Religion and Depression in Adolescence*

Jane Cooley Fruehwirth†, Sriya Iyer‡, Anwen Zhang§

December 2016

Abstract

Depression is the leading cause of illness and disability in adolescence. Many studies show a correlation between religiosity and mental health, yet the question remains whether the link is causal. The key issue is selection into religiosity. We exploit plausibly random variation in adolescents' peers to shift religiosity independently of individual-level unobservables that might affect depression, and show conditions such that an individual effect of religiosity is separated from the potential direct effect of peers. Using a nationally representative sample of adolescents in the US, we find robust effects of religiosity on depression, that are particularly strong for the most depressed. We demonstrate that these effects are not driven by the school social context. We find that religiosity buffers against stressors, possibly through improved psychological resources and religion-based support structures. This has implications especially for effective mental health policy.

JEL Codes: I10, Z12

*Fruehwirth thanks the British Academy and the Leverhulme Trust's Philip Leverhulme Prize for financial support. We thank the editor and anonymous referees, Peter Arcidiacono, Donna Gilleskie, David Guilkey, Ju Hyun Kim, Brian McManus, Salvador Navarro, Alexei Onatski, Tiago Pires, Valentin Verdier and seminar participants at UNC, ASREC, SOLE, IRP summer research workshop, UNC-Greensboro, CPC, and University of Western Ontario for helpful comments and Naifu Zhang for excellent research assistance.

†University of North Carolina; jane_fruehwirth@unc.edu.

‡University of Cambridge, St. Catharine's College and IZA; si105@cam.ac.uk

§London School of Economics and Political Science; a.zhang6@lse.ac.uk

1 Introduction

Depression is the leading cause of illness and disability in adolescence worldwide, and the World Health Organization lists mental health in adolescence as a key issue that needs to be addressed (WHO, 2014). In the US, the incidence of a major depressive episode in adolescence has risen by more than a third over the past decade to 12.5 percent of adolescents as of 2015 (CBHSQ, 2016). This is a troubling trend for a number of reasons. First, depression during adolescence is correlated with a range of adverse outcomes, including lower academic achievement and non-cognitive development (Cook et al., 2009). Second, studies estimate that half of adults who suffer from mental health issues had symptoms that begin in adolescence (WHO, 2014).[1] Third, the economics costs are substantial. Between 1996 and 2006, mental health expenditure rose rapidly from $35.2 to $57.5 billion and from the 5th to the 3rd most costly medical condition in the US (AHRQ, 2014). Langa et al. (2004) estimate a yearly cost of about $9 billion for caregiving associated with depressive symptoms in elderly Americans, many of whom began to experience depression in adolescence. In this paper, we examine the role of one potentially important determinant of depression in adolescence — religiosity.

A contentious literature dating back to Freud in the early 1900s debates the role of religion in mental health and has been highly influential in the treatment of mental health problems (Levin, 2010).[2] Understanding the role of religion remains relevant today. More than 8 in 10 people identify with a religious group worldwide.[3] Sixty-five percent of Americans say religion plays an important part in their daily lives and a majority claim religion could address most or all of today's problems.[4] Among adolescents, 31 percent of twelfth graders attend church on a weekly basis, and 28 percent report that

[1] Williams et al. (2002) highlight adolescence as a key period of development that should be addressed due to its important consequences for mental health in adulthood.

[2] Discussion of these issues features in Freud (1927) and his other writings which examine religion and its effect on the human psyche.

[3] PewForum (2012).

[4] Crabtree (2010); Newport (2014).

religion plays a *very* important part in their lives.[5]

Considerable scientific evidence suggests that religiosity is positively correlated with mental health, yet the meaning of this correlation remains a puzzle (Ellison and Henderson, 2011; Levin, 2010). We contribute to the debates about religion and mental health in several ways. First, we explore whether the link between religiosity and depression can be interpreted as causal and whether the link is driven by individual practice or the social context of having school peers who are more religious. Second, we combine insight from economics and social psychology to explore how religiosity affects depression. The National Longitudinal Study of Adolescent to Adult Health in the United States, a nationally representative sample of approximately 20,000 adolescents in grades 7 to 12 in 1995 provides an excellent context for studying these questions, as it includes measures of depression, religiosity, and detailed information about the home, the school environment and associated stressors. Adolescence is a particularly critical time for studying mental health; Frank and McGuire (2000) points out that mental health issues are often chronic and tend to differ from physical health in beginning at earlier ages, from 15 to 30. Clinically, the effect of anitdepressants on reducing depression is successful in about one-fifth of cases. So it does seem necessary to examine other non-clinical ways in which the symptoms of depression might be reduced.

The key challenge with establishing a causal effect of religiosity is the issue of selection into religiosity. In our context, it could be that religiosity simply proxies difficult-to-measure aspects of family background and that it is family background rather than religiosity that leads to lower depression.[6] Further, it could be that people select into religiosity as a way of dealing with negative shocks to mental health (Maselko et al., 2012; Ferraro and Kelley-Moore, 2000). To address the issue of selection into religiosity, we exploit variation in peer religiosity, which plausibly shifts an adolescent's religiosity independently of unobserved individual attributes. This strategy relies on plausibly random variation in peer composition across cohorts within schools. We show that

[5] Child Trends Databank (2014a,b).

[6] See Wille et al. (2008) for a discussion of the importance of home environment.

this seems to hold based on observables in the data, and that our results are robust to a number of specification checks. These include controlling for the possibly confounding effects of selection based on influential local churches through controls for average religiosity (and trends) of same-denomination schoolmates.

Beyond random within-school variation in peer composition, a key threat to identification is that there may be an unobservable third factor that leads to both higher religiosity for the peer group and for the individual that is correlated with depression but is not caused by religiosity. We show this is unlikely to be a concern through a variety of specification checks. First, we use predetermined peer characteristics instead of peer religiosity as instruments and obtain estimates of the effects of religiosity of similar magnitude. Second, peer group depression and peer group characteristics, which would be correlated with any unobservable shared group effects, do not predict an individual's depression and estimated effects of religiosity are the same after conditioning on these peer variables. Third, we control directly for peer group fixed effects in the non-instrumented regression and show that shared unobservables explain almost none of the correlation between religion and depression. Finally, we try a number of alternative definitions of peer groups and show that in the over-identified cases that we pass the test of over-identifying restrictions, providing additional suggestive evidence that group-level unobservables are not driving the results.

Given random variation in peers and that an unobservable third factor does not seem to be driving our results, it remains unclear whether the effect of religiosity derives through having peers in the school who are more religious or through a direct effect of an individual's own religiosity. While arguably both effects are interesting as they indicate an effect of religion, we interpret the evidence of a lack of an effect of peer depression and peer characteristics on depression and the robustness of our estimated effect of religiosity when controlling for these peer variables to suggest that the results are driven by an individual effect of religiosity. Then, we ask the question of how religiosity affects depression, bringing together insight from economic theory and social

psychology. We consider whether religiosity bolsters psychological resources or coping mechanisms for dealing with stress, provides support structures that help compensate for lack of support in the home or school, and/or eliminates sources of stress.[7]

Our paper contributes methodologically to the literature in economics that addresses the difficult problem of disentangling a causal effect of religiosity (Iannaccone, 1998; Hungerman, 2011; Iyer, 2016). The method we use is similar in spirit to methods developed in Gruber (2005) and later applied in Mellor and Freeborn (2011) for studying the effects of religiosity in other context. These studies use variation in religiosity at the county level to shift individual religiosity, relying on insight from the competition literature on how density of churches affects attendance. We build instead on the power of within-school peers to shift religiosity.[8] What has received less attention in the economics of religion literature is whether the effect of religiosity derives through having a more religious social context or a direct effect of an individual's religiosity. Even the most convincing identification strategies, such as Gruber and Hungerman (2008), do not take the additional step of trying to separate an individual effect from the effect of social context.

A broad literature in psychology and sociology studies the link between religiosity, depression and other indicators of mental health (Koenig, 1998; Hackney and Sanders, 2003; Levin, 2010; Ellison and Henderson, 2011; Dein et al., 2012). Many empirical studies demonstrate a positive correlation between religion and mental health, but none of them have demonstrated a clear causal link between them (Hackney and Sanders, 2003). These recent overviews of the literature on religion and mental health support a need to better understand why religion improves mental health, and a number of studies consider why religiosity is linked to mental health problems based on correlational evidence (Ellison et al., 2001; Idler, 1987; Nooney, 2005). Economics brings a unique set of tools for helping to address the issues of causation that make it

[7] These theories are described in Ellison et al. (2001) and Ellison and Henderson (2011).

[8] This is different from county level instruments that are focused around the insight of church availability and competition, rather than social incentives. That peers affect religiosity is explored in Cheadle and Schwadel (2012) and Desmond et al. (2010).

difficult to disentangle how and why religion affects outcomes. Chiswick and Mirtcheva (2013) is the only paper we are aware of that studies the effect of religiosity on mental health in youth and treats the concern about selection into religiosity. They also find positive effects, but are not able to control for selection on unobservables. Becker and Woessmann (2011) also find a significant effect of religion on mental health and have a unique instrument for dealing with selection on unobservables, but in a very different context of 19th century Prussia and focusing on the question of Protestantism and suicide. Our study is also related to the growing literature in economics that recognizes the importance of non-cognitive aspects of child development for determining outcomes (Cunha et al., 2010; Cunha and Heckman, 2008; Heckman et al., 2006).

We find that religiosity has sizeable effects on depression in adolescence, which is understated by OLS estimates that do not deal with selection into religiosity. For example, a one standard deviation increase in religiosity decreases the probability of being depressed by 11 percent; or going to church one more time a month decreases the probability of being depressed by 3 percent. By comparison, increasing mother's education from no high school degree to a high school degree or more only decreases the probability of being depressed by about 5 percent. Our finding on the effects of religiosity on depression are robust to a number of specification checks that mitigate concerns about potential confounders such as unobserved shared influences, selection of peers, and simultaneity in choices.

We also find support for the theory that religiosity buffers against some kinds of stressors, and is particularly helpful when the adolescent lacks other support structures. The effects of religiosity are not driven by having school peers who are more religious, but may derive through improved psychological resources and coping skills for dealing with stress. These findings help inform contemporary policy debates about effective ways of addressing mental health problems among the young.

2 Data

We use data drawn from the restricted version of the National Longitudinal Study of Adolescent to Adult Health (Add Health).[9] Add Health interviewed a representative sample of U.S. adolescents in grades 7–12 (primarily aged 13–18) during 1994/95 academic year. A short in-school survey was conducted for every student in the sampled schools. Following the in-school survey, a random sample of students also participated in an in-home survey, which provides more detailed information about the child, including our primary variables of interest, religiosity and depression. This is supplemented with information about the child and his/her parent provided in the parent survey, and is based primarily on self-reports.[10] On average, there are 330 students per school who respond to the in-home survey.

Depression is measured on the Center for Epidemiological Studies Depression (CES-D) scale, one of the most common screening tests for depression and depressive disorder developed by Radloff (1977). The CES-D scale consists of a list of symptoms, to each of which respondents report how often they experience the feeling.[11] Responses are rated on a frequency scale ranging from 0 = never or rarely, to 3 = most or all the time. Response values are aggregated to create a point score, with higher scores indicating greater depressive symptoms. A score of 16 or above is considered to be indicative of depression (Radloff, 1977). Figure A.1 shows the distribution of the depression scale. The

[9]This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

[10]While there are additional follow-up waves, we focus on wave one. This is because we have only one additional year while the students are in school and the later wave does not include a parent survey.

[11]The original CES-D scale lists 20 items, only 19 of which appear in Wave I of Add Health. Add Health substitutes the CES-D item "You felt life was not worth living" for two questions on sleeping and crying spells. Appendix Table A.1 lists the questions.

distribution is skewed left with a long right tail and 24% showing symptoms of depression by this scale. While we primarily focus on the effect of religiosity on the CES-D scale, we also consider effects on the indicator of whether an adolescent is depressed by this definition, in order to get a better sense of magnitudes. While we use a binary depression variable initially, we then also exmaine how sensitive it is to the choice of threshold and to to alternative scales, as discussed in Section 4.2.2.

The data provides information on four aspects of religiosity: frequency of church attendance, importance of religion, frequency of praying, and frequency of attending youth religious activities. Each aspect is assessed on a scale of 0–3 or 0–4. We use the sum of these four aspects as our main measure of religiosity for our analysis.[12] Again, we assess the robustness of our results to alternative definitions of the religiosity variables. Previous literature suggests that it may be important to consider these measures separately (Iyer, 2016). Particularly, believing (measured through prayer and religious importance), which can be thought of as belief in religion privately, and belonging (measured through attendance) have been shown to have different types of effects on individual outcomes. This could easily be true in our setting as well. However, we find that these dimensions are not separable in our data.[13] A limitation of the data on religiosity is that only adolescents who report a religious affiliation were asked the more detailed religion-related questions. Therefore, we are only able to study the effect of religiosity on mental health for the religious affiliates, i.e., the intensive margin. We do however also include the non-religious sample when we test the robustness of our results, as shown in Table 7.

Table 1 describes our sample selection process. Non-responses to depression (column 2) and religious affiliation questions (column 3) constitute only a slight proportion of the full in-home sample (column 1). Less than 3% are dropped

[12] Although these values are ordinal, the three frequency variables for the most part approximately measure the number of times practicing each religious activity every month. The details are in Appendix A.1. We find similar results if we use an extracted factor as our variable of interest rather than our index of religiosity.

[13] Both a principal component analysis and exploratory factor analysis support a model where the different dimensions of religiosity load on a single factor.

from these selection processes.

Our identification strategy relies on defining a set of "similar"peers to which individuals are most likely to respond in choosing religiosity, based on children in the same school, grade, race, gender and denomination, as discussed further in Section 3. Among these characteristics, we consider peers of a similar religious denomination. This requires categorizing denominations. The in-home survey identifies 28 religious affiliations. We drop non-Christian affiliating, as they are arguably not largely substitutable across belief systems and no single affiliation has enough of a presence to be considered separately. The largest, Jewish, is only 0.7% of the sample. We group Christian faiths into Catholic, Liberal Protestant, Moderate Protestant, and Conservative Protestant, as based on the categorization in the Churches and Church Membership data which is associated with this survey.[14]

Individuals who report not having a religious affiliation are coded as having 0 religiosity. This assumption is necessary given that individuals who reported no religious affiliation were not asked the questions about their religiosity. For most of the analysis, we will focus on those who report a religious affiliation, though we include the non-religious in the calculations of the school social context. We discuss robustness to including the non-religious in our estimates of the effect of religiosity in Section 4.2.1.

Our main sample also drops non-Christians because they are a very heterogeneous group. We show robustness to including them in Table 7. After dropping non-affiliated and non-Christians, we are left with 81% of the whole sample (see column 4). The selected sample remains comparable to the whole sample, with only slightly higher religiosity and slightly lower depression. We also control for a range of covariates in our baseline specifications, taken primarily from the in-home and parent survey: individual characteristics such as age, sex, race, physical development, whether the respondent was interviewed

[14]The details of the categorization are summarized in Table A.2. The categorization is based on the Churches and Church Membership 1990 (CCM1990) data which collect county-level membership information on 133 Judeo-Christian church bodies in the US. Add Health categorizes these church bodies as Jewish, Catholic, Black Baptist, other liberal, other moderate and other conservative denominations in the Contextual Database.

during the school year session; parental background including whether mother or father was present, mother's education and household income in our baseline specification. Removing those with missing data on religiosity (column 5) and covariates (column 6) further reduce the sample by about 4.5%, but leads to trivial changes in depression and religiosity.

Our last step of sample selection is to exclude observations that do not have any peer respondent with the same school, grade, race, gender and denomination (column 7). This is needed to identify the effect of religiosity, as described in greater detail in Section 3, though we show robustness to including those with missing peer groups in Table 7. This leaves 62% of the full sample. In comparison, the selected sample are mentally healthier and more religious, but only marginally. Depression in the full sample is 11.39 compared to 11.10 in the selected sample. Religiosity is 8.49 in the full sample compared to 8.58 in the selected sample. Descriptive statistics of the final sample are summarized in Table A.3.

We observe considerable heterogeneity in depression and religiosity by race, denomination and family background in our sample. Table 2 examines depression by race, denomination, household income, and mother's education, as well as gender. On average, females report being more religious and more depressed than males. Blacks are the most religious ethnic group, while Hispanics are the least. Whites are the least depressed ethnic group, while those defined as other ethnicity (not white, black or Hispanic) are the most. Conservative Protestants are the most religious group by our measure, followed by Moderate and Liberal Protestants. Catholics are the least religious group. In terms of depression, Liberal Protestants suffer less than all three other religious denominations. There are small differences in religiosity by family background, but the differences in mental health are more pronounced, with disadvantaged children suffering much higher depression. In summary, if we look at the results by economic disadvantage, we see an ordering of depression and religiosity that might suggest positive selection into religiosity, i.e., that more advantaged children are more religious and less depressed. But, a similar ordering does not hold by race, where black students are more religious than

whites and more depressed. This suggests that selection into religiosity may follow less clear patterns than the selection we observe in schooling or other common areas of interest.

3 Empirical Strategy

Adolescent i's mental health (H_i) is determined by religiosity (R_i) and background characteristics (X_i),

$$H_i = \alpha_0 + \alpha_1 R_i + \alpha_2 X_i + \epsilon_i, \tag{1}$$

where ϵ_i denotes the residual.

The key concern with identifying an effect of religiosity is unobservable individual characteristics that affect mental health and make an individual more likely to be religious. For instance, religiosity may signal something about the home environment that affects mental health. Similarly, a shock, like the death of a friend or family member, could lead an individual to become more religious and also suffer from mental health issues. Reverse causality could also be a concern if individuals go to church as a way of dealing with poor mental health. It is thus ambiguous whether OLS estimates of equation (1) would over- or under-state the effect of religiosity and depends on the type of selection that dominates.

We address these endogeneity concerns using an instrument that arguably shifts an individual's religiosity independently of other individual background characteristics or individual-specific shocks that might affect mental health. The instrument we use is based on two ideas. First, friend religiosity affects adolescent choices of religiosity (Cheadle and Schwadel, 2012). Second, there is homophily in friendship formation (McPherson et al., 2001). Because friends are arguably selected based on unobservable attributes that are correlated with religiosity and mental health, they are not a valid exclusion. However, there exists plausibly random variation in the religiosity of "like" peers within schools that can be exploited to shift own religiosity independently of unobservable

individual background characteristics.

To formalize this, suppose $f(i)$ denotes friends of i and $\bar{R}_{f(i)}$ denotes average religiosity of friends excluding i. Consider a simple model where individuals choose religiosity and they care about mental health. In order to achieve the linear specification as above, assume utility takes the form

$$U_i = \gamma_1 \tilde{X}_i H_i - \frac{\gamma_2}{2} R_i^2 + \gamma_3 R_i \bar{R}_{f(i)},$$

where the complementarity in own and peer religiosity generates the incentive for conformity, a form also used in Brock and Durlauf (2001) and elsewhere.[15] $\tilde{X}_i = (X_i, v_i)$ denotes both observed and unobserved (to the econometrician) characteristics of the student. The residual in the mental health equation includes characteristics that are both observed and unobserved to the individual when choosing religiosity, i.e., $\epsilon_i = v_i + \eta_i$, where η_i is the shock to mental health (which is unobserved to the student at the time of choosing religiosity) and v_i is the observed component. Individuals make decisions simultaneously based on their information sets, Ω_i, which includes X_i and v_i and characteristics of peers in a way that will be made specific below. In this case, we can write down the individual's best response as

$$R_i = \frac{\alpha_1 \gamma_{1X}}{\gamma_2} X_i + \frac{\gamma_3}{\gamma_2} E(\bar{R}_{f(i)} | \Omega_i) + \frac{\alpha_1 \gamma_{1v}}{\gamma_2} v_i. \tag{2}$$

We assume that the religiosity that we see in the data is a result of optimizing behavior, and we omit $*$'s here for notational simplicity, though in reality we should distinguish between realized mental health outcomes that come from optimizing behavior and the production function of hypothetical outcomes.[16] Note that equation (2) suggests that average friendship religiosity may be a plausible exclusion for shifting own religiosity independently of unobservable characteristics v_i that cause R_i to be endogenous in the mental health equation. There are two key concerns with using this as an exclusion. First, if

[15] Note that it is trivial to introduce a direct utility of religiosity.

[16] Given that religiosity is bounded and the model is linear, we know that an equilibrium exists and is unique in this model.

individuals observe their friends' v's at the time of choosing religiosity, i.e., $\Omega_i = (X_i, v_i, X_{f(i)}, v_{f(i)})$, there is a simultaneity concern in that peer average religiosity reflects v_i. Second, though not modeled, friendships themselves are likely to be endogenous and may be determined by v_i and $v_{f(i)}$. Intuitively, the friendships of an individual who is prone to depression may look systematically different than one who is not. This is problematic when this is correlated with religiosity, for instance, if church attendance makes it easier to find friends.

We can use instead the average religiosity of "like" students, i.e., students at the same school, in the same grade, race, gender and religious affiliation, denoted $g(i)$. Previous work, including McPherson et al. (2001) and Currarini et al. (2010) highlight sorting on race, gender and age. Along with these dimensions, we also find in the data friendship sorting based on religious affiliation. Therefore, $\bar{R}_{g(i)}$ is correlated with $\bar{R}_{f(i)}$ given homophily, but not with unobservable individual level attributes that might determine religiosity, v_i (after conditioning on the student's own grade, race, gender and religious affiliation).[17] Furthermore, simultaneity at this level is less likely to be a concern, and we describe a number of specification checks in Section 4.2 that support this.[18]

Note that a key concern with this strategy, as in the seminal work of Hoxby (2000), is whether this variation in peer groups is plausibly random, something that we return to in Section 4.2. Intuitively, this argument is only likely to hold within schools. Some schools may have more religious students because they are in a neighborhood with more churches or a particularly influential church. The provision of mental health support at the school level, for instance, might also vary depending on the resources in the community, such as the number of churches. Thus, it is important for our strategy that we also control for school fixed effects to eliminate these potential biases.

[17] Note that Patacchini and Zenou (2015) use a similar strategy of "like" peers to instrument for friend religiosity, though with a different purpose, to identify the effect of friend religiosity on parental investment in religiosity.

[18] We expect simultaneity to be less of a concern with larger peer groups $g(i)$. We check robustness to dropping observations where the subgroup is small, just to be sure that this is not driving our results. Furthermore, we find that the estimated effects of peer religiosity in the first stage are not significantly larger with larger peer groups.

With these underlying mechanisms in mind, we estimate the following baseline model:

$$H_{is} = \alpha_0 + \alpha_1 R_{is} + \alpha_2 X_i + \alpha_s + \varepsilon_{is}, \tag{3}$$
$$R_{is} = \beta_0 + \beta_1 \bar{R}_{g(i)s} + \beta_2 X_i + \beta_s + u_{is},$$

where the s subscript denotes the school, α_s and β_s school fixed effects.

There are two remaining concerns with the identification strategy: (1) a direct effect of peers on mental health and (2) unobserved shared group characteristics that are correlated with peer religiosity and mental health.

Peers may directly affect mental health, either through their religiosity or mental health (which is determined in part by their religiosity). In this case, our instrumenting strategy would not identify the direct effect of an individual's religiosity, but the effect inclusive of peer religiosity on mental health. As far as we know, this is a characteristic that is shared by all the instrumenting strategies used to identify the effect of religiosity, it is just made more explicit in our context. For instance, Gruber and Hungerman (2008) have one of the most convincing identification strategies for studying the effect of religiosity. They use changes in Blue Laws, which ban shopping on Sundays, to identify an effect of religiosity on different outcomes. The argument follows that by changing the outside options for an individual, this would affect church attendance of that individual. Implicitly, this is also an equilibrium argument, as these laws affect whether everyone in the community goes to church on Sundays, and so any estimated effects of religiosity would be inclusive of peer religiosity and associated peer outcomes, like mental health in our context. Arguably, the effect of religiosity inclusive of social context is also of policy interest. However, we describe below assumptions that would make our instrument valid for identifying the direct effect of religiosity.

Suppose peer mental health has a direct effect on i's mental health. Then peer religiosity is no longer a valid exclusion (through its correlation with peer

mental health) unless we condition on peer mental health, i.e.,

$$H_{is} = \alpha_0 + \alpha_1 R_{is} + \alpha_2 X_i + \alpha_3 \bar{H}_{g(i)s} + \alpha_s + \zeta_{is}. \tag{4}$$

However, if individuals take into account their effect on peer mental health, we may introduce an additional problem of simultaneity of own and peer mental health, which would bias up our estimates of α_3. Similarly to our discussion of potential simultaneity in religiosity, we do not expect this to be as much of a concern at the level of peer group that we have defined, particularly after we exclude friends. But, we discuss this further in the robustness checks in Section 4.2. However, if α_3 is close to 0, this would suggest that the true effect of religiosity comes through a direct effect of the individual's religiosity.

The more challenging case is if $\bar{R}_{g(i)s}$ has a direct effect on mental health. Then, peer religiosity is not a valid exclusion for identifying a direct effect of own religiosity even conditional on peer mental health, absent strong assumptions on the endogeneity of religiosity. In this case, we need to be open to the interpretation of our findings as an effect of religiosity inclusive of having a more religious social context, as estimated elsewhere in the literature. However, we can test whether the social context is likely to play an important role by controlling for peer covariates. For instance, if we know that children of better-educated parents are more religious, we would expect to see that the percentage of peers who have better-educated parents should matter for depression if peer religiosity has a direct effect on depression. A similar argument holds for peer mental health. Given these arguments, among the robustness checks we will see whether the marginal effect of peer mental health and peer characteristics are non-zero and whether the estimated marginal effect of religiosity is robust to controlling for these characteristics.

The remaining concern is whether there is an unobservable third factor that simultaneously predicts peer religiosity and own mental health; this is an example of a correlated effect, in the language of Manski (1993). To be a threat to identification it would need to vary at the group level within the school (so that it is not controlled by the school fixed effect) and be correlated with (but

not determined by) peer religiosity.[19] We check for these potential confounders in a number of ways, as described further in Section 4.2. For instance, we check the robustness of our results to using peer group characteristics which are predetermined and so therefore cannot be determined by an unobservable third factor (rather than religiosity) as instruments for religiosity. Furthermore, we show that controlling for a peer group fixed effect in the non-instrumented regression explains little of the correlation between religiosity and depression. Furthermore,

4 Results

4.1 Baseline Results

In Table 3 we present the results for the OLS and IV estimation of the relationship between mental health and religiosity. In all specifications, we control for individual characteristics, family background, grade dummies, and school fixed effects. We start with the baseline specification in column (1) which does not instrument for religiosity. These results suggest that religiosity decreases depression by -0.16. Conditional on other covariates, Hispanic and other ethnicity students are significantly more depressed than white students. Religious denomination does not seem to play a significant role in determining mental health, except that liberal Protestants are less depressed than Catholics. Older students are more depressed, while physical development is negatively correlated with depression for boys but not girls. Adolescents are consistently mentally healthier during holidays relative to school term-time, suggesting either seasonal effects or a role of school stress. Family background seems an influential factor in determining adolescent mental health. Not living with father is associated with higher depression. Mothers with more education have children with lower depression. Household income is not predictive of mental health, conditional on other household characteristics. This could be because

[19] Note that if it is determined by peer religiosity it is part of the social context of having peers who are more religious.

of measurement error in income and that 25% of the sample does not report income.[20]

As discussed above controlling for school fixed effects helps eliminate concerns about fixed factors at the school or community level that might predict both religiosity and mental health. For example, the provision of mental health support at the church level may depend on the provision at the school level, creating correlations between the average religiosity of the school and the mental health of adolescents attending the school. School fixed effects also help control for differences at the community level in the availability of churches or mental health care. Results that do not control for school fixed effects (not reported) are surprisingly similar, with estimates of -0.15 for the effect of religiosity rather than -0.16 with school fixed effects. This suggests that fixed characteristics of the school that determine mental health are not correlated with the adolescent's religiosity in ways that bias our findings.

Column (2) presents results when we instrument for religiosity using the average religiosity of same grade, gender, race and denomination peers, and column (3) shows the first stage results. First, note that peer religiosity is significant and positively predicts own religiosity, with an F-statistic of 30.44, suggesting that we do not have a weak instrument problem. The estimated effect of religiosity on depression using our IV estimator is -0.70, over four times as large as the OLS estimates of -0.16, and it is statistically significant at the 5% level. In standardized terms, this indicates that a one standard deviation increase in religiosity leads to a 0.31 standard deviation reduction in the depression scale. That the IV estimates predict more negative effects of religiosity than OLS suggests there may be negative selection into religiosity, *i.e.,* more depressed adolescents participate in more religious activities, biasing OLS toward zero. One explanation for this selection is that adolescents may choose religion as a way of coping with depression or other difficult home circumstances that are correlated with depression. This is consistent with evidence in Maselko et al. (2012) and Ferraro and Kelley-Moore (2000), which

[20]We code missing values of log household income to 0 and include a dummy variable for not reporting household income so that we do not drop these observations.

show that some health problems lead to increased religiosity.[21] An alternative interpretation is that IV and OLS results may not be directly comparable if there is heterogeneity in the effect of religiosity on mental health, as OLS estimates the average treatment effect and IV a weighted local average effect for those adolescents whose religiosity is affected by their peers. We return to consider heterogeneity in treatment effects in Section 4.3.

The first stage results are of interest in their own right. We see that, conditional on other covariates, Conservative Protestant adolescents are the most religious, followed by Moderate Protestants. Catholic and Liberal Protestants do not differ in statistically significant ways. Also, black, Hispanic and other ethnicity adolescents are all more religious than whites. Adolescents whose mothers have a college degree or above are more religious than those with less educated mothers. Finally, adolescents whose fathers are not present at home are less religious.

To get an idea of the magnitude of these effects, we consider an indicator of whether the adolescent is depressed as an alternative dependent variable.[22] Columns (4) and (5) present OLS and IV results from the linear probability model respectively. Comparison between these two sets of results shows again that IV estimates predict more negative effects than OLS. Column (6) reports the average marginal effects from an IV probit model.[23] The estimated effects of religiosity in columns (5) and (6) are similar, suggesting that being one unit more religious decreases the probability of being depressed by 3% on average.[24] A one standard deviation increase in religiosity (or 3.3 units) decreases the probability of being depressed by 11%.

[21] Mellor and Freeborn (2011) also find that IV is higher than OLS estimates of the effect of religiosity on risky behavior.

[22] The rule of thumb for this is whether the CES-D exceeds 15 (Radloff, 1977).

[23] In the probit model, we control for school fixed effects using school dummies, though this is not consistent.

[24] A one unit increase in religiosity would for instance mean going to church one more time a month.

4.2 Potential Threats to Identification

In this section we check the robustness of the estimates to a number of potential threats to our identification strategy as discussed in Section 3.

One key concern with the proposed instrumental variable is that students may select peers based on religiosity, so that peer religiosity, as measured at the group level, may reflect other unobservable attributes of the student. School fixed effects control for selection into schools based on fixed characteristics of the peer group. Dating back to Hoxby (2000), the literature often exploits random variation within schools to identify peer effects. The idea is that while individuals may select schools and friends, the variation in peer composition across grades within schools is plausibly random variation that can be exploited. The resemblance with the typical peer effect specification in the literature can be made clear by considering the reduced form equation,

$$H_{is} = \delta_0 + \delta_1 \bar{R}_{g(i)s} + \delta_2 X_{is} + \gamma_s + \mu_{is}, \tag{5}$$

where $\mu_{is} = \epsilon_{is} + \alpha_1 u_{is}$. In our case, the random variation in cohort composition across grades within schools creates variation in average religiosity at the group level.

Comparable to other studies that use random variation in peer composition across cohorts, we check this assumption using balancing tests, to see whether peer religiosity predicts observable individual characteristics. The added complication in our context is that instead of just using variation across grades within schools, we are also using variation across gender, race, and denomination. The balancing tests should hold conditional on the full set of gender, race and denomination dummies that define the peer group and that we condition on in the main regressions. For instance, Hispanics are more religious, and they also have peers who are more religious by our definition. Hispanic is also correlated with lower income. Therefore a regression of income on average religiosity of same-race peers that did not control for individual race dummies would find (for the case of Hispanic students) that peer religiosity is negatively correlated with individual income by construction. The variation

that we isolate by controlling for the full set of gender, race and denomination dummies is instead random variation in the average religiosity of "like" peers within schools across grades.[25]

Note that one way that this test might fail is if parents select schools based on the average religiosity of specific cohorts of students, which would not be controlled with a school fixed effect. Another reason that balancing tests might fail is if denomination is endogenous, so that individuals select their own denomination to better match their own religiosity to the religiosity of students in the school. This is less likely to be a concern as our definitions of denomination are fairly broad, and furthermore Smith et al. (2015) show that individuals are more likely to change religious affiliation in young adulthood rather than adolescence. A final reason that these balancing tests might fail is if there is simultaneity in religiosity at the group level. We would expect that if any of these are problematic, we would see some evidence of it in terms of the observable characteristics that predict religiosity and mental health being correlated with peer religiosity.

Table 4 shows the results of these tests. Out of nine indicators for adolescent and family background characteristics, only one variable, mother not being present, seems to be correlated with peer religiosity and the size of the correlation is very small, at -0.002. Thus the observable covariates seem to be well balanced between adolescents facing peers who are more religious and those facing peers who are less religious, conditional on the group dummies. Though we cannot rule out selection of peer religiosity or simultaneity in peer religiosity based on unobservable characteristics, this provides supportive evidence that in terms of observables the assumption of random variation in peer religiosity appears to be valid.

In Table 5, we provide further tests to show that potential selection and simultaneity are not biasing our estimates. Given that simultaneity is more likely to occur at the friend level, in column (1) we exclude reported friends from the calculation of peer religiosity. The estimated effect of religiosity is

[25]Note that results are also robust if we control for the interactions of gender, race and denomination at the individual level.

robust. In column (2), we remove private schools from the analysis, as these are the schools that are most likely to be selected based on religiosity. Results again are very similar. In column (3), we replace the adolescent's denomination with the parent's denomination as both a control and to define the relevant peer group for the instrument. Parent's are even less likely than adolescents to choose denomination based on the adolescents' peers, so it provides a useful test for ruling out potential endogenous denomination choices. Results are still robust, though a bit noisier because of the smaller sample size.

Lastly, we consider a couple of overidentified versions of our model. In column (4) we allow individuals to be influenced by peers of the same school, grade, denomination, but opposite gender, as well as peers of the same gender. These results show that own religiosity is affected by both same-gender and opposite-gender peers, but relatively more by same-gender peers. The second-stage results are similar to those obtained from using only one instrument in Table 6. Assuming validity of one instrument, the over-identification tests show that we cannot reject validity of the other instrument, providing further support for the strategy. In column (5) we then consider using both same- and cross-denomination peers. Again results are similar and the test of overidentifying restrictions supports that the additional instrument is not endogenous. These results are also interesting as they show the the main peer effects of religiosity derive through same-denomination friends.

Given selection and simultaneity in peer religiosity do not seem to be a concern, it remains to disentangle whether the estimated effect of religiosity derives through the social context of peers who are more religious (and associatedly less depressed) or through a direct effect of an individual's own religiosity on mental health. First, in column (1) of Table 6, we check that our results are not driven by school contextual variables that vary across grades and are used to define our subgroups, including the percentage female, the percentage belonging to different racial subgroups and the percentage belonging to different denominations. None of these are individually or jointly significant in determining mental health. Most importantly, this does not affect our estimate of the effect of religiosity on mental health. In column (2), we add

in controls for peer characteristics at the subgroup level. Note that if peer mental health or peer religiosity were important direct determinants of mental health, we would expect to see that some of these observable characteristics of the peer group matter, particularly the ones that are relevant at the individual level for determining mental health and religiosity. A similar argument holds if there are unobserved shared group characteristics that jointly determine religiosity and mental health. However, none of these peer characteristics are individually or jointly significant and controlling for them does not change our estimates of the effect of religiosity.

In columns (3) and (4), we control for peer depression, both alone, column (3), and with other peer characteristics in column (4). Recall from the discussion in Section 3 that peer mental health may be biased upward due to simultaneity. The coefficient on peer depression is close to 0 in both cases, suggesting that simultaneity is unlikely to be a driving concern. We also see that peer characteristics in column (4) remain jointly insignificant, providing additional support that peer unobservable characteristics are unlikely to be driving the link between religiosity and depression. In all cases, our estimated effects of religiosity are similar.

Despite the robustness of our results to different contextual variables, there may be remaining concerns about unobserved shared group effects. A particular type of this shared group effect could come from the presence of an influential local church which may encourage greater religiosity for students in a given denomination and also positively affect mental health. Already the similarity of our basic results with and without school fixed effects suggest that this may be unlikely. However, the school fixed effects do not control directly for these effects as the effect of a church would likely vary depending on the race and denomination of the student. We check that this is not driving our results by controlling for average religiosity of same-denomination peers and same-race peers. The latter helps deal with the fact that church attendance is often segregated along racial lines. The results in column (5) suggest that neither average race or average denomination religiosity predicts mental health, conditional on own religiosity, and the effect of own religiosity remains robust.

However, the average religiosity of the same denomination peers is a strong predictor of own religiosity and does weaken the first stage, though the F-statistic remains strong at 14.5. Results are similar with and without controls for average peer depression. Finally, column (6) addresses concerns about selection into schools based on trends in average same-denomination religiosity by controlling for same-denomination religiosity interacted with grade. Results are again similar.[26] Together these results provide strong support that unobserved factors at the denomination level are not biasing our findings.[27]

Despite the evidence that controlling for peer depression and peer characteristics do not affect our estimates of the effect of religiosity on depression and do not appear to be correlated with depression, there may still reasonably be concern that unobservable characteristics at the peer group level may be driving our results. Recall that these characteristics are problematic if they are correlated with peer religiosity but not caused by peer religiosity. To address this concern, we also use peer group characteristics rather than peer religiosity as an alternative instrument for religiosity and find estimates of the effect of religiosity that are similar in magnitude. This provides useful supportive evidence that an unobserved shared factor is not driving our findings given that these alternative instruments are predetermined and so could not be affected directly by this type of unobservable.

An additional test we perform for this is to rerun our baseline results in column (1) of Table 3 with peer group fixed effects, i.e., fixed effects at the same school, grade, race, gender, denomination level. This absorbs all group

[26] Note that in this last regression we do not control for peer characteristics. When we do include these as controls the point estimates of religiosity are similar but the p-values are just about 0.1. Given that peer characteristics are neither individually or jointly significant determinants of depression and we only have one year of data from which to estimate trends, we think controlling for peer characteristics along with average same-denomination trends in religiosity is asking a lot of the data, so do not find this surprising.

[27] Out of concern that there may be racial segregation across churches, so that for instance black and white students of the same denomination may face different church influences, we also attempt a specification where we control for average religiosity of the same school, race, denomination peers. In this case, there is again no effect of average same school, race, denomination religiosity on depression, suggesting this type of unobserved group effect is not a concern. However, the first stage loses power because it is a strong predictor of own religiosity.

level unobservables. Note that we cannot use our instrumenting strategy with the group fixed effects because that is that level at which our instrument is varying. That said, we can compare our estimates of the effect of religiosity in the model with school fixed effects to the model with group level fixed effects. If group level unobservables are important, we would expect that our estimates of religiosity would be significantly smaller. In fact, we find that our estimated effect of religiosity drop from -0.16 with school fixed effects, as reported in Table 3, to -0.14 with group fixed effects. This is a very small difference, particularly given the size of the standard errors. Furthermore, this test likely overstates the importance of the group level unobservables to the extent that these group fixed effects might also help control for individual level unobservables that are correlated with religiosity and matter for depression.

4.2.1 Sample Selection

In Table 7, we further test how sample selection affects our results. Column (1) is comparable to the results in column (2) of Table 3, but with the addition of 77 observations that were missing some peer characteristics and were dropped from the main sample. Column (2) adds in the "other" religion subgroup. Again, results are similar with an estimated effect of religiosity of -0.62. Column (3) attempts to deal with the problem of dropping observations for individuals due to missing peer groups. For these individuals, we assigned the peer religiosity at the school-grade-gender-denomination level, if available and if not at the school-grade-race-gender level.[28] These modifications incorporate most of the students who report a religious affiliation, 15,961 out of a total sample of 16,169 whose other relevant variables are not missing. In the specification, we also include a control for the students who are missing observations of school-grade-race-gender-denomination peer average religiosity and allow for the effect of the peer religiosity to be different for these students. The first stage (not reported) shows that the main effect of peer religiosity is 0.11 and this is reduced to about 0.04 for the subgroups where we do not observe

[28]Results are comparable if we replace missings first with school-grade-race-gender average religiosity and then school-grade-gender-denomination average religiosity.

peer religiosity at the school-grade-race-gender-denomination level, so our instrument is much weaker for this subgroup. That said, the estimated effects of religiosity with this bigger sample is still sizable, though smaller at -0.43. Students who are missing peer religiosity appear to be more depressed than the rest of the peer group, though this is not statistically significantly different from 0. Furthermore, we show that we pass the test of over-identifying restrictions, which provides further support that unobservables about these students with missing peer groups do not present additional endogeneity concerns.

A final sample selection concern is the exclusion of the non-religious from the sample. Ideally, we would like to find an instrument that shifts whether a student reports a religious affiliation, the extensive margin, as well as religiosity so that we could jointly estimate the selection into religion and religiosity. We tried a number of instruments based on within-school peer variation, including the percentage of peers that are non-religious using different definitions of peer groups and allowing for higher order terms. We could not find a robust predictor of whether a student was religious or not. One interpretation of this is that peers do not directly affect the choice to be religious, which is in line with previously cited work by Smith et al. (2015) showing that most transitions in religious affiliation occur in young adulthood rather than the teenage years. Thus, instead we treat whether a student is religious to be exogenous and include the non-religious in the regression, with a control for being non-religious and defining peer religiosity for these students at the school-grade-race-gender level.[29] This increases the sample to 18,137 out of a total possible sample of 18,420. The estimated effect of religiosity is robust at -0.48.

Overall, these results provide suggestive evidence that the effect of religiosity is bigger for the Christian sample, though the estimates in the larger sample are still sizable. We explored the extent to which the results for Christians might be driven by larger peer groups providing a more supportive effect of religion. We test this by interacting peer religiosity with the size of the peer group, the percentage of peers of the same race-gender-denomination in

[29]Note that if we define religiosity at the denomination level, peer religiosity is 0 and perfectly predicts own religiosity.

the student's grade in separate specifications. The estimated effect of peer religiosity on own religiosity is not significantly larger when the student has a larger peer group. Furthermore, the estimated effect of religiosity is not significantly different for students with larger peer groups (either in levels or percentages) and the effect of religiosity remains at around -0.7 independent of the size of the peer group. This provides further supportive evidence that our estimates are unlikely to be driven by social effects.

4.2.2 Scale

While the CES-D 20 is a well-recognized, validated scale, we remain concerned about the extent to which the scale is driving our results. In Table 8, we check robustness to adjustments to the scale. In column (1), we standardize the scale to have mean 0 and standard deviation 1 to offer a point of comparison across scales. In column (2), we remove 3 questions from the CES-D that are more social in nature, including "You felt that you were just as good as other people", "You felt that people disliked you", and "People were unfriendly to you". The estimated effect of peer religiosity remains similar. This helps support that our estimates are not driven by reference effects or social effects. Column (3) uses a reduced scale based on the CES-D 10, another well-recognized scale in the literature. The CES-D 20 includes 8 of the 10 questions in this scale and most notably does not include the questions that are more social in nature.[30] The estimated effect of religiosity drops sligly from -0.09 to -0.08 with the 8 questions. The choice to assign equal weights to the different questions was also arbitrary. Columns (4) to (6) include the same specifications in columns (1) to (3), except extracting a factor from the different questions included in each of the scales, thus allowing the data to tell us which questions are more indicative of depression. Again, results are similar, though slightly smaller in magnitude, ranging from -0.07 to -0.05.

An additional concern is that results could be driven by reference effects. For instance, it could be that individuals who have more religious peers report being less depressed simply because of their reference group. Already,

[30]The included questions are numbers 1, 5, 6, 8, 10, 11, 13, 18, as listed in Table A.1.

the results in table 6 suggests that this may not be important, given that controlling for peer depression does not affect estimates of the effect of religiosity. Table 9 provides further support by controlling for peer depression at various levels to check for reference effects at different levels, including the school-grade, school-race, school-gender, school-denomination, school-grade-race, school-grade-gender, school-grade-denomination, school-race-gender, school-race-denomination and school-grade-race-gender. The estimated effects of peer religiosity are remarkably robust across these specifications.[31]

Finally, Table 10 shows the robustness of the estimated effects of religiosity on being depressed, using different cutoffs for depression and both the linear probability model (Panel A) and the IV probit model (Panel B). The estimated effect of religiosity on depression at this standard cutoff, as previously reported in Table 3 and column (2) of Table 10 is -0.03. The estimated effect of religiosity is even stronger at higher cutoffs, with a high of -0.05 at a cutoff on 18 and drops to -0.03 (linear probability model) or -0.04 (IV probit model) at a cutoff of 24. For cutoffs below 16, the estimated effect of religiosity is smaller and not significantly different from 0. This seems reasonable given that the cutoff of 16 identifies 24% of students as depressed, which seems overly generous. We provide further evidence of non-linear effects of religiosity in section 4.3.

4.3 Heterogeneity in Effects

The effects of religiosity may vary depending upon the individual's unobservable propensity for being depressed. We estimate how the effects of religiosity differ across the conditional quantiles of the depression index, using a version of the two-step control function approach, as developed in Imbens and Newey (2009) and Lee (2007). We estimate the first stage regression as before, but obtain the residual from this regression rather than the predicted value of religiosity. We then include the residual as an additional regressor in our sec-

[31] Note that in some cases the peer depression takes an unexpected negative sign. This is a mechanical result that occurs when there are few groups per school, as in the case of school-gender in particular and school-race, on top of school fixed effects.

ond stage regression to control for the endogeneity of religiosity and estimate the second stage as a quantile regression.[32] Figure 1 shows that the effect of religiosity is higher for people who are more depressed — comparing the 0.1 quantile to the 0.8 quantile, we see that the estimated effect of religiosity increases from about -0.26 to -1.47.[33]

It is interesting to compare our findings to the alternative findings on the effectiveness of clinical treatments for depression. Evidence on psychotherapy, and particularly cognitive based therapy (a primary method of treatment for depression in the United States) is generally accepted to be effective for mild to moderate depression (Gloaguen et al., 1998). There seems to be a broad consensus that more severely depressed individuals may need a combination of psychotherapy and antidepressant medication (TADS, 2007), as suggested by the guidelines posted by the National Institute for Mental Health. That psychotherapy alone is less effective for the severely depressed then offers an interesting contrast to the role of religiosity in these contexts.

We also explore nonlinear effects of religiosity on mental health based on how religious the individual is. Some studies argue that the effect of religiosity on mental health is U-shaped, with average religiosity individuals being hurt and those with high or low religiosity being helped (McFarland, 2010; Schnittker, 2001). Others have argued that the effect is reverse U-shaped (for instance, see Eliassen et al., 2005). Part of the theory underlying this is that individuals on either extreme of religiosity may be more at risk of mental health problems, whereas those in the middle have the potential to benefit the most. We test this using a control function approach and try a number of different specifications of polynomials in religiosity. We find little evidence of

[32]There is no accepted way in the literature for incorporating fixed effects into quantile models. We report results that predict the school fixed effects from the mean 2SLS regression and then control for these in the quantile regression. Standard errors are block bootstrapped at the school level. Estimates are qualitatively similar if we instead include school dummies.

[33]The estimates at the 0.9 quantile (not pictured) are even larger, -2.4, but not statistically significantly different from 0, likely because of the possibly large disparities at this quantile in the severity of depression. This is also consistent with the literature on depression, which struggles with recommendations for treating the most severely depressed, as discussed below.

heterogeneity by degree of religiosity.[34]

5 Mechanisms

Ellison and Henderson (2011) discuss how a stress process model might explain the link between religiosity and mental health, based on a synthesis of the existing literature. They highlight several different mechanisms through which religiosity can affect mental health. First, religiosity may affect psychological resources, such as self-esteem, which may lead to better mental health (Smith et al., 1979). Second, religiosity might help provide coping tools for dealing with stressful life events (Sherkat and Reed, 1992). For instance, it may reduce the extent to which people engage in active problem solving in response to a stressful situation by encouraging a more fatalistic attitude. Third, religiosity might reduce exposure to stressors that can be linked with depression, for instance, by helping to foster more stable home environments. Fourth, religiosity may provide alternative support structures, such as helpful friendship or direct financial assistance, which help individuals deal with stressful situations in healthy ways.

A few other papers have studied directly the potential for the stress process model to explain the link between religiosity and mental health. Nooney (2005) highlights the role of stressors, such as school stress and health stresses, as well as perceived support and self-esteem as mediating the relationship between religiosity and mental health. Eliassen et al. (2005) find that social support and stress exposure largely explains the relationship between religiosity and mental health. Causality remains a concern however, as it is difficult to disentangle the role of religiosity and stressors from selection. We hope to add to this discussion by isolating a causal channel.

[34] One potential concern is whether this could be a result of the instrument we are using, in that peer religiosity does not shift over the full distribution of religiosity. To test this, we also estimate a quantile regression version of the first stage and find that peer religiosity has significant effects on all but the most religious (0.9 quantile of the conditional religiosity distribution), which is likely due to a ceiling effect. The estimated effects of peer religiosity are also fairly homogeneous across the conditional quantiles.

5.1 Psychological resources and coping tools

Self-esteem is one focal point in the literature on psychological resources that can help individuals cope with stress in healthy ways. Psychologists hypothesize that self-esteem can develop through the positive regard of others one holds in esteem. The church community can play a role in this, either positively or negatively, by imposing a different value system than adolescents experience in school, i.e., valuing moral integrity over scholastic achievement. Furthermore, it is hypothesized that relationship with a divine other may help provide a sense of worth. Importantly, the arguments for why religiosity could support self-esteem could also be turned to suggest reasons that religiosity could hurt self-esteem. For instance, relationship with a divine other that is seen largely as punitive could plausibly hurt self-esteem (Ellison and Henderson, 2011).

We consider whether religiosity affects self-esteem using an index based on 4 questions in the Add Health, which parallels Rosenberg's global self esteem scale that is widely used in the literature (Rosenberg, 1989; Nooney, 2005) and are detailed in Appendix Table A.4. The first 2 columns of Table 11 consider the effect of religiosity on self-esteem. Column (1) shows that consistent with the literature described in Ellison and Henderson (2011) religiosity is positively correlated with self-esteem using an OLS regression. Column (2) shows that when we instrument for religiosity to control for selection and potential reverse causality, the estimated effect of religiosity increases from 0.075 to 0.15. The standard errors are fairly large so that our IV results are not statistically significantly different from zero. Given the size of the coefficient, one interpretation of this could be that religiosity matters for self-esteem, but the effects vary across individuals; this makes sense given the wide variety of religious experiences.

A second related theory is that religiosity affects how people cope with difficult situations or problems. Pargament and Brant (1998) provide support of this, based on a detailed survey of the literature. For instance, different scholars have suggested that religion can lead one to engage in more passive problem-solving, in part by inspiring a more fatalistic perspective on life. We use the definition of passive problem solving in Nooney (2005) to capture this,

which is an index of several self-reported measures of how adolescents approach problems, as described in detail in Appendix Table A.4.

The second set of results in Table 11 show that OLS estimates of the effect of religiosity on passive problem solving are positive and significant. IV estimates again are larger, but not statistically significantly different from zero. As in the case of self esteem, one interpretation is that religiosity has an effect on passive problem solving, but estimates are noisy given heterogeneity in effects.

The final 3 columns show what happens to our estimated effect of religiosity on depression when we control for these measures of psychological resources and coping skills. Column (7) shows that controlling for self esteem and passive problem solving reduces the estimated effect of religiosity on depression to -0.41 (from around -0.70 in other estimates). The estimated effect of religiosity is no longer statistically significant though the point estimate is still sizable. Furthermore, the strong F-statistic of 31.3 suggests that our first stage still has power when we control for self-esteem and passive problem solving. Both passive problem solving and self-esteem help reduce depression.

Together these findings suggest that the effect of religiosity could derive through psychological resources and coping skills. These results are particularly interesting, given the possibility that some aspects of psychological resources and coping skills that matter for depression may not be adequately captured by our measures.

5.2 Stressors

There is a considerable literature which suggests that religiosity reduces exposure to stressors that may be correlated with mental health. In the case of adolescents, who may be transitioning from early family life and experiencing stress or distress, the anchor that religious commitment provides may help them deal better with negative influences such as anger or conflict, which are thought to emerge from a lack of trust within the home and established family routines (Eliassen et al., 2005, p. 189). Divorce, domestic violence and

chronic health problems are some types of stressors that the literature links to religiosity (Ellison and Henderson, 2011).

We consider a broad set of potential stressors for adolescents and present in Table 12 a subset selected based on whether we find them to be correlated with depression — GPA, whether a family member or friend has committed suicide in the past 12 months and general health.[35] Columns (1)–(3) show the instrumented effect of religiosity on each of these stressors. In none of these cases, does religiosity appear to have a causal effect, suggesting that religiosity does not reduce exposure to these types of stressors.

Columns (4)–(6) then consider whether there is evidence of stress-buffering effects of religiosity. If religiosity does provide better ways of dealing with stress as evidenced in the previous section, we would expect to see that more religious adolescents respond less to the stressor, as captured by interacting religiosity and the stressor in the depression regression. We instrument for religiosity and the interaction of religiosity and the stressor using our measure of peer religiosity and peer religiosity interacted with the stressor.[36] We find that the stress-buffering hypothesis does seem to hold for the suicide of someone close to the adolescent and general health, but not for GPA. This effect could derive through the improved psychological resources as described in the previous section or through improved support structures, which we consider next.

5.3 Support Structures

Another hypothesis is that religiosity provides alternative support structures to deal with stressful situations, often referred to as social resources in the literature. There is some evidence in the literature supporting this hypoth-

[35] We consider a number of other stressors in the literature that seem also applicable in our setting, including parental divorce, whether the parents fight, whether parents have other marriage difficulties or financial problems, but these are not significantly related to depression.

[36] Note that this is easiest to interpret when the stressor is exogenous, which may not be plausible here. Bun and Harrison (2014) describe conditions under which the interaction can be interpreted as exogenous even if the stressor itself is endogenous. The key condition in our context is that the covariance of peer religiosity and the unobservable determinants of mental health do not vary systematically with the stressor.

esis. Ellison and Henderson (2011) discuss how religious congregations offer financial aid and other tangible services, along with direct counsel on how to deal with problems and informal networks that provide support during difficult times. Bradley (1995) shows that there is a positive relationship between more frequent church-going and the size of one's social network, the frequency of contact by telephone and in-person, the support received and the perception of the quality of those supportive relationships.

We do not have data on the churches students attend, hence we cannot test the hypothesis of churches providing support directly. However, we test this hypothesis indirectly by considering whether adolescents who have less support in other key places, like in the home, school or neighborhood, experience larger effects of religiosity. In Table 13, we consider three indicators of these types of support structures that are correlated with depression — whether the adolescent is from a single parent home, protective factors that include questions related to how much the adolescents feels they are cared for (see appendix Table A.4) and an index of neighborhood resources (see appendix Table A.4) indicating how much people in the neighborhood know and look after each other.[37] The interaction is significant for the case of coming from a single parent home and for protective factors and supports the theory that religiosity matters more when other support structures at school and in the home are weaker. However, as in the previous table, this could also be indicative of better coping or psychological resources associated with religiosity.

One related hypothesis that we consider is whether religion confers the same benefits as participation in any sort of club, through a sense of belonging and associated social support (Michaelson et al., 2014). If this is the case, then club participation and religiosity might act as substitutes. The Add Health data include information about club participation, but not an intensity measure as in the case of religiosity, such as how often the club meets, etc. In Table 14, we consider whether there is evidence of substitutability, in that

[37] We consider a number of other indicators, including different measures of the number of friends the individual has and parental involvement. As in the case of stressors, we chose to include in the table the measures that were correlated with depression.

more religious students participate less in clubs or sports. Columns (1) to (3) suggest that this is not the case.

Using the same strategy as in the previous two tables, we test for evidence of substitutability by considering whether religiosity matters less if the adolescent is participating in a school club or other activity. Columns (4) to (6) show that there is no evidence of this regardless of what measure of activities we use. Furthermore, while there is a large negative correlation between sports participation and depression, the correlation is much smaller for other school club participation and in neither case are the point estimates significantly different from zero. This evidence suggests that religiosity offers something unique for supporting mental health from what is offered by other typical school-related activities in which adolescents participate.

6 Conclusion

In this paper, we find that religiosity positively affects depression. In particular, a one unit increase in religiosity, e.g., attending church one more time a month, decreases the probability of being depressed by 3% out of a probability of 24%. To put this estimate in context, an increase in mother's education from no high school degree to a high school degree or more is correlated with only a 5% reduction in the probability of being depressed. Our estimated effect of religiosity is bigger than what is found in OLS, suggesting negative selection into religiosity, i.e., that individuals may select into religiosity to deal with depression or shocks associated with depression.

Our results are robust to a large number of specification checks, helping us to rule out potential confounders such as selection into peer groups and unobservable shocks that affect the group as a whole. In particular, we show that estimates are similar when pre-determined peer characterstics are used as instruments for religiosity rather than peer religiosity, helping to alleviate concerns about an unobservable third factor driving the correlations. Results are robust to controlling for average same denomination religiosity and trends in average same-denomination religiosity, which helps alleviate concerns about

unobservables at the denomination level and differential selection into schools based on an influential local church. Furthermore, we show evidence that the variation we isolate is plausibly random within-school variation using the typical balancing tests.

Interestingly, while the effects of religiosity on depression do not vary by how religious the individual is, there is considerable heterogeneity in the effect of religiosity across the distribution of depression. More depressed individuals benefit significantly more from religiosity than the least depressed. This offers a startling contrast to evidence on the effectiveness of cognitive based therapy, one of the most recommended forms of treatment, which is generally less effective for the most depressed individuals.

We consider potential mechanisms for why religiosity may affect depression. We find that the benefits of religiosity do not derive from a more religious social context in the school. We also do not find evidence that religiosity reduces exposure to stressors. We find instead that religiosity helps to buffer against stressors and that individuals who have fewer support structures in place at home and in school have bigger effects of religiosity. We also find evidence that part of the effect of religiosity derives through improved self-esteem and coping skills.

In contrast, we do not see any substitution effect of club or athletic participation. Neither of these alternative activities directly affects depression, and the effect of religiosity is similar for those who participate in clubs/athletics and those who do not. This suggests that the social support and/or sense of meaning provided by club and athletic participation does not substitute for religiosity.

The method we use to identify a causal effect of religiosity relies on variation in peer composition within schools across time and homophily in friendship formation. Determining a causal effect of religiosity is a notoriously difficult problem, and we hope that our method can be applied more generally to infer an effect of religiosity in other settings.

A limitation of our study is that we cannot explore the potentially important margin of selection into having a religious affiliation, given that peer

measures of religiosity do not shift the extensive margin. While research suggests that this may be because adolescence is not a key time for changes in religious affiliation, to the extent that the extensive margin is important, we may understate the benefits of religiosity for depression. That said, we explore including the non-religious in the regression (treating the choice to be religious as exogenous) and find results of the effect of religiosity are not markedly different. Furthermore, it is important to emphasize that our main analysis focuses on Christians because this is the dominant faith in the US and the subgroup for which our identification strategy works best. That said, we show that the negative estimated effects of religiosity on depression remain when we incorporate non-Christian students.

Overall, our findings have important implications for policies related to improving mental health in adolescence. This is particularly true given the apparent power of religiosity to help the more severely depressed, who are traditionally difficult to treat. Given our evidence on social support, self esteem and coping skills, and that other school activities do not appear to act as substitutes for religiosity, future work would benefit from more detailed information on churches and other places of worship that adolescents attend to determine in more detail the mechanisms driving these effects.

References

AHRQ, "Mental Health: Research Findings," Agency for Healthcare Research and Quality (AHRQ), 2014. Available at http://www.ahrq.gov/research/findings/factsheets/mental/mentalhth/index.html.

Becker, Sascha O. and Ludger Woessmann, "Knocking on Heaven's Door? Protestantism and Suicide," IZA Discussion Papers 5773, Institute for the Study of Labor (IZA) June 2011.

Bradley, E., "Religious Involvement and Social Resources: Evidence from the Data Set 'Americans' Changing Lives'," *Journal for the Scientific Study of Religion*, 1995, *34* (2), 259–267.

Brock, William A. and Steven N. Durlauf, "Discrete Choice with Social Interactions," *The Review of Economic Studies*, April 2001, *68* (2), 235–260.

Bun, Maurice J.G. and Teresa D. Harrison, "OLS and IV Estimation of Regression Models Including Endogenous Interaction Terms," School of Economics Working Paper Series 2014-3, LeBow College of Business, Drexel University January 2014.

CBHSQ, "Key substance use and mental health indicators in the United States: Results from the 2015 National Survey on Drug Use and Health (HHS Publication No. SMA 16-4984, NSDUH Series H-51)," Center for Behavioral Health Statistics and Quality (CBHSQ) 2016. Available at http://www.samhasa.gov/data/.

Cheadle, Jacob E. and Philip Schwadel, "The 'Friendship Dynamics of Religion,' or the 'Religious Dynamics of Friendship'? A Social Network Analysis of Adolescents Who Attend Small Schools," *Social Science Research*, 2012, *41* (5), 1198–1212.

Child Trends Databank, "Attendance at Religious Services," Child Trends, 2014. Available at http://www.childtrends.org/?indicators=religious-service-attendance, last accessed February 2016.

_ , "Religiosity among Youth," Child Trends, 2014. Available at http://www.childtrends.org/?indicators=religiosity-among-youth, last accessed February 2016.

Chiswick, Barry R. and Donka M. Mirtcheva, "Religion and Child Health: Religious Affiliation, Importance, and Attendance and Health Status among American Youth," *Journal of Family and Economic Issues*, 2013, *34* (1), 120–140.

Cook, Mary N., John Peterson, and Christopher Sheldon, "Adolescent Depression: An Update and Guide to Clinical Decision Making," *Psychiatry*, 2009, *6* (9), 17–31.

Crabtree, Steve, "Religiosity Highest in World's Poorest Nations," GALLUP, 2010. Availabel at http://www.gallup.com/poll/142727/religiosity-highest-world-poorest-nations.aspx, last accessed February 2016.

Cunha, Flavio and James J. Heckman, "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formations," *Journal of Human Resources*, 2008, *43* (4).

_ , _ , and Susanne M. Schennach, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 05 2010, *78* (3), 883–931.

Currarini, Sergio, Matthew O Jackson, and Paolo Pin, "Identifying the roles of race-based choice and chance in high school friendship network formation," *Proceedings of the National Academy of Sciences*, 2010, *107* (11), 4857–4861.

Dein, Simon, Christopher C. H. Cook, and Harold Koenig, "Religion, Spirituality, and Mental Health: Current Controversies and Future Directions," *Journal of Nervous and Mental Disease*, 2012, *200* (10), 851–855.

Desmond, Scott A., Kristopher H. Morgan, and George Kikuchi, "Religious Development: How (And Why) Does Religiosity Change From Adolescence to Young Adulthood?," *Sociological Perspectives*, 2010, *53* (2), 247–270.

Eliassen, A. Henry, John Taylor, and Donald A. Lloyd, "Subjective Religiosity and Depression in the Transition to Adulthood," *Journal for the Scientific Study of Religion*, 2005, *44* (2), 187–199.

Ellison, Christopher G. and Andrea K. Henderson, "Religion and Mental Health: Through The Lens of The Stress Process," in "Toward a Sociological Theory of Religion and Health" 2011, pp. 11–44.

_ , Jason D. Boardman, David R. Williams, and James S. Jackson, "Religious Involvement, Stress, and Mental Health: Findings from the 1995 Detroit Area Study," *Social Forces*, 2001, *80* (1), 215–249.

Ferraro, Kenneth F. and Jessica A. Kelley-Moore, "Religious Consolation among Men and Women: Do Health Problems Spur Seeking?," *Journal for the Scientific Study of Religion*, 2000, *39* (2), 220–234.

Frank, Richard G. and Thomas G. McGuire, "Economics and Mental Health," in Anthony J. Culyer and Joseph P. Newhouse, eds., *Handbook of Health Economics*, Vol. 1, Part B, Elsevier, 2000, chapter 16, pp. 893–954.

Freud, Sigmund, *The Future of an Illusion*, Garden City, NY: Double Day, 1927. Trans. by W. D. Robson-Scott.

Gloaguen, Valerie, Jean Cottraux, Michel Cucherat, and Ivy-Marie Blackburn, "A Meta-Analysis of the Effects of Cognitive Therapy in Depressed Patients," *Journal of Affective Disorders*, 1998, *49* (1), 59–72.

Gruber, Jonathan and Daniel M. Hungerman, "The Church Versus the Mall: What Happens When Religion Faces Increased Secular Competition?," *The Quarterly Journal of Economics*, 2008, *123* (2), 831–862.

Gruber, Jonathan H., "Religious Market Structure, Religious Participation, and Outcomes: Is Religion Good for You?," *Advances in Economic Analysis & Policy*, 2005, *5* (1).

Hackney, Charles H. and Glenn S. Sanders, "Religiosity and Mental Health: A Meta-Analysis of Recent Studies," *Journal for the Scientific Study of Religion*, 2003, *42* (1), 43–55.

Heckman, James J., Jora Stixrud, and Sergio Urzua, "The Effects of Cognitive and Noncognitive Abilities on Labour Market Outcomes The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 2006, *24* (3), 411–482.

Hoxby, Caroline, "Peer Effects in the Classroom: Learning from Gender and Race Variation," Working Paper 7867, National Bureau of Economic Research August 2000.

Hungerman, Daniel M., "Rethinking the Study of Religious Markets," in Rachel M. McCleary, ed., *The Oxford Handbook of the Economics of Religion*, Oxford University Press, 2011.

Iannaccone, Laurence R., "Introduction to the Economics of Religion," *Journal of Economic Literature*, 1998, *36* (3), 1465–1495.

Idler, Ellen L., "Religious Involvement and the Health of the Elderly: Some Hypotheses and an Initial Test," *Social Forces*, 1987, *66* (1), 226–238.

Imbens, Guido W. and Whitney K. Newey, "Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity," *Econometrica*, 2009, *77* (5), 1481–1512.

Iyer, Sriya, "The New Economics of Religion," *Journal of Economic Literature*, 2016, *forthcoming.*

Kleibergen, Frank and Richard Paap, "Generalized Reduced Rank Tests Using the Singular Value Decomposition," *Journal of Econometrics*, 2006, *133* (1), 97–126.

Koenig, Harold, *Handbook of Religion and Mental Health*, Academic Press, 1998.

Langa, Kenneth M, Marcia A. Valenstein, A. Mark Fendrick, Mohammed U. Kabeto, and Sandeep Vijan, "Extent and cost of informal caregiving for older Americans with symptoms of depression," *American Journal of Psychology*, 2004, *161* (857-863).

Lee, Sokbae, "Endogeneity in Quantile Regression Models: A Control Function Approach," *Journal of Econometrics*, December 2007, *141* (2), 1131–1158.

Levin, Jeff, "Religion and Mental Health: Theory and Research," *International Journal of Applied Psychoanalytic Studies*, 2010, *7* (2), 102–115.

Manski, Charles, "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 1993, *60* (3), 531–542.

Maselko, Joanna, R. David Hayward, Alexandra Hanlon, Stephen Buka, and Keith Meador, "Religious Service Attendance and Major Depression: A Case of Reverse Causality?," *American Journal of Epidemiology*, 2012, *175* (6), 576–583.

McFarland, Michael J., "Religion and Mental Health Among Older Adults: Do the Effects of Religious Involvement Vary by Gender?," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 2010, *65B* (5), 621–630.

McPherson, Miller, Lynn Smith-Lovin, and James M Cook, "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, 2001, *27* (1), 415–444.

Mellor, Jennifer M. and Beth A. Freeborn, "Religious Participation and Risky Health Behaviors among Adolescents," *Health Economics*, 2011, *20* (10), 1226–1240.

Michaelson, Valerie, Peter Robinson, and William Pickett, "Participation in Church or Religious Groups and Its Association with Health: A National Study of Young Canadians," *Journal of Religion and Health*, 2014, *53* (5), 1353–1373.

Newport, Frank, "Majority Still Says Religion Can Answer Today's Problems," GALLUP, 2014. Available at http://www.gallup.com/poll/171998/majority-says-religion-answer-today-problems.aspx, last accessed February 2016.

Nooney, Jennifer G., "Religion, Stress, and Mental Health in Adolescence: Findings from Add Health," *Review of Religious Research*, 2005, *46* (4), 341–354.

Pargament, Kenneth I. and Curtis R. Brant, "Religion and Coping," in Harold G. Koenig, ed., *Handbook of Religion and Mental Health*, Academic Press, 1998.

Patacchini, Eleonora and Yves Zenou, "Social Networks and Parental Behavior in the Intergenerational Transmission of Religion," July 2015. Manuscript. Available at http://www.ne.su.se/polopoly_fs/1.244701.1440154904!/menu/standard/file/Patacchini_Zenou_REV_QE_July_2015.pdf. Last accessed January 2016.

PewForum, "The Global Religious Landscape," Report, Pew Research Center's Forum on Religion & Public Life (PewForum), 2012. Retrieved from http://www.pewforum.org/files/2014/01/global-religion-full.pdf, last accessed February 2016.

Radloff, Lenore Sawyer, "The CES-D Scale: A Self-Report Depression Scale for Research in the General Population," *Applied Psychological Measurement*, 1977, *1* (3), 385–401.

Rosenberg, Morris, *Society and the Adolescent Self-Image*, Wesleyan University Press, 1989.

Schnittker, Jason, "When Is Faith Enough? The Effects of Religious Involvement on Depression," *Journal for the Scientific Study of Religion*, 2001, *40* (3), 393–411.

Sherkat, DarrenE. and MarkD. Reed, "The Effects of Religion and Social Support on Self-esteem and Depression among the Suddenly Bereaved," *Social Indicators Research*, 1992, *26* (3), 259–275.

Smith, Christian, Melinda Lundquist Denton, Robert Faris, and Mark Regnerus, "Mapping American Adolescent Religious Participation," *Journal for the Scientific Study of Religion*, 2015, *41* (4), 597–612.

Smith, Christopher B., Andrew J. Weigert, and Darwin L. Thomas, "Self-Esteem and Religiosity: An Analysis of Catholic Adolescents from Five Cultures," *Journal for the Scientific Study of Religion*, 1979, *18* (1), 51–60.

TADS, "The Treatment for Adolescents with Depression Study (TADS): Long-term Effectiveness and Safety Outcomes.," *Archives of General Psychiatry*, 2007, *64* (10), 1132–1143.

WHO, "Health for the world's adolescents: a second chance in the second decade," Technical Report, World Health Organization (WHO) 2014.

Wille, Nora, Susanne Bettge, and Ulrike Ravens-Sieberer, "Risk and Protective Factors for Children's and Adolescents' Mental Health: Results of the BELLAstudy," *European Child & Adolescent Psychiatry*, 2008, *17* (1), 133–147.

Williams, Paula G., Grayson N. Holmbeck, and Rachel Neff Greenley, "Adolescent health psychology," *Journal of Consulting and Clinical Psychology*, 2002, *70* (3), 828–842.

Table 1: Sample selection criteria and sample means of key variables

	Sample selection criterion						
	(1) Full in-home sample	(2) Mental health not missing	(3) Religious affiliation not missing	(4) Excluding no and other religion	(5) Religiosity not missing	(6) Covariates not missing	(7) Peer variables not missing
Mental health							
Depression	11.390 (20,662)	11.390	11.366	11.228	11.226	11.169	11.099
Religiosity							
Religiosity	8.493 (17,748)	8.492 (17,725)	8.495 (17,706)	8.558 (16,736)	8.558	8.555	8.578
Religious attendance	1.973 (17,801)	1.974 (17,776)	1.974 (17,757)	1.993 (16,781)	1.994	1.994	2.002
Youth religious activities	1.204 (17,804)	1.204 (17,780)	1.204 (17,760)	1.218 (16,783)	1.218	1.218	1.219
Praying	2.967 (17,799)	2.967 (17,775)	2.968 (17,755)	2.992 (16,781)	2.993	2.991	2.999
Religious importance	2.345 (17,799)	2.345 (17,775)	2.345 (17,755)	2.353 (16,780)	2.353	2.353	2.358
Sample size	20,745	20,662	20,312	16,806	16,736	15,869	12,945
% of full sample	100.00	99.60	97.91	81.01	80.67	76.50	62.40

Notes Sample means of key variables under each sample selection criterion are reported in cells. The number of observations for each variable is reported in parenthesis when it is different from the total sample size under that selection criterion.

Table 2: Heterogeneity in religiosity and mental health

	N	Religiosity		Depression	
		Mean	SD	Mean	SD
Gender					
Female	6,666	8.89	(3.18)	11.99	(8.03)
Male	6,279	8.25	(3.38)	10.15	(6.61)
Race					
White	6,826	8.17	(3.46)	10.06	(7.09)
Hispanic	2,243	8.07	(3.09)	12.80	(7.88)
Black	2,817	9.78	(2.75)	11.46	(7.42)
Other ethnicity	1,059	9.10	(3.09)	13.26	(7.49)
Denomination					
Catholic	4,275	7.66	(3.09)	11.53	(7.65)
Liberal Protestant	1,130	8.09	(3.56)	9.34	(6.46)
Moderate Protestant	2,506	8.48	(3.43)	10.98	(7.25)
Conservative Protestant	5,034	9.51	(3.08)	11.19	(7.49)
Household income					
Low income	1,951	8.61	(3.25)	12.45	(7.75)
Medium income	5,283	8.51	(3.36)	10.89	(7.35)
High income	2,496	8.49	(3.31)	9.71	(6.97)
Mother's education					
Mother no high school	2,039	8.36	(3.22)	13.21	(7.93)
Mother high school	7,320	8.48	(3.32)	10.91	(7.29)
Mother degree and above	2,914	9.15	(3.21)	9.82	(6.99)

Table 3: Baseline estimates of the effect of religiosity on adolescent mental health

	Dependent variable = depression			Dependent variable = depressed		
	(1) OLS	(2) IV	(3) First stage	(4) OLS LPM	(5) IV LPM	(6) IV Probit
Religiosity	−0.163***	−0.698**		−0.006***	−0.034**	−0.034**
	(0.024)	(0.289)		(0.001)	(0.016)	(0.016)
Peer religiosity			0.112***			
			(0.020)			
Black	0.526	0.918**	0.660***	0.025	0.045*	0.048*
	(0.372)	(0.455)	(0.120)	(0.021)	(0.025)	(0.025)
Hispanic	1.165***	1.515***	0.600***	0.035*	0.053**	0.053**
	(0.287)	(0.365)	(0.133)	(0.020)	(0.023)	(0.022)
Other ethnicity	2.240***	2.766***	0.864***	0.100***	0.128***	0.124***
	(0.393)	(0.561)	(0.212)	(0.022)	(0.031)	(0.028)
Liberal Protestant	−0.616*	−0.466	0.242	−0.049***	−0.041**	−0.046**
	(0.325)	(0.342)	(0.195)	(0.017)	(0.018)	(0.022)
Moderate Protestant	0.074	0.436	0.604***	−0.010	0.009	0.013
	(0.253)	(0.303)	(0.116)	(0.013)	(0.017)	(0.018)
Conservative Protestant	0.155	0.757*	1.006***	−0.015	0.016	0.020
	(0.251)	(0.392)	(0.134)	(0.015)	(0.023)	(0.025)
Female	0.826	1.132**	0.505**	0.053	0.069*	0.068**
	(0.511)	(0.558)	(0.208)	(0.033)	(0.036)	(0.034)
Age	1.405***	1.276***	−0.235***	0.073***	0.066***	0.063***
	(0.105)	(0.135)	(0.048)	(0.007)	(0.008)	(0.008)
School year in session	1.092***	1.146***	0.100	0.052***	0.055***	0.055***
	(0.149)	(0.162)	(0.064)	(0.008)	(0.008)	(0.008)
Puberty (male)	−0.108***	−0.119***	−0.022	−0.006***	−0.007***	−0.008***
	(0.032)	(0.034)	(0.014)	(0.002)	(0.002)	(0.002)
Puberty (female)	0.015	0.008	−0.014	0.000	−0.001	−0.001
	(0.031)	(0.032)	(0.010)	(0.002)	(0.002)	(0.002)
Mother not present	−0.181	−0.302	−0.206	−0.001	−0.007	−0.005
	(0.339)	(0.347)	(0.136)	(0.019)	(0.018)	(0.016)
Mother high school or some college	−1.100***	−1.035***	0.124	−0.051***	−0.048***	−0.042***
	(0.280)	(0.251)	(0.119)	(0.012)	(0.012)	(0.012)
Mother degree and above	−1.646***	−1.266***	0.718***	−0.072***	−0.053**	−0.051**
	(0.351)	(0.390)	(0.157)	(0.017)	(0.020)	(0.022)
Father not present	0.591***	0.292	−0.555***	0.030***	0.014	0.013
	(0.163)	(0.228)	(0.069)	(0.010)	(0.013)	(0.014)
Log household income	1.194	1.367	0.388	0.044	0.053	0.075
	(1.500)	(1.451)	(0.662)	(0.078)	(0.081)	(0.082)

Continued on next page ...

. . . continued from previous page

	Dependent variable = depression			Dependent variable = depressed		
	(1) OLS	(2) IV	(3) First stage	(4) OLS LPM	(5) IV LPM	(6) IV Probit
Log household income squared/10	−0.079	−0.087	−0.019	−0.003	−0.003	−0.005
	(0.073)	(0.071)	(0.033)	(0.004)	(0.004)	(0.004)
Household income missing	3.843	4.812	2.123	0.144	0.194	0.296
	(7.722)	(7.445)	(3.367)	(0.402)	(0.413)	(0.413)
Grade 8	−1.113***	−1.179***	−0.089	−0.049***	−0.052***	−0.044***
	(0.273)	(0.258)	(0.104)	(0.015)	(0.014)	(0.016)
Grade 9	−2.058***	−2.060***	0.044	−0.107***	−0.107***	−0.093***
	(0.443)	(0.420)	(0.163)	(0.024)	(0.023)	(0.025)
Grade 10	−3.092***	−3.070***	0.110	−0.161***	−0.160***	−0.141***
	(0.521)	(0.503)	(0.177)	(0.029)	(0.029)	(0.029)
Grade 11	−4.522***	−4.432***	0.242	−0.226***	−0.221***	−0.197***
	(0.601)	(0.597)	(0.213)	(0.034)	(0.034)	(0.033)
Grade 12	−6.310***	−6.198***	0.299	−0.329***	−0.323***	−0.295***
	(0.705)	(0.696)	(0.256)	(0.039)	(0.038)	(0.039)
School FE	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic			30.438			

Notes This table reports the OLS and IV estimates of religiosity on CES-D scale of depression and the probability of being depressed. Columns (1)-(5) report the coefficients, whereas column (6) reports the marginal effects. The omitted groups for race, religious denomination, and mother's education background are white, Catholic, and mother's education lower than high school respectively. ***, **, and * denote statistical significance at 0.01, 0.05, and 0.10 levels respectively. F-statistic on the excluded instrument refers to the Wald version of Kleibergen and Paap (2006) rk-statistic on the excluded instrumental variables for non-i.i.d. errors. The number of observations is 12,945 for all models except in column (6), where the number of observations is 12,913 due to that including school fixed effects perfectly predicts outcomes for 32 observations.

Table 4: Balancing test

	(1) Mother not present	(2) Father not present	(3) Mother no high school	(4) Mother high school	(5) Mother degree and above	(6) Log household income	(7) School year in session	(8) Puberty (male)	(9) Puberty (female)
Peer religiosity	-0.002^{**}	-0.003	0.001	0.002	0.000	-0.011	0.001	0.025	0.010
	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)	(0.023)	(0.002)	(0.021)	(0.021)
Female	-0.006	0.038^{***}	0.010	0.001	-0.005	-0.206^{**}	-0.036^{***}		
	(0.005)	(0.008)	(0.007)	(0.010)	(0.009)	(0.103)	(0.012)		
Black	-0.007	0.219^{***}	-0.003	0.013	-0.003	-0.785^{***}	0.068^{***}	-1.311^{***}	-0.375^{**}
	(0.008)	(0.016)	(0.020)	(0.023)	(0.021)	(0.182)	(0.021)	(0.132)	(0.180)
Hispanic	-0.007	0.038^{**}	0.352^{***}	-0.180^{***}	-0.165^{***}	-1.544^{***}	0.041^{*}	-0.353^{**}	-0.057
	(0.007)	(0.019)	(0.045)	(0.031)	(0.025)	(0.221)	(0.021)	(0.162)	(0.169)
Other ethnicity	-0.001	0.002	0.030	-0.143^{***}	0.114^{**}	-1.556^{***}	0.003	-1.057^{***}	-0.432^{**}
	(0.012)	(0.018)	(0.045)	(0.026)	(0.049)	(0.279)	(0.027)	(0.169)	(0.173)
Liberal Protestant	-0.018^{*}	0.001	-0.031^{**}	-0.037	0.086^{***}	0.377^{**}	0.035	0.367^{***}	0.060
	(0.009)	(0.015)	(0.013)	(0.027)	(0.025)	(0.178)	(0.023)	(0.131)	(0.166)
Moderate Protestant	0.004	0.037^{***}	-0.013	0.019	-0.009	-0.080	-0.005	0.201^{*}	0.297^{**}
	(0.006)	(0.010)	(0.016)	(0.023)	(0.015)	(0.148)	(0.013)	(0.107)	(0.148)
Conservative	0.001	0.055^{***}	0.022	0.013	-0.036^{***}	-0.148	0.021	0.185	0.260^{*}
Protestant	(0.007)	(0.013)	(0.016)	(0.019)	(0.012)	(0.129)	(0.015)	(0.112)	(0.147)
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grade dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	12,945	12,945	12,945	12,945	12,945	12,945	12,945	6,279	6,666

Notes: Clustered standard levels at the school level are in parentheses. ***, **, and * denote significance at 0.01, 0.05, and 0.10 levels respectively.

Table 5: Robustness checks

	(1) Exclude friends from peers	(2) Exclude private schools	(3) Substitute with parental denomination	(4) Same- and cross-gender peers		(5) Same- and cross-denomination peers	
				second stage	first stage	second stage	first stage
Religiosity	−0.711**	−0.693**	−0.736*	−0.894***		−0.712**	
	(0.297)	(0.296)	(0.435)	(0.296)		(0.311)	
Same-gender peer religiosity					0.118***		
					(0.022)		
Cross-gender peer religiosity					0.063***		
					(0.022)		
Same-denomination peer religiosity							0.109***
							(0.022)
Cross-denomination peer religiosity							−0.003
							(0.023)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic	28.442	29.849	12.114	20.029	20.029	12.409	12.409
Hansen's J-test					0.766		0.238
Observations	12,927	12,079	9,972	11,831	11,831	12,035	12,035

Notes Baseline controls are as in Table 3. Clustered standard errors at the school level are in parentheses. ***, **, and * denote statistical significance at 0.01, 0.05, and 0.10 levels respectively. F-statistic on the excluded instruments refers to the Wald version of the Kleibergen-Paap (2006) rk-statistic on the excluded instrumental variables for non-i.i.d. errors. Hansen's J-test reports the p-values of Hansen's J-test on overidentifying restrictions.

Table 6: Robustness of IV estimates of the effect of religiosity on depression

	(1)		(2)		(3)		(4)		(5)		(6)	
	2nd stage	1st stage	2nd stage	1st stage	2nd stage	1st stage	2nd stage	1st stage	2nd stage	1st stage	2nd stage	1st stage
Religiosity	−0.675**		−0.606**		−0.655**		−0.597*		−0.712*		−0.741*	
	(0.295)		(0.302)		(0.314)		(0.318)		(0.429)		(0.437)	
Peer depression					0.010	−0.009	0.005	−0.008	0.004	−0.008	0.007	−0.008
					(0.025)	(0.006)	(0.024)	(0.006)	(0.025)	(0.006)	(0.026)	(0.006)
SR religiosity[a]									−0.084	0.007	−0.567	0.111
									(0.119)	(0.064)	(0.404)	(0.189)
SD religiosity[b]									0.154	0.167**	1.046*	−0.401*
									(0.170)	(0.082)	(0.600)	(0.237)
SR religiosity[a]× grade											0.049	−0.012
											(0.045)	(0.020)
SD religiosity[b]× grade											−0.093	0.060**
											(0.067)	(0.026)
Peer religiosity		0.110***		0.108***		0.108***		0.107***		0.089***		0.088***
		(0.021)		(0.021)		(0.021)		(0.021)		(0.023)		(0.023)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SG contextuals[c]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Joint test 1[d]	0.879	0.052	0.858	0.073	0.880	0.062	0.861	0.085	0.890	0.067	0.906	0.096
Peer characteristics	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	No	No
Joint test 2[e]			0.886	0.517			0.882	0.543	0.895	0.514		
F-statistic		28.664		27.817		27.477		27.024		14.509		14.061

Notes All models include the baseline controls as in Table 3. Clustered standard levels at the school level are in parentheses. ***, **, and * denote statistical significance at 0.01, 0.05, and 0.10 levels respectively. F-statistic on the excluded instrument refers to the Wald version of the Kleibergen-Paap (2006) rk-statistic on the excluded instrumental variables for non-i.i.d. errors. The number of observations is 12,945 in all models.

[a] School-race average religiosity.

[b] School-denomination average religiosity.

[c] School-grade level contextual variables, including proportions of female, each race, and each denomination.

[d] This row reports the p-values of a joint significance test on all school-grade contextual variables.

[e] This row reports the p-values of a joint significance test on all peer characteristics.

Table 7: Addtional robustness checks using larger samples

	(1) Christians only, largest possible sample	(2) Including other religion	(3) Including those with missing peer religiosity	(4) Including non-religious
Religiosity	-0.762^{***}	-0.622^{**}	-0.431^{*}	-0.484^{*}
	(0.289)	(0.289)	(0.245)	(0.279)
Peer religiosity missing			0.304	0.174
			(0.186)	(0.170)
Baseline controls	Yes	Yes	Yes	Yes
F-statistic	33.483	38.777	22.767	18.734
J-test			0.618	0.662
N	13,022	13,475	15,961	18,137

Notes This table reports the IV estimates of the effect of religiosity on depression on larger samples. Column (1) replicates the results on the Christian sample, without dropping missing peer covariates. Column (2) also includes individuals who report other affiliated religions. Column (3) includes those who do not have a valid SGRXD (same school, grade, race, gender, denomination), by replacing their peer religiosity with SGXD (same school, grade, gender, denomination) peer religiosity and SGRX (same school, grade, race, gender) peer religiosity. The instruments in this model are the redefined peer religiosity, and its interaction with a dummy indicating missing SGRXD peer religiosity. Column (4) also includes those who are not religious. Peer religiosity for these invididuals are redefined as SGRX peer religiosity. The instruments in this model are the redefined peer religiosity, and its interaction with a dummy indicating missing SGRXD peer religiosity. Baseline controls include covariates as in Table 3. Additionally, a dummy indicating other religion is also controlled for in columns (2) and (3). Column (4) further controls for a dummy indicating no religion. Clustered standard levels at the school level are in parentheses. ***, **, and * denote statistical significance at 0.01, 0.05, and 0.10 levels respectively. F-statistic refers to the Wald version of the Kleibergen-Paap (2006) rk-statistic on the excluded instrumental variables for non-i.i.d. errors. J-test reports the p-values of Hansen's J-test on overidentifying restrictions.

Table 8: Robustness checks using standardized measure of various CES-D scales

	Scale			Factor		
	(1) Original 19 questions	(2) Remove 3 social questions	(3) Using 8 questions	(4) Original 19 questions	(5) Remove 3 social questions	(6) Using 8 questions
Religiosity	-0.092^{**}	-0.092^{**}	-0.075^{*}	-0.072^{**}	-0.072^{**}	-0.054^{*}
	(0.038)	(0.038)	(0.038)	(0.035)	(0.035)	(0.032)
F-statistic	30.438	30.438	30.438	30.438	30.438	30.438
N	12,945	12,945	12,945	12,945	12,945	12,945

Notes This tables reports the IV estimates for the effect of religiosity on the standardised CES-D scale using sample mean and standard deviation. The instrument is peer religiosity, where the peer group is defined as other students of the same grade (G), race (R), gender (X) and denomination (D) in the same school (S). All models control for covariates and school fixed effects as in Table 3. Clustered standard levels at the school level are in parentheses. F-statistic on the excluded instrument refers to the Wald version of the Kleibergen-Paap (2006) rk-statistic on the excluded instrumental variables for non-i.i.d. errors. ***, **, and * denote statistical significance at 0.01, 0.05, and 0.10 levels respectively.

Table 9: Robustness checks by controlling for potential reference effects at various group levels

	(1) Control for SG mean depression	(2) Control for SR mean depression	(3) Control for SX mean depression	(4) Control for SD mean depression	(5) Control for SGR mean depression	(6) Control for SGX mean depression	(7) Control for SGD mean depression	(8) Control for SRX mean depression	(9) Control for SRD mean depression	(10) Control for SXD mean depression	(11) Control for SGRX mean depression
Religiosity	-0.712^{**}	-0.772^{**}	-0.706^{**}	-0.722^{**}	-0.705^{**}	-0.663^{**}	-0.686^{**}	-0.712^{**}	-0.710^{**}	-0.704^{**}	-0.643^{**}
	(0.296)	(0.303)	(0.292)	(0.293)	(0.301)	(0.291)	(0.296)	(0.295)	(0.300)	(0.291)	(0.294)
Group mean CES-D	-0.050	-0.279^{*}	-0.418^{***}	-0.115	-0.010	0.078	0.013	-0.050	-0.021	-0.035	0.060^{*}
	(0.085)	(0.164)	(0.159)	(0.083)	(0.055)	(0.049)	(0.043)	(0.083)	(0.055)	(0.051)	(0.036)
F-statistic	29.469	29.395	30.446	30.666	28.724	29.389	29.518	30.401	29.779	30.664	29.044
N	12,945	12,945	12,945	12,945	12,945	12,945	12,945	12,945	12,945	12,945	12,945

Notes This tables reports the IV estimates for the effect of religiosity on the CES-D scale of depression, controlling for average depression at various group levels. The instrument is peer religiosity, where the peer group is defined as other students of the same grade (G), race (R), gender (X) and denomination (D) in the same school (S). Mean CES-D at different group levels (calculated excluding oneself) are controlled for. For instance, column (1) controls for mean CES-D at the SG (school-grade) level, calculated as the mean CES-D of all students in the same school-grade excluding the respondent. All models control for covariates and school fixed effects as in Table 3. Clustered standard levels at the school level are in parentheses. F-statistic on the excluded instrument refers to the Wald version of the Kleibergen-Paap (2006) rk-statistic on the excluded instrumental variables for non-i.i.d. errors. ***, **, and * denote statistical significance at 0.01, 0.05, and 0.10 levels respectively.

Table 10: Sensitivity analysis with different cutoffs on the CES-D scale for high depressive symptoms

	Cutoffs on the CES-D scale									
	(1) ≥ 15	(2) ≥ 16	(3) ≥ 17	(4) ≥ 18	(5) ≥ 19	(6) ≥ 20	(7) ≥ 21	(8) ≥ 22	(9) ≥ 23	(10) ≥ 24
Panel A: IV LPM, main sample, $F = 30.438, N = 12,945$										
Religiosity	-0.024	-0.034^{**}	-0.041^{**}	-0.051^{***}	-0.045^{***}	-0.040^{***}	-0.039^{***}	-0.030^{**}	-0.027^{**}	-0.027^{**}
	(0.016)	(0.016)	(0.016)	(0.016)	(0.015)	(0.014)	(0.013)	(0.012)	(0.011)	(0.011)
% above cutoff	27.2	23.6	20.7	17.8	15.2	13.0	11.0	9.4	8.0	6.7
Panel B: IV probit, main sample										
Religiosity	-0.024	-0.034^{**}	-0.041^{**}	-0.053^{***}	-0.048^{***}	-0.046^{***}	-0.046^{***}	-0.035^{**}	-0.035^{**}	-0.037^{*}
	(0.017)	(0.016)	(0.017)	(0.016)	(0.016)	(0.017)	(0.017)	(0.016)	(0.017)	(0.019)
% above cutoff	27.3	23.6	20.7	17.9	15.3	13.1	11.1	9.4	8.2	6.9
N	12,913	12,913	12,915	12,872	12,872	12,831	12,831	12,831	12,653	12,517

Notes This tables reports the IV estimates for the effect of religiosity on a series of binary variables indicating high depressive symptom using different cutoffs on the CES-D scale. The instrument is peer religiosity, where the peer group is defined as other students of the same grade, race, gender and denomination in the same school. Linear probability models (LPM) report the coefficients, whereas probit models report the marginal effects evaluated at the means. All models control for covariates and school fixed effects as in Table 3. School fixed effects in probit models are controlled for by including school dummies in the estimation. Clustered standard levels at the school level are in parentheses. F-statistic on the excluded instrument refers to the Wald version of the Kleibergen-Paap (2006) rk-statistic on the excluded instrumental variables for non-i.i.d. errors. ***, **, and * denote statistical significance at 0.01, 0.05, and 0.10 levels respectively.

Table 11: Religiosity, psychological resources and depression

	(a) Dependent variable = psychological resources				(b) Dependent variable = depression		
	Self-esteem	Self-esteem	Passive P-S	Passive P-S	Self-esteem	Passive P-S	Both
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) IV	(6) IV	(7) IV
Religiosity	0.075***	0.153	0.022***	0.113	−0.508*	−0.571**	−0.406
	(0.008)	(0.105)	(0.007)	(0.102)	(0.270)	(0.275)	(0.257)
Self-esteem					−1.234***		−1.228***
					(0.041)		(0.038)
Passive problem-solving						−0.725***	−0.689***
						(0.032)	(0.031)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic		30.399		31.916	30.117	31.644	31.331
N	12,931	12,931	12,900	12,900	12,931	12,900	12,889

Notes Columns (1)–(4) report the IV estimates for the effect of religiosity on psychological resources. Columns (5)–(7) report the IV estimates for the effect of religiosity on depression conditional on psychological resources. Baseline controls include covariates as in Table 3. Clustered standard levels at the school level are in parentheses. ***, **, and * denote statistical significance at 0.01, 0.05, and 0.10 levels respectively. F-statistic refers to the Wald version of the Kleibergen-Paap (2006) rk-statistic on the excluded instrumental variables for non-i.i.d. errors.

Table 12: Religiosity, stressors and depression

	(a) Dependent variable = stressor			(b) Dependent variable = depression		
	(1) GPA	(2) Family or friends suicide	(3) General health	(4) GPA	(5) Family or friends suicide	(6) General health
Religiosity	0.033	−0.006	−0.063	−0.667*	−0.643**	−1.436***
	(0.031)	(0.019)	(0.039)	(0.349)	(0.293)	(0.389)
Interaction				0.015	−0.598***	0.160**
				(0.088)	(0.197)	(0.072)
Stressor				−1.747**	8.214***	−3.050***
				(0.780)	(1.687)	(0.623)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic	30.425	30.284	30.416	14.615	14.914	16.010
N	12,838	12,888	12,944	12,838	12,888	12,944

Notes Columns (1)–(3) report the IV estimates for the effect of religiosity on exposure to stressors. Columns (4)–(6) report the IV estimates for the main and interaction effect of religiosity on depression conditional on stressors. Baseline controls include covariates as in Table 3. Clustered standard levels at the school level are in parentheses. ***, **, and * denote statistical significance at 0.01, 0.05, and 0.10 levels respectively. F-statistic refers to the Wald version of the Kleibergen-Paap (2006) rk-statistic on the excluded instrumental variables for non-i.i.d. errors.

Table 13: Religiosity, support structures and depression

	(a) Dependent variable = support structure			(b) Dependent variable = depression		
	(1) Single parent	(2) Protective factors	(3) Neighborhood resources	(4) Single parent	(5) Protective factors	(6) Neighborhood resources
Religiosity	0.014	0.163	−0.020	−0.575*	−1.316**	−0.963***
	(0.013)	(0.174)	(0.052)	(0.320)	(0.537)	(0.299)
Interaction				−0.322*	0.024*	0.086
				(0.177)	(0.014)	(0.056)
Support structure				2.630*	−0.805***	−1.427***
				(1.525)	(0.119)	(0.485)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic	28.102	32.337	30.324	14.120	16.172	15.338
N	10,504	12,675	12,750	10,504	12,675	12,750

Notes Columns (1)–(3) report the IV estimates for the effect of religiosity on support structures. Columns (4)–(6) report the IV estimates for the main and interaction effects of religiosity on depression conditional on support structures. Baseline controls include covariates as in Table 3. Clustered standard levels at the school level are in parentheses. ***, **, and * denote statistical significance at 0.01, 0.05, and 0.10 levels respectively. F-statistic refers to the Wald version of the Kleibergen-Paap (2006) rk-statistic on the excluded instrumental variables for non-i.i.d. errors.

Table 14: Religiosity, school activities and depression

	(a) Dependent variable = school activities			(b) Dependent variable = depression		
	(1) School club participation	(2) School sports participation	(3) School activity participation	(4) School club participation	(5) School sports participation	(6) School activity participation
Religiosity	0.016	−0.012	−0.017	−0.670**	−0.748***	−0.740**
	(0.017)	(0.021)	(0.020)	(0.313)	(0.284)	(0.298)
Interaction				−0.040	0.135	0.053
				(0.138)	(0.144)	(0.154)
School activities				−0.137	−1.708	−1.211
				(1.150)	(1.290)	(1.318)
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic	30.438	30.438	30.438	14.821	15.177	15.721
N	12,945	12,945	12,945	12,945	12,945	12,945

Notes Columns (1)–(3) report the IV estimates for the effect of religiosity on participation in school activities. Columns (4)–(6) report the IV estimates for the main and interaction effect of religiosity on depression conditional on participation in school activities. Baseline controls include covariates as in Table 3. Clustered standard levels at the school level are in parentheses. ***, **, and * denote statistical significance at 0.01, 0.05, and 0.10 levels respectively. F-statistic refers to the Wald version of the Kleibergen-Paap (2006) rk-statistic on the excluded instrumental variables for non-i.i.d. errors.

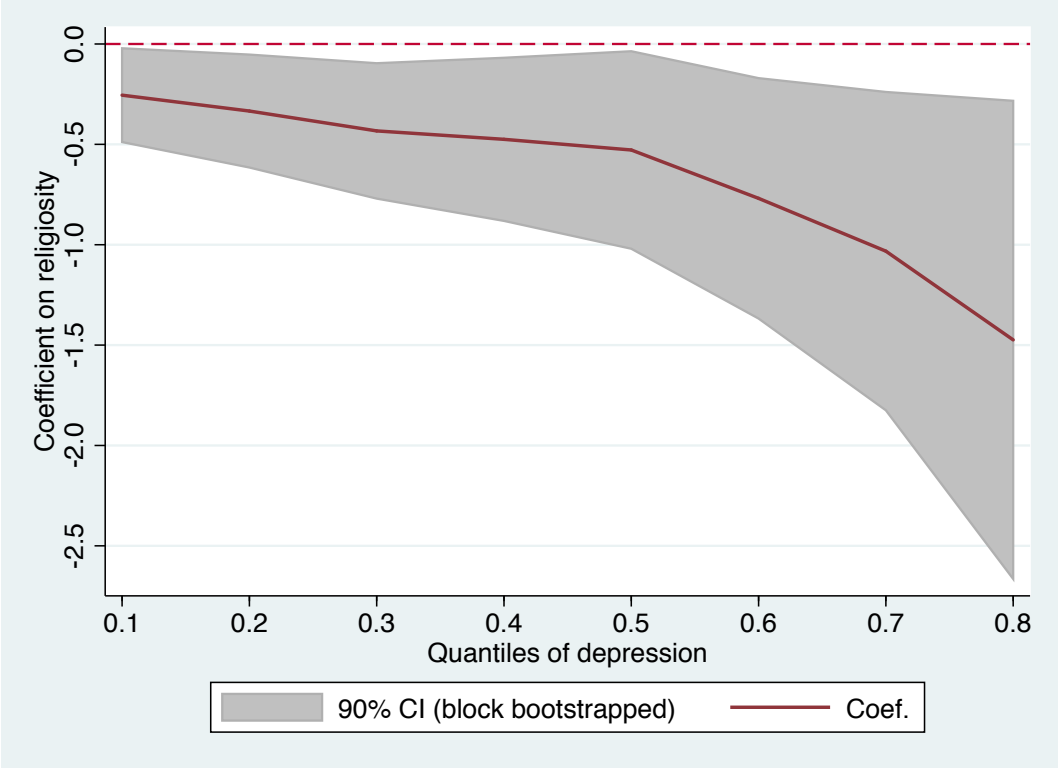

Figure 1: Effect of religiosity on different quantiles of the conditional depression distribution

A Appendix

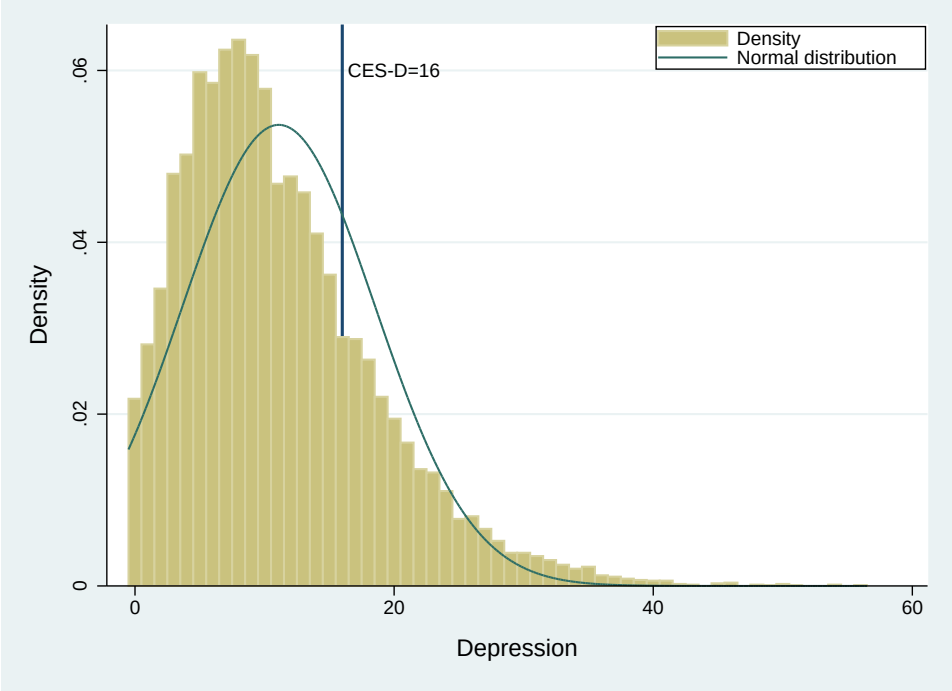

Figure A.1: Distribution of the CES-D scale of depression

Table A.1: Definition of key variables

No.	Question
Religiosity	
	Definition: sum over the following variables.
(1)	In the past 12 months, how often did you attend religious services? Responses: 0 = never, 1 = less than once a month, 2 = less than once a week/at least once a month, 3 = once a week or more.
(2)	Many churches, synagogues, and other places of worship have special activities for teenagers—such as youth groups, Bible classes, or choir. In the past 12 months, how often did you attend such youth activities? Responses: coded same as question (1) above.
(3)	How important is religion to you? Responses: 0 = not important at all, 1 = fairly unimportant, 2 = fairly important, 3 = very important.
(4)	How often do you pray? Responses: 0 = never, 1 = less than once a month, 2 = at least once a month, 3 = at least one a week, 4 = at least once a day.
Depression	
	Definition: sum over the following variables.
	Coding of responses: 0 = never/rarely, 1 = sometimes, 2 = a lot of the time, 3 = most/all of the time.
	How often was each of the following true during the last week?
(1)	You were bothered by things that usually don't bother you.
(2)	You didn't feel like eating, your appetite was poor.
(3)	You felt that you could not shake off the blues, even with help from your family and your friends.
(4)	You felt that you were just as good as other people.[a]
(5)	You had trouble keeping your mind on what you were doing.
(6)	You felt depressed.
(7)	You felt that you were too tired to do things.
(8)	You felt hopeful about the future.[a]
(9)	You thought your life had been a failure.
(10)	You felt fearful.
(11)	You were happy.[a]
(12)	You talked less than usual.
(13)	You felt lonely.
(14)	People were unfriendly to you.
(15)	You enjoyed life.[a]
(16)	You felt sad.
(17)	You felt that people disliked you.
(18)	It was hard to get started doing things.
(19)	You felt life was not worth living.

Notes

[a] Responses to these questions are reverse coded, such that 3 = never/rarely, 2 = sometimes, 1 = a lot of the time, 0 = most/all of the time.

Table A.2: Categorization of religious affiliations

Religious denomination	Religious affiliations
No religion	No religion
Catholic	Catholic
Liberal Protestant	Episcopal, Friends/Quaker, Methodist, Presbyterian, United Church of Christ, Unitarian
Moderate Protestant	Christian Church (Disciples of Christ), Lutheran, National Baptist, other Protestant
Conservative Protestant	Adventist, AME/AME Zion/CME, Assemblies of God, Baptist, Christian Science, Jehovah's Witness, Congregational, Holiness, Latter Day Saints (Mormon), Pentecostal
Other religion	Baha'i, Buddhist, Eastern Orthodox, Hindu, Islam, Jewish, other religion

Table A.3: Summary statistics

	Mean	Std. Dev.	Min.	Max.	N
Mental Health					
Depression	11.10	7.43	0.00	56.00	12,945
Depressed (CES-D ≥16)	0.24	0.42	0.00	1.00	12,945
Religiosity					
Religiosity	8.58	3.30	0.00	13.00	12,945
Religious attendance	2.00	1.07	0.00	3.00	12,945
Youth religious activities	1.22	1.24	0.00	3.00	12,945
Praying	3.00	1.26	0.00	4.00	12,945
Religious importance	2.36	0.75	0.00	3.00	12,945
Individual characteristics					
Female	0.51	0.50	0.00	1.00	12,945
White	0.53	0.50	0.00	1.00	12,945
Black	0.22	0.41	0.00	1.00	12,945
Hispanic	0.17	0.38	0.00	1.00	12,945
Other ethnicity	0.08	0.27	0.00	1.00	12,945
Catholic	0.33	0.47	0.00	1.00	12,945
Liberal Protestant	0.09	0.28	0.00	1.00	12,945
Moderate Protestant	0.19	0.40	0.00	1.00	12,945
Conservative Protestant	0.39	0.49	0.00	1.00	12,945
Age	16.17	1.68	11.42	21.25	12,945
School year in session	0.37	0.48	0.00	1.00	12,945
Puberty (male)	5.50	6.04	0.00	19.00	12,945
Puberty (female)	7.32	7.59	0.00	26.00	12,945
Parental background					
Mother not present	0.05	0.22	0.00	1.00	12,945
Mother high school or some college	0.57	0.50	0.00	1.00	12,945
Mother degree and above	0.23	0.42	0.00	1.00	12,945
Father not present	0.29	0.45	0.00	1.00	12,945
Log household income	7.85	4.57	0.00	13.81	12,945
Log household income squared/10	82.48	49.49	0.00	190.84	12,945
Household income missing	0.25	0.43	0.00	1.00	12,945
Peer mental health					
Peer depression	11.13	4.45	0.00	46.00	12,945
Peer religiosity					
Peer religiosity	8.57	2.24	0.00	13.00	12,945

Continued on next page ...

. . . continued from previous page

	Mean	Std. Dev.	Min.	Max.	N
Same-gender peer religiosity	8.57	2.24	0.00	13.00	12,945
Cross-gender peer religiosity	8.55	2.12	0.00	13.00	11,831
School-race and school-deonomination religiosity					
SR religiosity	8.53	1.32	2.50	13.00	12,945
SD religiosity	8.58	1.42	3.00	13.00	12,945
Psychological resources					
Self-esteem	16.37	2.53	4.00	20.00	12,931
Passive problem-solving	8.26	2.20	3.00	15.00	12,900
Stressors					
Most recent GPA	2.76	0.77	1.00	4.00	12,838
Friends/Family suicide	0.19	0.39	0.00	1.00	12,888
General health	3.90	0.90	1.00	5.00	12,944
Participation in school activities					
School club participation	0.44	0.50	0.00	1.00	12,945
School sports participation	0.42	0.49	0.00	1.00	12,945
School activity participation	0.61	0.49	0.00	1.00	12,945

Notes: Peer group is defined as the students in the same school-grade with the same gender, race, and religious denomination.
Source: Add Health Wave I.

Table A.4: Definition of additional variables

No.	Question
Self-esteem	
	Definition: sum over the following variables.
	Coding of responses: 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree.
(1)	You have a lot to be proud of.
(2)	You like yourself just the way you are.
(3)	You feel like you are doing everything just about right.
(4)	You have a lot of good qualities.
Passive problem-solving	
	Definition: sum over the following variables.
	Coding of responses: 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree.
(1)	You usually go out of your way to avoid having to deal with problems in your life.
(2)	Difficult problems make you very upset.
(3)	When making decisions, you usually go with your "gut feeling" without thinking too much about the consequences of each alternative.
Protective factors	
	Definition: sum over the following variables.
	Coding of responses: 1= not at all, 2 = very little, 3 = somewhat, 4 = quite a bit, 5 = very much.
(1)	How much do you feel that adults care about you?
(2)	How much do you feel that your teachers care about you?
(3)	How much do you feel that your parents care about you?
(4)	How much do you feel that your friends care about you?
(5)	How much do you feel that people in your family understand you?
(6)	How much do you feel that you want to leave home?
(7)	How much do you feel that you and your family have fun together?
(8)	How much do you feel that your family pays attention to you?
Neighborhood resources	
	Definition: sum over the following variables.
	Coding of responses: 1 = true/yes, 0 = false/no.
(1)	You know most of the people in your neighborhood.
(2)	In the past month, you have stopped on the street to talk with someone who lives in your neighborhood.
(3)	People in this neighborhood look out for each other.
(4)	Do you usually feel safe in your neighborhood?
(5)	On the whole, how happy are you with living in your neighborhood?[a]
GPA	
	Definition: average across the following variables.
	Coding of responses: 1 = D or lower, 2 = C, 3 = B, 4 = A.

Continued on next page ...

. . . continued from previous page

No.	Question
(1)	At the most recent grading period, what was your grade in each of the following subjects? English/Language Arts
(2)	At the most recent grading period, what was your grade in each of the following subjects? Mathematics
(3)	At the most recent grading period, what was your grade in each of the following subjects? History/Social Studies
(4)	At the most recent grading period, what was your grade in each of the following subjects? Science

Family/friends suicide

Definition: equals 1 if answer is "yes" to either question, and 0 otherwise.

Coding of responses: 1 = yes, 0 = no.

(1)	Have any of your family tried to kill themselves during the past 12 months?
(2)	Have any of your friends tried to kill themselves during the past 12 months?

General health

Definition: response to the following variable.

Coding of responses: 1 = poor, 2 = fair, 3 = good, 4 = very good, 5 = excellent.

(1)	In general, how is your health?

Notes

[a] Coded as: 1 = somewhat/quite a bit/very much, 0 = not at all/very little.