The Journal of Mental Health Policy and Economics
J Ment Health Policy Econ **14**, 165-186 (2011)

Isolating the Effect of Major Depression on Obesity: Role of Selection Bias

Dhaval M. Dave,[1*] Jennifer Tennant,[2] Gregory Colman[3]

[1]*Bentley University & National Bureau of Economic Research, Department of Economics, Waltham, MA, USA*
[2]*Ithaca College, Department of Economics, Ithaca, NY, USA*
[3]*Pace University & National Bureau of Economic Research, Department of Economics, New York, NY, USA*

Abstract

Background: There is suggestive evidence that rates of major depression have risen markedly in the U.S. concurrent with the rise in obesity. The economic burden of depression, about $100 billion annually, is under-estimated if depression has a positive causal impact on obesity. However, virtually the entire existing literature on the connection between the two conditions has examined merely whether they are significantly correlated, sometimes holding constant a limited set of demographic factors.

Aims of the Study: This study assesses whether, and the extent to which, the positive association between the two conditions reflects a causal link from major depression to higher BMI and obesity.

Methods: Individual-level data from three nationally-representative studies are utilized: (i) National Comorbidity Survey-Replication (N=3,229); (ii) National Longitudinal Survey of Youth-1979 (N=21,365); and (iii) Behavioral Risk Factor Surveillance System (N=2,858,973). Dependent variables include body mass index (BMI) and a dichotomous indicator for overweight or obese. We measure diagnosed major depression based on DSM-IV criteria and the CES Depression scale. While contemporaneous effects are considered, the study primarily focuses on the effects of past and lifetime depression to bypass reverse causality and further assess the role of non-random selection on unobservable factors. The effects of past and lifetime depression on obesity are estimated based on: (i) models that control for an extensive set of typically-unobserved factors, including parental history, family background, parental investments, risk-taking, and use of anti-depressants and other prescription medications; (ii) constrained selection models; and (iii) models controlling for family fixed effects.

Results: There are expectedly no significant or substantial effects of current depression on BMI or overweight/obesity, given that BMI is a stock that changes relatively slowly over time. Results also do not support a causal interpretation among males. However, among females, estimates indicate that past or lifetime diagnosis of major depression raises the probability of being overweight or obese by about seven percentage points. Results also suggest that this effect appears to plausibly operate through shifts in food consumption and physical activity.

Discussion: Unadjusted differences document a strong correlation between depression and obesity, both cross-sectionally and temporally. However, it remains unclear how much of this association is consistent with a causal link from depression to obesity and how much of it is being driven by non-random selection. We find evidence that past and lifetime depression raises the probability of being overweight or obese among females. We estimate that this higher risk of overweight and obesity among females could potentially add about 10% (or $9.7 billion) to the estimated economic burden of depression.

Implications for Health Policies: Estimates from this study suggest that the rising trend in obesity partly underlies the reported increased prevalence of depression, at least among women. Public health interventions which reduce major depression among women could therefore also further promote public health by reducing overweight and obesity.

Implications for Further Research: While this study points to some preliminary evidence that the effect of depression on obesity appears to operate through shifts in diet and physical activity, more research is required to inform the proximate and distant mediating pathways. Though this study focuses on gender differentials, differences based on race/ethnicity and educational status would further inform heterogeneous responses across individuals and population subgroups.

Received 10 June 2011; accepted 20 November 2011

*** Correspondence to:** Dhaval M. Dave, Bentley University & National Bureau of Economic Research, Department of Economics, 175 Forest St., AAC 195, Waltham, MA 02452, USA
Tel.: +1-212-817 7955
Fax: +1-212-817 1597
E-mail: ddave@bentley.edu

Source of Funding: The authors would like to thank their respective universities for research support.

Introduction

Until recently obesity in the U.S. was a fairly rare occurrence. However, between 1980 and 2000, the prevalence of obese adults doubled to 34%, and the prevalence of obesity for children almost tripled to 17%. Recent data indicate that obesity appears to have reached a plateau at these high rates.[1] Obesity is a significant public health concern due to the high level of morbidity, premature mortality and related economic burden. It has been identified as a contributing factor to approximately 100,000-400,000 deaths and imposes an economic burden of $140-$215 billion in direct (healthcare expenditures) and indirect (productivity losses) costs.[2,3]

Coincident with the rise in obesity, there is suggestive evidence that rates of major depression have also risen markedly, with increases noted for most socio-demographic subgroups of the U.S. population. From 1991-1992 to 2001-

2002, past year prevalence of major depression among adults has approximately doubled.[4] Major depression affects almost one out of every five adults in their lifetime, and has been linked to increased risk of stroke and of heart failure, reduced bone density among the elderly, higher rates of substance abuse, and higher mortality in general.[5-7] Its economic burden is correspondingly large. Recent estimates place the current cost of major depression in the U.S. due to lost productivity and increased healthcare expenditures at approximately $100 billion per year (inflation adjusted), being surpassed by only hypertension and heart disease.[8]

These costs of depression are under-estimated if depression has a positive causal impact on obesity, which in turn would further adversely affect productivity and raise healthcare spending. If depression plays a causal role in increasing the prevalence of obesity, then policy interventions aimed at promoting mental health may also have the indirect benefits of promoting a healthy bodyweight. One of the targets of the Healthy People 2010 initiative of the U.S. Department of Health and Human Services was to reduce the adult obesity rate to 15% and the childhood obesity rate to 5% by 2010. These goals were unmet, and Healthy People 2020 aims at a 10% improvement in the adult and childhood obesity rates.* Assessing the link between depression and obesity can be informative when measuring the effectiveness of mental health policy interventions in reducing obesity.

However, virtually the entire existing literature on the connection between the two conditions has examined merely whether they are significantly correlated, sometimes holding constant other factors, such as health status, exercise, and social isolation. No prior study has directly addressed potential endogeneity bias due to reverse causation and unobservable individual heterogeneity. The U.S. Surgeon General's report on mental health[9] summarizes estimates of the economic burden of depression and mental illness, though none of these reference any potential effects on obesity. Even the most recent studies,[8] which attempt to quantify how depression affects health care spending and related costs, do not account for the effects of depression on obesity because no credible estimates of a causal effect are available. The gap in knowledge with respect to the potential causal role of depression in affecting obesity casts some doubt on the estimated economic burden of depression, limits our understanding on the various determinants of obesity, and thus limits the design of effective public interventions that may directly or indirectly stem the rise in obesity prevalence.

The aim of this study is to address this gap. We apply alternate methodologies to individual records from three large-scale, nationally-representative datasets, both cross-sectional and longitudinal, in order to evaluate whether and the extent to which the observed association reflects a causal link from depression to obesity. We focus on the effects of past and lifetime depression to bypass reverse causality and assess the role of non-random selection on unobservable factors. The estimated marginal effects are used as inputs in deriving the first estimates of how much the increased risk of overweight and obesity adds to the economic costs of depression.

Background and Literature

A number of prior studies in the public health and epidemiology literature have explored connections between obesity and mental health. Several of these studies are based on small samples that focus on very specific questions or diseases. Solaroli *et al.*,[10] for instance, administer a questionnaire to obese patients, and conclude that eating disorders are common among obese individuals, especially in patients with poor health-related quality of life. McElroy *et al.*[11] assess the prevalence and clinical correlates of overweight status, obesity, and extreme obesity in a sample of bipolar patients and find that current body mass index (BMI) and bodyweight were each correlated with the number of weight gain-associated psychotropics to which patients had been exposed. They also find that the prevalence of obesity was common in their sample of bipolar patients, but are inconclusive about whether or not these rates were higher than the rates in the general population.

There has also been some previous research based on general population surveys. Onyike *et al.*[12] utilize data from the National Health and Nutrition Examination Survey (NHANES) to examine how obesity affects the risk of depression. They compare risks of depression in obese (BMI >= 30) and normal-weight (BMI 18.5 – 24.9) people as well as individuals stratified by the severity of obesity. This study finds that obesity is associated with depression mainly among persons with severe obesity (BMI >= 40). Ma & Xiao[13] also use the NHANES to study the link between obesity and depression in women. They find that BMI is positively associated with depressive symptoms, and severely obese women are at especially high risk for depression. Carpenter *et al.*[14] examine the relationship between relative body weight and depression or suicidal tendencies, and also find differences across gender. While obesity is associated with an increased risk of depression among females, it is associated with a decreased risk of depression among males. Using the National Comorbidity Survey – Replication (NCS-R), Simon *et al.*[15] find that obesity is significantly associated with increases in lifetime diagnosis of major depression, bipolar disorder, and panic disorder. They find that the association between obesity and mood disorder was strongest among non-Hispanic whites and college graduates. However, unlike some other studies, they find no difference in these associations between men and women.

Looking at this question in a more international context, Scott *et al.*[16] analyzed thirteen cross-sectional, general population surveys of the adult households in order to determine whether or not there was an association between obesity and mental disorders in various countries. They found that obesity and mental conditions had significant yet modest associations, particularly for females and those with

* www.healthypeople.gov

severe obesity. However, they noted that the findings did not ''clarify the direction or nature of the relationship observed''.

While these studies highlight that there is a link between obesity and depression, and also point to significant differences across gender, this is a typical conclusion in that they do not inform on the nature of this link. Several of the reviewed studies are based on small condition-specific samples (that is individuals who are overweight or obese), which limits their external validity. In addition, virtually all of the studies based on large population-based samples have considered how obesity (or overweight status) affects depression, rather than how depression impacts obesity. All of the previous studies have also focused on associations or correlations, and most do not control for other confounding factors that may be driving both the risk of obesity and depression. All of these studies have discussed the need for further research to clarify causal mechanisms.

Our study addresses this knowledge gap and uses alternate econometric methods and nationally-representative datasets with various strengths in order to assess whether, and the extent to which, there is a plausible causal link from depression to obesity. It should be noted that establishing causality from depression to obesity is not feasible with experimental or clinical data, and can only be evaluated with secondary data. The use of multiple, nationally-representative, individual-level datasets, combined with well-established methods, adds to their credibility and adds to the weight of the evidence bearing upon the question of how depression affects obesity.

Methods

Analytical Framework

There are several reasons to expect that major depression may be a significant input into the production of bodyweight and obesity. Weight gain is a relatively common problem during both acute and long-term treatment with antidepressants.[11,17] In addition, numerous studies have documented associations between depression (and other mental disorders) and binge eating, weight loss efficacy, and other eating disorders[18-21] though the net-effect on caloric intake is not always clear. While earlier studies had suggested that weight-loss is a symptom of depression, more recent studies have suggested that people can gain rather than lose weight when depressed.[21] Gluck *et al.*[22] for instance find that individuals with night-eating syndrome had higher rates of depression and conclude that increased caloric intake later in the day may be a contributor to poorer weight-loss outcomes. Epel *et al.*[23] study stress-response among women in a laboratory setting and find that women exposed to greater levels of stress consumed more calories on a given day relative to a control set of women.

Depression is also associated with exercise, physical activity, and caloric expenditure, though the direction of causality is not well-established and is often confounded by its link with obesity. Several studies show that exercise and physical activity can have a protective effect on the risk of depression and mental disorders.[24,25] Depression and mood disorders are also significant determinants of physical activity. Bauman *et al.*[26] review available research on the determinants relating to the adoption and maintenance of physical activity, and note several studies showing that mood disorders and poor psychological health reduce the likelihood of engaging in exercise and physical activity.

A large literature further indicates that depression and painful symptoms commonly co-occur.[27] To the extent that depression is also associated with certain types of pain, for instance headaches or joint pain, then a depressed individual may become sedentary and experience an increase in bodyweight.

Given that depression is correlated with the proximate causes of obesity – caloric intake and caloric expenditure – the objective of this study is to assess the extent to which depression impacts measures of obesity. Empirically identifying the causal effect of depression on obesity is complicated by two issues. First, depression may be structurally endogenous to obesity, that is, in addition to depression affecting obesity, the causality may also run in the other direction. Our focus on the effect of lifetime and past depression on current measures of obesity bypasses this simultaneity problem. Second, what we refer to as statistical endogeneity, wherein the risk of depression and obesity may depend on a common set of unobserved factors (for example, life history, early health investments, time and risk preference, divorce), is a more relevant concern for this study. Consider linear specifications of the structural production function for body mass index (BMI_{it}) and depression (D_{it-1}):*

$$(1)\ BMI_{it} = \alpha_1 D_{it-1} + a_2 X_{it} + a_3 \mu_i + e_{it}$$
$$(2)\ D_{it-1} = \lambda_1 BMI_{it-1} + \lambda_2 BMI_{it-2} + \lambda_3 X_{it-1} + \lambda_4 \mu_i + \nu_{it-1}$$

Equation (1) is a production function for body mass index, with past depression as an input. BMI is also a function of observable characteristics (X), such as age, gender, race/ethnicity, educational attainment, and unobservable characteristics (μ) pertaining to the individual, such as family background, tolerance towards risk, and the rate of time preference. While BMI is a stock, accumulated through past excess caloric intake, and therefore has a persistent element, past BMI is not included in equation (1) since one of the channels through which past depression (D_{it-1}) may impact current BMI is through its contemporaneous effect on past BMI. Similarly, equation (2) specifies a mental health production function, with contemporaneous and past BMI as inputs along with observable characteristics pertaining to the individual. The vector μ denotes unobserved determinants of BMI that may also influence mental health. The error terms are denoted by e and ν, representing potential unobserved time-varying influences of BMI and depression, respectively. Subscripts refer to the i^{th} individual observed in time period t.

* Both production functions are based on the physical health production function specified in Grossman;[28] see, specifically, Chou *et al.*[29] for a specification of the BMI production function, and Saffer and Dave[7] for a specification of the mental health production function. In alternate specifications, we also employ an indicator for overweight or obese defined as having a BMI of 25 or greater.

Our objective is to estimate α_1 in order to assess the existence and strength of a possible causal impact of past depression on BMI (and measures of obesity and overweight). While we also employ measures of current and lifetime depression in alternate models, we focus on past depression for three reasons. First, this bypasses the structural endogeneity (reverse causality) issue permitting us to address the remaining selection bias stemming from statistical endogeneity. Second, since body mass index and obesity are stock measures, accumulated through prior excess caloric intake over expenditure, it is expected that the impact of past depression on current BMI would be more reflective of a long-term effect. Indeed, we can use this as a specification check since, if depression causally affects obesity, the impact of past depression on current BMI should be larger than the contemporaneous impact. Third, the parameter α_1 captures the more policy-relevant net reduced-form effect of past depression on current obesity, operating through all potential pathways. This includes any channels operating through the persistence of depression and obesity over the life cycle – for instance, past depression may impact current obesity by (i) affecting past obesity, which in turn affects current obesity, (ii) raising current obesity directly, and / or (iii) raising the likelihood of current depression, which in turn affects current obesity.*

Ordinary least squares (OLS) estimation of Equation 1 yields biased estimates of α_1 however, due to unobserved non-random selection into both past depression and current obesity – that is, correlation between μ_i and $D_{it-1}(\lambda_4 \neq 0)$. We refer to this problem as selection on observables and selection on unobservables.[30]. These terms denote that respondents are not randomly sorted into past depression and current obesity. Selection on observables refers to observed factors (such as age, gender, and race) that are correlated with both outcomes. Selection on unobservables refers to possible factors that are not available in our dataset, and will therefore influence the effect of past depression on obesity.

Our estimation strategy proceeds in a stepwise fashion to address this selection bias and conforms to the strengths of each dataset. The National Comorbidity Survey (NCS-R) is cross-sectional, but has detailed information on the lifetime and current prevalence of major depressive disorder based on clinical guidelines set by the latest Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) published by the American Psychiatric Association. The NCS-R also contains extensive information on the respondent's life history, parental investments, and other risk-related behaviors, which are typically unobserved in other studies. The degree of selection on unobservable factors can be indirectly assessed by gauging the robustness of the estimated parameter α_1 as these typically-unobserved factors are included in the model.

Initially, we therefore begin with a parsimonious set of covariates, and then estimate models with an extended set of covariates including family history, parental characteristics, health investments, life shocks, proxies for risk tolerance, use of prescription medications and anti-depressants, and factors which are typically unobserved in other studies. Estimating both the basic and extended models allows us to evaluate how much of the association between past depression and obesity appears to be driven by selection on unobservables. If the magnitude of the effect of past depression on current obesity is highly sensitive to the inclusion of the additional covariates and typically-unobserved factors, then it is likely that factors which remain unobserved also play some role in this relationship.*

The degree of selection on the unobserved characteristics cannot be measured directly with non-experimental data. As a next step, we therefore rely on a novel approach proposed by Altonji *et al.*[30] to bound this latter effect, allowing us to draw inferences regarding the unbiased relationship between past depression and obesity. This involves obtaining estimates of the effect of past depression on obesity from a bivariate probit regression model in which the correlation between unobserved variables is fixed at various levels. This part of the analysis allows us to assess the sensitivity of the estimated effect of depression to the potential problem of additional selection on unobservables.

Specifically, we apply a bivariate probit methodology to equations (1) and (2), under which the disturbance terms (e and u) in both equations are distributed bivariate normally with a correlation of ρ.† First, we estimate the bivariate probit model without any identifying assumptions but with a constrained correlation coefficient, ρ. We constrain ρ to be 0.10 initially and then examine the effects of increasing ρ in increments of 0.10 to 0.20, 0.30, 0.40 and 0.50. Since it is also plausible that unobserved factors common to both past depression and obesity are negatively correlated, we then constrain ρ to be -0.10, -0.20, and -0.30. In this way, we impose on the model increasingly greater amounts of (positive and negative) correlation between unobservables and examine whether or not the effect of past depression on obesity is robust to such changes. This analysis allows one to determine the threshold of selection on unobservables, if any, at which depression no longer has a statistically significant effect on obesity.

The second step of this constrained-selection methodology computes the amount of selection into past depression and obesity based on observed variables, and obtains estimates of the effect of past depression under the assumption that the degree of sorting on unobserved variables is equal to the degree of sorting on observed variables. Altonji *et al.*[30] note

* We are careful not to ''over-control'' in the models for factors that could potentially be mechanisms through which depression may have an impact on obesity. It is validating that the magnitude of the effect of depression on obesity (and on exercise) diminishes when we control for pain comorbidities such as general pain, headaches, back and neck pain, and arthritis. Specifically, the effect sizes diminish by about 10-15% when these conditions are included. This suggests that such conditions (which are potential symptoms or comorbid correlates of depression) may underlie part of the channel through which depression may be exerting an influence on obesity. Our results are also robust to controlling for presence of diabetes. We prefer not to include it in our models, however, since diabetes is endogenous, being potentially affected by obesity, our dependent variable.

* The direction and magnitude, however, is unknown, depending on the nature of the joint distribution of the observed and unobserved characteristics.

† The model is estimated using maximum likelihood.[31,32]

that if the observable determinants of an outcome are truly just a random subset of the complete set of determinants, selection on observable characteristics should be equal to selection on unobservable characteristics. This assertion of equal selection is unlikely to be true, and in fact, given the specialized nature of the NCS-R where we are able to observe an extended set of individual and family characteristics, we would expect selection on observable factors to be greater than selection on unobservable factors. Thus, the causal effect of past depression is expected to be bounded between estimates obtained from the single-equation extended model (that assumes no additional selection on unobservable variables) and the constrained-selection model (that assumes additional selection on unobservable variables equals the selection on observables).

The advantage of the Altonji *et al.*[30] procedure is that it allows researchers to assess the possible existence and strength of a causal relationship without requiring the use of identifying assumptions that are often not credible – for example, the existence of valid instruments in an instrumental variables context or other ad hoc exclusion restrictions.* As a result, without any other identifying assumptions, researchers can estimate the degree of sorting on unobservable factors using the observed data, and identify bounds on the causal parameter estimate.

In addition to the NCS-R, we also utilize individual records from the National Longitudinal Survey of Youth-1979 (NLSY79) and apply panel-data methods to examine how depression impacts obesity. With the NLSY79, our goal is to estimate a regression model of the form

$$(4)\ \mathrm{BMI}_{iw} = \beta_0 + \beta\,\mathrm{Depressed}_{iw-1} + \mathrm{X}_{iw}\Omega + \mu_i + \varepsilon_{iw}$$

where X_{iw} represents observed time-varying characteristics of respondent i in survey wave w, μ_i represents observed and unobserved time-constant person-specific variables, and ε_{iw} is the error term, which includes, among other things, unobserved time-varying influences on BMI. The parameter of interest is β, which represents the impact of past depression on BMI. We denote survey wave with w rather than with time period t to emphasize that they are distinct, as depression is measured at alternate points in time for different individuals (described in the next section). The distinction between age and year allows us to control for both.†

As a comparison with our later fixed-effects estimates, and to see what a pure cross-sectional regression would reveal, we collapse equation (4) by respondent and estimate equation (5), the between-effects regression.

$$(5)\ \overline{\mathrm{BMI}}_i = \beta_0 + \beta\,\overline{\mathrm{Depressed}}_i + \overline{\mathrm{X}}_i\Omega + \mu_i + \overline{\varepsilon}_i$$

As noted with respect to equation (1), this will be consistent for β as long as $\overline{\mathrm{Depressed}}$ is uncorrelated with either μ_i or $\overline{\varepsilon}_i$. We then estimate equation (4) using a person-specific fixed effects model, which treats the μ_i as parameters to be estimated rather than as part of the error term. This controls for all person-specific time-invariant influences on BMI that may also be correlated with depression, such as family background, genetic predisposition to both obesity and depression, and traumatic events that occurred before the first observed wave. We also estimate equation (4) with first-differences to purge the unobserved time-invariant person-specific factors. One limitation of using person fixed effects and first-differences is that a respondent's BMI and depression status may not change much between waves of the survey, providing too little variation to identify the parameters. Therefore, we also estimate equation (4) using family fixed-effects, which is possible because about half of the respondents in our analysis sample are siblings to another NLSY79 respondent. The advantage of sibling or family fixed-effects is that there tends to be more (though perhaps still not sufficient) variation in BMI and depression among siblings than for a single person over time. The disadvantage is that important unobserved personal characteristics may possibly remain in the error term after the unobserved family characteristics have been removed by the family fixed effects model. We test whether our estimates of β in equation (4) using fixed-person-effects differ significantly from those estimated from equation (5). We interpret a rejection of the equality of coefficients as evidence that $\mathrm{Depressed}_i$ is correlated with μ_i, and that therefore the fixed effects estimator is more likely to provide consistent estimates of β.

Finally, any plausible causal effects of depression on obesity must necessarily operate through shifts in eating habits and physical activity – the proximate causes of obesity. We utilize multiple years of the Behavioral Risk Factor Surveillance System (BRFSS), a cross-sectional individual-level nationally-representative dataset, to assess whether depressive symptoms are associated with consumption of healthy and unhealthy foods and with physical activity. Due to the cross-sectional nature of the data, only a contemporaneous relationship can be examined. However, while obesity and BMI are long-term outcomes, changes in behaviors related to caloric intake and expenditure are expected to be more responsive in the short-term.[34] Furthermore, if data from the NCS-R and NLSY79 indicate an effect of depression on obesity, but depressive symptoms are not associated with changes in eating or physical activity, then this casts some doubt on a causal interpretation of the results. The use of these multiple datasets, samples, and methods, in addition to an exploration of proximate mechanisms, allows us to gauge robustness based on potentially reinforcing (or competing) results.

All models are estimated separately for males and females since prior studies point to significant gender differences in the prevalence of depression and its association with eating disorders and coping mechanism. Reported standard errors are corrected for arbitrary correlation within the appropriate cells: robust standard errors for models based on the NCS-R, clustered at the family level for models based on the NLSY79, and clustered at the state-year level for models

* See Rashad and Kaestner[33] for a discussion of identification issues with respect to instruments and exclusion restrictions in the bivariate probit framework.

† Also note that, if year effects are included, age is identified when equation (4) is estimated with person-specific fixed effects because the change in age between surveys varies among respondents. For example, someone who turns 40 in 1998 answers the CES-D questions 4 years after the 1994 survey, while someone who turns 40 in 2006 answers them 12 years later (see data description).

based on the BRFSS. Probit models are estimated for dichotomous outcomes, and OLS models are estimated for continuous outcomes.

Data

National Comorbidity Survey – Replication

The National Comorbidity Survey, fielded from 1990-1992, was the first nationally representative mental health survey in the U.S. to use a fully structured research diagnostic interview to assess the prevalence and correlates of DSM-III-R disorders.[35] The National Comorbidity Survey Replication (NCS-R) was conducted in 2001-2002 on a new sample of 10,000 respondents, with diagnoses of mental disorders, including major depressive episode, based on DSM-IV guidelines and representing the most accurate and detailed indicators of mental health available in any nationally-representative study. In addition, the NCS also includes extensive information on parental history, health investments, family background, and life events, which are typically unobserved in other datasets. We restrict the analyses to adults between the ages of 21 and 64.

In addition to the conventionally observed socio-demographic factors, the NCS-R also contains rich information on (i) parental history (for instance, parental education, history of depression and substance abuse, paternal occupation); (ii) family background (for instance, number of siblings, living arrangements when growing up, whether family received governmental assistance, number of times family moved when respondent was a child); (iii) parental investments (for instance, parental supervision, parental effort into upbringing); (iv) risk-taking (for instance, own alcohol and drug abuse, proxies for a risk-taking personality); (v) traumatic life events (for instance, being a victim of rape or sexual abuse, affected by the death of a parent or relative, etc); and (vi) use of anti-depressants and other prescription drugs, the latter being a proxy for health status and chronic health conditions. These measures (described in **Table 1**) are typically unobserved in other datasets; thus, comparing models with and without these factors allows us to assess the extent of the bias due to selection on these typically unobserved (but in our case, observed) factors*.

* We exclude the traumatic life events from our sample in order to maximize sample size and statistical precision since information on these life events is missing for about 20% of the sample. Incorporating these measures into our analyses does not materially affect the results or conclusions

Table 1. National Comorbidity Survey Replication (NCS-R) Weighted Sample Means

Variable	Description	Male		Female	
		No Lifetime Diagnosis of MDE	Lifetime Diagnosis of MDE	No Lifetime Diagnosis of MDE	Lifetime Diagnosis f MDE
BMI	Body mass index - kg/m^2	28.152	28.203	26.802	27.617
		(0.23)	(0.30)	(0.26)	(0.28)
Overweight/Obese	BMI >= 25	0.731	0.728	0.512	0.563
		(0.02)	(0.03)	(0.02)	(0.02)
Current Major Depressive Episode (MDE)	R had DSM-IV Major Depressive Episode in the last 12 months	_	0.399	_	0.431
			(0.03)		(0.02)
Past MDE	R had DSM-IV Major Depressive Episode more than a year ago	_	0.601	_	0.569
			(0.03)		(0.02)
Age	Age of respondent	40.427	41.868	41.045	41.596
		(0.48)	(0.65)	(0.45)	(0.46)
Black	African-American	0.075	0.052	0.111*	0.064*
		(0.01)	(0.01)	(0.01)	(0.01)
Other Race	Race not African-American or Non-Latino White	0.015*	0.047*	0.027	0.029
		(0.00)	(0.01)	(0.01)	(0.01)
Married	Respondent is married	0.678*	0.582*	0.638	0.560
		(0.02)	(0.03)	(0.02)	(0.02)
Divorced	Respondent is divorced	0.101*	0.197*	0.168*	0.253*
		(0.01)	(0.02)	(0.01)	(0.02)
High School grad	Respondent has a high school diploma	0.906	0.922	0.925	0.931
		(0.01)	(0.02)	(0.01)	(0.01)
College or more	Respondent has at least a bachelor's degree	0.327	0.296	0.337	0.310
		(0.02)	(0.03)	(0.02)	(0.02)
Any religious preference	Respondent has some religious preference	0.818	0.795	0.894	0.824
		(0.02)	(0.02)	(0.01)	(0.02)
Household income	Household income (in $1000's)	75.769	72.677	65.759	63.694
		(2.18)	(3.26)	(1.93)	(2.01)

➔

(continued)

Table 1. National Comorbidity Survey Replication (NCS-R) Weighted Sample Means

Variable	Description	Male		Female	
		No Lifetime Diagnosis of MDE	Lifetime Diagnosis of MDE	No Lifetime Diagnosis of MDE	Lifetime Diagnosis f MDE
Northeast	R lives in the Northeast	0.179	0.202	0.149	0.167
		(0.02)	(0.03)	(0.01)	(0.02)
Midwest	R lives in the Midwest	0.251	0.235	0.242	0.273
		(0.02)	(0.02)	(0.02)	(0.02)
South	R lives in the South	0.354	0.267	0.376	0.337
		(0.02)	(0.03)	(0.02)	(0.02)
Mom depressed	When R growing up, mother had periods of sadness 2+ weeks	0.144*	0.297*	0.185*	0.369*
		(0.01)	(0.03)	(0.02)	(0.02)
Mom drink or drugs	When R growing up, mother had problems with alcohol or drugs	0.046*	0.105*	0.058	0.101
		(0.01)	(0.02)	(0.01)	(0.01)
Dad depressed	When R growing up, father had periods of sadness 2+ weeks	0.076*	0.173*	0.074*	0.169*
		(0.01)	(0.02)	(0.01)	(0.02)
Dad drink or drugs	When R growing up, father had problems with alcohol or drugs	0.182*	0.309*	0.162*	0.274*
		(0.02)	(0.03)	(0.01)	(0.02)
Parents attempted suicide	At least one of R's parents attempted suicide	0.023	0.051	0.037	0.070
		(0.01)	(0.01)	(0.01)	(0.01)
Lived with biological parents	Lived with both biological parents until 16	0.781	0.767	0.783	0.787
		(0.02)	(0.02)	(0.02)	(0.02)
Number of siblings	Number of siblings	3.360	3.202	3.468	3.44
		(0.10)	(0.14)	(0.10)	(0.10)
Government assistance	R's family received government assistance for 6+ months in childhood	0.062	0.085	0.046	0.074
		(0.01)	(0.02)	(0.01)	(0.01)
Never unsupervised	R never left unsupervised in childhood	0.883*	0.759*	0.900*	0.783*
		(0.01)	(0.03)	(0.01)	(0.02)
Father years of education	R's father's years of education	11.569	11.839	11.548	11.598
		(0.18)	(0.21)	(0.14)	(0.15)
Parental effort	R's parents put "a lot" of effort into upbringing	0.935*	0.858*	0.909*	0.824*
		(0.01)	(0.02)	(0.01)	(0.02)
Times moved	Number of times R moved to a different neighborhood in childhood	2.090	1.907	1.671*	2.27*
		(0.14)	(0.15)	(0.09)	(0.13)
Current smoker	Currently a smoker	0.231*	0.328*	0.226	0.295
		(0.02)	(0.03)	(0.02)	(0.02)
Past smoker	Past smoker	0.238	0.235	0.207	0.248
		(0.02)	(0.03)	(0.02)	(0.02)
Alcohol abuse lifetime	DSM-IV Alcohol abuse lifetime	0.183*	0.397*	0.053*	0.158*
		(0.01)	(0.03)	(0.01)	(0.02)
Illicit drug abuse lifetime	DSM-IV drug abuse lifetime	0.101*	0.284*	0.033*	0.107*
		(0.01)	(0.03)	(0.00)	(0.01)
Take chances	Takes chances or does reckless things	0.173*	0.313*	0.056*	0.134*
		(0.01)	(0.03)	(0.01)	(0.01)
Urges	R says giving into urges gets him/her into trouble	0.378*	0.584*	0.307*	0.482*
		(0.02)	(0.03)	(0.02)	(0.02)
Antidepressants	R took antidepressants in last 12 months	0.038*	0.216*	0.091*	0.310*
		(0.01)	(0.02)	(0.01)	(0.02)
Any prescriptions	R took any prescription medication in last 12 months	0.080*	0.307*	0.153*	0.402*
		(0.01)	(0.03)	(0.01)	(0.02)
Post911	Interview took place after September 11, 2001	0.444	0.486	0.401	0.430
		(0.02)	(0.03)	(0.02)	(0.02)

Note: Weighted sample means are presented, with standard deviations reported in parentheses. Maximum observations are 3,229.. Asterisks denote a statistically significant difference in means between ever-depressed and never-depressed individuals, by gender, as follows: *p-value = 0.05

Table 1 presents weighted means for individuals who have been diagnosed with major depression at any point in their lifetime versus individuals with no lifetime diagnosis, across gender. The prevalence of major depression over one's lifetime is significantly higher among females (34.9%) relative to males (21.6%). Approximately 43.1% (39.9%) of females (males) with lifetime depression have a current or past-year diagnosis, and approximately 56.9% (61.1%) are not currently depressed but have had a diagnosis at some point beyond the past year. Women with lifetime depression are also more likely to be overweight or obese by about 5.1 percentage points (56.3% among women with lifetime depression versus 51.2% among women with no lifetime depression). In contrast to women, there are no significant differences in BMI or the prevalence of overweight/obesity among males by lifetime depression status. The figures further show that lifetime depression among females is correlated with other observed and typically- unobserved characteristics. For example, women with lifetime depression are less-educated, have lower household income, less likely to be married and more likely to be divorced, and less likely to have a religious affiliation. They are also more likely to have had parents who were depressed, had substance abuse issues, received governmental assistance, and relocated more when the respondent was a child; women with lifetime depression also reported reduced parental investments in the form of supervision and effort into their upbringing. Lifetime depression among women is also associated with other proxies for risk-taking and risky behaviors (smoking, drinking, illicit drug use, a more risk-taking personality), and as expected, both males and females with lifetime depression report greater use of anti-depressants and prescription medications. Thus, individuals who have an episode of major depression in their lifetime are not a random sample. They are also more likely to differ along characteristics that generally are associated with higher or lower bodyweight. The multivariate analyses account for these differences.

National Longitudinal Survey of Youth – 1979

The NLSY79 is a nationally representative sample of 12,686 young men and women who were 14-22 years old when they were first surveyed in 1979. Participants were interviewed annually from 1979 to 1994, and biennially since then. The primary variables of interest for this study are respondents' body-mass index (BMI) and the Center for Epidemiologic Studies Depression Scale (CES-D). Data on weight were collected in the surveys of 1981, 1982, and every survey from 1985 onwards, except 1987 and 1991. Data on height were collected in five surveys: 1981, 1982, 1985, 2006, and 2008. To compute BMI we use the average of the reported heights in 1985 and 2006, which for almost all respondents were identical.

Respondents were asked questions from the CES-D four times, the surveys of 1992, 1994, and as part of the Over-40 and Over-50 health modules. As all respondents had turned 40 by 2006 but only a few had turned 50 by 2008, we exclude the latter responses from our analysis. In 1992, respondents were asked all 20 CES-D questions, in 1994, 7 questions, and in the post-40 interview, 9 questions. Since we want to compare respondents' mental state over time, we use the overlapping 7 questions asked in 1994. This should not compromise our analysis as a subset of 5 CES-D questions, all of which are included in our group of 7, has been found to capture almost all of the information in the full set.[36]

The seven CES-D questions we use relate to poor appetite, trouble keeping one's mind on tasks, feeling depressed, feeling that everything takes extra effort, sleeping poorly, feeling sad, and feeling that one could not get going. The respondent was asked how often he or she experienced each problem. The answers are scored as zero for ''Rarely or none of the time'', 1 for ''Some of the time'', 2 for ''A moderate amount of the time'', and 3 for ''Most or all of the time.'' The sum of the responses for each condition gives the CES-D score. Thus scores can range from zero to 21. For comparison with the NCS-R, we compute a dichotomous indicator for depression if the score is 10 or greater.[37] Alternately, models (not reported) which utilize the continuous CES-D score yield similar patterns of results and conclusions.

Other continuous variables include age and family size. Dichotomous indicators are defined for marital status, highest grade completed, and region of residence. Since net family income is missing for many respondents (and preliminary analysis suggests that missing values are not missing at random), we include a variable indicating the dollar value of food stamps the respondent's family received in the previous calendar year. This variable has few missing values compared with poverty status and family income, and reflects the income of the whole family.

We organize the NLSY79 data into three waves. Wave 1 is 1992, Wave 2 is 1994, and Wave 3 is the year in which the respondent took part in the over-40 Health Module. Since fixed-effects models require at least two periods, our potential sample size is the number of respondents who contributed data to at least two of the three waves. Excluding respondents who were pregnant at the time of the surveys, 8,875 participants took part in at least two waves. Of these, 8,732, or 94 percent, have valid data for BMI and CES-D for at least two waves. After accounting for missing information on additional covariates, the ''analysis sample'' has 21,365 observations across three waves. When the analysis sample is further restricted to respondents providing valid data for all variables for all three waves, which we denote the ''balanced sample'', the sample size drops to 17,538.

We assess the external validity of the analysis and balanced samples, that is whether the smaller samples give the same results as would the full sample had not observations been lost to missing values. We gauge the importance of losing observations by regressing BMI on lagged CES-D depression status in all three samples and comparing the coefficients. We find that the coefficients are practically and statistically indistinguishable. The fixed effects coefficients are -0.15 (0.1) in the full sample,0.21 (0.11) in the analysis sample, and -0.28 (0.1) in the balanced sample. None of these pairs of coefficients is statistically significantly different at conventional levels, adding a level of confidence that the estimates and conclusions are not affected across the alternate samples.

Table 2 shows means and standard errors of our analysis sample by wave. Note that time-constant variables such as gender and race change slightly from wave to wave because the analysis sample is unbalanced. The means, as well as regressions discussed below, are calculated with probability weights supplied by the NLSY.

Table 2. National Longitudinal Survey of Youth-1979 (NLSY79) Weighted Sample Means

Variable	Description	Wave I (1992)	Wave II (1994)	Wave II (1998-2006)
Depressed	CES-D Score >= 10	0.102	0.094	0.089
		(0.303)	(0.292)	(0.285)
BMI	Body mass index	26.4	26.9	28.3
		(0.053)	(0.055)	(0.059)
Overweight / Obese	BMI >= 25	55.2	59.2	69.7
		(0.497)	(0.492)	(0.460)
Hispanic	Dichotomous indicator for Hispanic	18.4	18.6	18.7
		(0.387)	(0.389)	(0.390)
Non-Hispanic Black	Dichotomous indicator for Non-Hispanic black	29.1	29.2	29.8
		(0.454)	(0.455)	(0.458)
Other Race	Dichotomous indicator for other race	0.9	0.8	0.7
		(0.094)	(0.089)	(0.084)
Age	Age at interview	31.5	33.5	41.1
		(2.2)	(2.2)	(1.8)
Northeast	Dichotomous indicator for whether respondent resides in the Northeast	0.164	0.160	0.153
		(0.370)	(0.367)	(0.360)
Midwest	Dichotomous indicator for whether respondent resides in the Midwest	0.239	0.239	0.241
		(0.426)	(0.427)	(0.428)
South	Dichotomous indicator for whether respondent resides in the South	0.393	0.400	0.413
		(0.488)	(0.490)	(0.492)
Current Smoker	Dichotomous indicator for current smoker	0.338	0.333	0.295
		(0.473)	(0.471)	(0.456)
Food Stamps	Value of food stamps in $1000's	0.3	0.2	0.1
		(0.9)	(0.8)	(0.7)
Never married	Dichotomous indicator for never married	0.290	0.258	0.182
		(0.454)	(0.438)	(0.386)
Separated	Dichotomous indicator for separated	0.057	0.058	0.051
		(0.232)	(0.234)	(0.219)
Divorced	Dichotomous indicator for divorced	0.104	0.118	0.166
		(0.305)	(0.323)	(0.373)
Widowed	Widowed	0.004	0.006	0.011
		(0.067)	(0.075)	(0.103)
Highest Grade	Highest grade completed	13.0	13.0	13.3
		(2.4)	(2.4)	(2.5)
Family Size	Family size	3.2	3.3	3.3
		(1.7)	(1.6)	(1.6)
Mother High School	Dichotomous indicator for whether the respondent's mother is a high school graduate	0.422	0.423	0.431
		(0.494)	(0.494)	(0.495)
Mother Some College	Dichotomous indicator for whether the respondent's mother has had some college	0.095	0.095	0.094
		(0.294)	(0.293)	(0.291)
Mother College plus	Dichotomous indicator for whether the respondent's mother is a college graduate or above	0.075	0.075	0.076
		(26.3)	(26.4)	(26.5)
Savings Account	Dichotomous indicator for whether the respondent or their spouse had a savings account in 1985	0.620	0.617	0.619
		(0.485)	(0.486)	(0.486)
Life Insurance	Dichotomous indicator for whether respondent has life insurance through job	0.529	0.530	0.528
		(0.499)	(0.499)	(0.499)
Observations		7,179	7,369	6,817

Note: Weighted sample means are presented, with standard deviations reported in parentheses. Observations noted represent the maximum sample sizes; sample sizes in some models are lower due to missing information on some covariates. Asterisks denote a statistically significant difference in means between ever-depressed and never-depressed individuals, by gender, as follows: ***p-value = 0.01, **0.01<p-value = 0.05, *0.05<p-value = 0.10.

The figures show that on average BMI rose and the sum of CES-D responses declined slightly over the time period of the study. Thus a simple fixed-effects regression will find a negative correlation between the two variables. But since the respondents also aged and became less healthy, identification of the effect of mood on BMI depends on there being a significant proportion of respondents whose CES-D sum moved in the opposite direction to age and health, that is, whose CES-D sum rose over the sample period. Identification should therefore be possible since the CES-D sum rose from one wave to the next for about a third of the sample, and the average increase for these respondents was over 20%.

Behavioral Risk Factor Surveillance System

The Behavioral Risk Factor Surveillance System (BRFSS) is an individual-level data set representative of the population of the United States. As the largest telephone-based health survey available, the BRFSS has tracked health conditions and risk behaviors for adults 18 years of age and older in the U.S. The survey is conducted by state health departments in collaboration with the Centers for Disease Control. While only 15 states participated in 1984, the number grew to 33 in 1987, to 45 in 1990, and to all 51 states (including the District of Columbia) in 1996. More than 350,000 adults are interviewed each year, with response rates hovering around 50%.* The average number of interviews per state ranged from approximately 800 in 1984 to around 3,500 in more recent years. These data are publicly available from the Centers for Disease Control at http://www.cdc.gov/brfss, and provide information on a variety of personal characteristics, including gender, age, education, marital status, family income, and state of residence.

Measures of food consumption are included, although not consistently. Moreover, these variables are occasionally 'module' variables, asked of only a limited number of respondents, rather than 'core' variables, asked of all respondents. Consumption of carrots, fruit, fruit juice, green salad, and vegetables are asked consistently in years 1990-2007, with the exception of 2004 and 2006. The survey questions are generally phrased as follows: ''How often do you eat (FOOD)?'' Options are given for the respondent to record his/her answer in times per day, week, month, or year. Answers are converted to times per year for the purposes of this paper. While nutritionists caution using the terms 'healthy' and 'unhealthy' regarding foods in order to avoid classifying foods per se in preference for a focus on a balanced diet, we use the term healthy for the aforementioned foods as the food pyramid stresses their consumption. Moreover, most Continuing Surveys of Food Intakes by Individuals show consumption of fats, oils, and sweets, meant to be consumed sparingly, to be higher than recommended.† Consumption of snacks, hamburgers, hot dogs, French fries, and fried chicken is not as frequently observed in the BRFSS, yet we also analyze these outcomes in order to compare these results with those of our healthy food outcomes.

Physical activity is measured in the BRFSS based on reported activities, their intensity as measured by the metabolic equivalent (MET), and duration. Individuals are classified as: (i) physically inactive, if they report no physical activity; (ii) engaging in irregular activity, if they participate in any physical activity or pair of activities done for less than 20 minutes, or less than 3 times a week; (iii) engaging in regular activity if they report any physical activity of pair or activities done for 20 or more minutes, 3 or more times per week, less than 50 percent of capacity; and (iv) engaging in vigorous activity, if they report any physical activity or pair of activities that requires rhythmic contraction of large muscle groups at 50% functional capacity for 20 or more minutes, three or more times per week. Based on this classification, we define a dichotomous indicator for whether the individual currently engages in regular or vigorous physical activity during an average week.

Relative to the NCS-R and the NLSY79, which measure major depression based on clinical guidelines and a tested scale, the BRFSS includes only a single variable probing whether the individual's mental health, including stress, depression, and emotional problems, was ''not good'' during the past 30 days. We utilize a dichotomous indicator for whether the individual reported at least one day in the past month when mental health was not good, as indicative of depression symptoms, with the caveat that this measure may be more reflective of milder mental health issues such as daily stress and short-term depression than major depression.* As shown in **Table 3**, the mean number of days in the past month that respondents report feeling depressed, stressed, or experiencing emotional or mood problems is 3.4, and 32.8% of the BRFSS sample reports some mental health issues during the past month.

Measures of Depression

The alternate measures of depression across the three datasets likely capture differences in intensity and gradation. The measure of past year major depression in the NCS-R is based on DSM-III-R, the ''gold standard'' of measured depressive episodes in observational data. The measure of current (past week) depression in the NLSY-79 is based on the CES-D scale, also a highly reliable instrument used in observational data. It is only for our supplementary analyses based on the BRFSS, which assess potential mechanisms, where we are limited to self-reported depressed or sad state over the past month. In terms of ordering, the DSM-III-R measure captures the strictest cases of a major depressive episode, the CES-D scale also picks up milder forms of depression in

* Survey weights are included in the BRFSS to ensure that those included in the survey are reflective of the U.S. population. In addition, the study shows that means for those responding and the general population are comparable. (See http://www.cdc.gov/brfss.)

† See http://www.ars.usda.gov/Services/docs.htm?docid=14392.

* Results are not sensitive to using the continuous measure of the number of days in the past month that mental health was not good. Utilizing multiple indicators (0 days, 1-5 days, 6-10 days, 11-15 days, 15+ days) also indicates a dose-response relationship; greater frequency of mental health issues in the past month has larger effects on food consumption and physical activity. Results are available upon request.

Table 3. Behavioral Risk Factor Surveillance System (BRFSS) Weighted Means

Variable	Description	Mean
Days Depressed	Number of days in past month mental health was not good due to depression or stress	3.448 (7.465)
Depressed	Dichotomous indicator for whether mental health was not good at any time in past month	0.3280 (0.4695)
Fruits	Annual consumption of fruits	289.759 (305.511)
Fruit Juice	Annual consumption of fruit juice	248.672 (303.700)
Carrots	Annual consumption of carrots	96.096 (154.792)
Green Salad	Annual consumption of green salad	181.709 (187.952)
Vegetables	Annual servings of vegetables (not carrots, potatoes, or salad)	455.376 (346.701)
Snacks	Annual consumption of snacks	119.856 (167.966)
Hamburgers	Annual consumption of hamburgers, cheeseburgers, or meatloaf	76.944 (91.634)
Hot Dogs	Annual consumption of hot dogs	91.295 (201.530)
French Fries	Annual consumption of fries	64.594 (111.107)
Fried Chicken	Annual consumption of fried chicken	33.538 (68.163)
Physically Active	Dichotomous indicator for whether respondent currently engages in regular or vigorous physical activity in an average week	0.4283 (0.4948)
Male	Dichotomous variable that equals 1 if respondent is male, and 0 if respondent is female	0.496 (0.500)
Some High School	Dichotomous variable that equals 1 if respondent completed at least 9 but less than 12 years of school	0.072 (0.258)
High School	Dichotomous variable that equals 1 if respondent completed exactly 12 years of schooling	0.304 (0.460)
Some College	Dichotomous variable that equals 1 if respondent completed at least 13 but less than 16 years of school	0.269 (0.443)
College	Dichotomous variable that equals 1 if respondent graduated from college	0.316 (0.465)
Age	Age of respondent in years	39.468 (11.489)
Married	Dichotomous variable that equals 1 if respondent is married	0.643 (0.479)
Divorced	Dichotomous variable that equals 1 if respondent is divorced or separated	0.124 (0.330)
Widowed	Dichotomous variable that equals 1 if respondent is widowed	0.018 (0.134)
Black	Dichotomous variable that equals 1 if respondent is black and not Hispanic	0.098 (0.297)
Hispanic	Dichotomous variable that equals 1 if respondentis of Hispanic origin	0.118 (0.322)
Other Race	Dichotomous variable that equals 1 if respondent's race is other than white, black, or Hispanic	0.051 (0.220)
Work	Dichotomous variable that equals 1 if respondent is Employed	0.782 (0.413)
Unemployed	Dichotomous variable that equals 1 if respondent is out of work or unable to work	0.092 (0.289)
Family Income	Real household income in thousands of 1982-84 Dollars	33776.360 (26876.690)
Physical Health	Number of days in past month physical health not Good	2.915 (7.049)
Health Plan	Dichotomous variable that equals 1 if respondent currently has health insurance	0.829 (0.377)
Observations		2,858,973

Note: Weighted sample means are reported, and standard deviations are reported in parentheses. Observations denote the maximum sample size; sample sizes in some models are less due to missing information on some covariates.

addition to the strict cases, and the instrument in the BRFSS captures self-identified feelings of sadness and depression, which may capture the broadest cases.

As a cross-validation check, we redefined alternate depression measures to assess consistency and comparability across our two datasets which do not use the DSM-III-R. For instance, 9.1% of individuals report feeling depressed or sad in the past week for at least half the time, based on the NLSY-79. This compares to 9.3% in the BRFSS with respect to 15 or more days in the past month. The prevalence is higher among females (11.2% in the NLSY-79 and 11.1% in the BRFSS) relative to males (7.7% and 7.6%, respectively). The prevalence is also higher among non-Hispanic blacks relative to non-Hispanic whites. The stability in these patterns across alternate measures and datasets is reassuring.

Results

Table 4 and **Table 5** present estimates of the effects of current, past, and lifetime depression on BMI and the probability of being overweight or obese (BMI ≥ 25), separately for males and female respectively, using data from the NCS-R. As reported in **Table 4** (Panel A), we do not find any significant effects of current, past, or lifetime depression on the body mass index of males. In Panel B, current depression is negatively associated with the probability of being overweight or obese in the basic specification. Some prior studies, based on clinical samples, have reported that weight-loss may be a symptom of depression particularly among males,[38-40] noting that males may cope with depression differently than females, sometimes by raising their sports activity, raising smoking or alcohol use, and being affected more in their ability to work. However, this effect drops to zero and becomes insignificant in the extended specification. Similarly, neither lifetime nor past depression is significantly associated with the probability of being overweight or obese. The inconsistency of the estimates between the basic and extended models suggests the selection bias may be driving any observed association between depression and obesity among males.

Table 5 presents estimates for females. In the basic specifications, current depression is associated with a higher BMI and a higher probability of being overweight/obese (4.4 percentage points). However, these diminish substantially and become statistically insignificant in the extended specifications, suggesting that the association between contemporaneous depression and obesity is not very likely to be reflective of a causal relationship. This is to be expected given that bodyweight changes slowly over time, and if depression does raise bodyweight then this should be more apparent in the longer-term. This is indeed the case for females, where we find that past depression is more strongly associated with higher BMI (relative to males). Being diagnosed with depression at some point beyond the current year raises current BMI by about one point. Given that the effect of current depression is very small, the effect of lifetime depression mostly reflects the impact of past depression – raising BMI by about the same magnitude.

Past depression is also associated with a significant 5.5 percentage points increase in the probability of being overweight or obese. This represents an 11% increase in the probability relative to the mean prevalence of overweight/obesity among individuals with no lifetime depression. Also note that this estimated effect is slightly larger than the effect suggested by the simple difference in means (5.1 percentage points) reported in **Table 1**. Indeed, controlling for the typically unobserved factors in the extended specification raises the effect magnitude even further from 5.5 to 6.6 percentage points (13% increase relative to the mean prevalence). This is suggestive of negative selection. That is, there are several factors in our models that raise the likelihood of depression but reduce the likelihood of obesity and vice versa – for instance being single, non-Black, religious preference, household income, current smoker, maternal depression, parental suicide attempt, lived with both biological parents when young, and use of prescription drugs other than anti-depressants.

The effects of the other covariates on BMI and on the likelihood of being overweight or obese are generally consistent with those reported in the literature.[29] The BMI-age profile is quadratic, with BMI (and the probability of being overweight or obese) generally increasing with age up to the mid-50s and then declining, consistent with a loss in muscle mass after late adulthood. BMI is higher among individuals who are black or married and are low-educated. The latter is consistent with evidence suggesting that education has a causal protective effect on health, operating partly through educated individuals being more allocatively efficient and consuming healthier inputs.[41] Higher household income is also associated with a lower BMI among females, but not males. Being a current smoker also reduces BMI among both sexes. Among the extended covariates (not reported), maternal depression, living with both biological parents, having more educated parents, being a current smoker or abusing alcohol and/or drugs in one's lifetime, and taking prescription medications (other than anti-depressants) are associated with a lower probability of being overweight/obese. In contrast, paternal substance abuse, parental suicide attempt, moving more times when the respondent was a child, and having a risk-taking personality raise the probability of being overweight/obese. Also, as expected, the use of anti-depressants raises BMI and the probability of being overweight/obese.

While results among females consistently show a significant impact of past and lifetime depression on obesity after accounting for these extended covariates, the possibility of additional selection on unobservables remains. **Table 6** presents estimates from constrained-selection bivariate probit models to assess the sensitivity of the estimated effects from imposing additional amounts of selection on unobservables. Model 1 reports the single-equation probit estimates for baseline comparison, which assumes that correlation in the disturbance terms between the obesity and depression equations (equations 1 and 2) is zero. Models 2-6 constrain Rho incrementally from 0.1 to 0.5 in order to assess how the estimated effect of depression on obesity changes with increasingly positive selection on unobservable factors. That

Table 4. Effects of Major Depression on BMI & Overweight/Obesity among Males Data: NCS-R

Panel A	**Dependent Variable: BMI**					
Variable	1	2	3	4	5	6
Current Depression	−0.4639	−0.0063	–	–	–	–
	[0.40]	[0.48]				
Past Depression	–	–	0.5214	0.3702	–	–
			[0.34]	[0.36]		
Lifetime Depression	–	–	–	–	0.1071	0.2724
					[0.28]	[0.32]
Age	0.2875	0.3331	0.2778	0.3330	0.2786	0.3325
	[0.08]	[0.10]	[0.08]	[0.10]	[0.08]	[0.10]
Age Squared	−0.0029	−0.0034	0.0028	−0.0034	0.0028	−0.0034
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Black	0.3303	0.7998	0.3584	0.7936	0.3512	0.7992
	[0.44]	[0.63]	[0.44]	[0.63]	[0.44]	[0.63]
Other Race	−0.2129	−0.4489	−0.2570	−0.4842	−0.2529	−0.4982
	[0.75]	[0.83]	[0.75]	[0.84]	[0.74]	[0.83]
Married	1.0029	0.5531	1.0355	0.5468	1.0463	0.5619
	[0.37]	[0.44]	[0.37]	[0.44]	[0.37]	[0.44]
Divorced	0.1637	−0.1358	0.1159	−0.1655	0.1382	−0.1682
	[0.49]	[0.57]	[0.49]	[0.57]	[0.49]	[0.57]
High School Graduate	0.8859	0.5221	0.8682	0.4964	0.8983	0.5096
	[0.39]	[0.50]	[0.39]	[0.51]	[0.39]	[0.50]
College Graduate	−1.0913	−1.3132	−1.0913	−1.3196	−1.0785	−1.3125
	[0.28]	[0.31]	[0.27]	[0.31]	[0.28]	[0.31]
Any Religious Preference	0.9278	0.8349	0.9417	0.8416	0.9270	0.8315
	[0.32]	[0.34]	[0.32]	[0.34]	[0.32]	[0.34]
Household Income	0.0002	−0.0014	0.00014	−0.0015	0.0003	−0.0014
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Northeast	0.4707	0.4713	0.4827	0.4813	0.4685	0.4721
	[0.39]	[0.44]	[0.39]	[0.44]	[0.39]	[0.44]
Midwest	0.3894	0.3910	0.4069	0.4021	0.4002	0.3985
	[0.34]	[0.37]	[0.34]	[0.37]	[0.34]	[0.37]
South	0.326	0.1931	0.3658	0.2144	0.3468	0.2079
	[0.34]	[0.36]	[0.34]	[0.36]	[0.34]	[0.36]
Extended Covariates	No	Yes	No	Yes	No	Yes
R-squared	0.053	0.102	0.054	0.103	0.052	0.103
Observations	1669	1318	1669	1318	1669	1318

Note: Coefficients from OLS regression models are reported, with robust standard errors in brackets. Extended covariates include: mom depressed, mom drink or illicit drugs, dad depressed, dad drink or illicit drugs, lived with both parents, parent attempted suicide, number of siblings, government assistance to family when young, never unsupervised as a child, high parental effort, maternal/paternal education, times moved in childhood, current smoker, past smoker, lifetime diagnosis of alcohol abuse, lifetime diagnosis of illicit drug abuse, describes personality as ''taking chances'', troublesome urges, anti-depressant use, other prescription drug use, and interview took place after 9/11. Asterisks denote statistical significance of coefficients as follows: ***p-value = 0.01, **0.01<p-value = 0.05, *0.05<p-value = 0.10.

Panel B	**Dependent Variable: Overweight / Obese**					
Variable	1	2	3	4	5	6
Current Depression	−0.0516*	0.0053	–	–	–	–
	[0.03]	[0.03]				
Past Depression	–	–	0.0339	0.0254	–	–
			[0.03]	[0.03]		
Lifetime Depression	–	–	–	–	−0.0047	0.0212
					[0.02]	[0.03]
Extended Covariates	No	Yes	No	Yes	No	Yes
Observations	1965	1553	1965	1553	1965	1553

Note: Marginal effects from probit models are reported, with robust standard errors in parentheses. See notes for Panel A above.

Table 5. Effects of Major Depression on BMI & Overweight/Obesity among Females Data: NCS-R

Panel A	**Dependent Variable: BMI**					
Variable	1	2	3	4	5	6
Current Depression	0.8029**	0.0483	–	–	–	–
	[0.36]	[0.41]				
Past Depression	–	–	0.5471*	0.6164*	–	–
			[0.33]	[0.36]		
Lifetime Depression	–	–	–	–	0.8715***	0.5027
					[0.27]	[0.31]
Age	0.3069	0.3466	0.3063	0.3412	0.2944	0.3409
	[0.08]	[0.09]	[0.08]	[0.09]	[0.08]	[0.09]
Age Squared	−0.0029	−0.0034	−0.0029	−0.0034	−0.0027	−0.0034
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Black	2.6954	2.5001	2.6925	2.5523	2.7512	2.5226
	[0.43]	[0.55]	[0.43]	[0.55]	[0.43]	[0.55]
Other Race	1.0936	1.9154	1.0340	1.9065	1.0098	1.9034
	[0.77]	[0.84]	[0.77]	[0.85]	[0.77]	[0.85]
Married	0.8038	0.5662	0.75799	0.5789	0.8304	0.5970
	[0.40]	[0.45]	[0.40]	[0.45]	[0.40]	[0.45]
Divorced	0.4521	0.1750	0.4413	0.1562	0.4389	0.1645
	[0.46]	[0.54]	[0.46]	[0.54]	[0.46]	[0.54]
High School Graduate	−1.5292	−1.6663	−1.5900	−1.6910	−1.5808	−1.6883
	[0.49]	[0.71]	[0.49]	[0.71]	[0.49]	[0.71]
College Graduate	−1.2916	−1.0077	−1.2869	−0.9804	−1.2683	−0.9895
	[0.29]	[0.33]	[0.29]	[0.33]	[0.29]	[0.33]
Any Religious Preference	0.7886	0.6375	0.7935	0.6631	0.8330	0.6696
	[0.36]	[0.40]	[0.36]	[0.40]	[0.36]	[0.40]
Household Income	−0.01405	−0.0122	−0.0147	−0.0125	−0.0145	−0.0123
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Northeast	0.0973	0.3480	0.0991	0.3650	0.1201	0.3568
	[0.40]	[0.46]	[0.40]	[0.46]	[0.40]	[0.46]
Midwest	0.9233	0.9042	0.9236	0.8861	0.9017	0.8762
	[0.37]	[0.41]	[0.37]	[0.41]	[0.37]	[0.41]
South	0.2800	0.2359	0.2990	0.2324	0.2878	0.2302
	[0.35]	[0.40]	[0.35]	[0.40]	[0.35]	[0.41]
Extended Covariates	No	Yes	No	Yes	No	Yes
R-squared	0.086	0.123	0.085	0.125	0.0875	0.125
Observations	2485	1911	2485	1911	2485	1911

Note: Coefficients from OLS regression models are reported, with robust standard errors in brackets. Extended covariates include: mom depressed, mom drink or illicit drugs, dad depressed, dad drink or illicit drugs, lived with both parents, parent attempted suicide, number of siblings, government assistance to family when young, never unsupervised as a child, high parental effort, maternal/paternal education, times moved in childhood, current smoker, past smoker, lifetime diagnosis of alcohol abuse, lifetime diagnosis of illicit drug abuse, describes personality as ''taking chances'', troublesome urges, anti-depressant use, other prescription drug use, and interview took place after 9/11. Asterisks denote statistical significance of coefficients as follows: ***p-value = 0.01, **0.01<p-value = 0.05, *0.05<p-value = 0.10.

Panel B	**Dependent Variable: Overweight / Obese**					
Variable	1	2	3	4	5	6
Current Depression	0.0436*	0.0201	–	–	–	–
	[0.03]	[0.03]				
Past Depression	–	–	0.0551**	0.0656**	–	–
			[0.02]	[0.03]		
Lifetime Depression	–	–	–	–	0.0653**	0.0624**
					[0.02]	[0.02]
Extended Covariates	No	Yes	No	Yes	No	Yes
	2671	2040	2671	2040	2671	2040

Note: Marginal effects from probit models are reported, with robust standard errors in parentheses. See notes for Panel A above.

Table 6. Effects of Major Depression on Overweight/Obesity among Females Estimates from Constrained-Selection Models Data: NCS-R

Dependent Variable: Overweight / Obese										
Model	1	2	3	4	5	6	7	8	9	10
Constraint on Rho	Rho = 0	0.1	0.2	0.3	0.4	0.5	−0.1	−0.2	−0.3	Equal Selection[1]
Past Depression	0.0656*** [0.03]	−0.0014 [0.07]	−0.0702** [0.07]	−0.1398*** [0.07]	−0.2089*** [0.07]	−0.2762*** [0.07]	0.1299*** [0.07]	0.1905*** [0.07]	0.2471*** [0.07]	0.1565*** [0.07] Rho = −0.143
Lifetime Depression	0.0624*** [0.02]	−0.0020 [0.06]	−0.0672*** [0.06]	−0.1326*** [0.06]	−0.1977*** [0.06]	−0.2615*** [0.06]	0.1255*** [0.06]	0.1866*** [0.06]	0.2454*** [0.06]	0.1087*** [0.06] Rho = −0.073
Obs.	2040									

Note: Each cell represents a separate bivariate probit model, with alternate constraints on the correlation (Rho) between the error terms in the obesity and depression equations. Marginal effects are reported, with standard errors in brackets.

[1] Bivariate probit model is estimated by constraining rho such that the selection on unobservable factors is equal to the selection on observable factors.

is, if there are additional unobserved factors that positively affect both the probability of depression and obesity, how would this impact the estimated effect of depression on obesity?

Estimates from models 2-6 suggest that even small amounts of additional positive selection eliminate any positive impact of depression (both lifetime depression and past depression) on obesity. For instance, while the single-equation estimates suggest that lifetime- and past-depression raise the probability of being overweight/obese by about 6.5 percentage points, if additional unobserved factors are slightly positively correlated with the probability of both depression and obesity (Rho = 0.1) then these effects become essentially zero (− 0.001 to − 0.002) and insignificant. Additional amounts of positive selection (Rho = 0.2, 0.3, 0.4, 0.5) results in stronger negative effects of depression on the probability of being overweight or obese.

Models 7-9 impose additional amounts of negative selection by constraining Rho to be -0.1, -0.2, and -0.3. These models assess how the estimated impact of depression on obesity behaves if there are additional unobserved factors that raise (reduce) the probability of depression but reduce (raise) the probability of obesity. With additional amounts of negative selection on unobservables, the positive effect of depression on obesity remains robust and becomes increasingly stronger.

The last column (model 10) in **Table 6** uses selection on observable factors (including all the observed socio-demographic factors plus the typically unobserved factors that are available in the NCS-R) to inform on potential selection on unobserved factors. Specifically, Altonji et al. (2005) note that if the observed factors are a random subset of all factors affecting the outcome, then selection on unobservable factors should equal selection on observed factors. Model 10 applies this constraint and finds some evidence of negative selection (Rho = −0.07 to −0.14). Under this assumption of equal selection, past (lifetime) depression raises the probability of being overweight or obese by 15.6 (10.9) percentage points. Note that these effects are larger than those estimated by the single-equation probit models since the single-equation models do not account for selection on unobservables, and as the constrained bivariate probit models show, if there is additional negative selection effect then the estimated positive impact of depression on obesity becomes stronger.*

In summary, the results from the NCS-R suggest that past depression, in particular, is associated with a higher BMI and probability of being overweight and obese among females. Estimates from the extended models suggest that these effects are robust to controlling for various factors that are typically unobserved in other datasets. Constrained selection models indicate that additional negative selection on unobservables is likely, and that the positive effect of depression on obesity among females may be even stronger after accounting for this negative selection. In contrast, we do not find any significant and large positive effects of current depression on bodyweight or overweight/obesity. Estimates for males are generally insignificant and small in magnitude. This pattern of results suggests that a causal link from past depression to BMI and being overweight/obese is plausible among females but not males.

Table 7 presents the impact of lagged depression on BMI and on the probability of being overweight/obese among males, based on data from the NLSY79. Across all models, we do not find any substantial or significant effects of past depression on indicators or obesity among males, which is consistent with our previous results based on the NCS-R.

* As noted earlier, there are several factors in our models consistent with negative selection, that raise the likelihood of depression but reduce the likelihood of obesity and vice versa – for instance being single, non-Black, religious preference, household income, current smoker, illicit drug use, maternal depression, parental suicide attempt, lived with both biological parents when young, and use of prescription drugs other than anti-depressants. While there are also factors that lead to positive selection (for instance, age, education, other race, prior alcohol abuse), the degree of negative selection outweighs the positive selection on the net, reflected in the negative estimates of rho.

Table 7. Effects of Past Depression (CES-D scale) on BMI and Overweight/Obesity among Males Data: NLSY79

Panel A	**Dependent Variable: BMI**				
Model	1	2	3	4	5
Variable	Between Effects	Between Effects	Person Fixed Effects	First-Differences	Family Fixed Effects
Depressed	0.337	0.325	0.038	0.039	0.008
(Based on CES-D Scale)	(0.611)	(0.616)	(0.126)	(0.103)	(0.234)
Age at interview	−0.145	−0.241	0.411**	0.442**	0.090
	(0.670)	(0.676)	(0.135)	(0.142)	(0.237)
Age-squared	0.004	0.005	−0.003	−0.004*	0.001
	(0.010)	(0.010)	(0.002)	(0.002)	(0.003)
Lives in Northeast	0.657*	0.598*	−0.296	−0.499	−0.131
	(0.280)	(0.281)	(0.297)	(0.278)	(0.433)
Lives in Midwest	0.531*	0.482	0.391	0.122	−0.087
	(0.258)	(0.258)	(0.271)	(0.290)	(0.458)
Lives in South	0.624*	0.556*	0.011	−0.204	0.159
	(0.246)	(0.248)	(0.235)	(0.257)	(0.363)
Current Smoker	−1.749***	−1.711***	−0.135	0.027	−0.824***
	(0.219)	(0.220)	(0.110)	(0.096)	(0.218)
Food Stamps	0.537	0.564	−0.007	0.032	−0.141
	(0.453)	(0.451)	(0.093)	(0.066)	(0.099)
Never married	−0.483	−0.427	0.012	0.009	−0.523*
	(0.270)	(0.277)	(0.150)	(0.138)	(0.261)
Separated	−0.999	−0.990	−0.271	−0.090	−0.051
	(0.618)	(0.622)	(0.152)	(0.138)	(0.293)
Divorced	−0.791*	−0.775*	−0.210	−0.138	0.077
	(0.391)	(0.390)	(0.137)	(0.117)	(0.220)
Widowed	4.520*	4.481*	0.106	−0.201	1.166
	(2.005)	(2.012)	(1.157)	(0.818)	(0.801)
Highest grade completed	−0.284***	−0.254***	0.057	0.001	−0.226**
	(0.038)	(0.042)	(0.074)	(0.057)	(0.082)
Family size	−0.036	−0.046	−0.057*	−0.026	0.027
	(0.079)	(0.080)	(0.027)	(0.024)	(0.056)
Constant	31.577**	32.638**	16.399***	0.063	26.170***
	(11.351)	(11.448)	(2.615)	(0.057)	(4.437)
Extended Covariates	No	Yes	No	No	No
R-squared	0.047	0.050	0.201	0.042	0.113
Observations	3296	3296	10441	6361	10441
P-value Ho: RE valid	–	–	0.000	–	0.000

Note: Coefficients from OLS models are reported, standard errors clustered at the family-level in parentheses. Extended covariates include maternal education, and indicators of risk tolerance (savings account and purchase of life insurance). Asterisks denote statistical significance of coefficients as follows: ***p-value = 0.01, **0.01<p-value = 0.05, *0.05<p-value = 0.10.

Panel B	**Dependent Variable: Overweight / Obese**				
Model	1	2	3	4	5
Variable	Between Effects	Between Effects	Person Fixed Effects	First-Differences	Family Fixed Effects
Depressed	0.019	0.018	−0.013	0.003	−0.014
	(0.053)	(0.053)	(0.019)	(0.016)	(0.024)
Extended Covariates	No	Yes	No	No	No

Note: Marginal effects from probit models are reported. See additional notes above.

Table 8 presents similar estimates among females. Panel A presents estimates for BMI, and Panel B presents estimates for overweight/obesity. The basic specification (model 1) suggests that depression in the past wave is associated with about a one point increase in BMI. The estimate is not sensitive to controlling for the extended covariates (model 2), including maternal education and proxies for risk tolerance. However, the effect magnitudes become insignificant and essential drop to zero in the fixed effects and first-difference models. Similarly, in Panel B, we find that past depression raises the probability of being overweight or obese by between 6.7 (basic specification) and 7.6 percentage points

Table 8. Effects of Past Depression (CES-D scale) on BMI and Overweight/Obesity among Females Data: NLSY79

Panel A	**Dependent Variable: BMI**				
Model	1	2	3	4	5
Variable	Between Effects	Between Effects	Person Fixed Effects	First-Differences	Family Fixed Effects
Depressed	0.985*	1.091*	−0.187	−0.191	−0.100
	(0.591)	(0.585)	(0.129)	(0.108)	(0.243)
Age at interview	−3.267***	−3.178***	0.266	0.161	0.661*
	(0.818)	(0.816)	(0.173)	(0.186)	(0.319)
Age-squared	0.053***	0.052***	−0.001	−0.000	−0.007
	(0.012)	(0.012)	(0.002)	(0.002)	(0.004)
Lives in Northeast	−0.389	−0.410	−0.501	−0.556	−0.119
	(0.392)	(0.393)	(0.427)	(0.358)	(0.794)
Lives in Midwest	0.054	0.062	−0.522	−0.407	1.123
	(0.357)	(0.357)	(0.476)	(0.427)	(0.679)
Lives in South	0.431	0.376	−0.037	−0.241	−0.175
	(0.332)	(0.328)	(0.368)	(0.322)	(0.537)
Smokes yes/no	−1.516***	−1.548***	−0.114	0.048	−0.543
	(0.273)	(0.273)	(0.141)	(0.120)	(0.336)
Value of food stamps ($1000)	0.633**	0.550*	−0.036	−0.010	0.284*
	(0.217)	(0.221)	(0.060)	(0.049)	(0.115)
Never married	3.083***	2.954***	−0.394	−0.247	0.288
	(0.425)	(0.429)	(0.221)	(0.196)	(0.377)
Separated	0.992	0.869	−0.224	−0.030	−0.524
	(0.652)	(0.655)	(0.191)	(0.180)	(0.299)
Divorced	−0.221	−0.345	−0.174	−0.170	−0.181
	(0.440)	(0.434)	(0.153)	(0.140)	(0.272)
Widowed	1.322	1.276	0.460	0.413	−0.297
	(1.327)	(1.335)	(0.690)	(0.698)	(1.861)
Highest grade completed	−0.326***	−0.224***	0.141*	0.056	0.039
	(0.051)	(0.058)	(0.066)	(0.066)	(0.106)
Family size	0.295**	0.253*	−0.061	−0.080*	−0.076
	(0.107)	(0.107)	(0.039)	(0.037)	(0.073)
Constant	77.162***	74.820***	17.277***	0.172*	10.662
	(13.792)	(13.759)	(3.250)	(0.073)	(5.888)
R-squared	0.082	0.090	0.132	0.025	0.072
Observations	3217	3217	10924	6751	10924
P-value Ho: RE valid	–	–	0.000	–	0.000

Note: See **Table 7**.

Panel B	**Dependent Variable: Overweight / Obese**				
Model	1	2	3	4	5
Variable	Between Effects	Between Effects	Person Fixed Effects	First-Differences	Family Fixed Effects
Depressed	0.067*	0.076*	0.003	−0.002	0.026
	(0.040)	(0.040)	(0.015)	(0.013)	(0.019)
Extended Covariates	No	Yes	No	No	No

Note: See **Table 7**.

(extended specification). The effect magnitudes decline to zero in the person fixed effects (model 3) and first differenced models (model 4). While this may suggest that essentially all of the association between past depression and obesity is driven by unobserved person-specific factors, it should be noted that there is very limited variation in depression and obesity indicators within individuals over adjacent waves which may also lead to imprecision and inconsistent effects.

In order to maximize variation while still accounting for all unobserved time-invariant family characteristics (for instance, parental characteristics, family history, parental investments, past life events, genetic and health endowment, etc.), model 5 includes family-level fixed effects – suggesting that past depression raises the probability of being overweight/obese by 2.6 percentage points. This effect is biased downwards due to the measurement error in classifying an individual as depressed based on the CES-D score (versus using a more reliable indicator based on the DSM-IV guidelines as in the NCS-R data). Based on the level of misclassification (false positives and false negatives) associated with using a CES-D score threshold to diagnose major depression,[42] the estimate is biased downward by a factor of 3-4 times. Thus, the family fixed effects models suggests that past depression is associated with a 7.8-10.4 percentage points increase in the probability of being overweight or obese — these estimates are closely aligned with the single-equation and constrained bivariate probit estimates based on the NCS-R.

Table 9 and **Table 10** report estimates of the association of poor mental health with eating habits and physical activity. Among both males and females, we find that poor mental health in the past month reduces the consumption of healthy foods such as fruits and vegetables and increases the consumption on unhealthy foods such as snacks, French fries, hot dogs, and doughnuts. The estimates for measures of unhealthy foods are based on smaller sample sizes, rendering some of these effects imprecise; however, effect magnitudes for all measures of food consumption are generally robust across the basic and extended specifications. We also find that being of poor mental health in the past month is associated with a reduced probability of being physically active for both males and females. These proximate channels are consistent with depression leading to a higher BMI and raising the probability of being overweight and obese.

However, given the results from the NCS-R which suggest that past depression has a far stronger, positive effect on the bodyweight of females relative to males, we would expect stronger changes in food consumption and/or physical activity among females as a result of poor mental health. The effect magnitudes from the BRFSS are suggestive of this ''dose-response'' relation. For instance, scaling the effect magnitudes relative to the sample mean among mentally healthy individuals, poor mental health is associated with a 4-7% decline in healthy food consumption among females compared to a 2-4% decline among males. Substitution into high-calorie unhealthy foods is of similar magnitudes for both males and females, approximately a 4-12% increase. However, the decline in the prevalence of regular or vigorous physical activity is again far larger among females – 12.4% versus 4.3%. While these estimates should be interpreted with caution due to potential unobservables and simultaneity between food choices, physical activity and depression, the general pattern of results is consistent with a larger increase in BMI and overweight/obesity among females as a result of depression, relative to males, as noted earlier in the models estimated with the NCS-R.

Discussion

Unadjusted differences document a strong positive association between lifetime depression and overweight or obesity. The aim of this study was to examine whether, and the extent to which, this association is consistent with a causal link. Results are generally not supportive of a causal link among males. Among females, we find some evidence that past and lifetime depression significantly raise BMI and consequently the likelihood of being overweight or obese. These estimates are robust to controlling for a large vector of typically unobserved factors from the NCS-R as well as imposing an additional amount of selection on unobservable factors (based on the level of selection that is already present from observed factors). Expectedly, there are no significant or substantial positive contemporaneous effects of depression on bodyweight, given that bodyweight is a stock measure and responds over time to excess caloric intake. Panel data methods applied to the NLSY79 confirm this result. Robustness across multiple datasets and methods raises the level of confidence in the estimates.

In addition, these estimates appear to be plausibly driven by behavioral responses related to the proximate causes of obesity, that is shifts in food consumption and physical activity. Based on individual records from the BRFSS, we find that poor mental health is associated with lower healthy food consumption, part of which is substituted by higher consumption of unhealthy foods that are generally high in calorie and fat content. Poor mental health is also associated with a significantly lower likelihood of regular or vigorous physical activity. And, consistent with a stronger depression-obesity link among females, we find that the decrease in healthy food consumption and physical activity is also generally larger among females relative to males. Future research should focus on these gender differences and other distant mechanisms underlying the depression-obesity link.

Based on the NCS-R, 34.9% of adult women experience a major depressive episode at some point over their life. Healthy People 2020 aims at reducing this prevalence by 10% or by 3.5 percentage points. The estimates from this study suggest that this improvement in mental health could also help achieve another Healthy People 2020 objective aimed at increasing the proportion of adults who are at a healthy weight. Specifically, if the Healthy People 2020 objective relating to major depression is met, this would reduce overweight/obesity among women by between 0.22 and 0.55 percentage point, which represents about a 5% to

Table 9. Effects of Depression on Food Consumption and Physical Activity among Males Data: BRFSS

Panel A	Healthy Food Consumption									
Outcome	Fruit		Fruit Juice		Carrots		Green Salad		Vegetables	
Model	1	2	3	4	5	6	7	8	9	10
Depressed	−12.7027***	−11.9770***	−3.7174***	−5.3561***	−2.4712***	−2.6153***	−10.4296***	−7.9748***	−12.9913***	−12.2309***
	(0.8929)	(0.9602)	(0.9708)	(1.0974)	(0.4247)	(0.4567)	(0.5588)	(0.5842)	(1.1409)	(1.1137)
Extended Covariates	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
R-squared	0.037	0.037	0.032	0.031	0.015	0.014	0.051	0.055	0.047	0.046
Observations	783327	613530	785242	614545	775505	607725	786193	615388	779328	610885

Note: Coefficients from OLS models are reported. Standard errors are adjusted for arbitrary correlation within state-year cells and reported in parentheses. Extended covariates include: individual is working, individual is unemployed, real family income, number of days in the past month that physical health was good, and covered by health insurance. Asterisks denote statistical significance of coefficients as follows: ***p-value = 0.01, **0.01<p-value = 0.05, *0.05<p-value = 0.10.

Panel B	Unhealthy Food Consumption											
Outcome	Snacks		Hamburger		Hot Dogs		French Fries		Fried Chicken		Doughnuts	
Model	1	2	3	4	5	6	7	8	9	10	11	12
Depressed	11.1139***	14.6579***	3.5143	4.4522	2.1901	2.7331	10.0631***	10.0179***	0.8383	−1.8867	10.5293***	8.9737**
	(2.8784)	(3.5160)	(3.4916)	(3.2111)	(6.0840)	(7.2539)	(1.8727)	(2.1638)	(2.0000)	(1.8884)	(3.2989)	(3.5836)
Extended Covariates	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
R-squared	0.051	0.052	0.065	0.062	0.026	0.032	0.055	0.078	0.058	0.065	0.008	0.008
Observations	38878	6122	39169	6153	39093	6148	38975	6135	39057	6133	38941	6114

Note: See notes to Panel A above.

Panel C	Physical Activity	
Outcome	Regular / Vigorous Activity – Yes / No	
Model	1	2
Depressed	-0.0686***	-0.0190**
	(0.0070)	(0.0076)
Extended Covariates	No	Yes
R-squared	0.039	0.048
Observations	205836	178947

Note: Marginal effects from probit models are reported. See notes above.

Table 10. Effects of Depression on Food Consumption and Physical Activity among Females Data: BRFSS

Panel A	Healthy Food Consumption									
Outcome	Fruit		Fruit Juice		Carrots		Green Salad		Vegetables	
Model	1	2	3	4	5	6	7	8	9	10
Depressed	−25.1331***	−25.1531***	−10.0065***	−13.7200***	−5.1567***	−5.6103***	−14.9881***	−12.0212***	−20.7381***	−19.4144***
	(0.8712)	(0.8421)	(0.6294)	(0.6605)	(0.3685)	(0.3929)	(0.5009)	(0.4717)	(1.0062)	(1.0015)
Extended Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.058	0.061	0.042	0.043	0.022	0.022	0.040	0.047	0.055	0.054
Observations	1200345	894648	1200243	894122	1186800	886598	1201177	895431	1196205	892396

Note: Coefficients from OLS models are reported. Standard errors are adjusted for arbitrary correlation within state-year cells and reported in parentheses. Extended covariates include: individual is working, individual is unemployed, real family income, number of days in the past month that physical health was good, and covered by health insurance. Asterisks denote statistical significance of coefficients as follows: ***p-value = 0.01, **0.01<p-value = 0.05, *0.05<p-value = 0.10.

Panel B	Unhealthy Food Consumption											
Outcome	Snacks		Hamburger		Hot Dogs		French Fries		Fried Chicken		Doughnuts	
Model	1	2	3	4	5	6	7	8	9	10	11	12
Depressed	8.0648*	9.2276*	−0.0702	2.0424	4.4101*	3.0894	2.2634*	4.1046**	2.3393*	1.7749	8.4198*	11.1671**
	(4.5972)	(4.5696)	(3.1475)	(2.5322)	(2.4976)	(2.7177)	(1.2632)	(1.4883)	(1.2102)	(1.9432)	(4.5075)	(3.4023)
Extended Covariates	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.047	0.039	0.063	0.052	0.018	0.035	0.091	0.064	0.032	0.071	0.007	0.008
Observations	53374	8086	53627	8116	53438	8094	53429	8103	53541	8094	53429	8106

Note: See notes to Panel A above.

Panel C	Physical Activity	
Outcome	Regular / Vigorous Activity – Yes / No	
Model	1	2
Depressed	−0.1110***	−0.0519***
	(0.0059)	(0.0059)
Extended Covariates	No	Yes
R-squared	0.046	0.054
Observations	297999	245291

Note: Marginal effects from probit models are reported. See notes above.

12.6% improvement towards achieving the Healthy People 2020 healthy weight objective.*

Recent estimates place the economic burden of depression in the United States at $83.1 billion in the year 2000, or approximately $100 billion in current dollars.[43] Given the relative lifetime prevalence of major depression among females versus males in the NCS-R, about 61.8% of these costs are attributed to females.† In 2000-2003, based on the NCS-R, about 34.9% of women had suffered from major depression over their lifetime, which raises the probability of overweight/obesity by between 2.2 – 5.4 percentage points.

The total cost of adult obesity in the United States has been estimated to be as high as $215 billion.[2]. Cost estimates for overweight have not been established for the U.S. as a whole; however, a study for California[44] estimated the costs of overweight and obesity and found that overweight adds about a third more to the estimated economic burden of obesity. Given that the prevalence of overweight and obesity in California is reflective of the national prevalence, overweight imposes an additional $70 billion in economic burden for the U.S. Thus, the overall economic costs of obesity and overweight in the U.S. are as high as $285 billion annually. Of this, $134 billion can be attributed to female overweight and obesity, based on the relative prevalence across males and females. Since lifetime depression is estimated to raise overweight/obesity by 2.2 – 5.4 percentage points among females or about 4.2 – 10.2%, this higher risk of being overweight or obese could potentially add about $5.6 – $13.7 billion (or 6-14%) to the economic costs of depression.* These estimates suggest that public health interventions which reduce major depression among women could also further promote public health by reducing overweight and obesity.

Acknowledgments

We are grateful to Susan Averett and H. Elizabeth Peters for helpful comments and discussion.

* We assume that lifetime depression raises the probability of being overweight by between 6.24 – 15.6 percentage points, which is the range suggested by estimates for females reported in **Table 5**, **Table 6**, and **Table 8**.

† This allocation may be understated since depression among women is more likely to be chronic or more intensive, women are more likely to seek treatment and also more likely to attempt suicide, thus raising treatment or medical costs. This allocation may also be overstated since productivity losses due to depression among men may be relatively higher (due to greater labor force attachment and higher wages), fewer males seek treatment leading to potentially long-term adverse effects, and suicide mortality is higher among males relative to females.

* This calculation assumes that the marginal cost of overweight/obesity prevalence is approximately equal to the average cost.

References

1. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and trends in obesity among US adults, 1999-2008. JAMA 2010; **303**(3): 235-241.
2. Hammond RA, Levine R. The economic impact of obesity in the United States. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy* 2010; **3**: 285-295.
3. Wolf AM, Colditz GA. Current estimates of the economic cost of obesity in the United States. *Obes Res* 1998; **6**(2): 9–106.
4. Compton WM, Conway KP, Stinson FS, Grant BF. Changes in the prevalence of major depression and comorbid substance use disorders in the united states between 1991-1992 and 2001-2002. *Am J Psychiatry* 2006; **163**(1): 2141-2147.
5. Jonas BS, Mussolino ME. Symptoms of depression as a prospective risk factor for stroke. *Psychosom Med* 2000; **62**(4): 463.
6. Robbins J, Hirsch C, Whitmer R, Cauley J, Harris T. The association of bone mineral density and depression in an older population. *J Am Geriatr Soc* 2001; **49**(6): 732-736.
7. Saffer H, Dave D. Mental illness and the demand for alcohol, cocaine, and cigarettes. *Econ Inq* 2005; **43**(2): 229-246.
8. Greenberg PE, Kessler RC, Birnbaum HG, Leong SA, Lowe SW, Berglund PA, Corey-Lisle PK. The economic burden of depression in the United States: how did it change between 1990 and 2000? *J Clin Psychiatry* 2003; **64**(12): 1465–1475.
9. U.S. Department of Health and Human Services. *Mental Health: A Report of the Surgeon General—Executive Summary*. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health, 1999.
10. Solaroli A, Marchesini G, Baldassarri C, Boccaccini D, Biolcati R, Visani F, Migliorini S, Melchionda N. Eating behavior and health-related quality of life (HRQoL) in obesity. *Quality of Life Research* 1999; Vol. 8, No. 7, Abstracts: 6th Annual Conference of the International Society for Quality of Life Research: 633.
11. Mcelroy S, Frye M, Suppes T, Dhavale D, Keck P, Leverich G, Altshuler L, Denicoff K, Nolen W, Kupka R, Grunze H, Walden J, Post R. Correlates of overweight and obesity in 644 patients with bipolar disorder. *J Clin Psychiatry* 2002; **63**(3): 207-213.
12. Onyike C, Crum R, Lee H, Lyketsos C, Eaton W. Is obesity associated with major depression? Results from the Third National Health and Nutrition Examination Survey. *Am J Epidemiol* 2003; **158**(12): 1139-1147.
13. Ma J, Xiao L. Obesity and depression in US women: results from the 2005-2006 National Health and Nutritional Examination Survey. *Obesity* 2010; **18**(2): 347-353.
14. Carpenter K, Hasin D, Allison D, Faith M. Relationships Between Obesity and *DSM-IV* Major Depressive Disorder, Suicide Ideation, and Suicide Attempts: Results From a General Population Study. *Am J of Public Health* 2000; **90**(2): 251-257.
15. Simon G, Von Korff M, Saunders K, Miglioretti D, Crane P, van Belle G, Kessler R. Association between obesity and psychiatric disorders in the US adult population. *Arch Gen Psychiatry* 2006; **63**(7): 824-830.
16. Scott KM, Bruffaerts R, Simon GE, Alonso J, Angermeyer M, de Girolamo G, Demyttenaere K, Gasquet I, Haro JM, Karam E, Kessler RC, Levinson D, Medina Mora FE, Oakley Browne MA, Ormel J, Villa JP, Uda H, Von Korff M. Obesity and mental disorders in the general population: results from the World Mental Health Surveys. *Int J Obes* 2008; **32**(1): 192-200.
17. Fava M. Weight gain and antidepressants. *J Clin Psychiatry* 2006; **61**(11): 37-41.
18. Bulik CM. Anxiety, depression, and eating disorders. In Fairburn C.G. and Brownell K.D., eds. *Eating disorders and obesity: a comprehensive handbook*, New York, NY, The Guilford Press, 2002.
19. Cargill BR, Clark MM, Pera V, Niaura RS, Abrams DB. Binge eating, body image, depression, and self-efficacy in an obese clinical population. *Obes Res* 1999; **7**(4): 379-386.
20. Smith DE, Marcus MD, Lewis CE, Fitzgibbon M, Schreiner P. Prevalence of binge eating disorder, obesity, and depression in a biracial cohort of young adults. *Ann Behav Med* 1998; **20**(3): 227-232.
21. Frost RO, Goolkasian GA, Ely RJ, Blanchard FA. Depression, restraint,

and eating behavior. *Behav Res Ther* 1982; **20**(2): 113-121.
22. Gluck ME, Geliebter A, Satov T. Night eating syndrome is associated with depression, low self-esteem, reduced daytime hunger, and less weight loss in obese outpatients. *Obes Res*, 2001; **9**: 264-267.
23. Epel E, Lapidus R, McEwen B, Brownell K. Stress may add bite to appetite in women: a laboratory study of stress-induced cortisol and eating behavior. *Psychoneuroendocrinology* 2001; **26**: 37-49
24. Strohle A. Physical activity, exercise, depression, and anxiety disorders. *J Neural Transm* 2009; **116**(6): 777-784.
25. Strawbridge WJ, Deleger S, Roberts RE, Kaplan GA. Physical activity reduces the risk of subsequent depression for older adults. *Am J Epidemiol* 2002; **156**(4): 328-334.
26. Bauman AE, Sallis JF, Dzewaltowski DA, Owen, N. Toward a better understanding of the influences on physical activity: the role of determinants, correlates, causal variables, mediators, moderators, and confounders. *Am Journal Prev Med* 2002; **23**(2): 5-14.
27. Bair MJ, Robinson RL, Katon W, Kroenke K. Depression and pain comorbidity. *Arch Intern Med* 2003; **163**: 2433-2445.
28. Grossman M. On the concept of health capital and the demand for health. *J Polit Econ* 1972; **80**: 223-255.
29. Chou S, Grossman M, Saffer H. An economic analysis of adult obesity: results from the behavioral risk factor surveillance system. *J Health Econ* 2004; **23**(3): 565-587.
30. Altonji JG, Elder TE, Taber CR. Selection on observed and unobserved variables: assessing the effectiveness of catholic schools. *J Political Econ* 2005; **113**(1): 151-184.
31. Evans WN, Schwab RM. Finishing high school and starting college: do catholic schools make a difference? *Q J Econ* 1995; **110**: 941-974.
32. Goldman DP, Bhattacharya J, McCaffrey DF, Duan N, Leibowitz AA, Joyce GF, Morton SC. Effect of insurance on mortality in an HIV-positive population in care. *J Am Stat Assoc* 2001; **96**: 883-894.
33. Rashad I, Kaestner R. Teenage sex, drugs and alcohol use: problems identifying the cause of risky behaviors. *J Health Econ* 2004; **23**(3): 493-503.
34. Dave D, Kelly I. How does the business cycle affect eating habits. *Soc Sci Med* Available online 19 November 2011; Accessed at http://www.sciencedirect.com/science/article/pii/S0277953611006460
35. National Comorbidity Survey. National Comorbidity Survey (NCS) and National Comorbidity Survey Replication (NCS-R). 2010. Accessed at: http://www.hcp.med.harvard.edu/ncs/index.php
36. Bohannon RW, Maljanian R, Goethe J. Screening for depression in clinical practice: reliability and validity of a five-item subset of the ces-depression. *Percept Mot Skills* 2003; **97**(3): 855-861.
37. Pandya R, Metz L, Patten SB. Predictive value of the ces-d in detecting depression among candidates for disease-modifying multiple sclerosis treatment. *Psychosomatics* 2005; **46**(2): 131-134.
38. Glassman AH, Helzer JE, Covey LS, Cottler LB, Stetner F, Tipp JE, Johnson J. Smoking, smoking cessation, and major depression. *JAMA* 1990; **264**: 1546-1549.
39. Barrett-Connor E, von Muhlen DG, Kritz-Silverstein D. Bioavailable Testosterone and Depressed Mood in Older Men: The Rancho Bernardo Study. *J Clin Endocrinol Metab* 1999; **84**(2): 573-577.
40. Angst J, Gamma A, Gastpar M, Lepine JP, Mendlewicz J, Tylee A. Gender differences in depression: epidemiological findings from the European DEPRES I and II studies. *Eur Arch Psychiatry Clin Neurosci* 2002; **252**(5): 201-209.
41. Grossman M, Kaestner R. 1997. Effects of education on health. In *The social benefits of education*, edited by Jere R. Behrman and Nevzer Stacey. Ann Arbor, MI: University of Michigan Press, 69-123.
42. Thomas JL, Jones GN, Scarinci IC, Mehan DJ, Brantley PJ. The utility of the CES-D as a depression screening measure among low-income women attending primary care clinics. *Int J Psychiatry Med* 2001; **31**(1): 25-40.
43. Wang PS, Simon G, Kessler RC. The economic burden of depression and the cost-effectiveness of treatment. *Int J Methods Psychiatr Res* 2003; **12**(1): 22-23.
44. Chenoweth D. *The economic costs of physical inactivity, obesity, and overweight in California adults: health care, workers' compensation, and lost productivity*, California Department of Health Services, CA, April 2005.