OBESITY AND EDUCATIONAL ATTAINMENT

ABSTRACT:

This research estimates the influence of adolescent weight status on levels of educational attainment. Prior studies have found a causal role of obesity in other economic outcomes such as income. Given the crucial role of human capital investments for economic success, estimates of the causal influence of weight problems on education accumulation are provided. Models ignoring the potential endogeneity of education and weight status indicate that obese and overweight females obtain less education than their peers while weight problems have no influence on educational attainment for males. Estimates accounting for the endogeneity of weight status indicate that weight problems cause a reduction in the likelihood of high school graduation for males and a reduced likelihood of college attendance for overweight females. Potential mechanisms to explain these relationships and their implication for school-based programs to reduce obesity are discussed.

Timothy J. Classen[*,+]

Loyola University Chicago

January 2009

JEL Classifications: I12, J24, J62

Keywords: human capital, education, obesity

---

[*] email: tclass1@luc.edu, phone: +1 312-915-6184

[+] *Acknowledgements:* I thank Bobbi Wolfe, John Mullahy, Will Manning and Kosali Simon for many useful comments on earlier versions of this paper. Portions of this research were supported by a grant from the National Institutes of Health under Ruth L. Kirschstein National Research Service Award T32 MH18029-21 from the National Institute of Mental Health. Any remaining errors are the responsibility of the author.

1. Introduction

The goal of this research is to provide an estimate of the causal influence of an individual's weight status on their level of education accumulation. The influence of obesity on income, wealth and other measures of economic status has been documented in recent research. Given the essential role of human capital formation in determining economic success, the potential influence of weight problems on this process is estimated.

The relationship of health and education has generated a substantial literature since Michael Grossman's seminal work in the early 1970s. The direction of causality posited in many of the studies following Grossman is that higher levels of education are associated with (or cause) improved health outcomes.[1] However, the question asked in this paper is not whether adults with higher levels of education have better health outcomes, but whether an individual's health status (measured by the presence of elevated weight levels) in late adolescence and early adulthood has a causal influence on how much education they accumulate.

Educational attainment may be influenced by weight problems for several reasons.[2] For those children who are obese early in life, discrimination or social isolation from classmates or teachers may limit their desire to invest in education beyond the compulsory level.[3] Furthermore, parents may invest less in obese children's post-secondary education due to a reduced rate of return from lower expected future wages or a shorter lifespan over which to reap the rewards from such investments (given the large correlation between adolescent and adult

---

[1] See Grossman (2006) for a review of studies in this area.

[2] One of the first studies documenting the relationship of obesity and education accumulation is provided by Canning and Mayer (1966) who studied the effect of weight problems on admission to highly selective colleges in the Northeast. They found that access to more prestigious higher education may be limited for obese students.

[3] See Dietz (1998) for a discussion of the social effects of obesity on children.

obesity which may lead to serious health problems such as diabetes or heart disease). The contemporaneous relationship between educational attainment and weight status may also reflect relatively high discount rates that simultaneously reduce incentives to invest in two forms of human capital: knowledge and health.

The measure of health used in this paper, Body Mass Index (BMI),[4] is a useful summary of one dimension of health that has been shown to have a significant influence on several economic measures. The estimated role of elevated BMI levels on economic outcomes including wages (Cawley (2004), Baum & Ford (2004), Pagan & Davila (1997), Averett & Korenman (1996), Conley and Glauber (2006), Norton and Han (2008)), occupational choice (Pagan & Davila (1997), Morris (2006)), and wealth (Zagorsky (2005)) has been measured in previous studies. A majority of this literature finds asymmetric consequences of weight problems between genders with obese females generally more likely to suffer negative outcomes.

Using data from the National Longitudinal Survey of Youth 1979 (NLSY79) cohort of women and their offspring in the Children and Young Adults of the NLSY79 (YA NLSY79), estimates of the causal influence of adolescent weight status on educational attainment are presented in this paper. The relationship between these two variables is likely to be endogenous, potentially due to unobserved variables (such as time preference) that influence both outcomes. Concerns of reverse causality may also arise if the process of obtaining education exposes students to greater risks for developing weight problems (such as in studies of the relation of obesity to school lunch programs as in Schanzenbach (2009), or the so-called “Freshman 15”

---

[4] BMI is calculated as weight/(height^2) where weight is measured in kilograms and height is measured in meters. Adults with a BMI above 30 are typically classified as obese, while those with BMI levels above 25 are classified as overweight.

phenomenon for college students). To address these concerns, the estimated models include instrumental variables (IVs) that provide a source of exogenous variation in adolescent weight status and may be plausibly excluded from regressions predicting educational attainment.

The paper is organized as follows. Section 2 reviews previous research on the role of obesity in shaping educational attainment and achievement. Data from the NLSY used to estimate the relationships of interest are discussed in Section 3. Estimates of the causal role of weight problems in educational attainment are presented in Section 4. Finally, Section 5 discusses the findings of this research relative to previous work on the relationship of health and education. Implications for investments in programs to reduce child and adolescent obesity are considered and possibilities for future studies of the role of obesity in the determination of human capital investments are outlined.

2. Previous Research on Education and Obesity

This section reviews estimates of the influence of weight problems on educational outcomes. Most of the studies considered report negative correlations between education accumulation and BMI levels, but do not measure the causal influence of BMI on educational attainment. Disparities in the relationship of educational achievement or accumulation to weight status between genders (with females having stronger measured relationships) are generally found.

Sargent and Blanchflower (1994) estimate the relationship between obesity at age 16 and total years of education and wages at age 23 for participants in the British National Child Development Study. They find that individuals of either gender who were obese at age 16 obtained fewer years of education than their non-obese counterparts as well as a negative correlation between BMI at age 16 and wages at age 23 for females. In a similar study, Gortmaker et al. (1993) use data from the NLSY79 to estimate the association of overweight status in late adolescence to education accumulation seven years later. They find that overweight women obtain roughly 0.3 fewer years of education than their non-overweight peers.

The effect of high school graduation or GED completion on obesity and smoking patterns is studied using the NLSY79 data in Kenkel, Lillard & Mathios (2006). With the direction of causation hypothesized as higher levels of education leading to lower obesity levels, they use a novel set of education policy variables as instruments for schooling levels. They do not find consistent evidence that the completion of high school causes lower probabilities of being obese or overweight among either gender.

Karnehed, Rasmussen, Hemmingsson, & Tynelius (2006) study a large sample of Swedish men to estimate the influence of BMI at age 18 on the likelihood of obtaining a university education. Without attempting to control for the potential endogeneity of this relationship, they find that obese men in their sample are 52 percent less likely to complete a university degree than their peers with recommended weight levels. Using a similar measure for university education employed in this paper, the authors find that obese men are 37 percent less likely to attend at least one year of college.

Using data from the National Longitudinal Study of Adolescent Health (Add Health), Sabia (2007) considers the effect of elevated BMI levels on academic performance (measured by grade point average). He finds evidence of reduced academic performance among obese white females, controlling for unobserved individual heterogeneity that may influence both weight status and academic performance jointly. He employs both an instrumental variable approach (using perceived obesity status of the adolescent's mother and father as reported by one parent as an instrument for the adolescent's obesity) and an individual-level fixed effects approach in which changes in GPA are related to changes in BMI over multiple waves of the survey. The instrumental variable employed in the current paper improves on Sabia's IV with measures of actual maternal BMI levels and using observations of BMI at the same stage of life for both generations to capture primarily the genetic correlation, rather than the potentially confounding influence of a shared household environment, in determining weight outcomes for mothers and their children.

Kaestner and Grossman (2008) critique the validity of parental obesity status as an instrument for offspring obesity (as in Sabia (2007) and the current paper) due to the correlation of parental weight with family resources and unobserved genetic factors that likely influence academic achievement. They relate performance on Peabody Individual Achievement Tests (PIAT) among children of women in the NLSY79 to weight status using BMI levels to explain changes in achievement test scores in a fixed-effects model. They do not find any significant differences in educational achievement for obese or overweight children relative to peers with recommended BMI levels after controlling for individual unobserved heterogeneity with a variety of techniques.

Crosnoe (2007) uses data from the Add Health surveys and considers the influence of obesity on college attendance. Using indicator variables for the obesity of high school students, he allows for this influence to vary depending on the percentage of all students in a given school who are classified as obese. A significant reduction in the likelihood of college attendance is found for obese females and the size of this reduction is largest in schools where obesity is relatively uncommon. Crosnoe finds no significant effects of obesity status on education accumulation for male students.

This paper improves on previous studies by using maternal weight status as an instrumental variable for adolescent weight status in estimating the causal role of weight problems on educational outcomes. Measures of maternal weight status are based on actual BMI levels, rather than parental reports of whether the child's parents are obese, and are measured at the same stage in the lifecycle as the potentially endogenous measure of adolescent weight.[5] While the timing of the observation of an instrument is generally a weak justification for satisfaction of the typical exclusion criterion (i.e., a woman's weight status in early adulthood does not directly influence the education accumulation of her offspring), in this case it enhances the case for validity of the instrument since the child likely never observed their mother at the measured BMI level used as an IV. Of course, there is a strong correlation in BMI across the lifecycle, but insofar as weight problems (or any changes in weight) occur following this period of late adolescence and early adulthood, the realization of maternal weight status used as an instrument is unobserved by the child. Relative to previous concerns regarding the use of such an IV in studies of educational performance or ability, employing this instrument in models of educational attainment provides a source of exogenous variation in the offspring's weight status that is plausibly uncorrelated with whether or not the offspring graduates high school or attends college when controls for ability, maternal education and family income are included. This results from both the stage of life at which maternal weight status is measured (when the mother

---

[5] Averett and Stifel (2007) estimate the influence of weight problems on child achievement scores using a single observation of maternal BMI from the 1981 wave of the NLSY79 to instrument for a measure of child BMI in an IV approach similar to the one employed in this paper.

is age 16 to 24) and due to the included controls capturing much of the potential variation in education accumulation (rather than ability or achievement) resulting from the influence of maternal weight status on family resources or via genetic transmission of ability.

3. Description of Data

To estimate the causal influence of obesity on education accumulation, I use data from the National Longitudinal Survey of Youth 1979 (NLSY79) and the Children and Young Adults of the NLSY79 (YA NLSY79). The NLSY79 began in 1979 with interviews of a nationally-representative sample of people born between 1957 and 1964 (so that all respondents were between ages 14 and 22 at the beginning of the survey). This sample was interviewed annually until 1986 and interviews have been conducted biennially since. Beginning in 1986, data on the children of women in the NLSY79 sample were collected. In 1994, children aged 14 or older of women in the NLSY79 were surveyed in the YA NLSY79.

Among the 6,283 women initially surveyed in the NLSY79, approximately two-thirds have had children who have been included in at least one of the Child/YA surveys. The selection criterion (for educational attainment measures) for inclusion in this study requires that the child have reached at least age 20 by the 2004 wave of the YA NLSY79 survey.[6] These exclusions result in a data set of 3,166 children born to 1,937 women in the NLSY79 sample for analysis. Summary statistics for this sample are provided in Table 1.

---

[6] This requirement results in a selection bias toward younger NLSY mothers being included in this sample. For the youngest women in the NLSY79 (age 14 in 1979) to have a child included in the sample would require that the child have been born before the woman turned age 20 (so that the child would at least have turned age 20 by the 2004 NLSY survey year).

The racial and demographic composition of this sample is evidently not representative of the U.S. population. Due to the structuring of the data for this analysis, it is not possible to incorporate the sample weights from the NLSY in a logical fashion. Sample weights are provided in each Survey year to allow for a representative cross-section to be analyzed in a given year. There are also longitudinal weights for multiple years of the NLSY survey that account for sample attrition and other forms of non-response among the members of the panel. However, due to combining observations on mothers interviewed in the beginning of the NLSY79 panel (who were assigned sample weights in the first survey in 1979) and averaging over multiple Survey years (each of which has unique sampling weights) with observations on their children who are also assigned sample weights in each Survey, this sample generated from multiple NLSY surveys does not lend itself to using the NLSY weighting structure. The population represented oversamples relatively low socioeconomic status (SES) individuals and children born to mothers at younger ages, but this is exactly the population of most concern for reduced education accumulation resulting from weight problems. Thus, understanding how obesity affects investments in education for this relatively disadvantaged population may be more policy relevant than a properly representative sample.

3.1 Measures of Educational Attainment and Weight Status

In order to measure the effect of obesity on high school completion, only those respondents in the YA NLSY79 who graduated high school, but not those who have obtained a GED, are counted as completing high school. This is due to the potential for GED completion serving as an indication of individuals leaving high school due to social isolation related to weight problems. As well, those who graduate high school have significantly different measures of SES and weight status relative to those who obtain a GED (as seen in Tables 2 and 3 and discussed in Kenkel et al. (2006)). The measure of Any College is an indicator variable which equals one for an individual who ever attended college.

A desirable feature of the longitudinal structure of the NLSY is that multiple observations of height and weight allow for several observations of BMI to be averaged over the period of interest (ages 16 to 24) in the YA NLSY79 population. Due to the NLSY79 cohort containing women who were first interviewed at age 20 or older in 1979, I extend the range for BMI observations used to create the maternal weight status IV to age 24. While the period from age 16 to 24 is likely a time with substantial changes in BMI (perhaps due to individuals making food choices without supervision from parents for the first time), averaging BMI values over this range should accurately reflect weight status during this phase of development and reduce potential bias from only a single observation of BMI.[7] In classifying individuals as obese, overweight or underweight, I use the cutoffs for adults of BMI levels over 30 indicating obesity, BMI levels above 25 indicating overweight and BMI levels below 18.5 indicating underweight.[8] While a portion of the BMI observations included use measures of the height and weight of 16 to 18 year olds, the criterion that the individual has reached age 20 to be included in the sample makes this a reasonable approach.[9]

---

[7] BMI values for women who were pregnant at the time of the survey are set to missing, but these women may still be included in the sample if they were ever interviewed between ages 16 and 24 while not pregnant.

[8] The limitations of BMI in identifying obese individuals are discussed in Cawley and Burkhauser (2008).

[9] The 2000 CDC Growth Charts for BMI have 85th percentile cutoffs for the youngest 16 year old males and females of 24.2 and 24.7 and 95th percentile cutoffs of 27.6 and 28.9, respectively. These percentiles are generally used to classify adolescents as "at risk of overweight" and overweight, but often referred to as overweight (above 85th percentile) and obese (above 95th percentile). The 85th and 95th percentile cutoffs exceed 25 and 30 respectively for females by the end of their 17th year and for males early in their 19th year.

Conditional distributions of the educational attainment of the YA NLSY79 sample dependent on adolescent weight status (based on average BMI levels calculated between ages 16 and 24) are presented in Table 2. Table 2a shows that the rate of college enrollment among females with BMI levels in the recommended range is nearly double that of obese females, while the proportion of obese females who neither graduate high school nor pass the GED is nearly double that of females with recommended BMI levels. In contrast, Table 2b indicates that obese males in the YA NLSY79 are more likely to attend college and less likely to have obtained neither a GED nor a high school diploma compared to their peers with BMI levels in the recommended range. These conditional relationships without controls for other influential factors are similar to those found in previous studies of educational attainment discussed in Section 2, with a strong negative correlation between elevated weight levels and education accumulation for females, but no significant relationship between weight status and attainment for males.

Given the large intergenerational correlations in measured ability (likely due to genetics) and educational attainment (due to both genetics and investments of parental resources) controls for maternal education levels measured in the NLSY79 are included in models estimating the education accumulation of YA NLSY79 respondents. This will likely capture some of the variation in education outcomes for offspring related to the level of available parental resources that may otherwise be argued to arise due to maternal weight status (as Kaestner and Grossman (2008) contend). Table 3 displays the strength of this intergenerational relationship of educational attainment for women in the NLSY79 and their offspring in the YA NLSY79.

Table 3a presents the conditional probabilities of four levels of education for YA NLSY79 females as a function of the education accumulation of their mothers. As expected, daughters' education levels increase monotonically as maternal education increases. For females born to mothers who did not complete high school, only 10 percent have attended any college. In contrast, two-thirds of females born to women with at least a Bachelor's degree have attended college by age 20. Similarly, less than 12 percent of females born to college-educated women did not graduate high school while the rate for females born to women with less than a high school education is over 42 percent. Similar patterns of this intergenerational relationship for males are evident in Table 3b. While overall rates of high school graduation and college enrollment are lower for males than females in the YA NLSY79, disparities based on maternal education levels are similar. More than half of men born to mothers with less than a high school education do not graduate high school themselves. This compares to less than 12 percent of men born to women who graduated college.

As in the intergenerational correlation of education levels, the intergenerational relationship of BMI levels likely reflects both genetic factors and investments of parental resources, such as the purchase of (more expensive) healthier foods or time spent playing with children. Thus, there exists a strong correlation in BMI levels across generations, which is reflected in the results presented in Table 4.[10] The strength of this relationship provides justification (at least in terms of statistical power) for the use of maternal weight status as an instrumental variable for the weight status of their offspring in the relationships estimated in Section 4. The novel feature of the NLSY data is that weight status can be measured for both

---

[10] Classen (2008) estimates an intergenerational correlation of BMI of 0.38 between women and their daughters and 0.32 between women and their sons using a similar sample from the NLSY as the one used here.

generations at similar stages of development (rather than the contemporaneous correlation which likely reflects the shared household environment relatively more than genetic transmission which should provide a source of exogenous variation).[11] Furthermore, the strength of the maternal measures of weight status used in this paper as an IV is enhanced by the fact that children in the YA NLSY79 likely did not observe their mother's weight status between ages 16 and 24 (prior to their birth or when they were very young).

Table 4a displays the intergenerational relationship of weight status for females to their mothers when both generations are between ages 16 and 24. The rate of obesity for women born to mothers who were obese in early adulthood is four times as large as that for women born to mothers with BMI in the recommended range. Similarly, only one-quarter of women born to obese mothers had a BMI level in the recommended range by early adulthood compared to more than 60 percent of women born to mothers with recommended BMI levels. While rates of obesity among males in the YA NLSY79 sample are lower than those for females, two-thirds of men born to obese women become obese or overweight by late adolescence. Less than 30 percent of men born to obese mothers have BMI levels in the recommended range, while 36 percent of men born to women with recommended BMI levels develop weight problems (obese or overweight) reflecting the general increase in weight levels between the two generations.

3.2 Other Explanatory Variables of Educational Attainment

Measures of educational achievement for children in the YA NLSY79 samples were assessed via Peabody Individual Achievement Tests (PIAT) (as discussed in Kaestner and

---

[11] Of course, if the same genes that predict weight status are also influential in the determination of education accumulation then this transmission may weaken the exclusion restriction of maternal weight status as a valid instrument. Included controls for maternal education should account for some of this potential endogeneity.

Grossman (2008)). I include percentile scores from both the math and reading comprehension exams to provide a control for child ability at an earlier stage of life. When the PIAT exam was given to the same child several times over multiple survey years, I use scores from the exam at the oldest age so it is closest in age to the education outcome of interest. These measures provide substantial explanatory power for lifetime education accumulation and attenuate the potential omitted variable bias in the relationship of education accumulation to weight status resulting from unobserved child ability.

To control for relative levels of parental resources available for investment in the human capital formation of their children, I include dummy indicator variables for the quartile of family income for YA NLSY79 respondents. Household income is averaged over all available observations of income from when the offspring was age 2 to 18 (available from interviews of the mothers in the NLSY79) and adjusted for household size. The quartiles are determined based on the entire population of respondents in the Child/YA NLSY79 surveys. As seen in Table 1, this results in relatively more low-income households and relatively fewer households drawn from the highest income quartile due to the over-sampling of children born to relatively younger mothers.

Measures of the order of birth for an individual within their family and of the age of one's mother's at the time of one's birth provide explanatory power for education accumulation along several dimensions. Maternal age at the time of birth reflects a variety of potentially influential (unobserved) factors including the quality of parental investments in the human capital formation process of their offspring and other characteristics associated with educational attainment, such as rates of time preference. Order of birth within a family may indicate potential competition among siblings for household resources (including parental time or monetary resources in lower-income households).[12] A measure of whether female survey respondents became pregnant and gave birth to children prior to age 19 is included due to its substantial explanatory power in rates of high school graduation (and may also proxy for unobserved rates of time preference). The high rate of teenage pregnancy in the YA NLSY79 sample is likely due to the previously discussed over-representation of children born to younger (and lower SES) mothers in this sample.

4. Estimates of the Causal Influence of Weight Problems on Education Accumulation

Given the age range of respondents in the ongoing YA NLSY79 surveys, the observed educational attainment of this population is truncated relative to lifetime attainment. Thus, the measures of educational attainment considered in this analysis are whether respondents in the YA NLSY79 have graduated high school or enrolled in college by age 20. These educational attainment measures and the empirical approach employed are similar to that of Evans and Schwab (1995) who studied the influence of Catholic school attendance on education accumulation.

The model for estimation is based on the latent demand for education (EDU^*) that results in observed attainment of high school graduation or college enrollment (EDU).[13] Thus, models of the form

---

[12] The influence of birth order on education outcomes is discussed in Black, Devereux & Salvanes (2005).

[13] Binary indicator variables of educational attainment, rather than the number of years of education completed, are used based on the notion that a high school diploma has signaling value for job applicants in labor markets relative to someone with an equivalent number of years of education without a diploma (the so-called "sheepskin effect"). A discussion of this effect in the context of health and education is provided in Cowell (2006).

$$\begin{aligned} EDU^* &= \alpha + \beta Weight + \delta X + \varepsilon \\ EDU &= 1_{\{\varepsilon > -(\alpha + \beta Weight + \delta X)\}} \end{aligned} \tag{1}$$

are estimated, where *Weight* is a measure of weight status (binary variables indicating obesity or overweight) and X contains measures of the explanatory variables discussed in Section 3. If the weight status indicated by *Weight* is correlated with unobserved (latent) factors affecting educational attainment in ε, estimates of β will be biased due to this endogeneity. A potential remedy for this bias is provided by instrumental variable procedures that use maternal weight status (when the mother was 16 to 24 years old) as a source of exogenous variation that is strongly correlated with the determination of *Weight* and not directly influential in the observed realizations of *EDU*. Hence, we estimate models of education outcomes using indicator variables for maternal weight status (whether the mother was obese, overweight or underweight at ages 16-24, the same stage of life as observed weight status (*Weight*) for their offspring) as an instrumental variable for the offspring's obesity or overweight status. Thus, to control for the potential endogenity between *EDU* and *Weight*, we allow for the observed weight status indicators to result from the determination of *Weight** which is influenced by the variables in X as well as measures of maternal weight status in Z. This gives

$$\begin{aligned} Weight^* &= \mu + \gamma Z + \theta X + \eta \\ Weight &= 1_{\{\eta > -(\mu + \gamma Z + \theta X)\}} \end{aligned} \tag{2}$$

I first estimate the system (1) ignoring the potential endogeneity of *Weight* and *Edu* with univariate probit and linear probability models. Based on the results presented in Bhattacharya, Goldman & McCaffrey (2006) and recommendations of Terza, Bradford & Dismuke (2008) for the estimation of models of the form of (1) and (2) with limited dependent variables and a binary endogenous regressor, bivariate probit models of education accumulation are estimated using indicators of maternal weight status (Z) as instrumental variables to explain observed weight outcomes (*Weight*). In this case, it is assumed that ε and η are distributed bivariate normal with mean zero, variance equal to 1 and $\text{cov}(\varepsilon, \eta) = \rho \neq 0$. Likelihood-ratio tests of the null hypothesis that $\rho = 0$ (indicating no correlation between the error terms in the *EDU* and *Weight* equations) following the estimation of the bivariate probit models provides a simple test for the exogeneity of weight status in educational attainment. For comparison, two-stage linear probability models (2SLS LPM) are estimated (as advocated by Angrist (2001)), again using maternal weight status binary indicator variables to instrument for the offspring's weight status.

4.1 Influence of Weight Problems on Likelihood of High School Graduation

Using the models specified in the previous section, estimates of the influence of weight status and other measured covariates on the likelihood of high school graduation are presented in Tables 5 and 6. Given previous findings indicating asymmetric consequences of weight problems on economic success by gender, all models are estimated separately for females and males.

Columns (1) and (2) in Tables 5 and 6 display results of univariate probit models for high school graduation including binary variables indicating obese and overweight (including obese) females and males, respectively. Similarly, estimates of linear probability models of these outcomes are displayed in columns (5) and (6) in these tables. These results reveal differences in the direction of the association of weight status to high school graduation between genders. The univariate probit and LPM results (columns (1) and (5) of Tables 5 and 6) indicate that obesity is not associated with any statistically significant difference in the likelihood of completing high school, although the estimated coefficients for each gender are of opposite signs. Obese males are estimated to be more likely to graduate high school, but none of the estimates are statistically different from zero at the 5 percent level of significance. Including respondents with BMI levels between 25 and 30 (indicating overweight) in the measure of weight problems (columns (2) and (6)) produces a similar result for the likelihood of high school graduation of males with a positive, but statistically insignificant, estimate. Including overweight respondents in the binary indicator of weight problems for females increases the size of the estimated negative association of weight problems to high school graduation and increases the precision of the estimated coefficient but both remain statistically insignificant at the 5% confidence level.

Tables 5 and 6 also display estimates of bivariate probit and two-stage least squares linear probability models that include three binary indicators of maternal weight status (obese, overweight and underweight) as instrumental variables in the determination of the weight status of their offspring. The results for females in Table 5 show that the relationship of weight problems to the likelihood of high school graduation remains negative, but statistically insignificant. Likelihood ratio tests to determine whether the error terms of the weight status and high school graduation equations for females have a significant correlation in the bivariate probit models cannot reject the null hypothesis of $\rho = 0$ indicating the potential exogeneity of weight status to high school graduation for females. Support for the strength of the maternal weight status instruments in explaining the weight status of daughters is indicated by first-stage F-statistics over 10 in both the obese and overweight specifications in columns (7) and (8).

Introducing instrumental variables of maternal weight status to explain exogenous variation in the weight status of males in Table 6 indicates a potentially negative causal relationship of weight problems on the likelihood of high school completion for males. Likelihood ratio tests indicate the presence of an endogenous relationship in the joint determination of whether a male graduates high school and develops weight problems by late adolescence, with the hypothesis that $\rho = 0$ rejected at the 5% level in both bivariate probit specifications. The coefficient on the binary indicator of obesity is negative and significant at the 5% level while including overweight males reduces the size of the estimated effect, but still indicates a negative influence of weight status on high school graduation at the 10% level of significance. Similar results for the negative influence of weight problems on high school graduation for males are found in the 2SLS LPM estimates, although only the result using an indicator of obesity is statistically significant at the 10% confidence level. While the binary indicators of maternal weight status included as instruments in the first stage regressions have strong statistical significance and coefficients of the expected sign, the F-stat levels indicate potential concern for weak instruments.

Coefficients for other included covariates in the determination of high school graduation are generally of the expected sign and relative magnitude. Results in Tables 5 and 6 indicate that children born to mothers with higher levels of education are more likely to graduate high school. The one anomalous result is that daughters of women who graduated college appear no more likely to graduate high school than peers with mothers who did not graduate high school. This likely results from the relatively small sample of mothers who graduated college (less than 6 percent of the sample due to the oversampling of women who had children at relatively young ages) as well as the strength of the correlation between maternal education and achievement levels measured by PIAT percentiles. These PIAT measures from childhood or early adolescence have strong predictive power for the likelihood of high school graduation in all specifications. Being born to a relatively younger mother significantly reduces the likelihood of high school graduation, but being an older child within the distribution of siblings of a family has a positive influence on high school graduation. Not surprisingly, females who became pregnant and gave birth during high school have significantly lower likelihoods of high school graduation with the estimated marginal effect in the univariate probit estimates of reducing this likelihood by 22 percent. The included measures of family income indicate that students from families in the lowest income quartiles are less likely to graduate high school than their peers from families in the highest income quartile. The strength of this measured relationship in the second and third income quartiles is likely attenuated by the included measures of maternal education and PIAT scores.

4.2 Estimates of the Influence of Weight Status on College Enrollment

Tables 7 and 8 provide results for estimates of the likelihood of college enrollment by age 20 for respondents in the YA NLSY79. The univariate probit and LPM estimates in columns (1) and (5) of Table 7 both indicate significant decreases in the likelihood of college enrollment for obese females relative to their peers with BMI levels in the recommended range. Marginal effects from the estimated univariate probit model indicate that obese females are 8 percent less likely to enroll in college than women with BMI levels in the recommended range, controlling for maternal education and other included covariates. Including overweight females in the binary indicator variables of weight problems (columns (2) and (6)) increases the precision of the estimates while the magnitude of the estimated relationship does not change relative to including only obese females. Including measures of maternal weight status as instrumental variables to explain the weight status of their daughters in bivariate probit and 2SLS LPM estimates substantially increases the magnitude of the coefficient estimates (and standard errors) for the indicator variables of weight status. In IV models using overweight (BMI levels of 25 and above) indicator variables (columns (4) and (8)), the estimated influence of weight problems on the likelihood of college attendance is more than twice as large as in models that do not account for the potential endogeneity between weight status and college enrollment for females. Although controlling for endogeneity increases the standard errors, the p-value for the coefficient on the indicator variable of overweight is below 0.07 in both the bivariate probit and 2SLS LPM cases. For models including only obese females (columns (3) and (7)), the result from the 2SLS LPM estimate is marginally statistically significant at the 10% level (p-value equals 0.099). Whether these outcomes can be presumed to be exogenously related is tested via likelihood ratio tests which again indicate that the null hypothesis of no correlation between error terms in the bivariate probit cannot be rejected. Under this assumed exogeneity, the estimate marginal effects from the univariate probit models provide a lower bound on the size of the causal influence of obesity or overweight on the likelihood of college attendance for females.

As shown in Table 8, there is no statistically significant effect (at even the 10 percent level of significance) of weight status on the likelihood of college enrollment among the males in this sample for all specifications. Models controlling for the endogeneity of college enrollment and weight status produce coefficients of the opposite sign and larger magnitude than in the univariate probit and LPM estimates, but standard errors also increase substantially (potentially due to the issue of weak instruments discussed in the results for male high school graduation). These results that obese and overweight males are no less likely to enroll in college are consistent with evidence that obese males do not generally face a wage penalty after college relative to their non-obese peers.

For the estimated influence of the other included covariates, there are large and statistically significant effects of having a mother who completed college on the likelihood of college enrollment, with the daughters of women who graduated college having an increase in their likelihood of going to college of roughly 35 percent (in the univariate probit estimates) relative to their peers whose mothers did not complete high school. While the magnitude of the measured effect is somewhat smaller for males, there is consistent evidence that higher levels of maternal education accumulation significantly increase the likelihood of college enrollment.

Being born to a relatively older woman increases the likelihood of college attendance, while older siblings within a family are more likely to attend college (although this effect is not statistically significant for females). Females who had children during high school are again substantially less likely to make investments in education with a reduction in the likelihood of college attendance in the univariate probit model of 14 percent relative to females who did not give birth during high school. Measures of child ability from the PIAT exams again provide strong explanatory power for the likelihood of increased education accumulation. While income would be expected to have a significant influence on the likelihood of college enrollment, these results indicate that only those males from families in the lowest income quartile are significantly less likely to attend college than their peers from families in the highest income quartile. There do not appear to be significant differences in the likelihood of college attendance between females from families in the top and bottom income quartiles, but this relationship is likely confounded by the inclusion of measures of high school pregnancy, PIAT performance and maternal education which have significant power in predicting income quartiles for families of female students.

5. Discussion and Directions for Further Research

This paper adds to a growing literature on the role of weight problems in the determination of economic success. I study the causal influence of weight status on two measures of human capital accumulation, high school graduation and college attendance. Using instrumental variables that have strong predictive power in the development of weight problems among adolescents, I estimate bivariate probit and two-stage linear probability models due to the presence of binary regressors (indicator variables for obesity or overweight) that are likely endogenously related to the determination of educational attainment. These results add to a growing body of evidence on the asymmetric influence of obesity on economic outcomes between genders.

In contrast to previous studies that did not control for the likely endogenous relationship between the development of weight problems during adolescence and the decision to complete high school, I find negative consequences of being obese or overweight for males. I find that being obese or overweight in late adolescence causes a higher likelihood of dropping out of high school for males. In several specifications for the effect of weight status on high school graduation for females, I find no significant difference in the likelihood of graduating high school for obese or overweight females. Conversely, consistent with previous studies indicating a causal negative influence of obesity on wages for females, I find that the presence of weight problems during adolescence and early adulthood causes reductions in the likelihood of college attendance for females. None of the estimated models indicate a significant relationship between weight status and college enrollment for males which may reflect the absence of reduced returns to such human capital investments in the labor market for obese and overweight males estimated in prior research.

With the substantial growth in rates of childhood and adolescent obesity during the previous three decades, understanding the influence of such health behaviors in the process of human capital formation seems essential. While this study has used observations on education and weight status from a national sample, the local effects within a school or community of (relative) weight status on education accumulation may provide an enhanced understanding of this relationship. As well, the over-representation of relatively disadvantaged youth in this sample from the YA NLSY79 may for issues of identifying the results of interest due to higher rates of weight problems, teenage pregnancy and high school dropouts, but raise concerns about the external validity of the estimated effects for a nationally representative population.

Insofar as the development of weight problems reduces incentives to invest in education accumulation (as the results from this paper indicate), efforts to reduce the prevalence of elevated BMI levels in children and adolescents may mitigate this deleterious effect on human capital formation. Such an effect could provide a substantial (perhaps previously unaccounted for) benefit for school-based programs with a demonstrated ability to affect child and adolescent weight levels. Given the tremendous welfare loss resulting from reduced investments in human capital, such efforts may produce substantial returns in the form of increased educational attainment among students who would have developed weight problems in the absence of such programs.

Bibliography

Angrist, J. (2001) “Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice,” *Journal of Business and Economic Statistics*, vol. 19, no. 1, pp. 2-16.

Averett, S. L. and S. Korenman (1996) “The Economic Reality of the Beauty Myth,” *Journal of Human Resources,* vol. 31, no. 2, pp. 304-330.

Averett, S. L. and D. C. Stifel (2007) “Food for Thought: The Cognitive Effects of Childhood Malnutrition in the United States,” mimeo, Lafayette College.

Baum, C.L. and C.J. Ruhm (2007) “Age, Socioeconomic Status and Obesity Growth,” NBER Working Paper Number 13289

Bhattacharya, J.; D. Goldman and D. McCaffrey (2006) “Estimating probit models with self-selected treatments,” *Statistics in Medicine*, vol. 25, no. 3, pp. 389-413.

Black, S.; P. Devereux and K. Salvanes (2005) “The More the Merrier? The Effect of Family Size and Birth Order on Children’s Education,” *Quarterly Journal of Economics*, vol. 120, no. 2, pp. 669–700.

Canning, H and J. Mayer (1966) “Obesity - its possible effect on college acceptance,” *New England Journal of Medicine*, vol. 275, pp. 1172-1174.

Cawley, J. (2004) “The Impact of Obesity on Wages,” *Journal of Human Resources*, vol. 39, no. 2, pp. 451-474

Cawley, J. and R. Burkhauser (2008) “Beyond BMI: The Value of More Accurate Measures of Fatness and Obesity in Social Science Research," *Journal of Health Economics*, vol. 27, no. 2, pp. 519-529.

Classen, T. (2008) “Measures of the Intergenerational Transmission of Body Mass Index,” Working paper.

Conley, D. and R. Glauber. 2006. “Gender, Body Mass and Socioeconomic Status: New Evidence from the PSID.” *Advances in Health Economics and Health Services Research.* vol.17, pp. 255-280.

Cowell, A. J. (2006) “The relationship between education and health behavior: some empirical evidence,” *Health Economics*, vol. 15, no. 2, pp. 125 – 146

Crosnoe, R. (2007) “Gender, Education and Obesity,” *Sociology of Education*, vol. 80, no. 3, pp. 241-260

Dietz, W. H. (1998) “Health Consequences of Obesity in Youth: Childhood Predictors of Adult Disease,” *Pediatrics*, vol. 101, no. 3, pp. 518-525

Evans, W. N. and R. M. Schwab (1995) “Finishing High School and Starting College: Do Catholic Schools Make a Difference?” *Quarterly Journal of Economics*, vol. 110, no. 4, pp. 941-974.

Gortmaker, S.L.; A. Must; J. M. Perrin; A. M. Sobol; W. H. Dietz (1993) “Social and Economic Consequences of Overweight in Adolescence and Young Adulthood,” *New England Journal of Medicine*, vol. 329, no. 14, pp.1008-1012

Grossman, M. (2006) “Education and Non-Market Outcomes,” in *Handbook of the Economics of Education*, Vol. 1, E. Hanushek and F. Welch, eds., Amsterdam: Elsevier, pp. 577–633.

Kaestner, R. and M. Grossman (2008) “Effects of Weight on Children’s Educational Achievement,” NBER Working Paper No. W13764

Karnehed, N.; F. Rasmussen; T. Hemmingsson; P. Tynelius (2006) “Obesity and Attained Education: Cohort Study or More Than 700,000 Swedish Men,” *Obesity*, vol. 14, no. 8, pp. 1421-1428.

Kenkel, D.; D. Lillard; A. Mathios (2006) “The Roles of High School Completion and GED Receipt in Smoking and Obesity,” *Journal of Labor Economics*, vol. 24, no. 3, pp. 635-660.

Morris, S. (2006) “Body Mass Index and Occupational Attainment,” *Journal of Health Economics*, vol. 25, no. 2, pp. 347-364

Norton E.C. and Han, E. (2008) “Genetic Information, Obesity, and Labor market outcomes,” *Health Economics*, vol. 17, no. 9, pp. 1089−1104.

Pagan, J. A. and A. Davila (1997) “Obesity, Occupational Attainment, and Earnings,” *Social Science Quarterly*, vol. 78, no. 3, pp. 756-770

Sabia, J. J. (2007) “The Effect of Body Weight on Adolescent Academic Achievement,” *Southern Economic Journal,* vol. 73, no. 4, pp. 871-900.

Sargent, J. and D. Blanchflower (1994) “Obesity and Stature in Adolescence and Earnings in Young Adulthood. Analysis of a British Birth Cohort,” *Archives of Pediatrics and Adolescent Medicine*, vol. 148, no. 7, pp. 681– 687

Schanzenbach, D. W. (2009) “Does the Federal School Lunch Program Contribute to Childhood Obesity?” *Journal of Human Resources*, forthcoming.

Terza, J. V.; W. D. Bradford; and C.E. Dismuke (2008) “The Use of Linear Instrumental Variables Methods in Health Services Research and Health Economics: A Cautionary Note,” *Health Services Research*, vol.43, no. 3, pp. 1102 – 1120

Zagorsky, J. (2005) “Health and wealth: The late-20th century obesity epidemic in the U.S.,” *Economics and Human Biology*, vol. 3, no. 2, pp. 296-313.

Table 1 - Means of Variables for Children of NLSY79 Sample

(*Note:* Standard deviation for continuous variables in parentheses)

	Males (n = 1,551)	**Females** (n = 1,615)
Highest Educational Attainment		
Any College	23.8%	35.1%
High School Grad.	46.4%	44.6%
HS Grad (incl. Any College)	*68.9%*	*78.1%*
GED/No HS Grad.	13.7%	9.2%
No HS Grad./No GED	16.1%	11.0%
Weight Status, Age 16-24		
Obese *(BMI > 30)*	12.1%	16.2%
Overweight *(25 < BMI < 30)*	28.7%	23.2%
Obese or Overweight *(BMI > 25)*	*40.7%*	*39.4%*
Recommended *(18.5 < BMI < 25)*	57.9%	55.5%
Underweight *(BMI < 18.5)*	1.4%	5.1%
Average BMI	25.0 (4.4)	25.0 (5.4)
Maternal Education (Highest Level Attained)		
College Graduate	7.2%	5.9%
Some College	21.6%	25.8%
High School Graduate	55.3%	50.3%
Less than High School	15.9%	18.0%
Income Quartiles (relative to entire NLSY79 population)		
1st (Lowest) Quartile	35.3%	34.5%
2nd Quartile	28.8%	28.8%
3rd Quartile	23.5%	21.9%
4th (Highest) Quartile	12.5%	14.8%
Demographics		
Mother's Age at Birth	20.3 (2.9)	20.3 (2.9)
Birth Order	1.6 (0.8)	1.6 (0.9)
Pregnant by Age 19	NA	34.8%
Black	39.5%	41.2%
Hispanic	23.0%	21.6%
PIAT Percentiles		
Math test	44.9 (26.7)	42.5 (26.0)
Reading test	48.0 (30.8)	53.0 (28.9)
Maternal Weight Status, age 16-24 (IVs)		
Obese	5.5%	6.7%
Overweight	19.6%	20.7%
Recommended	69.3%	66.9%
Underweight	5.5%	5.7%
Average BMI	23.1 (3.9)	23.2 (4.0)

Note : Some respondents earned a GED and attended college, so HS Grad (incl. Any College) is less than the sum of High School Grad. and Any College.

Table 2 - Distributions of Educational Attainment Conditional on Weight Status

Table 2a - Highest Educational Attainment for Females

Weight Status (age 16-24)	No HS/GED	GED	HS Grad	Any College	# obs.
Underweight *BMI < 18.5*	6.1%	4.9%	45.1%	43.9%	82
Recommended *18.5 < BMI < 25*	8.9%	8.4%	41.3%	41.4%	898
Overweight *25 < BMI < 30*	12.9%	12.6%	47.5%	27.1%	373
Obese *BMI > 30*	17.2%	8.8%	51.9%	22.1%	262
# obs.	178	149	721	567	n = 1,615

Table 2b - Highest Educational Attainment for Males

Weight Status (age 16-24)	No HS/GED	GED	HS Grad	Any College	# obs.
Underweight	9.5%	4.8%	57.1%	28.6%	21
Recommended	17.4%	13.6%	45.7%	23.4%	899
Overweight	14.4%	15.1%	47.5%	23.0%	444
Obese	14.4%	12.3%	46.0%	27.3%	187
# obs.	249	213	720	369	n = 1,551

Table 3 - Distributions of Intergenerational Educational Attainment

Table 3a - Highest Educational Attainment for Female Offspring

Highest Level of Maternal Education	No HS/GED	GED	HS Grad	Any College	# obs.
Less than High School	29.0%	13.4%	47.6%	10.0%	290
High School Graduate	9.6%	8.6%	49.3%	32.5%	813
Some College	3.4%	8.2%	38.1%	50.4%	417
College Graduate	2.1%	6.3%	24.2%	67.4%	95
# obs.	178	149	721	567	n = 1,615

Table 3b - Highest Educational Attainment for Male Offspring

Highest Level of Maternal Education	No HS/GED	GED	HS Grad	Any College	# obs.
Less than High School	31.3%	20.7%	40.2%	7.7%	246
High School Graduate	15.6%	14.3%	48.6%	21.4%	858
Some College	9.3%	9.9%	49.3%	31.6%	335
College Graduate	6.3%	5.4%	34.8%	53.6%	112
# obs.	249	213	720	369	n = 1,551

Table 4 - Distributions of Offspring Weight Status Conditional on Maternal Weight Status

Table 4a - Distribution of Weight Status of Female Offspring (Age 16 - 24)

Maternal Weight Status (age 16-24)	Underweight	Recommended	Overweight	Obese	# obs.
Underweight *BMI < 18.5*	14.1%	70.7%	14.1%	1.1%	92
Recommended *18.5 < BMI < 25*	5.7%	62.4%	20.7%	11.1%	1080
Overweight *25 < BMI < 30*	1.2%	39.2%	32.0%	27.5%	334
Obese *BMI > 30*	2.8%	25.7%	26.6%	45.0%	109
# obs.	82	898	373	262	n = 1,615

Table 4b - Distribution of Weight Status of Male Offspring (Age 16 - 24)

Maternal Weight Status (age 16-24)	Underweight	Recommended	Overweight	Obese	# obs.
Underweight	4.7%	74.4%	16.3%	4.7%	86
Recommended	1.0%	62.6%	26.9%	9.5%	1075
Overweight	1.3%	44.7%	36.8%	17.1%	304
Obese	2.3%	30.2%	33.7%	33.7%	86
# obs.	21	899	444	187	n = 1,551

Table 5 - Estimates of Likelihood of High School Graduation for Females

(Dependent Variable = 1 if Female graduated high school, mean = 78.1%, n = 1,615)

	Univariate Probit		Bivariate Probit		Linear Probablity Model		2SLS LPM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Adolescent Weight Status (relative to non-obese or non-overweight)								
Obese	-0.11		-0.196		-0.028		-0.177	
BMI > 30	[0.104]		[0.393]		[0.029]		[0.118]	
Marginal Effect	-(0.030)							
Overweight (incl. obese)		-0.157		-0.323		-0.038		-0.11
BMI > 25		[0.082]		[0.295]		[0.022]		[0.087]
		-(0.042)						
Mother's Education (relative to Less than High School education)								
College Graduate	0.257	0.241	0.261	0.228	0.075	0.071	0.081	0.066
	[0.223]	[0.223]	[0.223]	[0.223]	[0.050]	[0.050]	[0.050]	[0.050]
	(0.060)	(0.057)						
Some College	0.392	0.387	0.394	0.381	0.117	0.116	0.119	0.114
	[0.133]**	[0.134]**	[0.133]**	[0.134]**	[0.039]**	[0.039]**	[0.039]**	[0.040]**
	(0.093)	(0.092)						
High School Graduate	0.24	0.241	0.239	0.24	0.091	0.091	0.091	0.091
	[0.111]*	[0.112]*	[0.111]*	[0.112]*	[0.037]*	[0.037]*	[0.037]*	[0.037]*
	(0.062)	(0.063)						
Income Quartile (relative to Highest Income Quartile)								
1st (Lowest) Quartile	-0.25	-0.241	-0.24	-0.222	-0.059	-0.058	-0.042	-0.05
	[0.152]	[0.151]	[0.158]	[0.155]	[0.034]	[0.033]	[0.036]	[0.034]
	-(0.068)	-(0.065)						
2nd Quartile	-0.07	-0.064	-0.062	-0.05	-0.006	-0.005	0.006	0.001
	[0.147]	[0.147]	[0.150]	[0.149]	[0.030]	[0.029]	[0.031]	[0.030]
	-(0.018)	-(0.017)						
3rd Quartile	-0.075	-0.069	-0.07	-0.058	-0.012	-0.011	-0.006	-0.007
	[0.152]	[0.152]	[0.153]	[0.153]	[0.028]	[0.028]	[0.028]	[0.028]
	-(0.020)	-(0.018)						
Demographics								
Mother's Age at Birth	0.048	0.047	0.047	0.046	0.011	0.011	0.01	0.01
	[0.017]**	[0.017]**	[0.017]**	[0.017]**	[0.004]**	[0.004]**	[0.004]*	[0.004]*
	(0.012)	(0.012)						
Birth Order	-0.211	-0.208	-0.211	-0.206	-0.059	-0.058	-0.059	-0.057
	[0.050]**	[0.050]**	[0.050]**	[0.050]**	[0.014]**	[0.014]**	[0.014]**	[0.014]**
	-(0.055)	-(0.054)						
Pregnant by age 19	-0.764	-0.759	-0.763	-0.751	-0.217	-0.216	-0.218	-0.213
	[0.077]**	[0.077]**	[0.077]**	[0.079]**	[0.023]**	[0.023]**	[0.023]**	[0.023]**
	-(0.218)	-(0.217)						
PIAT Percentiles								
Math Test	0.006	0.006	0.006	0.006	0.001	0.001	0.001	0.001
	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.000]**	[0.000]**	[0.000]**	[0.000]**
	(0.002)	(0.002)						
Reading Test	0.004	0.004	0.004	0.004	0.001	0.001	0.001	0.001
	[0.002]**	[0.002]**	[0.002]*	[0.002]*	[0.000]**	[0.000]**	[0.000]*	[0.000]**
	(0.001)	(0.001)						
Constant	-0.169	-0.129	-0.142	-0.037	0.511	0.524	0.556	0.563
	[0.361]	[0.361]	[0.377]	[0.388]	[0.090]**	[0.090]**	[0.096]**	[0.100]**
Pseudo R^2 / R^2	0.18	0.18			0.18	0.18	0.17	0.18
ρ			0.051	0.107				
X^2 value for L.R. test of ρ = 0			0.0602	0.3694		First-stage F-stat	12.63	16.66
p-value for L.R. test of ρ = 0			0.806	0.543				

Notes: Robust standard errors accounting for multiple observations from same mother in brackets. Marginal effects for Univariate Probit models in parentheses. Bivariate Probit and 2SLS LPM estimates use indicator variables for maternal weight status (obese, overweight, underweight) as instrumental variables for daughter's weight status. All regressions contain binary controls for whether the individual is black or Hispanic.

* significant at 5%; ** significant at 1%

Table 6 - Estimates of Likelihood of High School Graduation for Males

(Dependent Variable = 1 if Male graduated high school, mean = 68.9%, n = 1,551)

	Univariate Probit		Bivariate Probit		Linear Probablity Model		2SLS LPM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Adolescent Weight Status (relative to non-obese or non-overweight)								
Obese	0.111		-0.907		0.033		-0.382	
BMI > 30	[0.111]		[0.370]*		[0.034]		[0.212]	
Marginal Effect	(0.037)							
Overweight (incl. obese)		0.059		-0.561		0.017		-0.187
BMI > 25		[0.073]		[0.324]		[0.022]		[0.117]
		(0.020)						
Mother's Education (relative to Less than High School education)								
College Graduate	0.639	0.635	0.582	0.604	0.184	0.183	0.167	0.182
	[0.196]**	[0.197]**	[0.196]**	[0.190]**	[0.052]**	[0.053]**	[0.057]**	[0.052]**
	(0.178)	(0.177)						
Some College	0.464	0.467	0.455	0.441	0.168	0.168	0.174	0.167
	[0.131]**	[0.131]**	[0.130]**	[0.131]**	[0.044]**	[0.044]**	[0.046]**	[0.045]**
	(0.144)	(0.145)						
High School Graduate	0.33	0.33	0.308	0.309	0.127	0.127	0.126	0.125
	[0.104]**	[0.104]**	[0.104]**	[0.105]**	[0.038]**	[0.038]**	[0.040]**	[0.039]**
	(0.112)	(0.113)						
Income Quartile (relative to Highest Income Quartile)								
1st (Lowest) Quartile	-0.366	-0.367	-0.308	-0.311	-0.111	-0.111	-0.093	-0.098
	[0.139]**	[0.139]**	[0.136]*	[0.140]*	[0.040]**	[0.040]**	[0.042]*	[0.041]*
	-(0.127)	-(0.128)						
2nd Quartile	-0.149	-0.15	-0.108	-0.107	-0.028	-0.028	-0.012	-0.016
	[0.134]	[0.133]	[0.131]	[0.134]	[0.035]	[0.035]	[0.037]	[0.037]
	-(0.051)	-(0.052)						
3rd Quartile	-0.006	-0.005	0.033	0.024	0.011	0.012	0.029	0.021
	[0.138]	[0.138]	[0.135]	[0.136]	[0.033]	[0.033]	[0.036]	[0.035]
	-(0.002)	-(0.002)						
Demographics								
Mother's Age at Birth	0.04	0.041	0.046	0.037	0.012	0.013	0.016	0.012
	[0.015]*	[0.015]**	[0.015]**	[0.015]*	[0.005]**	[0.005]**	[0.005]**	[0.005]*
	(0.013)	(0.014)						
Birth Order	-0.152	-0.154	-0.172	-0.165	-0.05	-0.051	-0.062	-0.056
	[0.054]**	[0.054]**	[0.052]**	[0.052]**	[0.017]**	[0.017]**	[0.018]**	[0.017]**
	-(0.051)	-(0.052)						
PIAT Percentiles								
Math Test	0.008	0.008	0.007	0.008	0.002	0.002	0.002	0.002
	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.001]**	[0.001]**	[0.001]**	[0.001]**
	(0.003)	(0.003)						
Reading Test	0.005	0.005	0.005	0.005	0.002	0.002	0.002	0.002
	[0.001]**	[0.001]**	[0.001]**	[0.001]**	[0.000]**	[0.000]**	[0.000]**	[0.000]**
	(0.002)	(0.002)						
Constant	-0.774	-0.798	-0.737	-0.494	0.248	0.241	0.246	0.331
	[0.334]*	[0.336]*	[0.330]*	[0.378]	[0.103]*	[0.103]*	[0.109]*	[0.121]**
Pseudo R^2 / R^2	0.13	0.13			0.15	0.15	0.07	0.11
ρ			0.549	0.399				
X^2 value for L.R. test of $\rho = 0$			5.965*	4.065*		First-stage F-stat	5.28	6.05
p-value for L.R. test of $\rho = 0$			0.015	0.044				

Notes: Robust standard errors accounting for multiple observations from same mother in brackets. Marginal effects for Univariate Probit models in parentheses. Bivariate Probit and 2SLS LPM estimates use indicator variables for maternal weight status (obese, overweight, underweight) as instrumental variables for son's weight status. All regressions contain binary controls for whether the individual is black or Hispanic.

* significant at 5%; ** significant at 1%

Table 7 - Estimates of Likelihood of College Attendance for Females

(Dependent Variable = 1 if Female attended any college, mean = 35.1%, n = 1,615)

	Univariate Probit		Bivariate Probit		Linear Probablity Model		2SLS LPM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Adolescent Weight Status (relative to non-obese or non-overweight)								
Obese	-0.235		-0.319		-0.07		-0.201	
BMI > 30	[0.109]*		[0.353]		[0.029]*		[0.122]	
Marginal Effect	-(0.080)							
Overweight (incl. obese)		-0.233		-0.634		-0.071		-0.172
BMI > 25		[0.077]**		[0.339]		[0.022]**		[0.090]
		-(0.081)						
Mother's Education (relative to Less than High School education)								
College Graduate	0.914	0.892	0.918	0.848	0.239	0.231	0.245	0.224
	[0.192]**	[0.193]**	[0.192]**	[0.198]**	[0.054]**	[0.054]**	[0.055]**	[0.055]**
	(0.351)	(0.343)						
Some College	0.836	0.833	0.838	0.807	0.211	0.208	0.212	0.205
	[0.143]**	[0.144]**	[0.143]**	[0.146]**	[0.036]**	[0.037]**	[0.037]**	[0.037]**
	(0.312)	(0.311)						
High School Graduate	0.498	0.503	0.498	0.495	0.092	0.092	0.092	0.092
	[0.131]**	[0.132]**	[0.131]**	[0.132]**	[0.027]**	[0.027]**	[0.028]**	[0.028]**
	(0.175)	(0.177)						
Income Quartile (relative to Highest Income Quartile)								
1st (Lowest) Quartile	-0.169	-0.169	-0.16	-0.124	-0.068	-0.068	-0.053	-0.057
	[0.129]	[0.129]	[0.135]	[0.134]	[0.042]	[0.042]	[0.044]	[0.043]
	-(0.059)	-(0.059)						
2nd Quartile	-0.25	-0.248	-0.242	-0.21	-0.092	-0.092	-0.081	-0.084
	[0.118]*	[0.118]*	[0.122]*	[0.121]	[0.040]*	[0.039]*	[0.041]*	[0.040]*
	-(0.086)	-(0.085)						
3rd Quartile	-0.244	-0.24	-0.24	-0.213	-0.088	-0.087	-0.083	-0.082
	[0.115]*	[0.115]*	[0.116]*	[0.116]	[0.038]*	[0.038]*	[0.039]*	[0.039]*
	-(0.083)	-(0.082)						
Demographics								
Mother's Age at Birth	0.128	0.128	0.128	0.122	0.039	0.039	0.038	0.038
	[0.015]**	[0.015]**	[0.015]**	[0.017]**	[0.005]**	[0.005]**	[0.005]**	[0.005]**
	(0.045)	(0.045)						
Birth Order	-0.079	-0.076	-0.079	-0.07	-0.023	-0.022	-0.023	-0.02
	[0.052]	[0.053]	[0.052]	[0.053]	[0.015]	[0.015]	[0.015]	[0.015]
	-(0.028)	-(0.027)						
Pregnant by age 19	-0.418	-0.406	-0.418	-0.386	-0.119	-0.116	-0.119	-0.113
	[0.080]**	[0.081]**	[0.080]**	[0.084]**	[0.023]**	[0.023]**	[0.023]**	[0.023]**
	-(0.143)	-(0.139)						
PIAT Percentiles								
Math Test	0.009	0.008	0.009	0.008	0.003	0.003	0.003	0.003
	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.001]**	[0.001]**	[0.001]**	[0.001]**
	(0.003)	(0.003)						
Reading Test	0.004	0.004	0.004	0.004	0.001	0.001	0.001	0.001
	[0.002]**	[0.002]**	[0.002]**	[0.002]*	[0.000]**	[0.000]**	[0.000]*	[0.000]*
	(0.002)	(0.002)						
Constant	-3.815	-3.773	-3.79	-3.485	-0.614	-0.596	-0.575	-0.541
	[0.349]**	[0.351]**	[0.358]**	[0.458]**	[0.098]**	[0.099]**	[0.104]**	[0.112]**
Pseudo R^2 / R^2	0.203	0.204			0.23	0.24	0.22	0.23
ρ			0.051	0.264				
X^2 value for L.R. test of $\rho = 0$			0.0713	1.848		First-stage F-stat	12.63	16.66
p-value for L.R. test of $\rho = 0$			0.79	0.174				

Notes: Robust standard errors accounting for multiple observations from same mother in brackets. Marginal effects for Univariate Probit models in parentheses. Bivariate Probit and 2SLS LPM estimates use indicator variables for maternal weight status (obese, overweight, underweight) as instrumental variables for daughter's weight status. All regressions contain binary controls for whether the individual is black or Hispanic.

* significant at 5%; ** significant at 1%

Table 8 - Estimates of Likelihood of College Attendance for Males

(Dependent Variable = 1 if Male attended any college, mean = 23.8%, n = 1,551)

	Univariate Probit		Bivariate Probit		Linear Probablity Model		2SLS LPM	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Adolescent Weight Status (relative to non-obese or non-overweight)								
Obese	0.148		-0.623		0.039		-0.104	
BMI > 30	[0.115]		[0.393]		[0.031]		[0.140]	
Marginal Effect	(0.041)							
Overweight (incl. obese)		0.091		-0.298		0.017		-0.074
BMI > 25		[0.079]		[0.362]		[0.020]		[0.089]
		(0.025)						
Mother's Education (relative to Less than High School education)								
College Graduate	0.67	0.667	0.611	0.65	0.193	0.191	0.187	0.191
	[0.190]**	[0.190]**	[0.189]**	[0.188]**	[0.049]**	[0.049]**	[0.049]**	[0.049]**
	(0.220)	(0.218)						
Some College	0.336	0.342	0.336	0.331	0.061	0.062	0.064	0.061
	[0.164]*	[0.164]*	[0.160]*	[0.162]*	[0.033]	[0.033]	[0.033]	[0.033]
	(0.097)	(0.099)						
High School Graduate	0.245	0.247	0.233	0.238	0.032	0.032	0.031	0.031
	[0.153]	[0.152]	[0.149]	[0.151]	[0.024]	[0.024]	[0.024]	[0.024]
	(0.065)	(0.065)						
Income Quartile (relative to Highest Income Quartile)								
1st (Lowest) Quartile	-0.527	-0.524	-0.474	-0.493	-0.142	-0.141	-0.135	-0.135
	[0.141]**	[0.140]**	[0.143]**	[0.145]**	[0.039]**	[0.039]**	[0.040]**	[0.040]**
	-(0.131)	-(0.130)						
2nd Quartile	-0.185	-0.181	-0.149	-0.156	-0.083	-0.083	-0.078	-0.077
	[0.128]	[0.128]	[0.126]	[0.130]	[0.040]*	[0.040]*	[0.040]	[0.040]
	-(0.048)	-(0.047)						
3rd Quartile	-0.057	-0.053	-0.023	-0.036	-0.037	-0.037	-0.031	-0.033
	[0.123]	[0.123]	[0.120]	[0.123]	[0.040]	[0.040]	[0.041]	[0.041]
	-(0.015)	-(0.014)						
Demographics								
Mother's Age at Birth	0.135	0.138	0.137	0.134	0.035	0.035	0.036	0.035
	[0.017]**	[0.017]**	[0.016]**	[0.017]**	[0.004]**	[0.004]**	[0.004]**	[0.004]**
	(0.036)	(0.037)						
Birth Order	-0.166	-0.17	-0.185	-0.177	-0.042	-0.042	-0.046	-0.045
	[0.058]**	[0.058]**	[0.056]**	[0.057]**	[0.012]**	[0.012]**	[0.013]**	[0.013]**
	-(0.044)	-(0.045)						
PIAT Percentiles								
Math Test	0.01	0.01	0.01	0.01	0.003	0.003	0.003	0.003
	[0.002]**	[0.002]**	[0.002]**	[0.002]**	[0.001]**	[0.001]**	[0.001]**	[0.001]**
	(0.003)	(0.003)						
Reading Test	0.005	0.005	0.004	0.005	0.001	0.001	0.001	0.001
	[0.002]**	[0.002]**	[0.002]*	[0.002]**	[0.000]*	[0.000]*	[0.000]*	[0.000]*
	(0.001)	(0.001)						
Constant	-4.191	-4.248	-4.043	-3.996	-0.563	-0.571	-0.564	-0.531
	[0.406]**	[0.408]**	[0.418]**	[0.491]**	[0.092]**	[0.092]**	[0.092]**	[0.098]**
Pseudo R^2 / R^2	0.208	0.208			0.21	0.21	0.2	0.2
ρ			0.437	0.251				
X^2 value for L.R. test of $\rho = 0$			2.651	1.217		First-stage F-stat	5.28	6.05
p-value for L.R. test of $\rho = 0$			0.104	0.27				

Notes: Robust standard errors accounting for multiple observations from same mother in brackets. Marginal effects for Univariate Probit models in parentheses. Bivariate Probit and 2SLS LPM estimates use indicator variables for maternal weight status (obese, overweight, underweight) as instrumental variables for son's weight status. All regressions contain binary controls for whether the individual is black or Hispanic.

* significant at 5%; ** significant at 1%