Bodyweight and Academic Performance in High School

Jeff DeSimone[*] and Edward J. Schumacher[†]

November 2004

Abstract

Although childhood obesity has become a major public health issue in the U.S., little evidence exists regarding the impact of being overweight on academic performance. This paper accordingly examines the relationship between bodyweight and grades, using data on high school students in the 2001 and 2003 Youth Risk Behavior Surveys. We isolate a causal effect of bodyweight by specifying an instrumental variable model in which indicators of self-assessed weight relative to the ideal serve as instruments for actual weight. To control for likely sources of deviations between the instruments and actual bodyweight, we also include proxies for time preference and self-esteem. Empirically, our identification strategy works quite well for boys: extremely large partial correlations between the instruments and observed weight produce standard errors sufficiently small to allow for precise inferences, and little heterogeneity with residuals from the academic performance equations is evident. Estimates indicate that additional pounds, being overweight and obesity each reduce grades by a significant amount that exceeds the corresponding OLS effect. Results are similar for girls, though the validity of our procedure is suspect in some specifications.

---

[*] Department of Economics, University of South Florida, and NBER; E-mail: jdesimon@coba.usf.edu.
[†] Departments of Health Care Administration and Economics, Trinity University; E-mail: eschumac@trinity.edu.
The authors thank (in advance) participants in a 2004 Southern Economics Association session, particularly Richard Smith, for helpful suggestions.

1. Introduction

Childhood obesity is a significant public health problem in the U.S. Rapid increases in the prevalence among children of both obesity and consequent medical conditions that had previously been rare have led to abundant research and policymaking focus on the issue. Defining overweight as having a Body Mass Index (BMI) – weight in kilograms divided by the square of height in meters – of at least 25 (traditionally the 85th percentile) and obesity as a BMI of at least 30 (the 95th percentile), the American Obesity Association estimates that in 2000, 30 percent of youth age 12–19 were overweight and 16 percent were obese, the latter representing a threefold increase in obesity for this age group over the past two decades (Ogden et al. 2002). Accompanying this increase, cases of obesity-related diseases that once were considered adult conditions, such as asthma, diabetes (type 2), hypertension, orthopedic complications and sleep apnea, are being reported more frequently by pediatricians and researchers.

Along with these and other medical conditions as well as a variety of social problems, a potential consequence of being overweight for youth is poor academic performance. The limited previous research on this topic indicates that overweight adolescents have worse grades and ultimately attain fewer years of schooling than non-overweight adolescents. Morrill et al. (1991) found that among 100 14–15 year olds, GPAs were significantly lower among 50 obese students than 50 non-obese students. Gortmaker et al. (1993) found that among 10,039 16–24 year olds who were randomly selected in 1981, overweight women had completed an average of 0.3 fewer years of school by 1988, though there was no analogous relationship for men. In the Connecticut Youth Survey, which sampled 9,957 7th, 9th and 11th graders in 1996, Falkner et al. (2001) found that compared with their average weight counterparts, obese youth of both genders are more likely to consider themselves poor students, obese girls are more likely to be held back a grade, and obese boys are more likely to expect to quit school. Using the same data, Mellin et al. (2002) reported that overweight boys rated their school performance as lower than did non-overweight boys, and both overweight girls and boys reported lower expectations of their educational future than did non-overweight youth. Datar et al. (2004) studied 11,192 first time kindergartners from the 1998 Early Childhood Longitudinal Study, finding that math and reading test scores of kids who were overweight kindergarteners were significantly lower than were those of other kids both in kindergarten and in first grade, though except for boys' scores in kindergarten these differences became insignificant after including socioeconomic and behavioral variables.

With specific regard to high school students, the existing literature on the effect of being overweight on academic performance has two conspicuous weaknesses. The first is its paucity. Not only do the studies above appear to constitute the entire body of previous research on the subject, but furthermore only one of them examined grades, as opposed to self-rated relative performance or expected/actual completed schooling. The sample for that study contained only 100 early high school students, and the corresponding data are well over a decade old. This lack of related research belies the importance of academic performance among high school students, for whom poor grades can lower future earnings profiles through dropout or reductions in the quality of post-secondary educational opportunities.

The second weakness of the previous relevant literature is that it fails to establish the nature of the relationship between bodyweight and academic performance. It is unclear whether being overweight causes grades to fall relative to those of non-overweight students, or rather the negative relationship is simply a manifestation of other confounding factors that simultaneously influence bodyweight and grades. Being overweight might directly reduce achievement by limiting the energy available for studying or through other deleterious health or psychological effects. For instance, obesity-caused sleep apnea is linked to deficits in logical thinking. Alternatively, various attitudinal and psychological factors, such as a lack of either self-esteem or concern about the future, could be responsible for both increasing bodyweight and lowering academic performance. Or, higher-ability students who have better grades might also better understand the adverse health effects of becoming overweight and how to avoid excess weight gain through diet and exercise. On the other hand, students from higher socioeconomic status families might simultaneously perform better in school and consume more calories because of additional resources available to devote towards both educational achievement and food. This question of causality is important for policies designed to reduce the prevalence of overweight and obesity among adolescents. A causal relationship suggests that policies that effectively lower weight among the overweight and obese would carry the additional benefit of increasing academic performance, implying that more funds should be dedicated to such programs than if the relationship is merely an association created by other confounding factors.

In response to these limitations of existing studies, we examine the relationship between bodyweight and grades for high school students in the 2001 and 2003 Youth Risk Behavior Surveys (YRBS), which collects information on typical grades as well as heights and weights of respondents. The main goal is to establish whether this relationship has a causal component through which changes in bodyweight directly impact academic performance. To purge unobserved heterogeneity from the estimated partial correlation between bodyweight and grades and thereby isolate a causal impact, we estimate standard two-stage instrumental variable (IV) models. Our identification strategy is unique in that we use as instruments indicators of self-described bodyweight relative to the perceived "right" weight for the individual. This might seem counterintuitive at first, in that these instruments are not measures that can be exogenously manipulated to alter bodyweight. In fact, deviations between the instruments and actual weight might be produced by some of the same factors, self-esteem and time preferences, that are important sources of the heterogeneity that makes bodyweight endogenous in academic performance equations in the first place. To minimize the impact of these factors and thus legitimize the exclusion of the instruments from the grade equation, we also include proxies for time preference and self-esteem in our vector of controls.

Empirically, this identification strategy is quite effective for boys, as the substantial partial correlation between the instruments and observed weight yields standard errors small enough to allow for precise inferences, and little heterogeneity with residuals from the academic performance equations is apparent. Estimates indicate that weight gain and being overweight and obese each reduces grades by a significant amount that typically is statistically larger than the corresponding OLS effect. Results for girls are similar, though the validity of our procedure is suspect in some specifications, perhaps not surprisingly given well-known gender differences in the importance and impact of body image.

The next section describes the YRBS data that we analyze and outlines our econometric strategy. Section 3 presents the results, establishes the empirical validity of our methodology and offers potential reasons for differences between IV and OLS estimates. We conclude with implications and suggestions for further research in section 4.

2. Data and Methodology

We analyze data from the YRBS, a survey of U.S. high school students occurring biennially since 1991. Because information on grades in school was not collected until 2001, we utilize data from only the 2001 and 2003 waves.

The YRBS focuses on risky behaviors established during youth that result in significant health and social problems during both youth and adulthood (CDC 2004). Each wave employs a three-stage cluster sample design to produce a nationally representative sample of students in grades 9–12. First, 57 primary sampling units (PSUs), which consist of large counties, sub-areas of very large counties, or groups of small, adjacent counties, are chosen from a set of about 1,260. Second, slightly less than 200 schools located in these PSUs are chosen. Selection of both PSUs and schools occurs with probability proportional to school enrollment size, with schools containing large numbers of black and Hispanic students oversampled. Third, one or two classes of a required subject are selected at each school. Participation is anonymous and voluntary, with questionnaires self-administered in classrooms during regular class periods and local parental permission procedures followed. Overall response rates are around two-thirds, with non-response from schools being slightly more likely than non-response from students.

Our academic performance measures are constructed from a question asking students whether they would describe their grades during the past 12 months as mostly A's, mostly Bs, mostly Cs, mostly Ds or mostly Fs. We create two binary dependent variables from their responses: indicators for mostly A’s and for mostly Cs, mostly Ds or mostly Fs, with the former relevant for bodyweight impacts at the high end of the grade distribution and the latter relevant for impacts at the low end. To simultaneously consider weight effects at both ends of the grade distribution and provide estimates in units that are more natural than changes in probabilities of specific grade levels, we also construct a GPA variable. Using a four-point scale, the five grade categories are assigned values of 3.667, 2.833, 2, 1.167 and 0.333. This algorithm imposes symmetry and incorporates the expectation that, e.g., any non-B grades for those selecting "mostly Bs" will on average be worse than B (with the opposite true for "mostly Ds" and the analogous relationships holding by definition for "mostly A's" and "mostly Fs"). Relative to assigning integers from 4 to 0, this procedure generates conservative marginal effects. As long as the measurement error resulting from grades being bunched into a small number of categories is random, it will simply further reduce the magnitudes of the estimated effects. Because Cassady (2001) finds that self-reported GPA values are "remarkably similar to official records" and therefore are "highly reliable" and "sufficiently adequate for research use," we do not expect the use of self-reported rather than observed GPA to bias the results.

We construct bodyweight measures from questions about how tall respondents are and how much they weigh without shoes. The functional form of the weight-grade relationship is unclear, as the absence of an effect at average weights and a decreasing marginal effect at high weights are both plausible. Consequently, we employ a variety of specifications for bodyweight: linear and log forms of weight in pounds while also controlling for height, BMI (kilogram weight divided by meters height squared) and indicators of whether or not the respondent is overweight and is obese, defined as the traditional 85th and 95th percentiles, respectively, and specific to age and gender. Again, we do not anticipate the use of self-reported rather than measured bodyweight (and height) to be problematic, as Brener et al. (2003) show that self-reported height and weight are highly reliable measures of the actual quantities.

Although YRBS information on non-behavioral variables is limited, we are able to control for some basic exogenous sources of variation in weight and GPA, including indicators for each age from 15–18 (age 14 omitted), each grade from 10–12 (grade nine omitted), three race/ethnicities (white, black and Hispanic), urban and suburban residence (rural omitted), one of the two survey years and each state except one, along with real state per capita income and the average state unemployment rate over the previous 12 months. A limitation is that we do not observe the socioeconomic status of the student. Any bias from omission of family-level economic measures should be mitigated by controlling for state income and unemployment rate, but will not necessarily be eliminated because YRBS samples are not necessarily representative within states.

These data can be used in a straightforward manner to estimate the partial correlation between grades and bodyweight, holding constant the exogenous factors listed above. However, the estimated coefficient in an OLS regression is likely to be inconsistent because we do not observe many factors that might simultaneously determine weight and academic performance. The endogeneity problem is reflected by a nonzero correlation between e and ? in the equations

(1) $$W = a_0 + \mathbf{Za_1} + \mathbf{Xa_2} + ?,$$

(2) $$G = ß_0 + ß_1 W + \mathbf{Xß_2} + e$$

where W denotes bodyweight, G represents GPA, the vector **X** includes the previously described exogenous variables, the vector **Z** contains variables that affect weight but not GPA, ? and e are error terms, and the a and ß terms are parameters to be estimated. For the OLS estimate of $ß_1$, the coefficient of interest, to be consistent for the causal impact that bodyweight has on academic performance, e must be uncorrelated with (W, **X**). But this condition is violated when ? and e are correlated: substitution of W from equation 1 into equation 2 shows that ?, which is by definition correlated with W, is a component of e in (2).

Within this framework, a nonzero correlation between e and ?, and thus bias in the estimated $ß_1$, occurs whenever there are determinants of both bodyweight and academic performance that are not held constant in equation 2. Many such factors are conceivable. One example is being oriented towards the present rather than the future: students who heavily discount the future might perform worse in school and have less healthy eating and lifestyle habits because they are less concerned about the longer term consequences of poor grades or being overweight. Another is low self-esteem, which could diminish grades and increase bodyweight through a lack of confidence about having the capability to succeed in school and maintain a healthy weight. Likewise, unmeasured ability, for which we have no suitable proxy, could improve both school performance and awareness of the dangers of, and ways to avoid, becoming overweight. These types of unobserved heterogeneity will produce upward bias, in absolute terms, in the estimate of $ß_1$ by introducing negative spurious correlation between weight and grades. In contrast, our inability to adequately control for parental income and time that are available to the student might bias the estimated $ß_1$ in the opposite direction (downward in magnitude), as increases in these resources will raise both grades and net caloric intake under the presumption that both academic performance and calories are normal goods (but physical activity is not).

To obtain an unbiased estimate of $ß_1$ in the presence of these confounding factors, we employ a standard two-stage IV regression procedure. Both stages are estimated by OLS using YRBS sampling weights, with the first and second stages corresponding to equations 1 and 2, respectively, and standard errors in both equations are adjusted for state-level clustering and heteroskedasticity of unknown form. The vector **Z** in equation 1 contains the instrumental variables that identify the effect of bodyweight on academic performance through their exclusion from equation 2. As is well-known, for this IV procedure to yield a consistent estimate of $ß_1$, the instruments must jointly have a high partial correlation with bodyweight, as manifested by a highly significant coefficient vector $\mathbf{a_1}$, and have no further correlation with GPA, in the sense that e and $\mathbf{Z}$ are uncorrelated, i.e. $\mathbf{Z}$ truly does not belong in equation (2).

As is typical in this type of exercise, finding appropriate instruments is a non-trivial matter, particularly since most YRBS questions pertain to behaviors and self-assessments that are unlikely to be exogenous with respect to bodyweight. Ultimately, we take an unusual approach in choosing to construct our instrument set from one such question, which asks which of five categories respondents would use to describe their weight: very underweight, slightly underweight, about the right weight, slightly overweight or very overweight. We recode responses into a set of four binary indicators, one for each category except "about the right weight," which overidentifies the model and thus allows us to formally test the exclusion restrictions.

As opposed to the more conventional strategy of finding an instrument that is in some sense naturally exogenous to academic performance but might have a relatively weak correlation with bodyweight, our identification strategy emphasizes the strength of the instruments in explaining bodyweight variations and relies on a combination of additional precautions intended to purge any correlations between the instruments and grades (as described below) and the results of formal overidentification tests to provide a convincing validity argument. Conceptually, self-reported and actual bodyweight should be quite strongly correlated, as most students likely realize with some objectivity whether they are under- or overweight. This has the advantage of minimizing standard errors, which with more typical strategies are often too large to allow for any meaningful inference.

Conversely, our methodology is open to criticism because the instruments are not exogenous sources of variation for bodyweight. Two problems in particular could arise that would lead to upward bias in the magnitude of our IV estimate, i.e. an estimate that is “too negative.” The first is that a probable source of deviations between self-reported and actual weight is low self-esteem, which might separately influence academic performance. For example, boys with low self-confidence might incorrectly categorize themselves as underweight, and perhaps even more importantly, many girls with poor self-image are liable to describe themselves as overweight when in fact their weight is healthy or even too low. One way in which we address this problem is to conduct separate analyses by gender, with the expectation that heterogeneity with self-esteem will affect girls more than boys. Another is to include a proxy for self-esteem in the vector of exogenous variables **X**: an indicator that the student exercised to lose weight or avoid gaining weight during the past 30 days. Callaghan (2004) reviews an extensive literature that reports strong associations between exercise and self-esteem, Kirkcaldy et al. (2002) find that high school students who exercised regularly had higher self-esteem, and Mueller et al. (1995) show that 14–19 year olds who are concerned about overeating had low levels of both exercise and self-esteem.

Moreover, we empirically examine whether our instruments are endogenous with respect to academic performance, in the sense that their joint variation explains changes in grades even when the partial correlation with bodyweight is factored out, because of self-esteem or any other form of unobserved heterogeneity. Specifically, we conduct standard overidentification tests in which, after estimating the grade equation (2) by IV, the estimated residual is regressed on all the right hand side variables in (1), i.e. (1, **Z**, **X**), with a relatively high R-square indicating that the instruments cannot be jointly excluded from the grade equation. The test statistic is the sample size multiplied by R-square, which has a chi-square distribution with three degrees of freedom (the number of excluded instruments less one).

A second potential issue with our IV strategy is that, even if overidentification tests affirm its validity, whether it factors out effects of time preferences is unclear. For example, students who accurately report being very overweight might not be concerned with either their weight or their grades because they greatly discount the future. In that case, the IV estimates will be inflated because they at least partially incorporate time preference effects that are correlated with actual bodyweight. To alleviate this problem, we also include two proxies for time preference in the vector **X**, an indicator of any previous smoking and an estimate of the number of cigarettes smoked during the past 30 days. The latter is constructed by assigning midpoints to categorical responses to questions on days smoked and average number of cigarettes smoked on days smoked.[1] Evans and Montgomery (1994) and Fuchs (1982) showed that smoking behavior at early ages is highly related to rates of time preference.[2]

3. Results

We first discuss results for boys. Table 1 shows estimates for the identifying instruments, the self-described weight indicators, in the first stage bodyweight equations. Bodyweight is represented by the weight measures described in the previous section except for BMI, for which results are virtually identical to those for weight given height and are therefore not shown. As displayed under the instrument names, two percent of males consider themselves very underweight and another 16 percent describe themselves as slightly underweight. Meanwhile, only 24 percent consider themselves at least at least slightly overweight and only three percent

[1] The categories for days smoked are 0, 1–2, 3–5, 6–9, 10–19, 20–29 and all 30, while those for cigarettes smoked per day are less than 1, 1, 2–5, 6–10, 11–20, and more than 20, with a value of 30 assigned for the last category. Cigarette use might directly impact academic performance, but as cigarette use tends to have a relaxing effect on smokers and deleterious long-term effects of smoking are unlikely among this age group, in all likelihood any direct effect is small and perhaps even positive for current smoking and negligible for previous smoking.
[2] Results are consistent with our interpretation of these variables as proxies for the corresponding unobserved factors: in the academic performance equations, exercise enters positively while the two cigarette variables each enter negatively, with very high significance levels in each case.

describe themselves as very overweight, both of which underestimate the actual overweight and obesity prevalences of 31 and 16 percent, respectively.

Other than the very underweight indicator, which has a counterintuitive sign in the obese equation and is only marginally significant in the overweight equation, the individual instruments enter the weight equations as predicted and with very high significant levels. Boys describing themselves as slightly overweight weigh 36 pounds, or 23 percent more (based on the exact calculation for an indicator value change in the column 2 log equation), than those who consider themselves to be at the right weight, and are 55 percentage points more likely to be overweight and 38 percentage points more likely to be obese. Analogous differences between those describing themselves as very overweight and being at the ideal weight are 78 pounds (52 percent), 73 percentage points for overweight and 79 percentage points for obesity. F-statistics for tests of instrument joint significance are extremely large and significant at all conventional levels, which is the main benefit to using this identification strategy.

Table 2 presents estimates of the effect of bodyweight on academic performance for each combination of the four bodyweight measures for which first stage equation estimates were shown in Table 1 and the three previously described grade variables. The OLS estimates in columns 1 and 5 uniformly indicate a significant negative partial correlation between bodyweight and academic performance. The magnitudes of these estimates are nontrivial: e.g. in panel (b), implied elasticities at the mean are roughly one-third for the indicators of A and below B grades, and being overweight (obese) lowers the probability of A and above-C grades by about 15 (20) percent. In GPA terms, though, the estimated effects are rather small: a doubling of bodyweight lowers GPA by only about 0.2 points, and becoming overweight or obese reduces GPA by even smaller amounts.[3] Given the dearth of existing evidence on the subject, obtaining these results alone represents research progress. But the real question of interest is causality, for which, as demonstrated in the previous section, OLS estimates are not necessarily relevant.

To address the causality issue, we turn to the IV estimates in columns 2 and 6 of Table 2. These are always larger in magnitude than the corresponding OLS estimates. The exogeneity tests in columns 3 and 7 – in which the coefficient is the difference in the IV and OLS estimates, the standard error is the square root of the difference in the IV and OLS variances, and the test statistic is the t-ratio – show that the discrepancy between the IV and OLS estimates is statistically significant for all bodyweight specifications in the below B grade and GPA equations. The overidentification test statistics in columns 4 and 8 provide strong evidence that our IV estimator is consistent, as p-values are never below 0.2 and rarely below 0.5. Thus, we rely on the estimates from IV, rather than OLS, to make causal effect interpretations.

Given our arguments motivating IV estimation in the first place, neither the preference for IV over OLS nor the finding that causal effects are underestimated by OLS is surprising. In particular, while the proxies we include to control for variations in time preference and self-esteem will mitigate any impacts of these two potentially important sources of negative bias in OLS models, only the IV model avoids the possibility of positive bias from omitting controls for parental resources. On the other hand, not observing ability would presumably impart negative bias on OLS estimates if unmeasured ability not only increases school performance but also

[3] The relationship between estimated effects on probabilities of A and below B grades and those on GPA is b(GPA) = .833[b(A) + b(CDF)] + R, where b(.) is the coefficient for the corresponding grade measure and R > 0. This is because using our transformation from letter grade categories to GPA values, GPAs for those whose grade falls from A to below A, or at least B to below B, declines by .833 (at least). The R term is positive because grades might fall by more than one letter (e.g. from A to C or B to F) and decreases in grades within the below B range (e.g. from C to F) are captured by b(GPA) but not by b(A) or b(CDF).

makes students more cognizant of the risks and ways to keep from becoming overweight. Our results suggest that the impact of such ability bias on the OLS coefficients is small.

Consequently, the most likely explanation for the IV estimates exceeding those from OLS is seemingly a positive reverse causal effect that operates through a correlation between family socioeconomic status and academic performance: higher income families both devote more resources to the education of their children and spend more money on food. This is consistent with results we obtain for the state unemployment rate and per capita income variables. When included without the other to minimize collinearity, each enters both equations in the expected direction (negative for unemployment, positive for income), with high significance in the bodyweight equations. Furthermore, in state-level regressions in which only state and year indicators are held constant, per capita income has large positive effects on both BMI and grades that are marginally significant (at the .15 and .07 levels for BMI and GPA, respectively) and also has a highly significant negative impact on our exercise measure.

Further inspection of the IV results in Table 3 yields several notable conclusions. First, comparison of panels (a) and (b) shows that whether a constant or decreasing marginal effect of additional weight on grades is assumed makes no difference. Second, weight has significant effects at both ends of the grade distribution, with those at the lower end being slightly, but not significantly, larger than those at the upper end. Third, the estimated (semi-) elasticities reveal large effects on the likelihoods of achieving an A and at least a B. A 10 percent increase in weight decreases the former by 7 percent and the latter by 8 percent, while becoming overweight (obese) decreases the probability of A's by 33 (41) percent and of at least Bs by 38 (48) percent. Fourth, these translate to smaller, but economically significant, GPA impacts. A 10 percent weight gain reduces GPA by two percent (.05 points), and becoming overweight or obese reduces GPA by 9–11 percent (between .24 and .30 points). In sum, gaining weight, regardless of whether it results in crossing the overweight or obesity threshold, reduces academic performance by an amount that is neither unreasonably large nor inconsequentially small.

Table 3 shows the results of various sensitivity analyses in which we disaggregate the original sample into two groups and compare estimates from each. For ease of exposition we present the specification that uses the overweight indicator and GPA, but results for all other bodyweight and grade measures are nearly identical. First stage equation estimates and significance levels, which are also suppressed, are similar to those in Table 2.

The two upper panels each divide the sample by an exogenous factor. Panel (a) compares whites with nonwhites, motivated by child obesity rates being substantially higher for the latter. The impact of being overweight is roughly 50 percent larger for whites, which is consistent with unmeasured factors being more important, and causing a higher overweight prevalence, for nonwhites. Since conventional wisdom holds that the population shift into suburbia is an important reason for growing child obesity rates, panel (b) compares suburban residents with those in urban and rural locations. Estimated effects are similar for each, which is consistent with the lack of significant effects of urbanicity in our main sample first stage regressions.

The two lower panels each divide the original sample by one of the two unobserved factors for which we explicitly control in the regressions. Panel (c) compares past 30 day exercisers with non-exercisers. The presumption that the former have higher self-esteem, or at least more of some factor that is positively related with academic performance, is borne out by the fact that the mean GPA is slightly higher among exercisers even though overweight prevalence is about two and a half times as high. The estimated effect is about 50 percent larger for those with low self-esteem, which seems sensible given the above observation, i.e. self-esteem reduces the negative impact of weight on grades. Panel (d) compares current smokers and nonsmokers, with the former assumed to have higher future discount rates. The comparative sample means support this interpretation, as smokers have much lower GPAs, despite no obvious reason for smoking to directly diminish academic performance among an age group that by definition cannot yet be experiencing detrimental consequences of chronic smoking, and a slightly higher overweight prevalence despite the suppressant impact of smoking on weight. Accordingly, effects of being overweight are around 50 percent smaller among smokers.

A broader conclusion of Table 3 is that the estimated negative and significant impact of bodyweight on academic performance holds across a variety of different subgroups. This, and the consistent failure to reject overidentification tests (usually with large p-values), provides further support for our identification strategy and hence the interpretation of our IV estimates as causal effects.

Finally, Table 4 displays results for girls in models that specify GPA as the academic performance measure. We must modify our instrument set by omitting the two underweight indicators, which cause substantial overidentification problems when included. Because of the strength of the two overweight indicators, as indicated in columns 1 and 2, high joint significance of the instruments is maintained (column 3). This again produces small standard errors that allow for more precise inferences than typically possible using IV methods. But column 7 reveals that even reducing the number of exclusion restrictions to two fails to eliminate overidentification issues sufficient to temper any conclusions drawn from the results. In particular, though the overidentification test statistic is only significant at the .01 level in one of four regressions, it is significant at the .10 level in two other regressions and at a slightly weaker level in the fourth. This occurs despite neither identifying instrument ever being significant in the auxiliary GPA residual regressions that are estimated to conduct the overidentification tests.

The negative sign of the slightly overweight indicator in these auxiliary regressions indicates a residual negative correlation between considering oneself overweight and GPA, which suggests that our self-esteem proxy is inadequate for girls. This is not surprising in the sense that, e.g., Furnham et al. (2002) find that self-esteem was associated with body dissatisfaction for girls but not for boys among 16–18 year olds. In results not shown, we find that overidentification problems are even more severe for the grade A attainment indicator, with the slightly overweight indicator entering the auxiliary overidentification regression negatively and significantly, but are not an issue for the below grade B indicator. This implies that self-classification as slightly overweight reflects low self-esteem that reduces the incidence of very high, but not slightly above average, academic performance. Interestingly, column 1 shows that despite this, the slightly overweight indicator is a very strong predictor of actual bodyweight, including being overweight or obese.

Columns 4 through 6 provide OLS, IV, and exogeneity test estimates for the impact of bodyweight on GPA among girls. As for boys, the IV estimates exceed those from OLS in magnitude, and this difference is statistically significant for all bodyweight specifications except pounds in linear form. Estimated magnitudes are also similar to those for boys. A 10 percent increase in weight reduces GPA by .04 points, or just over one percent. Being overweight decreases GPA by .20 points (seven percent), while the effect of being obese is roughly another 50 percent larger.

4. Conclusion

This paper investigates the nature of the relationship between bodyweight and academic performance among high school students. We identify the causal impact on grades of weight changes and being overweight or obese using an IV model that specifies indicators of self-assessed weight relative to the ideal as instruments for observed weight. To account for the possibility that factors driving the differences between the instruments and actual bodyweight are inherently important reasons that weight satisfaction could influence academic performance separately from effects of actual weight, we also include proxies for time preference and self-esteem in our set of explanatory variables. Our identification strategy is particularly effective for boys, yielding precise estimates and no evidence that the instruments exert any influence on grades beyond that which works through observed bodyweight. The estimates show that gaining weight and being overweight or obese significantly decreases grades by an amount exceeding that indicated by the corresponding OLS effect. Although results are similar for girls, the legitimacy of our identification strategy is questionable in some models.

Given the almost complete absence of information on interactions between bodyweight and academic performance, our study contributes to the growing academic literature on overweight and obesity simply by showing the robustness of the strong negative partial correlation between the two quantities in recent nationally representative data on high school students. More importantly, we ascertain that the true causal impact of weight on grades is likely to be greater in magnitude than this negative correlation. Academic performance declines thus represent a short-run detriment of adolescent overweight and obesity that is not captured by health-related costs and potentially translates to substantial long-run costs, once effects on productivity, possibly through decreased quantity or quality of schooling, are considered.

A straightforward implication is that cost-benefit analyses of policies designed to address adolescent weight problems should incorporate the expected benefit of increased academic performance among overweight and obese students who lose weight because of such policies. This affects not only public funds provided to schools targeted towards improving nutrition education, the quality of foods served at meals, health and physical education curricula and the variety of extracurricular opportunities involving physical activity, but also simple behavioral changes initiated by parents or the adolescents themselves to exercise more often or eat more healthily. Explicit health benefit threshold levels necessary for such measures to pass cost-effectiveness tests might be much lower after positive academic performance effects are taken into account.

An important question that our research fails to address regards the mechanism through which the causal effect of bodyweight on academic performance operates. One possibility is that the impact occurs indirectly through health problems from which overweight students are more likely to suffer. Another is that carrying additional bodyweight literally reduces the energy available to allocate towards schoolwork. A third is that being overweight brings about social isolation, through avoidance or ridicule, which leads to psychological problems that impair school performance. More information on this issue would clearly help policymakers, school administrators, teachers, parents and students choose the optimal mix of policies to reduce the academic performance burden associated with being overweight or obese.

References

Brener, N. D., T. Mcmanus, D. A. Galuska, R. Lowry and H. Wechsler, "Reliability and Validity of Self-reported Height and Weight among High School Students," *The Journal of Adolescent Health*, April 2003, 32(4), 281–287.

Callaghan, P., "Exercise: A Neglected Intervention in Mental Health Care?" *Journal of Psychiatric & Mental Health Nursing*, August 2004, 11(4), 476–483.

Centers for Disease Control and Prevention (CDC), *Youth Risk Behavior Survey* data files, Atlanta: Division of Adolescent and School Health, 2002 and 2004.

Datar, Ashlesha, Roland Sturm and J.L. Magnabosco, "Childhood Overweight and Academic Performance: National Study of Kindergartners and First-graders," *Obesity Research*, January 2004, 12(1), 58–68.

Evans, William N., & Edward Montgomery, "Education and Health: Where There's Smoke There's an Instrument." NBER Working Paper 4949, 1994.

Falkner, Nicole H., Dianne Neumark-Sztainer, Mary Story, Robert W. Jeffery, Trish Beuhring and Michael D. Resnick, "Social, Educational, and Psychological Correlates of Weight Status in Adolescents," *Obesity Research*, January 2001, 9(1), 32-42.

Fuchs, Victor R., "Time Preference and Health: an Exploratory Study," in Victor R. Fuchs (Ed.), *Economic Aspects of Health*, Chicago: University of Chicago Press, 1982.

Furnham, Adrian, Nicola Badmin and Ian Sneade, "Body Image Dissatisfaction: Gender Differences in Eating Attitudes, Self-esteem, and Reasons for Exercise," *Journal of Psychology: Interdisciplinary & Applied*, November 2002, 136(6), 581–596.

Gortmaker, Steven L., Aviva Must, James M. Perrin, Arthur M. Sobol and William H. Dietz, "Social and Economic Consequences of Overweight in Adolescence and Young Adulthood," *New England Journal of Medicine*, 30 September 1993, 329(14), 1008–1012.

Kirkcaldy, B. D., R. J. Shephard, R. G. Siefen, "The Relationship between Physical Activity and Self-image and Problem Behaviour among Adolescents," *Social Psychiatry & Psychiatric Epidemiology*, 2002, 37(11), 544–550.

Mellin, Alison E., Dianne Neumark-Sztainer, Mary Story, Marjorie Ireland and Michael D. Resnick, "Unhealthy Behaviors and Psychosocial Difficulties among Overweight Adolescents: The Potential Impact of Familial Factors," *Journal of Adolescent Health*, August 2002, 31(2), 145–153.

Morrill, Correen M., Judy N. Leach, William C. Shreeve, Muriel R. Radebaugh et al., "Teenage Obesity: an Academic Issue," *International Journal of Adolescence & Youth*, 1991, 2(4), 245–250.

Mueller, Cynthia, Tiffany Field, Regina Yando, Jeff Harding et al., "Under-eating and Over-eating Concerns among Adolescents," *Journal of Child Psychology & Psychiatry*, September 1995, 36(6), 1019–1025.

Ogden, Cynthia L., Katherine Flegal, Margaret Carroll and Clifford L. Johnson, "Prevalence and Trends in Overweight Among US Children and Adolescents, 1999–2000," *JAMA*, 9 October 2002, 288(14), 1728–1732.

Table 1: Effects of Instruments on Bodyweight for Boys
(n = 11,152)

	(1) Weight (lb.) (µ = 162.5, s = 37.2)	(2) Log(weight)	(3) Overweight (µ = .312)	(4) Obese (µ = .155)
Very underweight (µ = .019)	-12.38 (3.055)	-.1141 (.0260)	-.0456 (.0284)	.0539 (.0257)
Slightly underweight (µ = .162)	-14.55 (.6217)	-.1002 (.0039)	-.1119 (.0084)	-.0174 (.0075)
Slightly overweight (µ = .209)	35.56 (1.221)	.2066 (.0066)	.5453 (.0143)	.3820 (.0172)
Very overweight (µ = .030)	78.25 (4.068)	.4196 (.0173)	.7271 (.0190)	.7936 (.0185)
F-statistic for joint significance	401.4	537.3	694.8	567.2

Entries are coefficients from OLS regressions of the observed bodyweight measure in the column heading on the variables listed in the text and the self-described weight indicators listed in the left column, with standard errors adjusted for heteroskedasticity and state-level clustering in parentheses. All F-statistics have p-values of less than .0001. Weighted sample means and standard deviations are indicated by µ and s, respectively.

Table 2: Effect of Bodyweight on Grades for Boys
(n = 11,152)

	(1) OLS	(2) IV	(3) Exog.	(4) OID	(5) OLS	(6) IV	(7) Exog.	(8) OID
	(a) Bodyweight (pounds) (µ = 162.5, s = 37.2)				(b) Log (bodyweight)			
Mostly A's	-.0007	-.0010	-.0003	1.951	-.0815	-.1618	-.0803	1.765
(µ = .235)	(.0002)	(.0004)	(.0003)	[.583]	(.0315)	(.0611)	(.0523)	[.622]
	[.000]	[.007]	[.348]		[.010]	[.008]	[.125]	
		-.69				*-.69*		
Mostly Cs, Ds, Fs	.0011	.0017	.0007	1.444	.1218	.2945	.1727	1.861
(µ = .367)	(.0002)	(.0003)	(.0003)	[.695]	(.0315)	(.0586)	(.0493)	[.602]
	[.000]	[.000]	[.013]		[.000]	[.000]	[.000]	
		.75				*.80*		
GPA	-.0019	-.0029	-.0011	2.996	-.2145	-.4932	-.2788	4.161
(µ = 2.64, s = .78)	(.0003)	(.0005)	(.0004)	[.392]	(.0449)	(.0958)	(.0846)	[.245]
	[.000]	[.000]	[.014]		[.000]	[.000]	[.001]	
		-.18				*-.19*		
	(c) Overweight (µ = .312)				(d) Obese (µ = .155)			
Mostly A's	-.0351	-.0780	-.0429	.462	-.0514	-.0968	-.0453	2.207
(µ = .235)	(.0112)	(.0280)	(.0256)	[.927]	(.0209)	(.0397)	(.0338)	[.531]
	[.002]	[.005]	[.094]		[.014]	[.015]	[.180]	
		-.33				*-.41*		
Mostly Cs, Ds, Fs	.0524	.1389	.0865	1.158	.0738	.1756	.1018	3.322
(µ = .367)	(.0107)	(.0245)	(.0220)	[.763]	(.0179)	(.0350)	(.0301)	[.345]
	[.000]	[.000]	[.000]		[.000]	[.000]	[.001]	
		.38				*.48*		
GPA	-.0975	-.2356	-.1381	1.586	-.1250	-.3005	-.1755	2.855
(µ = 2.64, s = .78)	(.0196)	(.0412)	(.0363)	[.663]	(.0316)	(.0593)	(.0502)	[.415]
	[.000]	[.000]	[.000]		[.000]	[.000]	[.000]	
		-.09				*-.11*		

Columns 1, 2, 5 and 6 show coefficients from least squares regressions of the grade variable in the row heading on the bodyweight measure in the panel heading and the additional factors listed in the text, using OLS in 1 & 5 and IV in 2 & 6. Standard errors adjusted for heteroskedasticity and state-level clustering are in parentheses, p-values are in brackets, and elasticities (semi-elasticities) are in italics in columns 2 and 6 of panels a and b (c and d). Columns 3 and 7 are exogeneity tests in which the coefficient is the difference between the corresponding IV and OLS estimates, the standard error is the square root of the difference in the IV and OLS variances, and the test statistic is the t-ratio. Columns 4 and 8 show chi-squared overidentification test statistics, formed by multiplying the sample size by the R-square from the regression of the IV grade equation residuals on all exogenous variables, and corresponding p-values with degrees of freedom equal to three (one less than the number of excluded instruments). Weighted sample means and standard deviations are indicated by µ and s, respectively.

Table 3: Effects of Being Overweight on GPA for Boys by Race, Urbanicity, Self-Esteem and Time Preference

(1) OLS	(2) IV	(3) Exog.	(4) OID	(5) OLS	(6) IV	(7) Exog.	(8) OID
(a) Race							
White (n = 5,442) μ(GPA) = 2.70, μ(OW) = .290				Nonwhite (n = 5,710) μ(GPA) = 2.51, μ(OW) = .359			
-.1054	-.2646	-.1591	1.752	-.0758	-.1719	-.0962	1.733
(.0300)	(.0607)	(.0528)	[.625]	(.0280)	(.0425)	(.0320)	[.630]
[.000]	[.000]	[.003]		[.007]	[.000]	[.003]	
(b) Urbanicity							
Suburban (n = 5,644) μ(GPA) = 2.66, μ(OW) = .315				Urban/rural (n = 5,508) μ(GPA) = 2.62, μ(OW) = .309			
-.1162	-.2551	-.1389	.161	-.0796	-.2270	-.1474	5.879
(.0301)	(.0655)	(.0582)	[.984]	(.0236)	(.0490)	(.0430)	[.118]
[.000]	[.000]	[.017]		[.001]	[.000]	[.001]	
(c) Self-esteem							
Exercised (n = 5,725) μ(GPA) = 2.65, μ(OW) = .439				Did not exercise (n = 5,427) μ(GPA) = 2.63, μ(OW) = .179			
-.1180	-.1952	-.0773	.983	-.0585	-.2976	-.2391	.527
(.0226)	(.0473)	(.0415)	[.805]	(.0387)	(.0647)	(.0519)	[.913]
[.000]	[.000]	[.063]		[.130]	[.000]	[.000]	
(d) Time preference							
Smoker (n = 2,904) μ(GPA) = 2.33, μ(OW) = .319				Nonsmoker (n = 8,248) μ(GPA) = 2.75, μ(OW) = .310			
-.0736	-.1654	-.0918	.660	-.0935	-.2419	-.1483	3.387
(.0386)	(.0790)	(.0689)	[.883]	(.0203)	(.0405)	(.0350)	[.336]
[.057]	[.036]	[.183]		[.000]	[.000]	[.000]	

Columns 1, 2, 5 and 6 show coefficients from least squares regressions of GPA on the overweight indicator and the additional factors listed in the text, using OLS in 1 & 5 and IV in 2 & 6. Standard errors adjusted for heteroskedasticity and state-level clustering are in parentheses and p-values are in brackets. Columns 3 and 7 are exogeneity tests in which the coefficient is the difference between the corresponding IV and OLS estimates, the standard error is the square root in the difference in the IV and OLS variances, and the test statistic is the t-ratio. Columns 4 and 8 show chi-squared overidentification test statistics, formed by multiplying the sample size by the R-square from the regression of the IV grade equation residuals on all exogenous variables, and corresponding p-values with degrees of freedom equal to three (one less than the number of excluded instruments). Weighted sample means for GPA and the overweight indicator are given by μ(GPA) and μ(OW), respectively.

Table 4: Effect of Bodyweight on GPA for Girls
(n = 11,762)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	First stage instrument effects			Second stage bodyweight effects			
	Slightly over	Very over	Joint sig.	OLS	IV	Exog.	OID
Bodyweight (lb.) (μ = 132.8, s = 27.4)	23.40 (.8134)	58.87 (2.844)	478.0	-.0025 (.0004) [.000]	-.0027 (.0005) [.000] *-.12*	-.0002 (.0002) [.361]	4.644 [.031]
Log (bodyweight)	.1742 (.0059)	.3850 (.0171)	510.1	-.2859 (.0521) [.000]	-.3854 (.0700) [.000] *-.13*	-.0995 (.0467) [.033]	2.765 [.096]
Overweight (μ = .213)	.3510 (.0153)	.7257 (.0282)	696.9	-.1170 (.0240) [.000]	-.1994 (.0349) [.000] *-.07*	-.0824 (.0253) [.001]	1.971 [.160]
Obese (μ = .079)	.1251 (.0114)	.5374 (.0327)	139.7	-.1531 (.0386) [.000]	-.2869 (.0761) [.000] *-.10*	-.1339 (.0655) [.041]	14.72 [.000]

Columns 1 and 2 display coefficients of the instruments indicated in the column headings in the first stage OLS regression of the weight variable listed in the row heading. Column 3 gives the F-statistic for joint significance of these two instruments in the regression. Columns 4 and 5 show coefficients from least squares regressions of GPA on the weight variables in the row headings and the additional variables listed in the text, using OLS in 4 and IV in 5. Standard errors adjusted for heteroskedasticity and state-level clustering are in parentheses, p-values are in brackets, and elasticities (semi-elasticities) are in italics in column 5 for the bodyweight (indicator) measures. Column 6 is the exogeneity test in which the coefficient is the difference between the corresponding IV and OLS estimates, the standard error is the square root of the difference in the IV and OLS variances, and the test statistic is the t-ratio. Column 7 shows chi-squared overidentification test statistics, formed by multiplying the sample size by the R-square from the regression of the IV GPA equation residuals on all exogenous variables, and corresponding p-values with degrees of freedom equal to one (one less than the number of excluded instruments). Weighted sample means are .311 for slightly overweight, .046 for very overweight and 2.88 for GPA.