NBER WORKING PAPER SERIES

SADNESS, SUICIDALITY AND GRADES

Jeffrey S. DeSimone

Working Paper 16239
http://www.nber.org/papers/w16239

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2010

I thank participants in a session at the 2010 ASHEcon conference, especially Pinka Chatterji, Dhaval Dave, Jason Fletcher and Cagatay Koc, for helpful comments. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Sadness, Suicidality and Grades
Jeffrey S. DeSimone
NBER Working Paper No. 16239
July 2010
JEL No. I10,I20

ABSTRACT

This study examines the past year relationship between GPA and experiencing a combination of two primary depression symptoms, feeling sad and losing interest in usual activities for at least two consecutive weeks, among high school students during 2001–2009. The GPA loss associated with sadness, as defined above, falls from slightly less than a plus/minus mark to around 0.1 point when commonly co-occurring behaviors are held constant. Nonetheless, this effect is significantly larger than those of having considered or planned suicide and equivalent to having attempted suicide, which seemingly signify more severe depression. Moreover, sadness lowers the probability of earning A grades, and raises that of receiving grades of C or below, by over 15%. Coefficient sizes are similar when comparison groups are restricted to students engaging in correlated behaviors and in matching and instrumental variable models, suggesting that sadness causally reduces academic performance.

Jeffrey S. DeSimone
Department of Economics
University of Texas at Arlington
701 S. West St.
Arlington, TX 76019
and NBER
jdesimone@uta.edu

1. Introduction

The 2008 National Survey on Drug Use and Health estimates that 12.9% of youth age 12–17 have experienced an episode of major depression in their lifetimes, and 8.3% have done so in the past year.[1] A major depressive episode (MDE) is defined by the DSM-IV as having at least five of nine potential symptoms during the same two-week period, including either depressed mood or markedly diminished interest in most activities, both for most of each day.[2] Adolescent depression has myriad potential adverse consequences, ranging from comparatively minor issues such as irritability and persistent aches and pains, to more serious problems such as anxiety, disruptive behavior, eating disorders, substance abuse, and the continuation of depression along with the onset of more severe illnesses in adulthood (Weissman et al., 1999).[3] Moreover, major depression is the most common risk factor for teen suicide attempts besides previous attempts (Lewinsohn et al., 1994), and characterizes most teens who commit suicide (Shaffer et al., 1996). In 2006, suicide was responsible for one in nine U.S. teenage deaths, making it the third leading cause of mortality after unintentional injury and homicide.

This paper studies one particular manifestation of depression recorded in national Youth Risk Behavior Survey (YRBS) data on high school students: responses to the question "During the past 12 months, did you ever feel so sad or hopeless almost every day for two weeks or more in a row that you stopped doing some usual activities?" This variable captures the two symptoms among which one must be present for a MDE diagnosis, along with the time frame.

---

[1] These were tabulated using the Inter-University Consortium for Political and Social Research online analysis system at www.icpsr.umich.edu/cgi-bin/SDA/SAMHDA/hsda?samhda+26701-0001.

[2] The other seven symptoms are (1) significant change in weight or appetite not related to dieting, (2) insomnia or excessive sleeping, (3) psychomotor agitation or retardation, (4) fatigue or energy loss, (5) feeling worthless or excessive guilt, (6) indecisiveness or diminished ability to concentrate, and (7) recurrent thoughts of death, suicide ideation or a suicide plan or attempt. Other than weight changes and a suicide plan or attempt, these must also occur nearly every day during the two-week period, represent a change in functioning, cause significant distress or functional impairment and not be attributable to substance use, a medical condition or bereavement.

[3] Information about depression, including specifically among adolescents, is available from the National Institute of Mental Health at www.nimh.nih.gov/health/topics/depression/.

Henceforth I refer to this as "feeling sad" to note the distinction with MDE, although it should be kept in mind that "sadness" incorporates the loss of interest symptom as well. Also, the analysis integrates data on two forms of suicidal thoughts as well as attempting suicide, any of which on its own would constitute another depressive symptom.

In the 2001, 2003 and 2009 YRBS data analyzed here, more than one-third of girls (36%) and one-fifth of boys (21%) report feeling sad. Overall, two out of seven respondents (29%) experienced sadness, equivalent to the rates in other YRBS years that contain information on sadness but not grades (1999, 2005 and 2007). By comparison with two other public health issues that have received substantial policy attention, 16% of respondents have ever smoked cigarettes daily for a month, and 31% describe themselves as overweight. Although the prevalence of sadness might therefore seem surprisingly high, it is consistent with information from Wilcox-Gok et al. (2004) that 8–10% of high school students are clinically depressed and another 15–20% have less severe depression. Moreover, at least three additional depression symptoms beyond sadness would be required for MDE diagnosis. In contrast, a MDE without sadness requires one of the two sadness conditions, in the absence of the other but in conjunction with at least four of the other seven depression symptoms.

The focus here is on how sadness relates to school performance, which can impact subsequent labor market outcomes through effects on immediate employment opportunities, high school graduation, and ensuing schooling quantity and quality. For each gender, YRBS students who report past year sadness have grade point averages (GPAs) about one mark lower on a plus/minus scale, e.g. B+ instead of A–. Among girls, the average past year four-point scale GPA was 2.82 for those who experienced sadness, but 3.14 for others, with analogous GPAs of 2.51 and 2.80 for boys.

Persistent sadness has the scope to lower academic performance. As Grossman (1973) argued for physical health, students with higher mental health status can more efficiently produce human capital through schooling. Thus, feeling sad could raise the marginal cost of school achievement. Similarly, feelings of hopelessness could increase the rate of future discounting, thus reducing the return to doing well in school. Sadness also likely reflects depression for some students, and several other depression symptoms, especially loss of energy and concentration capacity, would seemingly interfere with school performance. Even without further depression symptoms, sadness presumably takes time, effort and enthusiasm away from activities including schoolwork. Indeed, the "usual activities" that students discontinue might include homework and exam studying. Ding et al. (2009) conjectured that, in addition, teachers, parents and peers might treat visibly depressed students differently.

Of course, causation could simultaneously run in the reverse direction. Better students might more efficiently transform the same inputs into mental health (Grossman, 1973). Or, sadness might stem directly from poor school performance. For instance, nearly half of depressed students studied by Hysenbegasi et al. (2005) attributed their symptoms in part to low grades. Alternatively, students prone to sadness might tend to be those with lesser school achievement. Such a relationship is expected if inherited endowments or parental investments in human and mental health capital are positively correlated. Furthermore, many behaviors and characteristics that are apt to co-occur with sadness are also associated with lower grades.

This study quantifies the relationship between school performance and sadness in high school by estimating GPA regressions, with three themes emerging. First, the association between grades and feeling sad is not enormous. Even controlling for only the limited number of available exogenous determinants, sadness coefficients reflect just three-tenths of a GPA

standard deviation, and this falls to just over one-tenth when co-morbid behaviors and proxies for omitted factors are held constant. Many of these additional covariates have more pronounced relationships with grades than does feeling sad.

Second, the impact of sadness is nonetheless meaningful in several practical respects. Conceptually, one might envision that the GPA impact of a MDE, which involves at least one component of sadness along with additional depression symptoms, is at least as large. Empirically, even the smallest estimated effect of sadness is highly statistically significant, holding constant four measures of suicidality that constitute seemingly more severe depression symptoms: whether respondents have considered, planned, made, or been injured from a suicide attempt. Perhaps surprisingly, the negative feeling sad coefficient is larger than those of the suicidality variables, significantly so for all except attempting suicide. Feeling sad is also associated with a one-sixth decline in A grades and increase in grades of C or less.

Third, at least five pieces of evidence suggest that the ordinary least squares (OLS) sadness coefficient might signify a causal impact. One is that the GPA discrepancy between sad and non-sad students remains after controlling for not only suicidality, but also measures of substance use, anxiety, disruptive behaviors and low self-esteem, all of which often co-occur with adolescent depression and might influence school performance on their own. Another is that although I attempt to avoid interfering with a direct pathway from sadness to achievement by conditioning almost exclusively on covariates encompassing at least the past year period, it is impossible to eliminate the possibility that some variation in grades associated with these factors should instead be attributed to sadness. In addition, coefficients are reduced only slightly, if at all, when comparison groups are homogenized with respect to sadness by restricting samples to respondents who considered or planned (but did not attempt) suicide, or engaged in various other

strongly correlated behaviors. Furthermore, effects of sadness estimated in matching and instrumental variables (IV) models, which account for endogeneity even more systematically, are similar to those using OLS. Finally, the most comparable published study, the entire purpose of which is to construct an IV estimator to purge endogeneity from the relationship, reports an estimate considerably larger than mine.

2. Previous Literature

Grossman (1973) empirically established that less-healthy individuals obtain less schooling. Among draft-eligible white males in 1943, completed schooling increased as health during high school improved, albeit with an elasticity of only 0.03. However, respondents were unusually healthy and educated, in that all passed Air Force physical exams and achievement tests and graduated from high school, during which health was retrospectively rated as excellent by 87%. In a more representative group of men ages 19–29 in 1971, Perri (1984) found that health sufficiently poor to limit any activities was associated with reductions of 6% in current school enrollment and 0.7 years in education completed by those no longer enrolled.

As Grossman (1973) documents, these estimates might reflect more than just schooling reductions in response to health limitations. In particular, the more-educated might be able to produce better health, and unmeasured personal characteristics might simultaneously improve educational attainment and health status. Indeed, Grossman estimated that conditional on high school health status, an additional year of education was associated with health capital increases of up to 4% several decades after respondents had completed schooling. Furthermore, the addition of various correlates reduced the schooling coefficient by two-thirds. Specifically with regard to mental health, Chevalier & Feinstein (2007) showed in IV models that educational

attainment reduces the incidence of depression and malaise at ages 23, 33 and 42.

Economists have only recently begun to investigate the link between schooling outcomes and mental health in particular, focusing primarily on ADHD and depression.[4] ADHD appears to reduce academic performance. Using sibling fixed effects models for U.S. and Canadian children ages 4–11, Currie & Stabile (2006) found that ADHD symptoms have negative effects on test scores and schooling attainment 4–6 years later that are larger than those of physical health conditions. Also employing sibling fixed effects in a sample of U.S. children ages 5–12, Fletcher & Wolfe (2008) obtained similar positive effects of ADHD symptoms for grade repetition and special education by grades 7–12. They also showed deleterious impacts on high school suspensions and, without sibling fixed effects, GPA, expulsions and dropping out, along with reductions in college attendance and attainment. Instrumenting for ADHD using genetic code differences between siblings, Fletcher & Lehrer (2009) estimated very large but imprecise negative effects of ADHD on verbal test scores among 7^{th}–12^{th} graders.

Depression also has been estimated to reduce schooling and test scores, but not as consistently across genders and outcomes as ADHD. Wilcox-Gok (2004) found that the onset of depression at an age when school attendance is still compulsory increases high school dropout among men but not women. Currie & Stabile (2007) uncovered effects of depression on grade repetition but not test scores, enrollment or delinquency. Fletcher (2008) showed that adolescent depressive symptoms in 7^{th}–12^{th} grade are negatively associated with having completed high school and enrolled in a four-year college six years later, but only for females. Fletcher & Lehrer (2009) obtained results for depression similar to those described above for ADHD.

[4] Claessens et al. (2009) studied an overlapping, but differently disaggregated, set of socioemotional skills, estimating that neither externalizing nor internalizing problem behaviors, as of kindergarten entry or spring of 1^{st} or 3^{rd} grade, significantly predicted reading or math achievement test scores in 1^{st}, 3^{rd} or 5^{th} grade. A composite measure of interpersonal skills and self-control does significantly affect subsequent test scores in some specifications, but always in a counterintuitive direction.

Three studies have examined GPA responses, two among college students. Hysenbegasi et al. (2005) revealed that GPAs of Western Michigan University undergraduates were 0.49 points lower among those concurrently diagnosed with depression, but receipt of drug treatment eliminated 0.44 points of this difference. Eisenberg et al. (2009) reported that depression substantially increases dropout from college and graduate school over the next three years and has a significant negative cross-sectional association with GPA, although the two-year change in depression incidence affects the two-year GPA change only when co-occurring with anxiety.

Most comparable to my study is Ding et al. (2009), which analyzed the cross-sectional GPA impact of one-year lagged depression among 807 students observed in each of 10th, 11th and 12th grade, along with 86 additional students surveyed in at least one of those grades. They control for obesity, as do I, and ADHD in a time-invariant manner, and their sample period of 2001–2003 overlaps mine. In OLS models, they estimated a negative impact of both depression and ADHD, with the latter about 50% larger. Whereas ADHD coefficients are nearly identical by gender, the negative effect of depression is comparable to that of ADHD for boys, but only about one-fourth the size and less significant for girls. In IV models identified by genetic markers, the deleterious impact of depression is 3½ times larger than that using OLS. However, the estimates are large but insignificant for each gender, while the effect of ADHD is insignificant and much smaller for girls but positive for boys and overall.

My study is the only one besides Ding et al. (2009) to examine the impact of depression symptoms on GPA among high school students. While Ding et al. (2009) constructed a clinical depression indicator based on reported symptoms, I study an amalgam of a symptom pair that occurs more widely than do MDEs. Our approaches differ in that they rely primarily on IV supplemented by parental characteristic controls, whereas I focus on holding constant a large set

of co-morbidities and use IV as a robustness check. Also, the YRBS is a substantially larger and more nationally representative data set than their sample of five high schools from the same northern Virginia county. I discuss the Ding et al. (2009) results further after presenting mine.

3. Data

I analyze data from the 2001, 2003 and 2009 editions of the national YRBS (Brener et al., 2004), a national school-based survey administered from February through May of every odd-numbered year since 1991 (www.cdc.gov/HealthyYouth/yrbs/), because these are the only three years in which academic performance information was collected. The sampling scheme is designed to yield a nationally representative group of students in grades 9–12. Schools, 150 in 2001 and 158 in 2003 and 2009, were selected from primary sampling units (PSUs, comprising sub-areas of very large counties, single large counties or groups of small, adjacent counties), with probability proportional to enrollment. From each school, one or two classes of a required subject were chosen randomly from every grade level.[5]

Table 1 lists the analysis variables. Students are asked about their grades in school during the past 12 months, with five choices corresponding to "mostly" getting a letter grade of A, B, C, D or F. I construct a four-point GPA variable by coding A = 4, B = 3, C = 2, D = 1 and F = 0. This corresponds to category midpoints if A and F are treated like the others by imagining students could receive grades of A+ or F–, thus making the (hypothetical) bounds equal to –0.5 and 4.5 rather than the observed values of 0 and 4. This partially accounts for censoring, which later is explicitly addressed using an interval regression model which recognizes that only the interior grade category endpoints are observed.

[5] All students in these classes were eligible to participate, with student response rates of 88% in 2009 and 83% in the earlier waves. Local parental permission procedures were followed, and students completed self-administered questionnaires in their classrooms during a regular class period.

As described already, the sadness variable is an indicator coded to one for students who felt sufficiently sad or hopeless, almost daily for at least two consecutive weeks sometime during the past year, that they stopped doing usual activities. The analysis also integrates information on four indicators of past year suicidal behavior, each constructed from the response to a separate YRBS question. These are having seriously considered, made a plan about how to attempt, actually attempted, and had to be treated by a doctor or nurse as the result of attempting suicide.

The YRBS records a small set of variables that I label as exogenous: indicators for gender, age, grade level and racial/ethnic group, along with height and weight each interacted with gender.[6] In addition, although school identifiers are not reported, regressions control for PSU-by-year fixed effects. Before imposing the sample selection criteria outlined below, 53 PSUs contribute data in 2001 and 55 do so in 2003 and 2009, implying that each PSU includes an average of fewer than three surveyed schools.[7] Also, all schools in the same PSU are from the same survey stratum, meaning they have similar black and Hispanic enrollment and MSA status.

Remaining variables are incorporated because they reflect conditions commonly co-occurring with depression. All are represented in the analysis using sets of binary indicators.

Substance use often occurs alongside depression, because depressed individuals use alcohol or drugs as a way to escape their condition, substance use or withdrawal leads to depression, or underlying factors cause both. Hallfors et al. (2004) found that among 7^{th}–12^{th} graders, smoking, drinking and illegal drug use were all associated with depression as well as suicide ideation and attempts. Included substance use covariates measure previous lifetime use

[6] Weight is included to complement height and provide a specific interpretation for the self-described weight variable that is also utilized. The "exogenous" label for this covariate set serves primarily an organizational purpose, and results are identical regardless of whether weight is held constant.

[7] In 2001 and 2003, whether the school is urban, suburban or rural is also observed. Many PSUs encompass at least two urbanization categories, ultimately resulting in 175 PSU-by-urbanization-by year combinations containing an average of 1.8 schools. Results omitting 2009 data are virtually identical when "school group" fixed effects are alternatively specified using this additional information.

of cigarettes, alcohol, glue or paints/sprays to get high, steroids without a prescription, marijuana, cocaine, heroin and methamphetamines. Separate indicators are specified for having smoked cigarettes at all and daily for at least a month, as well as for seven different intervals regarding the number of days alcohol was consumed and times marijuana was used.

Anxiety disorder also commonly accompanies depression. Although the YRBS does not ask directly about anxiety, it does report information on several other behaviors that are likely to create angst. One such behavior is sexual activity, inasmuch as it reflects romantic relationships, presumably a leading cause of anxiety among high school students. Two-thirds of depressed students in Hysenbegasi et al. (2005) reported relationship problems as a cause, while Hallfors et al. (2004) uncovered an association between sexual activity and depression. Consequently, models include an indicator for each potential number of lifetime sex partners up to six or more.

Two other related regressors are having been hit, slapped or physically hurt on purpose by a boyfriend or girlfriend in the past year, and having ever been physically forced to have sex. Posttraumatic stress disorder (PTSD), a specific anxiety disorder that can result after physical or sexual assault victimization, is strongly linked with depression. Kilpatrick et al. (2003) estimated that being physically or sexually assaulted increases MDE risk among 12–17 year olds.

The YRBS likewise contains information related to disruptive behaviors, which also tend to co-occur with depression. A direct measure is involvement in a physical fight, expressed as indicators for once and at least twice. Other experiences potentially reflecting disruptive behavior, at least by peers if not the respondent (and presumably anxiety-producing as well), include having been absent from school because of feeling unsafe there or en route, threatened or injured with a weapon on school property, and offered, sold or given an illegal drug on school property. All encompass the past year except for feeling unsafe, which refers to only the past

month.[8] Any of these variables could also signal having witnessed violence, another potential PTSD trigger which Kilpatrick et al. (2003) also found to increase MDE risk.

Kostanski & Gullone (1998) showed that body image dissatisfaction is related to both depression and low self-esteem among 12–18 year olds. In the context of already controlling for bodyweight, indicators for describing oneself as underweight or overweight, relative to being "about the right weight," are intended to capture body image displeasure and thus poor self-esteem. Finally, indicators of whether respondents played on one, two or at least three sports teams in the past year also serve as proxies for self-esteem, confidence, or perhaps other unobserved factors inversely correlated with depression, as Sanders et al. (2000) found that high school seniors with greater sports participation had lower depression levels.

The original data files include 45,225 students: 13,601 from 2001, 15,214 from 2003 and 16,410 from 2009. I drop 4,180 respondents for which GPA or sadness is not observed, 3,433 with unreported values for at least one core personal characteristic (including 88 students who are only 12 or 13 years old), 4,271 missing information on one of the four suicidality variables,[9] 1,282 for whom cigarette use is unobserved, 2,405 with no information on use of other substances, and 2,228 missing values on other variables, for an analysis sample size of 27,426 observations representing 155 of the 163 originally surveyed PSUs.[10]

Table 1 also provides variable means. Column 1 contains means for the full sample. The average GPA is one-tenth of a point below a B. Two in seven respondents felt sad or hopeless.[11]

[8] Although I have otherwise avoided conditioning on behaviors reflecting time frames more recent than the past year period covered by GPA and sadness, I include feeling unsafe as a potentially important source of sadness.
[9] The injury variable, the least important suicidality measure, is responsible for only 196 of these exclusions.
[10] Each of the other 8 PSUs is excluded because at least one (but typically several) of the questions yielding information that I use is not included on the surveys in the PSU.
[11] Among the 13,619 respondents who are excluded despite reporting both GPA and sadness information, sadness is more frequent than in the analysis sample, but by only 1.1 percentage points. However, mean GPA is lower by 0.18 points, i.e. 20% of a standard deviation. Accordingly, the difference in mean GPA by sadness status is 0.054 points (20%) less than that reported in column 1 of table 2.

One-sixth of students considered suicide, one-eighth made a plan to attempt it, more than 7% actually did attempt it and just over one-quarter of those students were injured sufficiently to necessitate medical attention.

Columns 2 and 3 of table 1 display means for separate samples stratified by sadness status. Respondents who did not feel sad have GPAs that are higher by 0.27 points. Students who felt sad are disproportionately female, Hispanic rather than white non-Hispanic, and likely to experience all the listed risky behaviors and outcomes. Note also that the more serious suicidality categories do not automatically imply the less serious or even sadness, as 21% of those who attempted suicide, 29% who considered it and 31% who planned it did not feel sad.[12]

4. Results

Most of the analysis relies on OLS GPA regressions, with standard errors adjusted for heteroskedasticity.[13] Table 2 reports sadness coefficients for a preliminary set of models. Column 1 reiterates that the mean difference estimate is 0.27 and confirms its high significance. The coefficient increases by 10% controlling for exogenous personal characteristics in column 2, and decreases only slightly when PSU-by-year fixed effects are inserted in column 3. Even these maximal estimates are not overwhelmingly large, in the sense that they represent about 30% of a GPA standard deviation or slightly less than one plus/minus mark.

The first behavioral variables added to the regressions, in column 4, are the suicidality measures, in order to ensure that the relationship between GPA and sadness is not in fact attributable to an extreme form of depression accompanied by thoughts of suicide. As expected,

[12] Similarly, 22% of students who planned suicide did not seriously consider it, and reciprocally 39% who considered suicide did not make a plan for attempting it. The interaction between the two is thus identified, but was highly insignificant in the regressions and thus omitted from the analysis.

[13] For the OLS models, *t*-statistics calculated using standard errors clustered at the PSU-by-year level are essentially identical to the heteroskedasticity-robust *t*-statistics appearing in the tables.

the sadness coefficient falls in magnitude, but by only about 20%. Moreover, each suicidality coefficient is smaller than that of feeling sad, all substantially so except for attempting suicide. Effects of planning and being insured attempting suicide are statistically insignificant, and that of considering suicide is only one-third the size of the sadness effect.

Controlling for co-morbidities

Table 3 displays estimates when controls for various behaviors and conditions that tend to co-occur with depression are sequentially inserted into the regressions. I begin by holding constant cigarette use, which seems unlikely to directly affect GPA and has been argued to signal a high rate of future discounting (e.g. Farrell & Fuchs, 1982). The sadness coefficient falls by 30%, a larger decline than from adding the suicidality measures, with concordant or greater reductions in the suicidality coefficients. This suggests that unobserved heterogeneity is present.

However, further conditioning on the long list of remaining confounders in columns 2–5 decreases the sadness effect by only about the same fraction as the two cigarette indicators by themselves. Sexual activity, including sexual or physical assault victimization, interferes little with the relationship between GPA and sadness in column 4, whereas the four remaining violence-related disruptive behaviors appear important in column 3. The minimal sadness coefficient reductions from adding the extensive set of alcohol and drug use controls in column 2, and the self-described weight and sports participation variables in column 5, encouragingly suggests limited additional prospective omitted variable bias, especially since table 4 shows that many of these covariates have large and significant associations with GPA. Notwithstanding, there is no guarantee that all relevant unobservables have been exhausted, especially since three-quarters of the variation in GPA remains unexplained.

Column 5 represents the benchmark specification to which subsequent robustness checks are compared. Even conditional on the myriad included personal characteristics, behaviors and experiences along with location-by-year, sadness lowers GPA by a very highly significant 0.11 points, about 12% of the GPA standard deviation. In comparison, the coefficients of both considering and planning suicide are small and insignificantly different from zero, but significantly smaller at the 1% level than that of feeling sad. The decrease in GPA associated with attempting suicide is also smaller quantitatively, albeit equivalent statistically (*p*–value = 0.58).[14] Unexpectedly, grades of those seriously injured attempting suicide are 0.1 point higher than those emerging uninjured from an attempt, i.e. no lower than those of non-attempters.

Other covariates

Table 4 gives results for other coefficients. GPA declines with age but rises with grade level, is substantially higher for females but lower for blacks and Hispanics, increases with height but decreases with weight, is lower for cigarettes smokers, moderate but not heavy drinkers, and users of glue and steroids, falls with increasing use of marijuana, is negatively related to violence-related disruptive behaviors, number of sex partners and describing oneself as underweight, and rises with number of sports teams.

Lifetime smoking, having used marijuana at least three times, missing school because of feeling unsafe, involvement in multiple physical fights, having at least two previous sex partners and playing sports all have larger associations with school performance than does feeling sad. However, these and the other behavioral covariates are included specifically to reduce endogeneity in the relationship between GPA and sadness, and their coefficients should not necessarily be interpreted as causal effects. For instance, it seems doubtful that one use of

[14] When added to this model, the interaction of sadness with attempting suicide is insignificant (*p*-value = 0.15).

marijuana would lower grades as much as feeling sad or that each additional sport played would do the opposite. Still, GPA differences by gender and race, and between the age and grade level extremes, are 2–5 times larger than that between sad and non-sad students.

Samples stratified by suicidality

Table 5 provides estimates from samples stratified by suicidality experiences, in an effort to make the control group of students who did not feel sad more comparable to those who did. Column 1 excludes the relatively small minority of respondents who considered, planned or attempted suicide, meaning that no other observed depression-related differences exist between sad students and their non-sad counterparts. The coefficient on feeling sad is almost the same as that from the full sample in column 5 of table 3.

In column 2, the strategy is nearly the opposite, in that only students who considered or planned suicide, but did not attempt it, are included. Thus, students in the non-sad comparison group are similar to sad students in also experiencing suicidal thoughts but not attempting suicide. Yet, the impact of feeling sad has declined by only 20% and, despite the drastically reduced sample size, remains significant at 1%. The invariance of the sadness coefficient to these two different approaches for homogenizing the comparison groups based on suicidal thoughts and actions lends support to the notion that feeling sad directly reduces school performance. Predictably, when the columns 1 and 2 samples are combined in column 3, the effect of sadness remains nearly identical to that from the full sample. Considering and planning suicide remain highly insignificant both individually and jointly.

The column 4 results pertain to a sample consisting only of those who attempted suicide. Although the wide confidence interval encompasses the full-sample estimate, the much smaller

and insignificant sadness coefficient suggests this group is different from non-attempters with regard to the question of interest. This is not surprising, given that among students attempting suicide, 79% felt sad and mean GPA is 2.57, compared to respective values of 24% and 2.94 among non-attempters.[15] Because retaining this group in the sample has no impact on the sadness coefficient, though, I do so in the subsequent analyses.

Using alternative comparison groups

Table 6 continues the theme of table 5 by slicing the sample in various ways to increase the similarity between sad and non-sad students. The specific approach parallels column 2 of table 5, in that each row of table 6 reflects a subsample that excludes all students who are not characterized by a specific behavior or condition that is highly correlated with sadness. As columns 2 and 3 indicate, all the table 6 subsamples have lower average grades and higher rates of sadness, often substantially, than overall. Nonetheless, the effect of feeling sad is virtually identical to the benchmark in four of the nine panels, including among the relatively small set of students who feel unsafe or have been threatened or physically or sexually assaulted, which has by far the highest sadness incidence and close to the worst school performance. In the other five restricted samples, the sadness coefficient is smaller, but never by more than 30%, while maintaining high significance despite a considerably smaller number of respondents in each case. Thus, feeling sad deters academic performance even compared to non-sad students who have lower baseline GPAs in association with experiences that are common among sad students.

[15] This GPA discrepancy likely contributes to the odd result that considering and planning suicide are jointly significant at 5%, with those who do both predicted to have a 0.22 point higher GPA. Among attempters, 73% both considered and planned suicide, and only 6% failed to at least consider it. Those attempting suicide without initial consideration or planning might be more impulsive or have lesser cognitive skills.

Ordered logit model

Although converting GPA to a four-point scale is convenient for estimating OLS regressions with easily interpretable parameter estimates, the unaltered reported categorical information on grades can be easily examined using an ordered discrete choice model. The coefficient on sadness in one such model, an ordered logit, is –.253 (t = 8.57).

Table 7 provides corresponding marginal effects of feeling sad on the incidence of each grade category. Column 1 shows that three in seven respondents attain mostly B's while two in seven achieve A's, five in sex of remaining students receive C's and just under a quarter of the rest report F's. In column 2, all marginal effects are highly significant, even for the B category where outflows to C or below and inflows from A are virtually balanced. As column 3 reveals, associated semi-elasticities are 16–17% for all grade levels besides B. This implies that on net, feeling sad effectively shifts one-sixth of the grade distribution from A to C or below.

To compare grade point implications with OLS, column 4 gives the predicted GPA change implied by the product of the marginal effect and associated points for each grade level. The sum across categories is –.107, identical to the sadness coefficient in the benchmark OLS model. It thus appears that adapting the GPA measure to use with OLS does not alter the results.

Other robustness checks

Results for three additional models are omitted from the tables for brevity. Incorporating the YRBS sampling weights, the OLS sadness coefficient falls slightly to –.099 (t = 5.97), meaning that unweighted OLS estimate is reasonably representative of the relationship in the national high school population. Interval regression uses only information on the interior grade category thresholds, i.e. 1, 2 and 3 points, assuming censoring below and above and normally

distributed grades. The sadness coefficient rises to –.125 (t = 8.87), suggesting that if anything, relying on grade category midpoints for use with OLS biases the effect of feeling sad towards zero, even when the hypothetical boundaries are extended to –0.5 and 4.5.

I also estimate nearest neighbor propensity score matching models (Rosenbaum & Rubin, 1982; Dehejia & Wahba, 2002). A probit regression of feeling sad on all observables is estimated, with the predicted probability of sadness termed the propensity score. Each sad student is then matched to the non-sad student with the closest propensity score, discarding sad respondents who have either no match within the narrow score range of 0.0001 or scores outside the non-sad extremes. The average GPA difference between matched pairs of sad and non-sad students is –.103 (t = 5.19), with about two in seven sad students (2,186 of 7,705) unmatched and thus not used to construct the estimator. Compared to the full sample mean difference estimator (i.e. column 1 of table 2), matching narrows the effect of sadness primarily by constructing a control group with an average GPA that is lower (by 0.12 points) than that of all students who did not feel sad, whereas the GPA of the matched sad group is about 0.05 points higher excluding the unmatched students. The groups are well-balanced in the sense that the highest standardized bias among the 210 covariates is 4.7, i.e. well below the commonly-mentioned threshold of 10.[16] Also, the pseudo R-squared of the regression of sadness on all observables falls from .207 in the unmatched sample to .011 in the matched sample, with an associated p-value of .990 in the latter, meaning the set of observables is unrelated with whether matched students are sad or not.[17] If unobservables are distributed similarly (e.g. Altonji et al. 2005), the

[16] The standardized bias is the difference in means between sad and non-sad students, as a percentage of the square root of the average variances in each group.

[17] Using a tenfold wider caliper of 0.001, the set of observables is significant at 1% in the regression of feeling sad in the matched sample. With an intermediate caliper of 0.0005, the estimated effect of sad is –.086 (t = 4.19), only 7.6% of sad students are unmatched, and the p-value for joint observable significance in the matched sample regression is 0.271.

matching estimator, and thus the benchmark OLS model with a nearly identical sadness coefficient, reflects a causal effect of feeling sad on GPA.

Instrumental variables models

Although the propensity score estimate is encouraging, matching models do not explicitly control for unobserved heterogeneity. I thus proceed to estimate IV models. Eisenberg et al. (2009, pg. 2) note that "it is difficult to imagine variation in mental health problems that is clearly exogenous with respect to academic outcomes." My instrument choices are therefore empirically motivated. In particular, I select variables that are significant in the OLS counterpart to the sadness propensity score regression, but have absolute *t*-statistics less than one in the table 4 regression of GPA on sadness.

The primary instruments are considering and planning suicide, which table 3 showed to be insignificantly related to GPA. Clearly, neither is generated by a process exogenous to the link between sadness and school performance, and it seems unlikely that any causal mechanism would operate in the direction from suicidal thoughts to sadness. A potential intuitive justification is that feeling sad is the manifestation of depression that impacts grades, with suicidal thoughts constituting another depression symptom that is highly correlated with sadness, but evidently not otherwise associated with GPA. This could be because sadness encompasses at least two consecutive weeks and interrupted participation in usual activities, thus blunting the potential for suicide attempt consideration and planning to interact with grades via signaling more severe depression or occupying time that might otherwise have been devoted to academics. Eisenberg et al. (2009) analogously estimated a small and highly insignificant impact of suicidal thoughts on GPA, holding constant other depression symptoms, while likewise finding that loss

of interest is the symptom with the largest GPA effect.

Beyond the suicidal thoughts measures, several other variables satisfy the empirical criteria for use as instruments: having used heroin, used methamphetamines, been physically abused by a boyfriend or girlfriend, and forced to have sex, along with self-describing as overweight. The first stage sadness regression coefficients for the instruments are listed in the footnote to table 8. Considering suicide is by far the strongest instrument, with a *t*-statistic above 27, and planning suicide also has a *t*-statistic of 10. Their coefficients imply that students who considered and planned suicide are 31 and 12 percentage points, respectively, more likely to experience sadness. Among the other five instruments, the associated difference in sadness is at least 3.7 percentage points (considering oneself overweight) and *t*-statistics are no smaller than 3.7 (heroin use), with the latter implying that overidentification tests have reasonable power to distinguish differences in IV estimates across instrument sets.

Table 8 shows the IV regression results, with models estimated by GMM. Each column pertains to a distinct identification strategy using a different set of variable exclusions. Column 1 is the baseline model using just suicide consideration and planning as instruments. In the top row, the IV estimate is slightly larger in magnitude than the benchmark OLS estimate and is significant at 5%. The next two rows report that the procedure passes two diagnostic tests for instrument exogeneity, as *p*-values for both the overidentification *J*-statistic and the joint *F*-statistic for the two IVs in the GPA regression of table 4 are highly insignificant. The fourth and fifth rows indicate that, consistent with the individual instrument *t*-statistics, their joint significance in the first stage regression is extremely large. The *F*-statistic implies essentially no finite sample bias relative to OLS, and the partial *R*-squared value means that considering and planning suicide explain nearly 7% of the variance in sadness even after holding constant the

wide array of remaining covariates. Notably, standard errors are four times as large with IV than OLS, even using instruments that are markedly stronger than typically found in IV studies.

Remaining columns add to the basic instrument set the two drug use variables (column 2), the physical and sexual assault variables (column 3), the overweight variable (column 4) and finally all five supplementary instruments (column 5). Results are similar in all respects. In column 5, the model using all seven instruments produces the most conservative sadness coefficient, but even this is well within 10% of the benchmark OLS estimate. Exogeneity test *p*-values are above 0.9 and the partial *R*-squared of the instruments is over 8%.

In sum, the IV estimates provide further evidence that the OLS sadness effect might be interpretable as causal. Moreover, unlike many IV estimators in the broader empirical microeconomics literature, those obtained here not paradoxically much larger than OLS.

Samples stratified on exogenous covariates & smoking

Table 9 shows that the effect of feeling sad is similar across groups stratified by exogenous demographic characteristics. Sadness lowers GPA slightly more for females than males and over time, but nearly identically for white non-Hispanics and others, the latter even though Hispanics and non-whites have considerably lower grades and higher sadness incidence. The impact of feeling sad first rises and then falls with increasing age and grade in school, more sharply for the latter. This could in part be an artifact of selection, if students for whom school performance is most adversely affected by depression eventually are either held back in school or drop out sometime after reaching age 16.

Discussion

One important argument in favor of interpreting the benchmark OLS estimate as a direct effect of sadness on GPA is that it is smaller in magnitude than those from two earlier-cited studies of how depression affects GPA. Interestingly, my estimate is similar to that obtained by Ding et al. (2009) using OLS, although they found a much larger effect among males than females. However, their IV estimate is over six times greater than mine in terms of GPA standard deviations. In principle, this could simply reflect a large local average treatment effect among students for whom differences in depression incidence are produced by genetic code variation. Otherwise, the unexpected implication for the OLS estimate is that bias towards zero, primarily from measurement error, dominates upward bias (in absolute value) from reverse causation and most forms of unobserved heterogeneity. If anything, their IV estimate implies that mine is conservative, even though my main identification concern would seem to be the inability to verify having controlled for all relevant unobservables.

Although we study different populations, estimates from the Eisenberg et al. (2009) cross-sectional model that is most similar to mine are also somewhat larger. The GPA decline associated with an increase in their depression score corresponding with moving from low to severe symptoms is over twice greater than mine in standard deviation terms, and grows to about six times greater when co-occurring with anxiety, despite controlling for prior GPA and admission test scores along with other mental disorders.[18]

An important difference between these other studies and mine is that I look at sadness, not depression. Because many sad students are not depressed whereas most depressed students are sad, as mentioned previously, my analysis might be expected to yield a smaller estimate. However, even adding the coefficient of attempting suicide to that of sadness (and ignoring the

[18] A smaller change in score, roughly equal to the difference in mean scores of those below and at least the commonly used cutoff for a positive screen, for students without anxiety yields an impact that is about 20% of a standard deviation, which is still larger but more comparable to my estimate.

positive interaction term coefficient when added to the benchmark OLS model) produces a smaller GPA impact than in these other two studies, even though sad students who attempted suicide are less frequent (5.8%) than depressed adolescents at large and are likely depressed themselves given that they report all three measured symptoms. Moreover, in another cross-sectional specification which included each of the nine depression symptoms as a separate regressor, Eisenberg et al. (2009) found that the only symptom entering significantly is the loss of interest in usual activities, one of the two symptoms embodied by feeling sad. A change from experiencing loss of interest "not at all" to "nearly every day" over the previous two weeks reduces GPA by over twice as much as sadness in my study.[19]

The other argument in support of a causal interpretation is that, if anything, my estimates seem far more likely to be biased towards than away from zero. Most of the analysis was geared towards showing robustness of the OLS sadness coefficient to unobserved heterogeneity. The remaining concern is if low grades directly raise sadness propensities. However, the dependent variable reflects performance in a variety of subjects over the course of a full year and thus is not prone to being affected by temporary grade shocks. Reverse causation could manifest itself only if, say, students who consistently receive Bs (or Cs) are depressed over not being A (or B) students. However, it is hard to imagine chronic below-expected achievement among students who care about school performance enough for it to have a prolonged impact on their moods. A related possibility is that GPA reflects innate cognitive ability that protects against feeling sad. But this type of feedback effect, which is more literally an example of unobserved heterogeneity,

[19] The estimate from a third study, Hysenbegasi et al. (2005), is five times larger than mine, although the coefficients cannot be strictly compared because their GPA standard errors are not reported. Two reasons for this size discrepancy could be that all those categorized as depressed in their sample were diagnosed at their campus health center, and the effect was among those who did not fill any prescribed medications whereas there was essentially no GPA impact among the diagnosed who filled a prescription. A potentially more comparable strategy, in which pairs of diagnosed and undiagnosed students were matched, yielded a difference-in-difference estimator of –0.28 points.

is controlled for in part by holding constant height-by-gender (e.g. Case & Paxson, 2008), and is further conditional on a slew of co-morbid conditions against which GPA should also protect.

In contrast, for several reasons my estimate might instead be interpreted as a lower bound for the true GPA effect of sadness. One is measurement error that is random, or instead involves a tendency for either lower-performing students to falsely report not feeling sad or sad students to overestimate their grades. A second is if some variables in the control set, such as substance use or suicidality, lie along the causal pathway from sadness to grades, in which case some of the GPA reductions that in fact stem from sadness are instead attributed to co-morbidities. Third, Hysenbegasi et al. (2005) observed that the academic performance of depressed undergraduates reached a nadir at diagnosis, and had significantly improved by the 4th month after diagnosis. If this pattern also prevailed among YRBS students, onset of sadness recently before the interview would not allow for sufficient time to detect the full deleterious GPA impact, especially if reflective of an undiagnosed MDE, whereas effects of sadness that was treated early in the reporting period might have dissipated well before the interview.

5. Conclusion

This study has documented a negative relationship between prolonged activity-limiting sadness and academic performance among high school students that is small, yet highly significant and quantitatively meaningful. A confluence of evidence, including OLS models controlling for an abundance of co-morbidities and using comparison groups of students engaging in behaviors highly correlated with sadness along with matching and IV models, is consistent with this relationship signifying a causal impact of sadness on grades. On the other hand, Eisenberg et al. (2009, pg. 2) argued that "a careful descriptive analysis is the only feasible

approach to learn about this relationship" without randomized trials. By this interpretation, the analysis has, at the very least, identified a highly idiosyncratic form of endogeneity that must be present in order to deny a causal interpretation.

The results have at least two implications for mental health policy. First, improved school performance is a possible additional benefit from successful treatment of the depressive symptoms reflected by sadness. Alleviating sadness can thus convey long-term economic gains, given the multiple links between GPA and eventual labor market outcomes. Second, in terms of grades, addressing "less severe" depressive symptoms embodied by sadness is at least as important as confronting suicidality, even though the latter seems more extreme and tangible.

A limitation is that the analysis cannot identify adverse shocks that create transitory sadness. For instance, the death of a close relative or parental divorce is expected to temporarily, but substantially, disrupt the life of a student. The unavoidability of such events, and the ensuing sadness and impaired academic performance, suggests minimal role for mental health policy. Interventions to alleviating sadness would seem more appropriate, in terms of effects on grades, when sadness is chronic or not caused by an obvious external circumstance. This still implies a role for policy when sadness caused by a negative shock lasts longer than is typical or "healthy."

As highlighted earlier, research on how depressive symptoms affect school performance among college students is even scarcer than that on pre-college age adolescents, but a similar approach to that utilized here could be pursued in data such as the Cooperative Institutional Research Program and National College Health Assessment surveys. Possibly more challenging with available data would be investigating whether cognitive ability reduces the academic performance diminutions of poor mental health among adolescents, information on which could improve the ability to target mental health treatments more efficiently.

References

Altonji, Joseph G., Todd E. Elder & Christopher R. Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," Journal of Political Economy, February 2005, 113(1), 151–184.

Brener, Nancy D., Laura Kann, Steven A. Kinchen, Jo Anne Grunbaum, Laura Whalen, Danice Eaton, Joseph Hawkins & James G. Ross, "Methodology of the Youth Risk Behavior Surveillance System," Morbidity and Mortality Weekly Report, 24 September 2004, 53(RR-12), 1–13.

Case, Anne & Christina Paxson, "Stature and Status: Height, Ability, and Labor Market Outcomes," Journal of Political Economy, June 2008, 116(3), 499–532.

Chevalier, Arnaud & Leon Feinstein, "Sheepskin or Prozac: The Causal Effect of Education on Mental Health," Geary Institute, University College Dublin, Working Paper 2007-15.

Claessens, Amy, Greg Duncan & Mimi Engel, "Kindergarten Skills and Fifth-Grade Achievement: Evidence from the ECLS-K," Economics of Education Review, August 2009, 28(4), 415–427.

Currie, Janet & Mark Stabile, "Child Mental Health and Human Capital Accumulation: The Case of ADHD," Journal of Health Economics, November 2006, 25(6), 1094–1118.

Currie, Janet & Mark Stabile, "Mental Health in Childhood and Human Capital," National Bureau of Economic Research, Working Paper 13217, July 2007.

Dehejia, Rajeev H. & Sadek Wahba, "Propensity Score Matching Methods for Nonexperimental Causal Studies," Review of Economics and Statistics, February 2002, 84(1), 151–161.

Ding, Weili, Steven F. Lehrer, J. Niels Rosenquist & Janet Audrain-McGovern, "The Impact of Poor Health on Academic Performance: New Evidence Using Genetic Markers," Journal of Health Economics, May 2009, 28(3), 578–597.

Eisenberg, Daniel, Ezra Golberstein & Justin B. Hunt, "Mental Health and Academic Success in College," B.E. Journal of Economic Analysis and Policy: Contributions to Economic Analysis and Policy, 2009, 9(1).

Farrell, Phillip & Victor R. Fuchs, "Schooling and Health: The Cigarette Connection," Journal of Health Economics, December 1982, 1(3), 217–230.

Fletcher, Jason M., "Adolescent Depression: Diagnosis, Treatment, and Educational Attainment," Health Economics, November 2008, 17(11), 1215–1235.

Fletcher, Jason M. & Steven F. Lehrer, "Using Genetic Lotteries within Families to Examine the Causal Impact of Poor Health on Academic Achievement," National Bureau of Economic Research, Working Paper 15148, July 2009.

Fletcher, Jason & Barbara Wolfe, "Child Mental Health and Human Capital Accumulation: The Case of ADHD Revisited: Comment," Journal of Health Economics, May 2008, 27(3), 794–800.

Grossman, Michael, "The Correlation between Health and Schooling," 147–211, in Terleckyj, Nestor E., ed., Household Production and Consumption, New York: Columbia University Press, 1973.

Hallfors, Denise D., Martha W. Waller, Carol A. Ford, Carolyn T. Halpern, Paul H. Brodish & Bonita Iritani, "Adolescent Depression and Suicide Risk: Association with Sex and Drug Behavior," American Journal of Preventive Medicine, October 2004, 27(3), 224–231.

Hysenbegasi, Alketa, Steven L. Hass & Clayton R. Rowland, "The Impact of Depression on the Academic Productivity of University Students," Journal of Mental Health Policy and Economics, September 2005, 8(3), 145–151.

Kilpatrick, Dean G., Kenneth J. Ruggiero, Ron Acierno, Benjamin E. Saunders, Heidi S. Resnick & Connie L. Best, "Violence and Risk of PTSD, Major Depression, Substance Abuse/Dependence, and Comorbidity: Results from the National Survey of Adolescents," Journal of Consulting and Clinical Psychology, August 2003, 71(4), 692–700.

Kostanski, Marion & Eleonora Gullone, "Adolescent Body Image Dissatisfaction: Relationships with Self-esteem, Anxiety, and Depression Controlling for Body Mass," Journal of Child Psychology and Psychiatry, February 1998, 39(2), 255–262.

Lewinsohn, Peter M., Paul Rohde & John R. Seeley, "Psychosocial Risk Factors for Future Adolescent Suicide Attempts," Journal of Consulting and Clinical Psychology, April 1994, 62(2), 297–305.

Perri, Timothy J., "Health Status and Schooling Decisions of Young Men," Economics of Education Review, 1984, 3, 207–213.

Rosenbaum, Paul R. & Donald B. Rubin, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," Biometrika, April 1983, 70(1), 41–55.

Sanders, Christopher E., Tiffany M. Field, Miguel Diego & Michele Kaplan, "Moderate Involvement in Sports is Related to Lower Depression Levels among Adolescents," Adolescence, Winter 2000, 35(140), 793–797.

Shaffer, David, Madelyn S. Gould, Prudence Fisher, Paul Trautman, Donna Moreau, Marjorie Kleinman & Michael Flory, "Psychiatric Diagnosis in Child and Adolescent Suicide," Archives of General Psychiatry, April 1996, 53(4), 339–348.

Weissman, Myrna M., Susan Wolk, Risë B. Goldstein, Donna Moreau, Philip Adams, Steven Greenwald, Claudia M. Klier, Neal D. Ryan, Ronald E. Dahl & Priya Wickramaratne, "Depressed Adolescents Grown Up," Journal of the American Medical Association, 12 May 1999, 281(18), 1701–1713.

Wilcox-Gok, Virginia, Dave E. Marcotte, Farah Farahati & Carey Borkoski, "Early Onset Depression and High School Dropout," 27–51, in Marcotte, Dave E. & Virginia Wilcox-Gok, ed., The Economics of Gender and Mental Illness, Oxford: Elsevier, 2004.

Table 1: Variable list & means

	(1)	(2)	(3)
Sample	Overall	Sad	Not sad
Sample size	27,426	7,705	19,721
GPA past year	2.91 (0.91)	2.72 (0.96)	2.99 (0.88)
Felt sad or hopeless for 2+ weeks past year	.281	1	0
Considered suicide past year	.159	.399	.065
Planned suicide past year	.123	.303	.053
Attempted suicide past year	.074	.208	.022
Injured attempting suicide past year	.019	.060	.004
Female	.520	.650	.469
14 years old	.088	.081	.090
15 years old	.219	.219	.219
16 years old	.258	.264	.255
17 years old	.273	.274	.272
18+ years old	.163	.162	.163
9^{th} grade	.225	.228	.224
10^{th} grade	.245	.243	.245
11^{th} grade	.261	.269	.257
12^{th} grade	.269	.259	.273
White non-Hispanic	.518	.472	.537
Black	.151	.153	.151
Hispanic	.252	.291	.237
Non-white, black or Hispanic	.079	.085	.076
Female height (meters)	1.63 (0.07)	1.63 (0.07)	1.63 (0.07)
Male height (meters)	1.76 (0.08)	1.76 (0.08)	1.76 (0.08)
Female weight (kilograms)	61.0 (13.0)	61.6 (13.9)	60.6 (12.5)
Male weight (kilograms)	73.9 (16.7)	74.5 (17.0)	73.8 (16.7)
Smoked cigarettes in lifetime	.570	.692	.523
Smoked cigarettes daily for month in lifetime	.143	.217	.114
Never drank alcohol in lifetime	.217	.122	.254
Drank alcohol 1–2 days in lifetime	.141	.133	.143
Drank alcohol 3–9 days in lifetime	.180	.184	.179
Drank alcohol 10–19 days in lifetime	.123	.137	.117
Drank alcohol 20–39 days in lifetime	.119	.138	.112
Drank alcohol 40–99 days in lifetime	.103	.124	.095
Drank alcohol 100+ days in lifetime	.117	.161	.099
Never used marijuana in lifetime	.569	.459	.612
Used marijuana 1–2 times in lifetime	.097	.111	.091
Used marijuana 3–9 times in lifetime	.086	.105	.078
Used marijuana 10–19 times in lifetime	.049	.057	.046
Used marijuana 20–39 times in lifetime	.051	.064	.045
Used marijuana 40–99 times in lifetime	.047	.064	.040
Used marijuana 100+ times in lifetime	.102	.139	.087

Table 1: Variable list & means (continued)

	(1)	(2)	(3)
Sample	Overall	Sad	Not sad
Sample size	27,426	7,705	19,721
Used cocaine in lifetime	.088	.149	.064
Sniffed glue, paint or spray to get high in lifetime	.120	.208	.086
Used heroin in lifetime	.018	.037	.011
Used methamphetamines in lifetime	.062	.117	.041
Used steroids without prescription in lifetime	.034	.060	.023
Missed school because felt unsafe past 30 days	.046	.097	.027
Threatened with weapon at school past year	.068	.126	.046
Did not have physical fight past year	.680	.580	.719
Had 1 physical fight past year	.145	.177	.133
Had 2+ physical fights past year	.175	.243	.148
Offered/sold/given illegal drug at school past year	.263	.364	.223
Boy/girlfriend hit/slapped/phys. hurt past year	.092	.167	.062
Physically forced to have sex in lifetime	.072	.151	.040
Never had sex in lifetime	.497	.385	.541
Had 1 sex partner in lifetime	.180	.204	.171
Had 2 sex partners in lifetime	.099	.125	.088
Had 3 sex partners in lifetime	.069	.087	.062
Had 4 sex partners in lifetime	.041	.049	.037
Had 5 sex partners in lifetime	.030	.036	.028
Had 6+ sex partners in lifetime	.085	.113	.073
Describe self as underweight	.133	.140	.130
Describe self as about the right weight	.565	.496	.592
Describe self as overweight	.302	.364	.278
Did not play on sports team past year	.439	.501	.415
Played on 1 sports team past year	.253	.244	.257
Played on 2 sports teams past year	.168	.147	.176
Played on 3+ sports teams past year	.140	.109	.152

Parentheses contain standard deviations for non-binary variables. One indicator for each of age, grade level, race/ethnicity, alcohol use, marijuana use, physical fights, sex partners, self-described weight and sports teams is omitted from the regressions, which also include fixed effects for all but one of 155 PSU-by-year combinations.

Table 2: OLS regressions of GPA on sadness without co-morbidities

	(1)	(2)	(3)	(4)
Felt sad or hopeless 2+ weeks	–.269	–.296	–.289	–.227
	(21.4)	(23.9)	(23.7)	(17.4)
Considered suicide				–.074
				(3.42)
Planned suicide				–.025
				(1.11)
Attempted suicide				–.162
				(5.24)
Injured attempting suicide				–.063
				(1.24)
R-squared	.018	.092	.131	.137
Other covariates:				
Personal characteristics	No	Yes	Yes	Yes
PSU-by-year fixed effects	No	No	Yes	Yes

The sample size is 27,426. Parentheses contain heteroskedasticity-robust absolute *t*-statistics. All regressions include a constant term. Personal characteristics are indicators for age, gender, grade level and race/ethnicity, along with gender-specific height and weight.

Table 3: OLS regressions of GPA on sadness with co-morbidities

	(1)	(2)	(3)	(4)	(5)
Felt sad or hopeless 2+ weeks	–.159	–.142	–.120	–.118	–.107
	(12.5)	(11.1)	(9.42)	(9.17)	(8.44)
Considered suicide	–.044	–.026	–.020	–.018	–.014
	(2.12)	(1.24)	(0.98)	(0.86)	(0.69)
Planned suicide	–.006	.005	.013	.008	.016
	(0.27)	(0.21)	(0.62)	(0.36)	(0.74)
Attempted suicide	–.125	–.104	–.085	–.082	–.081
	(4.16)	(3.50)	(2.86)	(2.76)	(2.76)
Injured attempting suicide	.014	.061	.086	.097	.096
	(0.28)	(1.24)	(1.76)	(1.98)	(1.99)
R-squared	.193	.213	.221	.224	.241
Other covariates:					
Cigarette smoking	Yes	Yes	Yes	Yes	Yes
Alcohol and drug use	No	Yes	Yes	Yes	Yes
Violence	No	No	Yes	Yes	Yes
Sexual activity	No	No	No	Yes	Yes
Self-image & sports participation	No	No	No	No	Yes

The sample size is 27,426. Parentheses contain heteroskedasticity-robust absolute *t*-statistics. All regressions include a constant term, indicators for age, gender, grade level, race/ethnicity and PSUs-by-year, and gender-specific height and weight. Each category of "other covariates" is as listed in table 1, with physical and sexual assault considered measures of sexual activity rather than violence, i.e. not added until column 4.

Table 4: Coefficients of other covariates in OLS GPA regressions with co-morbidities

Female	.230 (20.4)
14 years old	.198 (6.06)
15 years old	.135 (5.35)
16 years old	.061 (3.48)
18+ years old	–.097 (5.84)
9th grade	–.265 (10.3)
10th grade	–.159 (8.82)
12th grade	.215 (12.9)
Black	–.259 (14.5)
Hispanic	–.244 (14.8)
Non-white, black or Hispanic	–.006 (0.29)
Female height (meters)	.435 (4.41)
Male height (meters)	.529 (5.10)
Female weight (kilograms)	–.004 (6.14)
Male weight (kilograms)	–.004 (6.68)
Smoked cigarettes in lifetime	–.170 (13.4)
Smoked cigarettes daily for month in lifetime	–.118 (6.39)
Drank alcohol 1–2 days in lifetime	–.097 (5.99)
Drank alcohol 3–9 days in lifetime	–.048 (3.02)
Drank alcohol 10–19 days in lifetime	–.061 (3.26)
Drank alcohol 20–39 days in lifetime	–.026 (1.29)
Drank alcohol 40–99 days in lifetime	–.029 (1.30)
Drank alcohol 100+ days in lifetime	–.016 (0.66)
Used marijuana 1–2 times in lifetime	–.094 (5.28)
Used marijuana 3–9 times in lifetime	–.131 (6.68)
Used marijuana 10–19 times in lifetime	–.189 (7.35)
Used marijuana 20–39 times in lifetime	–.164 (6.35)
Used marijuana 40–99 times in lifetime	–.208 (7.31)
Used marijuana 100+ times in lifetime	–.312 (12.3)
Used cocaine in lifetime	–.014 (0.58)
Sniffed glue, paint or spray to get high in lifetime	–.053 (2.98)
Used heroin in lifetime	.010 (0.18)
Used methamphetamines in lifetime	.020 (0.70)
Used steroids without prescription in lifetime	–.090 (2.71)
Missed school because felt unsafe past 30 days	–.155 (5.79)
Threatened with weapon at school past year	–.075 (3.28)
Had 1 physical fight past year	–.098 (6.73)
Had 2+ physical fights past year	–.176 (11.2)
Offered/sold/given illegal drug at school past year	–.031 (2.42)
Boy/girlfriend hit/slapped/phys. hurt past year	.002 (0.12)
Physically forced to have sex in lifetime	.018 (0.78)
Had 1 sex partner in lifetime	–.091 (6.49)
Had 2 sex partners in lifetime	–.166 (8.93)
Had 3 sex partners in lifetime	–.169 (7.79)
Had 4 sex partners in lifetime	–.131 (4.73)
Had 5 sex partners in lifetime	–.148 (4.60)
Had 6+ sex partners in lifetime	–.193 (8.13)
Describe self as underweight	–.033 (2.13)
Describe self as overweight	.004 (0.30)
Played on 1 sports team past year	.137 (11.1)
Played on 2 sports teams past year	.229 (16.6)
Played on 3+ sports teams past year	.348 (23.3)

The sample size is 27,426. Parentheses contain heteroskedasticity-robust absolute *t*-statistics. The regression corresponds to column 5 of table 3 and includes a constant term and PSU-by-year indicators. Omitted categories are those most prevalent: 17 years old, 11th grade, white non-Hispanic, never drank, never used marijuana, did not have a fight, never had sex, about the right weight and did not play on a sports team.

Table 5: OLS regressions of GPA on sadness in samples stratified on suicidality

Sample:	(1) Did not consider, plan or attempt suicide	(2) Considered or planned but did not attempt suicide	(3) Did not attempt suicide, i.e. (1) + (2)	(4) Attempted suicide
Sample size:	22,257	3,141	25,398	2,028
Felt sad or hopeless 2+ weeks	–.110	–.088	–.108	–.038
	(7.60)	(2.65)	(8.20)	(0.75)
Considered suicide		–.047	–.019	.127
		(1.08)	(0.87)	(1.49)
Planned suicide		–.004	.006	.096
		(0.10)	(0.25)	(1.80)
Injured attempting suicide				.041
				(0.79)

Parentheses contain heteroskedasticity-robust absolute *t*-statistics. The covariate set matches that for column 5 of table 3, with empty cells in this table reflecting omissions of students exhibiting the associated suicidality conditions from the given samples.

Table 6: OLS effects of sadness on GPA using alternative comparison groups

Panel	Sample	(1) Sample size	(2) Mean GPA	(3) Mean sad	(4) Effect of feeling sad on GPA
A.	Smoked cigarettes in life	15,646	2.72	.341	–.083 (4.51)
B.	Drank alcohol 10+ days in life	12,686	2.76	.341	–.094 (6.18)
C.	Used marijuana in life	11,811	2.63	.353	–.074 (3.88)
D.	Offered or used other drug	9,760	2.67	.400	–.087 (4.19)
E.	Had physical fight last year	8,786	2.62	.369	–.107 (4.74)
F.	Unsafe, threatened or assaulted	5,772	2.63	.503	–.108 (4.17)
G.	Had sex in life	13,794	2.71	.343	–.077 (4.56)
H.	Describe as under- or overweight	11,931	2.87	.325	–.105 (5.48)
I.	Did not play on sports team	12,048	2.78	.320	–.111 (5.67)

Parentheses contain heteroskedasticity-robust absolute *t*-statistics. The covariate set is identical to that for column 5 of table 3.

Table 7: Effects of sadness in ordered logit grade category regressions

Grade category	(1) Frequency	(2) Marginal effect of feeling sad	(3) Semi-elasticity, i.e. (2) ÷ (1)	(4) Predicted GPA change, i.e. (2) x GPA
A = 4.0	.283	–.045 (8.84)	–15.8%	–.179
B = 3.0	.421	–.005 (4.97)	–1.3%	–.016
C = 2.0	.236	.040 (8.41)	17.0%	.080
D = 1.0	.047	.008 (7.93)	16.8%	.008
F = 0.0	.014	.002 (7.46)	15.5%	0
Total	1			–.107

The sample size is 27,426. Parentheses contain heteroskedasticity-robust absolute *t*-statistics. The specification is identical to that for column 5 of table 3, except the dependent variable is the grade category measure listed in the left-hand column above instead of the numerical GPA used in the OLS regressions.

Table 8: IV regressions of GPA on sadness

	(1)	(2)	(3)	(4)	(5)
IV effect of feeling sad	–.117	–.115	–.105	–.115	–.100
	(2.38)	(2.35)	(2.28)	(2.36)	(2.23)
J-statistic for overidentification	0.70	1.32	1.40	0.80	2.11
	[.404]	[.725]	[.706]	[.669]	[.910]
F-statistic for IV in reduced form model (table 3, column 5)	0.37	0.33	0.35	0.27	0.30
	[.694]	[.856]	[.846]	[.845]	[.953]
First-stage IV *F*-statistic	697.8	358.9	426.6	484.5	257.8
First-stage partial *R*-squared	.068	.069	.078	.070	.081
Instruments:					
Considered suicide	Yes	Yes	Yes	Yes	Yes
Planned suicide	Yes	Yes	Yes	Yes	Yes
Used heroin	No	Yes	No	No	Yes
Used methamphetamines	No	Yes	No	No	Yes
Boy/girlfriend physically abused	No	No	Yes	No	Yes
Physically forced to have sex	No	No	Yes	No	Yes
Describe self as overweight	No	No	No	Yes	Yes

The sample size is 27,426. Parentheses contain heteroskedasticity-robust absolute *t*-statistics and brackets contain *p*-values. The covariate set is identical to that for column 5 of table 3, with regressions estimated using two-stage GMM. Instrument coefficients in the first-stage feeling sad regression are: considered suicide (.308, $t = 27.5$), planned suicide (.116, $t = 9.93$), used heroin (–.085, $t = 3.74$), used methamphetamines (.054, $t = 3.92$), boy/girlfriend physically abused (.093, $t = 9.42$), forced to have sex (.087, $t = 7.62$), and overweight (.037, $t = 5.45$).

Table 9: OLS effects of sadness on GPA in samples stratified by exogenous factors

Panel	Sample	(1) Sample size	(2) Mean GPA	(3) Mean sad	(4) Effect of feeling sad on GPA
A.	Female	14,263	3.03	.351	–.115 (7.22)
	Male	13,163	2.78	.205	–.095 (4.47)
B.	Age 14–15	8,419	2.92	.275	–.108 (4.42)
	Age 16	7,070	2.88	.288	–.114 (4.44)
	Age 17–18	11,937	2.93	.281	–.094 (5.18)
C.	9^{th} grade	6,172	2.87	.285	–.102 (3.47)
	10^{th} grade	6,715	2.86	.279	–.134 (4.99)
	11^{th} grade	7,150	2.90	.290	–.086 (3.67)
	12^{th} grade	7,389	3.00	.271	–.083 (3.60)
D.	Non-Hispanic white	14,214	3.04	.256	–.104 (5.67)
	Hispanic or non-white	13,212	2.77	.308	–.110 (6.17)
E.	Year 2001	8,430	2.87	.292	–.093 (4.14)
	Year 2003	8,759	2.90	.290	–.105 (4.77)
	Year 2009	10,237	2.96	.264	–.119 (5.51)

Parentheses contain heteroskedasticity-robust absolute *t*-statistics. The covariate set is identical to that for column 5 of table 3.