What Can We Learn About the Effect of Mental Health on Labor Market Outcomes Under Weak Assumptions? Evidence from the NLSY79.

Giuseppe Germinario[1]; Vikesh Amin[2]; Carlos A. Flores[3]; Alfonso Flores-Lagunes[4]

[1] Syracuse University
[2] Central Michigan University
[3] California Polytechnic State University
[4] Syracuse University, IZA, and GLO

Abstract

We employ a nonparametric partial identification approach to bound the causal effect of poor mental health on employment and earnings using the National Longitudinal Study of Youth 1979. Our approach allows us to provide bounds on the population average treatment effect based on relatively weak, credible assumptions. We find that being categorized as depressed decreases employment by 10% and earnings by 27% at most, but we cannot statistically rule out a zero effect. We also provide insights into the heterogeneity of the effects on labor market outcomes at different levels of adverse mental health experienced (no, little, mild, moderate, and severe depressive symptoms). We find that going from having no (little) to severe depressive symptoms reduces employment by 3-18% (3-16%) and earnings by 11-44% (12-36%). The estimated bounds statistically rule out null effects for earnings but not for employment.

Key Words: depression; mental health; employment; earnings; partial identification; bounds

We thank Editor Daniel Bennett and a reviewer for helpful and constructive comments. Thanks also to Gary Engelhardt, Osea Giuntella, Hugo Jales, Jeffrey Kubik, Eunhye Kwak, Shirlee Lichtman-Sadot, Analisa Packham, Emily Wiemers, Maria Zhu, and participants at the 2022 SOLE meetings for useful comments and discussions. We acknowledge research funding from NIH grant number 1R01HD094011-01. Amin also acknowledges research funding from the College of Business Administration at Central Michigan University. Flores-Lagunes and Germinario gratefully acknowledge financial support from the Melvin A. Eggers fund at the Maxwell School of Citizenship and Public Affairs.

1. Introduction

It is well-documented that poor mental health is correlated with worse labor market outcomes. A review of early studies by Marcotte and Wilcox-Gök (2001) concluded that 5–6 million workers in the US between the ages of 16 and 54 lose, fail to seek, or cannot find employment due to mental illnesses, and that mental illness decreases annual income by $3,500-$6000, conditional on working. There are several channels through which poor mental health can lead to worse labor market outcomes. Mental health might affect employment through beliefs and preferences. Depression is associated with negative beliefs about oneself and pessimistic expectations about the future, which may lead to reduced motivation and effort in job searches, as well as easily becoming discouraged with setbacks. Some studies have found that anxiety is associated with lower risk-taking (Giorgeta et al. 2012; Maner et al. 2007), which may affect individuals' decisions to apply for jobs. Employment could also be affected through mental health stigma and discrimination. Depressed individuals may want to avoid situations that hold a high risk of stigmatization and simply may not apply for jobs. Stigma may exclude depressed individuals from social networks that provide economic opportunities. Employers may discriminate and not hire depressed individuals due to stereotypes such as being incompetent and unreliable. Depression can affect labor market earnings through labor supply and productivity. Depression typically involves disrupted sleep, fatigue and concentration problems which might lead to higher rates of absenteeism, lower labor supply, lower productivity, and thus lower earnings. At the same time, depression may decrease the marginal utility of leisure (Nimrod et al. 2012), resulting in depressed individuals potentially working more since they receive less utility from non-work activities.

Despite multiple pathways linking poor mental health to labor market outcomes and empirical associations which do the same, establishing causality is complicated because of omitted variable bias and reverse causality that lead to endogeneity of mental health. Omitted variable bias arises due to "third" factors (e.g., genetic endowments, cognitive ability, childhood circumstances, personality) that are correlated with both mental health and labor market outcomes, while reverse causality occurs when lack of employment or reduced earnings worsen mental health. To identify causal effects, several studies (summarized in Section 2) have used instrumental variables (IV) strategies to isolate variation in mental health that is argued to be orthogonal to unobserved factors. Instruments used have pertained to childhood and parental mental health, religiosity, death of a close friend, and availability of social services and social support. Other studies have used panel data and individual fixed-effect regression specifications to control for time-invariant individual characteristics correlated with mental health and labor market outcomes. The consensus from prior studies is that poor mental health reduces employment but effects for earnings are mixed.

While IV and panel data approaches are an improvement upon naïve OLS estimates, the assumptions needed for causality may be violated. Valid instruments rely on the exclusion restriction—that the instruments only affect labor market outcomes indirectly through their effect on mental health. One can think of reasons why the exclusion restriction may not be satisfied for each of the instruments used in the literature. Childhood/adolescent mental health may violate the exclusion restriction because poor adolescent mental health is associated with lower educational attainment (Fletcher 2008) and worse labor market outcomes (Fletcher 2013a; Lundberg et al. 2014; Mousteri et al 2019). Parental mental illness has been found to increase the probability of high school dropout of children (Farahati et al. 2003), which could in turn affect labor market outcomes. Religiosity is likely correlated with earnings (Gruber 2005) and could affect labor market outcomes through other indirect channels (e.g., attending religious services could be helpful for career networking). Furthermore, in the presence of heterogeneous effects, IV identifies a local average treatment effect (LATE) for those individuals whose treatment (mental health) is affected by the instrument used.[1] Individual fixed effects do not control for time-varying factors correlated with mental health and labor market outcomes, such as stress and health shocks. Individual fixed-effect estimates are identified from individuals who experience changes in mental health and labor market outcomes. If within-person variation in mental health and labor market outcomes are randomly distributed across all individuals, then individual fixed-effect estimates will identify the average treatment effect (ATE) on the population, which is arguably a more policy-relevant parameter. Alternatively, the estimates could reflect an effect for a specific subpopulation if the within-person variation is present in certain subgroups of the population. Moreover, if there is insufficient within-person variation in mental health and labor market outcomes, then estimates may lack statistical power to detect effects of mental health.

We estimate the effect of mental health (based on a count of depressive symptoms from the Center for Epidemiologic Studies Depression Scale) in the 1992 round (when individuals are 30 years old) of the National Longitudinal Study of Youth 1979 (NLSY79) on employment and earnings in 1993. Our key contribution is that we provide novel evidence by employing a nonparametric partial identification approach (Manski and Pepper 2000) to bound the causal effect, which has at least four advantages. First, it accounts for the endogeneity of mental health that may arise from, e.g., omitted variables or reverse causality. Second, it provides bounds on the population ATE as opposed to a subpopulation treatment effect. Third, it relies on relatively weak assumptions, which are, arguably, more credible. These advantages, though, come at a cost of obtaining a range of possible values for the causal population ATE rather than a point estimate of it. The assumptions we use are three: (1) monotone treatment selection (MTS) which posits that, on average, individuals "selected" into worse mental health have lower latent employment probabilities and earnings; (2) monotone treatment response (MTR) which imposes the restriction that worse mental health does not improve labor market outcomes; (3) we employ adolescent test scores as a monotone instrumental variable (MIV)—a variable that is assumed to have a weakly increasing mean relationship with potential outcomes—to help tighten the bounds under the MTS and MTR assumptions. In our context, the MIV assumption states that individuals with higher test scores have no lower average latent employment probabilities and earnings than those with lower test scores. The MIV assumption is weaker than the exclusion restriction in IV models, since the MIV is allowed to have a direct impact on potential outcomes and may be non-random itself. We discuss and assess these assumptions below. Fourth, the bounds allow us to explore the strength of the estimated effects at different levels of depression symptoms experienced (no, little, mild, moderate, and severe). The nonparametric partial identification method that we employ has been used in other contexts but to our knowledge has not been applied to bound effects of mental health on labor market outcomes.[2]

Our estimated bounds indicate that being categorized as depressed decreases employment by 10% and earnings by 27% at most, but we cannot statistically exclude a zero effect at conventional significance levels. Mental health, however, is inherently a continuous condition, but an indicator variable for being depressed or having a psychiatric condition assumes away any heterogeneity in the way the intensity of depressive symptoms affects labor market outcomes. We examine the effects at different levels of depressive symptoms (none, little, mild, moderate, and severe), and find that going from having no (little) to severe depressive symptoms reduces employment by 3-18% (3-16%) and earnings by 11-44% (12-36%). The 95% confidence

[1] In this sense, estimates derived from random shocks such as death of a close friend are potentially less informative for policy.

[2] Examples of previous applications include bounding the causal effect of parents' schooling on children's schooling (De Haan 2011), unemployment on mental health (Cygan-Rehm et al. 2017), English proficiency on labor market outcomes (Gonzalez 2005), criminal convictions on labor market outcomes (Richey 2015), education on social support (Huang et al. 2012), and social activities on cognition (Christelis and Dobrescu 2020).

intervals statistically rule out null effects for earnings but not employment. Therefore, our results provide evidence of a statistically significant average causal effect of worsening mental health from no/little to severe depression symptoms on labor market earnings, while also ruling out relevant magnitudes. The effect for earnings is economically significant, comparable to IV estimates of an additional year of schooling on wages which is between 6-15% (Card 1999).

The remainder of the paper is organized as follows. Section 2 provides a review of the literature and Section 3 gives a description of the data. Our econometric approach is explained in Section 4 and the results are presented in Section 5. Finally, Section 6 concludes.

2. Literature Review

Most of the early studies on the topic documented associations based on OLS regressions. Bartel and Taubman (1979) were the first to document negative associations between poor mental health and labor market outcomes. They used a sample of white male veteran twins (National Academy of Science-National Research Council twin panel) and found that a psychiatric diagnosis (either psychoses or neuroses) between 1968-1973 was associated with a 27% reduction in 1973 earnings. Bartel and Taubman (1986) extended their original analysis by distinguishing between psychoses and neuroses, and by using earnings from 1954-1973. They found that both psychoses and neuroses were associated with lower earnings, with the association lasting as long as 15 years.[3] Frank and Gertler (1991) found that mental distress was associated with a 21% decrease in earnings for men from the Baltimore Epidemiological Catchment Area (ECA) Study.[4] Mitchell and Anderson (1989) used data on adults aged 50 and over from the Baltimore, Durham and Los Angles ECA sites, and found that having more mental health symptoms was associated with a lower probability of employment among men but not women. Miller and Kelman (1992) used data from four ECA sites and found that schizophrenia was associated with a 10-35% decrease in earnings, while anxiety disorders were associated with a 3-10% reduction in earnings. They also found that affective disorders were associated with higher earnings, which they attribute to the endogeneity of mental health.

To study the causal effect of mental health on labor market outcomes, several papers have estimated IV models using the National Comorbidity Survey (NCS), a nationally representative survey of 15-54-year-olds designed to study the prevalence, causes and consequences of comorbidity between substance abuse disorders and nonsubstance abuse psychiatric disorders. Using the 1990-1992 NCS and instrumenting adult mental health with childhood and parental mental illnesses, Ettner et al. (1997) found that a diagnosis of any psychiatric disorder in the last year reduced the probability of being employed by 11 percentage points. They also found negative effects of a past-year psychiatric disorder on earnings, though IV estimates were only statistically significant for women. Marcotte and Wilcox-Gök (2003) estimated the effect of depression, dysthymia, anxiety disorder and anti-social personality disorder on earnings in the 1990-1992 NCS using measures of parental mental health as instruments. They only found a statistically significant negative effect for anxiety disorder on earnings for women, with IV estimates showing that an anxiety disorder in the past year reduced earnings by 49%. Based on IV quantile regressions, they did, however, find large negative effects of all four disorders at the lower tail of the earnings distribution. Chatterji et al. (2011) estimated the effect of a past-year psychiatric disorder on employment and earnings in the 2001-2003 National Comorbidity Survey-Replication, using childhood mental health and religiosity as instruments. Their IV estimates showed that a past-year psychiatric disorder reduced the

[3] A diagnosis of psychoses (neuroses) 11-15 years prior to date of earnings was associated with a 32% (14%) reduction in earnings.

[4] Individuals were defined as being in mental distress if they had at least two of the following: a score of 4 or more on the 20-item General Health Questionnaire, a self-reported disability day, DSM-III diagnosis (based on the diagnostic interview schedule and mini-mental state examination scores).

likelihood of employment by 17 percentage points for men and 9 percentage points for women.[5,6] Unlike Ettner et al. (1997) and Marcotte and Wilcox-Gök (2003), Chatterji et al. (2011) did not find any evidence of negative effects for earnings. They conjecture that the difference in findings could be due to the US economy being stronger in 2001-2003 compared to 1990-1992.

Other IV studies for the US not using the NCS are Alexandre and French (2001), Chatterji et al. (2007) and Ojeda et al. (2010). Alexandre and French (2001) used survey data collected from a low-income neighborhood in Miami-Dade County, 1996-1997. Using religiosity and social supports as instruments, they found that depression reduces employment by 19 percentage points. Chatterji et al. (2007) estimated labor market effects of a past-year psychiatric disorder in the 2002-2003 National Latino and Asian American Study using childhood mental health and religiosity as instruments. They found large effects for Latinos— a past-year psychiatric disorder decreased employment by 8 percentage points for men and 26 percentage points for women— and smaller negative effects of 1-3% on employment for Asians. Ojeda et al. (2010) estimated the effect of the K6 Scale of Mental Illness on labor supply for natives and immigrants in the 2002 National Survey on Drug Use and Health using social support as the IV. They found that immigrants' labor supply is less responsive to mental health problems than the labor supply of natives.

Other papers have used panel data methods to account for the influence of unobserved factors. Using the 2004-2009 US Medical Expenditure Panel Survey and correlated random effects regressions, Peng et al. (2015) found that depression reduced the probability of employment by 2.6 percentage points but didn’t affect earnings. Cseh (2008) estimated regressions with individual fixed effects using the 1992, 1994 and 2004 rounds of the NLSY79, and found that depression didn’t affect wages for women, but reduced wages of men by 3.4-4.3%. Frijters et al. (2014) estimated the effect of mental health on employment using 10 waves of the Household, Income and Labour Dynamics in Australia survey. They constructed a mental health index based on nine questions from the short-form general health survey. They used death of a close friend as an instrument and estimated IV regressions with individual fixed effects. They found that a one standard deviation decrease in the mental health index decreased employment by 30 percentage points. Bryan et al. (2022) used nine waves of the UK Household Longitudinal Study and estimated individual fixed-effects regressions which showed that depression reduces employment by 1.6 percentage points.

Another strand of the literature has conducted randomized control trials (RCTs) mostly in lower- and middle-income countries to assess the effectiveness of interventions (e.g., cognitive behavioral therapy, antidepressant access) to improve mental health. Some of these RCTs have also examined labor market outcomes of individuals after interventions. For example, Patel et al. (2017) conducted an RCT of a brief behavioral activation therapy program on individuals with moderately-severe to severe depression in Goa, India. They found that treated individuals were 60% more likely to be in remission three months after the intervention and reported being able to work 2.3 more days per month compared to individuals in the control group. Lund et al. (2018) conducted a meta-analysis of 39 RCTs in lower- and middle-income countries, where they found some suggestive evidence that improvements in mental health have positive effects on economic

[5] These IV estimates are from appendix tables 2A and 2B in their NBER 2008 working paper (Chatterji et al. 2008). The IV results for employment are not shown in their published paper (Chatterji et al. 2011).

[6] Findings were similar based on bivariate probit models using the methods in Altonji et al. (2005), where identification comes from making assumptions about the correlation between unobserved and observed factors determining the outcome and endogenous variable, and the functional form. The bivariate probit estimates showed that a past-year psychiatric disorder decreases employment by 14 percentage points for men and 13 percentage points for women.

outcomes.[7] While such RCTs have not been conducted in high-income countries, Biasi, et al. (2021) leveraged the approval of lithium as a treatment for bipolar disorder in Denmark in 1976 as a natural experiment to identify the causal effect of access to treatment on career earnings. They compared those with access at age 20—the usual age for onset of bipolar disorder—to those born prior to 1956, using administrative data. They found that access to treatment reduced the observed earnings difference of 38% between those with bipolar disorder and those without by 28% and reduced the probability of zero earnings by 33%.

In sum, previous studies have consistently found negative effects of poor mental health on employment in the US and internationally, with mixed findings for earnings. IV estimates typically show larger detrimental effects of poor mental health on employment than individual fixed-effect estimates.

3. Data

We use the NLSY79, which is a nationally-representative panel study that follows 12,686 individuals from 1979, when they were aged 14–22, through the present. Of the 12,686 individuals, (1) 6,111 are from a cross-sectional sample designed to be nationally representative of the noninstitutionalized young people in the US in 1979 (2) 5,295 are from an oversample of Hispanic, Black, and economically disadvantaged, non-Black and non-Hispanic youths and (3) 1,280 are from a military sample. We measure mental health using the Center for Epidemiologic Studies Depression Scale (CES-D), which is widely used to measure symptoms of depression, and has been tested in multiple settings for validity and reliability (Cosco et al. 2017; Radloff 1977).[8] The CES-D assesses depressive symptoms as a continuum and measures the frequency of depressive symptoms over one week with 20 items. For example, respondents are asked how often in the last week they were bothered by things not normally bothersome; could not shake the blues; and felt like they were not as good as others. With four response categories (0 = none to 3 = almost every day), the maximum score of the scale is 60. A score of 16 or more on the CES-D indicates that one has a higher risk of clinical depression. Individuals can also be classified as having little-or-no depressive symptoms (0≤CES-D≤15), mild depressive symptoms (16≤CES-D≤20) moderate depressive symptoms (21≤CES-D≤25) and severe depressive symptoms (26≤CES-D≤60) (Unützer et al. 2002). We further differentiate between individuals with no depressive symptoms and those with little depressive symptoms (1≤CES-D≤15). The original 20-item CES-D was first administered at the 1992 round, when individuals were 30 years old on average. A total of 9,016 individuals were interviewed in 1992, of which 8,978 have valid data on the CES-D score.

We estimate the effect of mental health measured in 1992 on employment and annual earnings in the following year, 1993, to reduce concerns of reverse causality, following prior literature (e.g., Ringdal and Rootjes 2022). However, our bounding approach addresses endogeneity arising from reverse causality. Employment is measured by a dummy variable equal to one if the individual reports working positive hours, and zero if the individual reports working zero hours. Individuals who report being employed but work zero hours, and individuals in the armed forces are coded as missing. A useful aspect of the NLSY79 is that it contains Armed Forces Qualification Test (AFQT) scores from the 1981 round when respondents were between

[7] Their analysis documenting that improvements in mental health have positive effects on economic outcomes is based on using the intervention content as an IV for mental health at the level of the study (see the description in their mechanism analysis).

[8] While widely used (e.g., Fletcher 2013b; Koford and Cseh 2015), the CES-D is not without criticism (e.g., Mukamal 2015; Lu et al. 2017). Some researchers prefer the nine-item Patient Health Questionnaire (PHQ-9) scale for certain subpopulations (e.g., Milette et al. 2010; Zhang et al. 2015), although other studies report the two scales as comparable (e.g., Engel et al. 2020). The CES-D is the only available depression scale in the NLSY79.

16-21 years old. As we describe in the next section, we use the AFQT score as a monotone instrumental variable. In total we have a sample of 7,665 individuals with non-missing information on depressive symptoms, employment, earnings and the AFQT score. Throughout our analysis we use the 1993 panel weights provided in the NLSY79.[9]

Summary statistics for our estimation sample are shown in Table 1. In the full sample (column 1), the average age in 1992 is 31 years and 51% are female. White individuals make up 80% of the sample, and there is a higher proportion of Black individuals (14%) than Hispanic individuals (6%).[10] The average CES-D score in 1992 is 9.04 and 18% are classified as depressed (CES-D≥16). Most individuals (72%) have little depressive symptoms (1≤CES-D≤15), while 10% have no depressive symptoms, 8% have mild, 4% have moderate, and 6% have severe depressive symptoms. Women have worse mental health than men—22% of women are depressed whereas 15% of men are depressed. Black and Hispanic individuals are also more likely to be depressed than whites. The 1993 employment rate in the full sample is 88%, but there are differences by gender and race/ethnicity. The employment rate of men (94%) is substantially higher than women (82%). Similarly, white individuals have a higher employment rate (89%) than Black (80%) and Hispanic (84%) individuals. A similar pattern is also observed for earnings. Average earnings (in 1993 dollars) are higher for men than women, and for white individuals compared to Black and Hispanic individuals.

4. Methodology and Identifying Assumptions

Let every individual i have a response function $Y\,(\cdot)\colon T \rightarrow Y$ which maps treatments $t \in T$ into potential outcomes $y_i\,(t)\, \epsilon\, Y$. In our context, the treatment t is depressive symptoms measured by the CES-D score in the form of either a binary indicator for depression or a discretized version consisting of five levels: no, little, mild, moderate, and severe depressive symptoms. The outcome Y is employment or earnings. Let S_i denote the realized treatment received by individual i, so that $Y_i \equiv \sum_{t \in T} 1\{S_i = t\} \cdot Y_i\,(t)$ is the associated observed outcome.

To illustrate how the bounds are obtained, in this section we will focus on the case of two treatment levels $t_1,$ and t_2 which may correspond to any of the treatments analyzed: no depression/depression, or no, little, mild, moderate, and severe depressive symptoms. We are interested in the population ATE of worsening mental health from t_1 to t_2 on labor market outcomes defined as:

$$(1)\quad \Delta\,(t_2, t_1) = E[Y(t_2)] - E[Y(t_1)]$$

Estimation of the ATE is complicated because the potential outcome $Y(t_2)$ is unobserved for individuals with a different treatment level from t_2, and the potential outcome $Y(t_1)$ is unobserved for individuals with different treatment level from t_1. This identification problem can be seen by using the law of total probability to write the expected potential outcome $E[Y(t_2)]$ as:

$$(2)\; E[Y(t_2)] = E[Y(t_2)|S = t_2] * P(S = t_2) + E[Y(t_2)|S \neq t_2] * P(S \neq t_2)$$

[9] The information on primary sampling units (PSUs) for the NLSY79 is part of the restricted geocode data, which presents a complication for the bootstrapping inside our estimation and inference procedure described below. We use the respondent's Census region as the PSU, which is the finest level of geographical information available in the public use dataset. We also follow Altonji et al. (2012) and stratify based on race/ethnicity and sex for the bootstrap samples.

[10] Respondents reported their race/ethnicity at baseline, which was mapped into three groups: Black, Hispanic, and non-Black/non-Hispanic. We refer to the last group as white, although it is a heterogenous group.

The data identify the sample analogues of all the right-hand side quantities except of the counterfactual $E[Y(t_2)|S \neq t_2]$. We thus need to impose identifying assumptions about this missing counterfactual. One can similarly analyze $E[Y(t_1)]$. Manski (1989) suggested a bounded support assumption, whereby one imposes minimum (Y_{min}) and maximum (Y_{max}) values for the outcome variable in place of $E[Y(t_2)|S \neq t_2]$ (or $E[Y(t_1)|S \neq t_1]$ in the case of $E[Y(t_1)]$). Often, an outcome has natural boundaries, in which case the assumption is innocuous. An indicator for employment trivially has $Y_{min} = 0$ and $Y_{max} = 1$, and earnings also naturally are no less than zero. A choice of Y_{max} for earnings, on the other hand, is less obvious. In such cases, the bounded support assumption can instead serve to define the population of interest. In this vein, we adopt the in-sample maximum earnings in the NLSY as Y_{max}, such that we restrict attention to the space of earnings reported in the data. Bounded support gives Manski's (1989) "no-assumption" bounds (once Y_{min} and Y_{max} are appropriately substituted for the missing counterfactual $E[Y(t_2)|S \neq t_2]$):

$$(3) E[Y|S = t_2] * P(S = t_2) + Y_{min} * P(S \neq t_2) \leq E[Y(t_2)] \leq E[Y|S = t_2] * P(S = t_2) + Y_{max} * P(S \neq t_2)$$

Figure 1 in Panel A illustrates graphically the construction of the no-assumption bounds on $E[Y(t_2)]$, as in De Haan (2011), where the x-axis represents levels of the realized depressive symptoms (S) while the y-axis represents values of the outcome (Y).

The no-assumption lower (respectively, upper) bound on the ATE $\Delta\,(t_2, t_1)$ is calculated by subtracting the upper (lower) bound of $E[Y(t_1)]$ from the lower (upper) bound of $E[Y(t_2)]$. That is, no-assumption bounds on $\Delta\,(t_2, t_1)$ are:

$$(4)\ E[Y|S = t_2] * P(S = t_2) + Y_{min} * P(S \neq t_2) - E[Y|S = t_1] * P(S = t_1) - Y_{max} * P(S \neq t_1)$$
$$\leq \Delta\,(t_2, t_1) \leq$$
$$E[Y|S = t_2] * P(S = t_2) + Y_{max} * P(S \neq t_2) - E[Y|S = t_1] * P(S = t_1) - Y_{min} * P(S \neq t_1)$$

Bounds for other treatment effects such as $\Delta\,(t_5, t_1)$ or $\Delta\,(t_4, t_2)$ are computed analogously. In practice, the no-assumption bounds are typically wide and contain zero by construction. They can, however, potentially rule out large positive or large negative values of the ATE. To tighten the no-assumption bounds, we employ three monotonicity assumptions introduced in Manski (1997) and Manski and Pepper (2000): (1) monotone treatment selection, (2) monotone treatment response, and (3) monotone instrumental variable.

4.1 Monotone Treatment Selection (MTS)

The MTS assumption captures the notion that individuals who "selected" into higher CES-D scores have lower latent employment probabilities and earnings. Formally, the non-positive MTS assumption states that, for each $t \epsilon T$ and two treatment levels μ_1 and μ_2,

$$(5)\ \mu_2 > \mu_1 \Rightarrow E[Y(t) \mid S = \mu_2] \leq E[Y(t) \mid S = \mu_1]$$

In our context, the MTS assumption requires that, on average, individuals who "self-select" into worse mental health have weakly lower potential labor market outcomes (at every level of depressive symptoms) than individuals who "self-select" into better mental health. This assumption is untestable since some potential outcomes are counterfactual. However, we consider it plausible since there is evidence that poor mental health is associated with traits such as lower measured intelligence and higher neuroticism which are in turn associated with lower employment and earnings (Lin et al. 2018; Fletcher 2013b). Moreover, we can indirectly shed light on the plausibility of the MTS assumption in our sample by computing average characteristics of individuals in different categories of depressive symptoms. Intuitively, we would expect to see that individuals with worse depressive symptoms show traits correlated with worse labor market outcomes.[11] Table 2 provides summary statistics for some observed characteristics by the level of depressive symptoms. Average parental education is lower for individuals with higher depressive symptoms. For example, mothers' (fathers') average years of education are 11.94 (12.21) for individuals with no depressive symptoms, whereas it is 11.70 (11.94) for individuals with little depressive symptoms. One exception is that individuals with severe depressive symptoms have, on average, more educated parents than individuals with moderate depressive symptoms. However, the corresponding differences are not statistically different from zero, and thus do not contradict the weak inequality in the MTS assumption. Individuals with higher depressive symptoms also have lower AFQT scores in adolescence and are less likely to expect to graduate from college.[12] Education attained in adulthood (by 1993) may be influenced by mental health, but it is suggestive that individuals with higher depressive symptoms have fewer years of education and are less likely to be college graduates. Given the observed differences in those characteristics in Table 2, individuals with worse depressive symptoms appear negatively selected in terms of labor market outcomes (e.g., generally have parents with lower years of education, have lower ability, and have lower education), which is consistent with the MTS assumption.[13]

In spite of the indirect evidence in Table 2, the MTS assumption could still be violated if, on average, individuals with worse mental health have better potential labor market outcomes than individuals with better mental health (i.e., positive selection). For this to happen, for instance, it would have to be the case that the statistically insignificant difference in parent's education between those with severe and moderate depressive symptoms is meaningful and that it gives rise to mean potential labor market outcomes for these two groups that violate MTS. One possibility could be that more educated parents are more likely to have their children diagnosed with more severe depression symptoms and, simultaneously, they help their children have better labor market potential outcomes. In our view, while this could be the case for *some* of the individuals, it is hard to think that this situation is so pervasive so as to change the direction of the average potential outcomes stated in MTS in equation (5).

To illustrate the derivation of the bounds on the ATE $\Delta\,(t_2, t_1)$ under the MTS assumption, consider bounding the term $E[Y(t_2)]$ in equation (1). Using the law of total probability, we can write $E[Y(t_2)]$ as:

$$(6)\ E[Y(t_2)] = E[Y(t_2)|S < t_2] * P(S < t_2) + E[Y(t_2)|S = t_2] * P(S = t_2) + E[Y(t_2)|S > t_2] * P(S > t_2).$$

Figure 1 Panel B illustrates how MTS tightens the no-assumption bounds. For the conditional mean potential outcome $E[\,Y(t_2)\mid S < t_2\,]$, the no-assumption bounds in Figure 1 Panel A could only conclude that this object is bounded below by Y_{min}. As shown in Figure 1 Panel B, the MTS

[11] In principle, this exercise should use characteristics that are not themselves affected by the individual's mental health. Some of the characteristics in Table 2 satisfy this condition (e.g., mother's and father's education), while others may not (e.g., years of education). We still present averages for the latter variables for reference.

[12] At the 1981 round in the NLSY, respondents were asked "as things now stand, what is the highest grade or year you think you will actually complete?". We classify individuals as expecting to graduate college if they report 4th year of college, or 5th year of college, or 6th+ year of college.

[13] Corresponding summary statistics by gender and by race/ethnicity are given in Online Appendix Table A1. A similar pattern is observed to that in Table 2. The one notable point is that for Hispanic individuals average parental education, AFQT scores, college expectations and educational attainment are all higher for individuals with severe depressive symptoms compared to individuals with moderate depressive symptoms. However, these differences are not statistically different from zero, and thus they remain consistent with the weak inequality in the MTS assumption. The one exception is father's years of education, where the difference is statistically different from zero.

equation (5) implies that this cannot be less than $E[\,Y(t_2)\mid S = t_2\,]$, which is point identified by the observed mean for those receiving t_2 ($E[Y|S = t_2]$). Similarly, for the conditional mean potential outcome $E[\,Y(t_2)\mid S > t_2\,]$, bounded support alone (Figure 1 Panel A) could only yield an upper bound using Y_{max}. As shown in Figure 1 Panel B, using equation (5), MTS implies that the unidentified quantity can be no larger than $E[\,Y(t_2)\mid S = t_2\,]$, which is point identified. Then, the MTS bounds on $E[Y(t_2)]$ are given by (Manski and Pepper, 2000):

$$(7)\ E[Y|S = t_2] * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + Y_{min} * P(S > t_2) \\ \leq E[Y(t_2)] \leq \\ Y_{max} * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + E[Y|S = t_2] * P(S > t_2)$$

As before, the lower (respectively, upper) bound on the ATE $\Delta\,(t_2, t_1)$ is calculated by subtracting the upper (lower) bound of $E[Y(t_1)]$ from the lower (upper) bound of $E[Y(t_2)]$:

$$(8)\ E[Y|S = t_2] * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + Y_{min} * P(S > t_2) \\ - Y_{max} * P(S < t_1) - E[Y|S = t_1] * P(S = t_1) - E[Y|S = t_1] * P(S > t_1) \\ \leq \Delta\,(t_2, t_1) \leq \\ Y_{max} * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + E[Y|S = t_2] * P(S > t_2) \\ - E[Y|S = t_1] * P(S < t_1) - E[Y|S = t_1] * P(S = t_1) - Y_{min} * P(S > t_1)$$

and other comparisons of interest are obtained likewise.

4.2 Monotone Treatment Response (MTR)

The non-positive MTR assumption we employ imposes the restriction that worse mental health (higher CES-D scores) does not improve labor market outcomes. Formally, for each individual and any treatment levels t_k and t_j:

$$(9)\ t_j > t_k \Longrightarrow Y(t_j) \leq Y(t_k)$$

As noted in the Introduction and Section 2, the vast majority of channels through which poor mental health can affect labor market outcomes worsens them. All of these channels are consistent with the MTR assumption. Theoretical models also imply that poor health leads to worse labor market outcomes. For example, the Grossman (1972) health investment model shows that poor health reduces time available for work because of increased time spent being ill, increased preferences for leisure, or increased time needed to maintain health. Health is also an input into the production function for human capital, and those with worse health have lower human capital investments. Given the positive correlation of human capital investments and labor market outcomes, as well as the economic theory behind the positive effect of human capital accumulation on labor market outcomes (e.g., Card 1999), MTR is consistent with those theories.

However, the MTR assumption must hold for all individuals, and it is possible that it may not. For example, depression might decrease the marginal utility of leisure, as depressed individuals derive less pleasure from things they would normally enjoy (e.g., Nimrod et al. 2012), leading potentially to higher labor supply and earnings. Another possible violation could stem from employer provided health insurance—depressed individuals may continue to work because they rely on health insurance provided by their employer. Mental health is also linked with creativity, with studies showing that individuals with bipolar disorder work in creative occupations (Kyaga et al. 2011; Tremblay et al. 2010). This would violate MTR to the extent that creativity is associated with better labor market outcomes. While these conceptual threats to the MTR assumption are possible, we have been unsuccessful in finding studies that convincingly document their link to labor market outcomes. Conversely, there are numerous studies documenting average negative associations between poor mental health and labor market outcomes.[14] Note that, at the individual level, channels of poor mental health affecting labor market outcomes, both negatively and positively, can be present—MTR does not rule out the latter—however, MTR implies that the channels negatively affecting labor market outcomes dominate.

A key implication from MTR is that, for example, $E[Y(t_2)|S = t_\ell] \leq E[Y(t_1)|S = t_\ell]$ for any ℓ, given that $t_2 > t_1$. Figure 1 Panel C illustrates how MTR tightens the no-assumption bounds. For $t < t_2$, MTR implies that the conditional mean $E[Y(t_2)|S = t]$ is no greater than $E[Y(t)|S = t]$, which is identified by $E[Y|S = t]$. This decreases the upper bound on $E[Y(t_2)]$ compared to the no-assumption upper bound in Figure 1 Panel A. Further, for $t' > t_2$, MTR implies that the conditional mean $E[Y(t_2)|S = t']$ cannot lie below $E[Y(t')|S = t']$, which is identified by $E[Y|S = t']$. This raises the lower bound on the unconditional mean $E[Y(t_2)]$ when compared to the no-assumption lower bound in Figure 1 Panel A.

Following the above intuition, the MTR bounds on $E[Y(t_2)]$ are given by (Manski, 1997):

$$\begin{gathered}(10)\ Y_{min} * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + E[Y|S > t_2] * P(S > t_2) \\ \leq E[Y(t_2)] \leq \\ E[Y|S < t_2] * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + Y_{max} * P(S > t_2)\end{gathered}$$

The MTR lower (respectively, upper) bound on the ATE $\Delta(t_2, t_1)$ is calculated by subtracting the upper (lower) bound of $E[Y(t_1)]$ from the lower (upper) bound of $E[Y(t_2)]$. Thus, we have

$$\begin{gathered}(11)\ Y_{min} * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + E[Y|S > t_2] * P(S > t_2) \\ - E[Y|S < t_1] * P(S < t_1) - E[Y|S = t_1] * P(S = t_1) - Y_{max} * P(S > t_1) \\ \leq \Delta(t_2, t_1) \leq \\ E[Y|S < t_2] * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + Y_{max} * P(S > t_2) \\ - Y_{min} * P(S < t_1) - E[Y|S = t_1] * P(S = t_1) - E[Y|S > t_1] * P(S > t_1)\end{gathered}$$

Under the non-positive MTR assumption, the upper bound on $\Delta(t_2, t_1)$ is never above zero, because MTR rules out the possibility that worse mental health improves labor market outcomes.

The MTR and MTS assumptions can be combined to provide tighter bounds on the mean potential outcomes. The MTR-MTS bounds on $E[Y(t_2)]$ are a combination of Panel B and Panel C of Figure 1, and are given by (Manski and Pepper, 2000):

$$\begin{gathered}(12)\ E[Y|S = t_2] * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + E[Y|S > t_2] * P(S > t_2) \\ \leq E[Y(t_2)] \leq \\ E[Y|S < t_2] * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + E[Y|S = t_2] * P(S > t_2)\end{gathered}$$

As before, the MTR+MTS lower (upper) bound on the ATE $\Delta(t_2, t_1)$ is calculated by subtracting the upper (lower) bound of $E[Y(t_1)]$ from the lower (upper) bound of $E[Y(t_2)]$, as follows:

$$\begin{gathered}(13)\ E[Y|S = t_2] * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + E[Y|S > t_2] * P(S > t_2) \\ - E[Y|S < t_1] * P(S < t_1) - E[Y|S = t_1] * P(S = t_1) - E[Y|S = t_1] * P(S > t_1) \\ \leq \Delta(t_2, t_1) \leq \\ E[Y|S < t_2] * P(S < t_2) + E[Y|S = t_2] * P(S = t_2) + E[Y|S = t_2] * P(S > t_2) \\ - E[Y|S = t_1] * P(S < t_1) - E[Y|S = t_1] * P(S = t_1) - E[Y|S > t_1] * P(S > t_1)\end{gathered}$$

[14] Those studies have been reviewed in Section 2. Two other studies are Finkelstein et al. (2013) who find that depression reduces the marginal utility of consumption, thus reducing the incentives to work. The findings in Ringdal and Rootjes (2022) are also consistent with this notion.

Note that the MTS and MTR assumptions imposed together yield a testable implication that observed mean labor market outcomes are weakly decreasing in the CES-D score.[15] That is, for any two treatments t_k and t_j, $t_j > t_k$ implies that $E\left[Y\middle|S = t_j\right] \leq E[Y|S = t_k]$. Table 3 shows that average labor market outcomes are indeed decreasing as a function of depressive symptoms. The employment rate (respectively, average earnings) for individuals with no depressive symptoms is 92% ($26,211); in contrast, it is 84% ($17,456) for individuals with mild depressive symptoms and 72% ($11,912) for individuals with severe depressive symptoms.[16] While this is not necessarily evidence that the MTR+MTS assumptions jointly hold in the sample, it is consistent with their testable implication. In other words, the data does not reject the implication—in terms of the observable data—that stems from the combination of the two assumptions.

4.3 Monotone Instrumental Variable (MIV)

The MTR+MTS bounds can be further narrowed by using a monotone instrumental variable (MIV), which is a variable that has a monotone (weakly increasing or weakly decreasing) mean relationship with the potential outcomes $Y(t)$. The MIV assumption is weaker than the exclusion restriction in IV models, which requires the instrument to be mean independent of the outcome. The MIV assumption also does not require that the variable has a causal effect on the outcome. Specifically, a weakly increasing MIV Z satisfies:

$$(14)\quad m_1 \leq m \leq m_2 \;\Rightarrow\; E[Y(t)|Z = m_1] \leq E[Y(t)|Z = m] \leq E[Y(t)|Z = m_2]$$

for all treatment levels $t \in T$.

We use adolescent AFQT test scores as an MIV. Many studies (e.g., Lang and Manove 2011; Mansour and McKinnish 2014; Neal and Johnson 1996) have used the AFQT score as a proxy for cognitive ability. Using this measure of cognitive skills or ability as an MIV is easy to justify based on economic models of human capital (e.g., Ben-Porath 1967). These models imply that higher innate ability is related to higher labor market outcomes both directly and indirectly through education. Those models, as well as our MIV assumption, are consistent with the well-documented positive relationship between the AFQT score (and, in general, adolescent cognitive ability) and better labor market outcomes, as well as on genetic correlations between intelligence and household income.[17,18]

With a variable Z satisfying the MIV assumption, we can divide the sample into bins defined by the values of Z and compute the MTR+MTS bounds within each bin. In our case of a non-negative MIV, equation (14) implies that the lower bound on $E[Y(t_2)|Z = m]$ is no lower than

[15] In this context, if the testable implication is statistically rejected in the data, then this refutes the corresponding assumptions. If the data does not refute the assumptions in this way, then it is said that the data is only consistent with the assumptions. The latter evidence can be interpreted as making the assumptions more plausible relative to its statistical rejection, or if the testable implication were absent.

[16] Online Appendix Table A2 shows that mean employment and earnings are also decreasing in the CES-D score for men, women, white and Black individuals. Hispanic individuals with moderate depressive symptoms have a higher employment rate (78%) on average than individuals with mild depressive symptoms (77%), but the difference is not statistically significantly different from zero, and thus it is statistically consistent with the weak inequality in the testable implication.

[17] Using the UK Biobank, Hill et al. (2019) find a genetic correlation of 0.69 between intelligence and income.

[18] One can think of the relationship in equation (14) not being satisfied for certain individuals. For instance, a certain individual with a very high AFQT score may hold more stressful jobs and, as a result, tends to quit his/her jobs more often and have lower earnings. Or a certain individual might have a higher taste for leisure and thus lower latent employment and earnings. While these may be possible at the individual level, the fact that the MIV assumption need only hold at the mean implies that the proportion of individuals of this type would have to be implausibly high to reverse the direction of the weakly monotonic mean relationship imposed in the MIV assumption.

the lower bound on $E[Y(t_2)|Z=m_1]$, and its upper bound is no higher than the upper bound on $E[Y(t_2)|Z=m_2]$. For the bin where Z has a value of m, we can thus obtain a new lower bound by taking the largest lower bound over all bins where $Z \leq m$. Likewise, we can obtain a new upper bound by taking the smallest upper bound over all bins where $Z \geq m$. The MIV+MTR+MTS bounds are then obtained by taking the weighted average over all the conditional bounds (which follows from the law of iterated expectations):

$$(15) \sum_{m \in M} P(Z=m) * \left[max_{m_1 \leq m}\, LB_{E\left[Y(t_2)\middle|Z=m_1\right]}\right]$$
$$\leq E[Y(t_2)] \leq$$
$$\sum_{m \in M} P(Z=m)\ * \left[min_{m_2 \geq m}\, UB_{E\left[Y(t_2)\middle|Z=m_2\right]}\right]$$

where LB denotes the MTR+MTS lower bound from equation (12) on $E[Y(t_2)]$ at values $Z=m_1$ of the MIV. Likewise, UB represents the MTR+MTS upper bound on $E[Y(t_2)]$ conditional on values $Z=m_2$ of the MIV. The MIV+MTR+MTS lower (upper) bound on the ATE $\Delta\,(t_2,t_1)$ is calculated once again by subtracting the upper (lower) bound of $E[Y(t_1)]$ from the lower (upper) bound of $E[Y(t_2)]$:

$$(16) \sum_{m \in M} P(Z=m) * \left[max_{m_1 \leq m}\, LB_{E\left[Y(t_2)\middle|Z=m_1\right]}\right]$$
$$-\sum_{m \in M} P(Z=m) * \left[min_{m_2 \geq m}\, UB_{E\left[Y(t_1)\middle|Z=m_2\right]}\right]$$
$$\leq \Delta\,(t_2,t_1) \leq$$
$$\sum_{m \in M} P(Z=m)\ * \left[min_{m_2 \geq m}\, UB_{E\left[Y(t_2)\middle|Z=m_2\right]}\right]$$
$$-\sum_{m \in M} P(Z=m) * \left[max_{m_1 \leq m}\, LB_{E\left[Y(t_1)\middle|Z=m_1\right]}\right]$$

4.4 Estimation and Inference

The no-assumption, MTS, MTR, and MTR+MTS bounds are all estimated by plugging in sample analogues for the expectations and probabilities in the corresponding bounds' expressions. Inference is undertaken by constructing Imbens and Manski (2004) confidence intervals. Estimation and inference under the MIV+MTR+MTS bounds require that we deal with two issues. The first is that the plug-in estimators of Equation (15)—an example of so-called intersection bounds—suffer from bias in finite samples that makes them narrower relative to the corresponding true identified set. The bias then carries over to estimated bounds on the average treatment effects of interest in (16). The second, related issue is that the corresponding confidence intervals do not have the expected coverage at the desired level. Both of these issues arise because of the non-concavity and non-convexity, respectively, of the min and max operators.

We address both issues in the bounds involving the MIV assumption by employing the estimation and valid-inference procedure in Chernozhukov et al. (2013; hereafter, CLR) for intersection bounds, as implemented in the Stata package by Germinario et al. (2022).[19] The CLR procedure allows us to obtain lower and upper bound estimators that satisfy a half-median unbiasedness property; that is, the estimated lower/upper bound will fall below/above the true

[19] See also Flores and Flores-Lagunes (2013) for additional discussion on the CLR procedure.

lower/upper bound with a probability of at least one-half asymptotically. This property is important because Hirano and Porter (2012) showed that there exist no locally asymptotically unbiased estimators of parameters that contain min and max operators, implying that methods aimed at reducing bias (such as the bootstrap) cannot completely eliminate it and reducing bias too much eventually leads the variance of such methods to increase significantly. The details on our implementation of the CLR procedure can be found in Online Appendix B.

5. Results

5.1 Effect of Depression on Employment and Earnings

Table 4 provides results for the effect of being depressed on employment and earnings. The OLS estimates in column 1 indicate that depressed individuals are almost 11 percentage points less likely to be employed, and their earnings are $7,931 less than non-depressed individuals. OLS estimates are unlikely to be causal because of unobserved factors that are correlated with mental health and labor market outcomes. Columns 2-6 provide bounds and 95% confidence intervals on the ATE of being depressed using the nonparametric partial identification procedure described in Section 3. The no-assumption bounds in column 2 are wide, which is typical of this kind of bounds. They indicate that the true causal effect of depression on employment ranges from -77 to 23 percentage points, and from -$34,516 to $66,441 for earnings. Still, they rule out, without assumptions beyond bounded support of the outcome, detrimental effects from depression that are larger than 77 percentage points and $34,516 for employment and earnings, respectively. Adding the MTS assumption—that depressed individuals have lower average potential labor market outcomes than individuals who are not depressed—substantially increases the lower bounds. The MTS bounds in column 3 indicate that the effect of depression is to at most reduce employment by 11 percentage points and earnings by $7,935. The MTS bounds are still wide and null effects as well as relatively large positive effects cannot be ruled out. Column 4 shows bounds under the MTR assumption, which states that worse mental health does not improve labor market outcomes. This means that the upper bound under the MTR assumption is zero, while the lower bound is the same as in the no-assumption bounds. The combination of the MTR and MTS assumptions in column 5 provides considerably tighter bounds. The bounds for employment and earnings are [-0.11, 0.00] and [-$7,935, $0.00], respectively. Adding the MIV assumption to the MTR+MTS bounds further narrows the bounds. The MIV+MTR+MTS bounds in column 6 indicate that being depressed decreases employment by at most 9 percentage points and earnings by at most $6,082. These estimates correspond to effects of at most 10% for employment and 27% for earnings, relative to average employment and earnings for individuals who are not depressed.[20]

The true causal effect of depression on employment and earnings could be any value within the MIV+MTR+MTS bounds. Thus, the estimated bounds are consistent with null, small, and relatively large effects. The OLS estimates of the effect on both employment and earnings are ruled out by the estimated MIV+MTR+MTS bounds, though the OLS estimate is only statistically ruled out for earnings, as shown by the 95% confidence intervals. This is consistent with OLS having a statistically significant downward bias. Note, however, that it is not possible to ascertain whether the OLS estimates substantially overstate the true causal effect in general, given the range of values included in the estimated bounds. Finally, we note that the estimated bounds are also consistent with both relatively large effects (reported in IV studies) and small effects (reported in studies using individual fixed effects) of depression. However, it is important to keep in mind that in the presence of heterogeneous treatment effects, IV and individual fixed-effect methods estimate a parameter that is different from the ATE, which is the one we are bounding.

[20] Average employment and earnings for non-depressed individuals are 0.90 and $22,830, respectively.

5.2 Effects of Levels of Depressive Symptoms on Employment and Earnings

Figure 2 shows the MIV+MTR+MTS bounds, 95% confidence intervals and OLS estimates for the effect of worsening mental health at different levels of depressive symptoms on employment (full results are in Online Appendix Table A3). The estimated bounds indicate that going from having “no vs little” or “little vs mild” depressive symptoms reduces employment by at most 4%. The bounds are also consistent with null and smaller effects. The width of the bounds for “moderate vs mild” and “severe vs moderate” depressive symptoms is larger, indicating an effect to lower employment by at most 10% and 16%, respectively. The bounds, though, are again consistent with null and smaller effects. The estimated bounds on “moderate vs no”, “moderate vs little”, “severe vs no” and “severe vs little” depressive symptoms all exclude zero and the corresponding OLS estimate. For instance, going from having no (little) to severe depressive symptoms reduces employment by 3-18% (3-16%). However, in all cases, both zero and the corresponding OLS estimate are included in the 95% confidence interval. Lastly, note that it is difficult to make formal comparisons of the effects across different levels of depressive symptoms because the bounds overlap substantially—a common feature when using this methodology.

Figure 3 (Online Appendix Table A4) shows results for earnings. Findings are overall similar to those for employment in that most of the estimated bounds are consistent with null, small and plausibly large effects. Notably though, Figure 3 shows that a null effect is statistically excluded for “severe vs no” and “severe vs little” depressive symptoms. The bounds indicate that the population causal average effect of going from having no (little) to severe depression symptoms decreases earnings by 11-44% (12-36%). That these estimated bounds exclude a zero effect is meaningful as the bounds are obtained under relatively weak assumptions. To get a sense about the magnitude of these estimated effects, Figure 4 provides MIV+MTR+MTS bounds on the effect of education on employment and earnings for our estimation sample (full results are in Online Appendix Table A5). These estimated bounds closely mirror the original application in Manski and Pepper (2000). The MIV+MTR+MTS bounds on high school dropout vs high school graduate/some college do not exclude zero, but indicate that going from being a high school dropout to high school graduate/some college at most increases employment by 15% and earnings by 87%. The maximum estimated reduction in earnings from going from no/little to severe depressive symptoms is about half the size of the maximum estimated increase in earnings from going from being a high school dropout to high school graduate/some college. While this is a comparison of the upper bounds on the effects, it suggests that the effects of no/little to severe depression symptoms on earnings are non-negligible. Another point of comparison for the magnitude of our estimated effects is to consider traditional estimates of the returns to schooling. In a survey by Card (1999) of IV estimates of the return on earnings to one year of schooling, he finds that they range from about 6-15%. Our inference on the effect of no (little) to severe depressive symptoms on earnings of 11-44% (12-36%) implies that this effect is *at least* comparable to the earnings effect of an additional year of schooling, informally reinforcing that the magnitude of our estimated effects is not negligible (as before, we point out that the IV estimates surveyed in Card (1999) apply to a subpopulation defined by each of the instruments used in those studies).

Finally, our finding that there are negative effects of going from no/little to severe depressive symptoms on earnings appears robust to different classifications of depressive symptoms. In particular, we also conducted an analysis defining individuals as having no, mild (1≤CES-D≤15), moderate (10≤CES-D≤21) and severe (22≤CES-D≤60) depressive symptoms (Zhang et al. 2005). We find that going from no to severe depressive symptoms reduces employment by 2-14% and earnings by 6-40%, although a zero effect is not statistically excluded (see Online Appendix Table A18). That our results appear robust to different classifications of depressive symptoms is relevant given that there could be concerns of misclassification when the

CES-D is used to explore the severity of depression.[21] Misclassification may occur, for instance, since mental illness is a multifaceted disorder that may manifest in different ways for different people. Thus, the apparent robustness of results suggests that potential misclassification is not driving them.[22]

5.3 Results by Gender, and by Race/Ethnicity

In this section, we further analyze the effects of interest across different demographic groups. For ease of exposition, we focus on the effect of depression, "no vs moderate", "little vs moderate", "no vs severe" and "little vs severe" depressive symptoms. The results are summarized in Figures 5 and 6 (the full set of results are given in Online Appendix Tables A6-A17). It is difficult to make formal comparisons about the magnitude of the effects across groups because the bounds overlap. Still, these results lead to some interesting insights. First, in terms of the effects on employment (Figure 5), the estimated lower bounds for Black individuals are consistently smaller than those of other groups. Moreover, the estimated bounds on the effects of "no vs moderate" and "little vs moderate" depressive symptoms on employment for Black individuals statistically exclude a null effect. These bounds on "no vs moderate" ("little vs moderate") indicate detrimental effects of between 4-17% (4-11%). This is noteworthy given that this subgroup analysis entails a smaller number of observations per group. The other group whose estimated upper bounds often exclude a null effect for employment is that of men; however, the 95% confidence intervals include zero.

For earnings, Figure 6 shows that several of the estimated bounds exclude a null effect, all of them when looking at different depressive symptoms. However, only four sets of estimated bounds rule out zero when the sampling variability is taken into account. The estimated bounds for white individuals statistically rule out zero for "no vs severe" (effect bounded between 12% and 41%) and "little vs severe" (effect bounded between 13% and 35%) depressive symptoms. The other two sets of bounds statistically ruling out zero occur when considering the effect of "no vs moderate" depressive symptoms for men (effect bounded between 6% and 36%) and for Hispanic individuals (effect bounded between 11% and 46%).[23] Again, these statistically significant results are notable given the smaller sample sizes involved.

6. Summary

Credibly identifying the causal effect of poor mental health on labor market outcomes is challenging because of omitted variable bias and reverse causality. To address these issues, previous literature has used childhood and parental mental health, religiosity, death of a close friend, availability of social services and social support as instruments, but these instruments may not satisfy the exclusion restriction. Other studies have used individual fixed effects, which control for unobserved time-invariant characteristics, but they are unable to control for time-varying confounders, such as stress. In addition, in the presence of heterogeneous effects, IV and individual fixed-effect methods estimate effects for specific subpopulations that may or may not correspond to the population of interest. We contribute to the literature by using a nonparametric

[21] In addition to the two classifications discussed here, an earlier version of this manuscript employed an alternative classification based on 3 categories of depressive symptoms (no-to-mild with CES-D between 0-15, moderate with CES-D between 16-23 and severe with CES-D between 24-60), from which fairly similar results were obtained.

[22] If misclassification exists, it is also worth considering how it may impact the plausibility of our identification assumptions. We believe that, for reasonable levels of misclassification, the MTS and MIV assumptions would remain plausible as they are imposed at the mean level. Conversely, if some individuals with better labor market potential outcomes are misclassified as having more depressive symptoms, this could reduce the plausibility of the individual-level MTR.

[23] Another set of estimated bounds that statistically exclude a null effect but is not shown in Figure 6 is the "no vs mild" depressive symptoms for men where the effect bounded between 0.5% and 27%.

partial identification method to provide bounds on the population average effect of depression on employment and earnings. Although we do not point identify the causal effect, our approach uses relatively weak assumptions, provides bounds on the population average treatment effect, and allows us to analyze the effect for different levels of depressive symptoms. The assumptions are based on weak monotonicity relationships on the selection process (MTS), the treatment response (MTR), and the link between a variable (referred to as a monotone instrument) and the outcome (MIV).

When looking at a binary indicator for depression, our results under the MTS, MTR and MIV assumptions indicate that the effect of depression is at most 10% on employment and 27% on earnings, thus ruling out potentially plausible magnitudes for these effects. When considering the dose-response nature of depression symptoms, we find statistically significant and economically non-negligible effects of going from no and little to severe depression symptoms on earnings. Going from no (little) to severe depressive symptoms decreases earnings by 11-44% (12-36%). Interestingly, these bounds are somewhat comparable in magnitude (in absolute value) to corresponding bounds obtained for going from being a high school dropout to high school graduate/some college on earnings. Moreover, our estimated bounds rule out, for the population average effect, the magnitude of some of the effects obtained in prior studies, particularly very large detrimental effects on employment, and a zero effect on earnings (at least for going from no and little to severe depressive symptoms). Therefore, our findings, obtained using relatively weak assumptions, point to potentially significant detrimental causal effects of depression symptoms on earnings in the United States.

References

Alexandre, P. K., & French, M. T. (2001). Labor supply of poor residents in metropolitan Miami, Florida: The role of depression and the co-morbid effects of substance use. *Journal of Mental Health Policy and Economics*, *4*(4), 161-174.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy*, *113*(1), 151-184.

Altonji, J.G., Bharadwaj, P., & Lange, F. (2012). Changes in the characteristics of American youth: Implications for adult outcomes. *Journal of Labor Economics*, 30, 783-828.

Bartel, A., & Taubman, P. (1979). Health and labor market success: The role of various diseases. *The Review of Economics and Statistics*, 1-8.

Bartel, A., & Taubman, P. (1986). Some economic and demographic consequences of mental illness. *Journal of Labor Economics*, *4*(2), 243-256.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy*, *75*(4, Part 1), 352-365.

Biasi, B., Dahl, M.S. & Moser, P. (2021). Career effects of mental health. NBER Working Paper Number 29031.

Bryan, M. L., Rice, N., Roberts, J., & Sechel, C. (2022). Mental health and employment: a bounding approach using panel data. *Oxford Bulletin of Economics and Statistics,* 0305-9049.

Card, D. (1999). The causal effect of education on earnings. *Handbook of Labor Economics*, *3*, 1801-1863.

Chernozhukov, V., Lee, S., & Rosen, A. M. (2013). Intersection bounds: estimation and inference. *Econometrica*, *81*(2), 667-737.

Chatterji, P., Alegria, M., Lu, M., & Takeuchi, D. (2007). Psychiatric disorders and labor market outcomes: evidence from the National Latino and Asian American Study. *Health Economics*, *16*(10), 1069-1090.

Chatterji, P., Alegria, M., & Takeuchi, D. (2008). Psychiatric disorders and labor market outcomes: Evidence from the National Comorbidity Survey-Replication. NBER Working Paper Number 14404.

Chatterji, P., Alegria, M., & Takeuchi, D. (2011). Psychiatric disorders and labor market outcomes: Evidence from the National Comorbidity Survey-Replication. *Journal of Health Economics*, *30*(5), 858-868.

Christelis, D., & Dobrescu, L. (2019). The causal effect of social activities on cognition: Evidence from 20 European countries. *Social Science & Medicine, 247,* 1-9

Cosco, T. D., Prina, M., Stubbs, B., & Wu, Y. T. (2017). Reliability and validity of the Center for Epidemiologic Studies Depression Scale in a population-based cohort of middle-aged US adults. *Journal of Nursing Measurement*, *25*(3), 476-485.

Cseh, A. (2008). The effects of depressive symptoms on earnings. *Southern Economic Journal*, 383-409.

Cygan-Rehm, K., Kuehnle, D., & Oberfichtner, M. (2017). Bounding the causal effect of unemployment on mental health: Nonparametric evidence from four countries. *Health Economics*, *26*(12), 1844-1861.

De Haan, M. (2011). The effect of parents' schooling on child's schooling: a nonparametric bounds analysis. *Journal of Labor Economics*, *29*(4), 859-892.

Engel, M., Peternel, M., Jockel, K-H., Schramm, S, Slomiany, U., & Moebus, S. (2020). Sex-specific longitudinal comparison of CES-D and PHQ-9 depression scales. A concordance analysis using data from the population-based Heinz Nixdorf recall study. Research Square pre-print, https://doi.org/10.21203/rs.2.22423/v1.

Ettner, S. L., Frank, R. G., & Kessler, R. C. (1997). The impact of psychiatric disorders on labor market outcomes. *ILR Review*, *51*(1), 64-81.

Farahati, F., Marcotte, D. E., & Wilcox-Gök, V. (2003). The effects of parents' psychiatric disorders on children's high school dropout. *Economics of Education Review*, *22*(2), 167-178.

Finkelstein, A., Luttmer, E.F. & Notowidigdo M. J. (2013). What good is wealth without health? The effect of health on the marginal utility of consumption. *Journal of the European Economic Association*, 11(S1):221–258.

Fletcher, J. M. (2008). Adolescent depression: diagnosis, treatment, and educational attainment. *Health Economics*, *17*(11), 1215-1235.

Fletcher, J. (2013a). Adolescent depression and adult labor market outcomes. *Southern Economic Journal*, *80*(1), 26-49.

Fletcher, J. M. (2013b). The effects of personality traits on adult labor market outcomes: Evidence from siblings. *Journal of Economic Behavior & Organization*, *89*, 122-135.

Flores, C. A., & Flores-Lagunes, A. (2013). Partial identification of local average treatment effects with an invalid instrument. *Journal of Business & Economic Statistics*, *31*(4), 534-545.

Frank, R., & Gertler, P. (1991). An assessment of measurement error bias for estimating the effect of mental distress on income. *The Journal of Human Resources*, *26*(1), 154-164.

Frijters, P., Johnston, D. W., & Shields, M. A. (2014). The effect of mental health on employment: evidence from Australian panel data. *Health Economics*, *23*(9), 1058-1071.

Giorgetta, C., Grecucci, A., Zuanon, S., Perini, L., Balestrieri, M., Bonini, N., ... & Brambilla, P. (2012). Reduced risk-taking behavior as a trait feature of anxiety. *Emotion*, *12*(6), 1373.

Germinario, G., Flores, C.A., & Flores-Lagunes, A. (2022). mpclr: A Stata command for bias-corrected estimation and valid inference for nonparametric bounds on treatment effects. Mimeographed: Department of Economics, Syracuse University.

Gonzalez, L. (2005). Nonparametric bounds on the returns to language skills. *Journal of Applied Econometrics*, *20*(6), 771-795.

Grossman, M. (1972): On the concept of health capital and the demand for health. *Journal of Political Economy,* 80(2), 223–255.

Gruber, J. H. (2005). Religious market structure, religious participation, and outcomes: Is religion good for you?. *The BE Journal of Economic Analysis & Policy*, *5*(1).

Hill, W. D., Davies, N. M., Ritchie, S. J., Skene, N. G., Bryois, J., Bell, S., ... & Deary, I. J. (2019). Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. *Nature Communications*, *10*(1), 1-16.

Hirano, K., & Porter, J. R. (2012). Impossibility results for nondifferentiable functionals. *Econometrica*, *80*(4), 1769-1790.

Huang, J., van den Brink, H. M., & Groot, W. (2012). Does education promote social capital? Evidence from IV analysis and nonparametric-bound analysis. *Empirical Economics*, *42*(3), 1011-1034.

Imbens, G. W., & Manski, C. F. (2004). Confidence intervals for partially identified parameters. *Econometrica*, *72*(6), 1845-1857.

Koford, B.C., & Cseh, A. (2015). Older Americans, depression, and labor market outcomes. *Eastern Economic Journal*, *41*, 70-85.

Kyaga, S., Lichtenstein, P., Boman, M., Hultman, C., Långström, N., & Landen, M. (2011). Creativity and mental disorder: family study of 300 000 people with severe mental disorder. *The British Journal of Psychiatry*, *199*(5), 373-379.

Lang, K., & Manove, M. (2011). Education and labor market discrimination. *American Economic Review*, *101*(4), 1467-96.

Lin, D., Lutter, R., & Ruhm, C. J. (2018). Cognitive performance and labour market outcomes. *Labour Economics*, *51*, 121-135.

Lu, W., Lindsey, M.A., Irsheid, S. & Nebbitt, V. E. (2017). Psychometric properties of the CES-D among black adolescents in public housing. *Journal of the Society for Social Work and Research*, 8(4), 595-619.

Lund, C., Orkin, K., Witte, M., Davies, T., Haushofer, J., Bass, J., ... & Patel, V. (2018). *The economic effects of mental health interventions in low and middle-income countries*. Working paper.

Lundborg, P., Nilsson, A., & Rooth, D. O. (2014). Adolescent health and adult labor market outcomes. *Journal of Health Economics*, *37*, 25-40.

Maner, J. K., Richey, J. A., Cromer, K., Mallott, M., Lejuez, C. W., Joiner, T. E., & Schmidt, N. B. (2007). Dispositional anxiety and risk-avoidant decision-making. *Personality and Individual Differences*, *42*(4), 665-675.

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

Manski, C. F. (1997). Monotone treatment response. *Econometrica*, 65, 1311-1334.

Manski, C. F., & Pepper, J. V. (2000). Monotone instrumental variables with an application to the returns to schooling. *Econometrica,* 68(July), 997-1010

Mansour, H., & McKinnish, T. (2014). Who marries differently aged spouses? Ability, education, occupation, earnings, and appearance. *Review of Economics and Statistics*, *96*(3), 577-580.

Marcotte, D. E., & Wilcox-Gök, V. (2001). Estimating the employment and earnings costs of mental illness: recent developments in the United States. *Social Science & Medicine*, *53*(1), 21-27.

Marcotte, D. E., & Wilcox-Gok, V. (2003). Estimating earning losses due to mental illness: a quantile regression approach. *Journal of Mental Health Policy and Economics*, *6*(3), 123-134.

Milette, K, Hudson, M., Baron, M., Thombs, B.D. & Canadian Scleroderma Research Group (2010). Comparison of the PHQ-9 and CESD-D depression scales in systemic sclerosis: Internal consistency reliability, convergent validity and clinical correlates. *Rheumatology*, *49*, 789-796.

Miller, L. S., & Kelman, S. (1992). Estimates of the loss of individual productivity from alcohol and drug abuse and from mental illness. In Richard G. Frank and Willard G. Manning, Jr., eds., *Economics and Mental Health*. Baltimore: Johns Hopkins University Press, pp. 91-129.

Mitchell, J. M., & Anderson, K. H. (1989). Mental health and the labor force participation of older workers. *Inquiry*, 262-271.

Mousteri, V., Daly, M., Delaney, L., Tynelius, P., & Rasmussen, F. (2019). Adolescent mental health and unemployment over the lifespan: Population evidence from Sweden. *Social Science & Medicine*, *222*, 305-314.

Mukamal, K.J. (2015). Invited commentary: Fitness and fatness—causes of depression or of misclassification. *American Journal of Epidemiology*, 181(5), 321-324.

Neal, D. A., & Johnson, W. R. (1996). The role of premarket factors in black-white wage differences. *Journal of Political Economy*, *104*(5), 869-895.

Nimrod, G., Kleiber, D. A., & Berdychevsky, L. (2012). Leisure in coping with depression. *Journal of Leisure Research*, *44*(4), 419-449.

Ojeda, V. D., Frank, R. G., McGuire, T. G., & Gilmer, T. P. (2010). Mental illness, nativity, gender and labor supply. *Health Economics*, *19*(4), 396-421.

Patel, V., Weobong, B., Weiss, H. A., Anand, A., Bhat, B., Katti, B., ... & Fairburn, C. G. (2017). The Healthy Activity Program (HAP), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in India: a randomised controlled trial. *The Lancet*, *389*(10065), 176-185.

Peng, L., Meyerhoefer, C. D., & Zuvekas, S. H. (2016). The Short-Term Effect of Depressive Symptoms on Labor Market Outcomes. *Health Economics*, *25*(10), 1223-1238.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385-401.

Richey, J. (2015). Shackled labor markets: Bounding the causal effects of criminal convictions in the US. *International Review of Law and Economics*, *41*, 17-24.

Ringdal, C. & Rootjes, F. (2022). Depression and labor supply: Evidence from the Netherlands. *Economics and Human Biology*, *45*. https://doi.org/10.1016/j.ehb.2021.101103.

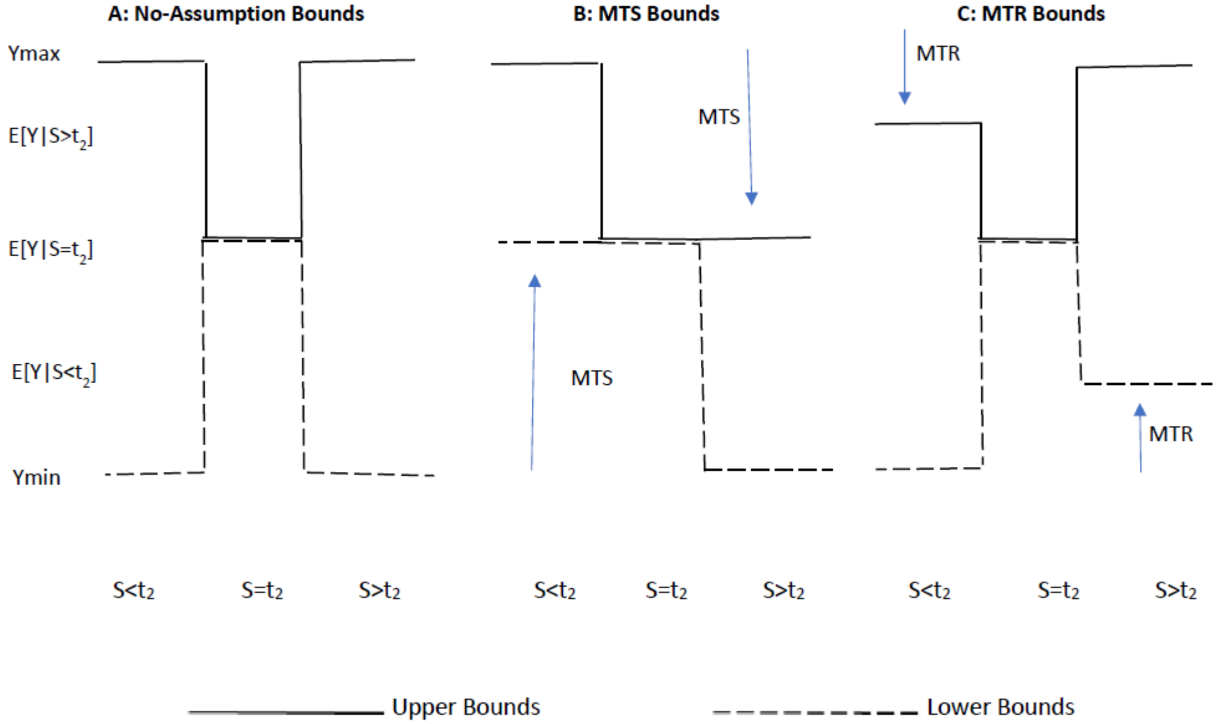
Tremblay, C. H., Grosskopf, S., & Yang, K. (2010). Brainstorm: Occupational choice, bipolar illness and creativity. *Economics & Human Biology*, *8*(2), 233-241.

Unützer, J., Patrick, D. L., Marmon, T., Simon, G. E., & Katon, W. J. (2002). Depressive symptoms and mortality in a prospective study of 2,558 older adults. *The American Journal of Geriatric Psychiatry*, *10*(5), 521-530.

Zhang, X., Norris, S. L., Gregg, E. W., Cheng, Y. J., Beckles, G., & Kahn, H. S. (2005). Depressive symptoms and mortality among persons with and without diabetes. *American Journal of Epidemiology*, *161*(7), 652-660.

Zhang, Y., Ting, R.Z.W., Lam, M.H.B., et al. (2015). Measuring depression with CES-D in Chinese patients with type 2 diabetes: The validity and its comparison to PHQ-9. *BMC Psychiatry*, *15*:198, 1-10.

Figure 1: Illustration of No-Assumption, MTS, and MTR Bounds for E[Y(t_2)]



Notes: Each panel illustrates graphically the construction of nonparametric bounds on $E[Y(t_2)]$ under a given assumption, drawing form De Haan (2011). The x-axis represents levels of the realized depressive symptoms (S) while the y-axis represents values of the outcome (Y). Panel A illustrates the construction of the no-assumption bounds where the maximum (Y_{max}) and minimum (Y_{min}) of the outcome are used to replace the missing counterfactuals for those with S<t_2 and S >t_2. In panel B, the MTS assumption tightens the lower and upper bounds relative to the no-assumption bounds by using the observed mean outcomes for those receiving t_2 (E[Y| S=t_2]) to replace the missing counterfactuals for those with S<t_2 and S>t_2. In panel C, the MTR assumption reduces the upper bound for those with S<t_2 by using the observed mean outcome for those with S>t_2 (E[Y|S>t_2]). Correspondingly, the MTR assumption increases the lower bound for those with S>t_2 by using the observed mean outcome for those with S<t_2 (E[Y|S<t_2]). The combination of panel B and panel C would correspond to the case when the MTS and MTR assumptions are imposed together.

Table 1: Descriptive Statistics

	Full	Men	Women	White	Black	Hispanic
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.51 (0.50)	—	—	0.50 (0.50)	0.52 (0.50)	0.50 (0.50)
White	0.80 (0.40)	0.80 (0.40)	0.80 (0.40)	—	—	—
Black	0.14 (0.35)	0.14 (0.34)	0.14 (0.35)	—	—	—
Hispanic	0.06 (0.24)	0.06 (0.24)	0.06 (0.24)	—	—	—
AFTQ (1981)	47.79 (28.85)	48.29 (30.03)	47.31 (27.65)	53.52 (27.44)	22.41 (20.62)	30.49 (24.67)
Age (1992)	31.06 (2.31)	31.03 (2.33)	31.07 (2.31)	31.07 (2.31)	31.01 (2.33)	30.93 (2.35)
CES-D score (1992)	9.04 (9.00)	8.07 (8.15)	9.99 (9.67)	8.49 (8.76)	11.42 (9.48)	10.82 (9.89)
Depressed	0.18 (0.38)	0.15 (0.35)	0.22 (0.42)	0.16 (0.37)	0.28 (0.45)	0.24 (0.43)
No depressive symptoms	0.10 (0.30)	0.11 (0.32)	0.09 (0.28)	0.11 (0.31)	0.05 (0.22)	0.09 (0.28)
Little depressive symptoms	0.72 (0.48)	0.74 (0.44)	0.69 (0.46)	0.73 (0.45)	0.67 (0.47)	0.67 (0.47)
Mild depressive symptoms	0.08 (0.27)	0.06 (0.24)	0.09 (0.28)	0.07 (0.26)	0.11 (0.31)	0.08 (0.26)
Moderate depressive symptoms	0.04 (0.21)	0.03 (0.18)	0.05 (0.23)	0.04 (0.19)	0.08 (0.27)	0.07 (0.25)
Severe depressive symptoms	0.06 (0.25)	0.05 (0.21)	0.08 (0.27)	0.06 (0.23)	0.09 (0.28)	0.10 (0.30)
Employed (1993)	0.88 (0.33)	0.94 (0.24)	0.82 (0.39)	0.89 (0.31)	0.80 (0.40)	0.84 (0.36)
Earnings (1993)	21,365 (19,862)	27,455 (21,575)	15,405 (15,916)	22,769 (20,594)	14,767 (14,823)	17,944 (16,405)
N	7,665	3,678	3,987	3,907	2,291	1,467

Notes: Summary statistics are weighted using the 1993 panel sampling weights. Standard deviations in parenthesis.

Table 2: Observed Family Background, Ability and Education by Depressive Symptom Severity, Full Sample

Variable (Survey Year)	None	Little	Mild	Moderate	Severe
	(1)	(2)	(3)	(4)	(5)
Demographics (1979)					
Female	0.45 (0.07)	0.49 (0.05)	0.59 (0.06)	0.63 (0.06)	0.61 (0.05)
White	0.87 (0.03)	0.81 (0.04)	0.74 (0.06)	0.66 (0.07)	0.72 (0.05)
Black	0.07 (0.02)	0.13 (0.04)	0.20 (0.06)	0.24 (0.07)	0.19 (0.05)
Hispanic	0.06 (0.01)	0.06 (0.01)	0.06 (0.01)	0.10 (0.02)	0.09 (0.02)
Family Background (1979)					
Mother's education	11.94 (0.08)	11.70 (0.08)	11.22 (0.16)	10.77 (0.15)	10.83 (0.25)
Father's education	12.21 (0.19)	11.94 (0.14)	11.57 (0.24)	10.82 (0.31)	11.12 (0.34)
Ability Measures (1981)					
AFQT Score	56.71 (1.53)	49.46 (1.36)	40.52 (2.35)	33.94 (2.26)	33.49 (1.65)
Expect to graduate college	0.47 (0.02)	0.40 (0.01)	0.33 (0.02)	0.25 (0.04)	0.24 (0.03)
Education (1993)					
Years of education	13.81 (0.09)	13.36 (0.06)	12.87 (0.10)	12.43 (0.14)	12.16 (0.14)
College graduate	0.32 (0.01)	0.25 (0.01)	0.19 (0.02)	0.13 (0.02)	0.10 (0.22)

Notes: Summary statistics are weighted by the 1993 panel sampling weights. The sample size for demographics and AFQT score is 7,665. The sample sizes for the other variables are lower due to missing values: (1) mother's education N=7,190; (2) father's education N=6,554; (3) expect to graduate college N=7,610; (4) Education N=7,654 Standard errors in parentheses.

Table 3: Labor Market Outcomes by Depressive Symptom Severity, Full Sample

Depressive Symptoms, t	$E[Employed \mid T = t]$	$E[Earnings \mid T = t]$	$\Pr[T = t]$	N_t
	(1)	**(2)**	**(3)**	**(4)**
None	0.92	26,211	0.09	677
Little	0.90	22,356	0.70	5,346
Mild	0.84	17,456	0.09	659
Moderate	0.80	14,821	0.05	418
Severe	0.72	11,912	0.07	565

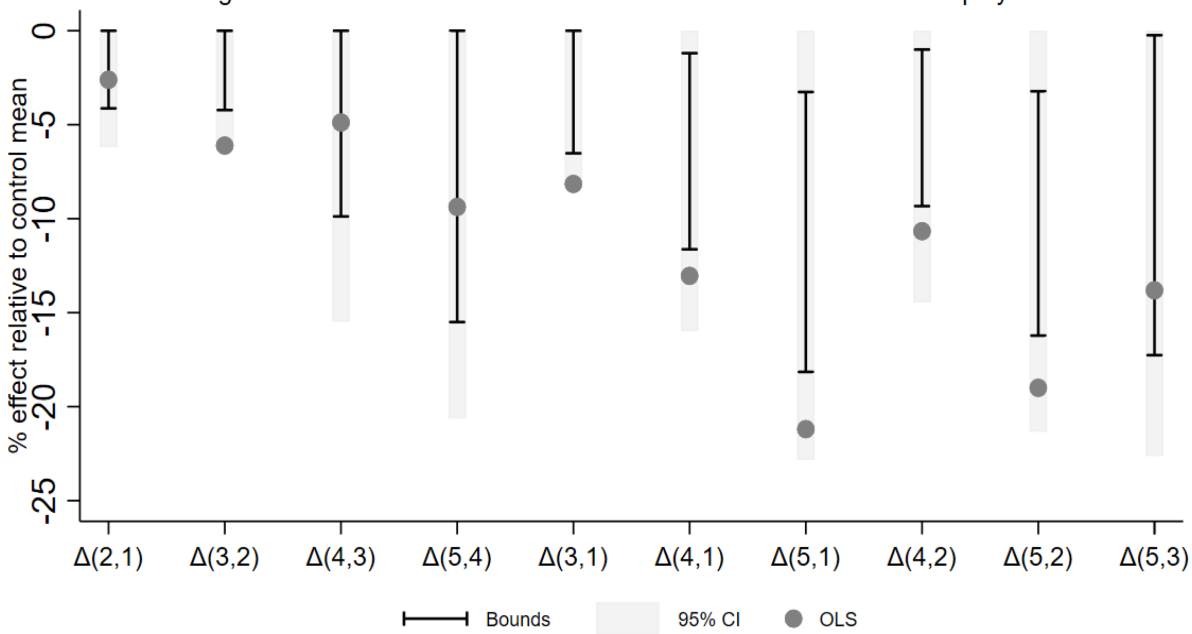
Notes: Summary statistics are weighted by the 1993 panel sampling weights.

Table 4: OLS Estimates and Estimated Bounds on the ATE of being Depressed on Labor Market Outcomes

	OLS	**No Assumption**	**MTS**	**MTR**	**MTR+MTS**	**MIV+MTR+MTS**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Panel A: Employment						
Depressed	-0.108*** (.019)	[-0.771, 0.229] (-0.788, 0.246)	[-0.108, 0.229] (-0.140, 0.246)	[-0.771, 0.000] (-0.788, 0.000)	[-0.108, 0.000] (-0.143, 0.000)	[-0.091, 0.000] (-0.121, 0.000)
Panel B: Earnings						
Depressed	-7931*** (749)	[-34,516, 66,441] (-35,868, 67,793)	[-7935, 66,441] (-9188, 67,793)	[-34,516, 0.000] (-35,868, 0.000)	[-7935, 0.000] (-9261, 0.000)	[-6082, 0.000] (-7520, 0.000)

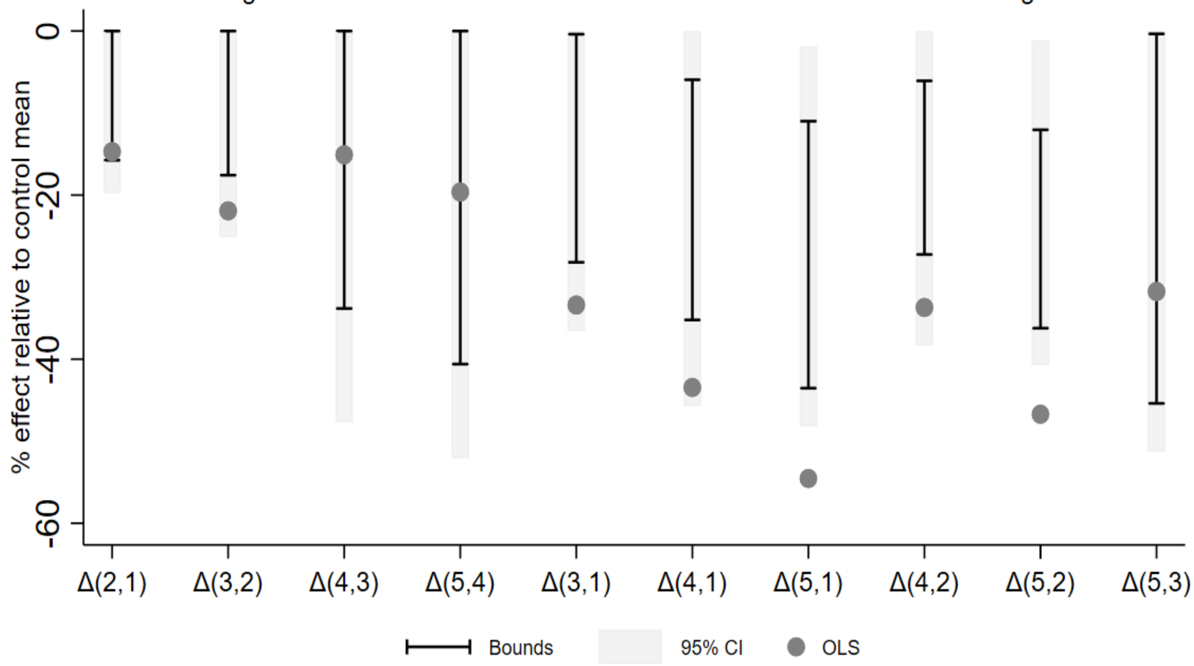
Notes: Robust standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. Earnings are bounded between $0 and $100,948. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds.

Figure 2: MIV+MTR+MTS Bounds on Effect of Mental Health on Employment



Notes: The figure presents the estimated bounds on the effects of different levels of depressive symptoms on employment, which are shown in the x-axis. The y-axis indicates the size of the effects in percentage terms relative to the corresponding "control" mean. The shaded boxes represent the valid 95% confidence intervals on the parameter of interest obtained using the CLR methodology described in the text.
Δ(2,1): Little vs No Depressive Symptoms, Δ(3,2): Mild vs Little Depressive Symptoms
Δ(4,3): Moderate vs Mild Depressive Symptoms Δ(3,1): Mild vs No Depressive Symptoms
Δ(4,1): Moderate vs No Depressive Symptoms, Δ(5,1): Severe vs No Depressive Symptoms
Δ(4,2): Moderate vs Little Depressive Symptoms, Δ(5,2): Severe vs Litte Depressive Symptoms
Δ(5,3): Severe vs Mild Depressive Symptoms

Figure 3: MIV+MTR+MTS Bounds on Effect of Mental Health on Earnings

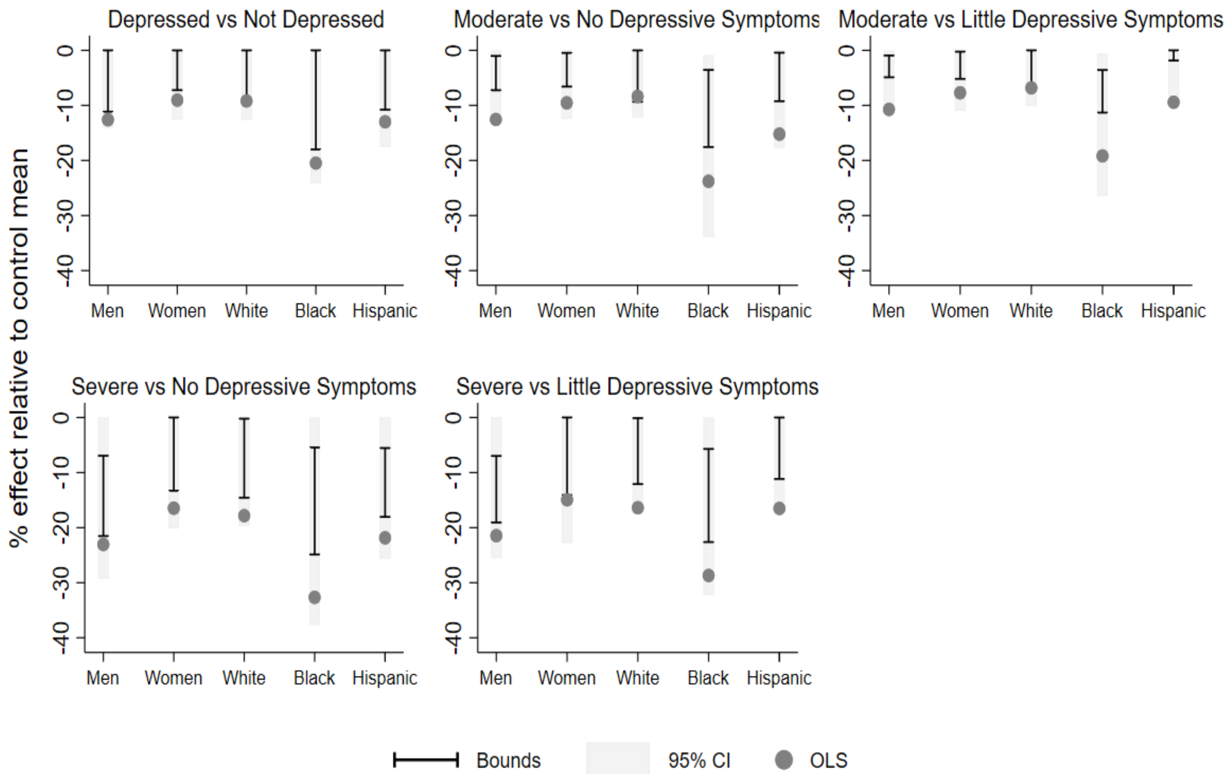


Notes: The figure presents the estimated bounds on the effects of different levels of depressive symptoms on earnings, which are shown in the x-axis. The y-axis indicates the size of the effects in percentage terms relative to the corresponding "control" mean. The shaded boxes represent the valid 95% confidence intervals on the parameter of interest obtained using the CLR methodology described in the text.
Δ(2,1): Little vs No Depressive Symptoms, Δ(3,2): Mild vs Little Depressive Symptoms
Δ(4,3): Moderate vs Mild Depressive Symptoms Δ(3,1): Mild vs No Depressive Symptoms
Δ(4,1): Moderate vs No Depressive Symptoms, Δ(5,1): Severe vs No Depressive Symptoms
Δ(4,2): Moderate vs Little Depressive Symptoms, Δ(5,2): Severe vs Litte Depressive Symptoms
Δ(5,3): Severe vs Mild Depressive Symptoms

Figure 4: MIV+MTR+MTS Bounds on Effect of Education on Labor Market Outcomes

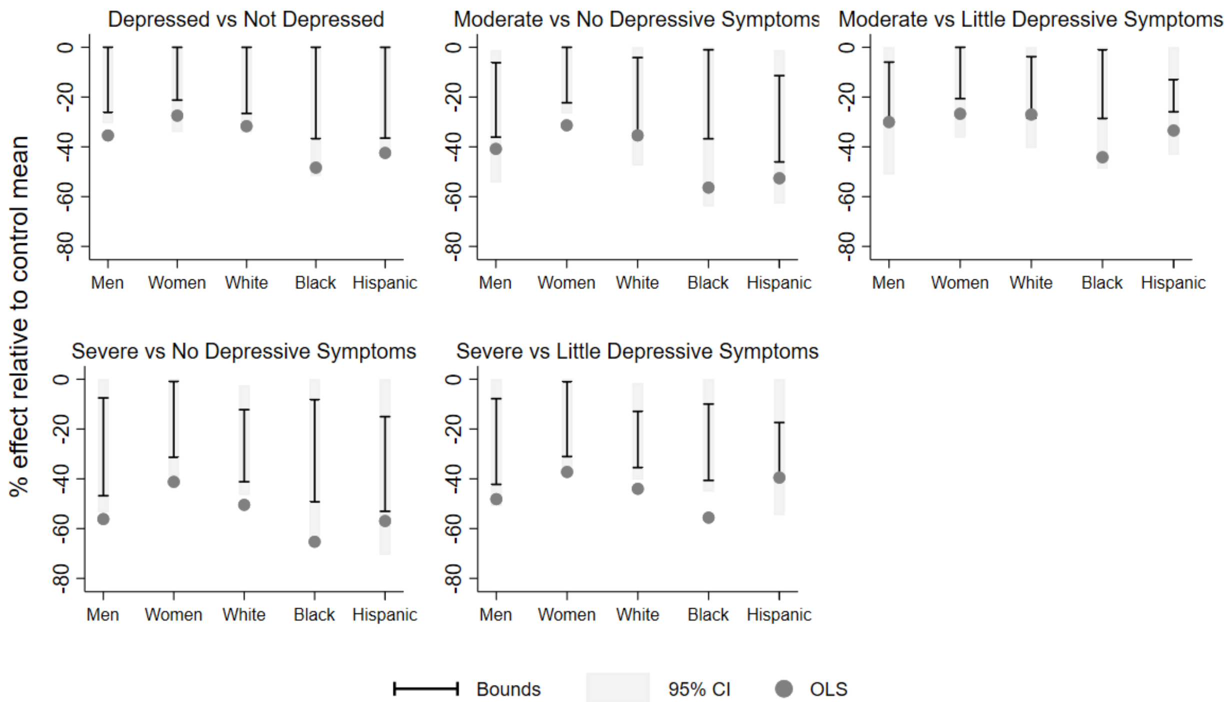
Notes: The figure presents the estimated bounds on the effects of different levels of schooling, which are shown in the x-axis, on employment (upper panel) and earnings (lower panel). The y-axis indicates the size of the effects in percentage terms relative to the corresponding "control" mean. The bands indicate the estimated bounds while the large dots represent the corresponding OLS estimate. The shaded boxes represent the valid 95% confidence intervals on the parameter of interest obtained using the CLR methodology described in the text. Δ(2,1): HS Grad/Some College vs HS Dropout, Δ(3,2): College Grad vs HS Grad/Some College, Δ(3,1): College Grad vs HS Dropout

Figure 5: MIV+MTR+MTS Bounds on Effect of Mental Health on Employment By Gender and Race



Notes: The figure presents the estimated bounds on the effects of different levels of depressive symptoms (shown in the different panels) on employment, by gender and race. The y-axis indicates the size of the effects in percentage terms relative to the corresponding “control” mean. The bands indicate the estimated bounds while the large dots represent the corresponding OLS estimate. The shaded boxes represent the valid 95% confidence intervals on the parameter of interest obtained using the CLR methodology described in the text.

Figure 6: MIV+MTR+MTS Bounds on Effect of Mental Health on Earnings, By Gender and Race



Notes: The figure presents the estimated bounds on the effects of different levels of depressive symptoms (shown in the different panels) on earnings, by gender and race. The y-axis indicates the size of the effects in percentage terms relative to the corresponding "control" mean. The bands indicate the estimated bounds while the large dots represent the corresponding OLS estimate. The shaded boxes represent the valid 95% confidence intervals on the parameter of interest obtained using the CLR methodology described in the text.

ONLINE APPENDIX – NOT INTENDED FOR PUBLICATION

Appendix A: Additional Tables

Table A1: Observed Family Background, Ability and Education by Depressive Symptom Severity, Gender, and Race/Ethnicity

Variable	None	Little	Mild	Moderate	Severe
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Panel A: Men					
Mother’s education	12.03 (0.10)	11.75 (0.13)	11.48 (0.31)	10.98 (0.30)	10.64 (0.47)
Father’s education	12.33 (0.33)	12.01 (0.22)	11.83 (0.44)	10.37 (0.55)	11.16 (0.78)
AFQT Score	58.75 (2.42)	49.20 (1.89)	41.37 (4.07)	30.30 (3.38)	32.33 (2.99)
Expect to graduate college	0.49 (0.03)	0.41 (0.02)	0.31 (0.03)	0.21 (0.03)	0.26 (0.05)
Years of education	13.82 (0.14)	13.26 (0.09)	12.85 (0.16)	12.39 (0.21)	11.88 (0.16)
College graduate	0.33 (0.02)	0.24 (0.01)	0.19 (0.02)	0.12 (0.03)	0.08 (0.01)
Panel B: Women					
Mother’s education	11.84 (0.14)	11.64 (0.11)	11.04 (0.17)	10.64 (0.16)	10.94 (0.29)
Father’s education	12.06 (0.17)	11.85 (0.17)	11.39 (0.28)	11.07 (0.34)	11.10 (0.29)
AFQT Score	54.17 (2.03)	49.74 (1.96)	39.93 (2.80)	36.11 (2.89)	34.23 (1.94)
Expect to graduate college	0.45 (0.02)	0.39 (0.02)	0.35 (0.03)	0.27 (0.05)	0.24 (0.04)
Years of education	13.80 (0.07)	13.46 (0.08)	12.90 (0.14)	12.45 (0.18)	12.34 (0.22)
College graduate	0.29 (0.02)	0.25 (0.01)	0.19 (0.03)	0.14 (0.04)	0.12 (0.04)
Panel C: White Individuals					
Mother’s education	12.18 (0.08)	12.02 (0.09)	11.67 (0.15)	11.36 (0.20)	11.31 (0.35)
Father’s education	12.47 (0.23)	12.34 (0.16)	12.23 (0.22)	11.53 (0.32)	11.54 (0.45)
AFQT Score	60.01 (1.56)	54.69 (1.11)	48.32 (1.60)	42.16 (1.70)	40.07 (2.23)
Expect to graduate college	0.47 (0.02)	0.40 (0.02)	0.35 (0.03)	0.24 (0.05)	0.23 (0.04)
Years of education	13.91 (0.10)	13.49 (0.08)	13.13 (0.12)	12.61 (0.19)	12.33 (0.21)
College graduate	0.33 (0.02)	0.27 (0.01)	0.23 (0.03)	0.16 (0.04)	0.13 (0.03)
Panel D: Black Individuals					
Mother’s education	11.86 (0.25)	11.07 (0.16)	10.71 (0.23)	10.30 (0.17)	10.26 (0.33)
Father’s education	11.32 (0.28)	10.48 (0.22)	9.80 (0.35)	9.59 (0.48)	9.88 (0.33)
AFQT Score	32.67 (1.98)	24.12 (1.48)	18.17 (1.81)	15.99 (2.07)	14.16 (1.08)
Expect to graduate college	0.51 (0.04)	0.41 (0.02)	0.31 (0.04)	0.27 (0.06)	0.28 (0.03)
Years of education	13.40 (0.21)	12.92 (0.07)	12.37 (0.14)	12.19 (0.21)	11.65 (0.09)
College graduate	0.23 (0.03)	0.15 (0.01)	0.10 (0.01)	0.06 (0.03)	0.03 (0.02)

Panel E: Hispanic Individuals					
Mother's education	8.10 (0.73)	8.37 (0.20)	7.42 (0.51)	7.61 (0.56)	8.29 (0.27)
Father's education	8.67 (0.79)	8.58 (0.28)	7.75 (0.35)	7.67 (0.56)	9.57 (0.49)
AFQT Score	36.08 (2.67)	33.00 (1.82)	20.41 (2.19)	21.91 (2.51)	21.82 (1.97)
Expect to graduate college	0.42 (0.03)	0.36 (0.02)	0.20 (0.03)	0.25 (0.05)	0.25 (0.03)
Years of education	12.80 (0.42)	12.47 (0.11)	11.47 (0.23)	11.74 (0.45)	11.87 (0.18)
College graduate	0.22 (0.02)	0.12 (0.02)	0.02 (0.01)	0.11 (0.05)	0.04 (0.02)

Notes: Summary statistics are weighted by the 1993 panel sampling weights. Standard errors in parentheses.

Table A2: Labor Market Outcomes by Depressive Symptom Severity, Gender and Race/Ethnicity

Depressive Symptoms t	$E[Employed \mid T=t]$	$E[Earnings \mid T=t]$	$\Pr[T=t]$	N_t
Panel A: Men				
None	0.98	33,356	0.10	361
Mild	0.96	28,213	0.72	2,651
Little	0.90	22,479	0.08	291
Moderate	0.85	19,742	0.04	159
Severe	0.75	14,624	0.06	216
Panel B: Women				
None	0.85	17,308	0.08	316
Mild	0.83	16,215	0.68	2,695
Little	0.80	13,936	0.09	368
Moderate	0.77	11,885	0.06	259
Severe	0.71	10,177	0.09	349
Panel C: White Individuals				
None	0.92	26,623	0.11	430
Little	0.91	23,556	0.73	2,833
Mild	0.88	19,421	0.07	274
Moderate	0.84	17,194	0.04	144
Severe	0.76	13,187	0.06	222
Panel D: Black Individuals				
None	0.90	20,994	0.05	118
Little	0.84	16,405	0.67	1,531
Mild	0.72	11,714	0.11	262
Moderate	0.68	9,160	0.08	174
Severe	0.60	7,288	0.09	206
Panel E: Hispanic Individuals				
None	0.92	26,520	0.09	129
Little	0.86	18,876	0.67	982
Mild	0.77	12,734	0.08	119
Moderate	0.78	12,566	0.07	100
Severe	0.72	11,420	0.09	137

Notes: Summary statistics are weighted by the 1993 panel sampling weights.

Table A3: OLS Estimates and Estimated Bounds on the ATE of Depressive Symptom Intensity on Employment, Full Sample

	OLS	No Assumption	MTS	MTR	MTR+MTS	MIV+MTR+MTS
	(1)	(2)	(3)	(4)	(5)	(6)
Δ(2,1)	-0.024* (.012)	[-0.352, 0.833] (-0.370, 0.848)	[-0.189, 0.814] (-0.211, 0.830)	[-0.206, 0.000] (-0.219, 0.000)	[-0.043, 0.000] (-0.060, 0.000)	[-0.038, 0.000] (-0.057, 0.000)
Δ(3,2)	-0.055** (.020)	[-0.861, 0.348] (-0.870, 0.366)	[-0.157, 0.241] (-0.185, 0.259)	[-0.771, 0.000] (-0.788, 0.000)	[-0.067, 0.000] (-0.100, 0.000)	[-0.038, 0.000] (-0.056, 0.000)
Δ(4,3)	-0.041 (.030)	[-0.952, 0.927] (-0.958, 0.934)	[-0.223, 0.229] (-0.265, 0.259)	[-0.823, 0.000] (-0.841, 0.000)	[-0.093, 0.000] (-0.146, 0.000)	[-0.083, 0.000] (-0.130, 0.000)
Δ(5,4)	-0.075** (.031)	[-0.945, 0.947] (-0.951, 0.954)	[-0.254, 0.235] (-0.302, 0.276)	[-0.849, 0.000] (-0.864, 0.000)	[-0.159, 0.000] (-0.208, 0.000)	[-0.124, 0.000] (-0.165, 0.000)
Δ(3,1)	-0.079*** (.012)	[-0.928, 0.896] (-0.933, 0.905)	[-0.170, 0.878] (-0.190, 0.886)	[-0.846, 0.000] (-0.856, 0.000)	[-0.088, 0.000] (-0.107, 0.000)	[-0.060, 0.000] (-0.074, 0.000)
Δ(4,1)	-0.120*** (.025)	[-0.957, 0.899] (-0.961, 0.910)	[-0.171, 0.886] (-0.209, 0.896)	[-0.910, 0.000] (-0.919, 0.000)	[-0.125, 0.000] (-0.168, 0.000)	[-0.107, -0.011] (-0.147, 0.000)
Δ(5,1)	-0.195*** (.025)	[-0.945, 0.890] (-0.951, 0.901)	[-0.195, 0.890] (-0.236, 0.901)	[-0.945, 0.000] (-0.951, 0.000)	[-0.195, 0.000] (-0.239, 0.000)	[-0.167, -0.030] (-0.210, 0.000)
Δ(4,2)	-0.096*** (.031)	[-0.890, 0.351] (-0.902, 0.369)	[-0.158, 0.248] (-0.207, 0.269)	[-0.835, 0.000] (-0.852, 0.000)	[-0.103, 0.000] (-0.159, 0.000)	[-0.084, -0.009] (-0.130, 0.000)
Δ(5,2)	-0.171*** (.031)	[-0.879, 0.342] (-0.893, 0.359)	[-0.182, 0.252] (-0.233, 0.272)	[-0.871, 0.000] (-0.885, 0.000)	[-0.174, 0.000] (-0.228, 0.000)	[-0.146, -0.029] (-0.192, 0.000)
Δ(5,3)	-0.116*** (.027)	[-0.941, 0.918] (-0.948, 0.925)	[-0.246, 0.233] (-0.294, 0.262)	[-0.858, 0.000] (-0.873, 0.000)	[-0.163, 0.000] (-0.214, 0.000)	[-0.145, -0.002] (-0.190, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds. Δ(2,1): little vs no depressive symptoms; Δ(3,2): mild vs little depressive symptoms; Δ(4,3): moderate vs mild depressive symptoms; Δ(5,4): severe vs moderate depressive symptoms; Δ(3,1): mild vs no depressive symptoms; Δ(4,1): moderate vs no depressive symptoms; Δ(5,1): severe vs no depressive symptoms; Δ(4,2): moderate vs little depressive symptoms; Δ(5,2): severe vs little depressive symptoms; Δ(5,3): severe vs mild depressive symptoms.

Table A4: OLS Estimates and Estimated Bounds on the ATE of Depressive Symptom Intensity on Earnings, Full Sample

	OLS	No Assumption	MTS	MTR	MTR+MTS	MIV+MTR+MTS
	(1)	(2)	(3)	(4)	(5)	(6)
Δ(2,1)	-3854*** (731)	[-77478, 42129] (-78711, 43151)	[-7989, 27607] (-9278, 28895)	[-74726, 0.000] (-76013, 0.000)	[-5236, 0.000] (-6575, 0.000)	[-4129, 0.000] (-5173, 0.000)
Δ(3,2)	-4900*** (1134)	[-43422, 78590] (-44542, 80008)	[-14670, 67299] (-16086, 68758)	[-34515, 0.000] (-35868, 0.000)	[-5764, 0.000] (-7632, 0.000)	[-3928, 0.000] (-5614, 0.000)
Δ(4,3)	-2634** (1193)	[-93918, 95810] (-94515, 96256)	[-71656, 76050] (-73440, 76939)	[-29471, 0.000] (-30592, 0.000)	[-7208, 0.000] (-9307, 0.000)	[-5905, 0.000] (-8328, 0.000)
Δ(5,4)	-2909** (1106)	[-96378, 94578] (-96872, 95330)	[-79704, 81366] (-81229, 83327)	[-26320, 0.000] (-27498, 0.000)	[-9641, 0.000] (-10500, 0.000)	[-6018, 0.000] (-7717, 0.000)
Δ(3,1)	-8754*** (1281)	[-92130, 91947] (-92959, 92494)	[-10654, 82898] (-12780, 84075)	[-90710, 0.000] (-91511, 0.000)	[-9233, 0.000] (-11484, 0.000)	[-7390, -101] (-9583, 0.000)
Δ(4,1)	-11389*** (1592)	[-92808, 94517] (-93682, 94996)	[-12349, 88988] (-14990, 89876)	[-92043, 0.000] (-92877, 0.000)	[-11584, 0.000] (-14463, 0.000)	[-9232, -1558] (-11983, 0.000)
Δ(5,1)	-14298*** (907)	[-92698, 92606] (-93504, 93205)	[-14303, 92606] (-15844, 93205)	[-92698, 0.000] (-93504, 0.000)	[-14303, 0.000] (-15890, 0.000)	[-11412, -2882] (-12633, -493)
Δ(4,2)	-7534*** (1286)	[-44100, 81159] (-45194, 82301)	[-16365, 73389] (-18202, 74553)	[-35848, 0.000] (-37231, 0.000)	[-8114, 0.000] (-10454, 0.000)	[-6090, -1358] (-8574, 0.000)
Δ(5,2)	-10444*** (552)	[-43990, 79248] (-45064, 80323)	[-18321, 77008] (-19643, 78149)	[-36503, 0.000] (-37901, 0.000)	[-10833, 0.000] (-11770, 0.000)	[-8099, -2692] (-9109, -249)
Δ(5,3)	-5543*** (1079)	[-93808, 93899] (-94430, 94587)	[-73614, 79670] (-75410, 80825)	[-30126, 0.000] (-31276, 0.000)	[-9927, 0.000] (-10845, 0.000)	[-7924, -60] (-8957, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. Earnings are bounded between $0 and $100,948. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds. Δ(2,1): little vs no depressive symptoms; Δ(3,2): mild vs little depressive symptoms; Δ(4,3): moderate vs mild depressive symptoms; Δ(5,4): severe vs moderate depressive symptoms; Δ(3,1): mild vs no depressive symptoms; Δ(4,1): moderate vs no depressive symptoms; Δ(5,1): severe vs no depressive symptoms; Δ(4,2): moderate vs little depressive symptoms; Δ(5,2): severe vs little depressive symptoms; Δ(5,3): severe vs mild depressive symptoms.

Table A5: OLS Estimates and Estimated Bounds on the ATE of Educational Attainment on Labor Market Outcomes

	OLS	No Assumption	MTS	MTR	MTR+MTS	MIV+MTR+MTS
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Employment						
High school dropout vs High school grad/some college	0.137*** (0.025)	[-0.398, 0.837] (-0.413, 0.851)	[-0.192, 0.166] (-0.207, 0.210)	[0.000, 0.825] (0.000, 0.840)	[0.000, 0.154] (0.000, 0.200)	[0.000, 0.106] (0.000, 0.162)
High school grad/some college vs College grad	0.071*** (.004)	[-0.697, 0.415] (-0.709, 0.430)	[-0.683, 0.170] (-0.694, 0.185)	[0.000, 0.332] (0.000, 0.344)	[0.000, 0.087] (0.000, 0.096)	[0.000, 0.069] (0.000, 0.080)
College grad vs High school dropout	0.208*** (.028)	[-0.749, 0.905] (-0.763, 0.921)	[-0.749, 0.208] (-0.763, 0.254)	[0.000, 0.905] (0.000, 0.921)	[0.000, 0.208] (0.000, 0.258)	[0.004, 0.148] (0.001, 0.196)
Panel B: Earnings						
High school dropout vs High school grad/some college	8355*** (1217)	[-78523, 46114] (-79547, 47951)	[-74105, 27625] (-75259, 30156)	[0.000, 30319] (0.000, 31709)	[0.000, 11829] (0.000, 13955)	[0.000, 9148] (0.000, 10971)
High school grad/some college vs College grad	14785*** (972)	[-39398, 72847] (-40891, 74082)	[-30214, 16896] (-31423, 18495)	[0.000, 71675] (0.000, 72851)	[0.000, 15723] (0.000, 17277)	[0.000, 8578] (0.000, 9392)
College grad vs High school dropout	23139*** (1530)	[-82948, 83984] (-84370, 84938)	[-82948, 23146] (-84370, 25543)	[0.000, 83984] (0.000, 84938)	[0.000, 23146] (0.000, 25606)	[307, 16428] (199, 17919)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. Earnings are bounded between $0 and $100,948. AFQT score is used as the MIV with 5 bins. Sample size is 7,659. Average employment and earnings in (1) high school dropouts are 0.74 and $10,480; (2) high school grad/some college are 0.88 and $18,835 and (3) college grads are 0.95 and $33,619. 1993 panel weights are applied to OLS regressions and estimated bounds.

Table A6: OLS Estimates and Estimated Bounds on the ATE of the being Depressed on Employment by Gender, and Race/Ethnicity

	OLS	**No Assumption**	**MTS**	**MTR**	**MTR+MTS**	**MIV+MTR+MTS**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Men	-0.121*** (.019)	[-0.841, 0.159] (-0.857, 0.175)	[-0.121, 0.159] (-0.154, 0.175)	[-0.841, 0.000] (-0.857, 0.000)	[-0.121, 0.000] (-0.156, 0.000)	[-0.107, 0.000] (-0.135, 0.000)
Women	-0.075** (.028)	[-0.703, 0.297] (-0.717, 0.312)	[-0.075, 0.297] (-0.122, 0.312)	[-0.703, 0.000] (-0.717, 0.000)	[-0.075, 0.000] (-0.127, 0.000)	[-0.060, 0.000] (-0.104, 0.000)
White Individuals	-0.080** (.023)	[-0.787, 0.214] (-0.807, 0.234)	[-0.080, 0.214] (-0.119, 0.235)	[-0.787, 0.000] (-0.807, 0.000)	[-0.080, 0.000] (-0.122, 0.000)	[-0.078, 0.000] (-0.110, 0.000)
Black Individuals	-0.174*** (.028)	[-0.704, 0.297] (-0.726, 0.319)	[-0.175, 0.297] (-0.235, 0.319)	[-0.704, 0.000] (-0.726, 0.000)	[-0.175, 0.000] (-0.240, 0.000)	[-0.153, 0.000] (-0.205, 0.000)
Hispanic Individuals	-0.118** (.033)	[-0.722, 0.279] (-0.761, 0.318)	[-0.118, 0.279] (-0.181, 0.320)	[-0.722, 0.000] (-0.761, 0.000)	[-0.118, 0.000] (-0.186, 0.000)	[-0.098, 0.000] (-0.160, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds.

Table A7: OLS Estimates and Estimated Bounds on the ATE of Depressive Symptom Intensity on Employment, Men

	OLS	No Assumption	MTS	MTR	MTR+MTS	MIV+MTR+MTS
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Δ(2,1)	-0.020*** (.004)	[-0.290, 0.858] (-0.302, 0.872)	[-0.162, 0.851] (-0.182, 0.865)	[-0.166, 0.000] (-0.176, 0.000)	[-0.038, 0.000] (-0.049, 0.000)	[-0.038, 0.000] (-0.056, 0.000)
Δ(3,2)	-0.056** (.021)	[-0.910, 0.286] (-0.922, 0.297)	[-0.137, 0.171] (-0.179, 0.185)	[-0.841, 0.000] (-0.857, 0.000)	[-0.068, 0.000] (-0.105, 0.000)	[-0.047, 0.000] (-0.070, 0.000)
Δ(4,3)	-0.047** (.021)	[-0.965, 0.938] (-0.971, 0.946)	[-0.175, 0.164] (-0.214, 0.206)	[-0.892, 0.000] (-0.906, 0.000)	[-0.102, 0.000] (-0.143, 0.000)	[-0.046, 0.000] (-0.094, 0.000)
Δ(5,4)	-0.103** (.051)	[-0.957, 0.959] (-0.965, 0.966)	[-0.238, 0.178] (-0.305, 0.221)	[-0.915, 0.000] (-0.930, 0.000)	[-0.196, 0.000] (-0.267, 0.000)	[-0.159, -0.016] (-0.223, 0.000)
Δ(3,1)	-0.077*** (.022)	[-0.940, 0.884] (-0.948, 0.899)	[-0.152, 0.876] (-0.195, 0.888)	[-0.874, 0.000] (-0.886, 0.000)	[-0.086, 0.000] (-0.125, 0.000)	[-0.077, 0.000] (-0.111, 0.000)
Δ(4,1)	-0.123*** (.025)	[-0.969, 0.885] (-0.972, 0.903)	[-0.167, 0.878] (-0.211, 0.894)	[-0.931, 0.000] (-0.939, 0.000)	[-0.129, 0.000] (-0.174, 0.000)	[-0.071, -0.010] (-0.126, 0.000)
Δ(5,1)	-0.226*** (.041)	[-0.959, 0.878] (-0.966, 0.894)	[-0.226, 0.878] (-0.294, 0.894)	[-0.959, 0.000] (-0.966, 0.000)	[-0.226, 0.000] (-0.301, 0.000)	[-0.211, -0.068] (-0.287, 0.000)
Δ(4,2)	-0.103*** (.026)	[-0.939, 0.288] (-0.946, 0.298)	[-0.151, 0.173] (-0.197, 0.192)	[-0.898, 0.000] (-0.910, 0.000)	[-0.110, 0.000] (-0.156, 0.000)	[-0.047, -0.009] (-0.095, 0.000)
Δ(5,2)	-0.206*** (.039)	[-0.929, 0.280] (-0.941, 0.294)	[-0.211, 0.173] (-0.279, 0.194)	[-0.927, 0.000] (-0.938, 0.000)	[-0.208, 0.000] (-0.280, 0.000)	[-0.183, -0.067] (-0.245, 0.000)
Δ(5,3)	-0.150*** (.049)	[-0.956, 0.930] (-0.965, 0.937)	[-0.235, 0.163] (-0.304, 0.209)	[-0.920, 0.000] (-0.935, 0.000)	[-0.200, 0.000] (-0.272, 0.000)	[-0.182, -0.022] (-0.244, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds. Δ(2,1): little vs no depressive symptoms; Δ(3,2): mild vs little depressive symptoms; Δ(4,3): moderate vs mild depressive symptoms; Δ(5,4): severe vs moderate depressive symptoms; Δ(3,1): mild vs no depressive symptoms; Δ(4,1): moderate vs no depressive symptoms; Δ(5,1): severe vs no depressive symptoms; Δ(4,2): moderate vs little depressive symptoms; Δ(5,2): severe vs little depressive symptoms; Δ(5,3): severe vs mild depressive symptoms.

Table A8: OLS Estimates and Estimated Bounds on the ATE of Depressive Symptom Intensity on Employment, Women

	OLS	No Assumption	MTS	MTR	MTR+MTS	MIV+MTR+MTS
	(1)	(2)	(3)	(4)	(5)	(6)
Δ(2,1)	-0.016 (.033)	[-0.412, 0.809] (-0.436, 0.832)	[-0.200, 0.772] (-0.254, 0.797)	[-0.244, 0.000] (-0.263, 0.000)	[-0.032, 0.000] (-0.087, 0.000)	[-0.033, 0.000] (-0.082, 0.000)
Δ(3,2)	-0.034 (.031)	[-0.813, 0.408] (-0.822, 0.433)	[-0.154, 0.308] (-0.193, 0.329)	[-0.703, 0.000] (-0.717, 0.000)	[-0.044, 0.000] (-0.095, 0.000)	[-0.030, 0.000] (-0.063, 0.000)
Δ(4,3)	-0.031 (.048)	[-0.940, 0.916] (-0.948, 0.929)	[-0.248, 0.277] (-0.318, 0.314)	[-0.756, 0.000] (-0.769, 0.000)	[-0.064, 0.000] (-0.148, 0.000)	[-0.041, 0.000] (-0.092, 0.000)
Δ(5,4)	-0.059 (.043)	[-0.932, 0.935] (-0.940, 0.945)	[-0.262, 0.270] (-0.331, 0.339)	[-0.785, 0.000] (-0.794, 0.000)	[-0.114, 0.000] (-0.182, 0.000)	[-0.103, 0.000] (-0.178, 0.000)
Δ(3,1)	-0.049* (.019)	[-0.916, 0.907] (-0.925, 0.919)	[-0.155, 0.881] (-0.193, 0.891)	[-0.819, 0.000] (-0.829, 0.000)	[-0.058, 0.000] (-0.093, 0.000)	[-0.016, 0.000] (-0.043, 0.000)
Δ(4,1)	-0.081** (.032)	[-0.945, 0.912] (-0.950, 0.926)	[-0.140, 0.894] (-0.191, 0.908)	[-0.890, 0.000] (-0.899, 0.000)	[-0.085, 0.000] (-0.141, 0.000)	[-0.056, -0.004] (-0.106, 0.000)
Δ(5,1)	-0.140*** (.028)	[-0.932, 0.903] (-0.938, 0.917)	[-0.140, 0.903] (-0.186, 0.917)	[-0.932, 0.000] (-0.938, 0.000)	[-0.140, 0.000] (-0.191, 0.000)	[-0.113, 0.000] (-0.171, 0.000)
Δ(4,2)	-0.065 (.047)	[-0.842, 0.413] (-0.855, 0.436)	[-0.139, 0.321] (-0.217, 0.341)	[-0.774, 0.000] (-0.790, 0.000)	[-0.071, 0.000] (-0.159, 0.000)	[-0.043, -0.002] (-0.091, 0.000)
Δ(5,2)	-0.124** (.044)	[-0.829, 0.403] (-0.844, 0.424)	[-0.139, 0.329] (-0.214, 0.346)	[-0.816, 0.000] (-0.829, 0.000)	[-0.126, 0.000] (-0.205, 0.000)	[-0.117, 0.000] (-0.189, 0.000)
Δ(5,3)	-0.091** (.034)	[-0.927, 0.906] (-0.932, 0.918)	[-0.248, 0.285] (-0.316, 0.319)	[-0.798, 0.000] (-0.808, 0.000)	[-0.118, 0.000] (-0.190, 0.000)	[-0.115, 0.000] (-0.187, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds. Δ(2,1): little vs no depressive symptoms; Δ(3,2): mild vs little depressive symptoms; Δ(4,3): moderate vs mild depressive symptoms; Δ(5,4): severe vs moderate depressive symptoms; Δ(3,1): mild vs no depressive symptoms; Δ(4,1): moderate vs no depressive symptoms; Δ(5,1): severe vs no depressive symptoms; Δ(4,2): moderate vs little depressive symptoms; Δ(5,2): severe vs little depressive symptoms; Δ(5,3): severe vs mild depressive symptoms.

Table A9: OLS Estimates and Estimated Bounds on the ATE of Depressive Symptom Intensity on Employment, White Individuals

	OLS	No Assumption	MTS	MTR	MTR+MTS	MIV+MTR+MTS
	(1)	(2)	(3)	(4)	(5)	(6)
Δ(2,1)	-0.015 (.014)	[-0.334, 0.831] (-0.357, 0.848)	[-0.164, 0.816] (-0.190, 0.835)	[-0.197, 0.000] (-0.214, 0.000)	[-0.027, 0.000] (-0.049, 0.000)	[-0.035, 0.000] (-0.055, 0.000)
Δ(3,2)	-0.027 (.022)	[-0.870, 0.334] (-0.881, 0.358)	[-0.120, 0.223] (-0.149, 0.247)	[-0.787, 0.000] (-0.807, 0.000)	[-0.037, 0.000] (-0.075, 0.000)	[-0.037, 0.000] (-0.069, 0.000)
Δ(4,3)	-0.035 (.034)	[-0.961, 0.932] (-0.965, 0.939)	[-0.185, 0.190] (-0.232, 0.220)	[-0.840, 0.000] (-0.861, 0.000)	[-0.064, 0.000] (-0.127, 0.000)	[-0.063, 0.000] (-0.092, 0.000)
Δ(5,4)	-0.087** (.033)	[-0.951, 0.955] (-0.958, 0.963)	[-0.229, 0.191] (-0.293, 0.238)	[-0.865, 0.000] (-0.883, 0.000)	[-0.143, 0.000] (-0.208, 0.000)	[-0.100, 0.000] (-0.145, 0.000)
Δ(3,1)	-0.042** (.011)	[-0.929, 0.891] (-0.934, 0.902)	[-0.125, 0.879] (-0.140, 0.888)	[-0.855, 0.000] (-0.865, 0.000)	[-0.050, 0.000] (-0.068, 0.000)	[-0.059, 0.000] (-0.076, 0.000)
Δ(4,1)	-0.077** (.027)	[-0.960, 0.894] (-0.965, 0.906)	[-0.126, 0.885] (-0.168, 0.896)	[-0.917, 0.000] (-0.927, 0.000)	[-0.082, 0.000] (-0.133, 0.000)	[-0.086, 0.000] (-0.112, 0.000)
Δ(5,1)	-0.164*** (.033)	[-0.948, 0.885] (-0.955, 0.897)	[-0.164, 0.885] (-0.220, 0.897)	[-0.948, 0.000] (-0.955, 0.000)	[-0.164, 0.000] (-0.226, 0.000)	[-0.134, -0.002] (-0.181, 0.000)
Δ(4,2)	-0.062 (.036)	[-0.901, 0.337] (-0.916, 0.360)	[-0.121, 0.229] (-0.177, 0.254)	[-0.848, 0.000] (-0.870, 0.000)	[-0.069, 0.000] (-0.135, 0.000)	[-0.064, 0.000] (-0.092, 0.000)
Δ(5,2)	-0.149*** (.039)	[-0.888, 0.328] (-0.907, 0.350)	[-0.160, 0.229] (-0.228, 0.252)	[-0.879, 0.000] (-0.898, 0.000)	[-0.151, 0.000] (-0.223, 0.000)	[-0.110, -0.001] (-0.158, 0.000)
Δ(5,3)	-0.122*** (.033)	[-0.948, 0.924] (-0.955, 0.931)	[-0.224, 0.190] (-0.288, 0.220)	[-0.871, 0.000] (-0.889, 0.000)	[-0.147, 0.000] (-0.214, 0.000)	[-0.109, 0.000] (-0.156, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds. Δ(2,1): little vs no depressive symptoms; Δ(3,2): mild vs little depressive symptoms; Δ(4,3): moderate vs mild depressive symptoms; Δ(5,4): severe vs moderate depressive symptoms; Δ(3,1): mild vs no depressive symptoms; Δ(4,1): moderate vs no depressive symptoms; Δ(5,1): severe vs no depressive symptoms; Δ(4,2): moderate vs little depressive symptoms; Δ(5,2): severe vs little depressive symptoms; Δ(5,3): severe vs mild depressive symptoms.

Table A10: OLS Estimates and Estimated Bounds on the ATE of Depressive Symptom Intensity on Employment, Black Individuals

	OLS	No Assumption	MTS	MTR	MTR+MTS	MIV+MTR+MTS
	(1)	(2)	(3)	(4)	(5)	(6)
Δ(2,1)	-0.053 (.041)	[-0.429, 0.848] (-0.453, 0.881)	[-0.286, 0.804] (-0.355, 0.846)	[-0.243, 0.000] (-0.286, 0.000)	[-0.101, 0.000] (-0.179, 0.000)	[-0.099, 0.000] (-0.192, 0.000)
Δ(3,2)	-0.122*** (.025)	[-0.814, 0.403] (-0.836, 0.427)	[-0.249, 0.313] (-0.296, 0.336)	[-0.704, 0.000] (-0.726, 0.000)	[-0.138, 0.000] (-0.183, 0.000)	[-0.075, 0.000] (-0.099, 0.000)
Δ(4,3)	-0.039 (.084)	[-0.916, 0.895] (-0.934, 0.912)	[-0.301, 0.345] (-0.447, 0.400)	[-0.753, 0.000] (-0.779, 0.000)	[-0.137, 0.000] (-0.289, 0.000)	[-0.096, 0.000] (-0.223, 0.000)
Δ(5,4)	-0.081 (.091)	[-0.923, 0.913] (-0.940, 0.926)	[-0.347, 0.343] (-0.448, 0.480)	[-0.781, 0.000] (-0.809, 0.000)	[-0.204, 0.000] (-0.298, 0.000)	[-0.150, 0.000] (-0.244, 0.000)
Δ(3,1)	-0.175*** (.024)	[-0.914, 0.922] (-0.930, 0.928)	[-0.294, 0.876] (-0.336, 0.890)	[-0.809, 0.000] (-0.835, 0.000)	[-0.189, 0.000] (-0.237, 0.000)	[-0.153, 0.000] (-0.225, 0.000)
Δ(4,1)	-0.214** (.080)	[-0.942, 0.929] (-0.957, 0.944)	[-0.274, 0.901] (-0.419, 0.933)	[-0.889, 0.000] (-0.903, 0.000)	[-0.222, 0.000] (-0.377, 0.000)	[-0.153, -0.032] (-0.305, -0.008)
Δ(5,1)	-0.295*** (.052)	[-0.942, 0.918] (-0.951, 0.937)	[-0.295, 0.918] (-0.392, 0.937)	[-0.942, 0.000] (-0.951, 0.000)	[-0.295, 0.000] (-0.401, 0.000)	[-0.244, -0.049] (-0.339, 0.000)
Δ(4,2)	-0.161* (.075)	[-0.842, 0.410] (-0.871, 0.435)	[-0.229, 0.338] (-0.377, 0.374)	[-0.784, 0.000] (-0.809, 0.000)	[-0.171, 0.000] (-0.322, 0.000)	[-0.095, -0.030] (-0.222, -0.005)
Δ(5,2)	-0.241*** (.049)	[-0.842, 0.399] (-0.870, 0.419)	[-0.250, 0.355] (-0.339, 0.378)	[-0.836, 0.000] (-0.863, 0.000)	[-0.244, 0.000] (-0.337, 0.000)	[-0.190, -0.048] (-0.271, 0.000)
Δ(5,3)	-0.120** (.044)	[-0.916, 0.885] (-0.927, 0.897)	[-0.322, 0.363] (-0.418, 0.408)	[-0.805, 0.000] (-0.831, 0.000)	[-0.211, 0.000] (-0.300, 0.000)	[-0.190, 0.000] (-0.270, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds. Δ(2,1): little vs no depressive symptoms; Δ(3,2): mild vs little depressive symptoms; Δ(4,3): moderate vs mild depressive symptoms; Δ(5,4): severe vs moderate depressive symptoms; Δ(3,1): mild vs no depressive symptoms; Δ(4,1): moderate vs no depressive symptoms; Δ(5,1): severe vs no depressive symptoms; Δ(4,2): moderate vs little depressive symptoms; Δ(5,2): severe vs little depressive symptoms; Δ(5,3): severe vs mild depressive symptoms.

Table A11: OLS Estimates and Estimated Bounds on the ATE of Depressive Symptom Intensity on Employment, Hispanic Individuals

	OLS	No Assumption	MTS	MTR	MTR+MTS	MIV+MTR+MTS
	(1)	(2)	(3)	(4)	(5)	(6)
Δ(2,1)	-0.059 (.030)	[-0.413, 0.827] (-0.478, 0.876)	[-0.267, 0.795] (-0.325, 0.858)	[-0.232, 0.000] (-0.301, 0.000)	[-0.086, 0.000] (-0.131, 0.000)	[-0.091, 0.000] (-0.164, 0.000)
Δ(3,2)	-0.099 (.057)	[-0.851, 0.402] (-0.880, 0.472)	[-0.237, 0.287] (-0.323, 0.341)	[-0.722, 0.000] (-0.761, 0.000)	[-0.107, 0.000] (-0.209, 0.000)	[-0.074, 0.000] (-0.152, 0.000)
Δ(4,3)	0.018 (.050)	[-0.929, 0.927] (-0.940, 0.936)	[-0.236, 0.325] (-0.306, 0.405)	[-0.762, 0.000] (-0.800, 0.000)	[-0.069, 0.000] (-0.150, 0.000)	[-0.009, 0.000] (-0.061, 0.000)
Δ(5,4)	-0.061 (.052)	[-0.916, 0.920] (-0.923, 0.935)	[-0.242, 0.265] (-0.336, 0.341)	[-0.800, 0.000] (-0.836, 0.000)	[-0.127, 0.000] (-0.200, 0.000)	[-0.067, -0.020] (-0.149, 0.000)
Δ(3,1)	-0.157** (.053)	[-0.935, 0.900] (-0.944, 0.927)	[-0.284, 0.862] (-0.361, 0.899)	[-0.812, 0.000] (-0.831, 0.000)	[-0.161, 0.000] (-0.247, 0.000)	[-0.131, 0.000] (-0.210, 0.000)
Δ(4,1)	-0.140*** (.043)	[-0.940, 0.903] (-0.953, 0.932)	[-0.216, 0.882] (-0.278, 0.918)	[-0.870, 0.000] (-0.884, 0.000)	[-0.146, 0.000] (-0.218, 0.000)	[-0.085, -0.004] (-0.163, 0.000)
Δ(5,1)	-0.201*** (.022)	[-0.923, 0.891] (-0.930, 0.925)	[-0.201, 0.891] (-0.248, 0.925)	[-0.923, 0.000] (-0.930, 0.000)	[-0.201, 0.000] (-0.252, 0.000)	[-0.166, -0.051] (-0.236, 0.000)
Δ(4,2)	-0.081 (.051)	[-0.856, 0.406] (-0.889, 0.473)	[-0.169, 0.308] (-0.249, 0.359)	[-0.780, 0.000] (-0.816, 0.000)	[-0.092, 0.000] (-0.185, 0.000)	[-0.016, 0.000] (-0.075, 0.000)
Δ(5,2)	-0.142*** (.039)	[-0.840, 0.393] (-0.866, 0.456)	[-0.154, 0.316] (-0.234, 0.364)	[-0.833, 0.000] (-0.863, 0.000)	[-0.148, 0.000] (-0.226, 0.000)	[-0.096, -0.045] (-0.145, 0.000)
Δ(5,3)	-0.044 (.066)	[-0.913, 0.915] (-0.923, 0.927)	[-0.222, 0.333] (-0.318, 0.419)	[-0.815, 0.000] (-0.848, 0.000)	[-0.124, 0.000] (-0.203, 0.000)	[-0.090, 0.000] (-0.138, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds. Δ(2,1): little vs no depressive symptoms; Δ(3,2): mild vs little depressive symptoms; Δ(4,3): moderate vs mild depressive symptoms; Δ(5,4): severe vs moderate depressive symptoms; Δ(3,1): mild vs no depressive symptoms; Δ(4,1): moderate vs no depressive symptoms; Δ(5,1): severe vs no depressive symptoms; Δ(4,2): moderate vs little depressive symptoms; Δ(5,2): severe vs little depressive symptoms; Δ(5,3): severe vs mild depressive symptoms.

Table A12: OLS Estimates and Estimated Bounds on the ATE of the being Depressed on Earnings by Gender, and Race/Ethnicity

	OLS	No Assumption	MTS	MTR	MTR+MTS	MIV+MTR+MTS
	(1)	(2)	(3)	(4)	(5)	(6)
Men	-9726*** (856)	[-36,708,64,253] (-38,861,66,406)	[-9730, 64,253] (-11,169,66,406)	[-36,708, 0.000] (-38,864, 0.000)	[-9730, 0.000] (-11,237, 0.000)	[-7184, 0.000] (-8392, 0.000)
Women	-4228*** (981)	[-32,374,68,586] (-34,234,70,446)	[-4234, 68,586] (-5917, 70,446)	[-32,374, 0.000] (-34,237, 0.000)	[-4234, 0.000] (-6119, 0.000)	[-3271, 0.000] (-5239, 0.000)
White Individuals	-7214*** (912)	[-33,887,67,071] (-35,467,68,651)	[-7219, 67,071] (-8780, 68,651)	[-33,887, 0.000] (-35,467, 0.000)	[-7219, 0.000] (-8901, 0.000)	[-6048, 0.000] (-7764, 0.000)
Black Individuals	-7141*** (1123)	[-37,338,63,618] (-38,584,64,863)	[-7148, 63,618] (-9348, 64,863)	[-37,338, 0.000] (-38,584, 0.000)	[-7148, 0.000] (-9568, 0.000)	[-5424, 0.000] (-7632, 0.000)
Hispanic Individuals	-7614*** (773)	[-36,347,64,607] (-37,335,65,595)	[-7619, 64,607] (-9076, 65,595)	[-36,347, 0.000] (-37,335, 0.000)	[-7619, 0.000] (-9159, 0.000)	[-6549, 0.000] (-7914, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. Earnings are bounded between $0 and $100,948. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds.

Table A13: OLS Estimates and Estimated Bounds on the ATE of Depressive Symptom Intensity on Earnings, Men

	OLS	No Assumption	MTS	MTR	MTR+MTS	MIV+MTR+MTS
	(1)	(2)	(3)	(4)	(5)	(6)
Δ(2,1)	-5143*** (837)	[-72481, 43384] (-74347, 45182)	[-9314, 32644] (-11006, 34613)	[-69652, 0.000] (-71633, 0.000)	[-6483, 0.000] (-8079, 0.000)	[-4892, 0.000] (-5811, 0.000)
Δ(3,2)	-5734*** (1400)	[-45703, 75083] (-47819, 76468)	[-15803, 65320] (-17926, 67578)	[-36708, 0.000] (-38861, 0.000)	[-6808, 0.000] (-9181, 0.000)	[-4856, 0.000] (-6575, 0.000)
Δ(4,3)	-2737 (2420)	[-95299, 96807] (-96099, 97351)	[-70630, 73537] (-74310, 74848)	[-33146, 0.000] (-35415, 0.000)	[-8471, 0.000] (-11848, 0.000)	[-8079, 0.000] (-14122, 0.000)
Δ(5,4)	-5117** (2178)	[-97497, 95919] (-98097, 96817)	[-79507, 77845] (-82515, 81746)	[-31094, 0.000] (-33380, 0.000)	[-13093, 0.000] (-14312, 0.000)	[-8755, 0.000] (-11735, 0.000)
Δ(3,1)	-10877*** (1536)	[-91926, 92210] (-93168, 92992)	[-12775, 85617] (-15615, 87014)	[-90526, 0.000] (-91673, 0.000)	[-11374, 0.000] (-14114, 0.000)	[-8909, -160] (-10816, -33)
Δ(4,1)	-13614*** (2569)	[-92697, 94490] (-93981, 95290)	[-14627, 90377] (-18785, 91285)	[-91956, 0.000] (-93137, 0.000)	[-13886, 0.000] (-18406, 0.000)	[-12045, -2068] (-18111, -430)
Δ(5,1)	-18732*** (742)	[-92615, 92829] (-93744, 93438)	[-18736, 92829] (-20017, 93438)	[-92615, 0.000] (-93744, 0.000)	[-18736, 0.000] (-20022, 0.000)	[-15613, -2500] (-17816, 0.000)
Δ(4,2)	-8471*** (1894)	[-46473, 77364] (-48404, 79159)	[-17652, 70082] (-20303, 72396)	[-38138, 0.000] (-40283, 0.000)	[-9318, 0.000] (-12780, 0.000)	[-8302, -1686] (-14409, 0.000)
Δ(5,2)	-13588*** (752)	[-46392, 75705] (-48312, 77870)	[-21766, 72535] (-23205, 74801)	[-38798, 0.000] (-41063, 0.000)	[-14170, 0.000] (-15249, 0.000)	[-11920, -2198] (-14375, 0.000)
Δ(5,3)	-7854*** (1537)	[-95218, 95148] (-96048, 96011)	[-74746, 75990] (-77808, 77512)	[-33806, 0.000] (-36203, 0.000)	[-13323, 0.000] (-14482, 0.000)	[-11675, 0.000] (-14203, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. Earnings are bounded between $0 and $100,948. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds. Δ(2,1): little vs no depressive symptoms; Δ(3,2): mild vs little depressive symptoms; Δ(4,3): moderate vs mild depressive symptoms; Δ(5,4): severe vs moderate depressive symptoms; Δ(3,1): mild vs no depressive symptoms; Δ(4,1): moderate vs no depressive symptoms; Δ(5,1): severe vs no depressive symptoms; Δ(4,2): moderate vs little depressive symptoms; Δ(5,2): severe vs little depressive symptoms; Δ(5,3): severe vs mild depressive symptoms.

Table A14: OLS Estimates and Estimated Bounds on the ATE of Depressive Symptom Intensity on Earnings, Women

	OLS	No Assumption	MTS	MTR	MTR+MTS	MIV+MTR+MTS
	(1)	(2)	(3)	(4)	(5)	(6)
Δ(2,1)	-1092	[-82370, 40903]	[-4681, 22170]	[-79693, 0.000]	[-2004, 0.000]	[-1082, 0.000]
	(886)	(-83477, 42190)	(-5860, 23402)	(-80917, 0.000)	(-3415, 0.000)	(-1994, 0.000)
Δ(3,2)	-2278	[-41193, 82023]	[-11603, 69086]	[-32374, 0.000]	[-2785, 0.000]	[-2022, 0.000]
	(1345)	(-42459, 84125)	(-12774, 71093)	(-34234, 0.000)	(-5132, 0.000)	(-3848, 0.000)
Δ(4,3)	-2051**	[-92569, 94835]	[-70749, 77079]	[-25878, 0.000]	[-4062, 0.000]	[-3266, 0.000]
	(803)	(-93638, 95519)	(-71691, 78195)	(-26997, 0.000)	(-6452, 0.000)	(-5797, 0.000)
Δ(5,4)	-1707	[-95284, 93266]	[-78996, 82956]	[-21653, 0.000]	[-5362, 0.000]	[-4508, 0.000]
	(1235)	(-95948, 94043)	(-80008, 84834)	(-22803, 0.000)	(-6441, 0.000)	(-6152, 0.000)
Δ(3,1)	-3371***	[-92331, 91691]	[-5218, 80188]	[-90892, 0.000]	[-3779, 0.000]	[-2416, 0.000]
	(786)	(-93489, 92548)	(-6429, 81854)	(-92067, 0.000)	(-5161, 0.000)	(-4604, 0.000)
Δ(4,1)	-5423***	[-92919, 94545]	[-6350, 87646]	[-92131, 0.000]	[-5561, 0.000]	[-3857, 0.000]
	(1219)	(-94170, 95199)	(-8311, 88930)	(-93365, 0.000)	(-7743, 0.000)	(-6773, 0.000)
Δ(5,1)	-7131***	[-92781, 92390]	[-7136, 92390]	[-92781, 0.000]	[-7136, 0.000]	[-5426, -159]
	(1072)	(-93986, 93379)	(-8943, 93379)	(-93986, 0.000)	(-9104, 0.000)	(-6729, 0.000)
Δ(4,2)	-4330**	[-41780, 84875]	[-12735, 76543]	[-33612, 0.000]	[-4567, 0.000]	[-3340, 0.000]
	(1643)	(-43102, 86167)	(-14766, 77772)	(-35520, 0.000)	(-7531, 0.000)	(-5878, 0.000)
Δ(5,2)	-6038***	[-41642, 82718]	[-13521, 81287]	[-34263, 0.000]	[-6138, 0.000]	[-5038, -165]
	(731)	(-42919, 83801)	(-15177, 82446)	(-36123, 0.000)	(-7457, 0.000)	(-6192, 0.000)
Δ(5,3)	-3759	[-92431, 92678]	[-71539, 81826]	[-26529, 0.000]	[-5634, 0.000]	[-4996, 0.000]
	(1137)	(-93528, 93392)	(-73365, 83442)	(-27600, 0.000)	(-6818, 0.000)	(-6133, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. Earnings are bounded between $0 and $100,948. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds. Δ(2,1): little vs no depressive symptoms; Δ(3,2): mild vs little depressive symptoms; Δ(4,3): moderate vs mild depressive symptoms; Δ(5,4): severe vs moderate depressive symptoms; Δ(3,1): mild vs no depressive symptoms; Δ(4,1): moderate vs no depressive symptoms; Δ(5,1): severe vs no depressive symptoms; Δ(4,2): moderate vs little depressive symptoms; Δ(5,2): severe vs little depressive symptoms; Δ(5,3): severe vs mild depressive symptoms.

Table A15: OLS Estimates and Estimated Bounds on the ATE of Depressive Symptom Intensity on Earnings, White Individuals

	OLS	No Assumption	MTS	MTR	MTR+MTS	MIV+MTR+MTS
	(1)	(2)	(3)	(4)	(5)	(6)
Δ(2,1)	-3066*** (809)	[-75728, 41855] (-77049, 43078)	[-6951, 29108] (-8278, 30587)	[-72970, 0.000] (-74420, 0.000)	[-4193, 0.000] (-5618, 0.000)	[-3664, 0.000] (-4928, 0.000)
Δ(3,2)	-4135*** (1365)	[-43399, 78115] (-44771, 79962)	[-14430, 67848] (-16165, 69569)	[-33886, 0.000] (-35466, 0.000)	[-4918, 0.000] (-7216, 0.000)	[-3820, 0.000] (-6025, 0.000)
Δ(4,3)	-2226 (1645)	[-94581, 96508] (-95016, 96919)	[-71324, 75460] (-73964, 76452)	[-29515, 0.000] (-30898, 0.000)	[-6256, 0.000] (-9416, 0.000)	[-6574, 0.000] (-9312, 0.000)
Δ(5,4)	-4006** (1528)	[-97115, 95259] (-97600, 96162)	[-79864, 79694] (-81889, 82701)	[-27073, 0.000] (-28537, 0.000)	[-9815, 0.000] (-10940, 0.000)	[-6662, 0.000] (-8251, 0.000)
Δ(3,1)	-7202*** (1419)	[-91462, 92303] (-92327, 92872)	[-9042, 84612] (-11399, 85704)	[-90071, 0.000] (-90920, 0.000)	[-7650, 0.000] (-10192, 0.000)	[-6516, -82] (-9166, 0.000)
Δ(4,1)	-9428*** (2067)	[-92198, 94966] (-93140, 95555)	[-10431, 90139] (-13923, 91198)	[-91437, 0.000] (-92333, 0.000)	[-9670, 0.000] (-13587, 0.000)	[-9209, -1102] (-12643, 0.000)
Δ(5,1)	-13435*** (977)	[-92068, 92981] (-92907, 93716)	[-13440, 92981] (-15107, 93716)	[-92068, 0.000] (-92907, 0.000)	[-13440, 0.000] (-15165, 0.000)	[-10965, -3253] (-12359, -637)
Δ(4,2)	-6361*** (1756)	[-44134, 80776] (-45475, 82246)	[-15818, 73373] (-18543, 74803)	[-35253, 0.000] (-36872, 0.000)	[-6937, 0.000] (-10358, 0.000)	[-6702, -892] (-9522, 0.000)
Δ(5,2)	-10368*** (713)	[-44005, 78790] (-45319, 80142)	[-18830, 76216] (-20444, 77635)	[-35884, 0.000] (-37514, 0.000)	[-10707, 0.000] (-11946, 0.000)	[-8358, -3041] (-9489, -376)
Δ(5,3)	-6233 (1381)	[-94452, 94524] (-94966, 95338)	[-74338, 78303] (-76588, 79505)	[-30146, 0.000] (-31568, 0.000)	[-10026, 0.000] (-11254, 0.000)	[-8208, -665] (-9339, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. Earnings are bounded between $0 and $100,948. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds. Δ(2,1): little vs no depressive symptoms; Δ(3,2): mild vs little depressive symptoms; Δ(4,3): moderate vs mild depressive symptoms; Δ(5,4): severe vs moderate depressive symptoms; Δ(3,1): mild vs no depressive symptoms; Δ(4,1): moderate vs no depressive symptoms; Δ(5,1): severe vs no depressive symptoms; Δ(4,2): moderate vs little depressive symptoms; Δ(5,2): severe vs little depressive symptoms; Δ(5,3): severe vs mild depressive symptoms.

Table A16: OLS Estimates and Estimated Bounds on the ATE of Depressive Symptom Intensity on Earnings, Black Individuals

	OLS	No Assumption	MTS	MTR	MTR+MTS	MIV+MTR+MTS
	(1)	(2)	(3)	(4)	(5)	(6)
Δ(2,1)	-4589 (2412)	[-85741, 43081] (-87054, 44481)	[-9132, 19744] (-13866, 21223)	[-83091, 0.000] (-84174, 0.000)	[-6483, 0.000] (-11740, 0.000)	[-4193, 0.000] (-7913, 0.000)
Δ(3,2)	-4690*** (1542)	[-42877, 79990] (-44313, 81192)	[-11064, 64442] (-14146, 65720)	[-37338, 0.000] (-38583, 0.000)	[-5524, 0.000] (-8564, 0.000)	[-3151, 0.000] (-5235, 0.000)
Δ(4,3)	-2554 (1698)	[-90303, 92578] (-92003, 93393)	[-67967, 76060] (-70883, 78208)	[-28700, 0.000] (-30266, 0.000)	[-6363, 0.000] (-9115, 0.000)	[-4627, 0.000] (-7749, 0.000)
Δ(5,4)	-1872 (1332)	[-93245, 92040] (-94076, 93046)	[-78557, 84389] (-80268, 85755)	[-22338, 0.000] (-23955, 0.000)	[-7644, 0.000] (-8273, 0.000)	[-6074, 0.000] (-8664, 0.000)
Δ(3,1)	-9279*** (2828)	[-95448, 89905] (-95863, 91222)	[-11224, 75219] (-16386, 76854)	[-94105, 0.000] (-94551, 0.000)	[-9880, 0.000] (-15774, 0.000)	[-6328, 0.000] (-10918, 0.000)
Δ(4,1)	-11834*** (3282)	[-96049, 92782] (-96627, 93621)	[-12656, 84743] (-18829, 86371)	[-95411, 0.000] (-95971, 0.000)	[-12017, 0.000] (-18858, 0.000)	[-7713, -207] (-13405, 0.000)
Δ(5,1)	-13706*** (2395)	[-96116, 91645] (-96726, 93118)	[-13720, 91645] (-18162, 93118)	[-96116, 0.000] (-96726, 0.000)	[-13720, 0.000] (-18622, 0.000)	[-10339, -1725] (-13935, 0.000)
Δ(4,2)	-7245*** (1456)	[-43478, 82868] (-44939, 84029)	[-12493, 73969] (-15686, 75786)	[-38644, 0.000] (-40008, 0.000)	[-7658, 0.000] (-10762, 0.000)	[-4693, -173] (-8007, 0.000)
Δ(5,2)	-9117*** (478)	[-43544, 81731] (-44906, 83252)	[-13557, 80871] (-15033, 82549)	[-39349, 0.000] (-40725, 0.000)	[-9361, 0.000] (-10293, 0.000)	[-6665, -1633] (-7399, 0.000)
Δ(5,3)	-4426 (1168)	[-90369, 91440] (-91898, 92618)	[-69033, 82964] (-70780, 85501)	[-29404, 0.000] (-30842, 0.000)	[-8065, 0.000] (-8670, 0.000)	[-6581, 0.000] (-7271, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. Earnings are bounded between $0 and $100,948. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds. Δ(2,1): little vs no depressive symptoms; Δ(3,2): mild vs little depressive symptoms; Δ(4,3): moderate vs mild depressive symptoms; Δ(5,4): severe vs moderate depressive symptoms; Δ(3,1): mild vs no depressive symptoms; Δ(4,1): moderate vs no depressive symptoms; Δ(5,1): severe vs no depressive symptoms; Δ(4,2): moderate vs little depressive symptoms; Δ(5,2): severe vs little depressive symptoms; Δ(5,3): severe vs mild depressive symptoms.

Table A17: OLS Estimates and Estimated Bounds on the ATE of Depressive Symptom Intensity on Earnings, Hispanic Individuals

	OLS	**No Assumption**	**MTS**	**MTR**	**MTR+MTS**	**MIV+MTR+MTS**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Δ(2,1)	-7643* (3039)	[-81660, 43535] (-83120, 45231)	[-12191, 23829] (-17223, 25437)	[-78741, 0.000] (-80180, 0.000)	[-9272, 0.000] (-14660, 0.000)	[-9029, 0.000] (-12584, 0.000)
Δ(3,2)	-6142*** (1766)	[-44935, 81602] (-47609, 83476)	[-15553, 65426] (-18342, 66536)	[-36347, 0.000] (-37335, 0.000)	[-6969, 0.000] (-9888, 0.000)	[-6328, 0.000] (-7273, 0.000)
Δ(4,3)	-168 (2006)	[-93410, 93998] (-94185, 94880)	[-68423, 75795] (-71375, 78854)	[-30625, 0.000] (-31930, 0.000)	[-5634, 0.000] (-8225, 0.000)	[-4401, 0.000] (-7513, 0.000)
Δ(5,4)	-1145 (2404)	[-93860, 91456] (-94521, 92564)	[-75015, 80958] (-77565, 83166)	[-25489, 0.000] (-26811, 0.000)	[-6644, 0.000] (-10269, 0.000)	[-2234, 0.000] (-7090, 0.000)
Δ(3,1)	-13786*** (2636)	[-93360, 91901] (-95516, 93213)	[-15892, 77413] (-20222, 80125)	[-91406, 0.000] (-93474, 0.000)	[-13938, 0.000] (-18626, 0.000)	[-13001, -70] (-16372, 0.000)
Δ(4,1)	-13954*** (1939)	[-93474, 92604] (-95736, 94309)	[-15177, 84072] (-18358, 86128)	[-92371, 0.000] (-94389, 0.000)	[-14075, 0.000] (-17600, 0.000)	[-12220, -3022] (-16638, -336)
Δ(5,1)	-15099*** (3699)	[-93222, 89948] (-95081, 91793)	[-15119, 89948] (-21342, 91795)	[-93222, 0.000] (-95081, 0.000)	[-15119, 0.000] (-22047, 0.000)	[-14070, -3978] (-18708, 0.000)
Δ(4,2)	-6310*** (1867)	[-45049, 82303] (-47583, 83934)	[-14840, 72087] (-16807, 73205)	[-37312, 0.000] (-38336, 0.000)	[-7109, 0.000] (-9792, 0.000)	[-4896, -2455] (-8140, 0.000)
Δ(5,2)	-7455*** (1357)	[-44799, 79646] (-47772, 81100)	[-14780, 77964] (-19560, 79269)	[-38164, 0.000] (-39357, 0.000)	[-8146, 0.000] (-11498, 0.000)	[-7465, -3286] (-10297, 0.000)
Δ(5,3)	-1313 (2594)	[-93159, 91342] (-93933, 92290)	[-68355, 81672] (-71351, 84770)	[-31476, 0.000] (-32902, 0.000)	[-6673, 0.000] (-10534, 0.000)	[-7048, 0.000] (-9566, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. Earnings are bounded between $0 and $100,948. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds. Δ(2,1): little vs no depressive symptoms; Δ(3,2): mild vs little depressive symptoms; Δ(4,3): moderate vs mild depressive symptoms; Δ(5,4): severe vs moderate depressive symptoms; Δ(3,1): mild vs no depressive symptoms; Δ(4,1): moderate vs no depressive symptoms; Δ(5,1): severe vs no depressive symptoms; Δ(4,2): moderate vs little depressive symptoms; Δ(5,2): severe vs little depressive symptoms; Δ(5,3): severe vs mild depressive symptoms.

Table A18: OLS Estimates and Estimated Bounds on the ATE of Depressive Symptom Intensity on Labor Market Outcomes using Zhang et al. (2005) Groupings

	OLS	No Assumption	MTS	MTR	MTR+MTS	MIV+MTR+MTS
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Employment						
Δ(2,1)	-0.024*	[-0.352, 0.833]	[-0.189, 0.814]	[-0.206, 0.000]	[-0.043, 0.000]	[-0.038, 0.000]
	(.012)	(-0.370, 0.848)	(-0.211, 0.830)	(-0.219, 0.000)	(-0.060, 0.000)	(-0.057, 0.000)
Δ(3,2)	-0.070***	[-0.854, 0.345]	[-0.162, 0.238]	[-0.771, 0.000]	[-0.079, 0.000]	[-0.050, 0.000]
	(.022)	(-0.864, 0.364)	(-0.194, 0.258)	(-0.788, 0.000)	(-0.118, 0.000)	(-0.073, 0.000)
Δ(4,3)	-0.066**	[-0.910, 0.905]	[-0.209, 0.232]	[-0.828, 0.000]	[-0.126, 0.000]	[-0.109, 0.000]
	(.023)	(-0.918, 0.914)	(-0.247, 0.260)	(-0.844, 0.000)	(-0.166, 0.000)	(-0.148, 0.000)
Δ(3,1)	-0.094***	[-0.920, 0.893]	[-0.175, 0.876]	[-0.846, 0.000]	[-0.101, 0.000]	[-0.074, 0.000]
	(.014)	(-0.926, 0.902)	(-0.194, 0.882)	(-0.856, 0.000)	(-0.123, 0.000)	(-0.092, 0.000)
Δ(4,1)	-0.160***	[-0.918, 0.884]	[-0.160, 0.884]	[-0.918, 0.000]	[-0.160, 0.000]	[-0.132, -0.014]
	(.019)	(-0.925, 0.895)	(-0.192, 0.895)	(-0.925, 0.000)	(-0.195, 0.000)	(-0.167, 0.000)
Δ(4,2)	-0.136***	[-0.851, 0.336]	[-0.147, 0.247]	[-0.843, 0.000]	[-0.139, 0.000]	[-0.110, -0.012]
	(.025)	(-0.867, 0.353)	(-0.189, 0.265)	(-0.859, 0.000)	(-0.184, 0.000)	(-0.149, 0.000)
Panel B: Employment						
Δ(2,1)	-3854***	[-77478, 42129]	[-7989, 27607]	[-74726, 0.000]	[-5236, 0.000]	[-4129, 0.000]
	(731)	(-78711, 43151)	(-9278, 28895)	(-76013, 0.000)	(-6575, 0.000)	(-5173, 0.000)
Δ(3,2)	-5258***	[-43269, 77683]	[-14809, 67233]	[-34515, 0.000]	[-6056, 0.000]	[-3848, 0.000]
	(1112)	(-44336, 79230)	(-16016, 68673)	(-35868, 0.000)	(-7884, 0.000)	(-5478, 0.000)
Δ(4,3)	-4149***	[-92398, 90850]	[-72510, 76912]	[-28716, 0.000]	[-8825, 0.000]	[-6934, 0.000]
	(961)	(-93204, 91816)	(-73980, 78135)	(-29793, 0.000)	(-10003, 0.000)	(-7966, 0.000)
Δ(3,1)	-9113***	[-91978, 91040]	[-10793, 82832]	[-90710, 0.000]	[-9525, 0.000]	[-7271, -101]
	(1247)	(-92804, 91752)	(-12815, 83989)	(-91511, 0.000)	(-11717, 0.000)	(-9413, 0.000)
Δ(4,1)	-13262***	[-92195, 89709]	[-13268, 89709]	[-92195, 0.000]	[-13268, 0.000]	[-10498, -1525]
	(1114)	(-93024, 90571)	(-15182, 90571)	(-93024, 0.000)	(-15280, 0.000)	(-11943, 0.000)
Δ(4,2)	-9467***	[-43487, 76349]	[-17285, 74110]	[-36001, 0.000]	[-9797, 0.000]	[-7145, -1333]
	(743)	(-44578, 77341)	(-18698, 75140)	(-37422, 0.000)	(-11135, 0.000)	(-8229, 0.000)

Notes: Standard errors in (.) in column 1. In columns 2-6 estimated bounds are in [.] and corresponding 95% confidence intervals in (.) are from 999 bootstrap replications. Earnings are bounded between $0 and $100,948. AFQT score is used as the MIV with 5 bins. 1993 panel weights are applied to OLS regressions and estimated bounds. Individuals are grouped into no, mild (1≤CES-D≤15), moderate (16≤CES-D≤21) and severe (22≤CES-D≤60) depressive symptoms. Δ(2,1): mild vs no depressive symptoms; Δ(3,2): moderate vs mild depressive symptoms; Δ(4,3): severe vs moderate depressive symptoms; Δ(3,1): moderate vs no depressive symptoms; Δ(4,1): severe vs no depressive symptoms; Δ(4,2): severe vs mild depressive symptoms .

Appendix B: Technical Appendix on the CLR Method

This technical appendix provides additional details about the Chernozhukov et al. (2013; CLR) methodology that yields bias-corrected estimated bounds and valid confidence intervals for intersection bounds. To provide some intuition on the CLR method, we first make explicit the notion of creating the bins of the MIV. We use below 5 MIV bins $\mathcal{B}_m$, $m = 1, \ldots, 5$, each spanning 20 percentiles of the empirical distribution of AFQT scores in the NLSY79 sample. The lower bound on $E[Y(t_2)]$ from Equation (15) can then be rewritten as:

$$(17)\ \sum_{m=1}^{5} P(Z \in \mathcal{B}_m) \cdot \max_{m_1 \leq m} LB^1_{m_1}$$

where the $LB^1_{m_1}$ are the MTR+MTS lower bounds in bins m_1 up through m, each obtained as in Equation (12).

Instead of expressions like (17) which comprise 5 different maxima, the CLR method requires that these be rewritten as a set of expressions under a single maximum (or minimum, for upper bounds), with each element inside the max operator called a bounding function. Intuitively, each bounding function represents one of the possible outcomes from evaluating (17) in the data. Finally, the full set of bounding functions is defined for the ATE, so we also perform all necessary subtractions. For example, the final bounding functions for the lower bound on $\Delta(t_2, t_1)$ are created from all possible subtractions of the $E[Y(t_1)]$ upper bound bounding functions from the $E[Y(t_2)]$ lower bound bounding functions. In total, each bound on each ATE implies $(2\text{^}\{5-1\})\text{^}2 = 256$ bounding functions, denoted $\theta^l(v)$ and $\theta^u(v)$, $v = 1, \ldots, 256$, for the respective lower and upper bounds.

The key aspect of the CLR procedure is that the steps for estimation of the bounds and for constructing confidence intervals are completed on the individual bounding functions prior to taking the associated maximum (or minimum). This is referred to as the *precision adjustment* and proceeds as follows.[24] Generally, the adjustment involves taking the product of a critical value $\kappa(p)$ and the pointwise standard error $s(v)$ of the bounding function. For lower bounds, this product is subtracted from the estimator $\widehat{\theta^l}(v)$; for upper bounds, it is added to $\widehat{\theta^u}(v)$. Then—depending on the choice of critical value p—the adjustment yields either the half-median unbiased estimator of the lower and upper bounds ($p = 0.5$), or the desired lower and upper limits of the confidence interval (see below). In this way, the CLR method offers the convenience that bias correction and inference are carried out within the same procedure. Also, we note that the resulting large number of bounding functions makes it crucial to implement the CLR procedure for estimation of the bounds and the construction of valid confidence intervals, as, in our experience, the amount of bias tends to increase with the number of bounding functions.

More specifically, the precision-corrected estimators of the bounding functions for each average treatment effect bound are given by:

$$(18)\ \widehat{\theta^l}(p) = \max_v\{\widehat{\theta^l}(v) - \kappa^l(p) \cdot s^l(v)\}$$

and

$$(19)\ \widehat{\theta^u}(p) = \min_v\{\widehat{\theta^u}(v) + \kappa^u(p) \cdot s^u(v)\}$$

[24] This process requires that the estimators of $\theta^l(v)$ and $\theta^u(v)$ are consistent and asymptotically normal. Since in our case these estimators are made up of sample means and sample proportions, this condition is met.

where $\widehat{\theta^l}(v)$ and $\widehat{\theta^u}(v)$ are the unadjusted estimators of the bounding functions, and $s^l(v)$ and $s^u(v)$ are their associated standard errors. The critical values $\kappa^l(p)$ and $\kappa^u(p)$ are computed via simulations as follows.

Let $\widehat{\boldsymbol{\gamma}^l}$ be a 256-dimensional column vector of all the unadjusted bounding functions for the lower bound, with $\widehat{\boldsymbol{\gamma}^u}$ defined likewise for the upper bounds. An initial step obtains from $B = 999$ bootstrap replications a consistent estimate $\widehat{\boldsymbol{\Omega}^l}$ of the asymptotic variance-covariance matrix of $\sqrt{N}\left(\widehat{\boldsymbol{\gamma}^l} - \boldsymbol{\gamma}^l\right)$ (an analogous process is followed for the upper bounds). With $\widehat{\boldsymbol{g}^l}(\boldsymbol{v})'$ the v^{th} row of $\widehat{\Omega^{1/2,l}}$, we can thus define $s^l(v) \equiv \frac{||\widehat{g^l}(v)||}{\sqrt{N}}$. Next, following CLR, we simulate $R = 100{,}000$ draws from a $\mathcal{N}(\mathbf{0}, \boldsymbol{I})$ distribution, where $\boldsymbol{I}$ is the 256×256 identity matrix. The draws are labelled $\boldsymbol{Z}_r$, $r = 1, \ldots, 100{,}000$, and are used to compute $Z_r^*(v) \equiv \widehat{\boldsymbol{g}^l}(\boldsymbol{v})'\boldsymbol{Z}_r / \left\|\widehat{g^l}(v)\right\|$ for each r and v. In each replication, we select the maximum over the set of $Z_r^*(1), \ldots, Z_r^*(256)$. From the resulting R values, we compute $\kappa^l(c)$, which is defined as the c^{th} quantile of the values, where $c \equiv 1 - (0.1/\log N)$. This value is used to construct the following set:

$$\widehat{V^l} = \left\{v \in \mathcal{V}^l : \widehat{\theta^l}(v) \geq \max_{\tilde{v} \in \mathcal{V}^l}\left[\widehat{\theta^l}(\tilde{v}) - \kappa^l(c) \cdot s^l(\tilde{v})\right] - 2\kappa^l(c) \cdot s^l(v)\right\}$$

where $\mathcal{V}^l$ is the indexing set for the lower bound bounding functions $\theta^l(v)$. From the values $Z_r^*(v)$, we next take the maximum from each replication r, this time restricting the search only to $v \in \widehat{V^l}$. The CLR critical value $\kappa^l(p)$ is taken as the p^{th} quantile of the R values, such that $\kappa^l(0.5)$ gives the half-median unbiased estimate of the lower bound $\widehat{\theta^l}(0.5)$ on the average treatment effect.

To obtain the lower bound on a $(1 - \alpha) \cdot 100\%$ confidence interval, we must make one final adjustment which accounts for the width of the identified set. Borrowing notation from CLR (2013), define:

$$\hat{\Gamma} \equiv \widehat{\theta^u}(0.5) - \widehat{\theta^l}(0.5)$$
$$\widehat{\Gamma^+} \equiv \max\{0, \hat{\Gamma}\}$$
$$\rho \equiv \max\{\widehat{\theta^u}(0.75) - \widehat{\theta^u}(0.25), \widehat{\theta^l}(0.25) - \widehat{\theta^l}(0.75)\}$$
$$\tau \equiv 1/(\rho \log N)$$
$$\hat{p} \equiv 1 - \Phi\left(\tau\widehat{\Gamma^+}\right) \cdot \alpha$$

where $\Phi(\cdot)$ is the standard normal CDF. We report 95% confidence intervals for the estimates based on, for example, $\widehat{\theta^l}(\hat{p})$, which uses the critical value $\kappa^l(\hat{p})$, with $\alpha = 0.05$ in the expression for $\hat{p}$.