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. "ADM" indicates linked administrative data sources.

Appendix Table A2: 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	2020/21
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)	BK	BK	BK	K	K	BK	B	B	
PedsQL measures - P1	K	BK	BK	BK	BK	BK	BK	B	
SDQ - P1	K	K	BK	BK	BK	BK	BK	B	
Weight - ITV	BK	BK	BK	BK	BK	BK	BK	BK	
Height - ITV	K	BK	BK	BK	BK	BK	BK	BK	
Waist circumference - ITV	K	BK	BK	BK	BK	BK	BK	BK	
Excellent health - P1	BK	BK	BK	BK	BK	BK	BK	BK	
Any ongoing condition - P1		BK	BK	BK	BK	BK	BK	BK	
Prescribed medicine - P1	BK	BK	BK	BK	BK	BK	BK	B	
MBS and PBS	BK	BK	BK	BK	BK	BK	BK	BK	
MR - ITV		K	K	BK	B	B			
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. SDQ = Strengths and Difficulties Questionnaire; MR = Matrix Reasoning; NAPLAN = National Assessment Program – Literacy and Numeracy test score; P1 - reported by Parent 1; P2 - reported by Parent 2; TC - reported by Teacher; SC – reported by Study Child; ITV – assessed by Interviewer.

Appendix Table A3: Raw correlations among sleep related variables in LSAC

Variable	Correlations												Summary statistics		
	Sleep duration (hour)	Sleep onset time (24-hour clock)	Wakeup time (24-hour clock)	SC sleep enough	SC's sleep quality	SC goes to bed at regular times	Bed time - School night (24-hour clock)	Bed time - No school next day (24-hour clock)	Sleep onset time - School night (24-hour clock)	Sleep onset time - No school next day (24-hour clock)	Wakeup time - School night (24-hour clock)	Wakeup time - No school next day (24-hour clock)	Mean	Min	Max
Sleep duration (hour) [a]	1.00												10.48	0.00	22.75
Sleep onset time (24-hour clock) [a]	0.29	1.00											20.08	0.00	23.98
Wakeup time (24-hour clock) [a]	0.25	0.23	1.00										6.94	0.00	17.25
SC sleep enough [b]	-0.08	-0.07	0.05	1.00									1.81	1.00	4.00
SC's sleep quality [c]	-0.07	-0.06	0.03	0.54	1.00								1.76	1.00	4.00
SC goes to bed at regular times [d]	-0.06	0.04	0.07	0.13	0.10	1.00							1.67	1.00	5.00
Bed time - School night (24-hour clock) [e]	-0.08	0.08	0.11			0.19	1.00						20.98	0.00	23.98
Bed time - No school next day (24-hour clock) [e]	0.09	0.22	-0.06	-0.15	-0.11	-0.09	0.15	1.00					19.69	0.00	23.98
Sleep onset time - School night (24-hour clock) [e]		0.23		-0.12	-0.11		0.49	0.26	1.00				20.98	0.00	23.98
Sleep onset time - No school next day (24-hour clock) [e]	0.12	0.27	-0.12	-0.21	-0.16	-0.14	0.06	0.70	0.27	1.00			18.51	0.00	23.98
Wakeup time - School night (24-hour clock) [e]	0.15	-0.03	0.33	0.03	0.03	0.14		-0.10	-0.09	-0.13	1.00		6.86	0.00	14.00
Wakeup time - No school next day (24-hour clock) [e]	0.05	-0.12	0.39	0.17	0.13	0.21	0.03	-0.27	-0.09	-0.33	0.34	1.00	8.57	0.00	15.00

Notes: Only correlation is statistically significant at 1% level is listed. [a] indicates variables which are derived from TUDs and described in the text.
[b] “SC sleep enough” is derived from responses to a question, asking the study child about “During the last month, do you think you usually got enough sleep?”. Responses are coded as: 1 Plenty; 2 Just enough; 3 Not quite enough; 4 Not nearly enough. This question is asked in waves 4 to 8 for K cohort and waves 6 to 8 for B cohort.
[c] “SC's sleep quality” is derived from responses to a question, asking the study child about “During the last month, how well do you feel you have slept in general?”. Responses are coded as: 1 Very well; 2 Fairly well; 3 Fairly badly; 4 Very badly. This question is asked in waves 4 to 8 for K cohort and waves 6 to 8 for B cohort.
[d] “SC goes to bed at regular times” is derived from responses to a question asking the corresponding parent about “Does the study child go to bed at regular times?”. Responses are coded as: 1 Always; 2 Usually; 3 Sometimes; 4 Rarely; 5 Never. This question is asked in waves 2 to 5 for K cohort and waves 2 to 7 for B cohort.
[e] These time variables are derived from responses to a respective question asking the study child “About what time do you go to bed on a usual school night?”, “About what time do you fall asleep on the nights you do not have school the next day?”, “About what time do you go to sleep on a usual school night?”, “About what time do you fall asleep on the nights you do not have school the next day?”, “About what time do you wake up in the morning on a usual school day?”, and “About what time do you wake up on the days you do not have school?”. This question is asked in waves 5 to 7 for K cohort and waves 6 to 8 for B cohort.

Appendix Table A4: Raw correlations among development outcome variables in LSAC

	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional	Conduct	Peer	SDQ Overall	BMI	Underweight	Overweight	Waist-for-height ratio	Excellent health	Any ongoing condition	Prescribed medicine	MBS ($1000)	PBS ($1000)	MBS and PBS ($1000)	Matrix reasoning	Reading	Writing	Spelling	Grammar	Grammar
Social development	1.00																									
Emotional development	0.52	1.00																								
Physical development	0.53	0.43	1.00																							
PedsQL Overall	0.81	0.75	0.86	1.00																						
Pro-sociality	0.22	0.21	0.17	0.24	1.00																					
Hyperactivity	0.31	0.30	0.20	0.32	0.33	1.00																				
Emotional	0.39	0.59	0.31	0.51	0.14	0.27	1.00																			
Conduct	0.28	0.36	0.20	0.33	0.42	0.49	0.29	1.00																		
Peer	0.52	0.34	0.27	0.44	0.28	0.31	0.42	0.32	1.00																	
SDQ Overall	0.50	0.53	0.34	0.54	0.63	0.75	0.62	0.73	0.66	1.00																
BMI	-0.07	-0.03	-0.08	-0.08	0.02	-0.02	-0.03	-0.03	-0.08	-0.04	1.00															
Underweight	-0.01		-0.01	-0.01			-0.02		-0.01	-0.01	-0.49	1.00														
Overweight	-0.08	-0.03	-0.09	-0.09	0.03		-0.05	-0.02	-0.09	-0.04	0.73	-0.13	1.00													
Waist-for-height ratio	-0.05	-0.06	-0.11	-0.09	-0.04	-0.07	-0.04	-0.15	-0.11	-0.12	0.60	-0.19	0.51	1.00												
Excellent health	0.21	0.20	0.24	0.27	0.11	0.14	0.23	0.12	0.19	0.23	-0.05	-0.04	-0.07	-0.07	1.00											
Any ongoing condition	-0.16	-0.18	-0.14	-0.19	-0.04	-0.12	-0.23	-0.08	-0.14	-0.18	0.04		0.04	-0.03	-0.19	1.00										
Prescribed medicine	-0.11	-0.12	-0.10	-0.13	-0.05	-0.09	-0.14	-0.07	-0.11	-0.14	0.03		0.04		-0.20	0.28	1.00									
MBS ($1000)	-0.14	-0.17	-0.14	-0.18	-0.05	-0.09	-0.21	-0.05	-0.13	-0.15	0.03	0.01	0.04	-0.01	-0.14	0.23	0.22	1.00								
PBS ($1000)	-0.02	-0.02	-0.02	-0.02			-0.01		-0.02	-0.02					-0.02	0.02	0.05	0.07	1.00							
MBS and PBS ($1000)	-0.06	-0.06	-0.05	-0.07	-0.01	-0.03	-0.06	-0.02	-0.05	-0.06			0.01		-0.06	0.09	0.11	0.38	0.95	1.00						
Matrix reasoning	0.10	0.02	0.05	0.07	0.05	0.16	0.08	0.12	0.08	0.15	-0.05		-0.06	-0.08	0.05	-0.03	-0.03	-0.03	-0.04	-0.04	1.00					
Reading	0.11	0.03	0.08	0.09	0.07	0.25	0.04	0.22	0.07	0.20	-0.02	0.02		-0.16	0.03	0.05		0.04			0.33	1.00				
Writing	0.14	0.06	0.09	0.12	0.10	0.29	0.06	0.23	0.12	0.25	-0.03	0.02	-0.02	-0.16	0.05	0.03	-0.02	0.02			0.26	0.71	1.00			
Spelling	0.11	0.05	0.07	0.09	0.05	0.26	0.03	0.21	0.07	0.19		0.02		-0.15	0.02	0.05		0.05			0.28	0.78	0.75	1.00		
Grammar	0.11	0.04	0.08	0.09	0.07	0.26	0.04	0.22	0.09	0.21	-0.04	0.02	-0.03	-0.16	0.04	0.04		0.03			0.34	0.82	0.72	0.80	1.00	
Numeracy	0.12	0.06	0.09	0.11	0.03	0.22	0.06	0.20	0.08	0.18			-0.02	-0.17	0.02	0.05		0.04			0.36	0.80	0.69	0.78	0.78	1.00

Notes: Only correlation is statistically significant at 1% level is listed.

Appendix Table A5: Daylight duration does not affect the time diary completion probability

	Probit estimates
	(1)
Daylight duration (hour/day)	0.07
	(0.06)
Child age	3.99*
	(2.32)
Child age squared	-0.01*
	(0.01)
Male	-10.78***
	(2.69)
Aboriginal	-20.46***
	(7.60)
Low birthweight	-4.31
	(5.39)
Mother with certificate/diploma [(a)]	4.62
	(3.40)
Mother with bachelor or higher degree [(a)]	15.05***
	(3.80)
Mother ESB migrant [(b)]	-2.03
	(4.90)
Mother NESB migrant [(b)]	-8.15*
	(4.24)
Number of siblings	-1.86
	(1.22)
Living with both parents	22.02***
	(3.00)
Observations	15,302
Number included in the sample	13,001
Pseudo R2	0.05
P value from a Wald test	0.28

Notes: Results are from a probit model and multiplied by 100 for aesthetic purposes. The dependent variable is equal to one if the child is in our final sample and zero otherwise. Sample: children of both cohorts surveyed in waves 4 to 6. [(a)] and [(b)] denote no qualification and native as the base group, respectively. Other explanatory variables include local socio-economic background variables, state/territory dummies, year dummies, month dummies, and day-of-week dummies. P value from a Wald test: P value of a Wald test for whether the estimate of daylight duration variable equals to zero. Robust standard errors are clustered at the individual level. The symbol *denotes significance at the 10% level, **at the 5% level, and ***at the 1% level.

Appendix Table A6: First stage regression results

Specification	POLS	FE
	(1)	(2)
Daylight duration (hour/day)	-3.65***	-3.51***
	[0.76]	[0.82]
Child age	-10.38***	-1.26
	[1.45]	[2.06]
Child age squared	0.22***	0.20***
	[0.04]	[0.05]
Male	-1.79	
	[1.11]	
Aboriginal	1.16	
	[4.02]	
Low birthweight	6.10***	
	[2.27]	
Mother with certificate/diploma [(a)]	-2.89**	5.13
	[1.45]	[3.50]
Mother with bachelor or higher degree [(a)]	-3.98***	2.78
	[1.52]	[4.53]
Mother ESB migrant [(b)]	-2.74	
	[1.83]	
Mother NESB migrant [(b)]	-6.60***	
	[1.99]	
Number of siblings	-0.03	-4.70***
	[0.56]	[1.25]
Living with both parents	-1.43	1.10
	[1.49]	[2.86]
Second quarter [(c)]	3.58	5.55
	[3.58]	[3.80]
Third quarter [(c)]	7.87**	6.34*
	[3.50]	[3.78]
Fourth quarter [(c)]	8.84**	5.14
	[3.61]	[4.00]
Monday [(d)]	-22.13***	-20.53***
	[1.59]	[1.71]
Tuesday [(d)]	-26.37***	-26.15***
	[1.54]	[1.65]
Wednesday [(d)]	-25.55***	-25.13***
	[1.52]	[1.61]
Thursday [(d)]	-27.61***	-28.07***
	[1.63]	[1.71]
Friday [(d)]	-52.85***	-51.98***
	[1.72]	[1.84]
Saturday [(d)]	-32.57***	-32.84***
	[1.71]	[1.76]
Holidays	7.57***	6.57***
	[1.38]	[1.46]
Observations	45,523	45,137
Number of unique individuals		8,222

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 60 for aesthetic purposes. [(a)], [(b)], [(c)] and [(d)] denotes having year 12 or below qualification, Australian born mother, first quarter and Sunday as the base group, respectively. Other variables include local socio-economic background variables, state/territory dummies, and TUD wave dummies. For OLS regression, we also control for child gender, Aboriginal status, low birthweight status, cohort dummy, and maternal migration statuses. Robust standard errors clustered at the individual level are in squared brackets. The symbol *denotes significance at the 10% level, **at the 5% level, and ***at the 1% level.

Appendix Table A7: Second stage regression results

	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional symptoms	Conduct	Peer problem	SDQ Overall	BMI	Underweight	Overweight
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Child age	-8.20***	-3.93**	4.21*	-1.43	23.00***	-5.63***	-2.24	15.50***	4.77**	9.11***	-7.45***	0.60	-1.79**
	[2.05]	[1.99]	[2.19]	[2.13]	[2.19]	[1.92]	[2.33]	[2.13]	[2.30]	[1.86]	[1.98]	[0.52]	[0.85]
Child age squared	0.51***	0.22***	-0.24***	0.11*	-1.13***	0.25***	0.10	-0.76***	-0.03	-0.41***	0.26***	-0.01	0.06**
	[0.06]	[0.06]	[0.07]	[0.06]	[0.06]	[0.05]	[0.06]	[0.05]	[0.06]	[0.05]	[0.06]	[0.02]	[0.03]
Mother education: Certificate [(a)]	-13.28***	-7.04**	-9.66**	-13.36***	-3.08	0.68	-7.27*	-1.92	-8.32**	-5.40*	6.76*	-0.20	2.03
	[3.65]	[3.54]	[3.82]	[3.80]	[3.76]	[3.31]	[3.91]	[3.91]	[3.75]	[3.23]	[3.77]	[0.79]	[1.56]
Mother education: Graduate [(a)]	-16.83***	-13.19***	-3.39	-13.70***	-4.07	3.27	-9.68*	-4.47	-8.19	-6.06	3.71	0.42	0.54
	[4.68]	[4.77]	[4.97]	[5.04]	[4.91]	[4.37]	[5.32]	[4.80]	[5.13]	[4.31]	[4.81]	[1.09]	[2.11]
Number of siblings	-0.92	0.90	4.77***	3.12*	-6.54***	3.50**	4.45***	-3.47**	-1.76	-0.60	-2.53	0.00	-0.90
	[1.66]	[1.57]	[1.60]	[1.72]	[1.64]	[1.44]	[1.72]	[1.65]	[1.73]	[1.43]	[1.58]	[0.37]	[0.68]
Living with both parents	13.70***	22.09***	10.97***	18.79***	14.71***	7.40***	16.73***	4.06	9.33***	15.29***	-8.18***	-0.04	-4.62***
	[3.06]	[3.09]	[3.19]	[3.13]	[3.19]	[2.78]	[3.37]	[3.13]	[3.15]	[2.75]	[3.09]	[0.73]	[1.32]
Second quarter [(b)]	-1.31	2.82	-3.27	-2.49	1.75	-0.18	-0.86	2.72	1.34	1.24	2.13	-1.20	1.30
	[4.25]	[3.71]	[4.60]	[4.35]	[3.62]	[3.10]	[3.88]	[3.94]	[3.81]	[3.12]	[3.80]	[1.06]	[1.60]
Third quarter [(b)]	-3.03	3.05	-5.25	-4.11	3.29	-0.20	-0.91	4.74	-1.38	1.45	3.01	-1.28	0.61
	[4.30]	[3.77]	[4.61]	[4.39]	[3.63]	[3.07]	[3.87]	[3.90]	[3.79]	[3.12]	[3.83]	[1.08]	[1.62]
Fourth quarter [(b)]	-4.55	2.49	0.36	-0.73	2.38	-0.40	-3.34	3.23	-5.89	-1.11	5.06	-1.19	1.85
	[3.65]	[3.55]	[4.06]	[3.95]	[3.48]	[3.00]	[3.78]	[3.73]	[3.74]	[3.01]	[3.49]	[0.95]	[1.48]
Monday [(c)]	4.40	5.78	10.40**	10.41**	-0.82	0.18	7.33	4.64	-1.79	2.74	6.30	0.49	2.65
	[4.68]	[4.19]	[5.29]	[4.71]	[4.23]	[3.73]	[4.51]	[3.98]	[4.51]	[3.58]	[4.25]	[1.14]	[1.87]
Tuesday [(c)]	3.60	5.55	9.64	10.36*	-1.88	0.38	8.11	4.77	-5.17	1.91	10.59**	-0.26	2.71
	[5.86]	[4.92]	[6.23]	[5.75]	[5.27]	[4.62]	[5.55]	[4.93]	[5.57]	[4.43]	[5.21]	[1.38]	[2.29]
Wednesday [(c)]	4.06	4.63	8.65	9.51*	-0.59	0.55	7.90	3.86	-4.62	2.17	8.99*	0.18	2.79
	[5.63]	[4.88]	[6.11]	[5.54]	[5.14]	[4.52]	[5.51]	[4.85]	[5.47]	[4.36]	[5.11]	[1.36]	[2.25]
Thursday [(c)]	4.11	5.53	11.19	11.46*	-1.45	0.06	8.18	4.56	-3.53	2.29	8.51	0.33	3.17
	[6.28]	[5.39]	[6.82]	[6.15]	[5.70]	[5.01]	[6.08]	[5.38]	[6.07]	[4.84]	[5.77]	[1.54]	[2.54]
Friday [(c)]	9.32	14.55	20.82*	23.57**	-2.14	0.58	15.54	11.61	-8.97	4.88	17.96*	0.31	6.33
	[11.45]	[9.72]	[12.30]	[11.22]	[10.35]	[9.07]	[10.99]	[9.74]	[10.99]	[8.75]	[10.14]	[2.70]	[4.47]
Saturday [(c)]	5.49	8.53	13.09*	14.55**	-1.05	0.17	10.76	6.07	-5.47	3.10	11.30*	-0.05	3.86
	[7.29]	[6.11]	[7.78]	[7.23]	[6.95]	[6.11]	[7.41]	[6.54]	[7.39]	[5.89]	[6.39]	[1.70]	[2.81]
Holidays	-1.77	-1.13	-2.86	-3.61*	-0.62	-1.64	-3.16	-1.54	0.64	-1.91	-2.29	-0.18	-1.51**
	[1.80]	[1.63]	[2.03]	[1.98]	[2.30]	[2.00]	[2.47]	[2.13]	[2.45]	[1.93]	[1.60]	[0.42]	[0.71]
Observation	45,137	46,141	45,132	43,539	40,421	40,414	40,418	40,419	40,421	40,407	46,599	46,637	46,637
Number of unique individuals	8,222	8,264	8,210	8,114	7,962	7,960	7,961	7,962	7,962	7,959	8,321	8,324	8,324

Notes: Results are from the FE-IV regression. Coefficient estimates and standard errors are multiplied by 100 for aesthetic purposes. [(a)], [(b)], and [(c)] denotes having year 12 or below qualification, first quarter and Sunday as the base group, respectively. Other variables include local socio-economic background variables, state/territory dummies, and TUD wave dummies. Robust standard errors clustered at the individual level are in squared brackets. The symbol *denotes significance at the 10% level, **at the 5% level, and ***at the 1% level.

Appendix Table A7: Second stage regression results (continued)

	Waist-for-	Excellent	Any	Prescribed	MBS	PBS	MBS and	Matrix	Reading	Writing	Spelling	Grammar	Numeracy
	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)
Child age	-1.41***	-0.27	-1.98*	-1.79*	-2.98**	-2.02	-5.00**	-7.58	13.37**	13.81*	8.45*	11.39*	4.47
	[0.10]	[1.28]	[1.11]	[0.92]	[1.26]	[1.72]	[2.05]	[11.50]	[5.81]	[7.73]	[4.65]	[6.45]	[5.02]
Child age squared	0.08***	0.01	0.04	0.08	0.14**	0.11	0.25**	0.17	-0.76***	-0.85**	-0.47**	-0.62**	-0.40*
	[0.00]	[0.07]	[0.03]	[0.05]	[0.06]	[0.09]	[0.10]	[0.65]	[0.26]	[0.37]	[0.22]	[0.29]	[0.23]
Mother education: Certificate [a]	0.39**	-1.89	2.82	1.38	1.60	-0.43	1.17	-6.47	-6.96*	-7.47	-4.79	-7.31*	-5.06
	[0.19]	[1.76]	[1.96]	[1.23]	[1.42]	[2.34]	[2.65]	[7.93]	[3.96]	[5.53]	[3.39]	[4.42]	[3.31]
Mother education: Graduate [a]	0.22	-3.34	0.26	2.35	4.27**	0.58	4.84	-12.67	-9.79*	-7.81	-1.71	-8.61	-6.67
	[0.25]	[2.33]	[2.58]	[1.63]	[2.06]	[2.48]	[3.11]	[10.67]	[5.52]	[7.09]	[4.69]	[6.18]	[4.18]
Number of siblings	-0.20**	0.35	-0.13	-0.02	0.48	1.93*	2.40**	-1.31	-2.16	0.01	0.47	-1.37	-0.94
	[0.08]	[0.68]	[0.86]	[0.48]	[0.63]	[1.01]	[1.16]	[3.33]	[1.52]	[2.30]	[1.38]	[1.98]	[1.39]
Living with both parents	-0.52***	5.12***	-2.81*	-0.42	-3.02**	2.19	-0.82	-2.73	2.36	6.31	4.83*	2.30	5.51**
	[0.15]	[1.46]	[1.66]	[0.97]	[1.42]	[2.15]	[2.52]	[6.74]	[3.30]	[5.38]	[2.75]	[3.80]	[2.76]
Second quarter [b]	0.07	-3.76**	3.34	1.59	0.54	0.89	1.42	4.53	3.55	7.15	6.40**	9.42**	4.14
	[0.20]	[1.73]	[2.19]	[1.18]	[1.48]	[1.42]	[1.87]	[9.86]	[3.19]	[5.19]	[2.55]	[3.80]	[2.95]
Third quarter [b]	-0.02	-5.32***	3.73*	1.70	0.62	1.75	2.36	9.08	1.23	3.77	3.95	5.90	3.79
	[0.20]	[1.76]	[2.20]	[1.21]	[1.57]	[2.97]	[3.25]	[10.22]	[3.13]	[5.26]	[2.51]	[3.75]	[2.89]
Fourth quarter [b]	-0.18	-2.36	6.18***	1.46	1.57	-1.53	0.01	11.48	3.98	5.89	5.78**	8.24**	2.72
	[0.19]	[1.79]	[2.12]	[1.27]	[1.56]	[1.48]	[2.00]	[11.10]	[3.47]	[5.61]	[2.78]	[4.08]	[3.19]
Monday [c]	0.24	0.44	0.63	-1.19	4.18**	-1.68	2.49	5.01	3.39	-0.91	4.10	1.13	-2.55
	[0.21]	[1.95]	[2.39]	[1.34]	[1.95]	[2.52]	[3.06]	[4.24]	[3.00]	[4.39]	[2.54]	[3.45]	[2.58]
Tuesday [c]	0.40	0.98	1.75	-1.97	5.41**	-0.17	5.23**	7.64	2.81	-1.29	5.36*	2.65	-1.06
	[0.26]	[2.37]	[2.82]	[1.66]	[2.38]	[1.42]	[2.58]	[5.43]	[3.47]	[4.93]	[2.92]	[3.92]	[2.98]
Wednesday [c]	0.38	1.06	0.75	-1.91	5.60**	-3.10	2.49	10.49**	2.79	0.35	4.20	2.42	-2.10
	[0.26]	[2.38]	[2.77]	[1.65]	[2.31]	[3.97]	[4.44]	[5.02]	[3.30]	[4.75]	[2.78]	[3.75]	[2.82]
Thursday [c]	0.42	0.49	1.00	-1.55	5.23*	-2.29	2.94	6.94	3.52	1.08	5.34*	2.41	-0.99
	[0.29]	[2.68]	[3.01]	[1.87]	[2.69]	[3.69]	[4.38]	[5.02]	[3.38]	[5.00]	[2.89]	[3.91]	[3.00]
Friday [c]	0.69	1.61	2.57	-2.67	10.45**	-5.52	4.92	11.63	7.84	-1.59	10.22*	5.34	-2.48
	[0.51]	[4.71]	[5.33]	[3.27]	[4.48]	[6.96]	[7.97]	[8.17]	[6.44]	[9.47]	[5.44]	[7.38]	[5.58]
Saturday [c]	0.39	0.28	1.68	-2.23	6.51**	-3.10	3.40	10.72	3.56	-1.00	5.55	3.29	-2.51
	[0.32]	[3.02]	[3.42]	[2.11]	[2.93]	[4.30]	[4.99]	[7.01]	[4.52]	[6.55]	[3.84]	[5.19]	[3.92]
Holidays	-0.24***	0.35	0.59	0.91**	0.20	0.00	0.20	4.61	0.30	0.16	-1.15	-0.64	1.88**
	[0.08]	[0.58]	[0.98]	[0.43]	[0.61]	[0.69]	[0.89]	[3.10]	[1.06]	[1.59]	[0.91]	[1.20]	[0.91]
Observation	46,495	53,691	41,362	53,686	53,000	53,001	53,000	14,384	18,854	18,849	18,881	18,876	18,742
Number of unique individuals	8,311	8,699	8,109	8,699	8,546	8,546	8,546	3,519	5,503	5,506	5,510	5,509	5,472

Notes: Results are from the FE-IV regression. Coefficient estimates and standard errors are multiplied by 100 for aesthetic purposes. [a], [b], and [c] denotes having year 12 or below qualification, first quarter and Sunday as the base group, respectively. Other variables include local socio-economic background variables, state/territory dummies, and TUD wave dummies. Robust standard errors clustered at the individual level are in squared brackets. The symbol *denotes significance at the 10% level, **at the 5% level, and ***at the 1% level.

Appendix Table A8: Impact of sleep duration 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	
Sleep duration	0.73**	29.36*	3.03***	-1.47	0.71**	22.85	1.55***	24.07	1.17***	-4.51
(hour/day)	[0.30]	[16.81]	[0.31]	[14.34]	[0.30]	[15.79]	[0.31]	[15.80]	[0.34]	[14.25]
Adjusted p-value	{0.02}	{0.40}	{0.00}	{0.96}	{0.02}	{0.40}	{0.00}	{0.40}	{0.00}	{0.93}
Observations	45,526	45,965	46,517	46,975	45,543	45,997	43,969	44,393	40,933	41,302
Mean of dep. variable	0.03	0.03	0.01	0.01	0.03	0.03	0.03	0.03	0.00	0.00
F-statistic of IV		16.99		20.14		17.80		17.84		21.87
Hausman test (p value)		0.06		0.77		0.14		0.13		0.69
	Hyperactivity		Emotional symptoms		Conduct		Peer problem		SDQ Overall	
Sleep duration	1.46***	9.19	1.56***	11.57	1.74***	15.90	1.04***	-2.57	2.06***	9.16
(hour/day)	[0.33]	[14.28]	[0.34]	[13.95]	[0.34]	[14.02]	[0.33]	[14.36]	[0.33]	[13.91]
Adjusted p-value	{0.00}	{0.80}	{0.0}	{0.75}	{0.00}	{0.52}	{0.00}	{0.96}	{0.00}	{0.80}
Observations	40,927	41,296	40,930	41,299	40,931	41,300	40,933	41,302	40,920	41,289
Mean of dep. variable	0.04	0.04	0.04	0.05	0.02	0.02	0.03	0.03	0.04	0.04
F-statistic of IV		21.87		21.80		21.77		21.89		21.86
Hausman test (p value)		0.59		0.46		0.30		0.81		0.60

Notes: POLS results are from the regression (1) without controlling for individual FE. IV results from models (1) and (2) without controlling for individual FE. Results (coefficient estimates and standard errors, which are reported in squared brackets) are multiplied by 100 for aesthetic purposes. "Adjusted *p*-values" to account for multiple inference issue calculated using the Simes-Benjamini-Hochberg method are in curly brackets. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration. Other explanatory variables include child age (and its square), child gender, Aboriginal status, low birthweight status, cohort dummy, maternal completed qualification, maternal migration statuses, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. The symbol *denotes significance at the 10% level, **at the 5% level, and ***at the 1% level.

Appendix Table A9: Impact of sleep duration 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-for-height ratio		Excellent health	
Sleep duration (hour/day)	-1.99***	32.27*	0.24***	-1.18	-0.54***	8.90	-0.06***	1.53*	0.47***	2.22
	[0.35]	[18.61]	[0.07]	[3.52]	[0.13]	[6.71]	[0.01]	[0.81]	[0.14]	[6.18]
Adjusted p-value	{0.00}	{0.40}	{0.00}	{0.93}	{0.00}	{0.42}	{0.00}	{0.40}	{0.00}	{0.93}
Observations	46,965	47,430	47,002	47,467	47,002	47,467	46,866	47,329	54,000	54,523
Mean of dep. variable	0.46	0.46	0.06	0.06	0.22	0.22	0.48	0.48	0.54	0.54
F-statistic of IV		18.74		18.80		18.80		18.71		24.35
Hausman test (p value)		0.04		0.70		0.14		0.03		0.77
	Any ongoing condition		Prescribed medicine		MBS ($1000)		PBS ($1000)		MBS and PBS ($1000)	
Sleep duration (hour/day)	-0.34**	0.34	-0.05	-0.56	-0.34*	10.37**	-0.17	-8.75	-0.51**	1.61
	[0.15]	[7.39]	[0.10]	[4.29]	[0.20]	[4.87]	[0.15]	[11.05]	[0.25]	[12.02]
Adjusted p-value	{0.03}	{0.96}	{0.59}	{0.96}	{0.09}	{0.40}	{0.25}	{0.75}	{0.04}	{0.96}
Observations	41,788	42,155	53,995	54,518	53,271	53,782	53,272	53,783	53,271	53,782
Mean of dep. variable	0.40	0.40	0.14	0.14	0.24	0.24	0.03	0.03	0.27	0.27
F-statistic of IV		18.08		24.55		25.38		25.39		25.38
Hausman test (p value)		0.92		0.90		0.02		0.43		0.86

Notes: POLS results are from the regression (1) without controlling for individual FE. IV results from models (1) and (2) without controlling for individual FE. Results (coefficient estimates and standard errors, which are reported in squared brackets) are multiplied by 100 for aesthetic purposes. "Adjusted *p*-values" to account for multiple inference issue calculated using the Simes-Benjamini-Hochberg method are in curly brackets. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration. Other explanatory variables include child age (and its square), child gender, Aboriginal status, low birthweight status, cohort dummy, maternal completed qualification, maternal migration statuses, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. The symbol *denotes significance at the 10% level, **at the 5% level, and ***at the 1% level.

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

	POLS	IV	POLS	IV	POLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
	Matrix reasoning		Reading		Writing	
Sleep duration (hour/day)	-1.44***	20.61	-1.22***	17.73	-1.30***	16.03
	[0.48]	[13.72]	[0.37]	[11.51]	[0.38]	[12.12]
Adjusted p-value	{0.00}	{0.40}	{0.00}	{0.40}	{0.00}	{0.42}
Observations	18,241	18,402	20,124	20,261	20,121	20,260
Mean of dep. variable	0.04	0.03	0.18	0.18	0.20	0.20
F-statistic of IV		28.88		23.75		23.80
Hausman test (p value)		0.10		0.09		0.14
	Spelling		Grammar		Numeracy	
Sleep duration (hour/day)	-0.70**	26.35**	-0.92**	17.63	-1.62***	5.38
	[0.35]	[12.09]	[0.37]	[11.78]	[0.33]	[10.32]
Adjusted p-value	{0.05}	{0.40}	{0.02}	{0.40}	{0.00}	{0.87}
Observations	20,150	20,289	20,146	20,285	20,038	20,176
Mean of dep. variable	0.19	0.19	0.18	0.17	0.23	0.22
F-statistic of IV		23.79		23.82		24.18
Hausman test (p value)		0.01		0.10		0.51

Notes: POLS results are from the regression (1) without controlling for individual FE. IV results from models (1) and (2) without controlling for individual FE. Results (coefficient estimates and standard errors, which are reported in squared brackets) are multiplied by 100 for aesthetic purposes. "Adjusted *p*-values" to account for multiple inference issue calculated using the Simes-Benjamini-Hochberg method are in curly brackets. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration. Other explanatory variables include child age (and its square), child gender, Aboriginal status, low birthweight status, cohort dummy, maternal completed qualification, maternal migration statuses, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. The symbol *denotes significance at the 10% level, **at the 5% level, and ***at the 1% level.

Appendix Table A11: Impact of sleep duration on child development – Robustness checks

	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional symptoms	Conduct	Peer problem	SDQ Overall	BMI	Underweight	Overweight
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel A: Baseline													
Sleep duration (hour/day)	8.92	15.94	22.94	26.34**	-4.27	-0.90	16.45	8.98	-9.50	3.16	22.96*	-0.14	8.38
	[13.19]	[11.25]	[14.26]	[13.02]	[11.29]	[9.89]	[12.00]	[10.62]	[12.00]	[9.55]	[11.76]	[3.13]	[5.18]
Observations	45,137	46,141	45,132	43,539	40,421	40,414	40,418	40,419	40,421	40,407	46,599	46,637	46,637
Individuals	8,222	8,264	8,210	8,114	7,962	7,960	7,961	7,962	7,962	7,959	8,321	8,324	8,324
F-statistic of IV	18.13	24.30	19.20	20.10	21.98	22.11	21.88	21.93	22.02	22.02	20.07	20.28	20.28
Hausman test (p value)	0.51	0.18	0.08	0.03	0.71	0.91	0.17	0.43	0.41	0.78	0.03	0.95	0.08
Panel B1: Using different instrument - Sunrise time													
Sleep duration (hour/day)	11.32	14.99	22.96	25.90**	6.41	-3.36	19.14*	6.59	-6.18	6.52	20.49*	-1.68	8.73
	[13.22]	[11.14]	[14.78]	[12.87]	[10.57]	[9.19]	[11.30]	[9.73]	[11.04]	[8.94]	[12.19]	[3.38]	[5.51]
Observations	45,137	46,141	45,132	43,539	40,421	40,414	40,418	40,419	40,421	40,407	46,599	46,637	46,637
Individuals	8,222	8,264	8,210	8,114	7,962	7,960	7,961	7,962	7,962	7,959	8,321	8,324	8,324
F-statistic of IV	18.42	24.20	18.07	20.47	25.71	25.73	25.63	25.67	25.75	25.63	17.89	18.07	18.07
Hausman test (p value)	0.39	0.21	0.09	0.03	0.53	0.69	0.09	0.54	0.56	0.50	0.06	0.61	0.08
Panel B2: Using different instrument - Sunset time													
Sleep duration (hour/day)	5.86	17.14	22.91	26.90	-20.21	2.73	12.44	12.53	-14.46	-1.82	25.68*	1.56	7.99
	[17.65]	[15.23]	[18.09]	[17.59]	[18.04]	[14.66]	[17.82]	[16.91]	[18.32]	[14.29]	[14.86]	[3.78]	[6.31]
Observations	45,137	46,141	45,132	43,539	40,421	40,414	40,418	40,419	40,421	40,407	46,599	46,637	46,637
Individuals	8,222	8,264	8,210	8,114	7,962	7,960	7,961	7,962	7,962	7,959	8,321	8,324	8,324
F-statistic of IV	9.83	13.61	11.66	11.01	9.67	9.80	9.61	9.65	9.70	9.76	12.77	12.91	12.91
Hausman test (p value)	0.75	0.28	0.17	0.09	0.24	0.87	0.50	0.47	0.40	0.87	0.05	0.69	0.17
Panel C: Clustering at the individual and postcode levels													
Sleep duration (hour/day)	8.92	15.94	22.94*	26.34**	-4.27	-0.90	16.45	8.98	-9.50	3.16	22.96*	-0.14	8.38
	[13.19]	[11.09]	[13.29]	[12.52]	[11.33]	[10.01]	[11.79]	[10.53]	[11.95]	[9.52]	[11.74]	[3.14]	[5.23]
Observations	45,137	46,141	45,132	43,539	40,421	40,414	40,418	40,419	40,421	40,407	46,599	46,637	46,637
Individuals	8,222	8,264	8,210	8,114	7,962	7,960	7,961	7,962	7,962	7,959	8,321	8,324	8,324
F-statistic of IV	17.29	22.70	18.01	19.31	19.68	19.82	19.58	19.64	19.73	19.73	18.95	19.33	19.33
Hausman test (p value)	0.50	0.18	0.07	0.02	0.71	0.91	0.16	0.42	0.41	0.78	0.03	0.95	0.08

Notes: Results are from models (1) and (2), unless stated otherwise. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration, unless stated otherwise. Unless stated otherwise, other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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: Impact of sleep duration on child development – Robustness checks (continued)

	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional symptoms	Conduct	Peer problem	SDQ Overall	BMI	Underweight	Overweight
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel D1: Excluding individual and household level variables (except child age and its square)													
Sleep duration (hour/day)	8.72	15.03	20.89	24.71*	-4.26	-1.93	14.87	8.94	-9.27	2.37	23.61**	-0.04	8.71*
	[13.05]	[11.14]	[13.99]	[12.74]	[11.17]	[9.81]	[11.78]	[10.51]	[11.86]	[9.46]	[11.75]	[3.10]	[5.17]
Observations	45,258	46,262	45,250	43,657	40,532	40,525	40,529	40,530	40,532	40,518	46,742	46,784	46,784
Individuals	8,246	8,288	8,234	8,138	7,985	7,983	7,984	7,985	7,985	7,982	8,347	8,350	8,350
F-statistic of IV	18.46	24.66	19.53	20.51	22.47	22.60	22.37	22.43	22.51	22.51	20.31	20.61	20.61
Hausman test (p value)	0.51	0.20	0.11	0.04	0.71	0.82	0.22	0.43	0.42	0.84	0.02	0.98	0.07
Panel D2: Excluding day of the week dummies													
Sleep duration (hour/day)	8.76	16.77	24.34	28.52*	-4.83	-1.33	17.96	9.60	-10.91	3.08	24.79*	-0.03	8.67
	[14.71]	[12.46]	[15.97]	[14.68]	[12.69]	[11.11]	[13.59]	[11.95]	[13.54]	[10.71]	[13.15]	[3.43]	[5.71]
Observations	45,018	46,022	45,015	43,423	40,310	40,303	40,307	40,308	40,310	40,296	46,472	46,510	46,510
Individuals	8,201	8,244	8,190	8,093	7,943	7,941	7,942	7,943	7,943	7,940	8,299	8,302	8,302
F-statistic of IV	14.30	19.58	15.23	16.06	17.03	17.13	16.94	16.99	17.12	17.06	16.35	16.58	16.58
Hausman test (p value)	0.56	0.19	0.09	0.03	0.71	0.89	0.18	0.44	0.40	0.80	0.03	0.98	0.10
Panel E1: Adding more variables - Personal care time													
Sleep duration (hour/day)	8.17	14.40	20.54	24.07**	-3.65	-0.67	14.27	7.93	-8.06	2.89	20.28**	-0.12	7.38*
	[11.87]	[9.93]	[12.56]	[11.57]	[9.68]	[8.50]	[10.15]	[9.11]	[10.25]	[8.21]	[10.07]	[2.75]	[4.46]
Observations	45,137	46,141	45,132	43,539	40,421	40,414	40,418	40,419	40,421	40,407	46,599	46,637	46,637
Individuals	8,222	8,264	8,210	8,114	7,962	7,960	7,961	7,962	7,962	7,959	8,321	8,324	8,324
F-statistic of IV	24.60	33.81	26.31	26.82	32.65	32.74	32.45	32.52	32.63	32.59	28.38	28.89	28.89
Hausman test (p value)	0.49	0.17	0.08	0.03	0.72	0.92	0.17	0.41	0.42	0.76	0.03	0.95	0.08
Panel E2: Adding more variables - School time													
Sleep duration (hour/day)	10.48	18.79	26.89	30.59**	-4.73	-0.80	18.59	9.69	-10.15	3.71	27.23*	-0.34	9.78
	[15.25]	[13.09]	[16.84]	[15.37]	[12.74]	[11.16]	[13.70]	[12.00]	[13.55]	[10.77]	[14.25]	[3.66]	[6.18]
Observations	45,137	46,141	45,132	43,539	40,421	40,414	40,418	40,419	40,421	40,407	46,599	46,637	46,637
Individuals	8,222	8,264	8,210	8,114	7,962	7,960	7,961	7,962	7,962	7,959	8,321	8,324	8,324
F-statistic of IV	13.88	18.76	14.66	15.59	17.48	17.59	17.41	17.45	17.55	17.53	14.89	15.07	15.07
Hausman test (p value)	0.50	0.16	0.08	0.02	0.72	0.92	0.17	0.44	0.43	0.77	0.03	0.92	0.08

Notes: Results are from models (1) and (2), unless stated otherwise. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration, unless stated otherwise. Unless stated otherwise, other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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: Impact of sleep duration on child development – Robustness checks (continued)

	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional symptoms	Conduct	Peer problem	SDQ Overall	BMI	Underweight	Overweight
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel E3: Adding more variables - Physically active													
Sleep duration (hour/day)	11.74	19.49	29.44	33.66*	-5.24	-1.02	20.25	11.03	-11.03	4.09	27.40*	-0.07	10.08
	[16.72]	[13.95]	[18.45]	[17.19]	[13.56]	[11.84]	[14.65]	[12.81]	[14.44]	[11.46]	[14.93]	[3.82]	[6.51]
Observations	45,137	46,141	45,132	43,539	40,421	40,414	40,418	40,419	40,421	40,407	46,599	46,637	46,637
Individuals	8,222	8,264	8,210	8,114	7,962	7,960	7,961	7,962	7,962	7,959	8,321	8,324	8,324
F-statistic of IV	11.78	16.72	12.69	13.16	15.74	15.89	15.66	15.72	15.81	15.79	13.84	13.94	13.94
Hausman test (p value)	0.49	0.17	0.07	0.02	0.71	0.91	0.16	0.41	0.42	0.76	0.03	0.97	0.08
Panel E4: Adding more variables - Media time													
Sleep duration (hour/day)	9.32	16.88	24.10*	27.43**	-3.43	-0.08	17.35	9.12	-8.35	4.31	22.29*	-0.25	7.98
	[13.11]	[11.23]	[14.26]	[13.07]	[11.16]	[9.79]	[11.94]	[10.53]	[11.83]	[9.46]	[11.58]	[3.09]	[5.09]
Observations	45,137	46,141	45,132	43,539	40,421	40,414	40,418	40,419	40,421	40,407	46,599	46,637	46,637
Individuals	8,222	8,264	8,210	8,114	7,962	7,960	7,961	7,962	7,962	7,959	8,321	8,324	8,324
F-statistic of IV	18.39	24.60	19.46	20.30	22.42	22.54	22.31	22.37	22.46	22.45	20.50	20.70	20.70
Hausman test (p value)	0.48	0.15	0.06	0.02	0.77	0.97	0.15	0.42	0.46	0.69	0.03	0.92	0.09
Panel E5: Adding more variables - Corresponding parent's general health (5-point scale indicating if general health is excellent, very good, good, fair													
Sleep duration (hour/day)	8.33	16.03	23.33	26.49**	-4.81	-1.36	17.80	8.50	-7.38	3.66	24.51**	-0.55	8.14
	[13.14]	[11.27]	[14.26]	[12.97]	[11.78]	[10.32]	[12.57]	[11.02]	[12.44]	[9.97]	[11.89]	[3.11]	[5.14]
Observations	45,019	46,018	45,013	43,423	40,161	40,154	40,158	40,159	40,161	40,147	46,012	46,050	46,050
Individuals	8,216	8,257	8,203	8,108	7,936	7,934	7,935	7,936	7,936	7,933	8,263	8,266	8,266
F-statistic of IV	18.11	24.06	19.25	20.19	20.29	20.42	20.19	20.25	20.34	20.33	20.20	20.42	20.42
Hausman test (p value)	0.53	0.17	0.07	0.03	0.69	0.88	0.16	0.47	0.53	0.75	0.02	0.85	0.09
Panel E6: Adding more variables - Corresponding parent's mental health (K6 mental health scores)													
Sleep duration (hour/day)	6.60	14.74	22.81	25.03**	-2.83	1.58	17.50	7.62	-7.38	4.99	19.24*	0.48	7.62
	[12.96]	[10.92]	[13.99]	[12.73]	[11.34]	[9.92]	[12.09]	[10.60]	[11.98]	[9.57]	[10.80]	[2.94]	[4.82]
Observations	44,638	45,619	44,637	43,066	39,937	39,932	39,934	39,935	39,937	39,925	45,721	45,759	45,759
Individuals	8,193	8,234	8,183	8,083	7,926	7,924	7,925	7,926	7,926	7,923	8,250	8,253	8,253
F-statistic of IV	18.34	25.16	19.87	20.44	21.68	21.80	21.58	21.64	21.73	21.71	22.59	22.81	22.81
Hausman test (p value)	0.62	0.20	0.08	0.03	0.81	0.89	0.15	0.51	0.52	0.64	0.05	0.88	0.09

Notes: Results are from models (1) and (2), unless stated otherwise. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration, unless stated otherwise. Unless stated otherwise, other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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: Impact of sleep duration on child development – Robustness checks (continued)

	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional symptoms	Conduct	Peer problem	SDQ Overall	BMI	Underweight	Overweight
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel E7: Adding more variables - Corresponding parent's work status (full-time employed, part-time employed, or unemployed)													
Sleep duration (hour/day)	8.74	15.09	22.37	25.72**	-4.53	-0.99	15.57	8.90	-9.41	2.81	23.03**	-0.33	8.28
	[13.11]	[11.14]	[14.07]	[12.90]	[11.24]	[9.84]	[11.88]	[10.56]	[11.95]	[9.50]	[11.56]	[3.07]	[5.07]
Observations	45,093	46,095	45,086	43,494	40,374	40,367	40,371	40,372	40,374	40,360	46,548	46,586	46,586
Individuals	8,218	8,260	8,206	8,109	7,956	7,954	7,955	7,956	7,956	7,953	8,317	8,320	8,320
F-statistic of IV	18.34	24.59	19.57	20.30	22.20	22.32	22.10	22.16	22.25	22.23	20.82	21.02	21.02
Hausman test (p value)	0.51	0.20	0.08	0.03	0.70	0.90	0.20	0.43	0.41	0.81	0.02	0.90	0.08
Panel E8: Adding more variables - Household income (weekly income, measured in 2004 price)													
Sleep duration (hour/day)	8.76	15.86	22.94	26.27**	-4.42	-0.80	16.15	8.90	-9.75	2.98	23.12**	-0.22	8.58*
	[13.18]	[11.24]	[14.26]	[13.01]	[11.30]	[9.89]	[11.98]	[10.62]	[12.01]	[9.55]	[11.79]	[3.13]	[5.19]
Observations	45,132	46,136	45,127	43,534	40,414	40,407	40,411	40,412	40,414	40,400	46,591	46,629	46,629
Individuals	8,222	8,264	8,210	8,114	7,960	7,958	7,959	7,960	7,960	7,957	8,320	8,323	8,323
F-statistic of IV	18.15	24.34	19.22	20.12	22.00	22.13	21.90	21.96	22.05	22.04	20.04	20.24	20.24
Hausman test (p value)	0.51	0.18	0.08	0.03	0.70	0.91	0.18	0.43	0.40	0.80	0.03	0.93	0.07
Panel E9: Controlling for weather conditions on TUD date - Daily maximum temperature (and its square) and precipitation													
Sleep duration (hour/day)	18.25	8.21	29.86	30.04*	-2.49	-14.12	17.80	-0.45	-0.34	-1.10	18.46	2.33	7.69
	[17.66]	[14.42]	[19.37]	[16.81]	[13.88]	[12.72]	[14.77]	[13.05]	[14.08]	[11.66]	[14.23]	[3.97]	[6.47]
Observations	45,137	46,141	45,132	43,539	40,421	40,414	40,418	40,419	40,421	40,407	46,599	46,637	46,637
Individuals	8,222	8,264	8,210	8,114	7,962	7,960	7,961	7,962	7,962	7,959	8,321	8,324	8,324
F-statistic of IV	10.89	13.83	11.24	12.81	14.36	14.54	14.31	14.31	14.39	14.50	12.42	12.68	12.68
Hausman test (p value)	0.28	0.63	0.08	0.04	0.87	0.24	0.23	0.93	0.97	0.89	0.15	0.56	0.20
Panel E10: Controlling for cumulative weather conditions in the 365 days before the survey date - Number of days with daily maximum temperature exceeding given thresholds and number of rainy days													
Sleep duration (hour/day)	10.19	23.50**	23.61*	30.65**	-4.18	2.61	14.14	11.78	-1.84	6.52	21.48**	-1.79	7.64
	[11.60]	[10.85]	[12.69]	[12.24]	[9.85]	[8.53]	[10.26]	[9.03]	[10.37]	[8.35]	[10.64]	[2.84]	[4.72]
Observations	41,576	42,422	41,537	40,161	36,553	36,546	36,550	36,551	36,553	36,539	42,763	42,791	42,791
Individuals	8,063	8,110	8,057	7,958	7,801	7,799	7,801	7,801	7,801	7,798	8,165	8,166	8,166
F-statistic of IV	23.43	28.62	24.72	24.67	28.43	28.56	28.30	28.37	28.46	28.46	24.60	24.76	24.76
Hausman test (p value)	0.39	0.03	0.04	0.00	0.69	0.77	0.18	0.20	0.85	0.46	0.02	0.52	0.08

Notes: Results are from models (1) and (2), unless stated otherwise. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration, unless stated otherwise. Unless stated otherwise, other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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: Impact of sleep duration on child development – Robustness checks (continued)

	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional symptoms	Conduct	Peer problem	SDQ Overall	BMI	Underweight	Overweight
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Panel F: Reduced form													
Daily daylight duration (hour)	-0.52	-1.07	-1.39*	-1.65**	0.28	0.06	-1.08	-0.59	0.62	-0.21	-1.41**	0.01	-0.52*
	[0.76]	[0.73]	[0.80]	[0.73]	[0.74]	[0.65]	[0.75]	[0.69]	[0.77]	[0.63]	[0.65]	[0.19]	[0.30]
Observations	45,137	46,141	45,132	43,539	40,421	40,414	40,418	40,419	40,421	40,407	46,599	46,637	46,637
Individuals	8,222	8,264	8,210	8,114	7,962	7,960	7,961	7,962	7,962	7,959	8,321	8,324	8,324

Notes: Results are from models (1) and (2), unless stated otherwise. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration, unless stated otherwise. Unless stated otherwise, other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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: Impact of sleep duration on child development – Robustness checks (continued)

	Waist-for-height ratio	Excellent health	Any ongoing condition	Prescribed medicine	MBS ($1000)	PBS ($1000)	MBS and PBS ($1000)	MR	Reading	Writing	Spelling	Grammar	Numeracy
	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)
Panel A: Baseline													
Sleep duration (hour/day)	0.92	1.24	1.18	-3.73	13.56**	-7.48	6.06	15.13	6.58	0.06	11.55**	3.59	-2.53
	[0.59]	[6.26]	[5.74]	[4.36]	[6.00]	[9.76]	[11.05]	[10.35]	[6.46]	[9.42]	[5.44]	[7.41]	[5.65]
Observations	46,495	53,691	41,362	53,686	53,000	53,001	53,000	14,384	18,854	18,849	18,881	18,876	18,742
Individuals	8,311	8,699	8,109	8,699	8,546	8,546	8,546	3,519	5,503	5,506	5,510	5,509	5,472
F-statistic of IV	20.69	21.84	26.65	21.92	22.00	22.04	22.00	30.47	24.40	24.49	23.96	23.95	23.49
Hausman test (p value)	0.09	0.88	0.80	0.37	0.01	0.44	0.54	0.12	0.32	0.96	0.02	0.67	0.64
Panel B1: Using different instrument - Sunrise time													
Sleep duration (hour/day)	1.56**	-0.42	-4.84	-1.04	14.71*	-0.69	14.02	13.97	8.61	6.57	12.05**	0.40	-2.20
	[0.69]	[6.79]	[6.32]	[4.82]	[7.95]	[5.52]	[9.56]	[10.12]	[6.61]	[9.48]	[5.51]	[7.67]	[5.69]
Observations	46,495	53,691	41,362	53,686	53,000	53,001	53,000	14,384	18,854	18,849	18,881	18,876	18,742
Individuals	8,311	8,699	8,109	8,699	8,546	8,546	8,546	3,519	5,503	5,506	5,510	5,509	5,472
F-statistic of IV	18.38	17.66	22.04	17.72	18.10	18.10	18.10	33.35	23.38	24.06	23.06	23.09	22.73
Hausman test (p value)	0.01	0.92	0.46	0.82	0.04	0.94	0.10	0.15	0.19	0.46	0.02	1.00	0.68
Panel B2: Using different instrument - Sunset time													
Sleep duration (hour/day)	0.21	2.93	7.53	-6.45	12.37*	-14.46	-2.12	16.75	4.07	-8.13	10.93	7.54	-2.94
	[0.69]	[7.70]	[6.99]	[5.37]	[6.46]	[14.86]	[15.59]	[13.26]	[8.12]	[12.25]	[6.89]	[9.32]	[7.19]
Observations	46,495	53,691	41,362	53,686	53,000	53,001	53,000	14,384	18,854	18,849	18,881	18,876	18,742
Individuals	8,311	8,699	8,109	8,699	8,546	8,546	8,546	3,519	5,503	5,506	5,510	5,509	5,472
F-statistic of IV	13.17	13.93	18.99	13.98	13.77	13.81	13.77	16.69	15.16	14.74	14.76	14.72	14.40
Hausman test (p value)	0.75	0.73	0.25	0.20	0.02	0.31	0.92	0.17	0.63	0.52	0.08	0.44	0.67
Panel C: Clustering at the individual and postcode levels													
Sleep duration (hour/day)	0.92	1.24	1.18	-3.73	13.56**	-7.48	6.06	15.13	6.58	0.06	11.55**	3.59	-2.53
	[0.64]	[6.64]	[5.98]	[4.49]	[5.85]	[9.73]	[10.91]	[11.61]	[6.55]	[9.27]	[5.14]	[7.57]	[5.49]
Observations	46,495	53,691	41,362	53,686	53,000	53,001	53,000	14,384	18,854	18,849	18,881	18,876	18,742
Individuals	8,311	8,699	8,109	8,699	8,546	8,546	8,546	3,519	5,503	5,506	5,510	5,509	5,472
F-statistic of IV	19.56	20.16	25.56	20.27	20.58	20.61	20.58	30.69	25.00	25.57	25.04	25.04	23.56
Hausman test (p value)	0.13	0.89	0.81	0.39	0.01	0.44	0.53	0.16	0.33	0.96	0.02	0.67	0.62

Notes: Results are from models (1) and (2), unless stated otherwise. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration, unless stated otherwise. Unless stated otherwise, other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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: Impact of sleep duration on child development – Robustness checks (continued)

	Waist-for-height ratio	Excellent health	Any ongoing condition	Prescribed medicine	MBS ($1000)	PBS ($1000)	MBS and PBS ($1000)	MR	Reading	Writing	Spelling	Grammar	Numeracy
	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)
Panel D1: Excluding individual and household level variables (except child age and its square)													
Sleep duration (hour/day)	0.98*	0.96	1.57	-3.94	13.67**	-7.51	6.15	13.87	6.57	0.30	11.56**	3.76	-2.64
	[0.59]	[6.26]	[5.73]	[4.36]	[6.00]	[9.82]	[11.10]	[10.29]	[6.53]	[9.52]	[5.48]	[7.47]	[5.72]
Observations	46,640	53,830	41,467	53,825	53,157	53,158	53,157	14,432	18,905	18,901	18,933	18,928	18,791
Individuals	8,337	8,727	8,132	8,727	8,573	8,573	8,573	3,530	5,518	5,521	5,525	5,524	5,486
F-statistic of IV	21.06	21.84	26.81	21.91	22.02	22.05	22.02	30.50	23.90	24.17	23.67	23.66	23.02
Hausman test (p value)	0.07	0.91	0.75	0.35	0.01	0.44	0.53	0.15	0.32	0.94	0.02	0.65	0.63
Panel D2: Excluding day of the week dummies													
Sleep duration (hour/day)	1.01	0.99	1.46	-3.74	14.48**	-8.11	6.36	14.59	7.40	0.86	12.45**	4.61	-1.52
	[0.66]	[6.72]	[6.25]	[4.68]	[6.52]	[10.64]	[11.98]	[10.22]	[6.85]	[9.92]	[5.86]	[7.88]	[5.94]
Observations	46,368	53,548	41,279	53,543	52,857	52,858	52,857	14,353	18,814	18,809	18,841	18,836	18,702
Individuals	8,289	8,675	8,092	8,675	8,522	8,522	8,522	3,511	5,491	5,494	5,498	5,497	5,460
F-statistic of IV	16.82	18.79	22.08	18.86	19.06	19.08	19.06	31.23	21.44	21.28	20.81	20.73	20.56
Hausman test (p value)	0.10	0.92	0.78	0.41	0.01	0.44	0.55	0.13	0.28	0.90	0.02	0.59	0.78
Panel E1: Adding more variables - Personal care time													
Sleep duration (hour/day)	0.83	1.14	0.95	-3.11	10.76**	-6.37	4.39	14.42	5.89	0.00	10.08**	3.18	-2.14
	[0.52]	[5.11]	[5.14]	[3.53]	[4.57]	[8.18]	[9.15]	[9.75]	[5.66]	[8.20]	[4.62]	[6.46]	[5.04]
Observations	46,495	53,691	41,362	53,686	53,000	53,001	53,000	14,384	18,854	18,849	18,881	18,876	18,742
Individuals	8,311	8,699	8,109	8,699	8,546	8,546	8,546	3,519	5,503	5,506	5,510	5,509	5,472
F-statistic of IV	28.30	35.96	35.87	36.11	36.14	36.21	36.14	37.80	33.78	34.58	33.93	33.78	31.60
Hausman test (p value)	0.09	0.86	0.83	0.35	0.01	0.43	0.61	0.11	0.30	0.96	0.02	0.67	0.67
Panel E2: Adding more variables - School time													
Sleep duration (hour/day)	1.08	1.50	1.37	-4.39	15.25**	-8.47	6.77	16.67	7.46	0.11	12.94**	3.59	-2.72
	[0.70]	[7.20]	[6.68]	[5.04]	[7.05]	[11.06]	[12.54]	[11.49]	[7.28]	[10.62]	[6.27]	[8.36]	[6.39]
Observations	46,495	53,691	41,362	53,686	53,000	53,001	53,000	14,384	18,854	18,849	18,881	18,876	18,742
Individuals	8,311	8,699	8,109	8,699	8,546	8,546	8,546	3,519	5,503	5,506	5,510	5,509	5,472
F-statistic of IV	15.59	16.79	20.05	16.85	17.06	17.08	17.06	25.25	19.51	19.40	18.97	18.94	18.58
Hausman test (p value)	0.09	0.87	0.81	0.37	0.01	0.44	0.54	0.12	0.31	0.96	0.02	0.70	0.65

Notes: Results are from models (1) and (2), unless stated otherwise. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration, unless stated otherwise. Unless stated otherwise, other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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: Impact of sleep duration on child development – Robustness checks (continued)

	Waist-for-height ratio	Excellent health	Any ongoing condition	Prescribed medicine	MBS ($1000)	PBS ($1000)	MBS and PBS ($1000)	MR	Reading	Writing	Spelling	Grammar	Numeracy
	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)
Panel E3: Adding more variables - Physically active													
Sleep duration (hour/day)	1.08	2.49	1.16	-5.05	16.88**	-9.64	7.22	15.44	7.06	0.02	12.71**	3.86	-2.87
	[0.74]	[7.96]	[7.06]	[5.60]	[8.03]	[12.54]	[14.13]	[10.58]	[7.12]	[10.49]	[6.17]	[8.28]	[6.21]
Observations	46,495	53,691	41,362	53,686	53,000	53,001	53,000	14,384	18,854	18,849	18,881	18,876	18,742
Individuals	8,311	8,699	8,109	8,699	8,546	8,546	8,546	3,519	5,503	5,506	5,510	5,509	5,472
F-statistic of IV	14.34	13.82	18.12	13.88	14.08	14.11	14.08	30.83	20.88	20.36	19.84	19.79	20.16
Hausman test (p value)	0.10	0.79	0.83	0.35	0.01	0.44	0.57	0.12	0.32	0.96	0.02	0.67	0.63
Panel E4: Adding more variables - Media time													
Sleep duration (hour/day)	0.88	1.93	0.95	-3.77	12.90**	-6.89	6.01	15.13	6.45	0.10	11.46**	3.42	-2.66
	[0.58]	[6.07]	[5.71]	[4.22]	[5.75]	[9.09]	[10.40]	[10.41]	[6.46]	[9.45]	[5.44]	[7.42]	[5.66]
Observations	46,495	53,691	41,362	53,686	53,000	53,001	53,000	14,384	18,854	18,849	18,881	18,876	18,742
Individuals	8,311	8,699	8,109	8,699	8,546	8,546	8,546	3,519	5,503	5,506	5,510	5,509	5,472
F-statistic of IV	21.12	23.33	26.99	23.41	23.45	23.48	23.45	30.11	24.33	24.36	23.84	23.82	23.42
Hausman test (p value)	0.10	0.78	0.83	0.35	0.01	0.45	0.51	0.12	0.33	0.96	0.02	0.68	0.62
Panel E5: Adding more variables - Corresponding parent's general health (5-point scale indicating if general health is excellent, very good, good, fair													
Sleep duration (hour/day)	0.90	0.58	1.26	-4.22	14.06**	-7.64	6.41	15.16	6.73	0.39	12.56**	3.03	-2.07
	[0.59]	[6.27]	[5.80]	[4.40]	[6.01]	[9.78]	[11.06]	[10.81]	[6.69]	[9.62]	[5.70]	[7.66]	[5.83]
Observations	45,914	53,030	40,903	53,025	52,333	52,334	52,333	14,236	18,657	18,647	18,679	18,674	18,548
Individuals	8,252	8,639	8,062	8,639	8,483	8,483	8,483	3,491	5,453	5,454	5,458	5,457	5,423
F-statistic of IV	20.73	21.69	26.46	21.77	22.42	22.46	22.42	28.14	22.95	22.93	22.43	22.43	22.09
Hausman test (p value)	0.10	0.96	0.79	0.32	0.01	0.43	0.51	0.13	0.32	0.94	0.01	0.73	0.71
Panel E6: Adding more variables - Corresponding parent's mental health (K6 mental health scores)													
Sleep duration (hour/day)	0.80	1.69	1.60	-3.70	13.52**	-2.85	10.67	15.39	6.68	1.55	11.75**	3.56	-1.43
	[0.55]	[6.05]	[5.65]	[4.20]	[5.73]	[5.39]	[7.67]	[10.56]	[6.51]	[9.50]	[5.44]	[7.43]	[5.62]
Observations	45,621	52,664	40,648	52,658	51,981	51,982	51,981	14,076	18,552	18,542	18,574	18,569	18,443
Individuals	8,240	8,621	8,050	8,621	8,469	8,469	8,469	3,482	5,451	5,452	5,456	5,455	5,420
F-statistic of IV	23.13	23.48	28.10	23.71	24.46	24.50	24.46	30.15	24.50	24.58	24.05	24.04	24.00
Hausman test (p value)	0.12	0.81	0.74	0.36	0.01	0.62	0.12	0.12	0.31	0.84	0.02	0.66	0.78

Notes: Results are from models (1) and (2), unless stated otherwise. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration, unless stated otherwise. Unless stated otherwise, other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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: Impact of sleep duration on child development – Robustness checks (continued)

	Waist-for-height ratio	Excellent health	Any ongoing condition	Prescribed medicine	MBS ($1000)	PBS ($1000)	MBS and PBS ($1000)	MR	Reading	Writing	Spelling	Grammar	Numeracy
	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)
Panel E7: Adding more variables - Corresponding parent's work status (full-time employed, part-time employed, or unemployed)													
Sleep duration (hour/day)	0.94	1.20	1.36	-3.50	13.93**	-7.69	6.23	14.53	6.85	-0.02	11.47**	3.26	-2.49
	[0.58]	[6.33]	[5.73]	[4.39]	[6.02]	[10.03]	[11.28]	[10.29]	[6.57]	[9.53]	[5.50]	[7.49]	[5.72]
Observations	46,444	53,636	41,333	53,631	52,938	52,939	52,938	14,381	18,821	18,818	18,850	18,845	18,711
Individuals	8,307	8,696	8,108	8,696	8,543	8,543	8,543	3,519	5,488	5,492	5,496	5,495	5,458
F-statistic of IV	21.48	21.45	26.87	21.54	22.15	22.18	22.15	30.54	23.78	24.03	23.48	23.47	23.05
Hausman test (p value)	0.08	0.89	0.78	0.41	0.01	0.44	0.53	0.13	0.30	0.97	0.02	0.70	0.64
Panel E8: Adding more variables - Household income (weekly income, measured in 2004 price)													
Sleep duration (hour/day)	0.93	1.36	1.24	-3.73	13.60**	-7.50	6.09	14.91	6.68	0.04	11.54**	3.63	-2.49
	[0.59]	[6.27]	[5.74]	[4.36]	[6.00]	[9.77]	[11.07]	[10.28]	[6.44]	[9.38]	[5.42]	[7.38]	[5.63]
Observations	46,487	53,686	41,359	53,681	52,995	52,996	52,995	14,380	18,852	18,847	18,879	18,874	18,740
Individuals	8,310	8,699	8,109	8,699	8,545	8,545	8,545	3,518	5,503	5,506	5,510	5,509	5,472
F-statistic of IV	20.67	21.82	26.68	21.89	21.99	22.02	21.99	30.85	24.60	24.70	24.15	24.14	23.68
Hausman test (p value)	0.09	0.86	0.79	0.38	0.01	0.44	0.53	0.12	0.31	0.96	0.02	0.66	0.64
Panel E9: Controlling for weather conditions on TUD date - Daily maximum temperature (and its square) and precipitation													
Sleep duration (hour/day)	1.15	8.11	2.67	-2.53	16.41**	-8.13	8.30	19.47	7.22	-0.21	11.99**	3.95	-2.56
	[0.74]	[7.93]	[6.88]	[5.31]	[7.33]	[12.25]	[13.61]	[12.87]	[6.48]	[9.39]	[5.48]	[7.40]	[5.76]
Observations	46,495	53,691	41,362	53,686	53,000	53,001	53,000	14,384	18,854	18,849	18,881	18,876	18,742
Individuals	8,311	8,699	8,109	8,699	8,546	8,546	8,546	3,519	5,503	5,506	5,510	5,509	5,472
F-statistic of IV	13.18	14.00	18.16	13.99	13.53	13.51	13.53	20.81	24.55	24.65	24.02	24.02	22.48
Hausman test (p value)	0.08	0.31	0.67	0.62	0.00	0.51	0.50	0.10	0.27	0.98	0.02	0.63	0.64
Panel E10: Controlling for cumulative weather conditions in the 365 days before the survey date - Number of days with daily maximum temperature exceeding given thresholds and number of rainy days													
Sleep duration (hour/day)	1.03*	6.53	1.17	-3.64	9.04*	-2.47	6.57	15.23	6.61	0.16	11.51**	3.57	-2.56
	[0.53]	[6.10]	[5.75]	[4.12]	[5.06]	[6.14]	[8.03]	[10.47]	[6.45]	[9.41]	[5.44]	[7.41]	[5.65]
Observations	42,680	45,118	41,362	45,112	44,583	44,584	44,583	14,384	18,854	18,849	18,881	18,876	18,742
Individuals	8,150	8,288	8,109	8,287	8,159	8,159	8,159	3,519	5,503	5,506	5,510	5,509	5,472
F-statistic of IV	26.09	25.18	26.59	25.49	25.43	25.45	25.43	29.78	24.42	24.52	23.96	23.96	23.56
Hausman test (p value)	0.03	0.30	0.80	0.36	0.05	0.70	0.37	0.12	0.31	0.95	0.02	0.67	0.63

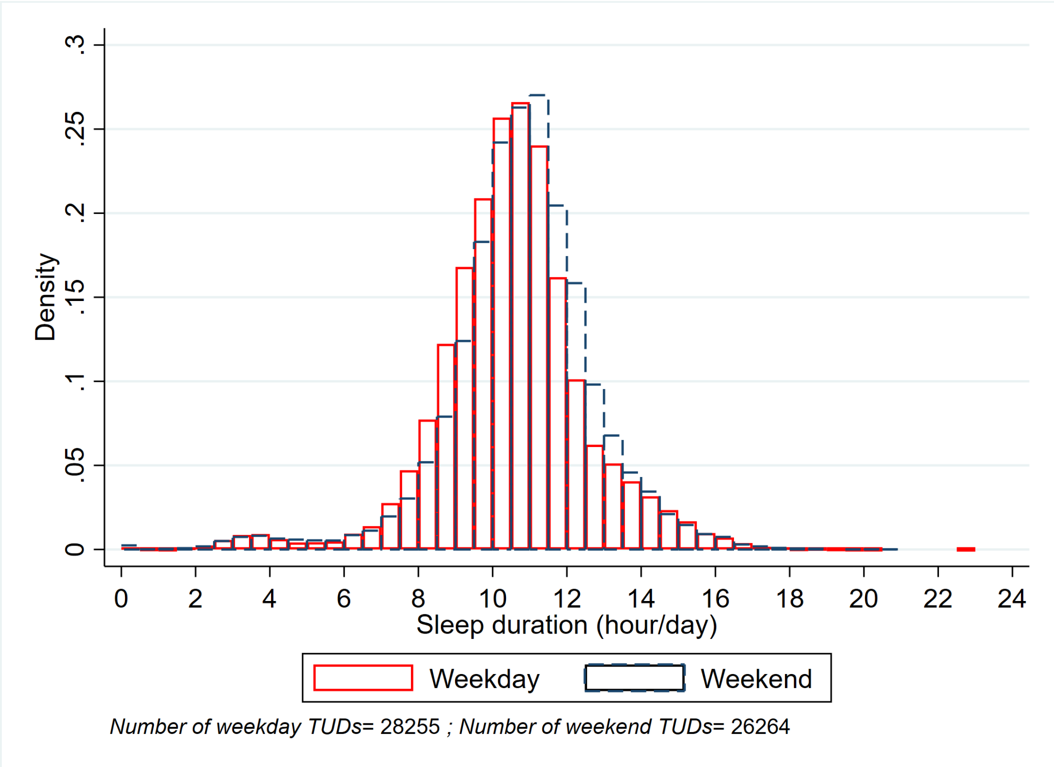
Notes: Results are from models (1) and (2), unless stated otherwise. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration, unless stated otherwise. Unless stated otherwise, other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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: Impact of sleep duration on child development – Robustness checks (continued)

	Waist-for-height ratio	Excellent health	Any ongoing condition	Prescribed medicine	MBS ($1000)	PBS ($1000)	MBS and PBS ($1000)	MR	Reading	Writing	Spelling	Grammar	Numeracy
	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)
Panel F: Reduced form													
Daily daylight duration (hour)	-0.06*	-0.08	-0.09	0.24	-0.87***	0.48	-0.39	-2.60	-0.69	-0.01	-1.19**	-0.37	0.26
	[0.03]	[0.40]	[0.43]	[0.27]	[0.34]	[0.62]	[0.70]	[1.70]	[0.67]	[0.98]	[0.51]	[0.77]	[0.58]
Observations	46,495	53,691	41,362	53,686	53,000	53,001	53,000	14,384	18,854	18,849	18,881	18,876	18,742
Individuals	8,311	8,699	8,109	8,699	8,546	8,546	8,546	3,519	5,503	5,506	5,510	5,509	5,472

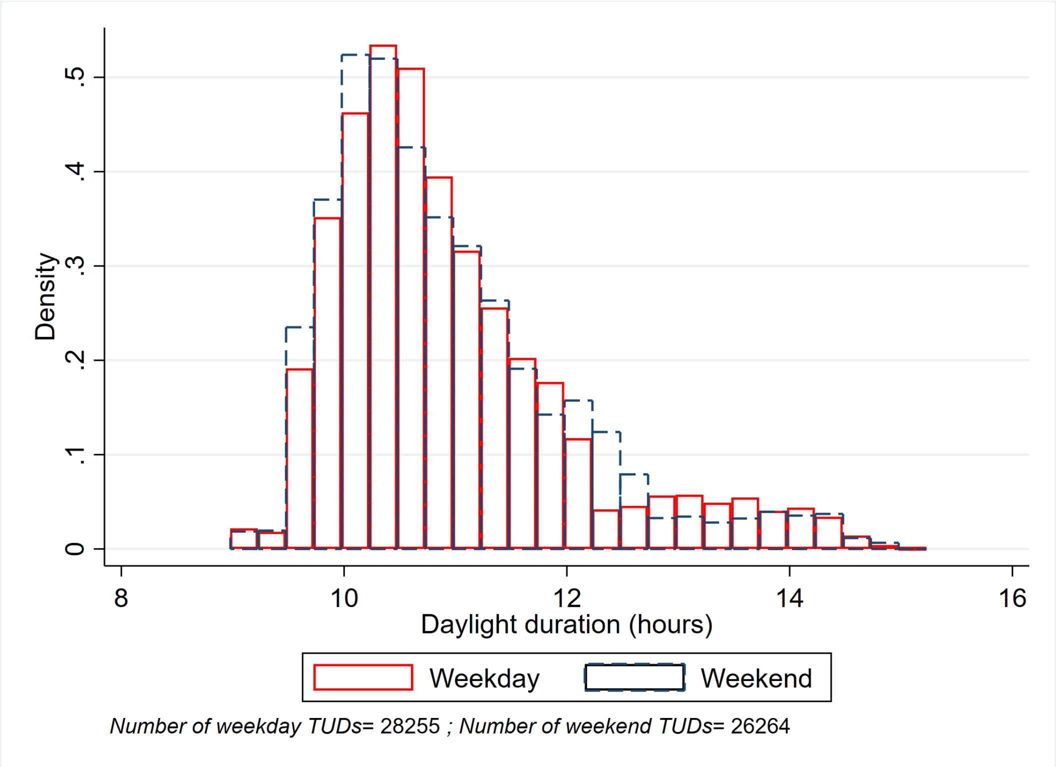
Notes: Results are from models (1) and (2), unless stated otherwise. F-statistic of IV denotes the F statistic for the excluded instrument in the first stage regression. Hausman test denotes p value from a Hausman test for endogeneity of the sleep duration variable in equation (2). Instrument: Daylight duration, unless stated otherwise. Unless stated otherwise, other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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: Distributions of sleep duration by weekdays/weekends



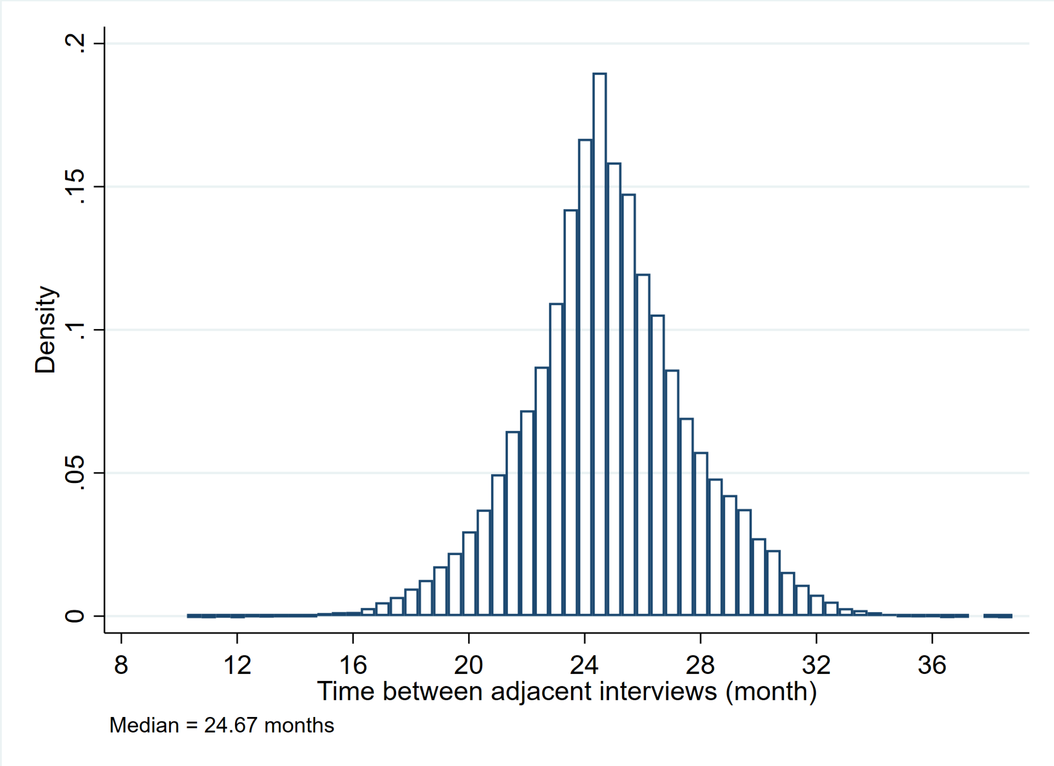
Notes: This figure reports sleep duration distribution for a pooled sample of all valid TUDs. Weekends include holidays.

Appendix Figure A2: Distributions of daylight duration recorded on TUD dates



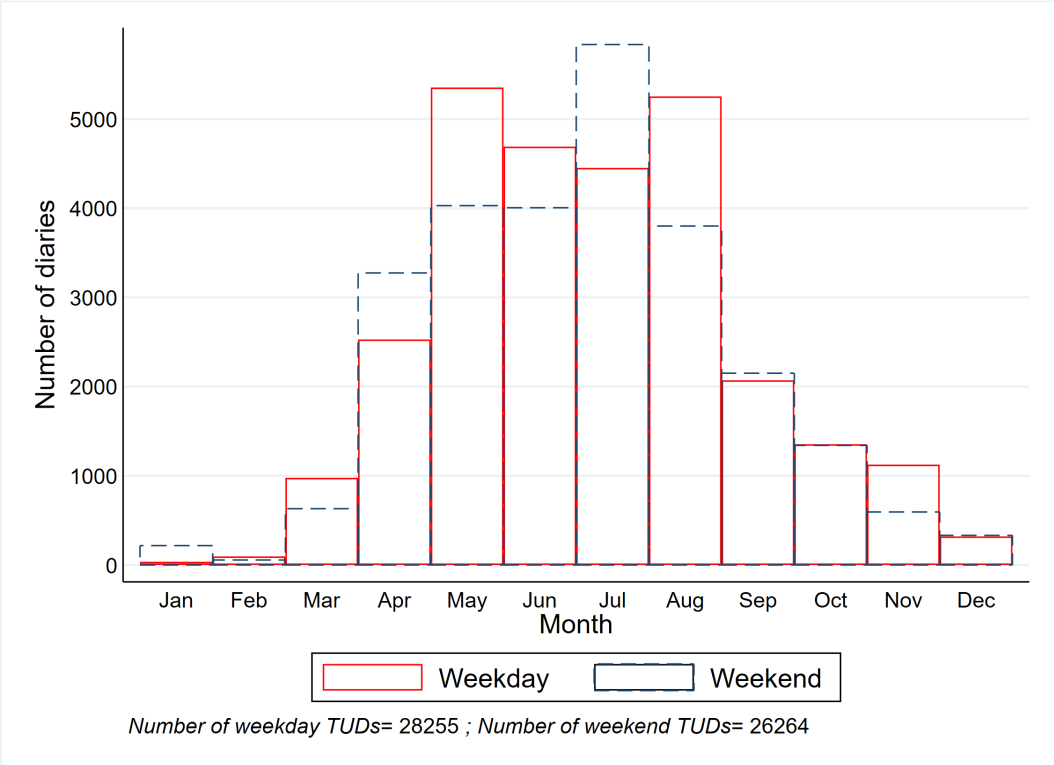
Notes: This figure reports daylight duration for a pooled sample of all valid TUDs. Weekends include holidays.

Appendix Figure A3: Distribution of time between adjacent interviews



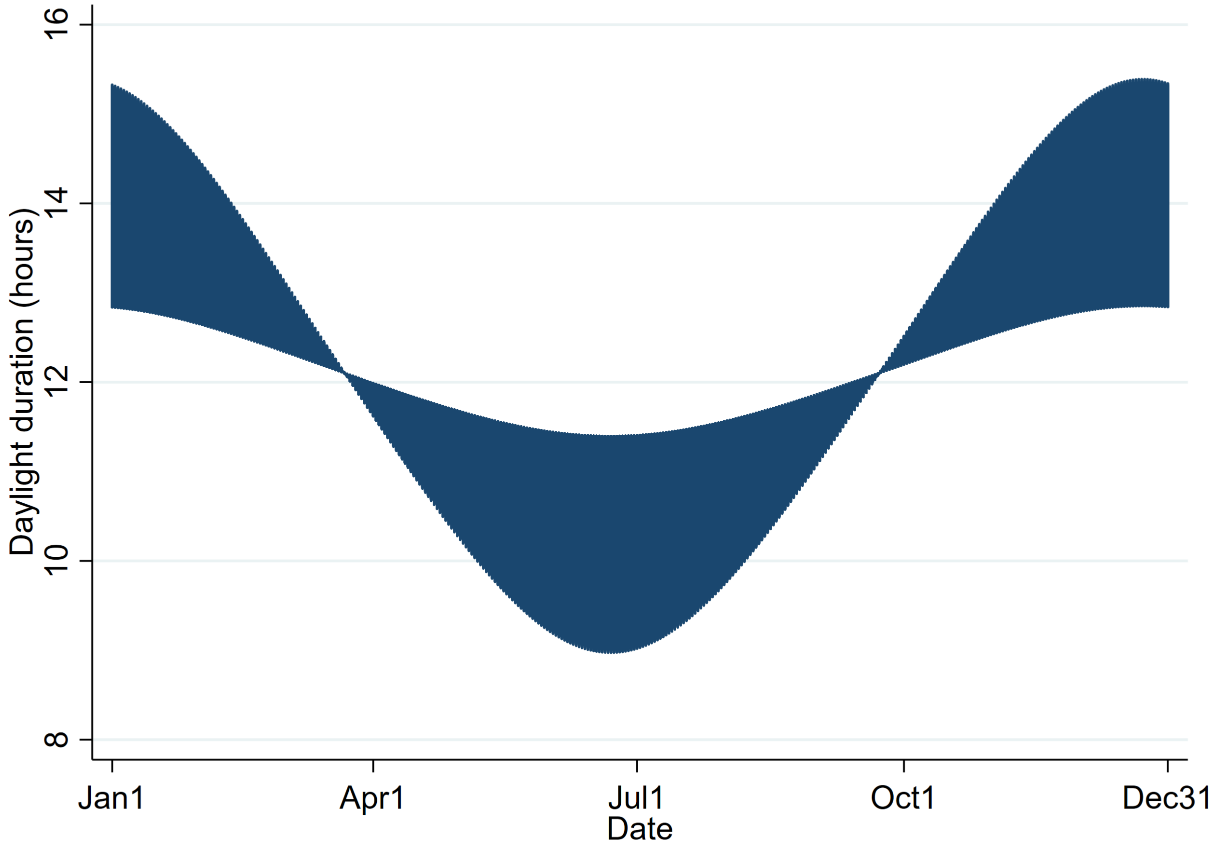
Notes: This figure reports distribution of time (in months) between two adjacent interviews for a pooled sample of all valid TUDs.

Appendix Figure A4: Distribution of time use diary months by weekdays/weekends



Notes: This figure reports the distribution of diary months for a pooled sample of all valid TUDs. Weekends include holidays.

Appendix Figure A5: Variation in daylight duration in LSAC TUDs



Notes: Each line in this figure shows daylight duration over a non-leap year for a postcode (among about 312 postcodes) sampled in LSAC TUDs.

Appendix B: Non-linear impact of sleep duration

Medical literature often documents a non-monotonic association between sleep duration and mortality (Cappuccio *et al.* 2010; Svensson *et al.* 2021) or BMI (Cappuccio *et al.* 2008). While our empirical model is not ideal to explore the potentially non-linear causal effect of sleep duration on child development,[1] in this section, we attempt to explore this possibility in three ways. First, we introduce the endogenous sleep duration variable in a quadratic form in Equation (1) and apply a FE regression method to estimate this modified model. As discussed earlier, a causal interpretation of the results obtained from this modified FE model requires a rather strong assumption that all individual time-variant unobserved characteristics are not simultaneously associated with sleep duration and child development. While this assumption cannot be formally tested in this case, the test results from the baseline FE-IV regressions provide some support for this approach because we found little evidence against this assumption for most outcomes (i.e., the p value of the Hausman test for exogeneity is greater than 0.1 in 18 out of 26 outcomes).

FE results, reported in Appendix Table B1 suggest no evidence of a non-linear relationship between sleep duration and most of the child development outcomes considered because estimates of the quadratic term of sleep duration are not statistically significant in almost all cases. There are two exceptions. First, the marginally statistically significant (at 10% level) and positive estimate of sleep duration variable and the statistically significant (at 5% level) and negative estimate of its quadratic term on Physical development suggests an inverted U-shaped relationship between sleep duration and Physical development. Numerically, the results suggest that children's physical

[1] One popular method to explore this possibility is to include sleep duration in a quadratic form in Equation (1). However, we cannot apply an IV approach to this modified model because of a lack of appropriate instruments to identify it. Specifically, to employ an IV approach to this modified model, we need at least two instruments, one for each of two potentially endogenous variables (i.e., sleep duration and its square). Theoretically, as suggested by Wooldridge (2010), this modified model can be identified by including the instrument (i.e., daylight duration in this case) in a quadratic form. This approach, however, does not work in practice because estimates of daylight duration and its square are not statistically significant in the first-stage regression. Probably due to the same unresolved identification issue, previous IV studies have not succeeded in drawing a non-linear causal impact of sleep duration either (Giuntella *et al*. 2017; Gibson & Shrader 2018).

development first increases with sleep duration, before starting to fall after 8 hours per day. By contrast, the statistically significant (at 1% level) and negative estimate of sleep duration and the statistically significant (at 1% level) estimate of its squared term on MBS expenditures indicate a U-shaped association between sleep duration and MBS expenditures. Specifically, children's MBS expenditures arrive at their minimum value when sleep duration reaches 11.5 hours per day, before increasing afterwards.

Second, to further explore the potential non-linear impact of sleep duration in a more flexible way, we categorize the daily sleep duration variable in the FE regression model. Specifically, we set the 10-11 sleep hour band, which includes the median of 10.5 daily sleep hours of all children in our sample, as the base, resulting in all other sleep duration band estimates being compared to the estimate of this sleep duration band.

The results, reported in Appendix Figure B1, suggest a non-linear relationship between sleep duration and selected outcomes.[2] For instance, the negative and statistically significant (at least at 5% level) estimates of the lowest sleep duration band (i.e., <8 hours) on Emotional development, PedsQL Overall, SDQ Emotional, SDQ Overall, and Excellent health indicator show that as compared with individuals sleeping from 10 up to 11 hours per day, those sleeping less than 8 hours daily have worse developmental outcomes in these domains. The statistically significant but opposite estimates of the highest sleep hour band on Emotional development and SDQ Peer indicate that as compared to individuals with 10-11 sleep hours per day, those sleeping 14 hours or more each day have a better outcome in Emotional development but worse in Peer. Moreover, the statistically significant and positive (negative) estimate of the 8-9 (<8) sleep hour band on BMI (Underweight) indicates weight gain associated with sleeping longer is mainly observed for individuals with these low sleep hours. Furthermore, the positive and statistically significant

[2] As discussed above, FE results in this exercise may not be interpreted as causal. Furthermore, results for some outcomes or sleep hour bands are statistically under-powered, possibly because of the small sample sizes.

estimates of the two top sleep hour bands on Grammar suggest that individuals who sleep 13 hours or more per day have greater grammar scores than those with a shorter sleep duration. Thus, the results from this exercise tend to indicate that the previously identified effects of sleep duration on these selected outcomes might have been driven by individuals at the two tails of the sleep duration distribution.

Third, motivated by sleep deprivation literature (Cappuccio *et al.* 2010), we dichotomize the sleep duration variable, using various cut-off points with a 30-minute increment, and use each of these newly created dummy variables in place of the continuous sleep duration variable in the baseline FE-IV model. We still use daily daylight duration as the sole instrument in this modified FE-IV model. Because the instrument is only sufficiently statistically significantly (i.e., F statistic from the first stage regression >10) associated with sleep binary variables identified between a range from 10 to 12 hours, we apply this modified model to these selected sleep duration cut-offs. Comparing the estimates for individuals with different sleep duration cut-offs, e.g., individuals who sleep at least 11 hours per day and those who sleep at least 10.5 hours per day, may reveal evidence for whether sleep has a non-linear impact on child development.

Unreported results from this experiment show little evidence of non-linearity in the impact of sleep duration on almost all development outcomes considered because estimates of sleep duration cut-off variables are not statistically significant at any conventional level. Exceptions are noted and reported in Appendix Figure B2 for three outcomes: BMI, overweight and MBS expenditures. Specifically, estimates of sleep duration cut-offs on these outcomes are positive and statistically significant (at least at 10% level) over the whole cut-off points considered. Visually, the relationship between sleep duration cut-offs and each of these three outcomes follows a U-shaped pattern and lowest estimates are observed at the cut-off of 10.5 hours per day. The finding that weight gain and hence the risk of being overweight are higher for individuals at the two ends of the sleep duration spectrum is consistent with an oft-observed pattern of an increased risk of obesity amongst short sleepers in children (Cappuccio *et al.* 2008).

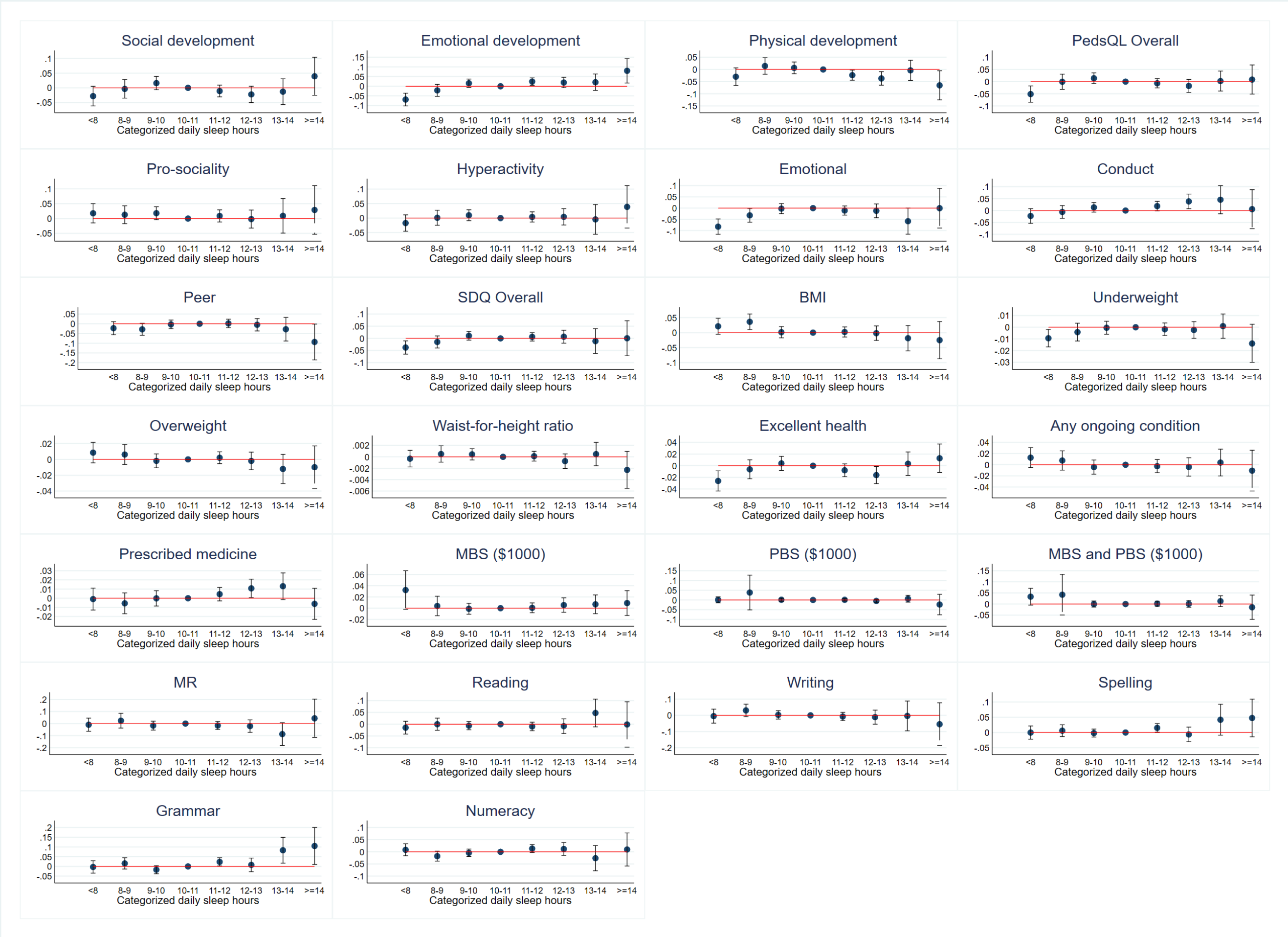
Overall, the results from this sub-section show some evidence of a non-linear relationship between sleep duration and selected general development, behavioural and health-related outcomes. However, the results indicate little evidence of such a non-linear relationship for almost all cognitive outcomes. To this end, our finding of a linear relationship between sleep duration and selected cognitive skills is in line with that in an experimental study by Lo *et al.* (2016) who find cognitive performance of adolescents is nearly-linearly correlated with accumulated duration of sleepiness over time.

Appendix Table B1: Non-linear impact of sleep duration – FE results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Social development	Emotional development	Physical development	PedsQL Overall	Pro-sociality	Hyperactivity	Emotional symptoms	Conduct	Peer problem
Sleep duration (hour/day)	1.46	0.58	1.84*	1.78*	-0.03	0.07	2.53**	-0.53	1.30
	[0.97]	[0.96]	[0.97]	[0.92]	[1.06]	[0.93]	[1.14]	[0.98]	[1.07]
Sleep duration squared	-0.07	0.04	-0.12**	-0.07	-0.01	0.01	-0.10	0.07	-0.06
	[0.05]	[0.05]	[0.05]	[0.05]	[0.06]	[0.05]	[0.06]	[0.05]	[0.06]
Observations	45,141	46,145	45,135	43,542	40,425	40,418	40,422	40,423	40,425
Individuals	8,223	8,265	8,211	8,115	7,963	7,961	7,962	7,963	7,963
	SDQ Overall	BMI	Underweight	Overweight	Waist-for-height ratio	Excellent health	Any ongoing condition	Prescribed medicine	MBS ($1000)
Sleep duration (hour/day)	0.94	0.13	0.22	-0.11	0.04	0.56	-0.81	-0.08	-1.39***
	[0.88]	[0.85]	[0.22]	[0.39]	[0.04]	[0.47]	[0.55]	[0.35]	[0.37]
Sleep duration squared	-0.02	-0.03	-0.01	-0.00	-0.00	-0.01	0.03	0.01	0.06***
	[0.05]	[0.05]	[0.01]	[0.02]	[0.00]	[0.02]	[0.03]	[0.02]	[0.02]
Observations	40,411	46,605	46,643	46,643	46,501	53,699	41,368	53,694	53,008
Individuals	7,960	8,322	8,325	8,325	8,312	8,700	8,110	8,700	8,547
	PBS ($1000)	MBS and PBS ($1000)	MR	Reading	Writing	Spelling	Grammar	Numeracy	
Sleep duration (hour/day)	0.68	-0.71	-0.54	0.75	-0.20	-0.42	-1.66	-0.39	
	[0.71]	[0.76]	[1.49]	[0.90]	[1.14]	[0.65]	[1.06]	[0.76]	
Sleep duration squared	-0.05	0.01	0.02	-0.03	-0.01	0.04	0.12**	0.03	
	[0.06]	[0.06]	[0.09]	[0.05]	[0.07]	[0.04]	[0.06]	[0.04]	
Observations	53,009	53,008	14,384	18,854	18,849	18,881	18,876	18,742	
Individuals	8,547	8,547	3,519	5,503	5,506	5,510	5,509	5,472	

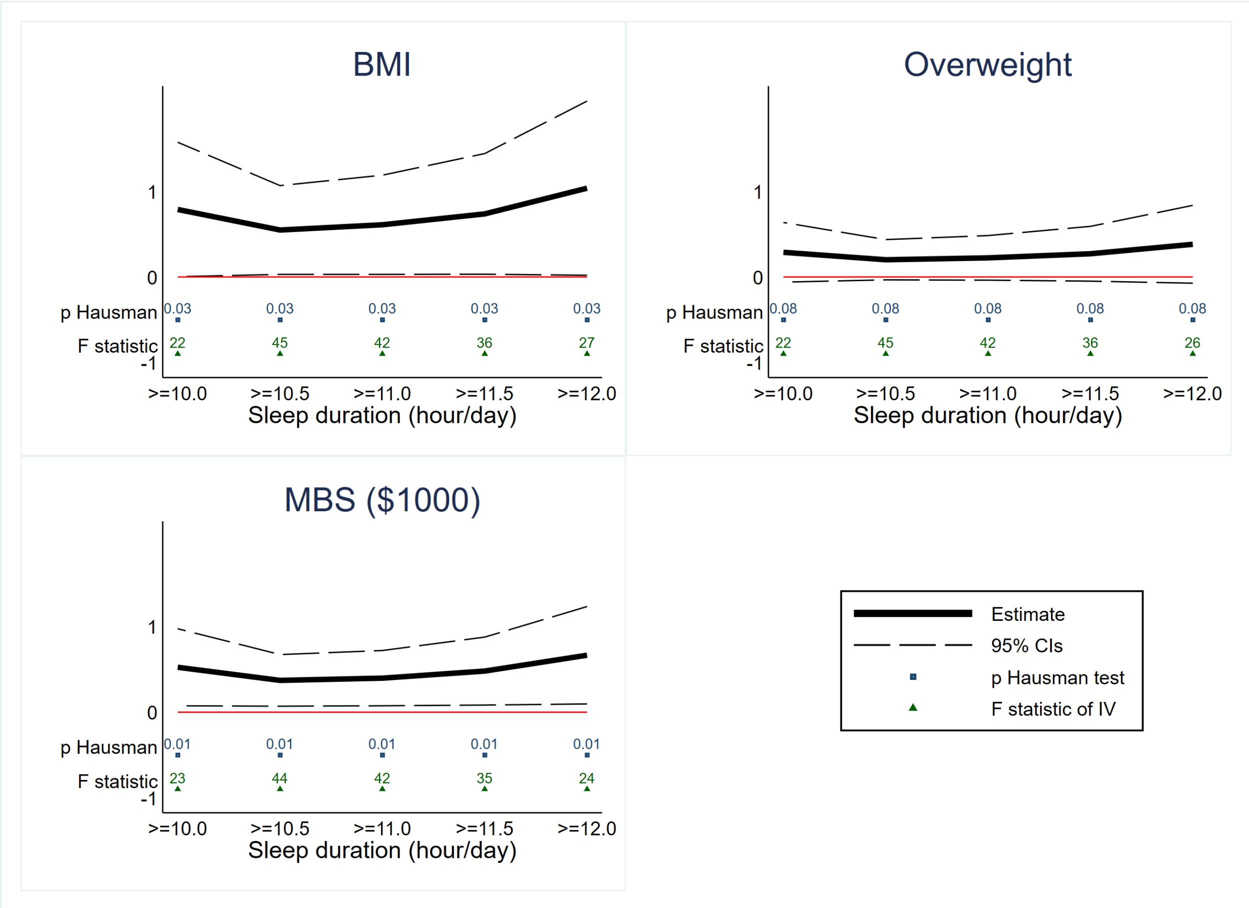
Notes: Results are from FE regression (1). Other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors clustered at the individual level are in squared brackets. 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 B1: Impact of sleep duration using categorized sleep hours



Notes: Results (in marginal effects) for each outcome are from a separate FE regression. Sleep duration is categorized with daily sleep duration between 10 and 11 hours is set as the base group. Other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors are clustered at the individual level.

Appendix Figure B2: Impact of sleep duration at different cut-offs



Notes: Results for each cut-off points are from a separate FE-IV regression. "F-statistic of IV" denotes the F statistic for the excluded instrument in the first stage regression. "P Hausman test" denotes p value from a Hausman test for endogeneity of the sleep duration cut-off variable in equation (2). Instrument: Daylight duration. Other explanatory variables include child age (and its square), maternal completed qualification, living with both parents, number of siblings; local socio-economic background variables, state/territory dummies, TUD wave dummies, TUD quarter dummies, TUD day-of-week dummies, and a holiday indicator. Robust standard errors are clustered at the individual level.