I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15612

Does College Selectivity Reduce Obesity? A Partial Identification Approach

Giorgio Brunello
Dimitris Christelis
Anna Sanz-de-Galdeano
Anastasia Terskaya

SEPTEMBER 2022

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15612

Does College Selectivity Reduce Obesity? A Partial Identification Approach

Giorgio Brunello
University of Padova and IZA

Dimitris Christelis
University of Glasgow, CSEF and CFS

Anna Sanz-de-Galdeano
University of Alicante and IZA

Anastasia Terskaya
University of Barcelona and IEB

SEPTEMBER 2022

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Does College Selectivity Reduce Obesity? A Partial Identification Approach*

We use data from the National Longitudinal Study of Adolescent to Adult Health to investigate whether the quality of tertiary education -measured by college selectivity- causally affects obesity prevalence in the medium run (by age 24-34) and in the longer run (about 10 years later). We use partial identification methods, which allow us, while relying on weak assumptions, to overcome the potential endogeneity of college selectivity as well as the potential violation of the stable unit treatment value assumption due to students interacting with each other, and to obtain informative identification regions for the average treatment effect of college selectivity on obesity. We find that attending a more selective college causally reduces obesity, both in the medium and in the longer run. We provide evidence that the mechanisms through which the impact of college selectivity on obesity operates include an increase in income, a reduction in physical inactivity and in the consumption of fast food and sweetened drinks.

JEL Classification: I14, I12, I26, C14
Keywords: obesity, college selectivity, partial identification

Corresponding author:
Giorgio Brunello
Department of Economics and Management
University of Padova
via del Santo 33
35123 Padova
Italy
E-mail: giorgio.brunello@unipd.it

* Sanz-de-Galdeano acknowledges financial support from Project PID2021-124237NB-I00 financed by MCIN/ AEI /10.13039/501100011033/ and by FEDER Una manera de hacer Europa. Terskaya acknowledges financial support from Project PID2020-120589RA-I00 financed by MCIN.

1. Introduction

In 2019, 42.8 percent of the US population was obese (OECD Health Statistics). Obesity is not only a health problem, as it also affects economic outcomes such as productivity, wages and skill formation (see for instance Böckerman *et al.* (2019); Brunello *et al.* (2009); Cawley (2004); Cawley (2013); Cawley (2015)). The negative correlation between education and obesity has been well documented: each additional year of schooling is associated with a two-percentage point reduction in the probability of being obese. Evidence of a causal protective effect of education on obesity is, however, weak at best (see Galama *et al.* (2018); Lochner (2011)).

Empirical research on the effects of education on obesity has mainly focused on years of schooling, a quantitative measure of education. Less has been done to investigate whether the quality of education matters.[1] Education quality -measured for instance by college admission standards or by instructional expenditures per student (see Brewer *et al.* (1999); Dale and Krueger (2002))- may affect obesity because it influences income and earnings (Brewer *et al.* (1999); Dale and Krueger (2002); Hoekstra (2009)), years of education, cognition, marital status and the quality of peers (Fletcher and Frisvold (2011)).

The small literature focusing on the effects of education quality on health outcomes includes Ross and Mirowsky (1999), who estimate the effects of attending a more selective college on physical functioning and self-rated health, and Fletcher and Frisvold (2014), who use data from siblings followed for over 50 years in the Wisconsin Longitudinal Survey and show that graduating from a more selective college reduces overweight and tobacco use.

[1] The importance of education quality has been recognized in the growth literature. Hanushek and co-authors have shown that —conditional on years of schooling— education quality, measured by attained test scores, significantly affects economic growth. In addition, years of schooling have no impact on growth once we condition on attained test scores (Hanushek and Kimko (2000); Hanushek and Wossmann, (2012)).

In the paper closest to ours, Fletcher and Frisvold (2011) use data from the National Longitudinal Study of Adolescent Health (Add Health) and present "suggestive" evidence[2] that college selectivity reduces obesity both during college and almost a decade later. This evidence, as the one provided in the papers cited above, is based on selection on observables, and, as the authors admit, is not robust to selection on unobservables. In particular, the latter needs to be less than 0.3 times the amount of the former to eliminate the estimated reduction in obesity induced by higher college selectivity.

Our paper contributes to this small literature in two ways. First, we use Add Health - the same data source used by Fletcher and Frisvold (2011) - and obtain *causal* estimates of the effects of college selectivity on obesity after college by implementing partial identification (PI) methods that locate non-parametrically these estimates in informative identification regions (Manski (1989) and (1997); Manski and Pepper (2000)). PI has several important advantages over selection on observables: i) it addresses the problem of selection on unobservables, both time-invariant and time-varying; ii) it relies on weak, and thus credible, assumptions (the price to pay being an increase in uncertainty); iii) it is completely transparent about how each assumption affects the results; iv) it makes no assumptions about preferences, functional forms, expectation formation, optimality of decision-making, or the joint distribution of errors and observable characteristics; v) it identifies the average treatment effect (ATE); vi) it is robust to the violation of the stable unit treatment value assumption (SUTVA), which is likely to be a problem in the context of education due to peer effects.

Second, from a policy perspective it is important to understand whether the causal effects of college selectivity persist or fade away as time goes by. We therefore estimate not only the medium-term *causal* effects of college

[2] The term "suggestive" is used by the authors.

selectivity on obesity (when individuals are aged 24 to 34) but also these effects in the longer term (when they are aged 33 to 44).[3]

We find that the causal effect of attending a more selective college on individual obesity 2 to 12 years after the typical graduation age (22) amount to *at least* a 2.0 percentage points reduction, which corresponds to a 7.7 percent of mean obesity (26.2%), a quantitatively important effect. The upper bound of the effect is less than half in absolute value than the estimate provided by Fletcher and Frisvold (2011) using the same dataset (-4.9), and significantly lower than the OLS estimate (-7.2 percentage points). About ten years later, when individuals are 11 to 22 years older than the typical age of graduation from college, the effect of attending a more selective college on obesity is still negative but slightly smaller (at least -1.6 percentage points, or a 4.4 percent reduction with respect to mean obesity).

These results indicate that attending a more selective college causally reduces obesity both in the medium and in the longer run. We consider physical activity (the lack of which is measured by hours spent watching TV and by a binary variable for no physical activity) and the consumption of fast food and sweetened drinks as mechanisms that could explain our findings. We find that attending a more selective college improves physical activity both in the medium and in the longer run. There is also evidence of a reduction in fast food and sweetened drinks consumption, especially in the medium run. Higher selectivity also leads to higher income, which facilitates the adoption of healthier eating habits, including those that we cannot observe in our data (e.g., the consumption of fruit and vegetables).

[3] Using selection on observables, Fletcher and Frisvold (2011) focus on the medium term by using the fourth wave of Add Health, and Fletcher and Frisvold (2014) look at the longer run in a small sample drawn from the Wisconsin Longitudinal Survey.

The remainder of the paper is organized as follows: Section 2 describes the data, Section 3 introduces our methodology and Section 4 presents the results. Conclusions follow.

2. Data and Variables.

2.1. The Add Health Dataset

We use data from the National Longitudinal Study of Adolescent to Adult Health (Add Health). Add Health is a longitudinal school-based study of a nationally representative sample of the U.S. adolescents who were in 7th to 12th grades in 1994-95 and have been followed until 2016-18. The first wave of Add Health (1994-95) was drawn from a stratified sample of 142 schools. Within each school and grade, approximately 17 males and 17 females were randomly selected for a detailed *in-home* interview. This *in-home* sample was subsequently interviewed in 1996 (Wave II), 2001-02 (Wave III), 2008 (Wave IV), and 2016-19 (Wave V). The *in-home* survey contains detailed information on self-reported health and health related behaviours, such as eating habits and physical activity. Moreover, starting from Wave II, anthropometric measurements were collected, and we use them to construct objective measures of the body mass index (BMI), obesity indicators, and waist circumference.

In Wave IV, when respondents were between 24 and 34 years old, they were asked if they had received a bachelor's degree and the institution from which they had acquired the degree. Add Health assigned an Integrated Postsecondary Education Data System (IPEDS) code to each institution and matched college information in the year of graduation from the IPEDS database. Using these codes, data containing information on college selectivity from the Mobility Report Card: The Role of Colleges in Intergenerational Mobility (Chetty *et al*. (2017)) were matched to Add Health respondents.[4]

[4] For further details, see the Add Health documentation available at https://cdr.lib.unc.edu/downloads/bc386p66s and https://cdr.lib.unc.edu/downloads/kh04dv36s?locale=en.

Since we are interested both in the medium and in the long-term effects of college selectivity on obesity, we use outcomes measured in Waves IV (age: 24-34) and V (age: 33-44) of Add Health. Hence, assuming that the usual age at college graduation is 22, we observe individuals between 2 and 12 years after graduation (the medium term) and between 11 and 22 years after graduation (the longer term).

Our estimation sample is constructed by applying the following filters. We start with a sample of 18,910 Wave I respondents with valid age and cross-sectional sample weights. We select 14,790 respondents who were interviewed again in Wave IV and further restrict the sample to 4,734 respondents who had completed college by Wave IV. Furthermore, we eliminate 922 respondents who were not matched to the IPEDS database or whose colleges do not offer a four-year degree. Finally, we retain the 3,578 individuals for whom we have both valid college selectivity information and the score in the Peabody Picture Vocabulary Test (PPVT), which are required for our identification strategy, as we explain below. For Wave V, the analysis is based on 2,856 respondents who have valid cross-sectional weights. Table A.1 in the Appendix reports the summary statistics for our estimation sample.

2.2. College Selectivity

Following Deming *et al.* (2015) and Chetty *et al.* (2017), our measure of college quality is the Barron's (2009) college selectivity index. Based on this index, Add Health classifies colleges into four categories: (1) elite, (2) highly selective, (3) selective, (4) non-selective. We define a binary variable "highly selective or elite college" that takes the value one if a college is classified as elite or highly selective, and zero otherwise. Table A.1 in the Appendix shows that 20 percent of the respondents in our sample have attended a highly selective or elite college.

2.2. Outcomes

2.2.1. Obesity

We are interested in the effect of college selectivity -as a measure of college quality- on obesity and health related behaviours. The body mass index (BMI) is defined as weight (kg.) divided by squared height (cm.), and obesity is defined as having a BMI of 30 or more. In Add Health, BMI is computed using anthropometric measures of weight and height collected in both waves IV and V. While anthropometric measures were collected from most respondents in Wave IV, they were collected in Wave V only for respondents who gave their consent to an in-home examination, which included taking measurements and collecting a blood sample. Because of this, BMI based on anthropometric measures is available in Wave V for only 1,548 respondents in our estimation sample.[5] Table A.2 in the Appendix reports that 26.2% and 36.6% of the respondents in our sample were obese in waves IV and V, respectively. The average BMI was 27.4 in Wave IV and 29.0 in Wave V.

An alternative measure of obesity is abdominal obesity, which is defined as having a waist circumference (in cm) above 88 cm for females and 102. cm for males.[6] Average waist circumference in our sample is 94.1 in Wave IV, and 93.7 in Wave V (see Table A.2 in the Appendix).

2.2.2 Health-Related Behaviours

Fast food consumption

Respondents were asked the following question in both waves IV and V: "*In the past 7 days, how many times did you eat food from a fast-food restaurant, such*

[5] Protocols for the collection of anthropometrics in Wave IV are available at https://addhealth.cpc.unc.edu/wp-content/uploads/docs/user_guides/Wave_IV_Cardiovascular_and_anthropometric_documentation.pdf, and in Wave V at https://addhealth.cpc.unc.edu/wp-content/uploads/docs/user_guides/WaveVAnthropometricsUserGuide.pdf.
[6] https://www.cdc.gov/healthyweight/assessing/index.html.

as McDonald's, Burger King, Wendy's, Arby's, Pizza Hut, Taco Bell, or Kentucky Fried Chicken or a local fast-food restaurant?". Using the answers to this question, we define a binary variable that takes the value one if respondents ate food from a fast-food restaurant more than once a week, and zero otherwise.

Sweetened drinks consumption

Using the answers to the following question in waves IV and V: "*In the past 7 days, how many regular (non-diet) sweetened drinks did you have? Include regular soda, juice drinks sweetened tea or coffee, energy drinks, flavored water, or other sweetened drinks?",* we define the variable "number of sugary drinks is the past 7 days".

Physical inactivity

We define the binary variable "no bouts of physical activity" in both waves using the following questions: *(1) "In the past 7 days, how many times did you bicycle, skateboard, dance, hike, hunt, or do yard work?", (2) "In the past 7 days, how many times did you roller blade, roller skate, downhill ski, snowboard, play racquet sports, or do aerobics?", (3) "In the past 7 days, how many times did you participate in gymnastics, weight lifting, or strength training?", (4) "In the past 7 days, how many times did you participate in individual sports such as running, wrestling, swimming, cross-country skiing, cycle racing, martial arts, or in strenuous team sports such as football, soccer, basketball, lacrosse, rugby, field hockey, or ice hockey?", (5) "In the past 7 days, how many times did you play golf, go fishing or bowling, or play softball or baseball?", (6) "In the past 7 days, how many times did you walk for exercise?".* The variable takes the value one if the respondent has not participated in any of these activities, and zero otherwise.

Hours spent watching TV

Using the following question asked in both waves: *"In the past seven days, how many hours did you watch television or videos, including VHS, DVDs or music videos?"*, we define the variable "number of hours spent watching TV in the past week".

3. Methodology[7]

Attending a more selective college is the outcome of decisions by both individuals and colleges and therefore cannot be treated as exogenous with respect to individual obesity. In particular, the unobservable characteristics that affect the attendance of a highly selective college may be correlated with the individual propensity to obesity.

For example, students attending a highly selective college may have higher earnings capacity, which could lead to better dietary habits and more exercising. They may also have higher human capital, which can be correlated with both college choice and dietary and exercise habits. And they could also be, on average, in better physical health, which again could affect both college choice and obesity later in life. Time varying un-observables may also matter (for instance, a shock that affects family income may affect both health and college choice).

Previous literature has addressed the problem that selection into college type is non-random by comparing students of selective colleges and students admitted to these colleges who attended instead non-selective institutions (Dale and Krueger (2002)); by explicitly modeling high school students' choice of college (Brewer *et al.* (1999)); by comparing twin pairs (Behrman *et al*. (1996)); or by exploiting discontinuities in admission rules and a regression discontinuity design (Hoekstra (2009); Zimmerman (2019)). These studies have focused mainly on labor market returns to college selectivity.

[7] The discussion in this section closely follows Christelis *et al.* (2020).

In the absence of plausible sources of exogenous variation in our data, we estimate the causal effect of college selectivity on individual obesity by using partial identification (henceforth PI), a methodology introduced by Manski (1989), (1990), (1994). PI methods bound non-parametrically the average treatment effect (henceforth ATE), that is, they locate the ATE in an identification region instead of producing a point estimate. In what follows, we give a brief overview on how we implement PI in our context and provide additional details in the online Appendix.

PI methods apply bounds to the counterfactual, and thus unobservable, average potential outcomes across sample units. For example, to estimate obesity prevalence when no individual in the sample has attended a more selective college, we need to calculate, keeping all other factors constant; i) the obesity prevalence in those who did not attend a more selective college, which is an observed magnitude; ii) the obesity prevalence in those who did attend a more selective college, had they not attended one. This latter term represents the counterfactual outcome. The endogeneity of attendance of a more selective college prevents us from replacing the counterfactual outcome with the observed obesity prevalence of those who did not attend a more selective college. This would have been possible, for example, if we had been conducting a randomized control trial.

We use PI to bound the counterfactual outcome, and thus we also bound the average potential outcome, that is, the obesity prevalence when nobody attends a more selective college. In an analogous fashion, we bound the other potential outcome, namely obesity prevalence when everybody attends a more selective college. Bounding the average potential outcomes also implies bounding their difference, that is, the ATE of interest, or the difference between average obesity when everybody attends a more selective college and when nobody attends such a college, keeping all other relevant factors constant. The ATE upper bound is equal to the upper bound of average obesity when everybody attends a more

selective college minus the lower bound of average obesity when nobody attends a more selective college. On the other hand, the ATE lower bound is equal to the lower bound of average obesity when everybody attends a more selective college minus the upper bound of average obesity when nobody attends a more selective college.

Without any assumptions, one can credibly use as the lower (upper) bound of counterfactual outcomes only the minimum (maximum) feasible values of the outcome, namely zero and one, given that obesity is a binary variable. As expected, these extreme values result in very wide and thus uninformative identification regions (we provide additional details on bounds using no assumptions in section A.2 of the online Appendix). Hence, to informatively bound unobserved counterfactual outcomes, PI uses assumptions that are, as we will see, much milder than those used in OLS and exogenous IV-based methods.

The first assumption is that of monotone treatment response (MTR henceforth; see Manski (1997), which in our context states that obesity prevalence is, on average, weakly decreasing in attending a more selective college, keeping everything else constant. This average weakly monotonic relationship holds for potential outcomes, and thus is unverifiable. MTR is, however, a reasonable assumption in our context because attending a more selective college is, on average, likely to reduce the likelihood of obesity in most individuals (Ross and Mirowsky (1999); Fletcher and Frisvold (2011) and (2014)). First, higher quality education makes it easier for individuals to be better informed about the negative consequences of obesity (Lochner (2011)); second, it affects health behaviors by increasing cognitive ability and the ability to process relevant information (Cutler and Lleras Muney (2010)); last but not least, it is associated with higher earnings (Dale and Krueger (2002)), which in turn improves eating and exercise habits.

We emphasize that the MTR assumption posits that the weakly negative relationship between attending a more selective college and obesity holds *on*

average. In other words, while there could be individuals who attend a more selective college but gain weight due to idiosyncratic reasons, the assumption postulates that these individuals are a minority in the population. Importantly, MTR posits a weakly negative association, and thus it is consistent with attendance of a more selective college to have, on average, no effect on obesity whatsoever.

Hence, this assumption does not guarantee that college selectivity negatively affects individual obesity, as will be shown in Table 1 below. To illustrate how the MTR assumption works, we would like to bound the counterfactual obesity prevalence in those who did not attend a more selective college, had they attended one. Under MTR, an upper bound of this counterfactual outcome is the actual obesity prevalence in those who did not attend a more selective college. Given that observed obesity prevalence is less than one, the MTR assumption leads to narrower identification regions compared to those obtained under no assumptions.

Under MTR, the upper bound of the average potential outcome, that is, obesity prevalence when everybody attends a more selective college is the observed obesity prevalence in the whole sample. This happens because in the whole sample we observe both individuals who have attended a more selective college and individuals who have not. Hence, MTR implies that in a situation in which, counterfactually, everybody has attended a more selective college, obesity prevalence would have been weakly lower than the observed one. Correspondingly, under MTR, the lower bound of the other average potential outcome, namely obesity prevalence when nobody has attended a more selective college, is again the observed obesity prevalence in the whole sample. This happens because MTR implies that in a situation in which nobody attended a more selective college, obesity prevalence would have been higher than in the actually observed sample, in which at least some individuals have attended a

more selective college. We further discuss the MTR assumption in section A.3 of the online Appendix.

The above discussion indicates that the upper bound of the ATE under MTR is equal to zero, as both potential outcome bounds are equal to the observed prevalence of obesity in the whole sample. This result is consistent with Manski (1997), who shows that the ATE identification region under MTR always includes zero.[8] Importantly, this upper bound shows the maximum change in the prevalence of obesity when everybody attends a more selective college, compared to when nobody does. Since in our context MTR implies a weakly negative treatment effect, we can equivalently state that the ATE upper bound under MTR denotes the minimum reduction in the prevalence of obesity when everybody attends a more selective college compared to when nobody does. Therefore, the ATE upper bound is the most important one we estimate, as it represents the most conservative acceptable estimate of the ATE. On the other hand, the ATE lower bound under MTR denotes by at most how much obesity prevalence changes from a situation in which nobody attends a more selective college to when everybody does.

The second assumption we use is the monotone instrumental variable (MIV henceforth) one, which was introduced by Manski and Pepper (2000) and serves to further narrow ATE identification regions. This assumption posits that, given the value of the treatment, the instruments are weakly monotonically associated with the average potential outcome, namely the prevalence of obesity. More in detail, it posits that both average potential outcomes are weakly decreasing for those with observed higher MIV values. This is a much milder assumption than that of exogeneity in a standard IV setup, and it is made even milder by the allowed possibility (under weak monotonicity) that the average potential outcomes (i.e., obesity prevalence) do not differ in subsamples defined by

[8] Manski (1997) discusses a setting in which the MTR assumption implies a weakly positive treatment effect, and thus zero is the lower bound of the ATE identification region. In our context, the treatment effect is assumed to be weakly negative, and thus zero is the upper bound of the identification region.

different values of the score. Crucially, the MIV assumption identifies the ATE and not the local average treatment effect (i.e., the effect for those whose attendance of a more selective college changes due to a change in the instrument). In fact, this assumption is silent about its association with the treatment. Given that MIV bounds the ATE, we can interpret PI results as applying to the whole population (see section A.4 of the online Appendix for further discussion of MIV).

We select as MIV the score in the Peabody Picture Vocabulary Test (PPVT hereafter), an age-specific test used to assess verbal ability and receptive vocabulary[9] administered in high school. We posit that individuals who have a higher cognitive ability (as indicated by a higher score in the test) are weakly less likely, on average, to be obese, given the value of the treatment (college selectivity). This assumption is consistent with previous empirical studies on the relationship between cognition and obesity. Cutler and Lleras Muney (2010), for instance, show that knowledge and measures of cognitive ability explain about 30 percent of the positive relationship between education and favorable health outcomes, known as the education gradient (see also Galama *et al.* (2018); Lochner (2011)). Smith *et al.*, (2011) review the evidence on the relationship between adiposity and cognitive performance and conclude that obesity is associated with cognitive deficits in children, adolescents and adults. One reason for the importance of cognition is that a higher cognitive ability makes it weakly easier to realize the negative consequences of obesity in their fullness.[10]

[9] The PPVT (Dunn and Dunn (2007)) is often considered as a measure of verbal intelligence and scholastic aptitude and is strongly correlated with the Wechsler Intelligence Test (Anderson and Flax (1968)) and the Armed Forces Qualifying Test. Amin *et al.* (2022) use maternal education as an MIV in their recent PI investigation of the effect of schooling on cognition later in life.

[10] The relationship between cognitive ability and health is reflected at the genetic level. For example, intelligence has a positive genetic correlation with longevity and a negative genetic correlation with various indicators of suboptimal lifestyle and physical health, such as hypertension, type 2 diabetes, body mass index (BMI) and smoking status (Deary *et al.*, 2019).

Importantly, as for MTR, the MIV assumption needs to hold only on average, that is, it allows for the possibility that a minority of individuals with a high PPVT score are more obese than some of their counterparts with a lower score. This assumption (like the MTR) refers to potential outcomes, and thus is unverifiable (as is the case of the exclusion restriction in standard IV estimates). In Add Health, as shown in Table A.3, there is a negative association between obesity and the PPVT score, which is precisely what the MIV assumption entails. Although this is no proof that the MIV assumption holds because this estimated negative association refers to observed data and not to potential outcomes, the observed association points to the same direction as the MIV assumption.

The identification of treatments such as college selectivity requires that the possible violation of the stable treatment unit value assumption (SUTVA) be considered. Such violation can happen because the obesity status of an individual's college peers -who attend the same selective/non-selective college- could also influence whether this individual becomes obese or not. The existing literature on peer effects in obesity indicates that they are indeed relevant both in high school (Trogdon *et al.* (2008), Brunello *et al.* (2020)) and in college (Yakusheva *et al.* (2011), (2014)). The influence of an individual's peers on his/her obesity status can operate through the social interactions that arise naturally from attending the same college, and can be due, among other things, to students trying to emulate each other or fit-in socially via their physical appearance (Carrell *et al.* (2011)).

A violation of the SUTVA would be a serious problem for identification, as it would invalidate results obtained using standard estimation methods such as OLS, panel data, propensity score matching or IV. An advantage of our approach is that we can address the SUTVA violation problem using PI. In particular, we can use the assumption of reinforcing interactions (Manski (2013)), which states that non-individualist treatment responses (which occur

when the SUTVA is violated) reinforce each other in terms of the outcome response. In our context, this implies that, for example, all students in a selective/elite college influence each other with respect to obesity in such a way that the prevalence of obesity is weakly lower than the prevalence that would have been obtained if the same students had been attending a less selective college. The reinforcing interactions (RI) assumption is plausible because the relatively lower prevalence of obese peers in very selective colleges is likely to induce individuals, through emulation and the desire to fit in socially, to take actions that reduce the risk of becoming obese. This happens over and above any negative influence other than through social interactions that a more selective education has on the likelihood to become obese. Analogously, in less selective colleges, reinforcing interactions are likely to increase obesity prevalence compared to very selective colleges, as the higher observed obesity prevalence in these colleges is likely to make obesity less of a problem when trying to fit in socially, and thus less of a taboo.[11] Once more, this happens over and above any positive influence other than through social interactions that a less selective education has on the likelihood to become obese.

In general, the violation of the SUTVA increases the width of the identification regions, thus making results more uncertain. This can be seen easily when evaluating the mean potential outcome under a treatment value for the subsample of the individuals who are observed having this particular value (e.g., the mean potential outcome of attending a more selective college for those who actually do so). When the SUTVA holds, this mean potential outcome can be consistently estimated using its sample analogue, that is, the observed mean outcome of those who are in this subsample. There is thus no need to examine what happens to those not in this subsample. On the other hand, when the

[11] Carrell *et al.* (2011), find that a one standard deviation increase in the high school fitness scores of all peers in a group raises the individual's college fitness score by 0.165 of a standard deviation, a large effect. Brunello *et al.* (2020), find that, in the short-run, a one standard deviation increase in peers' average BMI polygenic scores –which are a good predictor of peers' average BMI– raises the probability of obesity for females by 2.8% points. No significant effect is found for males. In the long-run, however, the social-genetic effect fades away.

response is non-individualistic, the treatment values of individuals outside this subsample potentially affect the outcome of those inside it. This is crucial because the definition of the mean potential outcome involves everybody (i.e., inside and outside of this subsample) taking the same treatment value. Hence, when the SUTVA is violated, the mean potential outcome in this subsample will be generally different when everybody outside this subsample has the same treatment value compared to when individuals outside this subsample have a variety of treatment values. Since the latter is what generally happens in actuality, one cannot use the observed mean outcome in this subsample to estimate the corresponding mean potential outcome. In other words, when the SUTVA is violated, one needs to bound a quantity for which one could obtain a consistent point estimate when response is individualistic.

However, there are some cases for which the assumption of reinforcing interactions produces the same bounds as those under MTR. This can be easily seen, for example, for the upper bound of the average potential outcome (i.e., the prevalence of obesity) when everybody goes to a more selective college. This upper bound, under MTR, is equal to, as already discussed, the observed prevalence of obesity in the whole sample. Given that the definition of the unobserved potential outcome entails everybody going to a very selective college, in this situation there are more students in such colleges compared to the number actually observed. Hence, the reinforcing interactions assumption implies that in the former situation each student can be influenced into not being obese by more individuals attending very selective colleges than in actuality. This should make the prevalence of obesity when everybody attends a very selective college weakly lower than the observed one. In other words, even when the SUTVA is violated, it is still the case that the observed obesity prevalence in the whole sample is an upper bound of the average potential obesity when everybody attends a more selective college, provided one uses the reinforcing interactions assumption.

Analogously, when the SUTVA is violated, the lower bound of the average potential outcome of a less selective college education, is, under the reinforcing interactions assumption, the same as the one obtained under MTR with the SUTVA holding, that is, the observed prevalence of obesity in the whole sample. This is so because the prevalence of obesity when everybody attends less selective colleges should be weakly higher than the actual one, as reinforcing social interactions when all students attend less selective colleges make it more likely for any student to be obese compared to when only some students go to such colleges. The bounds produced by the reinforcing interactions assumption are further discussed in section A.5.1 of the online Appendix.

In our empirical analysis, we always contrast the PI estimates with those obtained under exogenous treatment selection (ETS henceforth), which posits that respondents receiving different treatments are not systematically different from one another. In other words, ETS implies that college selectivity is as good as randomly assigned to individuals. Under ETS, the ATE is equal to the difference in the observed prevalence of obesity, conditional on the two different values of the college selectivity variable.

In conclusion, we stress that the advantages of using PI are considerable, as discussed in further detail in section A.6 of the online Appendix. First, it addresses the problem of selection on un-observables, both time-invariant and time-varying. Second, it relies on weak, and thus credible, assumptions, namely MTR, MIV, and that of reinforcing interactions, in contrast to the strong assumptions needed in OLS, panel data or IV estimation using exogenous instruments. Third, it is completely transparent about how each assumption affects results. Fourth, it makes no additional assumptions about preferences, functional forms, expectation formation, optimality of decision-making, or the joint distribution of errors and observable characteristics. Fifth, it identifies the average treatment effect (ATE) rather than the local average treatment effect

(LATE), while allowing for unlimited heterogeneity of the treatment effect across sample units. Sixth, it is robust to the violation of the stable unit treatment value assumption (SUTVA), which is likely to be a problem in our context due to peer effects.

The price to pay for these advantages is an increase in the uncertainty affecting results, as PI can sometimes lead to identification regions that are wide. As Manski (1994) notes, however, the point identification obtained using the assumptions of exogeneity in OLS and panel data may give false confidence about empirical results, because the reduction in uncertainty is obtained through strong and untestable assumptions that might not hold in the real world.

4. Results

4.1. OLS Results

Table A.4 in the Appendix shows the OLS estimates of the association between college selectivity, obesity and health behaviors. In each regression we control for age, gender, race, and PPVT. The results in columns (1) to (3) indicate that college selectivity is associated with a 7.2 percentage points reduction in the prevalence of obesity (both for waves IV and V), with a 1.2 to 1.3 points reduction in BMI, which corresponds to a 4.5 to 4.7 percent decline with respect to the mean, and with a reduction of 2.7 to 3.8 centimeters in waist circumference. The results in columns (4) to (7) suggest that individuals who went to a more selective college engage less in unhealthy behaviors associated with a higher risk of obesity. For instance, highly selective college graduates are less likely to consume fast food and sugary drinks.

4.2. PI Results

4.2.1. Statistical considerations

Our magnitude of interest, for which we estimate identification regions, is the ATE of attending a more selective college on the probability of being obese.

Each set of results shown below corresponds to a particular combination of assumptions. For each such combination, we report the ATE lower and upper bounds (or, in the case of exogenous treatment selection, ETS, the point estimate), as well as the 95% confidence intervals (CIs) for the lower and upper ATE bounds.

Following Kreider and Pepper (2007) and Manski and Pepper (2009), we show bias-corrected bootstrap estimates of the bounds (CIs are the same for both bias-corrected and bias-uncorrected bound estimates). The difference between bias-uncorrected and bias-corrected estimates is non-trivial only for the results that use the MIV assumption, as the latter uses minimization and maximization operators (this is further discussed in section B.4 of the online Appendix), which can induce bias when combined with the bootstrap.[12]

On the one hand, the estimate of the bias can be volatile, which is why Efron and Tibshirani (1993) recommend using bias-uncorrected estimates, especially when the bias is small (they suggest a 25% threshold) compared to the standard error. This volatility can also occasionally make the bias-corrected estimates fall outside the CI, which is not the case with bias-uncorrected estimates. When this problem occurs, a more reliable estimate of the ATE bound can be obtained by using the CI boundary.

On the other hand, we uniformly obtain bias-corrected estimates of the most important ATE bound, namely, the upper bound (denoting the minimum reduction in the probability of obesity induced by attendance of a more selective college), that are closer to zero, and thus more conservative, than bias-uncorrected ones. Hence, to be on the conservative side, we consider bias-corrected estimates as the preferred ones.

4.2.2. PI estimates for obesity

[12] As Manski and Pepper ((2009), p. S211) point out, Imbens-Manski (2004) confidence intervals are not applicable when using MIVs.

In Table 1 we report our findings for the probability of being obese. Results for waves IV and V are shown in Panel A and B respectively. We first examine the ETS estimates, which are equal to those obtained by running a weighted OLS regression on a constant and a dummy variable denoting the attendance of a more selective college. For Wave IV we find that this attendance reduces the prevalence of obesity by 7.81 percentage points, a precise effect, as documented by the relatively narrow CIs.

PI relaxes the assumption of exogeneity of attending a more selective college, under which ETS estimates are consistent. When using no assumptions whatsoever, and thus bounding counterfactual outcomes with the minimum and maximum feasible outcome values (equal to 0 and 1, respectively), we predictably obtain very wide and uninformative ATE identification regions: the lower bound implies that attending a more selective college decreases the probability of being obese by 38.3 percentage points, while the upper bound denotes an increase of this probability by 61.7 percentage points. We report these results only to illustrate the point that one cannot draw any useful conclusions about causal effects without any further identifying assumptions.

When we introduce the MTR assumption, the upper ATE bound becomes zero, while the lower ATE bound remains uninformative. Hence, MTR on its own does not allow us to reject the null hypothesis that attendance of a more selective college has no effect on the probability of being obese. We also note that the CI of the ATE lower bound has both its upper and lower bound equal to zero, which implies that the constraint that the ATE is non-positive, imposed by the MTR assumption, is binding in at least 95% of the bootstrap runs.

When we add the MIV to the MTR assumption, the ATE upper bound declines below 0, and is equal to about -2.0 percentage points (95% CI: -6.9, -1.1). In other words, attending a more selective college reduces the probability of being obese by at least about 2 percentage points. This result is quantitatively important, given that the prevalence of obesity in our sample is about 26.2

percent. On the other hand, the ATE lower bounds remain uninformative. In the absence of an informative lower bound one could use the ETS estimate of -7.81 percentage points as a substitute, under the assumption that it overestimates (in absolute value) the ATE.

To gauge how a lower ATE bound of -2 percentage points compares with the OLS results shown in Appendix Table A.4, we first note that the OLS point estimate is equal to -7.2, very close to the ETS estimate, and outside the CI of the ATE upper bound under MTR+MIV. Second, since the ATE lower bound under MTR+MIV is uninformative, the OLS estimate is included in the ATE identification region under MTR+MIV. We note, however, that this ATE identification region is robust to treatment selection when the maintained assumptions are valid, while the OLS estimate is not, and thus it is not clear how to interpret it.

Importantly, we note that the ATE upper bounds using MTR and MTR+MIV are, as already discussed in Section 3, robust to the violation of the SUTVA, when one uses the RI assumption instead of MTR.[13] This implies that the most important result, namely by at least how much attending a more selective college reduces the prevalence of obesity, is robust to the violation of the SUTVA due to students influencing other fellow students' eating habits. On the other hand, as discussed in appendix A.5.1, the ATE lower bound when the SUTVA is violated is equal to -100 percentage points, that is, completely uninformative.

When comparing the identifying power of the MTR assumption (or of the RI assumption, for the ATE upper bound under the violation of the SUTVA) with that of the MIV assumption, it is clearly the case that the former assumption

[13] The RI assumption, as discussed in Section 3, states that non-individualist treatment responses reinforce each other in terms of the outcome response. In our context, this implies that all students in a highly selective (less selective) college influence each other with respect to obesity in such a way that the prevalence of obesity is weakly lower (higher) than the prevalence that would have been obtained if the same students had attended a less selective (more selective) college. See section A.5.1 of the online Appendix for further details on the definition of the RI assumption and the bounds it produces.

shrinks the identification region from above by far more than the latter. However, without the MIV assumption it is not possible to obtain an ATE upper bound that is lower than zero, and thus both assumptions are needed to obtain statistically significant results. We also note that the bias-uncorrected ATE upper bound under MTR+MIV is equal to about -5.6 percentage points, and thus much more negative than the bias-corrected one.

The finding that college selectivity causally reduces obesity prevalence could be a temporary effect that fades away in the longer run. We investigate whether this is the case by using data from Wave V and considering individuals who are 11 to 22 years older than the typical age of graduation from college. As reported in Panel B of Table 1, we find that results for Wave V are quite similar to those for Wave IV. The bias-corrected estimate of the ATE upper bound under MTR+MIV is equal to about 1.6 percentage points reduction, showing that the effects of attending a more selective college on obesity are almost unchanged at least a decade after college graduation.

4.2.3. Robustness checks

We examine the robustness of our results by using abdominal obesity, based on waist circumference (the threshold for being considered obese is 102 cm for men and 88 cm for women). The results in Table 2 are in line with those for BMI based obesity. We find that the bias-corrected ATE upper bound under MTR+MIV is about -3.1 percentage points in Wave IV, and 0 percentage points in Wave V. In the latter case, we encounter the aforementioned problem of the bias-corrected estimate falling outside the bias-corrected CI. In this circumstance, the upper boundary of the CI, equal to about -1 percentage point, is a more reliable estimate of the ATE upper bound .

As an additional check, we re-estimate the ATE identification regions using the subsample of individuals in Wave IV for whom we have an objective measurement of obesity in Wave V. Notice that PI estimates are consistent in the presence of sample selection because a maintained assumption in PI

estimation is that of treatment endogeneity for any reason. Our results, shown in Appendix Table A.5, are quite similar to those obtained using the larger Wave IV sample. In particular, the bias-corrected estimate of the ATE upper bound under MTR+MIV in Wave IV is equal to -2.1 percentage points.

4.2.4. Mechanisms

In this section we discuss the effects of attending a more selective college on health behaviors that can affect obesity. We consider: i) eating fast food more than once in the last 7 days; ii) watching TV in the last 7 days; iii) the number of sweetened drinks respondents had in the last 7 days; iv) being physically inactive in the last 7 days. All these behaviors are expected to increase obesity either because they increase calorie intake or because they reduce calorie consumption. We also consider the effects of college selectivity on income (transformed using the inverse hyperbolic sine transformation, which still allows the interpretation of results as semi-elasticities) because a higher income makes it easier to buy healthier foods, which are on average more expensive.

Our estimates are reported in Table 3. Here, we discuss only the ATE upper bounds, which denote the minimum change in outcomes induced by attending a more selective college. These bounds are the most informative once we introduce the MTR assumption, which posits that attending a more selective college weakly decreases the probability of frequently eating fast food, the hours spent watching TV, the number of sweetened drinks, and the probability of being inactive. We consider this assumption to be reasonable, as a more selective education should make individuals more conscious of the negative consequences of unhealthy habits and physical inactivity, as well as lead to higher incomes.

For all outcomes, we select the PPVT score as a MIV, the reasonable assumption being that a higher score - or a higher human capital - is negatively associated with outcomes i)-iv), and positively associated with income. Considering Wave IV, we generally find non-trivial effects of attending a more

selective college on all health behaviors. The ATE upper bounds suggests that the probability of eating fast food more than once is reduced by at least about 1.9 percentage points, watching TV declines by at least about 0.3 hours, the number of sweetened drinks falls by at least 0.4 and the probability of being physically inactive is reduced by at least 2 percentage points. Moreover, a more selective college education increases income by at least 6.5 percent.

For Wave V, our results are generally a bit weaker for outcomes i)-iii), while for physical inactivity we still find a quantitatively important ATE upper bound of -2 percentage points. We also find that a more selective education considerably increases income, by at least 10 percent.

These results indicate that attending a more selective college improves physical activity both in the medium and in the longer run. By increasing calorie consumption, physical activity is expected to reduce obesity. There is also evidence of a reduction in fast food and sweetened drinks consumption, especially in the medium run. The associated reduction in calorie consumption should also reduce obesity. The observed effects of college quality on health behaviors could be both direct and indirect. For instance, we find that higher selectivity increases income, and higher income is typically associated with more physical activity (Kari et al. (2015)).[14] Higher income also facilitates the adoption of healthier eating habits, including those that we cannot observe in our data (e.g., the consumption of fruit and vegetables).[15]

5. Conclusions

We have presented evidence that education quality -measured by college selectivity- causally reduces individual obesity. We have shown that the size of

[14] In contrast, the variation in adult fast-food consumption across income and wealth groups is small (Zagorsky and Smith (2017)).

[15] Low income is associated with lower food expenditures, low fruit and vegetable consumption, and lower-quality diets (Drewnowski *et al.*, 2004). A higher intake of fruit and vegetables typically reduces obesity. See for instance He *et al.* (2004).

the effect -at least a 2.0 percentage points reduction in the medium term and a 1.6 percentage points reduction in the longer term– is far from negligible, although smaller than the non-causal estimates found in the literature.

Our study has several potential important implications. First, finding a causal effect of college selectivity on obesity suggests that, when measuring the returns to an elite college, it is important to consider non only monetary returns but also non-monetary ones. Second, the relationship between education and obesity may be more complex than it is often considered as it involves not only the quantity but also the quality of education. Third, the effects of college quality – measured by college selectivity– on obesity decline in size but do not fade away as time goes by.

In the US, children from low-income families have on average worse health outcomes than those from wealthier families (Case *et. al.* (2002)). Chetty *et. al* (2017) show that access to colleges varies substantially across the income distribution, so children whose parents are in the top of the income distribution are much more likely to attend an Ivy League college than those whose parents are in the bottom of the income distribution. If college quality has a positive impact on health, unequal access to elite or more selective colleges may reinforce health inequalities in the US. Moreover, to the extent that obesity hampers productivity, unequal access to elite colleges may exacerbate socioeconomic inequalities as well. Therefore, increasing access to such colleges for children from low-income families may contribute to the reduction of health and economic inequalities.

References

Amin, V., Behrman, J., Fletcher, J., Flores, C., Flores-Lagunes, A. and H.P. Kohler (2022), "Does schooling improve cognitive abilities at older ages: causal evidence from nonparametric bounds", IZA Discussion Paper No. 15371.

Anderson, D. E. and M. E. Flax. 1968, "A Comparison of the Peabody Picture Vocabulary Test with the Wechsler Intelligence Scale for Children," *Journal of Educational Research*, 62(3),114–116.

Barron's Educational Series, Inc. College Division (2009), *Barron's Profiles of American Colleges*. Barron's Educational Series, Incorporated.

Behrman, J.R., Rosenzweig, M.R. and P. Taubman (1996), "College choice and wages: estimates using data on female twins," *Review of Economics and Statistics*, 78(4), 672-685.

Böckerman, P., Cawley, J., Viinikainen, J., Lehtimäki, T., Rovio, S., Seppälä, I., ... and O. Raitakari (2019), "The effect of weight on labor market outcomes: An application of genetic instrumental variables," *Health Economics,* 28(1), 65-77.

Brewer, D. J., Eide, E. R., and R. G. Ehrenberg (1999), "Does it pay to attend an elite private college?" *The Journal of Human Resources*, 34(1), 104-123.

Brunello, G., and B. d'Hombres (2007), "Does body weight affect wages?: Evidence from Europe," *Economics & Human Biology*, 5(1), 1-19.

Brunello, G., P.-C. Michaud, , and A. Sanz-de-Galdeano (2009), "The rise of obesity across the Atlantic: an economic perspective," *Economic Policy*, 24(59), 551-596.

Carrell, S. E., Hoekstra, M., and J. E. West (2011), "Is poor fitness contagious?: Evidence from randomly assigned friends," *Journal of Public Economics*, 95(7-8), 657-663.

Case, A., Lubotsky, D., and C. Paxson (2002), "Economic status and health in childhood: The origins of the gradient," *American Economic Review*, 92(5), 1308-1334.

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

Cawley, J. (2015), "An economy of scales: A selective review of obesity's economic causes, consequences, and solutions," *Journal of Health Economics*, 43, 244-268.

Chetty, R., Friedman, J. N., Saez, E., Turner, N., and D. Yagan (2017), "Mobility report cards: The role of colleges in intergenerational mobility," NBER Working Paper No. 23618

Christelis, D., D. Georgarakos, T. Jappelli, and M. van Rooij (2020), "Consumption Uncertainty and Precautionary Saving," *The Review of Economics and Statistics*, 102, 148-161.

Cutler, D. M., A. Lleras-Muney (2010), "Understanding differences in health behaviors by education," *Journal of Health Economics*, 29(1), 1-28.

Dale, S. B., and A. B. Krueger (2002), "Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables," *The Quarterly Journal of Economics*, 117(4), 1491-1527.

Deary, I.J., Harris, S.E. and W.D. Hill (2019), "What genome-wide association studies reveal about the association between intelligence and physical health, illness and mortality", *Current Opinion in Psychology*, 27, 6-12.

De Haan, M. (2011), "The Effect of Parents' Schooling on Child's Schooling: A Nonparametric Bounds Analysis," *Journal of Labor Economics*, 29, 859-92.

Deming, D. J., Goldin, C., Katz, L. F., and N. Yuchtman (2015), "Can online learning bend the higher education cost curve?" *American Economic Review*, 105(5), 496-501.

Drewnowski, A. and S.E. Specter (2004), "Poverty and obesity: the role of energy density and energy costs", *The American Journal of Clinical Nutrition*, 79, 1, 6-16.

Dunn, L. M. and L. M. Dunn (2007), "The Peabody Picture Vocabulary Test," Bloomington, MN: NCS Pearson, Inc., 4th ed.

Efron, B. (1982), *The Jackknife, the Bootstrap, and Other Resampling Plans*, Philadelphia: Society for Industrial and Applied Mathematics.

Fletcher, J. M., and D. E. Frisvold (2011), "College selectivity and young adult health behaviors," *Economics of Education Review*, 30(5), 826-837.

Fletcher, J. M., and D. E. Frisvold (2014), "The long run health returns to college quality," *Review of Economics of the Household*, 12(2), 295-325.

Galama, T. J., Lleras-Muney, A., and H. Van Kippersluis (2018), "The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence," NBER Working Paper No. 24225.

Hanushek, E and L. Wossmann (2012), "Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation," *Journal of Economic Growth*, 17, 267–321

Hanushek, E. A., and D. D. Kimko (2000), "Schooling, labor force quality, and the growth of nations," *American Economic Review*, 90(5), 1184–1208.

He, K., Hu, F.B., Colditz, G.A., Manson, J.E., Willett, W.C. and S. Liu (2004), "Changes in intake of fruit and vegetables in relation to risk of obesity and weight gain among middle-aged women", *International Journal of Obesity*, 28, 1569-1574.

Hoekstra, M. (2009), "The effect of attending the flagship state university on earnings: A discontinuity-based approach," *The Review of Economics and Statistics*, 91(4), 717-724.

Imbens, G. and C. F. Manski (2004), "Confidence intervals for partially identified parameters," *Econometrica*, 72(6), 1845–57.

Kari, J, Pekkonen, J, Hirvesalo, H, Yang, X, Hutri-Kahonen, N, Raitakari, O and Tammelin, T (2015), "Income and Physical Activity among Adults: Evidence from Self-Reported and Pedometer-Based Physical Activity Measurement," *Plos One*, 1-15.

Kreider, B., and J. V. Pepper, (2007), "Disability and Employment," *Journal of the American Statistical Association*, 102(478), 432-441.

Lochner, L. (2011), "Non-Production Benefits of Education: Crime, Health, and Good Citizenship," NBER Working Paper No. 16722.

Manski, C. F. (1989), "Anatomy of the selection problem," *Journal of Human resources*, 343-360.

Manski, C. F. (1990), "Nonparametric Bounds on Treatment Effects," *American Economic Review Papers and Proceedings*, 80, 319-323.

Manski, C. F. (1994), "The Selection Problem," in *Advances in Econometrics: Sixth World Congress*, C. Sims ed. Cambridge, UK: Cambridge University Press.

Manski, C. F. (1997), "Monotone Treatment Response," *Econometrica,* 65, 1311-1334.

Manski, C. F. (2013), "Identification of treatment response with social interactions," *The Econometrics Journal*, 16(1), S1-S23.

Manski, C. F., and J. V. Pepper (2000), "Monotone Instrumental Variables, with an Application to the Returns to Schooling," *Econometrica,* 68, 997-1012.

Manski, C. F., and J. V. Pepper (2009), "More on Monotone Instrumental Variables," *Econometrics Journal*, 12, Suppl. (S1), S200-S216.

Ross, C. E., and J. Mirowsky (1999), "Refining the association between education and health: the effects of quantity, credential, and selectivity," *Demography*, 36(4), 445-460.

Smith, E., Hay, P., Campbell, L: and J.NB. Trollor (2011), "A review of the association between obesity and cognitive function across the lifespan: implications for novel approaches to prevention and treatment", *Obesity Reviews*, 12, 740-755.

Trogdon, J. G., Nonnemaker, J., and J. Pais (2008), "Peer effects in adolescent overweight," *Journal of Health Economics*, 27(5), 1388-1399.

Yakusheva, O., Kapinos, K. A., and D. Eisenberg (2014), "Estimating heterogeneous and hierarchical peer effects on body weight using room-mate assignments as a natural experiment," *Journal of Human Resources*, 49(1), 234–261.

Yakusheva, O., Kapinos, K., and M. Weiss (2011), "Peer effects and the freshman 15: Evidence from a natural experiment," *Economics and Human Biology*, 9(2), 119–132.

Zagorsky, Y and Smith, P (2017), "The Association between Socio-economic Status and Adult Fast Food Consumption in the US," *Economics and Human Biology*, 12-24

Zimmerman, S. D. (2019). Elite colleges and upward mobility to top jobs and top incomes. American Economic Review, 109(1), 1-47.

Table 1. PI lower and upper bounds of the effect of college selectivity on obesity. Waves IV and V of Add Health

Assumptions	Estimates		Lower Bound Low 95% CI	Lower Bound Upper 95% CI	Upper Bound Low 95% CI	Upper Bound Upper 95% CI
	Lower bound	Upper bound				
Panel A. Wave IV						
Exogenous treatment selection	-0.0781		-0.1233		-0.0297	
No assumptions	-0.3831	0.6169	-0.4223	-0.3490	0.5777	0.6510
MTR	-0.3831	0.0000	-0.4223	-0.3490	0.0000	0.0000
MTR + MIV (bias-corrected)	-0.3088	-0.0196	-0.3175	-0.2247	-0.0688	-0.0112
MTR + MIV (bias-uncorrected)	-0.2485	-0.0556	-0.3175	-0.2247	-0.0688	-0.0112
Number of observations	3,525					
Panel B. Wave V						
Exogenous treatment selection	-0.0973		-0.1783		-0.0162	
No assumptions	-0.4503	0.5497	-0.4942	-0.4138	0.5058	0.5862
MTR	-0.4503	0.0000	-0.4942	-0.4138	0.0000	0.0000
MTR + MIV (bias-corrected)	-0.3601	-0.0159	-0.4145	-0.2632	-0.0867	-0.0081
MTR + MIV (bias-uncorrected)	-0.3013	-0.0688	-0.4145	-0.2632	-0.0867	-0.0081
Number of observations	1,520					

Note: MTR: monotone treatment response, MIV: monotone instrumental variable (Peabody Picture Vocabulary Test score).

Table 2. PI lower and upper bounds of the effect of college selectivity on abdominal obesity

Assumptions	Estimates		Lower Bound Low 95% CI	Lower Bound Upper 95% CI	Lower Bound Low 95% CI	Upper Bound Upper 95% CI
	Lower bound	Upper bound				
Panel A. Wave IV						
Exogenous treatment selection	-0.0970				-0.1593	
No assumptions	-0.4832	0.5168	-0.5102	-0.4565	0.4898	0.5435
MTR	-0.4832	0.0000	-0.5102	-0.4565	0.0000	0.0000
MTR + MIV (bias-corrected)	-0.3055	-0.0306	-0.3710	-0.2284	-0.0837	-0.0254
MTR + MIV (bias-uncorrected)	-0.3151	-0.0709	-0.3710	-0.2284	-0.0837	-0.0254
Number of observations	3,547					
Panel B. Wave V						
Exogenous treatment selection	-0.1085				-0.1855	
No assumptions	-0.4914	0.5086	-0.5302	-0.4535	0.4698	0.5465
MTR	-0.4914	0.0000	-0.5302	-0.4535	0.0000	0.0000
MTR + MIV (bias-corrected)	-0.4110	0.0000	-0.4275	-0.3228	-0.0683	-0.0096
MTR + MIV (bias-uncorrected)	-0.3444	-0.0578	-0.4275	-0.3228	-0.0683	-0.0096
Number of observations	1,537					

Note: MTR: monotone treatment response, MIV: monontone instrumental variable (Peabody Picture Vocabulary Test score).

Table 3. Mechanisms

Outcomes	Estimates		Lower Bound Low 95% CI	Lower Bound Upper 95% CI	Lower Bound Low 95% CI	Upper Bound Upper 95% CI
	Lower bound	Upper bound				
Panel A. Wave IV						
Eaten fast food (bias-corrected)	-0.3456	-0.0186	-0.4029	-0.2610	-0.0753	-0.0182
Eaten fast food (bias-uncorrected)	-0.3209	-0.0615	-0.4029	-0.2610	-0.0753	-0.0182
Hours spent watching TV (bias-corrected)	-9.8746	-0.2808	-11.2038	-8.3985	-1.0435	-0.1111
Hours spent watching TV (bias-uncorrected)	-9.7400	-0.7407	-11.2038	-8.3985	-1.0435	-0.1111
Number of sweet drinks (bias-corrected)	-8.2882	-0.4022	-9.9039	-6.5639	-1.2377	-0.2565
Number of sweet drinks (bias-uncorrected)	-7.6330	-1.0451	-9.9039	-6.5639	-1.2377	-0.2565
Being physically active (bias-corrected)	-0.1278	-0.0208	-0.1777	-0.0737	-0.0611	-0.0206
Being physically active (bias-uncorrected)	-0.1107	-0.0524	-0.1777	-0.0737	-0.0611	-0.0206
Log of income (bias-corrected)	0.0652	1.5779	0.0748	0.2055	1.3438	1.7568
Log of income (bias-uncorrected)	0.1742	1.4887	0.0748	0.2055	1.3438	1.7568
Panel B. Wave V						
Eaten fast food (bias-corrected)	-0.3262	-0.0039	-0.3828	-0.2405	-0.0565	-0.0092
Eaten fast food (bias-uncorrected)	-0.3037	-0.0496	-0.3828	-0.2405	-0.0565	-0.0092
Hours spent watching TV (bias-corrected)	-11.1460	-0.1734	-13.0552	-9.3666	-1.0683	0.0000
Hours spent watching TV (bias-uncorrected)	-11.1184	-0.5067	-13.0552	-9.3666	-1.0683	0.0000
Number of sweet drinks (bias-corrected)	-5.0991	0.0000	-5.8405	-4.2930	-0.3522	-0.0191
Number of sweet drinks (bias-uncorrected)	-4.8462	-0.2930	-5.8405	-4.2930	-0.3522	-0.0191
Being physically active (bias-corrected)	-0.1003	-0.0198	-0.1552	-0.0491	-0.0581	-0.0119
Being physically active (bias-uncorrected)	-0.0944	-0.0419	-0.1552	-0.0491	-0.0581	-0.0119
Log of income (bias-corrected)	0.1034	1.0242	0.0278	0.2809	0.8270	1.1988
Log of income (bias-uncorrected)	0.1620	1.0132	0.0278	0.2809	0.8270	1.1988

Note: MTR (monotone treatment response) and MIV (monotone instrumental variable using the Peabody Picture Vocabulary Test score as an instrument) assumptions are imposed.

Online Appendix

A. Partial identification[1]

A.1. Bounds on potential outcomes

As in Manski (1997), let every individual i have a response function $y_i(\bullet): D \rightarrow Y$ that maps mutually exclusive and exhaustive treatments $d \in D$ into outcomes $y_i(d) \in Y$. Importantly, the response functions $y_i(\bullet)$ can differ across individuals in arbitrary ways, thus allowing for unlimited response heterogeneity. Let w_i denote the realized treatment received by i, and $y_i \equiv y_i(w_i)$ the associated observed outcome. In our case, the main outcome is obesity, while the treatment variable is a binary variable denoting attendance of a more selective college, with the value one denoting attendance of elite and very selective colleges, and zero attendance of selective and non-selective colleges. To keep notation simple, we omit conditioning expectations of outcomes and probabilities of treatments on observables X, but all results go through also after such conditioning.

Let $y_i(d_1)$ and $y_i(d_2)$ be two possible values of the outcome for individual i as a function of two different levels of college selectivity d_1 and d_2, with $d_2 > d_1$. We would like to estimate the ATE of higher college selectivity on our outcomes, that is,

$$ATE(d_2 - d_1) = E[y(d_2)] - E[y(d_1)] \tag{A.1}$$

Note that the ATE in (A.1) represents the difference in the two mean outcomes, which are both evaluated using all population units while taking the distribution of all other observable and unobservable variables as given (Manski 1997, p. 1322). In our context, these two mean outcomes denote probabilities of being obese under two different values of college selectivity.

By the law of iterated expectations, and given that $E[y(d)|w = d] = E(y|w = d)$, the expected outcome, when the treatment is equal to d, is

$$E[y(d)] = E(y|w = d)P(w = d) + E[y(d)|w \neq d]P(w \neq d) \tag{A.2}$$

where $P(w = d)$ denotes the probability that $w = d$. Note that the term $E[y(d)|w \neq d]$ in the right-hand side of (A.2) is an unobserved counterfactual one. The remaining three terms on the right-hand side of (A.2), however, have sample analogues that are observed in the data. Given that $E[y(d)|w \neq d]$ is unobserved, the unconditional expectation

[1] This Appendix draws from Christelis et al. (2020), and Christelis and Dobrescu (2020).

$E[y(d)]$ is also unobserved. Hence the ATE in (1) is equal to the difference between two average unobserved outcomes, and thus cannot be calculated without further assumptions.

If one assumes that the counterfactual conditional expectation $E[y(d)|w \neq d]$ is equal to the observed one when the treatment received is equal to d, that is, if

$$E[y(d)|w \neq d] = E(y|w = d) \tag{A.3}$$

then from (A.2) it follows that

$$E[y(d)] = E(y|w = d) \tag{A.4}$$

Equation (A.4) states that the unobserved potential outcome under d is equal to the mean outcome when the treatment in fact received is d. As the sample analogue of the latter is observed in the data, one can estimate the unobserved potential outcome $E[y(d)]$, and then the ATE from equation (A.4) as

$$ATE(d_2 - d_1) = E(y|w = d_2) - E(y|w = d_1) \tag{A.5}$$

We refer to the ATE estimate in (A.5) as the one under exogenous treatment selection (ETS henceforth) because it is derived under the assumption that (A.3) holds, which in turn implies that respondents attending colleges that differ in selectivity are not systematically different from one another. In other words, (A.3) implies that selection into treatment is exogenous.

Equation (A.3) is likely to hold in the case of a randomized control trial, in which treatment assignment is indeed exogenous. In observational data, however, (A.3) is unlikely to hold because treatment assignment is typically not random. In our context, attendance of a more selective college might be affected by unobservable variables that also affect obesity. Hence, the latter is likely to differ among population groups defined by attending colleges of different selectivity. This holds for any value d of the college selectivity variable.

Once one rules out the application of (A.3), the problem of estimating the unobservable potential outcome $E[y(d)]$ arises. As a solution, Manski (1990) suggested bounding this outcome from above and below by bounding the counterfactual potential outcome $E[y(d)|w \neq d]$ in (A.2). Let us denote the lower and upper bounds on $E[y(d)]$, computed using a set of assumptions *M*, as $LB^M(d)$ and $UB^M(d)$, respectively. Given

that $LB^M(d) \leq E[y(d)] \leq UB^M(d)$, Manski (1990) points out that equation (A.1) in turn implies that one can bound the ATE using a set of assumptions *M* as follows:

$$LB^M(d_2) - UB^M(d_1) \leq ATE(d_2 - d_1) \leq UB^M(d_2) - LB^M(d_1) \qquad \text{(A.6)}$$

The interval between the lower and the upper bound on the $ATE(d_2 - d_1)$ (which are denoted by $LB^M_{ATE}(d_2 - d_1)$ and $UB^M_{ATE}ATE(d_2 - d_1)$, respectively) is its identification region. Since it is an interval, the ATE is only partially identified.

A.2. Bounds using no assumptions

When calculating the upper and lower bounds on $E[y(d)]$, a natural starting point is to assume that, for any value d of the treatment, the outcome space Y is bounded below and above by two finite values, Y_{min} and Y_{max}, respectively. These values can be used to bound $E[y(d)|w \neq d]$. In our context, and since obesity is a binary variable, we use zero and one as Y_{min} and Y_{max}, respectively.

Given that zero and one are the most conservative feasible bounds on $E[y(d)|w \neq d]$ (which denotes a probability of being obese in our context), and that no other assumptions are used, the resulting identification regions of $E[y(d)]$ and the ATE can be considered as being derived under no assumptions (NA henceforth).

As in Manski (1990), one can replace the counterfactual term $E[y(d)|w \neq d]$ in (A.2) by $Y_{min} = 0$ and $Y_{max} = 1$, and thus bound $E[y(d)]$ from below and above as follows:

$$\begin{gathered} E(y|w = d)P(w = d) + Y_{min}\,P(w \neq d) = E(y|w = d)P(w = d) \\ \leq E[y(d)] \leq \\ E(y|w = d)P(w = d) + Y_{max}P(w \neq d) = E(y|w = d)P(w = d) + P(w \neq d) \end{gathered} \qquad \text{(A.7)}$$

The NA bounds can be readily calculated using their sample analogues, as these are observed in the data. As Manski (1989) points out, taking sample averages leads to consistent estimates of $E(y|w = d)$, $P(w = d)$ and $P(w \neq d)$.

A.3. The MTR assumption

The NA identification regions are typically very wide, and thus uninformative (i.e., they always include zero, as Manski (1990) shows), as one would expect when trying

to draw conclusions by using only the data without any additional assumptions imposed. It is possible, however, to narrow the NA identification region by making further assumptions. The first such assumption is that of monotone treatment response (MTR henceforth; see Manski, 1997). In our context, the MTR assumption implies attending a more selective college has a weakly negative effect on the probability of being obese.

In the case of a weakly negative treatment response, the MTR assumption as discussed by Manski (1997) states that for all sample units i, and for any two treatment values $d_1 \in D$ and $d_2 \in D$ such that $d_2 > d_1$,

$$y_i(d_2) \leq y_i(d_1) \tag{A.8}$$

Importantly, (A.8) holds irrespective of the treatment actually received, and for all sample respondents. Given that at each point in time one observes only one outcome for every sample respondent, one cannot test for the validity of (A.8) in isolation using the data at hand. As already discussed, however, there are various reasons, also supported by considerable evidence, why one would expect attendance of a more selective college to have a weakly negative effect on the likelihood of becoming obese.

In practice, we use a weaker, and thus more conservative, version of the MTR assumption than the one in (A.8). This weaker version states that for any treatment value $d \in D$, and any two values $d_1 \in D$ and $d_2 \in D$ such that $d_2 > d_1$,

$$E[y(d_2)|w = d] \leq E[y(d_1)|w = d] \tag{A.9}$$

Equation (A.9) implies that attending a more selective college has a weakly negative effect on the likelihood of becoming obesity *on average*, that is, not necessarily for every sample respondent. Furthermore, this average weak monotonicity holds for all subsamples that are defined by the treatment actually received.[2] Clearly, (A.8) implies (A.9), but the converse is not necessarily true.

Following the reasoning in Manski (1997) for the case of a weakly negative treatment response, the MTR assumptions (A.8) and (A.9) imply that the bounds on $E[y(d)]$ can be expressed as follows:

[2] Given that (A.9) holds for all values d of the observed treatment w, it is clearly the case that the weak monotonicity in (A.9) applies also to the unconditional expectation, that is, (A.9) implies that $E[y(d_2)] \leq E[y(d_1)]$. However, the converse need not be true.

$$Y_{min}P(w < d) + E(y|w = d)P(w = d) + E(y|w > d)\,P(w > d)$$
$$\leq E[y(d)] \leq \tag{A.10}$$
$$E(y|w < d)P(w < d) + E(y|w = d)P(w = d) + Y_{max}P(w > d)$$

This is so because under both (A.8) and (A.9) imply that $E(y|w > d)$ can be used as a lower bound for $E[y(d)|w > d]$ instead of Y_{min}. Similarly, both (A.8) and (A.9) imply that $E(y|w < d)$ can be used an upper bound for $E[y(d)|w < d]$ instead of Y_{max}. Given that $E(y|w > d)$ is likely considerably larger than Y_{min} and $E(y|w < d)$ considerably smaller than Y_{max}, the identification region defined in (A.10) should be considerably narrower, and thus more informative, than the one in (A.7) that is generated using no assumptions.

The above also imply that to obtain (A.10) one can use the weaker assumption (A.9) that states that MTR holds only on average for each subgroup defined by the treatment actually received instead of the stronger assumption (A.8) that states that MTR holds for every sample unit.

Importantly, Manski (1997) shows that in the case of a weakly increasing MTR the identification region of the ATE under MTR has a lower bound equal to zero because MTR rules out the possibility that a higher value of the treatment induces a lower mean outcome, while allowing for the possibility of a zero effect. In the case of a weakly decreasing MTR (as in our context), the corresponding result is that the MTR upper bound is equal to zero.

A particularly interesting instance of this result when examining the ATE of the change in the treatment from its minimum to its maximum value, denoted by d_{min} and d_{max}, respectively. Given that $(w < d_{min}) = P(w > d_{max}) = 0$ and $P(w \geq d_{min}) = P(w \leq d_{max}) = 1$, the MTR bounds in (A.10) imply that $UB^{MTR}(d_{max}) = LB^{MTR}(d_{min}) = E(y)$. In other words, the MTR assumption leads to the replacement of all counterfactual terms in $UB^{MTR}(d_{max})$ and $LB^{MTR}(d_{max})$ not multiplied with zero with observed outcomes, and, as a result, both these bounds become equal to the observed overall mean. This in turn implies that $UB^{MTR}_{ATE}(d_{max} - d_{min}) = UB^{MTR}(d_{max}) - LB^{MTR}(d_{min}) = 0$.

Clearly, this result applies to our context as well because we have a binary treatment, and thus $d_1 = d_{min}$ and $d_2 = d_{max}$.

A.4. The MIV assumption

One can further narrow the identification region of the ATE by using a considerably weaker kind of IV than the usual exogenous one, namely the MIV. MIVs were introduced by Manski and Pepper (2000), and they satisfy the following requirement for any pair of values z_1, z_2 of Z such that $z_2 > z_1$,

$$E[y(d)|Z = z_2] \geq E[y(d)|Z = z_1] \tag{A.11}$$

where X are a set of control variables. Equation (A.11) states that the MIV can influence the outcome in a particular direction, but also allows for the possibility of no influence whatsoever. Hence, this requirement is much weaker than that of an exogenous instrument which requires no direct relationship between the instrument and the outcome. It is important to note that (A.11) captures only a positive association of Z with Y; a causal relationship is neither implied nor required.

To better understand how MIVs work, we first note that we can always express the lower bound on $E[y(d)]$ under a set of assumptions M as

$$LB^M(d) = \sum_z LB^M(d|Z = z)\,P(Z = z) \tag{A.12}$$

Clearly, $P(Z = z)$ is a given magnitude in the data and thus cannot be changed. Hence, the lower bound $LB^M(d)$ can increase only by increasing the lower bound $LB^M(d|Z = z)$. Similarly, to decrease the upper bound $UB^M(d)$ one must decrease the upper bounds $UB^M(d|Z = z)$.

Let us first examine how an exogenous IV (XIV) – the IV type typically used in treatment effect estimation - can help narrow the identification range. Following Manski (1990), a variable Z is a XIV if $\forall\, d \in D, \forall\, z \in Z$,

$$E[y(d)|\, Z = z] = E[y(d)] \tag{A.13}$$

Equation (A.13) implies that conditioning on any value of the XIV does not change the mean potential outcome. Hence, all identification regions conditional on values of Z should provide identical lower and upper bounds on $E[y(d)]$. Therefore, the identification region of $E[y(d)]$ is the intersection of all identification regions conditional on Z. This intersection is defined as the region between the maximum of all lower bounds conditional on Z and the minimum of all upper bounds conditional on Z. Hence, we have

$$\max_z LB^M(d|Z = z) \leq E[y(d)] \leq \min_z UB^M(d|Z = z) \quad \text{(A.14)}$$

Hence, using XIVs implies that one searches for the maximum lower bound and the minimum upper bound on $E[y(d)]$ by partitioning the sample in cells defined by the XIV values and then comparing the extrema calculated in each cell. This search for the extrema is similar to the search for extrema of objective functions in a dynamic program, or of likelihood functions in econometric estimation, which, however, occur in subsets of the parameter space defined by the chosen grid and/or the optimization method. Clearly, different XIVs will define different partitions of the sample space, and thus likely yield different extrema.

There are, however, a couple of key difference between searching for extrema in the sample space versus the parameter space: i) the size of the sample partitions is in practice constrained by the number of observations in each cell, whereas there is no such constraint when partitioning the parameter space; and ii) local extrema of the bounds on $E[y(d)]$ lead to the estimation of perfectly valid identification regions, which, however are not as informative as when these extrema are global. In other words, using different valid XIVs and various possible combinations of their values will always produce valid identification regions, albeit not necessarily the most informative ones. In contrast, local extrema of objective functions in a dynamic program or in likelihood function estimation will typically yield estimates that are inconsistent. Hence, PI optimization delivers considerably more robust results than dynamic programming or likelihood optimization.

When using an MIV, equation (A.13) does not hold because (A.11) implies that the MIV is weakly monotonically correlated with the outcome. As a result, one cannot calculate the overall identification region as the intersection of all conditional identification regions, as was the case with XIVs. On the other hand, it is possible to exploit the fact that, by (A.11), a lower bound on $E[y(d)|\ Z = z_1]$ is also a lower bound on $E[y(d)|\ Z = z]$ for $z \geq z_1$, and, correspondingly, an upper bound on $E[y(d)|\ Z = z_2]$ is also a upper bound on $E[y(d)|\ Z = z]$ for $z \leq z_2$. Hence, one can potentially increase the lower bound $LB^M(d|Z = z)$ in (A.12) by using the maximum lower bound $LB^M(d|Z = z_1)$ over all $z_1 \leq z$. Correspondingly, one can potentially decrease the upper bound $UB^M(d|Z = z)$ by using the minimum upper bound $UB^M(d|Z = z_2)$ over all $z_2 \geq z$. Hence, we obtain

$$\max_{z_1 \leq z} LB^M[d|Z = z_1] \leq E[y(d)|Z = z] \leq \min_{z \leq z_2} UB^M[d|d|Z = z_2] \quad \text{(A.15)}$$

Once the bounds in (A.15) have been computed for all z, one can take the weighted average over all z using $P(Z = z)$ and bound the potential outcome $E[Y(d)]$ as follows:

$$\begin{aligned} &\sum_z \max_{z_1 \leq z} LB^M[d|Z = z_1]\, P(Z = z) \\ \leq &\sum_z E[y(d)|Z = z]P(Z = z) = E[y(d)] \leq \\ &\sum_z \min_{z \leq z_2} UB^M[d|Z = z_2]\, P(Z = z) \end{aligned} \quad \text{(A.16)}$$

In other words, by integrating Z out of the bounds on the conditional expectation $E[y(d)|Z = z]$, one can obtain bounds on $E[y(d)]$.

Clearly, the optimization operations in (A.15) take place over a restricted range of values of Z compared to (A.14), and thus the identifying power of the MIV assumption is smaller than that of the XIV one. This is to be expected, as the weak monotonicity of a MIV in (A.11) is a weaker assumption than the exogeneity of an XIV in (A.13). As with XIVs, this weak monotonicity assumption is imposed on the unobserved potential outcome $E[y(d)]$; hence, it cannot be tested using the observed data without imposing further assumptions.

As is the case with XIVs, valid MIVs generate valid identification regions, although not necessarily the most informative ones.

In our context, the MIV used, namely the PPVT score, is assumed to have a weakly negative association with obesity, as described in Section 3.

A.5. Treatment spillovers

When there are treatment spillovers and thus the SUTVA is violated, the potential outcomes can be expressed, following Manski (2013), as $y_i(d^J)$, that is, as a function of the treatment d received by a set J of respondents, here understood as consisting of the whole population. Specifically, the treatment received by any member of the set J can potentially affect the potential outcome of any other member i of J. In our context, this

implies, for example, that the likelihood of obesity for any given individual i who attends a very selective college, can be affected by that individual's peers who also attend a very selective college. This would not have been possible had the response function $y_i(\bullet)$ been individualistic.

The definition of the ATE now needs to be adapted to the non-individualistic nature of the response function as follows:

$$ATE(d_2^J - d_1^J) = E[y(d_2^J)] - E[y(d_1^J)] \quad \text{(A.17)}$$

Average potential outcomes under non-individualistic response are equal to:

$$E[y(d^J)] = E[y(d^J)|w = d]P(w = d) + E[y(d^J)|w \neq d]P(w \neq d) \quad \text{(A.18)}$$

In contrast to the case of individualistic responses, as described in (A.2), one cannot replace the term $E[y(d^J)|w = d]$ with $E(y|w = d)$, given that the latter is generically the outcome of the interaction of sample units that choose $w = d$ with sample units belonging in J that choose treatment values that can different from d.

Without making any further assumptions, both counterfactual terms $E[y(d^J)|w = d]$ and $E[y(d^J)|w \neq d]$ can be only bounded from below and above by $Y_{min} = 0$ and $Y_{max} = 1$, respectively, which thus become also the bounds on $E[y(d^J)]$, that is:

$$0 \leq E[y(d^J)] \leq 1 \quad \text{(A.19)}$$

A.5.1 Reinforcing interactions

The reinforcing interaction (RI henceforth) assumption states that, for any subsample defined by the actual treatment received (i.e., the selectivity of the college of actual attendance), if all individuals in the subsample had attended a very selective college, the average obesity prevalence would have been weakly lower than if those individuals had attended a less selective college. In other words, we have that, for any two treatment vectors d_1^J and d_2^J such that $d_2^k \geq d_1^k \; \forall \; k \in J$,

$$y_i(d_2^J) \leq y_i(d_1^J), \forall \; i \in J \quad \text{(A.20)}$$

where J in our context is understood to be the whole population. (A.20) implies, following the reasoning in Manski (2013, p. S10), that when RI holds, the likelihood of obesity

weakly decreases for any given person with the selectivity of the college that (s)he attends, and also with the selectivity of the college that other members of the population interacting with that person attend. Thus, as pointed out by Manski (2013), the treatments received by others reinforce a person's own treatment.

As was the case with MTR, we use a weaker, and thus more conservative, version of the RI assumption than the one in (A.20). This weaker version states that for any treatment value d, and any two treatment vectors d_1^J and d_2^J such that $d_2^k \geq d_1^k \; \forall \; k \in J$,

$$E[y(d_2^J)|w = d] \leq E[y(d_1^J)|w = d] \tag{A.21}$$

where, once more, J includes the whole population in our context. Compared to (A.20), (A.21) implies that, when RI holds, for any subsample of individuals choosing a particular treatment value, only the average prevalence of obesity in this subsample weakly decreases with the selectivity of the college that each subsample member attends, and also with the selectivity of the college that other individuals in the population interacting with the subsample members attend.

Using RI, one can narrow the identification region of $E[y(d^J)]$ when the treatment takes its minimum value $d_{min} = 0$. Specifically, both (A.20) and (A.21) imply that the counterfactual term $E[y(0^J)|w \geq 0]$ can be bounded below by $E(y|w \geq 0)$, instead of $Y_{min} = 0$. Intuitively, under RI, if nobody attends a very selective college, then the average obesity prevalence should be weakly higher than the observed one, which is the outcome of a situation in which at least some individuals attend a more selective college.

Given the above, and the fact that $P(w < 0) = 0$ and $P(w \geq 0) = 1$, the bounds for $E[y(0^J)]$ derived from (A.18) are as follows:

$$E(y|w \geq 0)\, P(w \geq 0) = E(y) \leq E[y(0^J)] \leq Y_{max} = 1 \tag{A.22}$$

Correspondingly, RI allows one to narrow the identification region of $E[y(d^J)]$ also when the treatment takes its maximum value d_{max} =1. Specifically, both (A.20) and (A.21) imply that $E(y|w < 1)$ can be used an upper bound for $E[y(1^J)|w < 1]$ instead of $Y_{max} = 1$. Intuitively, under RI, if everybody attends a very selective college, then the average obesity prevalence should be weakly lower than the observed one, which is the

outcome of a situation in which at least some individuals do not attend a more selective college.

Given the above, and that $P(w > 1) = 0$ and $P(w \leq 1) = 1$, the bounds for $E[y(1^J)]$ derived from (A.18) are as follows:

$$Y_{min} = 0 \leq E[y(1^J)] \leq E(y|w \leq 1)\, P(w \leq 1) = E(y) \tag{A.23}$$

Hence, we observe that under RI, we have that $UB^{RI}(d_{max}) = LB^{RI}(d_{min}) = E(y)$, just as was the case with MTR when the SUTVA was assumed to hold, as discussed in Section A.3 above. In other words, the observed obesity prevalence is both the RI upper bound of the obesity prevalence when everybody attends a very selective college and the RI lower bound of the obesity prevalence when everybody attends a less selective college.

(A.22) and (A.23) imply that the ATE under RI can be bounded as follows

$$\begin{gathered} LB^{RI}(1) - UB^{RI}(0) = Y_{min} - Y_{max} = -1 \\ \leq E[y(1^J)] - E[y(0^J)] \leq \\ UB^{RI}(1) - LB^{RI}(0) = E(y) - E(y) = 0 \end{gathered} \tag{A.24}$$

The MIV assumption under spillovers can be defined in a similar way as when the response is individualistic. In particular, we have that for any pair of values z_1, z_2 of Z such that $z_2 > z_1$,

$$E[y(d^J)|Z = z_2] \geq E[y(d^J)|Z = z_1] \tag{A.25}$$

The MIV assumption in the presence of treatment spillovers can be justified on the same grounds as when there are no such spillovers. In our context, this implies that, taking into account treatment spillovers, a higher human capital, as denoted by higher values of the PPVT score, should be associated with a lower prevalence or obesity, for any level of college selectivity.

The MIV assumption operates in the presence of treatment spillovers exactly as without them because it is always the case that terms involving unobserved potential outcomes are replaced in the bounds by observed magnitudes. Hence, there is no change in the way the optimization operations described in (A.15) and (A.16) work to narrow identification regions. What is different with treatment spillovers is that, in some cases,

the bounds on which the MIV assumption operates are not the same as those without treatment spillovers.

A.6. Advantages of PI

All in all, there are many reasons why one would prefer PI methods to other more commonly used ones (e.g., OLS-, IV- or panel data-based) when trying to estimate the causal effect of interest. First, PI methods are completely nonparametric, as they require only the calculation of sample averages of the outcome and the prevalence of the treatment.

Second, PI methods produce estimates of the ATE across all sample units, and not of the LATE as is the case with IV estimation when the treatment is heterogeneous. Thus, they allow for arbitrary forms of heterogeneity of the treatment effect because the ATE is just an average magnitude across sample units. Such unlimited heterogeneity of the treatment effect is not typically allowed for, as in most estimation methods one makes specific assumptions about how the treatment variable enters the specification. Moreover, if one is interested in the heterogeneity of the treatment effect in specific dimensions, then one can simply apply PI methods to subsamples defined by particular combinations of values of control variables.

Third, in PI one bounds the unconditional expectation $E[y(d)]$, taking as given the distribution of all observables and unobservables (other than the treatment) that might affect the outcome. Hence, one does not need to worry about i) which variables to add in the empirical specification; ii) the way they appear; and iii) whether they are endogenous.

Fourth, PI methods accommodate any form of endogeneity (e.g., due to both time-varying and time-invariant unobservables or selectivity), as they allow for any form of non-random selection into treatment. This also implies that one does not need to assume specific properties of the error term, as is the case with regression methods.

Fifth, PI uses very few and quite mild assumptions to narrow the identification region of the estimates. Importantly, it is completely transparent about how adding each assumption affects the identification region. In contrast, most commonly used estimation methods impose simultaneously many assumptions on the empirical model, and thus it is typically unclear how each of them affects estimates.

Sixth, PI methods allow the use MIVs that can tighten the identification regions. As is the case with standard IV estimation, the assumptions behind those variables cannot be tested without making further assumptions. However, MIVs - unusable in standard IV estimation - are required to be weakly monotonically related to the outcome, which is a much weaker assumption than the exogeneity required of standard IVs.

Seventh, the most important PI result (namely the upper bound of the ATE, that is, by at least how much attending a selective college reduces the prevalence of obesity) remains valid even when there are treatment spillovers, provided one uses the RI assumption, supplemented by the MIV assumption. This is in contrast to what happens with OLS, IV or panel data methods. This is an extremely useful result, given that treatment spillovers are likely to obtain in many contexts.

Eighth, PI can operate without problems on cross-sectional data, and thus panel data are not required. This is so because PI assumes that the treatment is endogenous, and this endogeneity can be due to time-varying or time-invariant unobservables, or both. One can accommodate dependencies among sample units (e.g., due to repeated observation of sample units or features of the sampling process) through the appropriate clustering and stratification when bootstrapping standard errors.

Appendix References

Christelis, D., D. Georgarakos, T. Jappelli, and M. van Rooij (2020), "Consumption Uncertainty and Precautionary Saving," *The Review of Economics and Statistics*, 102, 148-161.

Christelis, D., and L.I. Dobrescu (2020), "The causal effect of social activities on cognition: Evidence from 20 European countries," *Social Science and Medicine* 247, Article 112783.

Manski, C. F. (1989), "Anatomy of the selection problem," *Journal of Human Resources* 24, 343–60.

Manski, C. F. (1990), "Nonparametric Bounds on Treatment Effects," *American Economic Review Papers and Proceedings*, 80, 319-323.

Manski, C. F. (1994), "The Selection Problem," in *Advances in Econometrics: Sixth World Congress*, ed. by C. Sims. Cambridge, UK: Cambridge University Press, 143–170.

Manski, C. F. (1997), "Monotone Treatment Response," *Econometrica*, 65, 1311-1334.

Manski, C. F., and J. V. Pepper (2000), "Monotone Instrumental Variables: With an Application to the Returns to Schooling," *Econometrica* 68, 997-1012.

Manski, C. F., and J.V. Pepper (2018), "How Do Right-to-Carry Laws Affect Crime Rates? Coping with Ambiguity Using Bounded Variation Assumptions," *Review of Economics and Statistics* 100, 232–244.

Appendix: Figures and Tables

Table A1. Summary statistics

	Mean	Std. Dev.	N
Age (Wave IV)	28.25	1.81	3578
Age (Wave V)	37.24	1.87	2856
Female	0.54	0.50	3578
Hispanic	0.07	0.25	3578
Black	0.10	0.30	3578
PPVT	109.26	12.10	3578
Parent has a college degree	0.59	0.49	3554
Family SES	0.00	1.00	3474
Highly selective college	0.20	0.40	3578

Note: the means are weighted using Wave IV cross-sectional weights. PPVT is the Peabody Picture Vocabulary Test score.

Table A2. Summary of outcomes

	Mean	Std. Dev.	N
Panel A: Wave IV			
Obese	0.262	0.44	3525
BMI	27.356	6.27	3525
BMI self-reported	26.674	5.72	3536
Waist (cm)	94.073	15.03	3547
Ate fast food more than once in the past week	0.378	0.49	3576
# Sweet drinks in the past week	7.201	8.64	3576
No physical activity in the past week	0.093	0.29	3576
# Times watched TV in the past week	11.352	8.78	3567
Panel B: Wave V			
Obese	0.366	0.48	1520
BMI	29.004	6.86	1520
BMI self-reported	28.093	6.45	2839
Waist (cm)	93.706	17.10	1537
Ate fast food more than once in the past week	0.372	0.48	2805
# Sweet drinks in the past week	4.584	5.84	2800
No physical activity in the past week	0.077	0.27	2808
# Times watched TV in the past week	11.461	10.18	2791

Note: Wave IV cross-sectional weights are used in Panel A and Wave V cross-sectional weights are used in Panel B.

Table A3. Correlation between the PPVT index and the outcome variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Obese	BMI	Waist (cm)	Ate fast food more than once in the past week	# Sweet drinks in the past week	No physical activity in the past week	# Times watched TV in the past week
Panel A: Wave IV							
PPVT	-0.047***	-0.059***	-0.014	-0.093***	-0.042**	-0.006	0.001
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Panel B: Wave V							
PPVT	-0.093***	-0.094***	-0.056**	-0.101***	-0.038**	-0.024	0.029
	(0.027)	(0.027)	(0.027)	(0.019)	(0.019)	(0.019)	(0.019)

Note: Correlation coefficients between the Peabody Picture Vocabulary Test scores (PPVT) and outcome variables are reported. ∗ ∗ ∗: $p < 0.01$, ∗∗: $p < 0.05$.

Table A4. OLS estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Obese	BMI	Waist (cm)	Ate fast food more than once in the past week	# Sweet drinks in the past week	No physical activity in the past week	# Times watched TV in the past week
Panel A: Wave IV							
More selective college	-0.072***	-1.187***	-3.845***	-0.130***	-0.936**	-0.019	-1.147***
	(0.023)	(0.321)	(0.884)	(0.033)	(0.413)	(0.015)	(0.399)
Observations	3,525	3,525	3,547	3,576	3,576	3,576	3,567
R-squared	0.025	0.042	0.060	0.047	0.036	0.012	0.012
Panel B: Wave V							
More selective college	-0.072*	-1.332***	-2.703**	-0.087**	-1.321***	-0.017	-1.450**
	(0.042)	(0.499)	(1.326)	(0.035)	(0.342)	(0.014)	(0.595)
Observations	1,520	1,520	1,537	2,805	2,800	2,808	2,791
R-squared	0.033	0.063	0.096	0.059	0.015	0.029	0.020

Note: All regressions include age, age squared, gender and race indicators, and PPVT. Standard errors (in parentheses) are clustered at the school level. Estimates are weighted using Wave IV (Panel A) and Wave V (Panel B) weights. ∗ ∗ ∗: $p < 0.01$, ∗∗: $p < 0.05$, ∗: $p < 0.1$.

Table A5. Probability of being obese in wave IV, conditional on being observed in wave V

Assumptions	Estimates		Lower Bound Low 95% CI	Lower Bound Upper 95% CI	Lower Bound Low 95% CI	Upper Bound Upper 95% CI
	Lower bound	Upper bound				
Exogenous treatment selection	-0.1172		-0.1785		-0.0586	
No assumptions	-0.4094	0.5906	-0.4590	-0.3693	0.5410	0.6307
MTR	-0.4094	0.0000	-0.4590	-0.3693	0.0000	0.0000
MTR + MIV (bias-corrected)	-0.2674	-0.0209	-0.3314	-0.1674	-0.0953	-0.0100
MTR + MIV (bias-uncorrected)	-0.2374	-0.0676	-0.3314	-0.1674	-0.0953	-0.0100
Number of observations	1,506					

Note: MTR: monotone treatment response, MIV: monotone instrumental variable (Peabody Picture Vocabulary Test score).