Causal Effect of Sleep Duration on Body Weight in Adolescents

A Population-based Study Using a Natural Experiment

Young Kyung Do[a,b]

Background: A large number of observational epidemiologic studies have reported consistent associations between short sleep duration and increased body weight, particularly in children and adolescents. Causal evidence on the effect of sleep duration on body weight is still limited, however.

Methods: This study exploits a unique natural experiment that can be argued to have increased sleep duration in an adolescent population in South Korea. In March 2011, authorities in three of the 16 administrative regions decreed restricting the closing hours of *hagwon* (private tutoring institutes) to 10 PM. Assuming this policy change is a valid instrument for sleep duration, it allows investigation of the causal effect of sleep duration on body weight in a difference-in-differences and instrumental variable framework. We used a nationally representative sample of 191,799 in-school adolescents in 7th–12th grades surveyed in the 2009−2012 Korea Youth Risk Behavior Web-based Survey.

Results: The policy change was associated with sleep extension and body weight reduction in a subset of general high school 10th−11th graders (around 10% of the sample) whose sleep duration would otherwise have not increased. The main results suggested that a 1-hour increase in sleep duration was associated with a $0.56\,\mathrm{kg/m^2}$ reduction in body mass index (95% confidence interval = 0.07, 1.05) and a decreased risk of being overweight or obese by 4.2% points.

Conclusions: This study provides new population-level, causal evidence that corroborates consistent findings in the epidemiologic literature on the link between short sleep duration and increased body weight.

Keywords: Body weight; Causal inference; Instrumental variable; Natural experiment; Overweight and obesity; Sleep

A large number of observational epidemiologic studies have reported consistent associations between short sleep duration and increased body weight, particularly in children and adolescents.[1–12] Based on the robust associations in epidemiologic studies in combination with research findings on potential biologic and behavioral mechanisms in experimental studies,[11,13–22] promoting adequate sleep hours has been proposed for obesity prevention and weight reduction.[7,23–25] Causal evidence on the effect of sleep duration on body weight is still limited, however, posing a barrier to the potential translation of such research findings into effective public health interventions.[5,6,26–28]

Observational studies are prone to confounding and reverse causality in making causal inferences about the effect of sleep duration on body weight.[1,5,6,26–28] Both sleep duration and body weight may be correlated with a host of unmeasured factors.[4,29,30] For example, students who are strongly future-oriented may voluntarily choose to reduce sleep hours to increase their study time for better academic performance and future life chances. These same individuals are also more likely to watch their food intake and engage in physical activity and are therefore less likely to be obese.[31,32] Students who are less future-oriented may also choose to shorten sleep duration, albeit not for study, but for Internet games; these same students are also more likely to be obese.[32] Future orientation and other unobservable factors may be unmeasured confounders when estimating the effect of sleep duration on body weight in adolescents. Inadequate sleep can also be a symptom of weight problems, leading to reverse causation.[12,28,33,34] Previous

studies have attempted to address these issues with limited success.[6,28,34] An ideal study design would be a population-based, randomized prospective experiment,[5–7,12,21,35] in which a large number of participants are randomly assigned into two or more groups according to different manipulated sleep duration and subsequently followed up in free-living conditions for a sufficiently long period of time. It is hard to conduct such studies, for both ethical and practical reasons.[26,35,36]

In the absence of well-conducted population-based intervention studies, one alternative approach is to use a natural experiment.[37–40] This study exploits a unique natural experiment that can be argued to have increased sleep duration in an adolescent population in South Korea. In March 2011, amid growing concerns over the negative consequences of late-night classes at private tutoring institutes or cram schools (*hagwon*), authorities in three of the 16 administrative regions decreed restricting the closing hours of *hagwon* to 10 PM. Assuming this policy change is a valid instrument for sleep duration, it allows for investigating the causal effect of sleep duration on body weight in a difference-in-differences and instrumental variable (IV) framework. In doing so, this article aims to make the case that the policy change meets the instrumental conditions, with subject-matter knowledge and empirical evidence, though not all the conditions are verifiable.[41]

METHODS

Setting and the Natural Experiment

Secondary education in South Korea involves a structured progression from middle school (7th–9th grades) to either a general high school or a vocational high school (10th–12th grades). Just before completing the 12th grade (November), most general high-school students take the nationwide standardized College Scholastic Ability Test (CSAT). Challenged by fierce competition for better college and future life opportunities, approximately 75% of secondary-school students take extra after-school classes at *hagwon*, which operate until late at night, often even past mid-night. The negative impact of these practices on students prompted policymakers to propose a curfew on *hagwon* operating hours, to ensure the protection of adolescents' rights to health and happiness, including having enough time for leisure and sleep. After debates in the political sphere, an agreement was reached about the curfew, which would allow the education authority in each administrative region to autonomously set *hagwon* operating hours. In 2010, the central government proposed a 10 PM curfew in an effort to protect adolescents' right to sleep in particular.[42] On 1 March 2011, three regions implemented the 10 PM curfew for high-school students (two other regions adopted an 11 PM curfew). Seoul had already adopted the 10 PM rule for high-school students.

This policy change has not been part of efforts to address adolescents' weight problems. Although promoting adolescent health and sleep was included as one of the policy goals, weight problems were not specifically considered in the policy debates. Rather, the policy was primarily a product of complex regional politics involving the interests of the private tutoring industry advocating the freedom for business on the one hand, and the interests of civic groups advocating adolescents' right to sleep and health on the other hand. Although a stronger worry could have existed about students' sleep deprivation in these regions with the curfew change, there were no known other relevant regional policies concomitantly implemented, that might have affected adolescent energy balance (such as policies aimed at promoting physical activity and reducing fast-food intake).

The effect of this policy change on adolescents' sleep duration is likely to have varied by school type and grade. Many general high-school students in the 10th and 11th grades use *hagwon* to prepare for the CSAT and regular school exams, and tend to reduce their sleep hours to have more time for taking extra courses at *hagwon*. However, these students still have at least 1 year left before the CSAT, which is taken toward the end of 12th grade; hence, general high school 10th–11th graders may not be as desperate as 12th graders to sacrifice sleep hours for more study time. Therefore, many 10th–11th graders are likely to have slept more in response to the policy change.[41,43,44] This group is used as the main sample in the current study.

In contrast, three other groups of secondary-school students are likely to have been largely unaffected by the policy change. General high school 12th graders preparing for the upcoming CSAT may still want to spend a good amount of time on self-study even after returning home from *hagwon*. Vocational high-school and middle-school students are also unlikely to respond to the policy change, the reasons for which, however, are different from the explanation for the sample of general high-school 12th graders. Vocational students typically do not take late-night classes at *hagwon*. Middle-school students have at least 3 more years before taking the CSAT and therefore are far less likely to take late-night classes, even without the 10 PM rule. These three samples will be called hereafter "placebo" samples. For the reasons aforementioned, the policy change must be irrelevant to changes in sleep duration and body weight of these groups. By showing noneffect where no effect should be found, these placebo groups help remove possible alternative explanations for the causal inference being made.[45]

Statistical Analysis

This study implements the natural experiment design in a difference-in-differences and instrumental variable (IV) framework. As an initial step, we performed difference-in-differences analysis to examine whether the policy change affected average sleep duration and body weight, respectively, after controlling for common trends and baseline differences between the treatment regions and control regions. The validity of the difference-in-differences estimation critically

hinges on the assumption that, without the policy change, the treatment region would have experienced the same trend in sleep duration as the control region did.[46] Although this common trends assumption cannot be tested directly in the main sample, the three placebo samples allow for conducting a useful test for the assumption.[45,46]

Each of the difference-in-differences analyses can provide information on whether the policy change had effects on sleep duration and on body weight but does not answer the research question of the study: whether sleep duration affects body weight. The ratio of the two difference-in-differences estimates (i.e., for sleep duration and for body weight in each sample), the Wald/IV estimator,[46] can be thought of an intuitive approximation of the effect to be estimated in the study. This Wald/IV estimator captures the shift in average body weight per 1-hour increase in sleep duration induced by the policy change. The following IV estimation is, in essence, to extend this intuition behind the Wald/IV estimator to the IV regression framework with a full set of individual-level covariates.

Next, we conducted IV two-stage least squares analysis. The goal of this IV analysis is to estimate the causal effect of sleep duration (exposure) on body weight (outcome), exploiting the variation in sleep duration due to the policy change (IV). To be able to make causal inference, the IV estimation requires three important conditions:[41] (1) the IV is associated with the exposure variable, (2) the IV does not affect the outcome other than through the exposure variable, and (3) the IV does not share any causes with the outcome. Each of the three conditions is described more fully in the following paragraphs.

First, the policy change is associated with change in sleep duration. Had the policy change been associated with no or negligible increase in sleep duration, the IV estimation would fail. This condition can be tested using the partial *F* statistic in the first-stage regression (results can be found in eTable 1; http://links.lww.com/EDE/B577).

Second, the policy change should affect body weight only through sleep duration, which is the exclusion restriction condition for IV estimation.[41,46] If the policy change had affected, for example, eating and exercise behaviors directly (i.e., except through changes in sleep duration), our IV strategy would be problematic. Note, however, that it is not a concern here that changes in eating and exercise behaviors are due to changes in sleep duration, as the interest lies in the total effect of sleep duration, not the direct effect (i.e., after controlling for all downstream covariates). Stress could be another factor that directly mediates the effect of the policy change on body weight: Banning late night after-school classes itself could have lowered students' stress levels in the treatment region and, consequently, their average body weight could have decreased (with no relation to sleep duration). To probe the importance of stress in mediating the effect of the policy change on body weight, we conducted an additional analysis (results are available in eTable 2; http://links.lww.com/EDE/B577). Many more possibilities are conceivable. Within the

bounded time of 24 hours, any change in sleep duration must involve changes in time use on other activities, which might have consequences for body weight. For these reasons, the exclusion restriction condition is hard to justify conceptually. Unfortunately, this condition cannot be tested empirically.

Third, the policy change should not share any causes with body weight.[41,47] In the current study, this condition can be considered at two levels. At the regional level, there should be no relevant concomitant changes between the treatment and control regions, that affect body weight (e.g., coincidental nutritional or physical activity campaigns). The institutional background of the policy change aforementioned provides a justification based on subject-matter knowledge. Although this condition cannot be tested empirically for the main sample, the difference-in-differences analysis of the three placebo samples offers an opportunity to examine the validity of the condition: Had there been such coincidental changes, body weight in these placebo samples would show differential trends between the treatment and control regions because such "environmental" factors would have affected indiscriminately all the samples. Ultimately, the condition requires that at the individual level, the policy change be uncorrelated with unmeasured covariates in the IV regression model. This requirement cannot be tested directly, although previous studies have compared differences in measured covariates by the proposed IV versus by the treatment, as a way to explore potential confounding by unmeasured covariates and often to justify the IV strategy.[47] This article follows the method proposed in Jackson and Swanson[47] and presents the results (results are available in eFigure 3; http://links.lww.com/EDE/B577).

An additional requirement for IV estimation is the monotonicity assumption,[48,49] which means that, in this study, the policy change on *hagwon* curfew never decreases sleep duration. Several scenarios for violation of this assumption are presented (eContent 3-3; http://links.lww.com/EDE/B577), with empirical evidence in favor of the assumption (eFigure 4; http://links.lww.com/EDE/B577). This consideration also helps clarify what causal effect is being estimated among which subgroup of the study population. The causal effect estimated in this study is a weighted average per-unit treatment effect among those whose sleep duration increased because of the policy change.[49]

For regression analysis, we estimated three IV two-stage least squares models using BMI and two other weight outcomes. Next, we estimated four models of BMI, using each of four alternative sleep variables as the main explanatory variable. Finally, we performed a number of sensitivity analyses to examine how robust the main IV two-stage least squares results are to alternative sample definitions and the IV used. All analyses account for clustering of adolescents within clusters in the data.

Data

This study used data from the Korea Youth Risk Behavior Web-based Survey (KYRBS), an annually repeated

cross-sectional, school-based, online survey of a large, nationally representative sample of middle and high-school students (age 13–18 years) in South Korea. The main objective of this survey, administered in September and October by the Korea Centers for Disease Control and Prevention, is to monitor health and behavioral risk factors of South Korean adolescents. The KYRBS employs two-stage cluster sampling: The first stage consists of 400 middle schools and 400 high schools randomly selected within 131 districts, and the second sampling stage selects one class from each grade within each chosen school. All students in the selected classes, except for school dropouts, students with special needs, and those who have difficulty in reading comprehension, are surveyed. The KYRBS respondents complete an online self-administered questionnaire in a computer room in their school. The participation rates were above 95% between 2009 and 2012.

The current study used data from four annual surveys of KYRBS (2009−2012), which spanned two surveys for each of the pre- and postperiods (i.e., before and after the curfew change in March 2011) and used nearly identical survey questions for the study variables. All administrative regions were included initially. We omitted observations from two regions that had not had a *hagwon* curfew before but introduced an 11 PM curfew, to ensure the homogeneity of the natural experiment and sufficiently large variations in sleep duration induced by the policy change. After we excluded observations with missing values for the study variables included in the full IV regression model, the final main sample consisted of 52,585 general high-school 10th−11th graders. Three "placebo" samples included 25,892 general high-school 12th graders, 15,206 vocational high-school 10th−11th graders, and 98,116 middle-school students (a flow chart is available in eFigure 1; http://links.lww.com/EDE/B577).

Ethics approval to conduct this study was obtained from Seoul National University College of Medicine and Seoul National University Hospital Institutional Review Board (E-1501-048-639).

Variables

The main outcome variable, body mass index (BMI), was calculated with the use of self-reported body weight (in kilograms) and height (in centimeters), with both values reported up to one decimal place. In addition, a binary indicator variable of overweight/obesity was generated based on gender- and age-specific BMI cutoffs (≥85th percentile) in the 2007 Korean Growth Standards of Children and Adolescents.[50]

We defined sleep duration, the main explanatory variable of interest, as average weeknight sleep hours in the past week and calculated as the duration from bedtime to wake-up time, which the KYRBS asked about in terms of separate units of hours and minutes. In addition, we examined four alternative variables of sleep duration. Based on the distribution of sleep hours, we created two binary indicator variables using the 5- and 6-hour cut-off points. We also examined bedtime as

an alternative explanatory variable in this study, because late bedtime is found to be associated with weight problems.[51,52] We recoded the corresponding variable so that it ranges from 20 (8 PM) to 28 (4 AM). Finally, we examined perceived sleep adequacy, because this may be of interest for its direct behavioral implications. Despite its ordinal nature, this variable was treated as a continuous variable (with values of 1 = "very inadequate" to 5 = "very adequate").

We generated three indicator variables involving treatment and period status: Treatment for the three regions where the policy change in *hagwon* curfew occurred on 1 March 2011 (=1 vs. 0 for other regions); Post for the observations surveyed in 2011 and 2012 (=1 vs. 0 for 2009 and 2010); and the interaction term Treatment × Post (=1 if Treatment = 1 and Post = 1 vs. 0 otherwise). Other covariates included region and year dummies, gender, education levels of both parents, self-reported household economic status, and school grade. Table 1 presents how study variables were coded in this study.

RESULTS

Summary statistics of the study variables in the main sample of general high-school 10th−11th graders are presented in Table 1. The average BMI is 20.9. The proportion of overweight/obesity in the sample is 11%. The average sleep duration is 5.7 hours, and only 41% sleep 6 hours or more. The average bedtime is 00:54 AM. Observations from the treatment group account for 31% of the sample, while observations are nearly equally divided between pre- and postperiods. Observations are also balanced in terms of gender and school grade (Table 1).

Table 2 shows groups means of sleep hours and BMI by period and treatment status for each of the four samples. The (unadjusted) difference-in-differences estimate for the policy effect on sleep duration in the main sample of general high-school 10th−11th graders is 0.28 (A, column 3), indicating that the 10 PM curfew resulted in a relative increase in sleep duration by 0.28 hours, or approximately 17 minutes, on average (this result suggests that the policy change is strongly associated with sleep duration, satisfying the necessary condition of IV strength, which is supported by the large partial *F* statistic (33) in the first-stage regression of sleep hours,[53] eTable 1; http://links.lww.com/EDE/B577). The difference-in-differences estimate in column 6 suggests that the policy change decreased the average BMI by 0.11 (A, column 6). The ratio of these two difference-in-differences estimates (−0.39 in A, column 7) gives the Wald/IV estimator, an (unadjusted) IV estimate of the effect of a 1-hour increase in sleep duration on BMI. No similar pattern is observed for the three placebo samples (B−D). The overall estimates suggest that the policy change had effects on sleep duration and BMI in the main sample but virtually no effects in the three placebo samples (a more intuitive graphical representation is given in eFigure 2; http://links.lww.com/EDE/B577) (Table 2).

TABLE 1. Summary Statistics of Variables in the Sample of General High-school 10th–11th Graders

Variables	Mean
Weight-related outcomes	
Body mass index (kg/m^2)	20.9
Weight (kg)	58.8
Overweight/obesity (1 if yes, 0 otherwise)	0.11
Sleep variables	
Sleep hours (h)	5.7
Sleep hours 5 or more (1 if yes, 0 otherwise)	0.77
Sleep hours 6 or more (1 if yes, 0 otherwise)	0.41
Bedtime (range 20−28 [=8 PM−4 AM])	24.9
Perceived sleep adequacy (1−5: 5 = "very adequate")	2.7
Treatment and period variables	
Treatment (1 if curfew changed to 10 PM, 0 otherwise)	0.31
Post (1 if year 2011−2012, 0 if year 2009−2010)	0.53
Year dummies	
Y2009 (1 if 2009, 0 otherwise)	0.25
Y2012 (1 if 2012, 0 otherwise)	0.27
Instrumental variable	
Treatment × Post (1 if Treatment = 1 and Post = 1, 0 otherwise)	0.17
Other individual-level explanatory variables	
Female (1 if female, 0 if male)	0.49
11th grade (1 if yes, 0 if 10th grade)	0.51
Father's education level	
College or more	0.52
High school	0.37
Middle school	0.05
Do not know	0.06
Mother's education level	
College or more	0.38
High school	0.51
Middle school	0.05
Do not know	0.07
Household economic status	
High	0.05
Mid-high	0.23
Middle	0.50
Mid-low	0.18
Low	0.04

The sample comes from general high-school students in the 10th and 11th grades surveyed in the 2009−2012 Korea Youth Risk Behavior Web-based Survey (N = 52,585). Dummies for regions are not presented in this table.

The IV two-stage least squares estimates for the main sample suggest that a 1-hour increase in sleep duration decreases, on average, BMI by 0.56 kg/m^2, body weight by 1.63 kg, and the probability of being overweight or obese by 4.2 percentage points (Table 3), in a subset of general high-school 10th–11th graders, whose sleep duration would have increased because of the policy change. These estimates are larger in absolute magnitude than their respective ordinary least squares estimates from ordinary linear regressions (Table 3).

In the four models using the alternative sleep variables (Table 4), the IV two-stage least squares estimates are qualitatively consistent with the main IV estimate. For example, the adolescents who sleep 6 hours or more because of the policy change would have a BMI reduction of 1.19, compared with those who sleep less than 6 hours. Because two of the sleep variables are dichotomous, it is possible to estimate the size of the subgroup ("compliers") to which the IV estimates of this study pertain.[48] The proportion of 'compliers' was 12.6% for the 6-hour cut-off and 8.7% for the 5-hour cut-off (Supplemental Digital Content 3-3; http://links.lww.com/EDE/B577)). (Table 4)

Results from sensitivity analyses are close in magnitude to the main IV estimate of −0.56, albeit with varying magnitudes and a stronger effect in males (Table 5). Specification tests generally suggest that IV estimation is preferable to ordinary least squares in this study (eTables 3−5; http://links.lww.com/EDE/B577) (Table 5).

DISCUSSION

Using a natural experiment that affected adolescents with short sleep duration in South Korea, this study has provided new, causal evidence that sleep gain is associated with improvement of weight problems. The overall results are in line with previous epidemiologic studies. The stronger estimated effect found in males is also consistent with the gender difference reported in the literature.[2,10]

This contribution is meaningful in the context of the existing epidemiologic literature on the sleep–obesity link. There has been skepticism that further epidemiologic studies employing traditional research methods would be "superfluous" with no signs of improving causal inference in the current research settings.[6] In the meantime, recent years have seen in-laboratory experimental studies on the sleep–obesity link.[13,17–19,22,54] Although these studies collected detailed information on sleep patterns, they involve a small number of participants and are very expensive. The current article studied a large population group in a real-world setting, while enhancing the internal validity of causal inference using a plausible natural experiment and data spanning both pre- and postperiod. Furthermore, the natural experiment allowed for studying a population-wide sleep gain (away from the conventional investigations on individual sleep loss), thus providing more pertinent evidence for a potential effective public health intervention.

This study also sheds light on another key issue debated in the literature—whether the beneficial effect of increasing sleep duration on weight is large enough to justify promoting modification of sleep duration.[55] The relatively small magnitude of the association between sleep duration and body weight reported in previous epidemiologic studies has been challenged as having a negligible benefit to an individual obese patient.[27,28] The magnitude of the main IV estimate in the current study is a 0.56 BMI change per hour of sleep,

TABLE 2.   DD Estimates on Sleep Hours and BMI and Wald/IV Estimators

	Sleep Duration (Hours)			BMI (kg/m^2)			
	Change in *Hagwon* Closing Hours in Region			Change in *Hagwon* Closing Hours in Region			
Sample	No Change (Control)	10 PM Curfew (Treatment)	DD Estimate on Sleep	No Change (Control)	10 PM Curfew (Treatment)	DD Estimate on BMI	Wald/IV Estimator
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. GHS 10th−11th [N = 52,585]							
Pre (2009 and 2010)	5.71	5.52		20.91	20.83		
Post (2011 and 2012)	5.66	5.75		20.97	20.78		
			0.28			−0.11	−0.39 (−0.76, −0.02)
B. GHS 12th [N = 25,892]							
Pre (2009 and 2010)	5.24	5.14		21.29	21.15		
Post (2011 and 2012)	5.26	5.22		21.41	21.38		
			0.06			0.10	1.62 (−1.16, 4.40)
C. VHS 10th−11th [N = 15,206]							
Pre (2009 and 2010)	6.31	6.24		20.79	20.88		
Post (2011 and 2012)	6.30	6.21		20.99	21.05		
			−0.03			−0.03	1.26 (−7.93, 10.45)
D. MS 7th−9th [N = 98,116]							
Pre (2009 and 2010)	7.04	6.97		20.05	19.94		
Post (2011 and 2012)	7.13	7.08		20.13	20.01		
			0.02			−0.003	−0.18 (−5.19, 4.83)

Data come from the 2009−2012 Korea Youth Risk Behavior Web-based Survey. 95% confidence intervals are presented in parentheses.
DD indicates difference-in-differences; MS, middle school; GHS, general high school; OLS, ordinary least squares; VHS, vocational high school.

TABLE 3.   Effect of Sleep Duration (in Hours) on BMI and Other Outcomes

Outcome	OLS	IV-2SLS	First-stage Partial F	Interpretation: A 1-hour Increase in Sleep Duration Leads to a …
BMI (kg/m^2)	−0.10 (−0.13, −0.08)	−0.56 (−1.05, −0.07)	33	0.56 kg/m^2 reduction in BMI
Weight (kg)	−0.39 (−0.47, −0.31)	−1.63 (−3.10, −0.16)	33	1.6 kg reduction in body weight
Overweight/obesity (1 vs. 0)	−0.006 (−0.0008, −0.003)	−0.042 (−0.091, 0.007)	33	4.2 percentage-point reduction in probability of being overweight/obese

The main sample of general high-school students in the 10th and 11th grades (N = 52,585) is used. Coefficients and 95% confidence intervals (in parentheses) are presented. Confidence intervals are based on cluster-robust standard errors.
IV-2SLS indicates instrumental variable two-stage least squares; OLS, ordinary least squares.

TABLE 4.   Effect of Alternative Sleep Variable on Body Mass Index

Alternative Sleep Variable	OLS	IV-2SLS	First-stage Partial F
Sleep hours 5 or more (vs. <5)	−0.20 (−0.25, −0.14)	−1.80 (−3.45, −0.16)	30
Sleep hours 6 or more (vs. <6)	−0.17 (−0.23, −0.12)	−1.19 (−2.22, −0.15)	40
Bedtime[a] (20−28)	0.08 (0.05, 0.10)	0.69 (0.05, 1.33)	30
Perceived sleep adequacy[b] (1−5)	0.07 (0.05, 0.10)	−0.93 (−1.78, −0.07)	28

The main sample of general high-school students in the 10th and 11th grades (N = 52,585) is used. Coefficients and 95% confidence intervals (in parentheses) are presented. Confidence intervals are based on cluster-robust standard errors.
[a]Higher number denotes later bedtime, ranging from 20 (8 PM) to 28 (4 AM).
[b]Higher number denotes greater perceived adequacy, ranging from 1 ("very inadequate") to 5 ("very adequate").
IV-2SLS indicates instrumental variable two-stage least squares; OLS, ordinary least squares.

TABLE 5. Sensitivity Analysis: Effects of Sleep Duration (in Hours) on Body Mass Index

Sample Definition and IV	OLS	IV-2SLS	First-stage Partial *F*
Main study sample			
GHS 10th−11th graders in 2009−2012 (N = 52,585)	−0.10 (−0.13, −0.08)	−0.56 (−1.05, −0.07)	33
Alternative sample definition			
Excluding observations with sleep hours > 8 (N = 51,891)	−0.11 (−0.14, −0.09)	−0.61 (−1.13, −0.10)	34
Excluding observations with sleep hours > 7 (N = 47,816)	−0.11 (−0.14, −0.08)	−0.62 (−1.23, 0.00)	35
Excluding underweight students (N = 49,034)	−0.08 (−0.10, −0.06)	−0.49 (−0.94, −0.03)	33
Excluding nonmetropolitan control (N = 35,146)	−0.11 (−0.13, −0.08)	−0.54 (−1.01, −0.07)	40
Post period 2011 only (N = 38,461)	−0.11 (−0.14, −0.08)	−0.45 (−1.00, 0.10)	28
Post period 2012 only (N = 39,121)	−0.12 (−0.15, −0.09)	−0.67 (−1.29, −0.05)	21
Male students only (N = 26,835)	−0.13 (−0.16, −0.09)	−0.79 (−1.48, −0.10)	26
Female students only (N = 25,750)	−0.08 (−0.11, −0.05)	−0.33 (−0.92, 0.26)	19
Adding GHS 12th (N = 78,477)	−0.11 (−0.13, −0.09)	−0.27 (−0.89, 0.35)	20
Alternative IV and sample			
IV from DDD (adding VHS sample) (N = 67,791)	−0.10 (−0.12, −0.08)	−0.78 (−2.06, 0.50)	5.4

Coefficients and 95% confidence intervals (in parentheses) are presented. Confidence intervals are based on cluster-robust standard errors.

DDD indicates difference-in-differences-in-differences; GHS, general high school; IV-2SLS, instrumental variable two-stage least squares; OLS, ordinary least squares; VHS, vocational high school.

which would hardly seem large enough for an individual adolescent to consider increasing sleep duration for a sustained period of time. Young[55] highlights the importance of putting such estimates in the perspective of population health. In fact, the BMI reduction of 0.56 per hour of sleep translates to a 4.2% point decrease in overweight/obesity risks from 11.4% in this study. As shown in this particular case, sleep extension may confer a substantial reduction in weight problems at a population level but only a small reduction in BMI for each individual, a phenomenon known as the prevention paradox.[55,56] The population strategy of shifting the overall distribution of sleep duration to longer hours has greater relevance,[57] because short sleep duration is widespread and has pervasive negative health consequences.[58–60]

The results of this study should be interpreted with caution because of several limitations. First, as described in the "Methods" section in detail, the exclusion restriction condition for IV estimation is hard to satisfy conceptually. The policy change could have affected body weight via various other mechanisms except through increased sleep duration. To the extent that such other mechanisms are at work, the IV estimates of the effect of sleep extension on BMI reduction could have been biased. Given the results of IV versus ordinary least squares analysis, overestimation is a more relevant concern. This concern, however, is partially alleviated by contextual knowledge and some evidence. The policy change is unlikely to have substantially promoted physical activity among general high-school 10th−11th graders just because more time became available between 10 PM and mid-night. Eating more certainly could have occurred because of the policy change, but this possibility would have contributed to diminishing the estimated effect on BMI reduction. Analysis on stress, though improvised, showed that it may not be important as a possible

mechanism for the effect of the policy change on body weight in the current data. These arguments strengthen the case that the main effect of the policy change on BMI reduction is through sleep extension, rather than through other behavioral mechanisms independent of sleep, although the exclusion restriction still cannot be verified.

A second limitation is that the variable of sleep duration and all other variables were derived from self-reported responses, which are prone to nonignorable missingness and a variety of measurement errors. For a population-based survey, the data on sleep duration were based on relatively specific questions on sleep, which include several key elements to improve the accuracy of self-reports on sleep duration.[61] The issue of systematic measurement error for sleep duration is partially addressed by the IV method used in this study, where the variation in sleep duration is induced by the exogenous policy change. Third, the IV results should be interpreted in the specific population group and context where the natural experiment existed and cannot be generalized to other settings. In the current data, the proportion of "compliers," whose sleep duration crossed the cut-off of 5 or 6 hours because of the policy change, was around 10%. It should be noted that the IV estimates of this study are derived from this subset of the adolescent population. The apparent effect on BMI should also be interpreted as the effect of replacing late studying in *hagwon* by sleep in the specific social context of South Korea, which may not be the same effect as replacing other activities by sleep in other settings. Fourth, the study was not able to account for the duration of exposure under the reported sleep hours and any changes in sleep behavior. However, students in the sample are likely to maintain a relatively regular schedule largely shaped by school times and *hagwon* policies in South Korea. Finally, analysis of this study relies on the validity of

various assumptions for statistical modeling (e.g., functional forms of variables).

Despite these limitations, this study provides population-level, causal evidence that corroborates the consistent findings in the extant epidemiologic literature on the sleep–obesity link. The growing literature in sleep epidemiology will need to better address the causality issue among diverse population groups.[12,59,60,62]

REFERENCES

1. Cappuccio FP, Taggart FM, Kandala NB, et al. Meta-analysis of short sleep duration and obesity in children and adults. *Sleep*. 2008;31:619–626.
2. Chen X, Beydoun MA, Wang Y. Is sleep duration associated with childhood obesity? A systematic review and meta-analysis. *Obesity*. 2008;16:265–274.
3. Patel SR, Hu FB. Short sleep duration and weight gain: a systematic review. *Obesity (Silver Spring)*. 2008;16:643–653.
4. Magee L, Hale L. Longitudinal associations between sleep duration and subsequent weight gain: a systematic review. *Sleep Med Rev*. 2012;16:231–241.
5. Marshall NS, Glozier N, Grunstein RR. Is sleep duration related to obesity? A critical review of the epidemiological evidence. *Sleep Med Rev*. 2008;12:289–298.
6. Nielsen LS, Danielsen KV, Sørensen TI. Short sleep duration as a possible cause of obesity: critical analysis of the epidemiological evidence. *Obes Rev*. 2011;12:78–92.
7. Coughlin JW, Smith MT. Sleep, obesity, and weight loss in adults: is there a rationale for providing sleep interventions in the treatment of obesity? *Int Rev Psychiatry*. 2014;26:177–188.
8. Van Cauter E, Knutson KL. Sleep and the epidemic of obesity in children and adults. *Eur J Endocrinol*. 2008;159(suppl 1):S59–S66.
9. Watson NF, Buchwald D, Vitiello MV, et al. A twin study of sleep duration and body mass index. *J Clin Sleep Med*. 2010;6:11–17.
10. Knutson KL. Sex differences in the association between sleep and body mass index in adolescents. *J Pediatr*. 2005;147:830–834.
11. Knutson KL. Does inadequate sleep play a role in vulnerability to obesity? *Am J Hum Biol*. 2012;24:361–371.
12. Ogilvie RP, Patel SR. The epidemiology of sleep and obesity. *Sleep Health*. 2017;3:383–388.
13. Greer SM, Goldstein AN, Walker MP. The impact of sleep deprivation on food desire in the human brain. *Nat Commun*. 2013;4:2259.
14. Haghighatdoost F, Karimi G, Esmaillzadeh A, et al. Sleep deprivation is associated with lower diet quality indices and higher rate of general and central obesity among young female students in Iran. *Nutrition*. 2012;28:1146–1150.
15. Spiegel K, Tasali E, Penev P, et al. Brief communication: sleep curtailment in healthy young men is associated with decreased leptin levels, elevated ghrelin levels, and increased hunger and appetite. *Ann Intern Med*. 2004;141:846–850.
16. Taheri S, Lin L, Austin D, et al. Short sleep duration is associated with reduced leptin, elevated ghrelin, and increased body mass index. *PLoS Med*. 2004;1:e62.
17. Schmid SM, Hallschmid M, Jauch-Chara K, et al. Short-term sleep loss decreases physical activity under free-living conditions but does not increase food intake under time-deprived laboratory conditions in healthy men. *Am J Clin Nutr*. 2009;90:1476–1482.
18. Markwald RR, Melanson EL, Smith MR, et al. Impact of insufficient sleep on total daily energy expenditure, food intake, and weight gain. *Proc Nati Acad Sci USA*. 2013;110:5695–5700.
19. Hart CN, Carskadon MA, Considine RV, et al. Changes in children's sleep duration on food intake, weight, and leptin. *Pediatrics*. 2013;132:e1473–e1480.
20. Prats-Puig A, Grau-Cabrera P, Riera-Pérez E, et al. Variations in the obesity genes FTO, TMEM18 and NRXN3 influence the vulnerability of children to weight gain induced by short sleep duration. *Int J Obes (Lond)*. 2013;37:182–187.
21. Bayon V, Leger D, Gomez-Merino D, et al. Sleep debt and obesity. *Ann Med*. 2014;46:264–272.
22. Capers PL, Fobian AD, Kaiser KA, et al. A systematic review and meta-analysis of randomized controlled trials of the impact of sleep duration on adiposity and components of energy balance. *Obes Rev*. 2015;16:771–782.
23. Taheri S. The link between short sleep duration and obesity: we should recommend more sleep to prevent obesity. *Arch Dis Child*. 2006;91:881–884.
24. Taveras EM, McDonald J, O'Brien A, et al. Healthy habits, happy homes: methods and baseline data of a randomized controlled trial to improve household routines for obesity prevention. *Prev Med*. 2012;55:418–426.
25. Quan SF, Parthasarathy S, Budhiraja R. Healthy sleep education–a salve for obesity? *J Clin Sleep Med*. 2010;6:18–19.
26. Cizza G, Skarulis M, Mignot E. A link between short sleep and obesity: building the evidence for causation. *Sleep*. 2005;28:1217–1220.
27. Horne J. Too weighty a link between short sleep and obesity? *Sleep*. 2008;31:595–596.
28. Horne J. Obesity and short sleep: unlikely bedfellows? *Obes Rev*. 2011;12:e84–e94.
29. Broussard J, Knutson K. *Sleep and Metabolic Disease. Sleep, Health and Society: from Aetiology to Public Health*. Oxford: Oxford University Press; 2010:111–140.
30. Carter PJ, Taylor BJ, Williams SM, et al. Longitudinal analysis of sleep in relation to BMI and body fat in children: the FLAME study. *BMJ*. 2011;342:d2712.
31. Smith PK, Bogin B, Bishai D. Are time preference and body mass index associated? Evidence from the national longitudinal survey of youth. *Econ Hum Biol*. 2005;3:259–270.
32. Komlos J, Smith PK, Bogin B. Obesity and the rate of time preference: is there a connection? *J Biosoc Sci*. 2004;36:209–219.
33. Gupta NK, Mueller WH, Chan W, et al. Is obesity associated with poor sleep quality in adolescents? *Am J Hum Biol*. 2002;14:762–768.
34. Marshall N, Stranges S. *Sleep Duration: Risk Factor or Risk marker for Ill-Health. Sleep, Health, and Society*. New York, NY: Oxford, Oxford University Press; 2010.

35. Chaput JP. Short sleep duration as a cause of obesity: myth or reality? *Obes Rev*. 2011;12:e2–e3.

36. Cizza G, Marincola P, Mattingly M, et al. Treatment of obesity with extension of sleep duration: a randomized, prospective, controlled trial. *Clin Trials*. 2010;7:274–285.

37. Rosenbaum PR. *Observational Studies*. New York, NY: Springer; 2002;1–17.

38. Craig P, Cooper C, Gunnell D, et al. Using natural experiments to evaluate population health interventions: new medical research council guidance. *J Epidemiol Community Health*. 2012;66:1182–1186.

39. Bor J. Capitalizing on natural experiments to improve our understanding of population health. *Am J Public Health*. 2016;106:1388–1389.

40. Craig P, Katikireddi SV, Leyland A, et al. Natural experiments: an overview of methods, approaches, and contributions to public health intervention research. *Annu Rev Public Health*. 2017;38:39–56.

41. Swanson SA, Hernán MA. Commentary: how to report instrumental variable analyses (suggestions welcome). *Epidemiology*. 2013;24:370–374.

42. Ministry of Gender Equality and Family. The fourth basic plan for youth policy (revised and supplemented), 2008–2012. Seoul, South Korea. 2010.

43. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *J Am Stat Assoc*. 1996;91:468–472.

44. Harris KM, Remler DK. Who is the marginal patient? Understanding instrumental variables estimates of treatment effects. *Health Serv Res*. 1998;33(5 pt 1):1337–1360.

45. Jones AM. Identification of treatment effects in health economics. *Health Econ*. 2007;16:1127–1131.

46. Angrist JD, Pischke JS. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press; 2008.

47. Jackson JW, Swanson SA. Toward a clearer portrayal of confounding bias in instrumental variable applications. *Epidemiology*. 2015;26:498–504.

48. Swanson SA, Hernán MA. The challenging interpretation of instrumental variable estimates under monotonicity. *Int J Epidemiol*. 2018;47:1289–1297.

49. Angrist JD, Imbens GW. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *J Am Stat Assoc*. 1995;90:431–442.

50. Korean Pediatric Society. *Korean Growth Standards of Children and Adolescents*. Seoul: Korea Centers for Disease Control and Prevention; 2007.

51. Olds TS, Maher CA, Matricciani L. Sleep duration or bedtime? Exploring the relationship between sleep habits and weight status and activity patterns. *Sleep*. 2011;34:1299–1307.

52. Golley RK, Maher CA, Matricciani L, et al. Sleep duration or bedtime? Exploring the association between sleep timing behaviour, diet and BMI in children and adolescents. *Int J Obes (Lond)*. 2013;37:546–551.

53. Davies NM, Smith GD, Windmeijer F, et al. Issues in the reporting and conduct of instrumental variable studies: a systematic review. *Epidemiology*. 2013;24:363–369.

54. Klingenberg L, Chaput JP, Holmbäck U, et al. Sleep restriction is not associated with a positive energy balance in adolescent boys. *Am J Clin Nutr*. 2012;96:240–248.

55. Young T. Increasing sleep duration for a healthier (and less obese?) population tomorrow. *Sleep*. 2008;31:593–594.

56. Rose G. Strategy of prevention: lessons from cardiovascular disease. *Br Med J*. 1981;282:1847–1851.

57. Rose G. *Rose's Strategy of Preventive Medicine*. New York, NY: Oxford University Press; 2008.

58. Zee PC, Turek FW. Sleep and health: everywhere and in both directions. *Arch Intern Med*. 2006;166:1686–1688.

59. Bixler E. Sleep and society: an epidemiological perspective. *Sleep Med*. 2009;10(suppl 1):S3–S6.

60. Cappuccio F, Miller MA, Lockley SW. *Sleep, Health, and Society: From Aetiology to Public Health*. Oxford: Oxford University Press; 2010.

61. Matricciani L. Subjective reports of children's sleep duration: does the question matter? A literature review. *Sleep Med*. 2013;14:303–311.

62. Ferrie JE, Kumari M, Salo P, et al. Sleep epidemiology–a rapidly growing field. *Int J Epidemiol*. 2011;40:1431–1437.