Does Weight Status Affect Academic Performance? Evidence from Australian Children

Hui Shi
Victoria Institute of Strategic Economic Studies, Victoria University

Chuhul Li
Department of Econometrics and Business Statistics, Monash University

Working Paper No. 52
Victoria Institute of Strategic Economic Studies
Victoria University
Melbourne

December 2015

Does weight status affect academic performance? Evidence from Australian children

Hui Shi *[a] and Chuhui Li[b]

[a]Victoria Institute of Strategic Economic Studies, Victoria University
[b]Department of Econometrics and Business Statistics, Monash University

Abstract

This paper investigates the impact of children's weight status on their academic performance using the Longitudinal Survey of Australian Children (LSAC). Considering the endogeneity of the weight status variable (obesity or overweight), to obtain consistent estimates of its impact, we use body mass index (BMI) of biological parents as instrumental variables. The two stage estimation shows that obesity or overweight has significant negative impact on academic results, which is much larger than the OLS estimate. With estimation for separate year groups, we find that the academic impact of obesity or overweight especially on numeracy is generally larger for senior year students.

Keywords: Obesity; Overweight; Academic performance; Instrument variable; Australia.

JEL codes: I1, I2, C1.

1 Introduction

Obesity has been rising sharply worldwide in recent years. This epidemic problem not only occurs among adults, but also becomes more serious in children and adolescents. According to estimations by Ng et al. (2014), prevalence has increased substantially in children

*Corresponding author. Email address: hui.shi@vu.edu.au

and adolescents in developed countries. For example, in 2013, 24% of boys and 23% of girls were overweight or obese, compared with 17% of boys and 16% of girls in 1980. Specifically, the Centers for Disease Control and Prevention (CDC) reported a rapid 4-fold rise in child and adolescent obesity (aged 6–19) over the past 20 years in US (Taras and Potts-Datema, 2005). Australian Bureau of Statistics (ABS) also reports that one-quarter of all Australian children aged 5–17 in 2007–08 were overweight or obese, up four percentage points from 1995. Developing countries are facing the problem of childhood obesity as well. In total, the prevalence of overweight and obesity in children and adolescents has also increased from 8.1% to 12.9% for boys and from 8.4% to 13.4% for girls during 1980–2013 (Ng et al., 2014). Obesity is more common in new emerging countries. For example, China has experienced a rapid increase in obesity. By 2009, approximately 12% of children and adolescents were overweight and 3% of children and 1% of adolescents were obese (China Health and Nutrition Survey). India is facing the same burden of child obesity. A meta-analysis based on studies from 2000 to 2009 has found that the prevalence of childhood overweight was estimated to be 12.6% and that of obesity to be 3.4% (Midha et al., 2012).

Child obesity is attributed to individual choices and behaviours on one hand. On the other hand, biological and genetic factors, socio-economic conditions, combined with influences from individuals' social and physical surroundings, are responsible for creating 'obesogenic environments' which promote obesity (Butland et al., 2007; Jones et al., 2007; Swinburn et al., 1999; Swinburn and Egger, 2004). Lake and Townshend (2006) propose that both 'micro-environments' (e.g. school, home, neighborhood) and 'macro-environments' (e.g. education and health systems, government policy and society's attitudes and beliefs) work together to influence obesity.

Obesity can cause physical, social and emotional health problems in childhood and adolescence. Nevertheless, individuals, communities or even policy makers do not entirely understand the implications of obesity. In countries or societies such as Tonga and some middle east countries (Kiribati, Federated States of Micronesia, Libya, Qatar and Samoa) where estimated prevalence of adult obesity exceeded 50%, obesity or overweight is even regarded as a sign of affluence and health. Physically, obese children and adolescents suffer from an increased range of medical conditions such as the increasing rate of Type 2 diabetes in children and adolescents. Other problems include sleep disturbances, heart intolerance, breathlessness when active, and tiredness. It is very important to prevent and manage obesity in children as there is a high risk that the problem will persist into adulthood. According to Dietitians Association of Australia, obese children in Australia have a 25 to 50% chance of becoming obese adults, and this possibility can be as high as 78% for obese adolescents. Socially and emotionally, obese children tend to feel stressed, i.e., they are more likely to have bully behaviour and lower self-esteem, particularly for girls. They also have higher rates of anxiety disorders, depression and other mental illnesses.

Studies have found that poor health can severely weaken the learning capacity of students (Mayer-Foulkes, 2005; Bloom and Canning, 2009; Behrman, 1996; Glewwe and Jacoby, 1995; Glewwe et al., 2001; Wisniewski, 2010; Ding et al., 2009; Zhao et al., 2012; etc.) As obesity is considered to be a health concern, an important question in this study is whether obesity in childhood has any impact on a child's academic performance. Some studies look at the relationship between obesity and behaviours affecting school performance. Schwimmer et al. (2003) examine the association of obesity and school absenteeism. This association has reinforced the results by Falkner et al. (2001) that obese teenagers are more likely to view themselves as poor students and therefore hold back grades or drop out school. Gable et al. (2012) test the mediation of interpersonal skills and internalizing behaviours between weight status and math performance.

Some researches find connections between obesity and factors which contribute to low academic achievement. For example, Tobin (2013) has found that students who had a higher than average intake of fast food had between 5% and 16% lower reading test scores and between 6% and 18% lower math test scores. Yau et al. (2012) have shown that the prevalence of obesity is paralleled with a rise in metabolic syndrome which may affect cognitive abilities and therefore academic achievement. Another major factor in weight, physical activity, has been found to be positively correlated with GPA (Kantomaa et al., 2013). In addition, the poor school performance may be related to several social circumstances such as family income, ethnicity, the mother's education level and job status, and both parents' expectations for the child's performance in school and these factors may lead to the child obesity problem. Mikkilä et al. (2003) have found that obesity and weight dissatisfaction are associated mostly with disadvantageous health behaviours and low socio-economic status among Finnish adolescents.

Although connections between obesity and impact factors for academic performance has been found in the aforementioned researches, the direct association between obesity and academic performance has not been extensively established with a few studies such as Wendt and Kinsey (2009), Sabia (2007), and Crosnoe (2007). More important, the causal relationship between child obesity and school performance has not been well discussed [1]. Datar et al. (2004) find that overweight children had significantly lower math and reading test scores compared with non-overweight children in kindergarten and at the end of first grade, but they suggest that obesity is a marker, not a cause, of low academic performance. Gable et al. (2012) follow 6,250 children from kindergarten through fifth grade and find that those who were obese throughout that period scored lower on math tests than non-obese children. However, they do not test for causal relations among the variables.

This paper aims to investigate the impact of child weight status measured with obesity and overweight on educational outcomes, utilizing data from the Longitudinal Survey of Australian Children (LSAC). There are currently 5 waves of data collection for the same sample of children followed from 2004 (4–5 year old) to 2012 at age 12–13. With different age points, we could reveal the linkage between obesity and academic performance across different ages during childhood and adolescence, which has not been widely observed. We explore the effects of child and adolescent obesity or overweight on a few of educational outcome variables from test results in the five areas of reading, writing, spelling, grammar and numeracy from the National Assessment Program — Literacy and Numeracy (NAPLAN) in Australia. The test scores are objective reflection of cognitive capabilities. As the non-cognitive outcomes are valued by households and teachers, there exist some biases for the real reflection of children's skills, which we do not consider in this study.

A child's weight status can be endogenous, as it is related to socio-economic conditions and individual's social and physical surroundings. These contributors are also determinants of the child's academic performance. Therefore, it is complicated to identify the causal effect of child's weight status on educational outcomes. Economists have used different methods to empirically identify the causal relationship between health status and academic performance. Glewwe et al. (2001) use an achievement production function to examine the impact of the early childhood nutrition on academic achievement by focusing on sibling pairs of school age. Ding et al. (2009) use a set of genetic markers across individuals to examine the influence of health conditions on academic performance during adolescence. Zhao et al. (2012) use instrumental variable estimation to test the effect of youth smoking

[1] The reverse causality has been tested by Webbink et al. (2010)

on schooling. On the other side, Silles (2009) provides evidence of the causal effect of education on health status using TSLS estimates.

Based on methodologies from previous studies, we employ an instrumental variable approach to identify the impact of child obesity or overweight on academic performance. We instrument the endogenous children's weight variables with BMI of their biological parents. The results support for a negative impact of obesity or overweight on academic performance. Furthermore, the marginal effects from instrumental variable (IV) estimation are much larger than ordinary least square (OLS) estimation. When estimated by separate year groups, we find that the academic impact of obesity or overweight especially on numeracy is generally larger for senior year students.

To our knowledge, this is the first econometric study of the impact of weight status on academic performance for Australian children. By using the LSAC data with extensive information of children and households in Australia, we are able to explore the relationship of between children's educational outcomes and other variables. With this study, we could compare with aforementioned similar studies in other countries, which will provide some insights into how to improve children's school performance from different perspectives.

The rest of the paper is organized as follows. Section 2 presents a model specification for econometric estimation. Section 3 discusses the data and background in child obesity and overweight in Australia. Section 4 reports and compares results from different estimations in detail. Conclusion is provided in Section 5.

2 Model specification

2.1 Empirical strategy

This paper aims to identify empirically the causal impact of weight status on educational outcomes for children and adolescents. First of all, we present a simple dynamic framework from weight status to education for empirical analysis.

From Shi and Dzhumashev (2015), the optimal level of educational inputs is a function of health status. Let weight status (obesity or overweight) proxy the status of health, we have

$$e_t = g(W_t, X_t) \tag{1}$$

where W_t is the weight status and X_t denotes a vector of covariates that influence educational inputs such as parents' preferences and education level, socio-economic status of family, etc.

Educational achievement at time $t+1$, E_{t+1}, depends on the educational input at time t, e_t, and educational achievement at the beginning of period t, E_t.

$$E_{t+1} = \theta f(e_t, E_t) \tag{2}$$

Where θ represents the learning ability of individual children. Due to the heterogeneity of individuals' ability, it is possible to obtain different achievements when spending the same amount on a child's education. Hence the learning ability parameter θ is endogenous and captures the effects of all other factors such as innate gift, parents' education level and lifestyle on a child's ability. According to previous empirical studies (Datar et al., 2004; Wendt and Kinsey, 2009; Gable et al., 2012; Yau et al., 2012; etc.) we can state $d\theta/dW < 0$. But we assume that the individual is unaware of this negative impact of obesity on learning, similar to the assumption in Zhao et al. (2012). The assumption is plausible as child obesity or overweight is not paid as much attention as other diseases. Particularly in some countries, overweight or obesity is regarded as a sign of wealth and success.

It is noted from (1) and (2) that weight status does not explicitly appear as an input in education production function. It affects educational outcome through different channels. First, it may affect the physical energy level of a child which determines the learning productivity (Kantomaa et al., 2013). Second, it affects children's mental status that may have a direct impact on their academic performance. Third, discrimination or isolation against overweight children from peer effects shapes the unhealthy learning environment for a child. The first three channels are through direct impact on learning ability. However, the effect of obesity on educational outcomes may arise indirectly from reduced demand for educational input. Clinical studies have found that obese or overweight adolescents are more likely to develop weight-related illnesses and early death in adulthood (Dietitians Association of Australia). Parents with obese children may expect lower returns to education and thus may decrease the educational investment.

Combine (1) and (2), we have

$$E_{t+1} = \theta f(W_t, X_t, E_t) \tag{3}$$

The relationship derived in (3) implies that educational achievement is affected by child's weight status, educational input impact factors and educational achievement in the previous period.

On the other hand, the current weight status W_t is determined by parents' decision on health investment which is related to socioeconomic status and other spending including child bearing cost and educational investment which may crowd out health investment. And educational investment is often made by parents based on children's educational achievement, so weight status is affected by educational achievement E_t. Furthermore, at the beginning of period of t, children or adolescents choose whether or not to engage in behaviours leading to obesity such as consumption of fast food and less physical activity. This decision is also affected by their education level and social and physical surroundings(Webbink et al., 2010). As a result, the weight status can be expressed as a function of educational achievement and a vector of impact factors.

$$W_t = h(E_t, X_t) \tag{4}$$

The relationships reflected in (3) and (4) make it clear that weight status and educational outcomes are endogenous. Due to endogeneity of weight status variable, the academic performance may be correlated to other impact factors of weight status such as parents' socio-economic status and preference. These impacts may be entangled with the aforementioned channels. Ideally, we would like to specifically identify the effect of obesity on academic performance by distinguishing this from other impact factors responsible for obesity.

2.2 Estimation equations

Since θ is a function of W_t, we can rewrite (3) and (4) as

$$E_{t+1} = \theta(W_t, X_t, E_t) \tag{5}$$

Linearizing educational achievement function (5) yields:

$$E_{it} = \alpha_0 + \alpha_1 X_{it} + \alpha_2 W_{it} + \varepsilon_{it} \tag{6}$$

where the vector X contains individual and family characteristics (gender, race, income, etc.) that may influence learning outcomes and educational investment. The endogeneity of the weight variable makes W_t likely to be correlated with ε_{it} due to unobserved variables. For example, the supply of fast food and the availability of fitness facilities may lead to overweight or obesity of a child. Therefore, OLS estimates may suffer from a bias because of measurement errors in the weight status variable. IV methods are used to estimate the structural equation (6) for causal effect of obesity on educational achievement (α_2). To find a qualified IV, the variable should influence weight status but not be correlated with unobservable factors affecting educational achievement.

In addition, the weight status variable may suffer from a censoring problem. We use BMI as the measurement of weight status. According to the American Obesity Association, 'overweight' is defined as BMI greater than 25 for adolescents while "obesity" is measured as BMI greater than 30. The American CDC does not use the term "obesity" for children but "overweight" with BMI above 30.

To control for heterogeneity in learning abilities, we include personal characteristics, parents' education level and household income as explanatory variables. For example, gender has played a role in learning abilities. Researchers have long agreed girls have superior language abilities to boys. Parents' education level may affect the cognitive achievement of children (Nghiem et al., 2015). Household income is an important indicator of economic resources. It influences the nutrition level acquired by children which contributes to their cognitive skills (Glewwe et al., 2001).

3 Data and background

3.1 Data source

Data in this study is sourced from the Longitudinal Study of Australian Children (LSAC). The LSAC is a major study of 10,000 children and families from all parts of Australia. It is the first comprehensive, national Australian data on children as they grow up. The main aim of the study is to analyse children's physical health and social, cognitive and emotional development and investigate how Australia's unique social and cultural environment impacts on child development. A dual cohort was designed including families with children born between 2003 and 2004 (Cohort *B*) and families with children born between 1999 and 2000 (Cohort *K*). Each cohort contains approximately 5000 children. The families have been interviewed every two years since 2004. There are currently 5 waves of data collection for two cohorts of children. In this study, we mainly focus on children of Cohort *K* because measures of academic performance are more widely available for this cohort in wave 3, 4 and 5. The age of studied children are between 8 and 13 years old.

The reason to use Australian children as the study sample is due to epidemic overweight and obesity among children and adolescents in Australia. According to estimates by Australian Bureau of Statistics, between 1995 and 2007–08 there was a significant increase in the proportion of children who were obese. The rate of obesity for boys aged 5–17 years doubled from 5% in 1995 to 10% in 2007–08. The increase in overweight rate occurred for girls aged 13–17 years, up from 12% in 1995 to 20% in 2007–08. Specifically, children living in the areas of greatest relative disadvantage had more than double the rate of obesity (28%) of children living in areas with the lowest relative disadvantage (13%). In addition, Australia is a multi-cultural society made up of immigrants from all over the world. The influences of race and culture on child obesity and hence educational outcomes are of particular interest.

3.2 Variable selection

3.2.1 Dependent variable

The main educational outcomes of interest are the academic performance measured by the NAPLAN results for children from each interviewed family. NAPLAN is an annual assessment for students in year 3, 5, 7 and 9. It is made up of tests in the five areas of reading, writing, spelling, grammar and numeracy. The NAPLAN results are available from 2008 to 2012. As the oldest students in current available LSAC data (wave 5) are at 12–13 years, we only use the NAPLAN results for year 3, 5 and 7 students to match the studied sample.

Table 1 provides some comparison of the educational outcomes of obese and non-obese students in the three year groups. For comparison, the test scores of each grade are standardized by the means and standard deviations. It can be seen that the mean test scores of obese students are generally lower than non-obese students, but the differences are only statistically significant for reading, spelling, grammar and numeracy, not for writing.

Table 1. Academic outcomes of obese and non-obese children

NAPLAN Test	Mean	Non-Obese		Obese		p-value
		Obs.	Mean (M_0)	Obs.	Mean (M_1)	($H_0 : M_0 = M_1$)
Reading	519.01	1577	519.95	82	501.03	0.096*
Writing	507.33	1577	508.11	82	492.20	0.115
Spelling	505.70	1577	506.75	82	485.57	0.049**
Grammar	523.91	1577	524.97	82	503.41	0.061*
Numeracy	514.83	1577	515.90	82	494.22	0.058*

*** p<0.01, ** p<0.05, * p<0.1

3.2.2 Control variables

The BMI can be obtained from the weight and height of a child in the LSAC dataset. The child's weight for Cohort B in wave 1 was obtained by calculating the difference between the weight of parent 1 (or another adult) with the child and the weight of the parent/other adult on their own. For all subsequent waves of Cohort B and Cohort K at all waves, the child's weight was measured directly. Height is measured for children aged 2 years and older. It was measured twice, and if the two measurements differ by 0.5 cm or more a third measurement was taken. The average of the two closest measures is included on the data. Australian standard definitions for measuring overweight and obesity for children and adolescents are different to adults. The definition of child obesity and overweight based on BMI varies according to age and gender. For example, an 11 year old boy with BMI equal to 21 would be considered as overweight while a 7 year old girl with BMI equal to 20.5 would be considered obese. We use the classification from the Australian Department of Health[2] to define overweight and obesity for children and adolescents.

There has been a long list of literature discussing the endogeneity of BMI (Cawley, 2004; Kline and Tobias, 2008; Trogdon et al., 2008; Cawley and Meyerhoefer, 2012). To test the endogeneity of BMI here we first perform a Hausman test (see Table A.1 in the Appendix). As the definitions of obesity and overweight are based on the values of BMI, we also test the endogeneity of two measures of weight status (obesity and overweght). The p-values from Hausman test in Table A.1 are all close to zero. This demonstrates that at 5% confidence level we reject the null hypothesis that the variables representing weight status (obesity, overweight, and BMI) are exogenous, and instrumental variables are required to consistently estimate the impact of weight status (obesity or overweight) on children's academic performance.

The socio-economic factors in terms of parents' education level and income are sourced directly from the LSAC. Parents' education levels are represented by the level of the highest qualification completed by mother and father which is divided into five levels: postgraduate degree; graduate diploma/certificate; bachelor degree; advanced diploma/diploma; certificate and lower. Parents' income is represented by weekly income earned by the mother and the father, which has been adjusted by inflation index. The household income is the sum of mother's and father's weekly income. As Australia is a multicultural society, ethnicity is an important factor to be considered. Due to the eating habit, children from Asian family background are less likely to be obese, and from the tradition of prioritising education, their parents may invest more in children's education and therefore higher scores they may achieve. We also consider the time parents spent on the development of their children, which is captured by average working hours per week. In addition, educational outcomes are greatly affected by types of school that a child attends. The LSAC data contains such information, dividing schools into three types: government or public school,

[2] http://www.health.gov.au/internet/main/publishing.nsf/content/health-pubhlth-strateg-hlthwt-obesity.htm

catholic school and independent school.

Aside from socio-economic differences in education and income, area of residence may also determine the access to educational resources, physical activities and healthy food options. Around the world, levels of childhood obesity have been rising for a number of reasons including a reduction in the amount of time spent on physical activity and a shift in diet towards increased intake of foods that are high in fat and sugars.

It is possible that the common household environments shared by parents and the child may affect the behaviours and hence academic performance of the responding child, we may consider parents' current weight status in the regression. Parents can be biological parents, adopted or step parents as long as they live with children in the same home. In the sample, 99% of children have biological parent (either mother or father) living in the home, and nearly 80% have both biological parents at home. With only a sample of children and biological parents, it is impossible to identify the environmental effect from the genetic inheritance. To solve this, we organize a subsample of households which children and parents are unrelated and conduct a falsification test to explore whether weight status of children are associated with their non-biological parents' weight, following the method in Kinge and Morris (2015). With a sample of children and their adopted or step parents ($N = 348$), we regress non-biological parents' BMI on unrelated children's BMI. The coefficient is equal to 0.06 with a large p-value (0.12) of its t-statistic. The insignificant impact of shared household environment on weight has been proved by a few studies such as Cawley and Meyerhoefer (2012) and Maes et al. (1997). Therefore, we do not include the household environment impact in the model.

3.2.3 Instrumental variables

The choice of valid instruments is followed from previous studies. Some studies on endogeneity of obesity consider the use of peer effect as instrument. Cawley (2004) uses siblings' body weight as an instrument for sample persons' body weight to estimate the effect of obesity on the wage rate. Christakis and Fowler (2007) also find that a person's chances of becoming obese increased by 57% if he or she had a friend who became obese in a given interval. Asirvatham et al. (2014) use the weight of peers within the same grade as the instrument for childhood obesity as they find that a typical student's BMI z-score increases when facing heavier peers. In addition, environment influences attitudes and behaviours to food intake and exercise in the local population, area BMI is therefore expected to be a non-weak predictor of individual level BMI (James, 1995). Based on this, Morris (2006) uses area based measures to instrument BMI, which are respectively mean BMI across individuals and the prevalence of obesity in the health authority in which the respondent lives. Recent studies have used genetic markers for health status including obesity as instruments. Norton and Han (2008) argue that genetic information from specific genes linked to obesity provides strong exogenous variation in the BMI, and thus can be used as instrumental variables. Ding et al. (2009) exploit variation within specific genetic markers to identify the impact of poor health on academic performance. In case of unavailability of biomarkers data, Kinge and Morris (2015) employ BMI of biological parents as instrument for childhood obesity.

Based on previous studies and availability of data, we use biological parents' BMI (both mother's BMI and father's BMI) as instruments for a child's weight status. The instruments are constructed by matching parents' individual BMI in the LSAC sample with each child. It includes some genetic information even though parent and child do not live together. Comuzzie and Allison (1998) also believe the weight of a biological relative is a powerful instrument because roughly half of the variation in weight across people is of genetic origin. The instrument for obesity or overweight using the BMI of a biological relative has been tested in Norton and Han (2008) which use siblings' BMI as a valid instrument for an individual's obesity to study the impact of childhood obesity on health service utilization in UK.

When mother's BMI is missing, father's BMI can be used as the instrument, as some researchers have found that the gender of parents have no impact on genetic transmission. For example, Sørensen et al. (1992) find that the average correlation of adoptees with biological mothers was 0.17 and with biological fathers, 0.16. However, Lindeboom et al. (2010) have found that the variation in the association between parental BMI and child's BMI vary between mothers and fathers. Based on two different findings, we will conduct tests for the validity of the chosen instruments. The test results are presented in Table A.1 in the Appendix.

First, since we have two instrumental variables but only one endogenous variable, we conduct a Sargan (1958) test for over-identification. The p-values in Table A.1 are all greater than 0.10, which indicates that at 10% level we fail to reject the null hypothesis that there exist over-identification problem for the chosen instruments and we conclude that the two instrumental variables are valid.

Second, an F-test is run to examine whether the chosen instrumental variables are highly correlated with the endogenous variable conditional on other variables in the model. In the first stage regression, we conduct an F-test on all instruments to check if they are jointly significant in affecting the endogenous variable. Staiger and Stock (1997) suggest that the F-statistic of instrumental variables should be larger than 10 to ensure that the maximum bias in IV estimators to be less than 10%. From Table A.1, the F-statistics are all greater than 30, so the instrumental variables are not weak. To further investigate whether the association between parental BMI and child's BMI vary across between mothers and fathers, we also include control variables in both the first and second stage regressions for whether or not the instrument is significantly based on the father or the mother or both. The results are shown in Section 4.

3.2.4 Descriptive statistics of key variables

Following the above variable selection, the descriptive statistics of key variables in the analysis are presented in Table 2.

The age of children of LSAC sample used in the study ranges from 8 to 13 years old, which matches grade level from 3 to 7. The gender of studied children is evenly split between male and female, with 49% of girls and 51% of boys. On average, the BMI is 18.97, below the overweight level with BMI equals to 19.8 for an average 10 year old child (Australian standards). Of studied children, about 29% are overweight and nearly 8% are obese, in comparison with the national level that 25% of Australian children are overweight or obese (Australian Health Survey 2011–12, ABS). For parents in the survey, an average BMI higher than 25 implies the obesity epidemic of adults in the Australian society. It is very likely that the obesity of parents (biological) could be passed down to their children through genes.

Household income in this study is the sum of mother's and father's income. Some households are lone parent families with only one income. However, with a complete welfare system in Australia, most households with low income receive subsidy from government to raise children. According to the LSAC data, over 5% of households receive government

Table 2. Descriptive statistics of key variables

Variable	Definition	Obs.	Mean	Std. Dev.	Min	Max
reading	Reading test	10329	501.19	95.12	0	842
grammer	Grammar test	10326	506.63	98.03	0	839
spell	Spelling test	10329	491.88	88.99	0	751.9
write	Writing test	10313	489.84	84.18	89	807.2
num	Numeracy test	10294	496.92	90.74	0	922.8
age	Age	12451	10.27	1.76	7	13
female	= 1 if child's female	12456	0.49	0.5	0	1
cbmi	Child's BMI	12110	18.97	3.7	10.16	48.08
covrwt	= 1 if child is overweight	12105	0.29	0.45	0	1
cobsty	= 1 if child is obese	12105	0.08	0.27	0	1
fbmi	Father's BMI (biological)	7095	27.63	5.73	5.95	224.28
mbmi	Mother's BMI (biological)	10823	26.43	5.78	13.71	68.43
hincome	Household income	11932	1969.76	1516.94	0	25392.87
ethnity	= 1 if English speaking background	12234	0.92	0.28	0	1
metro	= 1 if live in metro city	12428	0.64	0.48	0	1
innerreg	= 1 if live in inner regional city	12428	0.21	0.41	0	1
hmeduc	= 1 if mother has postgraduate/graduate diploma	8949	0.22	0.42	0	1
lmeduc	= 1 if mother has certificate or lower	8949	0.4	0.49	0	1
hfeduc	= 1 if father has postgraduate/graduate diploma	8084	0.21	0.41	0	1
lfeduc	= 1 if father has certificate or lower	8084	0.47	0.5	0	1
fwhrs	Working hours of father	10022	46.57	13	0	168
mwhrs	Working hours of mother	9469	28.06	13.82	0	120
private	= 1 if study in private school	7336	0.34	0.48	0	1

payment in the sample. With the mean income closer to the minimum income level, we conclude that the distributions of households' income are right skewed with most values centerd in the left. Hence, we still use household income as a reliable indicator of household's economic resources though there exist large standard deviations.

In the sample, nearly 92% of children mainly speak english at home, which implies that ethnicity may not be a big concern in our analysis. For the area of residence, about 64% of studied households live in major cities and 21% live in inner regional cities, about 13% of children live in outer regional areas and only 2% of studied sample are from remote or very remote areas. An important impact factor is parents' education level. Mothers and fathers in the survey have exhibited similar education level, with about 22% of them having obtained postgraduate degree or graduate diploma. There is a higher percentage of mothers attaining bachelor degree or advanced diploma (37%) than fathers (32%). To contrast, nearly a half of interviewed fathers (47%) only obtain certificate or lower relative to 40% of mothers. The average working hours of mothers per week (28 hours) indicate the time spent on the development of children is much more than that of fathers (46.6 hours). This implies that fathers are main income earners and mothers may spend more time with their children. In terms of school type that a child attends, about 34% of children study in private schools including independent and catholic schools, and 66% of children attend public schools.

4 Results and discussions

4.1 Determinants of child weight status

The results of first stage regression are presented in Table 3. We report results for three measures of child weight status: (1) whether the child is obese; (2) whether the child is overweight; and (3) child's BMI. We estimate the coefficients of covariates that determinate child obesity and overweight in a probit model and BMI in a linear model. The regressions control for all available covariates and two instrumental variables — biological mother's BMI and biological father's BMI.

Comparing three regressions, we find that BMIs of biological mother and father are significantly associated with all measures of a child's weight status — obesity, overweight and BMI. The positive correlation implies that an increase in biological parents' BMI may increase the possibility of a child being obese or overweight and may also result in the rise of child's BMI. Since the estimation depends on the validity of chosen IVs, we conduct a likelihood ratio test for the explanatory power of the IVs. Under the null hypothesis that the chosen IVs have no explanatory power for child obesity or overweight, the likelihood ratio statistic has an asymptotic chi-square distribution with 2 excluded instruments. The F-statistic is calculated with likelihood ratio test divided by 2, which are 24.211 for child obesity and 41.006 for child overweight, both significant at the 5% level. So there are no weak instruments concerns.

In terms of other determinants of child obesity, we find that most coefficients of variables determining obesity are not significant, in part due to the small percentage of obese children (5%) among the whole sample. Among the few significant variables, ethnic background is a key impact factor for child obesity. Generally, children who speak English at home are less likely to be obese. This may be attributed to the fact that sports are put in a very important place in school's curriculum, it is likely that children with English speaking background are more involved in a variety of physical activities than children from

Table 3. First stage estimation of child's weight status

	Obesity, Probit	Overweight, Probit	BMI, Linear
Instrumental variables			
Father's BMI (biological)	0.019(0.008)**	0.020 (0.006)***	0.071 (0.013)***
Mother's BMI (biological)	0.057(0.009)***	0.052 (0.007)***	0.161 (0.014)***
Other explanatory variables			
Age	-0.522(0.422)	0.271(0.278)	0.513(0.555)
Age square	0.028(0.021)	-0.014(0.014)	0.004(0.028)
Gender (female=1)	0.034(0.108)	0.041(0.067)	0.176(0.135)
Household income	-0.011(0.041)	0.017(0.021)	0.077(0.042)*
English speaking background	-0.588(0.182)***	-0.303(0.125)**	-0.712(0.258)***
Metro city	0.026(0.175)	0.066(0.106)	0.014(0.215)
Inner regional city	0.511(0.185)***	0.176(0.120)	0.499(0.248)**
High education of mother	-0.012(0.146)	0.023(0.084)	0.060(0.167)
Low education of mother	0.069(0.128)	0.056(0.084)	0.073(0.172)
High education of father	-0.262(0.167)	-0.053(0.090)	-0.020(0.175)
Low education of father	0.155(0.127)	0.269(0.081)***	0.664(0.166)***
Working hours of father	0.000(0.004)	0.004(0.003)	0.008(0.006)
Working hours of mother	0.003(0.004)	0.006(0.002)**	0.012(0.005)**
School type (private =1)	-0.269(0.116)**	-0.039(0.069)	-0.240(0.139)*
No. of obs.	1830	1830	1797
Log likelihood	-317.587	-927.661	-
Weak instruments test[b]	48.422(0.000)	82.012(0.000)	-

a. Standard errors in parentheses

b. Log likelihood ratio tests against the explanatory power of excluded IVs.

*** p<0.01, ** p<0.05, * p<0.1

other ethnic backgrounds. The covariate — residing in inner regional cities is positively correlated with child obesity, significant at the 1% level. It is possible that children living in inner regional cities have less access to physical facilities compared to children in metro cities. On the other hand, shorter distance to schools makes children less active than those living in outer or remote areas. In addition, school type plays a significant role in determining the possibility of a child being obese at the 5% level. Children attending catholic or independent schools have lower risk of being obese. Compared to most public schools, private schools have more varieties of activities for students to get involved in.

When measured with child overweight, significant determinants include ethnicity, education level of father and weekly working hours of mother. Results have shown that only low education level of father has significantly positive impact on child overweight. As father is the main breadwinner in most families, his low education level may reflect low socio-economic status of the family. Longer working hours of mother per week increases the possibility of child being overweight, significant at the 10% level. Mother working longer hours may have less time and energy to focus on the diet and physical activities of her child. This reflect the importance of mother to the development and health status of a child to some extent. The regression with BMI results in all significant determinants included in the other two regressions except for household income. It is shown that household income has a positive impact on child's BMI, statistically significant at the 10% level. Higher household income increases child's BMI but not necessarily contributes to obesity with the negative coefficient of household income on child obesity. This is because wealthier families are more concerned about the nutrition level and consume more healthy food.

We will look at other variables affecting child weight status, though they are not significant in the model. First, age is an important indicator for weight status. It is shown that the possibility of obesity decreases with age, given other variables. This reflects the awareness of appearance or the negative effect of obesity when a child gets older and become more educated. However, the likelihood of overweight and value of BMI increases with age. According to the summary statistics of the sample, the overweight ratio increases by 26% from 8–9 year old (year 3) to 12–13 year old (year 7). Gender of children is a key impact factor. Results from all three regressions have shown that girls are more likely to be obese or overweight than boys for the same age. According to the Australian classification of overweight and obesity for children and adolescents, for 7–13 year old, BMI level equivalent to overweight and obesity for girls are generally higher than boys. It is seen from the survey sample that there are more girls with BMI over the standard level than boys. Household income is negatively associated with child obesity, but positively correlated with child overweight possibly due to the reason that overweigh is not considered by many parents as a potential health risk. To our surprise, the coefficients of household income are not significant for child obesity and overweight, probably due to large standard deviations in household income. Moreover, higher education level of mother and father (postgraduate or graduate diploma) has negative impact on child obesity, which reflects the fact that parents with high education are more aware of harmfulness of obesity. Both income and education level are important indicators of socio-economic status, and the inverse relationship between obesity and socio-economic status has been found by many studies (O'Dea et al., 2011; Wang and Lim, 2012; Bonnefond and Clément, 2014; O'Dea and Dibley, 2014). Combined with the fact that father's low education is positively associated with a child being overweight, the results indicate that improving parents' education level may have a preventative effect on child obesity or overweight. Unlike the significant impact of working hours of mother on child overweight, the positive but insignificant association between working hours of father and child's weight status implies father has less impact on the health status of a child than mother.

4.2 Child weight status and academic outcomes

To examine the impact of child weight status on academic performance, we compare regression results for two measures of child weight status — obesity and overweight from ordinary least square (OLS) and two stage least square (2SLS) estimations, controlling all covariates. The children's academic performances are measured by five subjects: reading, grammar, spelling, writing and numeracy. We present estimates of cross-sectional data[3] for combined grades (year 3, 5 and 7) and separate estimates of each grade in Table 4.

From OLS regression, the combined results of three grades show that when controlling other explanatory variables such as the child's age, gender, household income and parents' education, etc., child obesity has significant impact on academic outcomes, generally resulting in lower scores in all test subjects. For example, obese children, on average, have scored 16 points lower in reading, 17 points in grammar, 17 points in spelling, 24 points in writing and 20 points in numeracy than non-obese children. These results are similar to previous findings (Datar et al., 2004; Wendt and Kinsey, 2009; Yau et al., 2012) in other countries.

In particular, when considering the separate grade, the academic impacts of obesity are different. For grade 3 and grade 7 students, estimation has shown that obesity has significantly negative impact on NAPLAN tests except on reading. For grade 5 students, child obesity significantly results in lower scores in all five subjects.

Estimates of the effect of child overweight on academic outcomes are shown in Table 5.

[3] With the null hypothesis that the time-invariant fixed effects such as gender of child, education level of parents, area of residence, school type are correlated with other explanatory variables including household income and weight status, a Hausman test cannot reject it at the 5% significance level. This implies that fixed effect estimates are appropriate in this panel data. However, a problem with the panel estimation is that only three time points (year 3, 5 and 7) are available for academic outcomes, resulting in large standard errors for most variables, hence we do not attempt a fixed effects approach.

Table 4. Effect of child obesity on academic performance by grade – OLS

Variables	Reading	Grammar	Spelling	Writing	Numeracy
Combined grades					
Obesity	-16.05(6.475)**	-17.33(6.493)***	-16.60(5.569)***	-24.38(5.612)***	-19.62(5.758)***
Observations	2600	2600	2601	2596	2595
R-squared	0.303	0.294	0.320	0.307	0.344
Grade 3					
Obesity	-14.37(10.18)	-20.75(10.65)*	-27.08(9.173)***	-22.97(8.993)**	-20.58(8.879)**
Observations	1079	1079	1080	1078	1079
R-squared	0.111	0.101	0.089	0.093	0.085
Grade 5					
Obesity	-19.92(9.300)**	-19.32(8.978)**	-13.07(7.547)*	-30.35(7.693)***	-24.24(8.205)***
Observations	1310	1309	1309	1307	1304
R-squared	0.111	0.119	0.094	0.121	0.104
Grade 7					
Obesity	-10.50(7.655)	-19.25(8.143)**	-15.26(7.377)**	-14.56(7.808)*	-23.82(8.349)***
Observations	1290	1290	1290	1289	1284
R-squared	0.116	0.138	0.094	0.144	0.112

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

With overweight as the measurement of weight status, the impacts on five subject scores for three grades students are significant. The marginal effects imply that the fatter (or higher BMI) the child, the lower scores he or she achieved in all five subjects. The estimates by year group, however, are different from obesity estimates in Table 4. For grade 3 students, the impact of overweight on test scores of all subjects are insignificant except on writing. This may be due to the fact that overweight is not a big risk for younger children. For year 5 and year 7 students, overweight turns to be negatively associated with academic outcomes in all five subjects, significant at 1% to 10% level. The estimates for separate year group imply that the overweight measure has a greater impact on senior year group than junior group.

Detailed estimates of other determinants of academic outcomes from OLS regressions are presented in Table A.2 and Table A.3 in the Appendix. We will briefly discuss the results. First, age is a significant determinant for test results for all three grades. With the variation of age in the same grade (i.e. grade 3 includes students between 8 and 9 year old), the older students generally acquire better cognitive outcomes. In terms of gender, girls have more advantage in reading, grammar, spelling and writing, while boys are better in numeracy, which reflects a general phenomenon of gender difference in learning capabilities.

Table 5. Effect of child overweight on academic performance by grade — OLS

Variables	Reading	Grammar	Spelling	Writing	Numeracy
Combined grades					
Overweight	-9.057(3.583)**	-6.674(3.595)*	-7.979(3.083)***	-12.54(3.109)***	-11.50(3.186)***
Observations	2600	2600	2601	2596	2595
R-squared	0.303	0.293	0.320	0.306	0.344
Grade 3					
Overweight	-8.038(5.683)	-1.251(5.945)	-6.813(5.129)	-11.11(5.006)**	-7.503(4.951)
Observations	1079	1079	1080	1078	1079
R-squared	0.113	0.102	0.087	0.100	0.089
Grade 5					
Overweight	-10.29(5.045)**	-11.33(4.868)**	-10.29(4.086)**	-13.43(4.186)***	-14.77(4.450)***
Observations	1310	1309	1309	1307	1304
R-squared	0.110	0.120	0.097	0.119	0.106
Grade 7					
Overweight	-7.158(4.236)*	-14.47(4.547)***	-13.26(4.117)***	-12.76(4.364)***	-21.89(4.583)***
Observations	1290	1290	1290	1289	1284
R-squared	0.117	0.142	0.098	0.147	0.122

a. Standard errors in parentheses
b. *** p<0.01, ** p<0.05, * p<0.1

The adjusted household income — an important indicator of social-economic status has a significant positive impact on all test scores. Children speaking English at home, mostly Australian children or immigrants from English-speaking countries achieve significantly lower scores in all tests except reading, compared with non-English background children. It is true that immigrants from east asian countries pay more attention to education. Remoteness of residence is important for higher academic performance. Children living in major cities or inner regional cities with most excellent educational resources centered generally achieve higher scores in NAPLAN tests except writing than children living in remote areas. In addition, parents' education level is a key determinant of child's academic performance. In particular, mother or father with an education level of certificate or below has negative effect on child's test scores for all subjects. However, parents with high education level (postgraduate) do not significantly correspond to higher academic achievement of their children. In regard to parental time input on children, we find that mothers who work longer hours per week have significantly led to lower test scores in reading, grammar and writing. However, the impact of fathers' working hours on their children's test scores is not statistically significant.

With the endogenous weight status variable, the IV estimates for two measures (obesity and overweight) are shown in Table 6 and 7. BMIs of biological mother and father are employed as the instruments. All control variables used in OLS regressions are included in IV regressions for comparison.

Table 6. Effect of child obesity on academic performance by grade — IV

Variables	Reading	Grammar	Spelling	Writing	Numeracy
Combined grades					
Obesity	-134.9(47.90)***	-160.4(48.66)***	-153.0(41.95)***	-162.5(42.82)***	-157.8(44.00)***
Observations	1797	1797	1798	1795	1792
DWH test[b]	0.00395	0.00061	0.00009	0.00020	0.00027
Sargan test[c]	0.7740	0.4942	0.9843	0.8945	0.5159
Grade 3					
Obesity	-194.7(85.32)**	-180.5(89.27)**	-172.3(75.13)**	-153.1(73.51)**	-111.6(69.70)
Observations	730	730	731	730	729
DWH test	0.01478	0.02548	0.01657	0.04182	0.13743
Sargan test	0.7885	0.8326	0.5058	0.5740	0.467
Grade 5					
Obesity	-83.46(57.06)	-114.7(55.18)**	-109.2(47.00)**	-142.6(49.08)***	-148.2(53.39)***
Observations	916	916	916	915	912
DWH test	0.13332	0.05032	0.01325	0.00751	0.00602
Sargan test	0.7372	0.8418	0.9084	0.9636	0.4007
Grade 7					
Obesity	-141.1(48.70)***	-163.2(48.66)***	-138.4(44.02)***	-135.1(47.74)***	-199.7(53.49)***
Observations	921	918	918	917	916
DWH test	0.00414	0.00104	0.00241	0.00237	0.00005
Sargan test	0.1876	0.3428	0.3277	0.1177	0.2860

a. Standard errors in parentheses
b. Endogeneity test with the null hypothesis of exogenous regressor
c. Overidentification tests are obtained by assuming the first stage estimation as linear, large p-value represents the rejection of null hypothesis of overidentification of two instruments
d. *** p<0.01, ** p<0.05, * p<0.1

The regression results are in line with our expectations. With BMIs of biological parents as instruments, child obesity is negatively related to academic results, which are statistically significant at 1% level for all five subjects. As linear regression is used in the first stage regression, the coefficient of the weight status variable in the IV estimation represents its marginal effect on the dependent variable. It can be seen that obese children, on average, have scored 135 points lower in reading, 160 points in grammar, 153 points in spelling, 163 points in writing and 158 points in numeracy than non-obese children. The impact is much larger than that in the OLS estimation, implying that BMI of biological parents have indirect impact on child's academic results. In addition, the impact of obesity is identified from other impact factors such as socio-economic status of family and social and physical surroundings in IV estimation while weight status variable is entangled with these impacts in OLS resulting in underestimated coefficients. For example, higher household income which produces more educational investment may counteract the negative effect of obesity on child's cognitive development.

Looking at the separate year group regressions, for grade 3, obesity has negative effect on four subjects: reading, grammar, spelling and writing, all at the 5% significance level. However, the impact of obesity on numeracy performance is not significant, which is reflected by the large p-value of Durbin-Wu-Hausman (DWH) test. The OLS estimation is more suitable in this situation. Similarly, for grade 5 regression, obesity has significantly reduced the test scores of subjects except reading. The large p-value acquired from DWH test cannot reject the null hypothesis so endogeneity cannot be identified. Lastly, it is obvious that obesity has significantly negative impact on scores of all five subjects for grade 7 group, at the 1% level. Compared to results from OLS estimation, obesity has no significant impact on numeracy for grade 3 students and on reading for grade 5 students using IV estimation while reading scores for grade 3 and 7 tests are not significantly correlated with obesity in OLS regression.

The impacts of another measure — child overweight on academic performance in the NAPLAN test using the IV estimation are presented in Table 7. Similar to obesity, overweight is negatively associated with test results for all five subjects, statistically significant at the 1% level. The small p-value from DWH test and large p-value from Sargan test have verified that biological parents' BMIs are suitable instruments for the endogenous overweight variable. As biological parents' BMIs indirectly influence child's academic performance through the impact on child overweight, the marginal effects of overweight alone on test scores here are much larger than that in the OLS estimation.

Regressions with separate year group provide further information of the effect of child weight status on educational outcomes. For example, overweight has shown significantly negative effect on grade 3 students' test scores. In particular, overweight results in a lower score in numeracy, at the 10% significance level while obesity does not have significant impact on numeracy for grade 3 students. Compared to OLS estimation where the results for grade 3 students are not significant, the IV regression is more appropriate in this case due to the fact that the endogenous weight status variable is correlated with error term resulting in biased estimation. For year 5 students, obesity has negative association with all subject performance except reading, which is similar to the measurement of obesity. Moreover, for grade 7 students, the associations between overweight and academic performance are significant at the 1% level. From results of separate year group with both IV and OLS estimation, we have found that weight status has larger negative impact on academic performance especially in numeracy for senior year group students. This may be attributed to the fact that obesity or overweight poses higher mental and physical risk to older children or adolescents, leading to a poorer academic performance compared to younger children.

Table 7. Effect of child overweight on academic performance by grade — IV

Variables	Reading	Grammar	Spelling	Writing	Numeracy
Combined grades					
Overweight	-58.31(20.12)***	-69.14(20.47)***	-66.27(17.53)***	-70.69(17.90)***	-67.94(18.30)***
Observations	1797	1797	1798	1795	1792
DWH test	0.01529	0.00120	0.00016	0.00056	0.00098
Sargan test	0.7049	0.4239	0.8976	0.9902	0.4317
Grade 3					
Overweight	-78.93(34.76)**	-73.55(37.46)**	-76.20(32.11)**	-67.51(30.65)**	-50.95(29.73)*
Observations	730	730	731	730	729
DWH test	0.03618	0.02155	0.00500	0.05011	0.08109
Sargan test	0.4312	0.5141	0.8749	0.9174	0.7119
Grade 5					
Overweight	-37.61(25.99)	-53.20(24.88)**	-49.91(20.89)**	-65.74(22.27)***	-70.05(23.75)***
Observations	916	916	916	915	912
DWH test	0.29403	0.11408	0.05631	0.01167	0.01790
Sargan test	0.6249	0.9896	0.7158	0.7405	0.5602
Grade 7					
Overweight	-70.87(23.04)***	-85.17(24.25)***	-72.44(21.89)***	-71.49(23.41)***	-103.6(26.02)***
Observations	921	918	918	917	916
DWH test	0.00530	0.00181	0.00507	0.00661	0.00021
Sargan test	0.2907	0.5205	0.4871	0.1787	0.4941

a. Standard errors in parentheses
b. Endogeneity test with the null hypothesis of exogenous regressor
c. Overidentification tests are obtained by assuming the first stage estimation as linear, large p-value represents the rejection of null hypothesis of overidentification of two instruments
*** p<0.01, ** p<0.05, * p<0.1

IV estimations for other covariates of three time points are presented in Table A.4 and Table A.5 in the Appendix. The impacts of age, gender, household income and English background on test scores are similar to OLS estimation but with different magnitude. Remoteness of residence is negatively correlated with child's academic outcomes, which implies that living in metro cities or inner regional areas are helpful for a student's educational outcomes due to the availability of educational resources. Different from OLS estimation, living in metro cities will not significantly improve writing skills. In terms of parents' education level, it is surprising to find that whether using OLS or IV estimation parents with certificate or lower is negatively correlated to child's test scores but higher education level of parents does not necessarily result in higher academic performance of their children, especially father's higher education. However, different from OLS results, mothers who obtained postgraduate degrees could help their children achieve better not only in reading but also in grammar. Regarding to time spent on children, IV estimation with child overweight as the measurement makes longer working hours of fathers enhance child's academic performance on numeracy scores significant at the 10% level. A possible explanation lies in the fact that relative to other subjects numeracy needs more practice and tutoring. Father may bring more income by working harder and hence are able to invest more on child's education. In contrast, the longer hours a mother works each week, the lower scores her child achieves. This implies that mother's time input on child is more important for child's educational outcomes. It is surprising to find that attending private schools does not really contribute to a better academic performance. With OLS estimation, it only improves performance in reading and writing. Using IV estimation with obesity, it results in poor performance in all subjects. This result for children of 8-13 year old is consistent with the finding in Nghiem et al. (2015) that Catholic or independent primary school attendance does not significantly result in better cognitive outcomes.

5 Conclusion

This study has examined the impact of child weight status with the measures of obesity and overweight on academic performance. Using the LSAC dataset containing extensive information of Australian children and households, the study has shown that obesity or overweight has negative impact on academic results reflected by NAPLAN test. For example, an obese student tends to score lower with 16 points in reading, 17 points in grammar, 17 points in spelling, 24 points in writing and 20 points in numeracy than a non-obese student with the OLS estimation. After accounting for endogeneity of weight status, the impacts on test scores in IV estimation for all five subjects are about 8 times larger than

OLS estimators. This large difference could be attributed to the reason that BMIs of biological parents have further impacted on a child's academic results through indirectly affecting the endogenous weight variables. Another reason is that the impact of obesity is identified from other impact factors such as socio-economic status of family, social, and physical surroundings which are entangled with weight status variable leading to underestimate of the OLS estimators. Moreover, the impact on academic performance is different for each grade. From IV estimation results of three separate grades (year 3, 5 and 7 students), the impact on reading test is not significant for grade 5 students. In particular, we have found that the negative impact of obesity or overweight on NAPLAN tests especially the numeracy scores is greater for senior year students. A possible reason is that obesity or overweight adversely affect older children or adolescents mentally and physically, leading to a poorer academic performance relative to younger children. The difference in the academic impact of weight status between primary school and high school is of importance to the literature and of course to the interested parties.

Our results are consistent with previous studies in other countries (Datar et al., 2004; Wendt and Kinsey, 2009; Yau et al., 2012). It is the first study in Australia so far and will contribute to the existing empirical literature. More importantly, we have found appropriate instrumental variables for the endogenous regressor. Instead of using genetic markers of health status which are usually unavailable, we use biological parents' BMIs as instruments for a child's weight status and identify the environmental effect from genetic transmission of obesity to further support the validity of our chosen IVs.

Due to the limitation of LSAC data, we only have three waves of surveys making the fixed effect approach not appropriate. It is noted that a cross-sectional data used in this study cannot explain the sequential effect. More work with panel data could be done as further waves of data are available. In addition, the definition of child obesity and overweight follows the Australian standards and changes with age and gender. As a result, the prevalence of obesity or overweight is relatively higher than that using other standards such as the one from the American Obesity Association, which may overestimate the impact of child obesity or overweight on academic performance.

This study has important policy implications. We have known well about the inverse relationship between obesity and socio-economic status existing in many countries (see Offer et al., 2010; O'Dea et al., 2011; Wang and Lim, 2012; O'Dea and Dibley, 2014; Xiao et al., 2013; Bonnefond and Clément, 2014; Ma et al., 2014). However, except some medical risks, the consequences of obesity in other areas such as health care cost, labor market or human capital formation are not well studied. Investigation of the impact of child weight status on academic performance will provide some insights to identify policy opportunities for improving educational outcomes for children and adolescents.

Compliance with Ethical Standards

This study was funded by Victoria University Central Research Grant Scheme (grant number CRGS 15/15). The authors declare that they have no conflict of interest.

Appendix

Table A.1. Effect of child's weight status on academic performance by grade — IV

	Naplan Results	Endogeneity test Hausman statistic (p-value)	Over-identification test Sargan statistic (p-value)	F-statistic (p-value)
obesity	Read	14.29 (0.000)	0.76 (0.384)	36.61 (0.000)
	Grammar	14.25 (0.000)	0.24 (0.627)	36.61 (0.000)
	Spelling	14.04 (0.000)	1.16 (0.282)	36.61 (0.000)
	Write	13.72 (0.000)	2.34 (0.126)	36.61 (0.000)
	Num	17.65 (0.000)	1.08 (0.300)	36.61 (0.000)
overweight	Read	13.99 (0.000)	1.09 (0.296)	75.61 (0.000)
	Grammar	14.22 (0.000)	0.25 (0.618)	75.61 (0.000)
	Spelling	22.45 (0.000)	1.31 (0.252)	75.61 (0.000)
	Write	22.05 (0.000)	2.38 (0.123)	75.61 (0.000)
	Num	22.04 (0.000)	0.33 (0.564)	75.61 (0.000)
BMI	Read	12.56 (0.000)	0.17 (0.680)	111.67 (0.000)
	Grammar	11.41 (0.000)	0.00 (0.995)	111.67 (0.000)
	Spelling	12.12 (0.000)	0.37 (0.543)	111.67 (0.000)
	Write	12.30 (0.000)	1.27 (0.259)	111.67 (0.000)
	Num	14.10 (0.000)	0.29 (0.588)	111.67 (0.000)

Table A.2. Effect of child obesity on academic performance (grade 3, 5 and 7 combined) — OLS

VARIABLES	Reading	Grammar	Spelling	Writing	Numeracy
Obesity	-16.05**	-17.33***	-16.60***	-24.38***	-19.62***
	(-6.475)	(-6.493)	(-5.569)	(-5.612)	(-5.758)
Age	102.9***	121.0***	74.36***	53.90***	90.68***
	(-12.62)	(-12.64)	(-10.84)	(-10.94)	(-11.21)
Age square	-3.713***	-4.657***	-2.320***	-1.378**	-3.017***
	(-0.631)	(-0.632)	(-0.542)	(-0.547)	(-0.56)
Gender (female=1)	16.28***	20.35***	12.48***	20.71***	-7.349***
	(-3.051)	(-3.059)	(-2.624)	(-2.646)	(-2.715)
Household income	4.166***	5.015***	3.554***	3.216***	3.723***
	(-0.964)	(-0.966)	(-0.829)	(-0.836)	(-0.858)
English speaking background	-5.086	-14.71**	-27.25***	-16.67***	-11.59**
	(-5.777)	(-5.794)	(-4.97)	(-5.008)	(-5.15)
metro city	12.23***	16.49***	15.50***	11.25***	16.34***
	(-4.679)	(-4.703)	(-4.035)	(-4.066)	(-4.165)
inner regional city	11.10**	12.14**	11.87**	0.0935	10.72**
	(-5.407)	(-5.425)	(-4.653)	(-4.694)	(-4.814)
High education of mother	8.998**	6.29	1.889	3.128	3.367
	(-3.906)	(-3.918)	(-3.361)	(-3.388)	(-3.475)
Low education of mother	-21.24***	-19.63***	-10.95***	-16.02***	-16.52***
	(-3.765)	(-3.773)	(-3.235)	(-3.266)	(-3.347)
High education of father	0.367	4.832	5.354	3.852	3.434
	(-4.128)	(-4.142)	(-3.551)	(-3.583)	(-3.673)
Low education of father	-26.72***	-18.82***	-15.26***	-14.62***	-19.41***
	(-3.727)	(-3.736)	(-3.206)	(-3.232)	(-3.316)
Working hours of father	-0.138	-0.113	-0.0701	0.012	0.0688
	(-0.125)	(-0.125)	(-0.108)	(-0.109)	(-0.111)
Working hours of mother	-0.339***	-0.477***	-0.318***	-0.391***	-0.121
	(-0.113)	(-0.114)	(-0.0975)	(-0.0983)	(-0.101)
School type (private=1)	1.966	-5.706*	-5.667**	1.637	-0.546
	(-3.146)	(-3.154)	(-2.705)	(-2.728)	(-2.799)
Constant	-141.4**	-216.3***	-1.566	105.5*	-97.81*
	(-63.13)	(-63.29)	(-54.27)	(-54.78)	(-56.17)
Observations	2,600	2,600	2,601	2,596	2,595
R-squared	0.303	0.294	0.32	0.307	0.344

Table A.3. Effect of child overweight on academic performance (grade 3, 5 and 7 combined) — OLS

VARIABLES	Reading	Grammar	Spelling	Writing	Numeracy
Overweight	-9.057**	-6.674*	-7.979***	-12.54***	-11.50***
	(-3.583)	(-3.595)	(-3.083)	(-3.109)	(-3.186)
Age	104.2***	122.1***	75.59***	55.75***	92.35***
	(-12.62)	(-12.65)	(-10.85)	(-10.95)	(-11.21)
Age square	-3.777***	-4.713***	-2.382***	-1.471***	-3.100***
	(-0.631)	(-0.632)	(-0.542)	(-0.547)	(-0.56)
Gender (female=1)	16.25***	20.28***	12.44***	20.65***	-7.379***
	(-3.05)	(-3.061)	(-2.625)	(-2.647)	(-2.714)
Household income	4.176***	5.032***	3.566***	3.231***	3.734***
	(-0.964)	(-0.967)	(-0.829)	(-0.836)	(-0.858)
English speaking background	-5.072	-14.45**	-27.12***	-16.56***	-11.61**
	(-5.777)	(-5.798)	(-4.972)	(-5.01)	(-5.148)
metro city	12.52***	16.69***	15.74***	11.62***	16.70***
	(-4.68)	(-4.708)	(-4.038)	(-4.069)	(-4.165)
inner regional city	11.13**	11.94**	11.78**	0.0296	10.77**
	(-5.407)	(-5.429)	(-4.654)	(-4.696)	(-4.813)
High education of mother	8.863**	6.18	1.768	2.939	3.198
	(-3.906)	(-3.921)	(-3.362)	(-3.39)	(-3.475)
Low education of mother	-21.35***	-19.86***	-11.12***	-16.25***	-16.64***
	(-3.763)	(-3.774)	(-3.235)	(-3.265)	(-3.344)
High education of father	0.487	4.982	5.487	4.057	3.574
	(-4.128)	(-4.145)	(-3.553)	(-3.584)	(-3.672)
Low education of father	-26.45***	-18.82***	-15.12***	-14.32***	-19.05***
	(-3.734)	(-3.745)	(-3.213)	(-3.241)	(-3.322)
Working hours of father	-0.129	-0.106	-0.0613	0.0258	0.0812
	(-0.125)	(-0.126)	(-0.108)	(-0.109)	(-0.111)
Working hours of mother	-0.326***	-0.469***	-0.308***	-0.374***	-0.105
	(-0.114)	(-0.114)	(-0.0977)	(-0.0986)	(-0.101)
School type (private=1)	2.136	-5.566*	-5.510**	1.883	-0.331
	(-3.146)	(-3.157)	(-2.706)	(-2.73)	(-2.798)
Constant	-147.7**	-222.1***	-7.657	96.47*	-105.8*
	(-63.12)	(-63.33)	(-54.29)	(-54.81)	(-56.16)
Observations	2,600	2,600	2,601	2,596	2,595
R-squared	0.303	0.293	0.32	0.306	0.344

Table A.4. Effect of child obesity on academic performance (grade 3, 5 and 7 combined) — IV

VARIABLES	Reading	Grammar	Spelling	Writing	Numeracy
Obesity	-134.9***	-160.4***	-153.0***	-162.5***	-157.8***
	(-47.9)	(-48.66)	(-41.95)	(-42.82)	(-44)
Age	109.3***	122.0***	68.09***	54.88***	88.97***
	(-15.92)	(-16.21)	(-13.99)	(-14.31)	(-14.71)
Age square	-3.941***	-4.635***	-1.938***	-1.332*	-2.854***
	(-0.796)	(-0.811)	(-0.699)	(-0.716)	(-0.735)
Gender (female=1)	20.02***	21.09***	12.90***	21.19***	-6.233*
	(-3.863)	(-3.932)	(-3.393)	(-3.466)	(-3.568)
Household income	4.278***	4.035***	3.518***	2.511**	3.422***
	(-1.197)	(-1.217)	(-1.051)	(-1.074)	(-1.107)
English speaking background	-10.6	-18.55**	-32.51***	-21.97***	-15.13**
	(-7.829)	(-7.967)	(-6.876)	(-7.017)	(-7.251)
metro city	10.56*	17.35***	14.47***	7.752	15.73***
	(-6.132)	(-6.252)	(-5.397)	(-5.509)	(-5.668)
inner regional city	19.10**	23.00***	21.52***	3.23	18.67***
	(-7.605)	(-7.729)	(-6.668)	(-6.811)	(-7.017)
High education of mother	11.62**	8.444*	1.138	4.303	1.836
	(-4.78)	(-4.864)	(-4.199)	(-4.288)	(-4.417)
Low education of mother	-16.98***	-19.83***	-10.07**	-15.68***	-15.35***
	(-4.911)	(-5)	(-4.31)	(-4.408)	(-4.528)
High education of father	-2.304	1.759	1.846	1.22	0.739
	(-5.073)	(-5.165)	(-4.456)	(-4.554)	(-4.687)
Low education of father	-23.39***	-15.55***	-15.12***	-8.906**	-15.67***
	(-4.971)	(-5.053)	(-4.361)	(-4.452)	(-4.59)
Working hours of father	0.0111	0.148	0.152	0.0729	0.17
	(-0.163)	(-0.166)	(-0.143)	(-0.146)	(-0.15)
Working hours of mother	-0.311**	-0.391***	-0.174	-0.272**	-0.029
	(-0.143)	(-0.145)	(-0.125)	(-0.128)	(-0.132)
School type (private=1)	-0.953	-12.14***	-11.17***	-2.356	-5.425
	(-4.115)	(-4.19)	(-3.616)	(-3.694)	(-3.799)
Constant	-178.8**	-226.9***	24.37	104.8	-89.93
	(-79.93)	(-81.41)	(-70.22)	(-71.84)	(-73.87)
Observations	1797	1797	1798	1795	1792
R-squared	0.225	0.173	0.18	0.168	0.21

Table A.5. Effect of child overweight on academic performance (grade 3, 5 and 7 combined) — IV

VARIABLES	Reading	Grammar	Spelling	Writing	Numeracy
Overweight	-58.31***	-69.14***	-66.27***	-70.69***	-67.94***
	(-20.12)	(-20.47)	(-17.53)	(-17.9)	(-18.3)
Age	120.1***	134.9***	80.41***	67.80***	101.7***
	(-15.51)	(-15.8)	(-13.57)	(-13.83)	(-14.21)
Age square	-4.513***	-5.314***	-2.589***	-2.017***	-3.527***
	(-0.773)	(-0.788)	(-0.676)	(-0.69)	(-0.708)
Gender (female=1)	20.14***	21.22***	13.04***	21.27***	-6.131*
	(-3.742)	(-3.814)	(-3.274)	(-3.334)	(-3.428)
Household income	4.511***	4.297***	3.765***	2.774***	3.691***
	(-1.159)	(-1.18)	(-1.014)	(-1.033)	(-1.063)
English speaking background	-7.288	-14.57**	-28.74***	-18.02***	-11.24*
	(-7.272)	(-7.41)	(-6.361)	(-6.477)	(-6.68)
metro city	10.97*	17.83***	14.93***	8.173	16.21***
	(-5.943)	(-6.066)	(-5.209)	(-5.302)	(-5.447)
inner regional city	14.58**	17.55**	16.32***	-2.265	13.33**
	(-6.946)	(-7.071)	(-6.068)	(-6.18)	(-6.359)
High education of mother	12.59***	9.622**	2.258	5.494	2.997
	(-4.619)	(-4.706)	(-4.041)	(-4.115)	(-4.232)
Low education of mother	-16.29***	-18.96***	-9.269**	-14.94***	-14.56***
	(-4.791)	(-4.887)	(-4.187)	(-4.267)	(-4.379)
High education of father	-0.709	3.68	3.651	3.234	2.585
	(-4.856)	(-4.949)	(-4.247)	(-4.328)	(-4.448)
Low education of father	-21.28***	-13.09**	-12.74***	-6.232	-13.24***
	(-5.092)	(-5.179)	(-4.443)	(-4.538)	(-4.657)
Working hours of father	0.0723	0.223	0.225	0.151	0.243*
	(-0.159)	(-0.163)	(-0.14)	(-0.143)	(-0.146)
Working hours of mother	-0.242*	-0.308**	-0.0944	-0.186	0.0515
	(0.144)	(0.147)	(0.126)	(0.128)	(0.132)
School type (private=1)	1.697	-8.993**	-8.182**	0.887	-2.327
	(3.836)	(3.908)	(3.355)	(3.417)	(3.512)
Constant	-232.7	-290.9	-37.08	40.47	-153.4
	(77.41)	(78.92)	(67.74)	(69.05)	(70.95)
Observations	1797	1797	1798	1795	1792
R-squared	0.273	0.222	0.238	0.23	0.272

References

Asirvatham, J., R. M. Nayga, and M. R. Thomsen (2014). Peer-effects in obesity among public elementary school children: A grade-level analysis. *Applied Economic Perspectives and Policy*, doi: 10.1093/aepp/ppu011.

Behrman, J. R. (1996). The impact of health and nutrition on education. *The World Bank Research Observer 11*(1), 23–37.

Bloom, D. E. and D. Canning (2009). Population health and economic growth. In M. Spence and M. Lewis (Eds.), *Health and Growth*, pp. 53–75. NW, Washington DC: The World Bank.

Bonnefond, C. and M. Clément (2014). Social class and body weight among chinese urban adults: the role of the middle classes in the nutrition transition. *Social Science & Medicine 112*, 22–29.

Butland, B., S. Jebb, P. Kopelman, K. McPherson, S. Thomas, and J. Mardell (2007). *Tackling Obesities: Future Choices Report* (2 ed.). UK Government's Foresight Programme.

Cawley, J. (2004). The impact of obesity on wages. *Journal of Human Resources 39*(2), 451–474.

Cawley, J. and C. Meyerhoefer (2012). The medical care costs of obesity: an instrumental variables approach. *Journal of health economics 31*(1), 219–230.

Christakis, N. A. and J. H. Fowler (2007). The spread of obesity in a large social network over 32 years. *New England journal of medicine 357*(4), 370–379.

Comuzzie, A. G. and D. B. Allison (1998). The search for human obesity genes. *Science 280*(5368), 1374–1377.

Crosnoe, R. (2007). Gender, obesity, and education. *Sociology of education 80*(3), 241–260.

Datar, A., R. Sturm, and J. L. Magnabosco (2004). Childhood overweight and academic performance: national study of kindergartners and first-graders. *Obesity research 12*(1), 58–68.

Ding, W., S. F. Lehrer, J. N. Rosenquist, and J. Audrain-McGovern (2009). The impact of poor health on academic performance: New evidence using genetic markers. *Journal of Health Economics 28*(3), 578–597.

Falkner, N. H., D. Neumark-Sztainer, M. Story, R. W. Jeffery, T. Beuhring, and M. D. Resnick (2001). Social, educational, and psychological correlates of weight status in adolescents. *Obesity Research 9*(1), 32–42.

Gable, S., J. L. Krull, and Y. Chang (2012). Boys' and girls' weight status and math performance from kindergarten entry through fifth grade: a mediated analysis. *Child Development 83*(5), 1822–1839.

Glewwe, P. and H. G. Jacoby (1995). An economic analysis of delayed primary school enrollment in a low income country: the role of early childhood nutrition. *The review of Economics and Statistics*, 156–169.

Glewwe, P., H. G. Jacoby, and E. M. King (2001). Early childhood nutrition and academic achievement: a longitudinal analysis. *Journal of Public Economics 81*(3), 345–368.

James, W. (1995). A public health approach to the problem of obesity. *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity 19*, S37–45.

Jones, L., C. Wilson, and T. Wadden (2007). Lifestyle modification in the treatment of obesity: an educational challenge and opportunity. *Clinical Pharmacology & Therapeutics 81*(5), 776–779.

Kantomaa, M. T., E. Stamatakis, A. Kankaanpää, M. Kaakinen, A. Rodriguez, A. Taanila, T. Ahonen, M.-R. Järvelin, and T. Tammelin (2013). Physical activity and obesity mediate the association between childhood motor function and adolescents' academic achievement. *Proceedings of the National Academy of Sciences 110*(5), 1917–1922.

Kinge, J. M. and S. Morris (2015). The impact of childhood obesity on health and health service use: an instrumental variable approach. *Working paper*.

Kline, B. and J. L. Tobias (2008). The wages of bmi: Bayesian analysis of a skewed treatment-response model with nonparametric endogeneity. *Journal of Applied Econometrics 23*(6), 767–793.

Lake, A. and T. Townshend (2006). Obesogenic environments: exploring the built and food environments. *The Journal of the Royal society for the Promotion of Health 126*(6), 262–267.

Lindeboom, M., P. Lundborg, and B. van der Klaauw (2010). Assessing the impact of obesity on labor market outcomes. *Economics & Human Biology 8*(3), 309–319.

Ma, H. W., S. A. James, M. G. Merli, and H. Zheng (2014). An increasing socioeconomic gap in childhood overweight and obesity in china. *American journal of public health 104*(1), e14–e22.

Maes, H. H., M. C. Neale, and L. J. Eaves (1997). Genetic and environmental factors in relative body weight and human adiposity. *Behavior genetics 27*(4), 325–351.

Mayer-Foulkes, D. (2005). Human development traps and economic growth. In G. Lopez-Casasnovas, B. Rivera, and L. Currais (Eds.), *Health and Economic Growth*, pp. 115–141. Cambridge, MA: MIT Press.

Midha, T., B. Nath, R. Kumari, Y. K. Rao, and U. Pandey (2012). Childhood obesity in india: a meta-analysis. *The Indian Journal of Pediatrics 79*(7), 945–948.

Mikkilä, V., M. Lahti-Koski, P. Pietinen, S. M. Virtanen, and M. Rimpelä (2003). Associates of obesity and weight dissatisfaction among finnish adolescents. *Public Health Nutrition 6*(01), 49–56.

Morris, S. (2006). Body mass index and occupational attainment. *Journal of Health Economics 25*(2), 347–364.

Ng, M., T. Fleming, M. Robinson, B. Thomson, N. Graetz, C. Margono, E. C. Mullany, S. Biryukov, C. Abbafati, S. F. Abera, et al. (2014). Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: a systematic analysis for the global burden of disease study 2013. *The Lancet 384*(9945), 766–781.

Nghiem, H. S., H. T. Nguyen, R. Khanam, and L. B. Connelly (2015). Does school type affect cognitive and non-cognitive development in children? evidence from australian primary schools. *Labour Economics 33*, 55–65.

Norton, E. C. and E. Han (2008). Genetic information, obesity, and labor market outcomes. *Health Economics 17*(9), 1089–1104.

O'Dea, J. A. and M. J. Dibley (2014). Prevalence of obesity, overweight and thinness in australian children and adolescents by socioeconomic status and ethnic/cultural group in 2006 and 2012. *International journal of public health 59*(5), 819–828.

O'Dea, J. A., T. D. N. Hoang, and M. J. Dibley (2011). Plateau in obesity and overweight in a cross sectional study of low, middle and high socioeconomic status schoolchildren between 2004 and 2009. *International journal of public health 56*(6), 663–667.

Offer, A., R. Pechey, and S. Ulijaszek (2010). Obesity under affluence varies by welfare regimes: the effect of fast food, insecurity, and inequality. *Economics & Human Biology 8*(3), 297–308.

Sabia, J. J. (2007). The effect of body weight on adolescent academic performance. *Southern Economic Journal*, 871–900.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica: Journal of the Econometric Society*, 393–415.

Schwimmer, J. B., R. Deutsch, J. B. Rauch, C. Behling, R. Newbury, and J. E. Lavine (2003). Obesity, insulin resistance, and other clinicopathological correlates of pediatric nonalcoholic fatty liver disease. *The Journal of pediatrics 143*(4), 500–505.

Shi, H. and R. Dzhumashev (2015). Child health investment and human capital accumulation. *Working paper*.

Silles, M. A. (2009). The causal effect of education on health: Evidence from the united kingdom. *Economics of Education Review 28*(1), 122–128.

Sørensen, T., C. Holst, A. J. Stunkard, and L. T. Skovgaard (1992). Correlations of body mass index of adult adoptees and their biological and adoptive relatives. *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity 16*(3), 227–236.

Staiger, D. and J. H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65*(3), 557–586.

Swinburn, B., G. Egger, and F. Raza (1999). Dissecting obesogenic environments: the development and application of a framework for identifying and prioritizing environmental interventions for obesity. *Preventive medicine 29*(6), 563–570.

Swinburn, B. and G. J. Egger (2004). Influence of obesity-producing environments. *Handbook of obesity: Clinical applications*, 97–114.

Taras, H. and W. Potts-Datema (2005). Obesity and student performance at school. *Journal of School Health 75*(8), 291–295.

Tobin, K. (2013). Fast-food consumption and educational test scores in the usa. *Child: care, health and development 39*(1), 118–124.

Trogdon, J. G., J. Nonnemaker, and J. Pais (2008). Peer effects in adolescent overweight. *Journal of health economics 27*(5), 1388–1399.

Wang, Y. and H. Lim (2012). The global childhood obesity epidemic and the association between socio-economic status and childhood obesity. *International Review of Psychiatry 24*(3), 176–188.

Webbink, D., N. G. Martin, and P. M. Visscher (2010). Does education reduce the probability of being overweight? *Journal of Health Economics 29*(1), 29–38.

Wendt, M. and J. Kinsey (2009). Childhood overweight and school outcomes. In *Agricultural and Applied Economics Association Annual Meeting, Milwaukee*.

Wisniewski, S. L. (2010). Child nutrition, health problems, and school achievement in sri lanka. *World Development 38*(3), 315–332.

Xiao, Q., H. Arem, S. C. Moore, A. R. Hollenbeck, and C. E. Matthews (2013). A large prospective investigation of sleep duration, weight change, and obesity in the nih-aarp diet and health study cohort. *American journal of epidemiology 178*(11), 1600–1610.

Yau, P. L., M. G. Castro, A. Tagani, W. H. Tsui, and A. Convit (2012). Obesity and metabolic syndrome and functional and structural brain impairments in adolescence. *Pediatrics 130*(4), e856–e864.

Zhao, M., Y. Konishi, and P. Glewwe (2012). Does smoking affect schooling? evidence from teenagers in rural china. *Journal of health economics 31*(4), 584–598.