ASIA HEALTH POLICY PROGRAM

THE WALTER H. SHORENSTEIN
ASIA-PACIFIC RESEARCH CENTER

FREEMAN SPOGLI INSTITUTE
FOR INTERNATIONAL STUDIES

SHORENSTEIN APARC STANFORD

Stanford University
Walter H. Shorenstein Asia-Pacific Research Center
Asia Health Policy Program

Working paper series
on health and demographic change in the Asia-Pacific

The Effect of Sleep Duration on Body Weight in Adolescents: Evidence from a Natural Experiment

Young Kyung Do, *Department of Health Policy and Management, Seoul National University College of Medicine and Institute of Health Policy and Management, Seoul National University Medical Research Center*

Asia Health Policy Program working paper #38

February, 2014

http://asiahealthpolicy.stanford.edu

For information, contact: Karen N. Eggleston (翁笙和)
Walter H. Shorenstein Asia-Pacific Research Center
Freeman Spogli Institute for International Studies
Stanford University
616 Serra St., Encina Hall E311
Stanford, CA 94305-6055
(650) 723-9072; Fax (650) 723-6530
karene@stanford.edu

The Effect of Sleep Duration on Body Weight in Adolescents: Evidence from a Natural Experiment

Young Kyung Do*
Department of Health Policy and Management, Seoul National University College of Medicine and Institute of Health Policy and Management, Seoul National University Medical Research Center
103 Daehak-ro, Jongno-gu, Seoul 110-799, Korea
E-mail: ykdo89@snu.ac.kr

Abstract: Despite a large number of observational studies consistently reporting the association between shorter sleep duration and higher body weight, causality has yet to be established at a population level. This study aims to estimate the population-level causal effect of sleep duration on adolescent body weight, using an instrumental variable (IV) approach that exploits a unique natural experiment in the context of South Korea's highly competitive secondary education. In March 2011, amid growing concerns over the negative consequences of late-night tuition at private tutoring institutes (*hagwon*), authorities in 3 of the 16 administrative regions in South Korea decreed adjusting the closing hours of *hagwon* to 10 p.m. This policy change caused a substantial and plausibly exogenous variation in the sleep duration of the "marginal student," whose sleep duration is most likely to be affected by the policy. The IV estimation results on a sample of general high school 10th- and 11th-graders in the 2009−2012 Korea Youth Risk Behavior Web-based Survey show that a 1-hour increase in sleep duration led to a 0.56 kg/m^2 reduction in body mass index, or a 4.3 percentage-point decrease in overweight/obesity. Short sleep duration among adolescents may be an important contributor to increased body weight at the population level.

JEL Classification Codes: **I12; I18; I28**
Key Words: obesity; body weight; sleep; adolescent health; education; government policy; regulation; natural experiment; instrumental variable; difference-in-differences; Korea

---

* **Acknowledgments:** This study used data from the 5th through 8th Korea Youth Risk Behavior Web-based Survey (2009−2012), administered by the Ministry of Education, Science and Technology, Ministry of Health and Welfare, and Korea Centers for Disease Control and Prevention. Earlier versions of this paper were presented in seminars at Seoul National University, Hallym University, Duke-NUS Graduate Medical School Singapore, and the Korean Society for Preventive Medicine Annual Research Meeting. I thank Mary Ann Bautista, Ada Batcagan-Abueg, and Eunhae Shin for helpful assistance in the initial stage of the study. I also thank Marcel Bilger, Soo-yong Byun, Eric Finkelstein, Hai V. Nguyen, and Justin White for useful comments on earlier versions of this paper. All errors are my own.

1. Introduction

The escalating global epidemic of overweight and obesity poses a threat to public health (WHO, 2013a), contributing to a myriad of health problems, including diabetes and dyslipidemia (Crowley et al., 2002). The increasing prevalence of childhood obesity worldwide (WHO, 2013b) presents an alarming predicament, given that childhood overweight status is likely to persist over time, with overweight youth having an increased risk of being overweight during adulthood (Singh et al., 2008). From a policy perspective, children and adolescents are of particular interest, because they are at a critical period of developing health behaviors within the school and home environment (Kong et al., 2011; Taheri, 2006). Recently, short sleep duration has been associated with weight gain and an increased risk of overweight/obesity in children and adolescents (Agras et al., 2004; Do et al., 2013; Kong et al., 2011; Patel and Hu, 2008; Reilly et al., 2005), as well as adults (Cournot et al., 2004; Heslop et al., 2002; Patel and Hu, 2008; Patel et al., 2006; Shigeta et al., 2001; Vorona et al., 2005). A meta-analysis published in 2008 initially identified 696 studies that examined sleep and body weight in children and adults (Cappuccio et al., 2008). Results from this meta-analysis show that a 1-hour reduction in sleep duration increases the body mass index (BMI) among adults by 0.35 kg/m^2 and that the pooled odds ratio of short sleep duration for obesity is 1.89 and 1.55 among children and adults, respectively (Cappuccio et al., 2008). A systematic review of 20 longitudinal studies involving children and adults reported that short sleep duration consistently predicted weight gain in children, while the relationship was not strong in adults (Magee and Hale, 2012). Following the emergence of a large number of studies reporting the consistent association between short sleep duration and increased body weight, promoting adequate sleep hours has been proposed for obesity prevention (Taheri, 2006; Taveras et al., 2012).

Multiple mechanisms involving physiology, genetics, and behavior have been proposed to explain the link between short sleep duration and increased body weight. Shorter sleep duration can increase the caloric intake through extending waking hours, which may subsequently increase the opportunity to eat (Patel and Hu, 2008; Taheri, 2006), and increasing hunger signals (as indicated by elevations in the hunger-stimulating hormone ghrelin and reductions in the satiety hormone leptin) (Patel and Hu, 2008; Spiegel et al., 2004a; Spiegel et al., 2004b). Moreover, sleep deprivation was associated with poorer diet quality (Haghighatdoost et al., 2012) and an increase in the desire for high-calorie foods (Greer et al., 2013). Likewise, shorter sleep duration can reduce energy expenditure by modifying thermoregulation and affecting feelings of fatigue that lead to reductions in physical activity (Patel and Hu, 2008). Variations in obesity genes were found to increase the susceptibility of short sleep duration to weight gain in children (Prats-Puig et al., 2012). These effects of sleep loss (either through increased caloric intake or decreased energy expenditure, or both) ultimately contribute to a positive energy balance, where daily energy intake exceeds the energy expenditure of an individual (Markwald et al., 2013). Nonetheless, all things being equal, shorter sleep duration can lead to greater energy expenditure: individual energy expenditure during an extended wake-time is *at least* 1.0 Metabolic Equivalent Task (MET, defined as "the ratio of the work metabolic rate to the resting metabolic rate") (Ainsworth et al., 2011), which remains higher than the MET value of 0.95 during sleeping. All these findings illustrate the role of multiple intervening mechanisms and the complex interactions of biological and behavioral responses, surrounding the relationship of sleep duration with body weight. As such, attempts to quantify the contributions of a particular causal mechanism are unlikely to be feasible for practical purposes. From a public health perspective, it is important to determine the magnitude of the net causal effect of all pathways combined taking into account the relative magnitudes of the marginal changes in energy intake and energy expenditure with respect to sleep duration (Chaput et al., 2011; Crowley et al., 2002; Skidmore and Yarnell, 2004; Taheri, 2006).

Despite the accumulation of findings from initial investigations that aimed to clarify the link between sleep and body weight, the existing literature remains limited in providing a valid population-level estimate of the causal effect of sleep duration on body weight. The large body of population-based, observational epidemiological studies has yet to overcome the major issue of establishing causality between shorter sleep duration and increased body weight (Cappuccio et al., 2008; Taheri, 2006; Young, 2008). Having recognized the limited evidence to support a causal argument, several proposals for future research have been made, including improvements in the study design by controlling for more potential confounders in multiple regression models (Carter et al., 2011), and conducting prospective studies to better understand the relevant temporal sequence (Magee and Hale, 2012). Nevertheless, previously published population-based studies have attempted to address the issue of causality with limited success.

Experimental studies have focused on specific physiologic mechanisms and reported rather conflicting findings. One recent experimental study of 16 individuals in a laboratory setting found that 5 days of sleep deprivation (that is, 5 hours of sleep) increased total daily energy expenditure by approximately 5 percent, compared with the control group (that is, under 9 hours of sleep) (Markwald et al., 2013). However, the same study reported that shorter sleep caused a physiologic increase in food intake to maintain wakefulness (Markwald et al., 2013). The increased food intake was found to exceed the necessary amount to maintain energy balance, leading to an average weight gain of 0.82 kg. In another study on healthy male adults found that under free-living conditions, two nights' sleep deprivation decreased physical activity; but in laboratory conditions, short-term sleep deprivation did not increase food intake (Schmid et al., 2009). An experimental study on adolescents reported that three nights' sleep deprivation resulted in a negative energy balance, because of increased energy expenditure (longer wake time) and decreased energy intake (lower motivation to eat) (Matricciani, 2013). Another experimental study on school-age children found that increased sleep duration led to lower food intake and lower body weight (Hart et al., 2013).

Although laboratory-based experimental studies may address the concern of internal validity and elucidate potential causal mechanisms, the evidence gained from such short-term, small-scale studies cannot be easily generalized to the population. A controlled laboratory environment is also limited in capturing the total effect of complex behavioral responses to short sleep duration, which are sustained for a long-term period in the real-life setting (Horne, 2008). Furthermore, in the real world, most individuals would adjust their sleep duration at the margin, whereas most previous experimental studies have used study designs involving a 5-hour (4- vs. 9-hour) or 4-hour (5- vs. 9-hour) decrement or total sleep deprivation (Greer et al., 2013; Markwald et al., 2013; Matricciani, 2013), with the exception of the study by Hart and colleagues (2013). In this regard, an ongoing nonlaboratory-based, randomized trial of individuals with obesity investigates the feasibility of improved sleep duration through behavioral interventions as a possible strategy for addressing obesity (Cizza et al., 2010). This trial cited a relatively small sample size and nonblindness as its major limitations. In a recent study involving a 6-month health education program (Taveras et al., 2012), adequate sleep duration was identified as one of the target household routines promoted to prevent overweight/obesity. Despite such recommendations, experimental evidence to support that sleep restriction causes weight gain in free-living conditions is hard to obtain (Chaput, 2011; Nielsen et al., 2011). Hence, the central question of causality at the population level remains open, posing a major barrier to the potential translation of such research findings into public health action (Marshall et al., 2008).

The empirical challenge of inferring the causal effect of sleep duration on body weight using nonexperimental, population-level data is, in essence, concerned with the potential endogeneity of sleep duration, a time allocation decision. Biddle and Hamermesh (1990) present a theoretical model of the demand for sleep, where sleep duration is explicitly modeled as a choice variable determined by the economic variables affecting other uses of time. Consistent with this theoretical model, their empirical results show that sleep duration decreases with increasing education levels and wage rates among U.S. workers (Biddle and Hamermesh, 1990). Similar results are reported from a study on the time use of South African adults (Szalontai, 2006). Applying such a model to in-school adolescents would imply that an adolescent student's sleep duration is endogenously determined in a system involving time spent on study and leisure, as well as the marginal utilities of sleep, study, and leisure. For example, students who are strongly future-oriented (that is, those who have lower rates of time preference) may be more driven than their peers and may then voluntarily choose to reduce sleep hours at the margin to increase their study time for better academic performance and future life chances. These future-oriented individuals are also more likely to watch their food intake and engage in physical activity and are therefore less likely to be overweight/obese (Huston and Finke, 2003; Komlos et al., 2004; Smith et al., 2005). However, students who derive greater utility from immediate gratification (that is, those who have higher rates of time preference) may also choose to shorten sleep duration, albeit not for study, but for Internet games; these same students are also more likely to be overweight/obese (Komlos et al., 2004). Time preference is therefore an important but often unobservable determinant of time allocation and body weight in adolescents. Finally, inadequate sleep can be a symptom of obesity (Marshall et al., 2008), rather than a cause, suggesting the possibility of reverse causality. Ignoring such omitted variables and reverse causality in the model of body weight results in biases in the estimates of the effect of sleep duration on body weight.

The current study estimates the causal effect of sleep duration on body weight, exploiting a unique natural experiment in South Korea that produced an exogenous variation in sleep duration in adolescents. Informed by economic theory, special attention is paid to the potential endogeneity of sleep duration in guiding the empirical work and deriving policy implications.

Against the backdrop of the large volume of population-based studies (Cappuccio et al., 2008), this study tests the hypothesis that increased sleep duration leads to a reduction in BMI among in-school adolescents.

2. Background and the Natural Experiment

Secondary education in South Korea involves a structured progression from middle school (MS) (7^{th} through 9^{th} grades) to either a general high school (GHS) or a vocational high school (VHS) (10^{th} through 12^{th} grades). Just before completing the 12^{th} grade (in November each year), GHS students take the nationwide standardized College Scholastic Ability Test (CSAT), which constitutes the major component of the college admissions evaluation criteria along with their cumulative grade-point average during the 3-year high school period. Earning a slot in a high-ranking university, particularly the few elite universities in the country, is of paramount importance to adolescents and their families. Challenged by fierce competition for better college and future life opportunities, many secondary school students in South Korea often sacrifice sleep time for study. Approximately 75 percent of secondary school students take extra after-school classes at private tutoring institutes or cram schools (*hagwon*) providing shadow education (Korea National Statistical Office, 2011), which have been institutionalized into the South Korean education system (Mori and Baker, 2010), as with many other Asian countries (Bray and Lykins, 2012). Given this scenario, it is not surprising to find that the total household expenditure on private tutoring in South Korea amounts to approximately 2.8 percent of the GDP, which is equivalent to 80 percent of the government expenditure on education (Nam, 2007). Expenditures on tutoring at *hagwons* are a major source of economic burden for middle- and low-income families (Bae, 2009), suggesting its potential to propagate socioeconomic disparities in education (Byun, 2010). In Seoul, the quality of *hagwons* in each district—measured by the number of registered students who are accepted into prestigious universities—strongly predicts district-level housing price (Lee and Kim, 2003), as shown by the high demand for housing in districts with many high-performing *hagwons*.

Aside from the financial impact of *hagwons* on families, concerns have also been raised over the fact that late-night classes at *hagwons* crowd out adolescents' time for adequate leisure and sleep. The majority of *hagwons* operate until late at night, often even past midnight, primarily because students are required to stay in school quite late in the evening (spending another 2 to 3 hours after regular school hours in so-called self-study sessions that are practically mandatory). Seemingly, many students are not concerned with the negative consequences of spending hours on late-night tuition and consequently reducing sleep time (resulting in daytime sleepiness) that may also unduly compromise their academic performance. Ludicrous as it seems, *hagwons* are perceived by many to be more helpful than regular school education in equipping students with the needed test-taking skills to ensure better exam scores. The impact of these practices on students and their families prompted policymakers in South Korea 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 (Ministry of Gender Equality and Family, 2010). Such a government action to reduce the cost of shadow education is not unprecedented. A high school equalization policy was implemented in the early 1970s with the primary goal of eliminating high school entrance exams and potentially decreasing the demand for shadow education (Byun, 2010). Another prior initiative relevant to the *hagwon* curfew policy was the total ban on all kinds of private tutoring in the early 1980s, which was deemed unconstitutional in the 1990s (Kim and Lee, 2001). Despite its discontinuation, the policy has since remained a controversial topic of debate in the country, forcing the government to backtrack as enforcement became very difficult and the goal of reducing private tutoring was never really achieved (Kim and Lee, 2001).

The proposed regulation on *hagwon* curfew, led by the Presidential Council for National Future and Vision in early 2009, advised all metropolitan and provincial education authorities to set a standard curfew prohibiting the operation of *hagwons* beyond 10 p.m. (Kang, 2009a; Kang, 2009b, 2010). Determined to implement this regulation, the government guaranteed to mobilize the police for law enforcement. Despite the proposed restriction being a critical part of government efforts to reduce the household economic burden of private tutoring (Bae, 2009), it immediately stirred criticism from the education ministry and *hagwon* operators. The education ministry argued that it would be impossible to supervise and enforce the law for a large number of *hagwons* nationwide. *Hagwon* owners likewise argued that reining in students who choose to spend more time for study is simply absurd. Even the parents were pessimistic, asserting that the curfew might even stimulate an increase in the demand for private tutoring and subsequently increase fees, further widening the gap between high- and low-income earners (Bae, 2009; Byun, 2010), given that low-income households may not readily afford individual or small-group tutoring at home. As substantial resistance to the policy grew, debates on the *hagwon* curfew ensued and often evolved into political discourse between conservative and liberal camps.

Eventually, an agreement was reached about the curfew, which allows the education authority in each administrative region to autonomously set *hagwon* operating hours, giving the regional authorities the option to change ordinances according to their regional situations (Kang, 2009b). Seoul became the first region to adopt the "10 p.m. rule" for high school students in 2008. On March 1, 2011, 3 regions out of the 16 administrative regions implemented the 10 p.m. curfew for high school students. Two other regions adopted an 11 p.m. curfew in March 2011. The education ministry has sent officials to regional education offices to oversee and facilitate the regulations since 2009 (Kang, 2009c). Despite these arrangements, government officials still faced difficulties in monitoring and controlling the large number of *hagwons*. Thus, the government devised a civic monitoring system with a financial reward program that offers cash rewards to citizens reporting *hagwons* violating the regulation. These citizen patrols have come to be called the *hagparazzi* (*hagwon* paparazzi). A substantial number of violations (*hagwons* running past the curfew hours, if any, set by the administrative region) were reported by *hagparazzis*, who collectively received 3.9 billion Korean won (USD 3.49 million) between 2009 and 2012 (Na, 2012).

The theoretical model of the demand for sleep (Biddle and Hamermesh, 1990) suggests that sleep duration may be affected by changes in relevant exogenous factors, hinting at a potential empirical strategy to deal with the endogeneity issue in estimating the effect of sleep duration on body weight. Evidently, the change in *hagwon* curfew hours is an exogenous factor that shifts the demand for sleep and has two important features necessary for instrumental variable (IV) estimation. First, this regional policy is likely to have increased individual sleep duration among a certain group of adolescents. The overall magnitude of increase in sleep duration may be nontrivial and the large proportion of students attending *hagwon* classes. However, not all students are likely to comply with the policy and increase their sleep duration: some students may still choose to attend *hagwons* that illegally operate past the curfew and/or postpone bedtime for self-study at home or other activities (regardless of the new *hagwon* curfew). Because of this compliance issue, the predictive power of this regional policy change for individual sleep duration may vary by the extent to which a specific sample of adolescents is likely to contain the "marginal" student. This issue is elaborated in Section 3.1 and is empirically examined in Section 3.2.

The second important feature of this natural experiment relates to the validity of the IV. For the IV to be valid, the change in *hagwon* curfew hours should affect adolescent body weight only through inducing variations in the adolescent's sleep duration. The policy decision to change the *hagwon* curfew hours has not been part of efforts to address adolescents' weight problems. Although promoting adolescent health was included as one of the policy goals, weight problems were not specifically considered in the policy debates, possibly because the link between sleep duration and weight problems was simply not taken into account or the relevant information on the region-level prevalence of short sleep duration or weight problems was unlikely to be available. 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. Because the policymaking on *hagwon* curfew hours was left to regional authorities, the regulation policy is also free from the possible concern of selection by the central government based on certain regional characteristics—often a major methodological issue in the evaluation of government policies. Furthermore, there were no 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). Overall, this institutional background provides a conceptual justification for the exogeneity of the regional policy change in *hagwon* curfew to adolescent body weight and, consequently, the validity of the IV in the empirical work (empirical evidence is presented in Section 3.2).

Briefly, this unique natural experiment involving the policy change in *hagwon* curfew hours is likely to have caused an increase in sleep duration among a certain group of adolescents and is likely to provide a plausibly valid IV that is not directly correlated with body weight. This exogenous variation in sleep duration allows for an estimation of the causal effect of sleep duration on body weight.

3. Identification Strategy

3.1. Who Is the "Marginal Student"?

Based on the use of the term "marginal" in economics and health services research involving specific subpopulations that are likely to be affected by IVs (Gruber et al., 1999; Harris and Remler, 1998; McClellan et al., 1994), the "marginal student" in this study is defined as the student who is likely to increase his or her sleep duration in compliance with the change in *hagwon* curfew from midnight to 10 p.m.—that is, a *complier* in the parlance of the IV literature (Angrist et al., 1996). In this study, the successful implementation of the IV method rests on using the sample of students who are likely to be the marginal students; a well-developed understanding of the institutional context is critical in doing so.

GHS 10th- and 11th-graders are most likely to be the marginal students. Many GHS students in the 10th and 11th grades use *hagwons* to prepare for the CSAT and tend to reduce their sleep hours to have more time for taking extra courses at *hagwons*. However, these students still have one year left before the entrance exam, which is taken toward the end of 12th grade; hence, GHS 10th-and 11th-graders are not desperate enough to sacrifice sleep hours for more study time at every opportunity. Therefore, GHS 10th- and 11th-graders are likely to get more sleep in response to the policy restricting *hagwon* operation hours, providing the source of identification necessary for the IV estimation in this study.

GHS 12th-graders are much less likely than GHS 10th- and 11th-graders to be the marginal students. Although the policy is aimed at allowing more sleep hours among high school students, particularly 12th-graders, GHS 12th-graders may still want to spend a good amount of time on self-study even after returning home from a *hagwon* at 10 p.m. Some may also choose to take classes at *hagwons* that illegally operate past 10 p.m. The sample of GHS 12th-graders therefore consists largely of *never-takers* (Angrist et al., 1996), individuals who will not take up the treatment of more sleep regardless of assignment to a curfew region or a no-curfew region. Given that the policy may likely produce little variation in sleep duration among GHS 12th-graders, the IV estimation would suffer from a weak IV problem in this group.

Two other groups of secondary-school students are also likely to be unaffected by the policy change, but for reasons that are different from the aforementioned explanation for the sample of GHS 12th-graders. These groups include (1) VHS students, who typically do not take late-night classes at *hagwons* for college entrance, and (2) MS students (7th- through 9th-graders) who have at least three more years before taking the entrance exam and therefore are less likely to take *hagwon* classes after 10 p.m. Both of these groups are far less likely to display a behavioral response to a change in *hagwon* curfew hours. Because the policy change must be irrelevant to the sleep duration of these two groups, a valid IV should then have no effect on the students' body weight. Such a test, aptly termed as an "anti-test (or placebo test)," can provide "counter-evidence" for the validity of the IV (Jones, 2007), by showing noneffect where no effect should be found, thereby removing possible alternative explanations for the causal inference being made. If the policy change in *hagwon* curfew (the IV) was found to be directly correlated with body weight in these "IV-irrelevant" groups, the validity of the IV would be questionable. Therefore, examining these groups serves the important purpose of providing empirical evidence on the validity of the IV (Angrist, 1990), as will be described in detail in the next section (section 3.2).

3.2. Identifying Assumptions and the Wald Estimator

The identification strategy of this study can be presented in an intuitive and concise way, using a table that shows group means of the endogenous explanatory variable and the outcome variable according to IV status and then deriving the Wald estimator (Table 1) (Duflo, 2001). While time trends in sleep duration (see the third row in columns 1 and 2 for each panel) vary by group, the largest increase between the two periods is observed for the treatment group in the sample of GHS 10th- and 11th-graders (0.23 in column 2, Panel A). This increase is even more pronounced in contrast with the reduction (−0.05 in column 1) in the corresponding control group. Such differential trends are not evident in the other three panels. In contrast to these results on sleep duration, increases in BMI are consistent across various groups between the two periods (see the third row in columns 4 and 5 for each panel), with the only exception being that of the treatment group in the sample of GHS 10th- and 11th-graders (−0.05 in column 5, although this difference itself is not statistically different from 0). These findings suggest that increased sleep duration induced by the policy change in the sample of GHS 10th- and 11th-graders may explain the reduction in BMI in the affected group (Panel A), despite the overall time trends toward higher BMI among secondary-school students (Panels B−D). In addition to these time trends, baseline differences in sleep duration and BMI between the treatment and control groups should also be controlled to separate the relevant policy effects, which can be implemented in the difference-in-differences (DD) framework.

In the main study sample of GHS 10th- and 11th-graders (Panel A), the DD estimate for the policy effect on sleep duration in column 3 (with no individual-level control variables) is statistically significant and is of sizable magnitude (0.28, $p < 0.01$). This estimate indicates that the policy change in *hagwon* curfew at the regional level led to an increase in the average sleep duration by 0.28 hours at the individual level. In the same panel, the DD estimate in column 6 suggests that the policy change decreased the average BMI by 0.11 kg/m^2. The ratio of these two DD estimates (in columns 3 and 6) gives the Wald estimate of the effect of a 1-hour increase in sleep duration on BMI (Eq. 1), −0.39 in column 7.

$$\beta_{IV-Wald} = \frac{(\overline{BMI}_{1,1} - \overline{BMI}_{1,0}) - (\overline{BMI}_{0,1} - \overline{BMI}_{0,0})}{(\overline{SLEEP}_{1,1} - \overline{SLEEP}_{1,0}) - (\overline{SLEEP}_{0,1} - \overline{SLEEP}_{0,0})} \qquad \text{(Eq. 1)}$$

where $SLEEP_{1,1}$ denotes the average of sleep duration when *Treatment* = 1 and *Post* = 1, $SLEEP_{1,0}$ when *Treatment* = 1 and *Post* = 0, $SLEEP_{0,1}$ when *Treatment* = 0 and *Post* = 1, $SLEEP_{0,0}$ when *Treatment* = 0 and *Post* = 0, and the four terms in the numerator denote the average of BMI obtained in the same way as for the four terms in the denominator.

Results for the sample of GHS 12th-graders (Panel B) show that the policy change had a positive and statistically significant effect on their average sleep duration (0.06, $p < 0.05$, column 3). The magnitude of the effect, however, is much smaller than in the sample of GHS 10th- and 11th-graders, as expected. The DD estimate of the policy effect on BMI is statistically insignificant and even positive in direction (0.102, column 6), with the resulting Wald estimate being statistically insignificant (column 7).

For the samples of VHS 10th- and 11th-graders (Panel C) and MS students (Panel D), the DD estimates in columns 3 and 6 indicate no significant policy effects on sleep duration and BMI, both statistically and practically. These nonsignificant findings collectively provide supporting evidence on the second identifying assumption (the validity of the IV) for the Wald estimator in Panel A. The nonsignificant DD estimates in column 3 suggest that there were no important changes (other than the *hagwon* policy change) that could have affected sleep duration among adolescents in the treatment regions compared with the control regions.[1] Given no

[1]If the DD estimates on sleep duration (column 3) for these two groups (Panels C and D) were positive and significant (both statistically and practically), then part of the policy effect on sleep duration captured by the DD estimate in Panel A might reflect common region-level environmental influences that increased sleep duration among adolescents in general in the treatment regions. Another possibility would be a positive spill-over effect of the policy change on these two groups.

significant policy effects on sleep duration, no perceptible policy effects on body weight (nonsignificant DD estimates in column 6) add empirical evidence of the IV validity assumption. If the DD estimates on the policy effect on BMI (column 6) were found to be negative and statistically significant in these “IV-irrelevant” samples, then the reduction in BMI observed in the marginal student sample (Panel A) might not necessarily be due to the effect of increased sleep duration, but might be due to other environmental factors shared in the region. Given the large sample size (especially for the sample of MS students, N = 98,116) and the small magnitude of the four DD estimates (columns 3 and 6), their lack of statistical significance is unlikely to be attributable to a statistical power issue.

In summary, the results presented in Table 1 show the basic idea behind the IV estimation of this study, with evidence supporting the two key identifying assumptions: (1) IV strength in the marginal-student sample of GHS 10th- and 11th-graders (in comparison with the other three samples) and (2) IV validity from the two “IV-irrelevant” samples. Furthermore, the Wald estimate (−0.39 kg/m^2 per hour) in the marginal-student sample serves as a yardstick for the magnitude of the causal effect of sleep duration on BMI, which will be estimated in the full IV regression model controlling for other observable individual characteristics.

4. Methods

4.1. Empirical Model

The following linear regression model of body mass index (*BMI*) is estimated with sleep duration (*SLEEP*) and a set of control variables (*X*), in the main specification. β_1 is the coefficient of interest, α_1 the constant, and ε the error term.

$$BMI = \alpha_1 + \beta_1 SLEEP + X\beta + \varepsilon \qquad \text{(Eq. 2)}$$

The causal effect of sleep gain on weight reduction is tested by examining whether the estimated coefficient β_1 is negative and statistically significant.

The main econometric issue is that *SLEEP* is potentially endogenous, implying Corr(*SLEEP*, ε) ≠ 0 in Eq.2. As described earlier, ε may contain many unobservable factors, such as the adolescent’s time preference and the family environment, including parental discipline, which may influence both sleep duration and body weight. Therefore, the standard ordinary least squares (OLS) estimate of β_1 may suffer from omitted variable bias.

This study exploits the region-level change in the *hagwon* closing hours in South Korea as the IV. This IV represents adolescents' potential exposure to the policy change in *hagwon* curfew hours implemented as the interaction term between an indicator variable of the treatment regions (with the policy change) and an indicator variable of the post period (after the policy change). A similar identification strategy has been used in the literature (Duflo, 2001). The first-stage regression in the IV estimation can be specified as the following DD model (Eq. 3). In this equation, the coefficient $\gamma_{1,1}$ on the interaction term of *Treatment* and *Post* indicators (*Treatment* × *Post*) captures the increase in sleep duration generated in the treatment regions in the post period, while controlling for the baseline difference between the treatment and control regions (captured by $\gamma_{1,0}$ on *Treatment*) and the overall serial change across the pre and post periods (captured by $\gamma_{0,1}$ on *Post*).

$$SLEEP = \alpha_2 + \gamma_{1,1} Treatment \times Post + \gamma_{1,0} Treatment + \gamma_{0,1} Post + X\gamma + \nu \quad \text{(Eq. 3)}$$

If the estimated coefficient $\gamma_{1,1}$ is highly statistically significant in this first-stage regression, then the natural experiment satisfies the first condition of IV estimation that the IV (*Treatment* × *Post*) must have strong explanatory power for the potentially endogenous variable of *SLEEP*. Given the policy context and behavioral prediction, the expected sign of $\gamma_{1,1}$ is positive. This exogenous increase in sleep duration is then used to estimate the causal effect of *SLEEP* on *BMI* in IV two-stage least squares (IV-2SLS) estimation of the main model of *BMI* by circumventing the correlation of *SLEEP* with ε.

4.2. Data

This study uses data from the Korea Youth Risk Behavior Web-based Survey (KYRBS), a repeated cross-sectional, school-based, online survey of a large, nationally representative sample of middle-school and high school students (aged 13 through 18 years) in South Korea (Ministry of Education Science and Technology et al., 2012). The main objective of this annual survey, administered in September and October by the Korea Centers for Disease Control and Prevention (KCDC), 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 percent between 2009 and 2012.

The current study uses data from four annual surveys of KYRBS (2009−2012), which cover two surveys for each of the pre and post periods (that is, before and after the curfew change on March 1, 2011) and use nearly identical survey questions for the study variables. Out of the total sample of 298,133 secondary-school students included in the 2009–2012 KYRBS, the main IV estimation of this study focuses on the sample of GHS 10th- and 11th-graders. Observations from two regions that had not had a *hagwon* curfew but introduced an 11 p.m. curfew are not used, to ensure the homogeneity and strength of the IV. Further excluding observations with missing values for the study variables, the final analytic sample consists of 52,585 GHS 10th- and 11th-graders. Three other samples of secondary-school students are also examined to check the identifying assumptions for IV estimation: GHS 12th-graders (N = 25,892), VHS 10th- and 11th-graders (N = 15,206), and MS students (7th- through 9th-graders, N = 98,116).

4.3. Variables

The main outcome variable, BMI, is 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 is generated based on gender- and age-specific BMI cutoffs (≥ 85th percentile) in the 2007 Korean Growth Standards of Children and Adolescents (The Korean Pediatric Society, 2007).

Sleep duration, the main explanatory variable of interest, is defined as average weeknight sleep hours in the past week and is 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 data processing, for any observations with either bedtime falling outside 8 p.m. to 4 a.m. or wake-up time falling outside 4 a.m. to 9 a.m., both values were deemed abnormal and were recoded as missing in the KYRBS (Ministry of Education Science and Technology et al., 2012). Bedtime is also examined as an alternative explanatory variable in this study, because late bedtime was found to be independently associated with weight problems (that is, even after controlling for sleep duration) (Olds et al., 2011). The corresponding variable is recoded so that it ranges from 20 (8 p.m.) to 28 (4 a.m.) (that is, 24 = midnight). Based on the distribution of sleep hours, two binary indicator variables are created with the use of the 5- and 6-hour cut-off points. Perceived sleep adequacy is also examined as an alternative sleep variable, because this may be of interest for its direct behavioral implications. This variable is derived from responses to the following survey question: "Do you think your sleep hours in the past week were adequate for fatigue recovery?" ("Very adequate," "Adequate," "Moderate," "Inadequate," and "Very inadequate"). This variable is treated as a continuous variable (with values of 1 = "Very inadequate" to 5 = "Very adequate") despite its ordinal nature, to make the IV estimation easier and simpler than it would be if it were treated as an ordinal variable.

The indicator variable of *Treatment* is defined as 1 (vs. 0) for the three regions where the policy change in *hagwon* curfew to 10 p.m. occurred on March 1, 2011. Observations surveyed in 2011 and 2012 have a value of 1 (0 if observations from 2009 and 2010) for the indicator variable of *Post*. The interaction term *Treatment* × *Post* has a value of 1 (vs. 0) if both *Treatment* and *Post* have 1, and serves as the IV in this study.

Other explanatory variables include regions and year dummies (to account for the multiple regions in the treatment/control groups and two years in the pre and post periods), gender, a set of dummy variables for education levels of both parents, a set of dummy variables for self-reported household economic status ("High," "Mid-high," "Middle," "Mid-low," and "Low"), and a set of dummies for school grade.

Summary statistics of the study variables in the sample of GHS 10th- and 11th-graders are presented in Table 2. The average BMI is 20.9. The proportion of overweight/obesity in the sample is 11.4 percent. The average sleep duration is 5.7 hours, and only 41.3 percent sleep six hours or more. The average bedtime is 24.9, or 12:54 a.m. The mean value of perceived sleep adequacy is 2.7. Observations from the treatment group account for 31.4 percent of the sample, while observations are nearly equally divided between pre and post periods. Observations are also balanced in terms of gender and school grade. On average, a father's education level is higher than a mother's. The proportion for self-reported household economic status is highest in "Middle," followed by "Mid-high" and "Mid-low."

4.4. Statistical Analysis

The first-stage regression model of *SLEEP* (Eq. 3) is estimated for all of the four samples for comparison. A major goal for this stage is to check for IV strength. Because only the sample of GHS 10th- and 11th-graders turns out to have a sufficiently strong IV based on the *F*-statistics, the main IV estimation focuses on this study sample. This and all following regression analyses account for clustering of individual students in the KYRBS clusters and report cluster-robust standard errors.

The basic specification for the main IV-2SLS model has BMI as the outcome variable and sleep duration (in hours) as the explanatory variable of interest, with all other explanatory variables included. Two models of body weight (in kilograms) are also estimated, both with and without controlling for height. In addition, height is examined as an outcome, for a falsification test as another antitest (Jones, 2007). Given the average age of these students and little biological evidence on the effect of sleep on height growth in a relatively short period, height is unlikely to be affected by sleep duration in this study. Given the natural interest in the magnitude of the effect on overweight/obesity, an IV-Probit model of the binary indicator of overweight/obesity is also estimated. An additional four models of BMI are estimated, using each of the four alternative sleep variables (that is, five or more sleep hours; six or more sleep hours; bedtime; and perceived sleep adequacy) as the main explanatory variable.

A number of robustness checks are performed to examine how sensitive the main IV-2SLS estimation results are to alternative sample definitions and IV used. First, the main model is re-estimated twice, after excluding students with sleep hours greater than seven or eight, because there is some evidence to suggest a U-shaped relationship between sleep duration and body weight with a nadir of 7 to 8 hours, rather than a monotonic downward slope (Marshall et al., 2008). Second, re-estimation is made after excluding students who are classified as being underweight (< 5th percentile) (The Korean Pediatric Society, 2007). Third, analysis is restricted to adolescents in metropolitan or largely urban regions, excluding six nonmetropolitan provinces in the control group. Fourth, the model is estimated based on postperiod data from either 2011 or 2012 only (but still based on preperiod data from 2009−2010), given the difference in duration of exposure to the new *hagwon* curfew. Fifth, separate analyses are conducted for male and female students. Sixth, GHS 12th-graders are added to the main study sample, for the sake of completeness for GHS students, although these students as a group are far less likely to be marginal students. Finally, because VHS 10th- and 11th-graders are irrelevant to the policy change but are of the same age as those in the main study sample, they can serve as another dimension of control for the entire main IV estimation, offering the possibility of implementing the IV estimation in a difference-in-differences-in-differences (DDD) framework.

5. Results

The results from the first-stage regression show that the IV is positively and statistically significantly correlated with sleep duration among GHS 10^{th}- and 11^{th}-graders, but not for the other three samples (Table 3). The DD estimate of the effect of the *hagwon* curfew change on sleep duration is 0.25, comparable in magnitude to the corresponding DD estimate 0.28 in Table 3 (when no individual variables are controlled). Female students sleep less than their male counterparts do across the four samples. Students whose parents are not college-educated tend to sleep more, compared with those whose parents are college-educated. The constants reported at the bottom of Table 3 provide meaningful interpretation of the predicted values of sleep hours for adolescents in each model, with individual characteristics of the omitted reference categories (for example, 5.75 hours for the sample of GHS 10^{th}- and 11^{th}-graders).

The main IV estimate of the effect of sleep duration on body weight indicates that a 1-hour increase in sleep duration decreases adolescent BMI by 0.56 kg/m^2 (Table 4). The estimated marginal effect on weight is −1.63 kg and −1.51 kg with height uncontrolled and controlled, respectively. All of these IV estimates are larger in absolute magnitude than their respective OLS estimates, and exogeneity of sleep duration in these models is rejected ($p < 0.1$). As for the outcome variable of height, the IV estimate is statistically insignificant, whereas the corresponding OLS estimate is statistically significant ($p < 0.01$) and negative. The IV-Probit estimation result indicates that a 1-hour increase in sleep duration reduces the probability of overweight/obesity by 4.3 percentage points.

When four alternative sleep variables are used (Table 5), the IV-2SLS estimates are qualitatively consistent with the estimates in the model with sleep duration as the main variable of interest: BMI decreases with sleep duration greater than the cut-off points of five and six, earlier time to bed, and greater perceived sleep adequacy. The marginal student who would sleep five hours or more because of the change in *hagwon* curfew has a BMI reduction of 1.80 kg/m^2 (1.19 kg/m^2), compared with those who would sleep less than five (six) hours. Putting off bedtime one hour later, between 8 p.m. and 4 a.m., increases BMI by 0.69 kg/m^2. Higher levels of perceived sleep adequacy reduce BMI in IV estimation, while the OLS estimate indicates the opposite. The null hypothesis of exogeneity is rejected in all these models ($p < 0.1$). First-stage F-statistics show that the IV remains a strong predictor for these alternative sleep variables.

Results from robustness checks generally support the main IV estimation result (Table 6). Most of the statistically significant IV estimates are close in magnitude to the main IV estimate of −0.56 kg/m^2, with some differences by gender and postperiod year. The effect appears to be greater in magnitude and statistically significant for male students (versus female) and when postperiod is restricted to 2012 (versus 2011). Adding GHS 12th-graders or VHS students to the main study sample produces negative but statistically insignificant IV estimates. All IV estimates are larger in absolute magnitude than their corresponding OLS estimates, with endogeneity of sleep hours detected in the majority of the models ($p < 0.1$).

6. Discussion

This study provides population-level evidence of the causal effect of sleep duration on adolescent body weight. A 1-hour increase in sleep duration leads to a 0.56 kg/m^2 reduction in BMI and a 4.3 percentage-points decrease in overweight/obesity among adolescents in the 10th and 11th grades. As predicted by economic theory, sleep duration is found to be endogenous to body weight. Consistent with findings from a previous study (Olds et al., 2011), the effect of late bedtime has a greater magnitude than that of shorter sleep. Likewise, the current study describes gender differences in the effect of sleep duration on body weight, with the effect being stronger among males, results that are consistent with those from Knutson (2005).

Importantly, the current study addresses the central concern of causal inference in the extant epidemiological literature (Cappuccio et al., 2008), while still studying a large population-based sample in a real-life setting, and thereby overcoming the key limitations of laboratory-based experimental research (Markwald et al., 2013; Schmid et al., 2009). The study results are particularly interesting and potentially useful in informing relevant strategies and intervention studies (Skidmore and Yarnell, 2004), because the source of identification comes directly from exogenous sleep gain (away from the conventional investigations on sleep loss) induced by a natural experiment. Further discussion on the empirical strategy, as well as the research and policy implications of the study findings, are presented in proper balance with the study limitations.

6.1. Review of the Empirical Methods

This study demonstrates several key aspects of using a natural experiment for causal inference. The natural experiment of change in *hagwon* curfew hours is exploited to address the endogeneity of sleep duration in answering the important and still controversial research question of whether sleep duration has a causal effect on body weight. Both the endogeneity arguments discussed earlier and the IV method as a potential solution to address endogeneity issues are motivated by economic theory (Biddle and Hamermesh, 1990). The IV estimation is supported by data from a multiyear (spanning both pre- and postimplementation periods), nationwide (covering both treatment and control groups) study of a large sample of adolescents that contains the key study variables. A well-developed understanding of the institutional context and behavior allows for carefully checking the key identifying assumptions and permitting plausible causal inference. In doing so, presenting a table of group means with Wald estimators proves to be a simple and intuitive way of elucidating the basic workings of the IV estimation. The overall results from the main IV regression and multiple robustness checks support the causal argument.

The primary methodological strength of this paper lies in the validity of the IV. The IV derived from a region-level natural experiment is practically free from the concern that the IV itself may not be exogenous to individual choices. The IV validity is conceptually justified on the ground of the institutional context of region-level policymaking on *hagwon* curfew. The placebo test showing that there are no policy effects on body weight in the two “IV-irrelevant” samples of adolescents also provides suggestive empirical evidence. The statistically significant effect of sleep duration on weight (in kilograms) but insignificant effect on height in the IV estimation results (Table 4) further suggests that the IV is not directly correlated with individual characteristics and that it affects BMI (in kg/m^2) only through changes in sleep duration and weight (in kilograms).

The IV estimates in this study allow for offering more realistic and policy-relevant estimates of potential effects. From a public health perspective, the more important question would concern the effect of a marginal increase in sleep duration estimated on individuals who would catch more sleep under a different policy, rather than the average effect of sleep deprivation for a predefined number of hours estimated on the entire target population (Angrist and Krueger, 2001). Because the source of identification comes from the marginal student whose sleep duration was likely to change in response to the natural experiment, the IV estimates in this study are interpreted as the local average treatment effect (LATE) (Imbens and Angrist, 1994). One caveat, however, is that the IV estimates in this study can neither be generalized to other institutional settings nor directly compared with population average treatment effects.

6.2. Implications for Research

Overall, the results of this study highlight that in models of body weight, sleep duration is prone to endogeneity bias. Hence, the key methodological implication of this study relates to the importance of considering the endogeneity of sleep duration in investigating the link between sleep duration and body weight. Sleep duration may vary, not only by a random biological process but also by many individual choices involving other uses of time that are ultimately correlated with body weight. Not accounting for this unobserved heterogeneity can bias the estimates of the causal effect. All OLS estimates in this study are smaller in absolute magnitude than their corresponding IV estimates, suggesting that OLS estimation may underestimate the causal effect of sleep duration on body weight. A possible explanation for this finding is that many students with lower body weight may tend to sleep less (that is, not causally related), attenuating the true causal effect of sleep on weight in correlation and OLS estimation. The results of this study offer some suggestive evidence. Students whose parents are college-educated tend to sleep less (Table 3), but these same students also have lower body weights (results not shown). Additionally, future-oriented students may sacrifice their sleep hours for better academic performance and, at the same time, watch their food intake and include exercise time for better future health. Further insight may be drawn from the more striking estimation results, using the subjective measure of perceived sleep adequacy (Table 5). For example, future-oriented (and therefore less likely to be overweight/obese) and driven students are more likely than their present-oriented peers to perceive and report that their sleep is adequate despite having the same short sleep duration as their peers (that is, they have different standards of adequacy). These driven students might readily sacrifice their sleep because they perceive their short sleep is adequate even though their body already suffers from its negative consequences. Given this correlation, omitting time preference and other relevant variables would cause an upward bias in the estimate of perceived sleep adequacy, as shown by the switch in sign to (+) in OLS estimation (Table 5).

This study also shows that economic theory and methods have the potential to make significant contributions to the literature on sleep and obesity. Sleep is an allocation of a scarce resource: time (Biddle and Hamermesh, 1990; Szalontai, 2006). Over the past 50 years, sleep duration has been reported to decrease by as much as 2 hours in national surveys in the United States (Cappuccio et al., 2008). Another study reports that the prevalence of short sleeper among US adults increased from 7.6% in 1975 to 9.3% in 2006 (Knutson et al., 2010). The historical declines in sleep duration among children are supported by a comprehensive systematic review describing the consistent decrease in the sleep duration of children and adolescents over the last century (Matricciani et al., 2012). Several economic and lifestyle-related factors, including the expanding impact of technology and the Internet in recent years, may exacerbate the declining sleep duration observed among adolescents, because these factors crowd out time spent on sleeping. Against the backdrop of the well-known health and economic consequences of sleep loss, sleep emerges as an important area of health economics research. Adolescents who experience critical changes in sleep behavior (Carskadon et al., 2004) and other health behaviors related to body weight are then of particular interest in this respect. However, relatively little research has been conducted on sleep in this population group (Araujo et al., 2012; Taheri, 2006). Given the increasing interest in describing the early-life origins of obesity (Diethelm et al., 2011; O'Dea et al., 2012; Reilly et al., 2005), investigating sleep in children and adolescents from a health economics perspective is substantiated as an important topic for future research.

6.3. Implications for Policy and Practice

This paper sheds light on current academic and policy debates that focus on two related questions: whether reduced sleep duration is a significant contributor to increased body weight at a population level, and whether public health programs should recommend increasing sleep duration as part of obesity prevention and management (Horne, 2008; Marshall et al., 2008; Spiegel, 2008; Young, 2008). This study provides some insight to both questions by presenting empirical evidence that increased sleep duration, induced by an active public policy, causally reduced body weight at a population level. Despite a large number of observational studies consistently reporting the association between shorter sleep duration and higher body weight, the small magnitude of the association and the limited causal evidence have hindered efforts to narrow the gap between diverging perspectives on the role of sleep loss as a novel and significant risk factor for weight gain. In this study, the BMI reduction of 0.56 kg/m^2 per 1-hour increase in sleep duration is not a small magnitude for a population-level health effect. Rose (2008) highlights the importance of changing the overall distribution of risk factors at the population level in improving health, especially in disease prevention. This point has been reiterated in a contemporaneous commentary on increasing sleep duration to improve weight problems at a population level (Young, 2008). The current study findings support the view that short sleep duration among adolescents is an important public health concern that contributes to population-level weight gain. Nevertheless, this study does not suggest that short sleep duration is the main driver of the obesity epidemic. In fact, applying the study results to the clinical practice of individual-based obesity treatment would be less relevant. The relatively small magnitude of the effect of increased sleep duration on weight reduction reported in previous studies may even be challenged as having a practically insignificant benefit to an individual obese patient (Horne, 2008). The results of this study (along with the previous population-based studies) should be viewed from a population health perspective, focusing on prevention through risk reduction (Rose, 2008). Given the many other negative consequences of short sleep duration related to major chronic diseases, such as diabetes and hypertension (Spiegel, 2008), short sleep duration should then be given more attention in public health programs aimed to reduce major risk factors.

The study, however, points to some foreseeable challenges in formulating relevant public health programs, from an economic perspective. The economic theory of endogenous sleep duration, supported by the empirical findings of this study, implies that adolescents may already have optimized their time use on sleep and other activities. Thus, proposed sleep education programs (Arora et al., 2012; Quan et al., 2010) can only be economically justifiable and potentially effective to the extent that the consumption of sleep is distorted by information failure, which may not be the main market failure in obesity. Even if people are well informed of the detrimental effects of sleep loss, many individuals would find it less attractive to increase time spent sleeping (Chaput, 2011), because they would then have to sacrifice other uses of time (such as study, online social network activities, and computer games, among adolescents) that may give them a greater marginal utility than sleeping. This prediction challenges the popular notion that short sleep duration is a risk factor that can be easily modified without individual costs. Public health programs focusing only on providing sleep education may have limited effectiveness unless individuals can change the marginal rate of substitution between their time spent sleeping and other uses of time in favor of more sleep.

Because the marginal rate of substitution between sleep and other uses of time closely reflects social and environmental contexts, broader public policy may have an important role through affecting the marginal utility of other uses of time. For example, this study presents how education policy can make a substantial change in sleep duration among GHS 10^{th}- and 11^{th}-graders and, consequently, in their body weight. Interestingly, the policy change did not substantially increase sleep duration among 12^{th}-graders, probably because the marginal utility of study time was greater than that of sleep among 12^{th}-graders preparing for the upcoming CSAT. This clearly illustrates the importance of accounting for the specific context and the attractiveness of other uses of time besides sleep in any attempt to change sleep duration. Knutson (2012) highlights that identifying the social, cultural, and environmental determinants of sleep is key to addressing the pervasive negative consequences of inadequate sleep.

The next important question is: what is the market failure that would economically justify public policy intervention on sleep duration? Again, this question can best be answered in a specific social and economic context. Although, in general, there is no apparent market failure in the individual time use on sleep, the sleep duration among secondary school students in South Korea has a market failure related to the specific context of the nation's highly competitive secondary education. All else being equal, students have an incentive to spend *more* time on study than their peers, not just for good scores but also for *better* scores that would offer them an advantage in competition for better college and life chances. However, any further advantage to be gained from spending one more hour of study will always be relative to other students who also choose to spend more hours studying; the gain balances out in the end. Such a phenomenon offers a classical example of the positional arms race (Frank, 1985, 2008) and has been termed an educational arms race in the specific context of private supplementary tutoring in another Asian country (Gee, 2012). Tutoring at *hagwons* is inherently a positional good (Bray and Lykins, 2012). Each next student must race for even longer hours of tutoring at a better *hagwon* in order to remain competitive in securing a better seat in college entrance, even though the marginal increase in study time at *hagwons* may not improve social welfare any further beyond a certain point. If it is so, then the individual optimum of study is higher than the social optimum, and (given the trade-off between study and sleep) the individual optimum of sleep is lower than the social optimum, which leads to a welfare loss due to the positional externality (Frank, 2005). The *hagwon* curfew policy can then be viewed as a regulatory measure to correct the negative positional externality of private tutoring at *hagwons*.[2,3] Identifying and addressing such market failure would be a priority in formulating public policy on short sleep duration. Doing so may be more effective and welfare-improving at both individual and societal levels than promoting longer sleep duration for individuals, who might not be able or willing to bear the cost of reducing other uses of time. The study findings show that in-school adolescents do increase their individual optimum of sleep at the margin in response to a change in education policy.

6.4. Limitations

This study has several limitations. First, all variables used are derived from self-reported responses in a population-based survey, which are prone to a variety of measurement errors. Nevertheless, this study uses data on sleep duration derived from relatively specific questions on sleep (that is, weeknight bedtime and wake-up time in the past week), which include several key elements to improve the accuracy of reports on sleep duration (Matricciani, 2013). Although nonsystematic measurement errors arising from such a limitation might have diminished the magnitude and statistical significance of the estimated results, statistically significant effects are still reported in this study. The issue of systematic measurement error for sleep duration is partially addressed by the IV method used in this study, where the source of identification comes from the exogenous variation in sleep duration. Overall, the sleep duration measure used in this study has the desirable features of a population-based survey, and measurement errors, if any, are unlikely to affect the main conclusions. Still, the variable of sleep duration in this study may not necessarily reflect the duration of exposure under the reported sleep hours and any changes in sleep behavior, even based on the assumption that secondary-school students in the sample are likely to maintain a regular schedule. This study does not explore the specific mechanisms involved in decreased energy intake or increased energy expenditure, given the focus on the total net effect of sleep duration on body weight and the lack of a complete set of relevant variables.

---

[2] Despite the obvious burden on students and their families, some still view the educational arms race among secondary-school students as inevitable, even desirable, for developing an individual's intellectual capacity and strengthening the country's global market competitiveness. Investment in education is believed to have played a vital role in South Korea's rapid economic development in the second half of the twentieth century.

[3] Given that private supplementary tutoring is a positional good in the arena of a national competition for college entrance, the policy would have had a greater effect if it had been implemented nationwide.

Finally, the results of this study should still be interpreted in the specific context, even though the study has substantially improved the external validity of causal inference by examining a population group in a real-world setting (that is, beyond laboratory-based settings).

6.5. Concluding Remarks

This paper provides population-level evidence of the causal effect of sleep duration on body weight, filling an important gap in the extensive literature on sleep duration and body weight. This contribution is particularly relevant to both policy and practice, having demonstrated that a policy-induced sleep gain leads to reduced body weight at a population level. Economic theory applied to sleep also offers critical insights on the challenges and opportunities in the development of relevant public health programs and public policies. Given the ever-increasing use of the Internet and other technologies that may crowd out time spent sleeping, and the pervasive effects of sleep on energy intake and expenditure, sleep is an important factor worth exploring in the growing areas of economic research on obesity.

References

Agras WS, Hammer LD, McNicholas F, Kraemer HC. Risk factors for childhood overweight: A prospective study from birth to 9.5 years. *Journal of Pediatrics* 2004; 145; 20-25.

Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Jr., Tudor-Locke C, Greer JL, Vezina J, Whitt-Glover MC, Leon AS. 2011 Compendium of Physical Activities: A second update of codes and MET values. *Med Sci Sports Exerc* 2011; 43; 1575-1581.

Angrist JD. Lifetime earnings and the Vietnam Era draft lottery: Evidence from Social Security administrative records. *American Economic Review* 1990; 80; 313-336.

Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables: Rejoinder. *Journal of the American Statistical Association* 1996; 91; 468-472.

Angrist JD, Krueger AB. Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives* 2001; 15; 69-85.

Araujo J, Severo M, Ramos E. Sleep duration and adiposity during adolescence. *Pediatrics* 2012; 130; e1146-e1154.

Arora T, Hosseini-Araghi M, Bishop J, Yao GL, Thomas GN, Taheri S. The complexity of obesity in UK adolescents: Relationships with quantity and type of technology, sleep duration and quality, academic performance and aspiration. *Pediatric Obesity* 2012.

Bae J. "Will hagwon curfew work?" *Korea Times.* The Korea Times, 26 April 2009. Web. 12 March 2013.

Biddle JE, Hamermesh DS. Sleep and the allocation of time. *Journal of Political Economy* 1990; 98; 922-943.

Bray M, Lykins C. *Shadow Education: Private Supplementary Tutoring and Its Implications for Policy Makers in Asia*. Asian Development Bank: Mandaluyong City; 2012.

Byun S-Y. Does policy matter in shadow education spending? Revisiting the effects of the high school equalization policy in South Korea. *Asia Pacific Education Review* 2010; 11; 83-96.

Cappuccio FP, Taggart FM, Kandala NB, Currie A, Peile E, Stranges S, Miller MA. Meta-analysis of short sleep duration and obesity in children and adults. *Sleep* 2008; 31; 619-626.

Carskadon MA, Acebo C, Jenni OG. *Regulation of Adolescent Sleep: Implications for Behavior.* Annals of the New York Academy of Sciences 2004; 1021; 276-291.

Carter PJ, Taylor BJ, Williams SM, Taylor RW. Longitudinal analysis of sleep in relation to BMI and body fat in children: FLAME study. *British Medical Journal* 2011; 342; d2712.

Chaput JP. Short sleep duration as a cause of obesity: myth or reality? *Obesity Reviews* 2011; 12; e2-3.

Chaput JP, Lambert M, Gray-Donald K, McGrath JJ, Tremblay MS, O'Loughlin J, Tremblay A. Short sleep duration is independently associated with overweight and obesity in Quebec children. *The Canadian Journal of Public Health* 2011; 102; 369-374.

Cizza G, Marincola P, Mattingly M, Williams L, Mitler M, Skarulis M, Csako G. Treatment of obesity with extension of sleep duration: A randomized, prospective, controlled trial. *Clinical Trials* 2010;7; 274-285.

Cournot M, Ruidavets JB, Marquie JC, Esquirol Y, Baracat B, Ferrieres J. Environmental factors associated with body mass index in a population of Southern France. *European Journal of Cardiovascular Prevention and Rehabilitation* 2004; 11; 291-297.

Crowley VEF, Yeo GSH, O'Rahilly S. Obesity therapy: altering the energy intake-and-expenditure balance sheet. *Nature Reviews Drug Discovery* 2002; 1; 276-286.

Diethelm K, Bolzenius K, Cheng G, Remer T, Buyken AE. Longitudinal associations between reported sleep duration in early childhood and the development of body mass index, fat mass index and fat free mass index until age 7. *International Journal of Pediatric Obesity* 2011; 6; e114-123.

Do YK, Shin E, Bautista MA, Foo K. The associations between self-reported sleep duration and adolescent health outcomes: What is the role of time spent on Internet use? *Sleep Med* 2013; 14; 195-200.

Duflo E. Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review* 2001; 91; 795-813.

Frank RH. The demand for unobservable and other nonpositional goods. *American Economic Review* 1985; 75; 101-116.

Frank RH. Positional externalities cause large and preventable welfare losses. *American Economic Review* 2005; 95; 137-141.

Frank RH. Should public policy respond to positional externalities? *Journal of Public Economics* 2008; 92; 1777-1786.

Gee C. 2012. The Education Arms Race: All for one, loss for all. IPS Working Paper. Institute of Policy Studies: Singapore; 2012.

Greer SM, Goldstein AN, Walker MP. The impact of sleep deprivation on food desire in the human brain. *Nature Communications* 2013; 4.

Gruber J, Levine P, Staiger D. Abortion Legalization and Child Living Circumstances: Who Is the "Marginal Child"? *Quarterly Journal of Economics* 1999;114; 263-291.

Haghighatdoost F, Karimi G, Esmaillzadeh A, Azadbakht L. 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.

Harris KM, Remler DK. Who is the marginal patient? Understanding instrumental variables estimates of treatment effects. *Health Services Research* 1998; 33; 1337-1360.

Hart CN, Carskadon MA, Considine RV, Fava JL, Lawton J, Raynor HA, Jelalian E, Owens J, Wing R. Changes in children's sleep duration on food intake, weight, and leptin. *Pediatrics* 2013;132(6):e1473-80.

Heslop P, Smith GD, Metcalfe C, Macleod J, Hart C. Sleep duration and mortality: The effect of short or long sleep duration on cardiovascular and all-cause mortality in working men and women. *Sleep Med* 2002; 3; 305-314.

Horne J. Too weighty a link between short sleep and obesity? *Sleep* 2008; 31; 595-596.

Huston SJ, Finke MS. Diet choice and the role of time preference. *Journal of Consumer Affairs* 2003; 37; 143-160.

Imbens GW, Angrist JD. Identification and estimation of local average treatment effects. *Econometrica* 1994; 62; 467-475.

Jones AM. Identification of treatment effects in healtheconomics. *Health Economics* 2007; 16; 1127-1131.

Kang H. Plan to set curfew for hagwon scrapped. *Korea Times.* The Korea Times*, 18 May* 2009a. Web. 12 March 2013.

Kang S. Hagwon curfew ruled constitutional. *Korea Times.* The Korea Times, 29 Oct. 2009b. Web. 12 March 2013.

Kang S. Tipsters on illegal hagwon to get up to 2 mil. reward. *Korea Times.* The Korea Times, 6 June 2009c. Web. 12 March 2013.

Kang S. Hagwon curfew backsliding. *Korea Times* The Korea Times, 1 April 2010. Web. 12 March 2013.

Kim S, Lee J-H. *Demand for Education and Developmental State: Private Tutoring in South Korea*. University of Wisconsin, Milwaukee, USA; KDI School of Public Policy and Management, Seoul, Korea 2001.

Knutson KL. Sex differences in the association between sleep and body mass index in adolescents. *Journal of Pediatrics* 2005; 147; 830-834.

Knutson KL. Does inadequate sleep play a role in vulnerability to obesity? *American Journal of Human Biology* 2012; 24; 361-371.

Knutson KL, Van Cauter E, Rathouz PJ, DeLeire T, Lauderdale DS. Trends in the prevalence of short sleepers in the USA: 1975-2006. *Sleep* 2010; 33(1):37-45.

Komlos J, Smith PK, Bogin B. Obesity and the rate of time preference: Is there a connection? *Journal of Biosocial Science* 2004; 36; 209-219.

Kong AP, Wing YK, Choi KC, Li AM, Ko GT, Ma RC, Tong PC, Ho CS, Chan MH, Ng MH, Lau J, Chan JC. Associations of sleep duration with obesity and serum lipid profile in children and adolescents. *Sleep Med* 2011; 12; 659-665.

Korea National Statistical Office. The survey of private education expenditures in 2010. 2011.

Lee S-G, Kim M-R. 2003. An address some people will die for. *Korea JoongAng Daily*. JoongAng Ilbo, 12 June 2003. Web. 12 March 2013.

Magee L, Hale L. Longitudinal associations between sleep duration and subsequent weight gain: A systematic review. *Sleep Medicine Reviews* 2012; 16; 231-241.

Markwald RR, Melanson EL, Smith MR, Higgins J, Perreault L, Eckel RH, Wright KP, Jr. Impact of insufficient sleep on total daily energy expenditure, food intake, and weight gain. *Proceedings of the National Academy of Sciences USA 2013*; 5695-5700.

Marshall NS, Glozier N, Grunstein RR. Is sleep duration related to obesity? A critical review of the epidemiological evidence. *Sleep Medicine Reviews* 2008; 12; 289-298.

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

Matricciani L, Olds T, Petkov J. In search of lost sleep: secular trends in the sleep time of school-aged children and adolescents. *Sleep Med Rev* 2012;16(3):203-11.

McClellan M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *Journal of the American Medical Association* 1994; 272; 859-866.

Ministry of Education Science and Technology, Ministry of Health and Welfare, Korea Centers for Disease Control and Prevention. *User Guides for the Korea Risk Behavior Web-Based Survey (2005−2012)*; 2012.

Ministry of Gender Equality and Family. *Youths Basic Law*. Ministry of Gender Equality and Family; 2010.

Mori I, Baker D. The origin of universal shadow education: What the supplemental education phenomenon tells us about the postmodern institution of education. *Asia Pacific Education Review* 2010;11; 36-48.

Na J-J. Bounty hunters targeting hagwons. *Korea Times* 2012.

Nam K. Time-series analysis of the scale of private supplementary tutoring. *Economic Study of Educational Finance* 2007; 16; 57-79.

Nielsen LS, Danielsen KV, Sørensen TIA. Short sleep duration as a possible cause of obesity: Critical analysis of the epidemiological evidence. *Obesity Reviews* 2011; 12; 78-92.

O'Dea JA, Dibley MJ, Rankin NM. Low sleep and low socioeconomic status predict high body mass index: A 4-year longitudinal study of Australian schoolchildren. *Pediatric Obesity* 2012; 7; 295-303.

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.

Patel SR, Hu FB. Short sleep duration and weight gain: A systematic review. *Obesity* (Silver Spring) 2008; 16; 643-653.

Patel SR, Malhotra A, White DP, Gottlieb DJ, Hu FB. Association between reduced sleep and weight gain in women. *Am J Epidemiol* 2006;164; 947-954.

Prats-Puig A, Grau-Cabrera P, Riera-Pérez E, Cortés-Marina R, Fortea E, Soriano-Rodríguez P, de Zegher F, Ibánez L, Bassols J, López-Bermejo A. Variations in the obesity genes FTO, TMEM18 and NRXN3 influence the vulnerability of children to weight gain induced by short sleep duration. *International Journal of Obesity* 2012; 37; 182-187.

Quan SF, Parthasarathy S, Budhiraja R. Healthy sleep education--a salve for obesity? *Journal of Clinical Sleep Medicine* 2010; 6; 18-19.

Reilly JJ, Armstrong J, Dorosty AR, Emmett PM, Ness A, Rogers I, Steer C, Sherriff A. Early life risk factors for obesity in childhood: Cohort study. *British Medical Journal* 2005; 330; 1357.

Rose G. *Rose's Strategy of Preventive Medicines*. Oxford University Press: New York; 2008.

Schmid SM, Hallschmid M, Jauch-Chara K, Wilms B, Benedict C, Lehnert H, Born J, Schultes B. 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. *American Journal of Clinical Nutrition* 2009; 90; 1476-1482.

Shigeta H, Shigeta M, Nakazawa A, Nakamura N, Yoshikawa T. Lifestyle, obesity, and insulin resistance. *Diabetes Care* 2001; 24; 608.

Singh AS, Mulder C, Twisk JW, van Mechelen W, Chinapaw MJ. Tracking of childhood overweight into adulthood: A systematic review of the literature. *Obesity Reviews* 2008; 9; 474-488.

Skidmore PM, Yarnell JW. The obesity epidemic: Prospects for prevention. *QJM: An International Journal of Medicine* 2004; 97; 817-825.

Smith PK, Bogin B, Bishai D. Are time preference and body mass index associated? Evidence from the National Longitudinal Survey of Youth. *Economics of Human Biology* 2005; 3; 259-270.

Spiegel K. Sleep loss as a risk factor for obesity and diabetes. *International Journal of Pediatric Obesity* 2008; 3 Suppl 2; 27-28.

Spiegel K, Leproult R, L'Hermite-Baleriaux M, Copinschi G, Penev PD, Van Cauter E. Leptin levels are dependent on sleep duration: Relationships with sympathovagal balance, carbohydrate regulation, cortisol, and thyrotropin. *Journal of Clinical Endocrinology and Metabolism* 2004a; 89; 5762-5771.

Spiegel K, Tasali E, Penev P, Van Cauter E. 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* 2004b;141; 846-850.

Szalontai G. The demand for sleep: A South African study. *Economic Modelling* 2006; 23; 854-874.

Taheri S. The link between short sleep duration and obesity: We should recommend more sleep to prevent obesity. *Archives of Diseases in Childhood 2006*; 91; 881-884.

Taveras EM, McDonald J, O'Brien A, Haines J, Sherry B, Bottino CJ, Troncoso K, Schmidt ME, Koziol R. 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.

The Korean Pediatric Society. Committee of Korean Growth Standards Enactment. Korean growth standards of children and adolescents. Korea Centers for Disease Control and Prevention: Seoul; 2007.

Vorona RD, Winn MP, Babineau TW, Eng BP, Feldman HR, Ware JC. Overweight and obese patients in a primary care population report less sleep than patients with a normal body mass index. *Arch Intern Med* 2005; 165; 25-30.

WHO. 2013a. World Health Organization. Childhood overweight and obesity.

WHO. 2013b. World Health Organization. Controlling the global obesity epidemic.

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

Table 1. Group means of sleep hours and body mass index by period and region's change in *hagwon* closing hours

Sample	Sleep duration (hours)			Body mass index (kg/m^2)			
	Change in *hagwon* closing hours in region			Change in *hagwon* closing hours in region			
	No change (Control, C)	MN to 10 p.m. (Treatment, T)	Difference = T − C	No change (Control, C)	MN to 10 p.m. (Treatment, T)	Difference = T − C	Wald estimate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. GHS 10−11th (N=52,585)							
Pre (2009 & 2010)	5.708	5.518	−0.190**	20.912	20.833	−0.078*	
			(0.015)			(0.039)	
Post (2011 & 2012)	5.655	5.745	0.090**	20.971	20.783	−0.188**	
			(0.014)			(0.036)	
Difference = Post – Pre	−0.052**	0.227**	0.280**	0.059*	−0.051	−0.110*	−0.392*
	(0.011)	(0.017)	(0.020)	(0.030)	(0.043)	(0.053)	(0.190)
Panel B. GHS 12th (N=25,892)							
Pre (2009 & 2010)	5.235	5.127	−0.108**	21.285	21.151	−0.133*	
			(0.021)			(0.056)	
Post (2011 & 2012)	5.264	5.219	−0.045*	21.411	21.380	−0.031	
			(0.019)			(0.051)	
Difference = Post – Pre	0.029	0.092**	0.063*	0.127**	0.229**	0.102	1.619
	(0.016)	(0.023)	(0.028)	(0.042)	(0.063)	(0.076)	(1.419)
Panel C. VHS 10−11th (N=15,206)							
Pre (2009 & 2010)	6.306	6.243	−0.064	20.788	20.881	0.093	
			(0.035)			(0.081)	
Post (2011 & 2012)	6.298	6.209	−0.090**	20.991	21.052	0.060	
			(0.030)			(0.069)	
Difference = Post – Pre	−0.008	−0.034	−0.026	0.203**	0.171	−0.033	1.256
	(0.025)	(0.038)	(0.046)	(0.057)	(0.092)	(0.107)	(4.689)
Panel D. MS 7−9th (N=98,116)							
Pre (2009 & 2010)	7.039	6.974	−0.065**	20.054	19.942	−0.112**	
			(0.012)			(0.031)	
Post (2011 & 2012)	7.126	7.077	−0.049**	20.126	20.012	−0.114**	
			(0.011)			(0.028)	
Difference = Post – Pre	0.087**	0.103**	0.016	0.073**	0.070*	−0.003	−0.180
	(0.009)	(0.014)	(0.017)	(0.023)	(0.034)	(0.042)	(2.558)

Notes: Data come from the 2009−2012 Korea Youth Risk Behavior Web-based Survey. GHS = general high school; VHS = vocational high school; MS = middle school. Standard errors are in parentheses. ***p* < 0.01, **p* < 0.05.

Table 2. Summary statistics of variables in the sample of GHS 10^{th}- and 11^{th}-graders

Variables	Mean or Frequency	Std. Dev. or %
Weight-related outcomes		
Body mass index (kg/m^2)	20.9	2.8
Weight (kg)	58.8	10.5
Height (cm)	167.4	8.1
Overweight/obesity (1 if yes, 0 otherwise)	5984	11.4%
Sleep variables		
Sleep hours (h)	5.7	1.1
Sleep hours 5 or more (1 if yes, 0 if <5)	40408	76.8%
Sleep hours 6 or more (1 if yes, 0 if <6)	21714	41.3%
Bedtime (range 20−28 [= 8PM−4AM])	24.9	1.0
Perceived sleep adequacy (1−5: 5="Very adequate")	2.7	1.1
Treatment and period variables		
Treatment × Post (1 if Treatment=1 & Post=1)	9126	17.4%
Treatment (1 if curfew changed to 10PM, 0 otherwise)	16503	31.4%
Post (1 if Year 2011−2012, 0 if Year 2009−2010)	27588	52.5%
Year dummies		
Y2009 (1 if 2009, 0 otherwise)	12926	24.6%
Y2012 (1 if 2012, 0 otherwise)	14124	26.9%
Other individual-level explanatory variables		
Female (1 if female, 0 if male)	25750	49.0%
11^{th} grade (1 if yes, 0 if 10^{th} grade)	26545	50.5%
Father's education level		
College or more	27402	52.1%
High school	19333	36.8%
Middle school	2590	4.9%
Don't know	3260	6.2%
Mother's education level		
College or more	19819	37.7%
High school	26726	50.8%
Middle school	2571	4.9%
Don't know	3468	6.6%
Household economic status		
High	2579	4.9%
Mid-high	12211	23.2%
Middle	26340	50.1%
Mid-low	9291	17.7%
Low	2164	4.1%

Notes: The sample comes from general high school students in the 10^{th} and 11^{th} grades surveyed in the 2009−2012 Korea Youth Risk Behavior Web-based Survey (N = 52,585), excluding subjects in Jeonbuk and Incheon (Group 2 in Table 1). Dummies for region are not presented in this table.

Table 3. First-stage regression of sleep hours

	Main sample	Comparison sample		
	GHS 10−11th	GHS 12th	VHS 10−11th	MS 7−9th
Variables	[N=52,585]	[N=25,892]	[N=15,206]	[N=98,116]
Instrumental variable				
Treatment × Post	0.252**	0.025	0.034	0.011
	(0.044)	(0.050)	(0.083)	(0.027)
Treatment[a]	−0.045	0.131*	0.464**	0.114**
	(0.052)	(0.054)	(0.089)	(0.031)
Post (Years 2011−2012)	−0.021	0.082*	0.122*	0.107**
	(0.032)	(0.034)	(0.057)	(0.020)
Year dummies				
Y2009	0.043	0.092**	0.130*	0.012
	(0.028)	(0.032)	(0.062)	(0.018)
Y2010 (omitted)	-	-	-	-
Y2011 (omitted)	-	-	-	-
Y2012	−0.006	0.013	−0.111*	0.044**
	(0.027)	(0.028)	(0.047)	(0.017)
Female	−0.311**	−0.183**	−0.367**	−0.488**
	(0.015)	(0.019)	(0.027)	(0.010)
Father's education level				
College or more (omitted)	-	-	-	-
High school	0.104**	0.080**	0.128**	0.137**
	(0.013)	(0.016)	(0.029)	(0.010)
Middle school	0.203**	0.184**	0.204**	0.179**
	(0.027)	(0.034)	(0.043)	(0.022)
Don't know	0.185**	0.117**	0.163**	0.178**
	(0.028)	(0.043)	(0.052)	(0.013)
Mother's education level				
College or more (omitted)	-	-	-	-
High school	0.059**	−0.005	0.089**	0.109**
	(0.012)	(0.017)	(0.032)	(0.010)
Middle school	0.045	0.071*	0.122**	0.164**
	(0.025)	(0.035)	(0.046)	(0.023)
Don't know	0.192**	0.129**	0.263**	0.157**
	(0.027)	(0.046)	(0.048)	(0.013)
Other control variables[b]	Yes	Yes	Yes	Yes
Constant	5.750**	5.092**	6.105**	7.215**
	(0.040)	(0.045)	(0.099)	(0.027)
R-sq.	0.051	0.030	0.055	0.107
F-statistic	32.977**	0.250	0.164	0.172

Notes: GHS = general high school; VHS = vocational high school; MS = middle school. [a]A dummy indicator for regions where ordinance on *hagwon* closing hours changed from midnight to 10 p.m. on March 1, 2011. [b]Other control variables are a set of dummy variables for region, five-category household economic status, and school grade where applicable. Standard errors are in parentheses. ** $p < 0.01$, * $p < 0.05$.

Table 4. Effect of sleep duration (in hours) on body mass index and other outcomes

Outcomes	OLS/ Probit	IV-2SLS/ IV-Probit	Exogeneity test	Marginal effect
Body mass index (kg/m^2)	−0.103**	−0.556*	p=0.061	0.56 kg/m^2 reduction
	(0.012)	(0.250)		
Weight (kg)	−0.391**	−1.629*	p=0.090	1.6 kg reduction
	(0.039)	(0.752)		
Weight (kg) (height controlled)	−0.283**	−1.508*	p=0.078	1.5 kg reduction
	(0.034)	(0.716)		
Height (cm)	−0.144**	−0.163	p=0.963	Not statistically sig.
	(0.023)	(0.407)		
Overweight/obesity (1 vs. 0)	−0.031**	−0.215†	p=0.160	4.3%p reduction
	(0.007)	(0.127)		

Notes: The sample of general high school students in the 10th and 11th grades (N=52,585) is used. Coefficients and standard errors (in parentheses) are presented. Marginal effects of one additional hour of sleep are derived from IV estimates. **p < 0.01, *p < 0.05, †p < 0.1.

Table 5. Effect of alternative sleep variable on body mass index

Alternative sleep variable	OLS	IV-2SLS	Exogeneity test	First-stage F
Sleep hours 5 or more (vs. <5)	−0.196**	−1.802*	p=0.042	29.569**
	(0.029)	(0.839)		
Sleep hours 6 or more (vs. <6)	−0.174**	−1.187*	p=0.050	39.956**
	(0.026)	(0.528)		
Bedtime[a] (20−28)	0.075**	0.692*	p=0.042	30.494**
	(0.013)	(0.326)		
Perceived sleep adequacy[b] (1−5)	0.074**	−0.927*	p=0.014	27.772**
	(0.012)	(0.437)		

Notes: The sample of general high school students in the 10th and 11th grades (N = 52,585) is used. Coefficients and standard errors (in parentheses) are presented. [a]Higher number denotes later bedtime, ranging from 20 (8 p.m.) to 28 (4 a.m.). [b]Higher number denotes greater perceived adequacy, ranging from 1 ("Very inadequate") to 5 ("Very adequate"). **p < 0.01, *p < 0.05.

Table 6. Robustness checks: Marginal effect of sleep hours on body mass index

Sample definition and IV	OLS	IV-2SLS	Exogeneity test	First-stage F
Main study sample				
GHS 10th−11th-graders in 2009−2012	−0.103**	−0.556*	p=0.061	32.977**
[N=52,585]	(0.012)	(0.250)		
Alternative sample definition				
Excluding obs. with sleep hours > 8	−0.112**	−0.614*	p=0.049	33.896**
[N=51,891]	(0.012)	(0.265)		
Excluding obs. with sleep hours > 7	−0.112**	−0.617*	p=0.097	34.741**
[N=47,816]	(0.015)	(0.314)		
Excluding underweight students	−0.079**	−0.487*	p=0.069	32.702**
[N=49,034]	(0.011)	(0.233)		
Excluding nonmetropolitan control	−0.105**	−0.538*	p=0.066	40.075**
[N=35,146]	(0.014)	(0.238)		
Post period 2011 only	−0.111**	−0.446	p=0.222	28.167**
[N=38,461]	(0.014)	(0.280)		
Post period 2012 only	−0.119**	−0.666*	p=0.067	20.780**
[N=39,121]	(0.014)	(0.317)		
Male students only	−0.127**	−0.787*	p=0.048	26.267**
[N=26,835]	(0.018)	(0.352)		
Female students only	−0.079**	−0.333	p=0.403	19.028**
[N=25,750]	(0.015)	(0.300)		
Adding GHS 12th	−0.109**	−0.271	p=0.612	19.945**
[N=78,477]	(0.010)	(0.316)		
Alternative IV and sample				
IV from DDD (adding VHS sample)	−0.095**	−0.777	p=0.244	5.421*
[N=67,791]	(0.010)	(0.652)		

Notes: GHS = general high school; VHS = vocational high school; IV = instrumental variable; DDD = difference-in-differences-in-differences. Coefficients and standard errors (in parentheses) are presented. **p < 0.01, *p < 0.05.