QUEEN, WILL. Ph.D. Adolescent Depression and Adult Labor Market Outcomes. (2022)
Directed by Dr. Jeremy Bray. 134 pp.

The prevalence of adolescent depression has more than doubled in the past ten years. Youth with depression do worse in the labor market as adults, but it is unclear whether adolescent depression plays a causal role in determining labor market outcomes. This dissertation uses a mediation analysis framework to estimate the direct and indirect effects of adolescent depression on adult earnings and wages. Adolescent depression leads to many adverse outcomes, including poor performance in school and a higher risk of mental health disorders. I treat adult depression and educational attainment as mediators of the effect of adolescent depression on labor market outcomes.

First, I conduct a literature review on the relationship between adolescent depression and labor market outcomes. I establish a conceptual framework for thinking about the long-term impacts of adolescent depression and discuss the challenges in identifying causal effects. Next, I produce an in-depth descriptive analysis of how adolescent depression relates to adult earnings and wages. I use mediation analysis and the difference method to estimate the direct and indirect effects of adolescent depression. This approach reconciles many seemingly conflicting results in the previous literature. To address several issues of endogeneity, I then use instrumental variables and a system of equations to identify causal effects of adolescent depression on earnings and wages. I estimate the indirect effects through years of education and adult depression, as well as the leftover ‘direct’ effect.

I find that adolescent depression lowers educational attainment and increases the likelihood of adult depression, leading to lower average earnings of about 5% and lower average wages of about 3.7%. These results are robust to identification strategy and alternative

measurements of several variables. In contrast, using instruments for identification drops the direct effect of adolescent depression closer to zero, suggesting that findings of a direct pathway are driven by omitted variables bias. The findings of this dissertation imply that there are large economic benefits to better preventing and treating adolescent depression. The long-term labor market consequences of adolescent depression can also be avoided by addressing the effects of adolescent depression on mediating outcomes like educational attainment and adult depression.

ADOLESCENT DEPRESSION AND ADULT LABOR MARKET OUTCOMES

by

Will Queen

A Dissertation

Submitted to

the Faculty of The Graduate School at

The University of North Carolina at Greensboro

in Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

Greensboro

2022

Approved by

Dr. Jeremy Bray
Committee Chair

DEDICATION

To Leeanna, for your abundant love and support.

To my parents, family, friends, and colleagues, for all that you have done over the years to support me. I would not be where I am without you.

APPROVAL PAGE

This dissertation written by Will Queen has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair

Dr. Bray

Committee Members

Dr. Gicheva

Dr. van Hasselt

March 3, 2022

Date of Acceptance by Committee

March 1, 2022

Date of Final Oral Examination

ACKNOWLEDGEMENTS

I would like to thank Jeremy Bray for his mentorship, guidance, and support over the past five years. Most of what I have learned at UNCG, in one way or another, is thanks to him. I would also like to thank Dora Gicheva and Martijn van Hasselt for all that they have taught me, both in the classroom and during their time on my committee. I would like to thank Stephen Holland for his mentorship and support, especially during Green Fund projects.

I would like to thank Martin Andersen and Chris Swann for always letting me barge into their office with unwanted econometrics questions. I would also like to thank the rest of the faculty in the Department of Economics for their instruction, as well as Jess Saunders and Ashley Peters for their administrative guidance along the way. I would like to thank my many classmates and colleagues in the graduate program over the years, including Lorissa Pagan, Kelsi Hobbs, Cody Morris, Anurag Pant, Nina Davis, Sarah Rahbek, Satyaki Chakravarty, Rashed Sardar, Thea Liu, Hitanshu Pandit, Nana Addai, Danny Turkson, and Kayleigh Willis. I have learned a lot from working with you over the years, and I appreciate all of your help and support.

Additionally, I want to thank Stephen Sills, Ken Gruber, Sonja Frison, and all of my colleagues in the Center for Housing and Community Studies. They have taught me valuable lessons in what it means to do applied research. I would also like to thank Sean McInnes, Nihal Al Raees, and my colleagues in Facilities Operations.

Finally, I would like to thank my family for their support. I would especially like to thank my parents, Rebekah & Jeff Tozer and John Queen. Success in graduate school, and life more generally, would not be possible without your help. A special thanks to my brother and roommate, Ethan Queen. Our Barberitos trips and many nights laughing on discord have kept me sane. Lastly, I would like to thank my fiancée, Leeanna Spellman, for all of her support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

Adolescent depression is a serious health issue with significant consequences. Depression is one of the most common mental disorders in the United States and is becoming more prevalent among adolescents. Between 2004 and 2019, the percentage of youth ages 12-17 in the U.S. with a past year major depressive episode increased from 9.0% to 15.7% (Substance Abuse and Mental Health Services Administration, 2020). This compares to an increase from 7.6% to 8.9% among adults ages 26-49. Depression (Major Depressive Disorder) is a mood disorder that has symptoms such as feelings of worthlessness, irritability, depressed mood, anxiety, poor concentration, and decreased interest in regular activities (Tylee, 2005). Physical symptoms can also occur, including sudden changes in appetite, difficulty sleeping or oversleeping, and physical aches and pain. Social, genetic, and environmental factors can all contribute to the onset of depression (Belmaker & Agam, 2008). Rates of treatment for depression remain dangerously low among adolescents. Between 2004 and 2019, less than half of youth with depression ever received treatment (Substance Abuse and Mental Health Services Administration, 2020). This is especially troublesome considering the availability of effective treatment for adolescent depression and suicidal ideation, including cognitive behavioral therapy (Weersing et al., 2017).

Symptoms of depression generally have an immediate impact on day-to-day life, affecting relationships with friends and family, performance in school, and how you feel and act. Research over the past several decades provides evidence of several of these consequences. Symptoms of depression make it especially difficult to do well in school, leading to less classroom engagement, lower grade point averages and test scores, and lower educational attainment (Eisenberg et al., 2009; J. M. Fletcher, 2008; McLeod et al., 2012). Severe cases of depression can lead to increased risks of self-harm – trends in adolescent depression have

followed closely trends in rates of self-harm. For example, Mercado et al. (2017) found that emergency department visits for self-harm approximately tripled for 10 to 14-year-old girls from 2009-2015. Early-onset depression also more than doubles the likelihood of suffering from depression or other mental illness in adulthood, which implies long-lasting effects on health and mental health (Bohman et al., 2010; Gollan et al., 2005; Pine et al., 1999). Youth with depression are more likely to use various substances and have comorbid mental disorders than youth without depression (Hoffman et al., 2012). The wide array of adverse effects of adolescent depression can stretch to nearly every aspect of life.

Adolescents with depression do worse in the labor market as adults – they are less likely to be employed, have lower earnings or wages, and are more likely to rely on government services (Fergusson et al., 2007; J. Fletcher, 2013; Johar & Truong, 2014). While several consequences of adolescent depression are well-understood, the effects of adolescent depression on adult labor market outcomes are not. However, the many consequences of adolescent depression provide insight into how it may impact labor market outcomes. It could be that youth with depression achieve a lower level of education than they would have otherwise, which leads to lower paying opportunities in adulthood. It may also be that early-onset depression's effect on depressive symptoms later in life has an impact on labor supply in adulthood. Several other factors may confound the relationships between adolescent depression and adult labor market outcomes. Family characteristics (e.g., family income, parents' education), individual characteristics (e.g., race, gender, age), environmental characteristics (e.g., school, social groups, neighborhood, economic circumstances), and health characteristics (e.g., health behaviors, body mass index) could be driving the association between adolescent depression and success in the labor market. Differentiating between causal effects and confounding factors is key to

understanding whether and how adolescent depression affects adult labor market outcomes. A mediation analysis approach is essential to conceptualizing direct and indirect pathways, as it explicitly designates the roles of confounding and mediating variables.

Estimating the causal effects of adolescent depression on labor market outcomes is made especially difficult by several issues of endogeneity. First, adult depression may be endogenous in an earnings equation, resulting in bias in the estimated effect of adult depression on earnings. Second, educational attainment is commonly treated as endogenous in an earnings equation due to its relationship with unobserved ability, potentially biasing estimates of the returns to education. If these estimates are biased, estimates of the indirect effects of adolescent depression on earnings will be biased. Finally, estimates of the marginal (direct) effect of adolescent depression on earnings could be biased if there are important omitted variables, measurement error, or misspecification. For example, a family history of depression could be related to higher rates of adolescent depression and lower socioeconomic status. This sort of bias could have large implications for the types of effects that are estimated and how we interpret them. More creative and effective identification strategies are needed to better estimate the direct and indirect effects of adolescent depression on labor market outcomes.

Understanding the magnitude of the labor market consequences of adolescent depression is crucial to better motivating policy and interventions that address these consequences. Identifying relevant pathways of effects and their importance may inform us of creative or effective ways to address gaps in outcomes. For example, if adolescent depression lowers earnings entirely through its effect on educational attainment, then a recommended course of action differs greatly from one that would arise if adolescent depression lowers earnings entirely through its persistence into adulthood.

In this dissertation, I make two important contributions to our understanding of the effects of adolescent depression on adult labor market outcomes.

First, I use mediation analysis to define effects and contextualize the previous literature. I treat adult earnings and hourly wages as dependent variables and adolescent depression as the 'treatment' variable. I estimate the indirect effects of adolescent depression through two mediators: educational attainment and adult depression. This is the first analysis in the literature that treats both educational attainment and adult depression as mediators. All other factors are treated as confounders. I model direct and indirect effects using the difference method, and I estimate effects with Ordinary Least Squares and a school fixed effect. I argue that, conceptually, the direct effect of adolescent depression should be zero – any non-zero direct effects could be the result of omitted variables bias or measurement error. This approach stands in contrast to some of the previous literature but clarifies the pathways of effects that we are estimating and how we should interpret results. While estimating effects with the difference method, I replicate some previous literature, which clarifies seemingly conflicting results in the literature and reveals that my approach is consistent with previous findings.

Second, I use instrumental variables to identify the causal effects of adolescent depression on adult earnings and wages. I use a system of equations with jointly distributed error terms to model the direct and indirect effects. I use a measure of average cohort religiosity as an instrument for adolescent depression, which provides exogenous variation in adolescent depression, better identifying the direct effect and its effect on adult depression. I use measures of Census-tract level school enrollment and history as instruments for educational attainment, and I use two measures of traumatic events as instruments for adult depression. These instruments allow me to better identify the marginal effects of educational attainment and adult

depression on labor market outcomes. I estimate the system of equations with Limited Information Maximum Likelihood (LIML) and bootstrap standard errors for indirect effects. This approach provides improved estimates of direct and indirect effects that are conceptually consistent with how adolescent depression ought to affect labor market outcomes. After presenting results for earnings and wages, I explore several robustness checks and find that results are not sensitive to the measures of depression used, the measure of education used, or family-level heterogeneity.

The dissertation is laid out as follows. In Chapter II, I provide a survey of the previous literature on adolescent depression and adult labor market outcomes. In Chapter III, I set up the conceptual framework used for understanding relationships and discuss challenges of endogeneity. In Chapter IV, I detail the methods I use to identify parameters and estimate effects. In Chapter V, I give an overview of the data used in this dissertation. Chapter VI gives a detailed overview of descriptive results, including results using the difference method. Chapter VII presents results from the system of equations and instrumental variables, detailing estimates of direct and indirect effects. This chapter also explores the sensitivity of my results to several factors. Finally, Chapter VIII discusses the implications for the findings and concludes the dissertation.

CHAPTER II: BACKGROUND

This chapter accomplishes two purposes. First, it provides an overview of the previous literature focused on adolescent depression and adult earnings, which is a necessary background for this dissertation. Second, it discusses the strengths and shortcomings of the literature, which provides context for the contributions of this dissertation.

Adolescent Depression and Adult Labor Market Outcomes

Youth who experience depression are worse off in the labor market as adults. A growing literature studies whether and why symptoms of depression relate to various labor market outcomes.

Fergusson et al. (2007) uses a nationally representative sample of young adults in New Zealand to estimate how depressive episodes in adolescence relates to long-term psychiatric conditions, education, and economic outcomes. They use linear regression to estimate the effect of the number of depressive episodes on various outcomes. Controlling for confounding factors (e.g., gender, parental attachment, exposure to childhood abuse) and co-occurring disorders, they find that adolescent depressive episodes do not significantly affect yearly earnings. Consistent with previous literature on the persistence of depression, they find that depressive episodes in adolescence strongly predict depression, anxiety, and suicidal ideation in adulthood. They also observe that adolescent depression increases the likelihood of being unemployed later in life. While they find no direct effect of adolescent depression on earnings, they conclude that the persistence of depression and comorbidity of other psychiatric disorders may be responsible for an effect of adolescent depression on adult earnings. Their findings highlight the role of adult depression as a potential mediator, especially its impact on labor supply.

Smith & Smith (2010) uses the Panel Study of Income Dynamics to estimate the long-term economic costs of adolescent depression. They use recall data on whether respondents had depression in adolescence and use the level of yearly earnings where respondents with zero earnings are included. The authors mention several potential mediating pathways between adolescent depression and earnings, naming adult psychological problems as the primary pathway. Therefore, they estimate a reduced form linear model for earnings where measures of adult depression and education are excluded. Controlling for family characteristics and comorbid health conditions, they find adolescent depression is associated with average drops in yearly earnings of about four to five thousand dollars. To address the potential bias of unobserved family-level heterogeneities, they estimate regressions with a family fixed effect, which cuts this effect in half and leaves it statistically insignificant. This result suggests that the effect of depression on earnings may not be robust to family-level heterogeneities, although issues of omitted variables bias may still drive results. It is also possible that a relatively small sample and number of youth per family drives the imprecision of results.

While the effect of adolescent depression on earnings is sensitive to specification, Smith & Smith (2010) finds that its effect on weeks worked is statistically significant and robust to a family fixed effect. This is consistent with evidence linking psychiatric disorders to lower rates of employment and labor supply (Banerjee et al., 2017; Ettner et al., 1997; Frijters et al., 2014; Kessler et al., 2007). The authors assert that psychological disorders in adulthood are the "principal transmission pathway" by which adolescent mental disorders affect adult socioeconomic status, which affirms the findings of Fergusson et al. (2007).

Fletcher (2013) finds that adolescent depression may have a large direct effect on earnings. He uses data from the National Longitudinal Study of Adolescent to Adult Health to

estimate a linear model for log earnings with a family fixed effect. Fletcher controls for several confounding factors, including family income, gender, age, race/ethnicity, health characteristics, and adolescent substance use measures. Several specifications are run with and without measures of educational attainment and adult depression, as well as school and family fixed effects. The paper finds that when adult depression is omitted and a family fixed effect is included, adolescent depression leads to a statistically significant 21% average decrease in yearly earnings. Once adult depression is added to the model, the coefficient decreases to about 16% and becomes statistically insignificant. Fletcher also estimates the effect of adolescent depression on employment. Reduced form estimates find that adolescent depression reduces the probability of being employed by about 6 percentage points. Controlling for educational attainment and adult depression reduces this estimate to about 3-4 percentage points. This suggests that parts of the effect may occur through mediating factors, while part of the effect remains direct. Estimates are robust to family fixed effects.

In agreement with previous literature, Fletcher (2013) finds that adolescent depression explains earnings in part through its relationship with adult depression. In contrast to evidence from previous literature, a large portion of its effect remains after including a family fixed effect and controlling for adult depression. However, the robustness of these effects to family-level heterogeneities is still unclear. Neither Smith & Smith (2010) nor Fletcher (2013) compare family fixed effects estimates to OLS models run on the identifying sample (i.e., youth with within-family variation in adolescent depression), leaving the possibility that results are being driven in part by sample selection (Miller et al., 2019). In spite of this, Fletcher (2013) gives us reason to reconsider the previous findings in the literature.

Evensen et al. (2017) uses a Norwegian health survey to estimate the relationship between various adolescent mental health problems and adult earnings. They estimate linear models for log earnings with and without a family fixed effect and use a measure of internalizing problems for depression that includes both symptoms of anxiety and depression. All specifications do not include measures of education or adult mental health. Conditional on family and demographic characteristics, internalizing problems reduce earnings by about 6%. Adding a school fixed effect drops the effect down to about 3%; adding a family fixed effect instead drops it further to about 2.5% and it becomes statistically insignificant. This suggests that adolescent internalizing problems may still affect earnings even after holding constant school and family-level heterogeneities. It could also be that family characteristics not constant across siblings could be playing a role. Attention problems in adolescence appear to have a stronger and more robust effect on earnings.

Johar & Truong (2014) is the first paper to explicitly estimate how adolescent depression affects a labor market outcome through mediating variables. They posit that adolescent depression affects hourly wages through educational attainment, years of experience, and occupation choice. Using the National Longitudinal Study of Youth 97, they estimate marginal effects of depression on each outcome and calculate mediated effects. They find that 30% and 66% of the total effect of adolescent depression on wages (for males and females, respectively) flows through these indirect channels, which is relatively consistent with previous findings that about half of the effect is mediated. They do not control for adult depression, so estimates of the direct effect may also include an effect mediated by adult depression. Their estimates of the total effect range from a 10-15% drop in wages, on average. Although the measure of depression used is not externally valid, the authors explicitly highlight the significant mediating role of human

capital accumulation. This point was made as early as Mullahy & Sindelar (1993), which noted that ignoring how mental illness affects human capital is an important omission when estimating its effects on earnings. Johar & Truong (2014) make the important point that the direct effect is underestimating the total effect of adolescent depression on outcomes if the indirect effects are large.

Philipson et al. (2020) is the only study to explicitly treat adult depression as a mediator, use clinically relevant measures of depression, and track labor market outcomes over time. They use the Uppsala Longitudinal Adolescent Depression study to track how different types of depressive disorders in adolescence (i.e., ages 16-17), such as Persistent Depressive Disorder or Major Depressive Disorder, relate to earnings over the course of 10 years later in life (i.e., ages 31-40). The measures of depression used are determined by healthcare professionals, which makes results more clinically relevant than literature that uses self-reported measures. They use mediation analysis and linear regression to estimate the effect mediated through adult depression. They find that youth with Persistent Depressive Disorder earn about 17% less than their counterparts, and that this effect is persistent over time. About half of this relationship is mediated by depression in early adulthood, a mediated effect that is especially prevalent for women. In contrast, they find that Major Depressive Disorder does not have a statistically significant effect on earnings. Notably, the study does not adjust for educational attainment, leaving open the possibility that the estimated direct effect could be mediated by education or other measures of human capital. The authors note that effective interventions for adolescent depression are needed to avoid any loss of human capital. Lastly, this study is strengthened by its ability to use labor market outcomes from several points in time for each respondent, rather than a point-in-time measurement.

This section highlights the best evidence that we have that adolescent depression leads to poor labor market outcomes in adulthood.

Where Does the Literature Stand?

Previous literature presents evidence that adolescent depression leads to lower earnings, lower employment rates, and lower wages. However, the mechanisms driving these relationships are not agreed upon. Some studies emphasize the mediating role of adult mental illness, others estimate a direct effect, and others find that human capital accumulation plays a mediating role.

The literature varies in the types of effects that it estimates, as well as how it estimates, discusses, and frames results. Some papers estimate an earnings equation and discuss the coefficient on adolescent depression as the 'effect' of adolescent depression on earnings (e.g., Fletcher, 2013). As a larger group of right-hand side variables are added to the equation, the coefficient usually drops in magnitude until the authors reach a preferred specification. This approach generally ignores the several possible mediating pathways, which may or may not be controlled for on the right-hand side. Some papers use mediation analysis language and estimate what is deemed a direct effect, but they omit one or more mediating variables found to be important in previous literature (e.g., Philipson et al., 2020). For example, educational attainment may be omitted from the earnings equation, which means that the coefficient on adolescent depression can no longer be interpreted as a 'direct effect' unless education is independent of earnings and adolescent depression. A few papers explicitly estimate an indirect effect through one or more mediators. However, the few papers that do this also exclude mediators found to be important in previous literature, such as adult depression (e.g., Johar & Truong, 2014).

While there is a large peer-reviewed literature that motivates the presence of indirect effects (see: Chapter III), there is little evidence supporting the presence of a direct effect. This

reality is not communicated well in the literature. The economic significance and relevance of estimated effects are usually not discussed relative to other estimates in the literature. What difference does it make that part of the effect is through a mediating pathway? What reason do we have to think that a direct pathway exists between adolescent health conditions and adult labor market outcomes? Should we be interested in the direct pathway only, or should we focus on mediating pathways? It is usually assumed that the coefficient on adolescent depression is innately of interest, although its interpretation and significance varies depending on specification.

As a consequence of these points, the magnitude of the total effect of adolescent depression on earnings is also not consistent throughout the literature. Some studies find that any effect on earnings is confounded with family-level heterogeneities and fail to find a significant total effect. Some studies find modest effects ranging from 2-10%, while other studies find total effects as large as 15-20%. These differences are in part due to differences in methods and the types of effects estimated. However, it is not immediately clear whether differences in findings are due to heterogeneities in samples, measures, methods, or the type of effect estimated.

Fully embracing a mediation analysis framework would clarify and improve the current state of the literature. This approach is described at length in the next section. Using consistent language to describe, motivate, and discuss effects would better our understanding of the consequences of adolescent depression. Ultimately, we should be interested in both (1) the total impact of adolescent depression on earnings, and (2) the several pathways of effects. Having an idea of the total magnitude of the effect is important to establish whether there is a causal relationship and to motivate action. Understanding the relative importance of pathways of effects allows us to better prescribe ways to alleviate consequences.

Estimating the direct effect and several indirect effects at once would be a step in the right direction. Interpreting and discussing the estimate of the direct effect would also provide useful direction. If there are mediating variables that are unobserved or difficult to measure, then an estimate of the direct effect would likely include mediated effects. However, without an identification strategy that provides random variation in adolescent depression, unobserved mediated effects and omitted variables bias cannot be separated. Without this discussion and distinction, results could be confusing and uninformative. Finally, discussing the relative importance of pathways of effects has potential policy implications. For example, if educational attainment mediates a significant portion of the total effect, then school-based interventions could be especially attractive options. These implications have not been used to motivate or discuss results in the previous literature.

The literature has focused heavily on the potential endogeneity of adolescent depression in an earnings equation. What has received less attention, however, is the potential endogeneity of mediators in the earnings equation. For example, educational attainment is usually thought to be endogenous in an earnings equation. Individuals with higher innate ability, a higher return to education, and/or a higher socioeconomic background are more likely to both pursue additional schooling and be better off in the labor market. This results in an upward bias on the coefficient on educational attainment. Adult depression may also mediate part of the effect of adolescent depression on earnings, but it could be determined in part by labor market success. Occupation choice, workplace environment, earnings, and time spent at work could all influence symptoms of depression, resulting in bias. Even if the issues of endogeneity surrounding the coefficient on adolescent depression were addressed, estimates of mediated effects may still be biased. Any

discussion of endogeneity surrounding adolescent depression in an earnings equation must also consider the consistency of the coefficients on mediators.

We must also consider whether adolescent depression is endogenous with respect to educational attainment or adult depression. Several factors at the family or environmental level could impact success in school, depressive symptoms in adolescence, and depressive symptoms in adulthood. Even if the marginal effects of these variables on earnings are consistently estimated, omitted variables bias when estimating the impact of adolescent depression on mediators (e.g., education, adult depression) will lead to biased estimates of indirect effects.

In the context of an earning equation, school and family-level heterogeneities are the primary tools used to address endogeneity in the literature. For example, unobserved family history of mental illness could be related to both the likelihood of adolescent depression and success in the labor market. School characteristics such as the presence of a school psychologist could affect the likelihood of adolescent depression and learned skills that are helpful in the labor market. As a result, the coefficient on adolescent depression in an earnings equation could be biased away from zero, overestimating the effect. This issue could also bias other parameters of interest, such as the marginal effect of adolescent depression on educational attainment. In response to these issues, the literature has used fixed effects to account for time-invariant characteristics at the school and family levels.

While a family fixed effect has been used in several cases, the literature has yet to compare the relevant identifying samples. When using a fixed effect, a coefficient is identified by observations from the fixed effect groups that have within-group variation of that variable. When the variable of interest is binary and infrequent, this results in only a minority of the sample identifying the coefficient of interest. As a result, when we compare a coefficients from

models estimated with and without a family fixed effect, two things are changing: adding a family fixed effect and the identifying sample. This may not be a bad thing, but it could introduce a sample selection issue if groups with within-group variation differ in any significant way from groups without that variation. To address this, the 'comparison' model that runs OLS without a fixed effect should use the identifying sample from the fixed effect specification. This ensures that we are seeing the change in the coefficient due to addressing family-level heterogeneities and not due to sample differences. While this issue could be notable when using a school fixed effect, its especially worth addressing when there are only a few observations per group, such as with a family fixed effect.

An alternative approach could be to estimate a system of equations that endogenizes variables like adolescent and adult depression. Allowing error terms to be correlated across equations for adolescent depression, mediators, and labor market outcomes could inform us of whether endogeneity is present. More creative identification strategies using instrumental variables or regression discontinuity design could be used to reduce bias and paint a clearer picture of causal effects. These strategies will need to be employed to better identify several parameters at once if we want to improve estimates of direct and indirect effects. When successfully used, instrumental variables could provide exogenous variation in endogenous variables that gives us less biased estimates of marginal effects.

Obtaining better estimates of adolescent depression's direct, indirect, and total effects on earnings and wages would provide evidence immediately helpful in informing policy around adolescent depression. Understanding the relative importance of mediating pathways will emphasize what types of interventions would be most effective in reducing these consequences and clear up unclear results currently found in the literature.

CHAPTER III: CONCEPTUAL FRAMEWORK

With a review of the previous literature in mind, this section lays out a theoretical motivation and framework for understanding the relationship between adolescent depression and adult labor market outcomes. This is the underpinning of the descriptive and causal results of this dissertation. I then discuss some of the complexities and difficulties of estimating causal effects.

Understanding Effects Through a Mediation Analysis Framework

It is oftentimes difficult to pin down the source of correlation between two events, conditions, or outcomes that are separated by a long period of time. I refer to the earlier event as the *treatment* and the later event as the *outcome*.

Certainly, the outcome cannot affect the treatment, as the treatment was determined long before the outcome. This leaves three possibilities. First, there could be factors that are correlated with both the treatment and the outcome that spur a correlation. These are called *confounders* or *confounding variables.* If the entirety of the correlation between the treatment and the outcome is driven by confounders, then holding confounders constant should reveal that the treatment and the outcome are, in fact, uncorrelated. If there is a remaining correlation, then the other two possibilities are relevant.

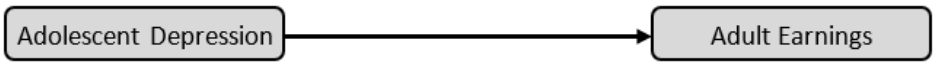
Second, conditional on confounding variables, the outcome could be determined by one or more *mediating variable*s (or *mediators*) that are themselves determined by the treatment. Mediating variables are determined at some point in time between the treatment and the outcome. A change in the treatment causes a change in a mediator, which in turn causes a change in the outcome. This option is especially likely when the treatment and the outcome are separated by a long period of time. Third, the outcome could be determined wholly or in part by observation of the treatment. This could be the case if the outcome is a direct or immediate

function of the treatment. This is called a *direct effect* – conditional on confounding variables, the treatment has a direct impact on the outcome.

This is the terminology of *mediation analysis*, a framework used to estimate the causal pathways between a treatment and an outcome that pays special attention to possible mediated pathways. In this section, I outline the relationship between adolescent depression and adult earnings using a mediation analysis framework. This will highlight the important underlying mechanisms of this relationship, the pathways we should be focused on understanding, and how to define and discuss these effects.

Adolescent depression and adult earnings are outcomes separated by a long period of time. Define adolescent depression as the treatment and adult earnings as the outcome, represented by the directed acyclic graph (DAG) in Figure 1 below. A solid line represents a causal relationship, while a dashed line represents a confounding relationship.

Figure 1: Adolescent Depression and Earnings



Consider the first possibility, whether confounding variables drive the negative correlation between adolescent depression. Dozens of studies suggest that adolescent depression and adult earnings are confounded by a potentially large set of confounding variables, both observed and unobserved. Females are more likely to be diagnosed with depression or score higher on measures of depressive symptoms, and there is a well-documented gender pay gap (Bishu & Alkadry, 2017; Van de Velde et al., 2010). Depending on context, minority populations are more likely to be depressed in adolescence and face more barriers to success in the labor market, resulting in lower earnings (Thompson, 2021; Wickrama et al., 2005).

Youth from families with a low socioeconomic status, particularly below the poverty line, are more likely to suffer from mental illness (Santos & Ribeiro, 2011). Family socioeconomic status is also strongly correlated with own socioeconomic status (Fox et al., 2016). Some of these correlations are unobserved, such as the connections and human capital that may impact the probability of receiving treatment for depression or being referred to a job opportunity.

Health behaviors such as alcohol, marijuana, and other substance use in adolescence are also positively related to depressive symptoms (Weller et al., 2004) and impact long-term labor market outcomes in a variety of ways (Bray, 2005; Ringel et al., 2006), suggesting they may confound the relationship between adolescent depression and earnings. Geographic variation, social characteristics, and school choice are all correlated with reported symptoms and detection of depression, as well as long-term earnings. Controlling for confounding variables will be an essential part to better understanding how adolescent depression affects adult earnings. Consider the updated DAG:

Figure 2: Adding Confounders

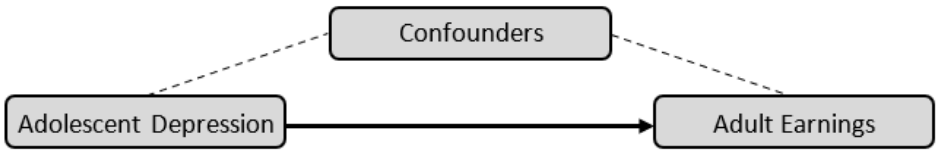


Figure 2 accounts for important confounding variables – the solid black line represents a more accurate estimate of the relationship between adolescent depression and adult labor market outcomes. Now we will consider what mediating variables may mediate the effect of adolescent depression on labor market outcomes.

Education is a clear contender to be a mediator, as there is mounting conceptual and quantitative evidence that it is a determinant of earnings and affected by adolescent depression.

There are large economic returns to education. The relationship between education and wages is one of the most widely studied topics in economics. Most research focuses on the role of education in determining wages, which are usually seen as a proxy for productivity. There are two general theories motivating why education affects wages. The first is known as the human capital model (Becker, 1962). In this approach, an additional year of education provides the recipient with increased knowledge and skills. These acquired skills translate to improved productivity in the workforce, which employers observe. Employers then reward higher productivity with a higher wage. This view of the returns to education generally views returns as a continuous function without kinks in the return to education.

The second approach to understanding the returns to education is known as signaling (Spence, 1973). In this approach, the value of education is not inherent in the skills or knowledge that you gain, but in the signal that you provide employers by demonstrating that you can accomplish a difficult task (e.g., graduating high school or graduating college). A degree or certification signals to employers that the employee has higher productivity than their non-degree-receiving counterparts. Employers do not observe true ability, so they reward signals of ability (i.e., a degree) with higher wages. In contrast to the human capital model, there are several kinks in the returns to education using signaling. For example, this approach points to the increased wages of someone with a bachelor's degree in contrast to someone who dropped out a semester before graduation as evidence of signaling.

Many ascribe to some combination of these approaches, and several recent advancements have been made in the understanding of the mechanisms of how education affects wages, including differences by occupation. In any case, there is robust empirical evidence that additional educational attainment leads to higher wages, which in turn results in higher earnings

(e.g., see Card, 2020). While the magnitude of this relationship and its underlying mechanisms are still up for debate, the presence of a premium on education is settled.

Educational attainment is a function of symptoms of depression. A large literature explores the potential underlying mechanisms of this relationship. Roeser et al. (1998) describes several hypotheses for why emotional difficulties (e.g., depression) are related to troubles in school. First, they present the academic difficulties hypothesis, which suggests that difficulties in school lead to emotional difficulties. This happens because youth begin to form negative beliefs about why they did poorly in school and their ability to do well in the future. Over time, these responses can foster symptoms of depression. An alternative hypothesis is the emotional difficulties hypothesis, where emotional difficulties cause academic difficulties. Emotional difficulties affect the ability to pay attention and the types of memories associated with schooling. The authors mention that emotional and academic difficulties may simultaneously affect one another, making their effects difficult to untangle. Through several empirical illustrations, they find that symptoms of depression are adversely associated with cognition, participation in school, and the overall quality of learning in school.

Alternative, but related, approaches can be formulated using the human capital and signaling models. In the human capital view, workers are incentivized to invest in education until the cost of an additional year of education (including foregone earnings, cost of schooling, etc.) is equivalent to the added benefit of receiving an additional year of education (i.e., an increase in lifetime wages). If the marginal benefit of a year of education is higher than the marginal cost, then they stay in school. Symptoms of depression could increase the marginal cost of attending school. Physically attending school may present additional anguish and unpleasantness that those

without depression may not experience. An increased cost of attending school reduces the optimal amount of schooling for the individual to receive, so they choose to get less schooling.

Symptoms of depression may also decrease the marginal benefit of attending school if symptoms of depression reduce the amount of knowledge retained or skills learned, thereby reducing the productivity gained from additional schooling. In the signaling view of education, symptoms of depression could make it more difficult to perform well in school, resulting in a decreased likelihood in finishing a degree. With this in mind, the expected benefit of attending school diminishes since the youth are less likely to be able to successfully signal in the labor market. Finally, the marginal benefit of schooling could decline if those with depression adjust their occupational preferences. For example, a youth with depression may shy away from occupations with high levels of stress or demanding schedules due to their current mental health state. This could lower their benefit of schooling and lower their optimal level of schooling.

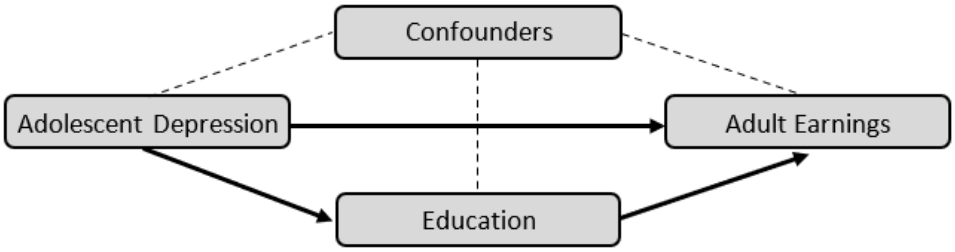
Several studies have found empirical evidence of these potential mechanisms. Eisenberg et al. (2009) finds that depression is associated with a 0.3 standard deviation drop in GPA in the same semester for college students. When analyzed longitudinally and with an individual fixed effect, the effect of depression is pronounced when depression cooccurs with generalized anxiety. They also find that a large increase in symptoms of depression significantly increases the probability of dropping out, which ultimately affects educational attainment. Frojd et al. (2008) finds that depression is related to a lower GPA and changes in several subjective measures of schooling, such as an increase in the perceived load of schoolwork, more frequent difficulties concentrating and paying attention, and difficulties doing homework and activities that require initiative. This suggests that depression is related to a variety of adverse schooling outcomes.

McLeod et al. (2012) finds that the cooccurrence of depression with anxiety, delinquency, and/or substance use is associated with significant drops in GPA and highest degree received, highlighting the role of comorbid disorders. Berndt et al. (2000) uses the context of a randomized clinical trial and finds that women with early-onset depression were significantly less likely to graduate college than women with late-onset depression or no history of depression.

Fletcher (2008) uses the National Longitudinal Study of Adolescent to Adult Health to study how diagnosis and treatment of depression relate to educational attainment. Controlling for several important family, neighborhood, and individual characteristics, the paper finds that depression in high school significantly predicts the probability of dropping out of high school for females and not males. Depression is also linked to a lower probability of enrolling in college, which is an especially strong relationship for females. In a follow-up study, Fletcher (2009) finds that a standard deviation increase in depressive symptoms leads to a 25-30 percentage point increase in the probability of dropping out of high school using a sibling sample. Smaller adverse effects are found on the probability of enrolling in college and all effects are robust to within-family heterogeneities.

The literature finds that several measures of childhood, adolescent, and young adult depression are related to lower GPA, poor performance in school, and lower educational attainment. Combined with evidence that educational attainment positively predicts earnings, we have sufficient conceptual and empirical evidence to suggest that educational attainment mediates part of the effect of adolescent depression on earnings, as is represented below in Figure 3.

Figure 3: Education as a Mediator



There is also strong conceptual and empirical evidence supporting adult depression as a mediator between adolescent depression and adult labor market outcomes. First, adult depression helps determine labor supply, which ultimately influences yearly earnings.

One approach to modeling the decision to work is to start with a model of a reservation wage. Consider a worker who is unemployed but looking for work. The reservation wage is the lowest wage the worker would be willing to accept in lieu of staying unemployed. If the worker is offered a wage below their reservation wage, then they would rather stay unemployed than work. If they are offered a wage at or above their reservation wage, then they would take the job. Among other things, the reservation wage can be a function of the worker's health or mental health. This can show up through their utility of leisure or their disutility of work. Increased symptoms of depression can then lead to an increase in the reservation wage, lowering the likelihood that the worker takes a job and becomes employed.

This approach could be extended to a search model, where the person is searching for a job every period. Rather than be randomly assigned to a job, workers generally have to search for a job and decide whether or not to take the job based on their reservation wage. The type and frequency of wage offers they receive is a function of search intensity, among other things, which bears a cost. Increased symptoms of depression can raise the cost of searching for a job

and impact the intensity of the job search, ultimately leading to fewer job offers and a lower probability that they receive an offer above their reservation wage.

Conditional on being employed, symptoms of depression could also impact earnings via changes in productivity and long-term impacts on the likelihood of being promoted. Just as symptoms of depression could affect one's ability to do well in school, symptoms of depression could also affect how productive one is at work. While the effects of these symptoms may not be immediately obvious to an employer, over time they can have an effect on their productivity, revealing lower productivity and leading to lower wages. If we assume that employers reward employees with promotions or career advancements by how productive they are, then increased symptoms of depression can decrease the likelihood of promotion, and thus wages. This could also have long-run implications for the likelihood of finding improved job prospects by switching firms.

Empirical evidence with a variety of methodological approaches highlights several of these potential mechanisms. Ettner et al. (1997) uses the National Comorbidity Survey to assess how psychiatric disorders impact employment, labor supply, and earnings. They find that psychiatric disorders have a small effect on work hours for males, while it has significant adverse effects on earnings for both males and females, conditional on employment. This supports the story that depression leads to lower earnings through a loss of productivity or advancement. Beck et al. (2011) also finds evidence of how adult depression leads to productivity loss at work. They use data on the severity of symptoms and productivity and impairment at work for patients suffering from depression who recently started antidepressants. They find that increased symptoms of depression were associated with significant losses in productivity at work across the scale of symptoms, not only for high levels of symptoms.

Banerjee et al. (2017) uses the US National Comorbidity Survey and the National Latino and Asian American Study with a structural model approach. Psychiatric disorders have significant negative effects on the probability of being employed and absenteeism, leading to large economic losses. Frijters et al. (2010) addresses the two-way relationship between mental health and work using instrumental variables methods. They find robust evidence that an increase in symptoms of mental illness leads to significant drops in the probability of employment. For example, a one standard deviation decrease in 'mental health' decreases the probability of labor force participation by 17 percentage points. Frank & Gertler (1991) use a community-level survey to estimate the impact of mental distress on average earnings. Consistent with individual-level analyses, they find that those with mental distress earn about 20% less than their counterparts, all else constant. Mitra & Jones (2017) finds significant adverse effects of drops in mental health on the probability of working, but no effects on earnings.

These sources are just a subset of the peer-reviewed evidence that adult depression and other mental illness has an effect on adult labor market outcomes, such as earnings, labor supply, and employment. While the exact mechanisms are not agreed upon, evidence suggests that we should expect a strong relationship between adult depression and earnings or wages.

A large literature in psychiatry, psychology, and related fields find strong links between adolescent and adult depression. For example, Pine et al. (1999) finds that symptoms of depression in adolescence are associated with a significant increase in the probability of depression in adulthood. A two-standard deviation increase in adolescent symptoms predicts a greater risk of a major depressive episode in adulthood by two to three-fold. Gollan et al. (2005) follows a sample of adults who recovered from depression after cognitive behavioral therapy and measures their depressive symptoms every six months for two years post-treatment. Holding

factors like age, education, treatment, and history of depression constant, those with early-onset depression (i.e., before age 20) experience significantly more depressive symptoms than the late onset group. The authors suggest that early-onset depression leads to a shorter time to 'relapse' of symptoms.

Several other papers confirm the relationship between early-onset depression and increased depressive symptoms and probability of depressive episodes later in life (e.g., McLeod et al., 2016; Fergusson et al., 2007).

Adolescent depression also has several significant comorbidities in adolescence and later in life. Hoffmann et al. (2012) uses administrative data on about 140,000 German adolescents to assess the prevalence of depression and its comorbidities. They find that about 63% of adolescents with depression also have at least one comorbid mental health disorder. Internalizing disorders, such as anxiety and emotional disorders, are the most common comorbidities. Using data from the Oregon Adolescent Depression Project, Lewinsohn et al. (1998) finds that youth with depression are significantly more likely to suffer from anxiety disorders, substance use disorders, conduct use disorder, and attention deficit hyperactivity disorder. An increase in comorbid disorders worsens symptoms, complicates treatment, and could lead to adverse long-term outcomes.

Bohman et al. (2010) uses a 15-year follow-up study of adolescents in Sweden to study long-term outcomes of adolescent depression. They find that youth with depression are significantly more likely to have inpatient stays due to mental health and substance use disorders in adulthood. Fergusson & Woodward (2002) finds that youth with depression are at a significantly increased risk of depression, anxiety, alcohol abuse, nicotine dependence, and suicidality in adulthood than youth without depression. Adolescent depression is also associated

with poorer physical health outcomes later in life, even after controlling for adult depression and treatment. Keenan-Miller et al. (2007) finds that adolescent depression is associated with poorer self-reported health, poorer interviewer-rated health, difficulty working due to health concerns, and a greater use of health care services.

Now that evidence has been presented that adult depression is a likely mediator between adolescent depression and adult labor market outcomes, consider the updated DAG below.

Figure 4: Considering Education and Adult Depression as Mediators

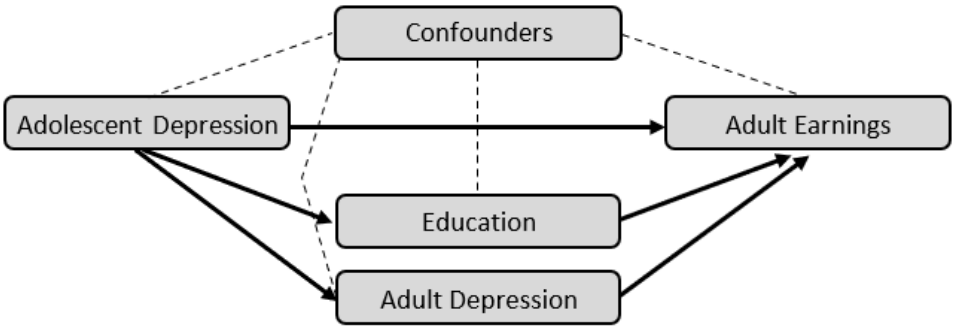


Figure 4 is the final DAG that I use to model relationships. Adolescent depression impacts adult earnings through its impacts on education and adult depression. Controlling for confounding variables is essential, as they may impact the desired effects. The aforementioned literature review reveals that different mechanisms could affect labor market outcomes differently. For this reason, I split up analyses on earnings and wages to better understand how much of effects on earnings are driven by impacts on productivity vs. labor supply. I expect education and adult depression impact earnings in different ways – education may primarily affect wages, while adult depression could play a stronger role in decisions on labor supply. In addition to earnings, looking more specifically at wages or labor supply may allow us to test some of these pathways.

To formalize this approach, consider the following set of recursive equations, where Y is log earnings, D_1 is adolescent depression, D_2 is adult depression, E is educational attainment, C contains confounding variables, S represents school-level factors, and the ε terms are error terms.

$$Y = y(D_1, D_2, E, C_Y, S, \varepsilon_Y) \quad (1)$$

$$D_2 = d_2(D_1, C_{D_2}, S, \varepsilon_{D_2}) \quad (2)$$

$$E = e(D_1, C_E, S, \varepsilon_E) \quad (3)$$

The *total effect* of adolescent depression on earnings can be written as the total derivative of earnings with respect to adolescent depression, $\partial Y_i/\partial D_{1i}$. If we only consider equation 1 and assume that all variables confound the relationship between adolescent depression and earnings, then the total effect is equal to the direct effect, $dY_i/dD_{1i} = \partial Y_i/\partial D_{1i}$. When we treat adult depression and education as mediators and consider equations 2-3, the expression becomes the following, where the first term is the direct effect, the second term is the indirect effect through adult depression, and the third term is the indirect effect through education:[1]

$$dY_i/dD_{1i} = \partial Y_i/\partial D_{1i} + (\partial Y_i/\partial D_{2i})(\partial D_{2i}/\partial D_{1i}) + (\partial Y_i/\partial E_i)(\partial E_i/\partial D_{1i}) \quad (4)$$

The total effect is comprised of three terms: a *direct effect*, an *indirect effect* via adult depression, and an *indirect effect* via education. The direct effect is the effect of adolescent depression on earnings, holding mediators and confounding variables constant. Confounding variables do not change with respect to adolescent depression; hence they do not show up in equation 4.

One relationship that has not been discussed at length yet is the direct effect. The direct effect is the immediate impact of adolescent depression on earnings, conditional on mediating

[1] If the reader would like a copy of this dissertation with equation formatting that is not terrible, please contact the author at willqueenphd@gmail.com.

and confounding variables. There is little reason to think that adolescent depression has a direct effect on earnings – a health condition in adolescence generally does not immediately impact an outcome 15 years later. However, estimates of the direct effect could be non-zero if there is omitted variables bias, measurement error, or misspecification when estimating equation 1. For example, if influential school or family characteristics relate to both adolescent depression and earnings, then the direct effect may be biased away from zero. Another option is that omitting an important mediating variable could drive the direct effect away from zero. For example, if we omitted educational attainment, then the direct effect would certainly appear larger in magnitude than it ought to be. If education is mismeasured, then the direct effect could also include part of the effect mediated by education that is not adequately measured.

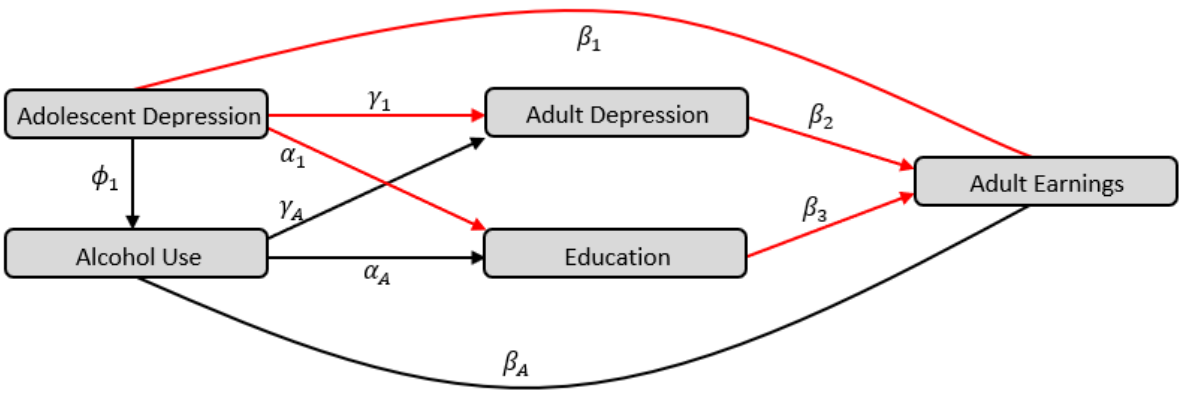
While there is no conceptual reason for the direct effect to be non-zero, it could be estimated as non-zero anyways, and it is important that we understand why. This point will be emphasized throughout this dissertation. Another way to view the direct effect is as the 'unexplained' effect of adolescent depression on earnings – the portion of the effect that we cannot explain. However, not knowing what portion of this effect is being driven by a true causal underlying effect and what portion is simply confounded from omitted variables renders the effect uninterpretable. Better understanding whether the direct effect is picking up on anything causal would be especially helpful for our understanding of pathways of effects.

Mechanisms of Mediated Effects

The true model for how adolescent depression affects earnings is likely much more complex than Figure 4. While education and adult depression are likely important mediating variables, some factors that we are treating as confounding variables may also act as mediators. Namely, there are other mediating variables that could mediate effects of adolescent depression

on other mediators and outcomes. For example, take adolescent alcohol use, A_1. There is evidence that adolescent depression may increase the likelihood of alcohol use, all else constant. Youth suffering from depression may elect to 'self-treat' with alcohol, so we might expect a causal effect flowing from adolescent depression to alcohol use. In turn, alcohol use might affect educational attainment, the likelihood of depression later in life, and even earnings. To represent these changes, we can alter equations 1-3 so that earnings, educational attainment, and adult depression are all also functions of A_1. Additionally, A_1 is also a function of D_1. The diagram now looks as follows:

Figure 5: Adding Another Mediator



In Figure 5, alcohol use effectively serves as a mediator of several relationships and complicates the total effect. The total effect can now be written as follows:

$$dY/dD_1 = (\partial Y/\partial D_2)(\partial D_2/\partial D_1 + (\partial D_2/\partial A_1)(\partial A_1/\partial D_1)) + (\partial Y/\partial E)(\partial E/\partial D_1 + (\partial E/\partial A_1)(\partial A_1/\partial D_1)) + (\partial Y/\partial A_1)(\partial A_1/\partial D_1) + \partial Y/\partial D_1$$

$$dY/dD_1 = \beta_2(\gamma_1 + \gamma_A\varphi_1) + \beta_3(\alpha_1 + \alpha_A\varphi_1) + \beta_A\varphi_1 + \beta_1 \qquad (6)$$

Accounting for the potential mediating pathways of alcohol use provides a more comprehensive understanding of the total effect. Several other factors could take the place of alcohol use or be added in addition to alcohol use. Marital status, involvement with the criminal

justice system, and occupation are a few examples. This dissertation does not explore every possible mediating pathway. Therefore, if any of the aforementioned factors act as mediators, then I inaccurately assume that these effects are zeros and ***underestimate*** the total effect of adolescent depression on earnings. I only estimate the terms that are bolded in equations 5-6, or colored red in Figure 5.

I take a narrow view of the effects mediated through adult depression and education for a few reasons. First, as outlined above, there is ample evidence for their role as mediators. There is also some evidence that they play some of the most important mediating roles here, as highlighted by their prominence in the literature. Second, adult depression and education lend themselves well to potential identification strategies that will be discussed in the following sections. Treating additional factors as mediators may make identification of these effects more difficult. Third, restricting the scope of mediated effects estimated does not bias the mediated effects that we *do* estimate. Mediating variables and confounding variables enter the outcome equation the same, regardless of how we treat them. To estimate the marginal effect of education on earnings, we must hold constant all mediating and confounding variables. In practice, I treat potential mediators (e.g., alcohol use, marital status) as confounding variables in my approach by controlling for them in regression analysis and ignoring the causal pathway between adolescent depression and these variables. While we will underestimate the total effect by ignoring the multitude of other potential mediated effects, we can still estimate the effects mediated through adult depression and education.

The distinction I make between confounding variables and mediating variables is different in concept than it is in practice. The implication of this approach is that the direct and indirect effects of adolescent depression that I estimate will be conditional on other mediating

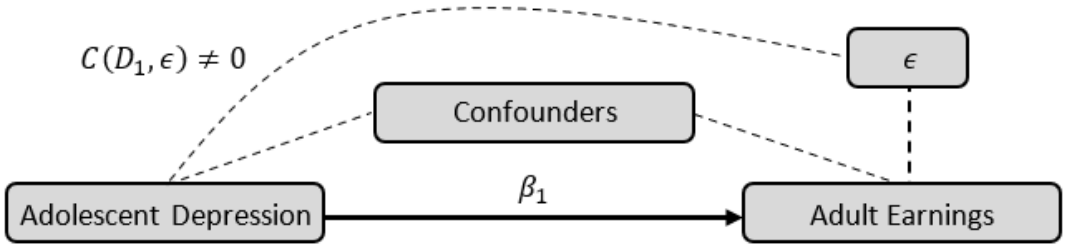
channels. The effect mediated by education will only include how adolescent depression directly affects education, conditional on alcohol use. It will not include the effect of adolescent depression on education via its impact on alcohol use, for example.

This approach is also advantageous compared to removing potential mediators from the model. By including all potential mediating and confounding variables in my approach, I narrow the focus to the relationships of interest. Removing some of these variables would cause the direct effect (β_1) to be biased away from zero. While this change in the direct effect could include some true unobserved mediated effect, it could also still be influenced by omitted variables bias, and we would be unable to separate the two. Including all mediating and confounding variables in my approach results in the most easily interpretable set of results.

The Endogeneity Problem

To estimate the causal effects of adolescent depression on earnings, we must have variation in adolescent depression that is random with respect to earnings. In practice, variation in adolescent depression is not random – it is influenced by a long list of factors that themselves could be related to adult earnings in some way. The fundamental problem in identifying causal effects is isolating variation in the variable of interest that is arguably random with respect to the outcome of interest. This is known as the endogeneity problem. A variable is endogenous when it is correlated with unobserved factors in the error term that are related to the outcome. This prevents us from consistently estimating the marginal effect of the variable on the outcome. Consider the following simplified diagram.

Figure 6: The Endogeneity Problem



Adolescent depression affects adult earnings through β_1, but it is also related to earnings through its association with confounding variables, such as family income. By holding confounding variables constant, we shut down any backdoor pathways connecting adolescent depression to earnings and can consistently estimate β_1. However, there could be unobservable characteristics in the error term ϵ that are related to both adolescent depression and adult earnings. For example, a family history of depression could increase the likelihood of depression in adolescence, while also lowering family income, which ultimately affects adult earnings. If we do not observe and control for potential confounders like family history of depression, then they remain in the error term and bias our estimate of β_1.

One way to approach causal effects is to use an experimental framework, such as a randomized controlled trial (RCT). In an RCT, the 'treatment' (or variable of interest) is explicitly randomized by the researcher. For example, we could assess the effectiveness of a new drug in treating a disease by randomly assigning the drug to a subset of patients who have the disease. This approach addresses the problem of endogeneity by manufacturing random variation in the treatment that is unrelated to other variables. If patients were instead left to decide to use the treatment on their own, then all of the factors that influence their decision to take the treatment could be related to their recovery, ultimately confounding our estimate of a causal effect with variation that we are not interested in. In practice, RCTs are expensive and oftentimes

unethical. For the purposes of this dissertation, we cannot randomize symptoms of depression or educational attainment. But an RCT highlights the gold standard that other methods are attempting to emulate.

Endogeneity is a problem that plagues several relationships of interest in this dissertation. First, adolescent depression and adult depression could be related through a variety of unobserved factors. If we lack measures of family history of depression, family environment, school and social environmental characteristics, and access to treatment for mental health disorders, we could leave important correlates of adolescent depression and adult depression in the error term. To give light to one example, youth who grow up in a community where there is a stigma around depression and receiving treatment could be more likely to experience depression in adolescence, as well as see those symptoms persist into adulthood. This could bias our estimate of the marginal effect of adolescent depression on adult depression. Adolescent depression's impact on educational attainment is also plausibly biased. Difficulties making or keeping friends could influence symptoms of depression and amplify difficulties in school. Socioeconomic status or hardships at home may make performing well in school especially difficult, while also increasing the likelihood of adolescent depression. These factors are difficult to measure and could bias the impact of adolescent depression on education.

Adult depression and educational attainment are also likely endogenous in an earnings regression. While adult depression may influence earnings, socioeconomic status could impact the likelihood of experiencing depression. There could also be unobserved outside events that affect the likelihood of depressive symptoms and labor supply at once, such as a mass layoff event. This could lead to bias in our understanding of how adult depression impacts earnings. Finally, the endogeneity of education in a wages regression is the seminal example of

endogeneity in labor economics. For example, unobserved innate ability could influence performance in both school and the labor market, leading to an overestimated effect of the return to a year of education. Even in a relationship as studied as schooling and earnings, endogeneity is a concern.

Addressing issues of endogeneity is essential to making contributions to our understanding of adolescent depression and adult labor market outcomes. In the following chapter, I detail two approaches I take to addressing these issues of endogeneity.

CHAPTER IV: EMPIRICAL MODEL AND IDENTIFICATION

I use two approaches to identify and estimate direct, indirect, and total effects. First, I use the difference method and linear regression. This approach relies on the conditional independence assumption to identify effects and is most comparable to previous literature. When assumptions fail, it provides an enhanced descriptive view of relationships. Next, I use instrumental variables and a system of equations. This approach addresses several issues of endogeneity by using plausibly random sources of variation in adolescent depression, adult depression, and years of education. To my knowledge, this is the first time that instrumental variables have been used to identify effects in the literature on adolescent depression and adult labor market outcomes.

Regression and the Difference Method

Adolescent depression is not randomly assigned. However, it could be that variation in adolescent depression is as good as random once we have conditioned a sufficiently large set of confounding variables. This is especially likely to be true if we observe data on many of the factors that we believe to be related to both adolescent depression and earnings. This approach to addressing endogeneity and identifying parameters aims to satisfy the Conditional Independence Assumption (CIA). The CIA assumes that adolescent depression is random with respect to the outcome, conditional on a set of confounding variables (C_i). For this assumption to hold, C_i must contain all confounding variables that relate to both adolescent depression and the outcome. For reasons discussed in the previous section, this assumption is unlikely to hold in practice, but this approach can still provide useful descriptive information about relationships.

I use linear regression to estimate direct and indirect effects. I aim to include as many confounding variables as possible to satisfy the CIA. I include rich data on adolescent health,

health behaviors, family background, socioeconomic status, geographic location, and more as confounding variables, which will be detailed in the following chapter. Characteristics at the environment and school level could also be driving issues of endogeneity, so I include an indicator variable for what school each respondent attended. This holds constant school-level factors that could be related to adolescent depression and outcomes. Consider the following earnings equation, where Y_i is log earnings, D_1 is adolescent depression, D_2 is adult depression, E is years of education, C_i is a vector of confounding variables, s is a school fixed effect, and ε_1 is an error term.

$$Y_i = \beta_0 + D_{1i}\beta_1 + D_{2i}\beta_2 + E_i\beta_3 + C_{Yi}\beta_4 + S_i + \varepsilon_1 \qquad (7)$$

If the CIA is satisfied, then β_1 in equation 7 identifies the direct effect of adolescent depression on earnings. Holding all else constant, including adult depression and education, β_1 represents the leftover effect of adolescent depression on earnings.

Conveniently, the CIA is also the identifying assumption for a common approach in mediation analysis: the difference method (T. J. VanderWeele, 2016). It is a simple approach in that it only requires modeling the outcome equation. The outcome equation is specified in several different ways to derive direct, indirect, and total effects. Consider the following four earnings equations.[2]

$$Y_i = \beta_0 + D_{1i}\beta_1 + D_{2i}\beta_2 + E_i\beta_3 + C_{Yi}\beta_4 + S_i + \varepsilon_1 \qquad (8)$$

$$Y_i = \beta_0 + D_{1i}\beta_1 + \qquad + E_i\beta_3 + C_{Yi}\beta_4 + S_i + \varepsilon_1 \qquad (9)$$

$$Y_i = \beta_0 + D_{1i}\beta_1 + D_{2i}\beta_2 + \qquad + C_{Yi}\beta_4 + S_i + \varepsilon_1 \qquad (10)$$

$$Y_i = \beta_0 + D_{1i}\beta_1 + \qquad\qquad + C_{Yi}\beta_4 + S_i + \varepsilon_1 \qquad (11)$$

[2] The numbers in the subscripts of the error terms in equations 8-11 indicate that these error terms differ from one another.

If the CIA holds, then there are no omitted variables related to earnings and either adolescent depression, adult depression, or educational attainment. Since there are multiple mediators, we must also assume that the mediators do not causally affect one another. If both assumptions hold, then we can estimate indirect effects by removing mediators from the earnings regression. Consider equations 9 and 10, where education and adult depression are each removed, respectively. In equation 9, adult depression is now in the error term, which changes the interpretation of coefficients, implying that that β_1^D equals the sum of the direct effect and the indirect effect through adult depression (T. VanderWeele & Vansteelandt, 2014). This identifies the indirect as follows: *Indirect Effect* $(D_2) = \beta_1^D - \beta_1$. To calculate the indirect effect through education, we follow a similar process as in equation 9. We remove educational attainment from the right-hand side, and β_1^E is the sum of the direct effect and the indirect effect through education: *Indirect Effect* $(E) = \beta_1^E - \beta_1$.

Finally, the total effect can be calculated in two ways using the difference method. First, the direct and indirect effects estimated by the above methods can be summed. Second, we can estimate equation 11, which excludes both adult depression and education. If the assumptions described above hold, then the coefficient on adolescent depression is the sum of the direct effect and both mediated effects. *Total Effect* $= \beta_1 + (\beta_1^D - \beta_1) + (\beta_1^E - \beta_1) = \beta_1^T$. If the estimates of the total effects are significantly different, it is an indicator of one or more of the assumptions being violated.

To estimate effects with the difference method with earnings as the dependent variable, I estimate equations 8-11 with Ordinary Least Squares (OLS). I cluster the standard errors at the school level to account for the school-based sampling design of the Add Health.

I also use the difference method to calculate effects with hourly wages as the outcome. If a non-representative sample has observed hourly wages, then estimates could be biased. To address this, I modify the difference method with a Heckman sample selection approach. This approach allows us to estimate parameters in the wage equation that are representative for the full sample, despite only observing wages for a subsample. Consider the following two equations, where Y_i is hourly wages, w_i^* is a latent variable for labor supply and Z_{wi} is an instrumental variable.

$$Y_i = \beta_0 + D_{1i}\beta_1 + D_{2i}\beta_2 + E_i\beta_3 + C_{Yi}\beta_4 + S_i + \varepsilon_{Yi} \qquad (12)$$

$$w_i^* = \gamma_0 + D_{1i}\gamma_1 + D_{2i}\gamma_2 + E_i\gamma_3 + C_{Yi}\gamma_4 + Z_{wi}\gamma_5 + S_i + \varepsilon_{wi} \qquad (13)$$

We observe or do not observe hourly wages as follows:

$$w_i = 1 \text{ if } w_i^* > 0$$

$$= 0 \text{ if } w_i^* \leq 0$$

I assume that the error terms in equations 12 and 13 are distributed joint normal as follows in Figure 7.

Figure 7: Distribution of Error Terms in Sample Selection Model

$$\begin{matrix} \epsilon_Y \\ \epsilon_w \end{matrix} \sim N \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_Y^2 & \sigma_{Y,w} \\ & \sigma_w^2 \end{bmatrix}$$

We can now write the expected value of wages given observed wages as follows, where λ_i is the Inverse Mills Ratio:

$$E(Y_i|w_i = 1)\ X_i\beta + \sigma_{Y,w}\,\sigma_Y\,\lambda_i \qquad (15)$$

Acknowledging the relationship between observing wages and the levels of wages allows us to write out the potential bias of estimates. Estimates are biased if and only if the covariance between the error terms ($\sigma_{Y,}w$) is non-zero. If error terms are uncorrelated, then we have no need

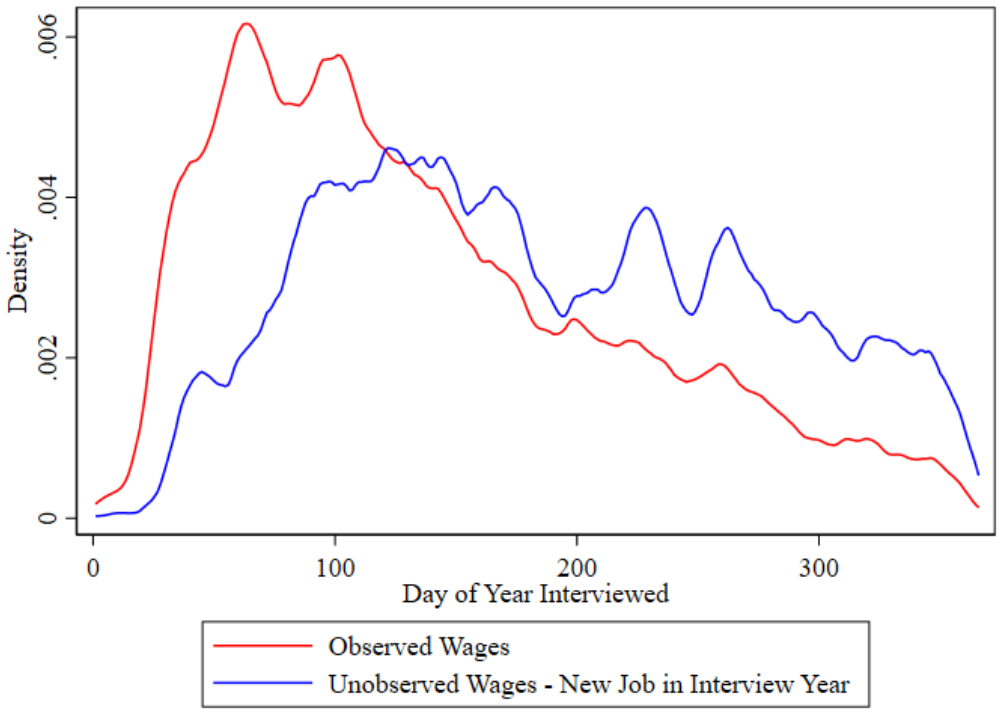
for a sample selection model. If there is sample selection bias, then the bias term is estimated and used to correct estimates.

Technically, equations 12-13 are identified by parametric assumptions alone. However, I use an instrumental variable to ensure that identification is based on more than just assumptions of normality. An instrumental variable in this context is a factor that is both (1) a significant predictor of the probability of observing wages, and (2) conditionally independent of hourly wages. An instrument addresses sample selection bias by using variation in whether or not an observation is in the wage sample that is unrelated to wages. However, most factors that affect labor supply also affect wages. I use a feature of the Add Health interview process and how they ask about labor supply to create an instrument. Whether or not a respondent is in the wage sample is in part due to the timing of their interview. The later in the year the respondent is interviewed, the more likely it is that they started a new job and we do not observe their wages.[3] Conversely, being interviewed at the beginning of the year increases the chances that a respondent who meets the criteria for the wage sample is still at their same job. With this in mind, I use the wave 4 interview date (saved as a number between 1 and 366) as an instrument in the sample selection equation. The structure of the survey makes interview date relevant in the sample selection equation. There is a strong case that interview date is conditionally random with respect to wages. The interview date is determined after last-year's earnings and wages are determined and Add Health interviewers do not observe the respondent's earnings until after the interview date is established (i.e., during the interview date).

[3] Recall that earnings are the previous calendar year's earnings, and hours worked is only asked about for the respondent's current or most recent job. If they started a new job in the interview year, then we cannot match earnings and labor supply.

As Figure 8 shows, a high density of those with observed wages were interviewed towards the beginning of the year. Those who started a new job during the interview (i.e., those on the margin, whose observed wages may depend on interview date) are much more likely to be interviewed later in the year. This is descriptive evidence that interview date addresses the

Figure 8: Interview Date Density by Observed Wages



Note: The density of interview date is plotted for two groups of respondents: those with observed wages, and those with unobserved wages who started a new job in the interview year.

sample selection issue for respondents who would have had observed wages if they had been interviewed earlier in the year.

I estimate equations 12-13 using the Conditional Mixed Process (CMP) package in STATA, which uses Limited Information Maximum Likelihood (LIML) to obtain estimates (Roodman, 2011). In contrast to the two-stage Heckman approach, all parameters and corrections are calculated in one step. A simple test of the presence of sample selection bias is noting

whether the covariance between the error terms is non-zero. The null hypothesis is that the covariance is zero and that there is no sample selection bias. I then use the difference method with this two-equation system to calculate direct and indirect effect. With the two-equation system , I remove a mediator from both equations when calculating indirect effects. I do this to avoid incidentally treating a mediator as an instrumental variable; doing so would severely bias estimates in the wage equation.

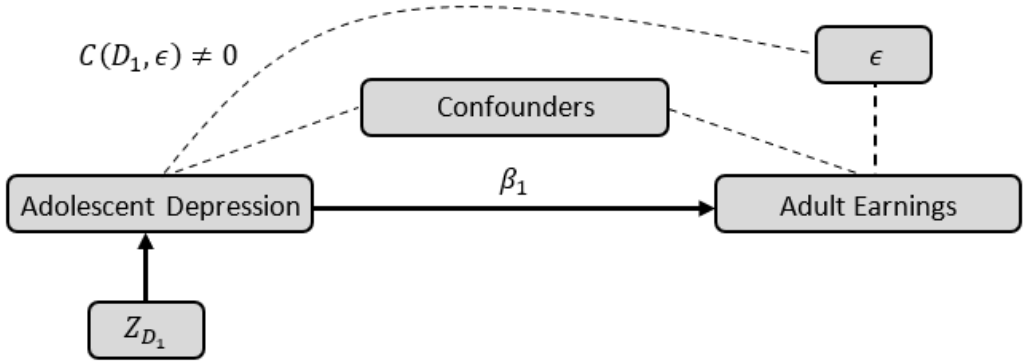
Lastly, the difference method used in this section relies on the CIA holding. If this strong assumption holds, parameter estimates will represent causal effects. In the more likely case that we do not observe all relevant confounding variables and the CIA does not hold, then parameter estimates can be viewed as improved descriptive statistics that account for a wide range of factors. These results will be discussed in Chapter VI: Descriptive Results. Regardless of whether the CIA holds, results will be especially useful due to their similarity to previous literature on this topic. Difference method results will build a bridge between previous literature and the instrumental variables methods described in the next section, which will allow room for discussion of causal effects.

Instrumental Variables and the Product Method

In the case that we do not observe all confounding variables, endogeneity must be addressed in a more creative way. I use instrumental variables and a system of equations as my preferred identification strategy. Instrumental variables do not directly address the omitted variables problem. Instead, an instrument is a variable that explains a significant amount of variation in the endogenous variable that is conditionally random with respect to the outcome. In this way, it ignores how or why the endogenous variable and outcome were related in the first

place, and it uses a completely independent source of variation for identification. Figure 9 shows how an instrument fits into our diagrams.

Figure 9: Instrumental Variables DAG



Since the instrumental variable is the source of identifying variation for parameters, the marginal effect changes in interpretation. Rather than an average marginal effect, using an instrumental variable identifies the Local Average Treatment Effect (LATE). The LATE is the average marginal effect for units whose values of the endogenous variable were influenced by the instrument. In other words, if the instrumental variable only influences a subset of the sample, then the LATE only captures the marginal effect for that subset of the sample.

To discuss the assumptions needed for identification with instruments, consider two equations: one for the endogenous variable and one for the outcome. Three assumptions are needed.

Assumption 1 (A1) – The instrumental variables are orthogonal to all error terms. E.g., $E(Z\,\varepsilon) = 0$, where $Z = \{Z_1, Z_2, Z_3\}$ and $\varepsilon = \{\varepsilon_Y, \varepsilon_{D_2}, \varepsilon_E, \varepsilon_{D_1}\}$.

Assumption 2 (A2) – The instrumental variables are relevant in the endogenous variable's equation (i.e., the rank condition holds). E.g., $\varphi_2 \neq 0$ in equation 19.

Assumption 3 (A3) – The relationship between the instrument and the endogenous variable is weakly monotonic.

Considered together, instruments produce consistent estimates if they are relevant predictors of the endogenous variable and as good as random in the outcome equation conditional on covariates. The instrument must also influence the endogenous variable in the same direction, or not at all, for all units in the sample. If instruments are not sufficiently strong, then the first-stage F-statistic is small and estimates are biased. If the instruments are not excludable, then estimates suffer from omitted variables bias that could be worse than not using instruments at all. Failing the monotonicity assumption precludes consistent estimation of the LATE.

I combine the use of instrumental variables with another way to estimate effects in mediation analysis: the product method. The product method requires modeling equations for the outcome and each mediating variable, but arguably is more intuitive than the difference method. Consider the following system of equations:

$$Y_i = \beta_0 + D_{1i}\beta_1 + D_{2i}\beta_2 + E_i\beta_3 + C_{Yi}\beta_4 + S_i + \varepsilon_{Yi} \qquad (16)$$

$$D_{2i}^* = \gamma_0 + D_{1i}\gamma_1 + C_{D_2i}\gamma_2 + Z_{D_2i}\gamma_3 + S_i + \varepsilon_{D_2i} \qquad (17)$$

$$Y_i = \beta_0 + D_{1i}\beta_1 + D_{2i}\beta_2 + E_i\beta_3 + C_{Yi}\beta_4 + S_i + \varepsilon_{Ei} \qquad (18)$$

$$D_{1i}^* = \varphi_0 + C_{D_2i}\varphi_1 + Z_{D_2i}\varphi_2 + S_i + \varepsilon_{D_1i} \qquad (19)$$

The product method identifies the direct effect (β_1) identically to the difference method – the CIA must hold. The indirect effect is calculated as the product of the marginal effect of adolescent depression on the mediator and the marginal effect of the mediator on earnings. In the case of adult depression, this is $f(\gamma_1)\beta_2$, where $f(\gamma_1)$ is the average marginal effect. For education, this is $\alpha_1\beta_3$. Without instrumental variables, the CIA is a sufficient assumption to identify all parameters for effects.

Although the difference and product methods both require the same initial identifying assumptions, the product method allows instrumental variables to be used for identification. I add an instrument in the adolescent depression equation to identify β_1 and γ_1. I add an instrument in the adult depression equation to identify β_2. I add an instrument in the education equation to identify β_3. When the CIA does not hold and instruments are used, identifying assumptions are replaced by A1, A2, and A3. Each marginal effect needed for direct and indirect effects can be targeted and identified with instruments in the product method, which is a clear advantage over the difference method. Equation 20 maps parameters in equations 16-19 to direct and indirect effects to be identified.

$$dE(Y_i)/dD_{1i} = \beta_1 + \beta_2\gamma_1 + \beta_3\alpha_1 \qquad (20)$$

Instruments

This subsection discusses the several instruments I use for identification. The details of the dataset these instruments come from are discussed in the following chapter.

Instrumenting for Adolescent Depression

I use average cohort religiosity as an instrument for adolescent depression, where cohorts are defined within schools by grade, gender, and race. Religiosity is a positive variable with a range of 0-13 that is the sum of scores from four questions about the religiousness of the respondent (Fruehwirth et al., 2019).[4] Someone who reports no religious affiliation is not asked about their religiosity and is assigned a score of zero. All religious respondents are questioned about how important religion is to them, how often they pray, how frequently they attend church,

[4] Youth who indicate having a religion are asked the following four questions: (1) "In the past 12 months, how often did you attend religious services?" (2) "How important is religion to you?" (3) "How often do you pray?" (4) "In the past 12 months, how often did you attend such youth (religious) activities?" The respondents choose from a Likert scale of responses (e.g., very important, fairly important, fairly unimportant, not important at all). Responses are assigned integer values and summed up, with a higher sore indicating being more religious.

and how frequently they attend religious events designed for youth. Average cohort religiosity is calculated by averaging up the religiosity score for all other respondents in the same school, grade, gender, and race/ethnicity category as the youth.[5] For example, a respondent who is a white male in 8th grade is assigned the average religiosity score of all other white males in 8th grade at their school.

Youth are significantly more likely to be friends with those who are similar to them, known as homophily. Youth consider observable characteristics when selecting friends, which makes it more likely that they are the same age, gender, and race/ethnicity as their friends (McPherson et al., 2001; Shrum et al., 1988). Youth behavior is also substantially influenced by friends after selecting into friendships (Kandel, 2021; Prinstein, 2007). The cohorts I define make up likely candidates for friends of each respondent, implying that cohort characteristics may have significant impacts on own behaviors.

With this in mind, average cohort religiosity may be a relevant predictor of adolescent depression for three reasons. First, changes in peer religiosity motivate changes in own religious behavior (Cheadle & Schwadel, 2012). Increased religious involvement plays a protective effect against mental illness, including depression (Levin, 2010; McCullough & Larson, 1999). Second, peers who are more religious are also less likely to be depressed (Fruehwirth et al., 2019), and symptoms of depression spread between close friends (Conway et al., 2011; Giletta et al., 2012; Prinstein, 2007). This increases the likelihood of own depression. Third, religiosity includes questions regarding attending religious events. Increased attendance to these events represents

[5] Race/ethnicity categories are as follows: white, black, non-white Hispanic, and other.

increased social cohesion of the peers the youth are likely friends with. Increased social connections and cohesion can lower the likelihood of depressive symptoms in youth.

Within-school variation in cohort religiosity is plausibly random in the earnings and adult depression equations after conditioning on the variables used to make the cohorts (i.e., grade, gender, race) and potential mediating variables (e.g., adult religiosity, educational attainment). Although the relationship between cohort religiosity and own depression may not be causal, identifying variation comes from variation in depressive symptoms across cohorts within schools, satisfying A1. I do not address the potential endogeneity of adolescent depression in the years of education equation. Variation in cohort religiosity could be associated with unobserved peer characteristics that affect educational attainment, violating A1 when it is used to identify the effect of adolescent depression on years of education. For this reason, I only use cohort religiosity as an instrument to identify the effect of adolescent depression on earnings and adult depression.

Instrumenting for Adult Depression

I use two measures of potentially traumatic events in the past 12 months as instruments for adult depression: whether a family member or close friend has attempted suicide, and whether a parent or sibling has passed away. Traumatic and stressful life events are an exogenous shock to one's mental health and can increase symptoms of depression (Ettner et al., 1997; Hammen, 2005). Therefore, these events may increase the respondent's symptoms of depression, making A2 likely to hold.

To ensure that A1 holds, the likelihood of these traumatic events must be conditionally random with respect to earnings. These events could impact the respondent's labor supply, ultimately affecting earnings and violating A1. To address this, I control for two measures of

how family influences work. First, I control for the extent to which the respondent reports cutting back on hours due to family responsibilities in the past 12 months. Second, I control for how strongly the respondent agrees/disagrees that family responsibilities interfered with their ability to work in the past 12 months. Conditional on these measures and confounding variables, variation in traumatic events is plausibly random with respect to earnings. These instruments are used to identify the effect of adult depression on earnings.

Instrumenting for Educational Attainment

I use two census tract-level variables as instruments for years of education in the earnings equation. First, I use the tract-level proportion of those aged 16-19 currently enrolled in school. The decisions youth make about their education are influenced by the paths taken by their peers and role-models. Peer networks have significant impacts on the decision to enroll and stay enrolled in school, and students are more likely to conform to their peers' decisions about college applications (Bobonis & Finan, 2009; Rosenqvist, 2018). As a result, local school enrollment is a relevant predictor of own education and A2 holds. Second, I use the tract-level proportion of those aged 25+ who have a bachelor's degree. Evidence suggests that the educational attainment of youth is influenced by the proportion of neighbors with high educational attainment or high-status jobs (Ainsworth, 2002; Ginther et al., 2000). Neighborhood characteristics affect the types of role models that youth are shaped by, ultimately affecting their own educational attainment, and making A2 likely to hold.

To satisfy A1, within-school variation in tract-level characteristics must be plausibly random with respect to earnings. It is possible that families select into neighborhoods within school districts based on their own earnings, education, or race/ethnicity, which are all potentially related to the adolescent's educational attainment. I address these backdoor pathways

by controlling for log family income, mother's education, and race/ethnicity. Neighborhood education characteristics could also be correlated with the propensity to commit crime, ultimately affecting earnings. I address this by controlling for whether the respondent has ever been to jail. Within-school variation in urban/rural status could also be driving long-term earnings differences, so I also control for Rural-Urban Commuting Area Codes. Conditional on the aforementioned variables, within-school variation in tract-level education characteristics is likely random with respect to earnings.

Estimation

I estimate the following four equations as a recursive system using Limited Information Maximum Likelihood (LIML) via the Conditional Mixed-Process Models (CMP) package in Stata (Roodman, 2011):

$$Y_i = \beta_0 + D_{1i}\beta_1 + D_{2i}\beta_2 + E_i\beta_3 + C_{Yi}\beta_4 + S_i + \varepsilon_{Yi} \qquad (21)$$

$$D_{2i}^* = \gamma_0 + D_{1i}\gamma_1 + C_{D_2 i}\gamma_2 + Z_{D_2 i}\gamma_3 + S_i + \varepsilon_{D_2 i} \qquad (22)$$

$$Y_i = \beta_0 + D_{1i}\beta_1 + D_{2i}\beta_2 + E_i\beta_3 + C_{Yi}\beta_4 + S_i + \varepsilon_{Ei} \qquad (23)$$

$$D_{1i}^* = \varphi_0 + C_{D_1 i}\varphi_1 + Z_{D_1 i}\varphi_2 + S_i + \varepsilon_{D_1 i} \qquad (24)$$

I assume that the error terms are distributed joint normally as shown in Figure 10.

Figure 10: Error Distribution in Earnings System

$$\begin{matrix} \epsilon_Y \\ \epsilon_{D_2} \\ \epsilon_E \\ \epsilon_{D_1} \end{matrix} \sim N \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_Y^2 & \sigma_{Y,D_2} & \sigma_{Y,E} & \sigma_{Y,D_1} \\ & 1 & \sigma_{D_2,E} & \sigma_{D_2,D_1} \\ & & \sigma_E^2 & \sigma_{E,D_1} \\ & & & 1 \end{bmatrix}$$

This method allows parameters across equations to be related through the jointly distributed error structure; all variables show up on the right-hand side as observed, not as predicted values. This approach produces consistent estimates in a recursive system given the

assumptions above (Maddala & Lee, 1976; Roodman, 2011). Although I use discrete measures of adolescent and adult depression, I model their latent variables D_1^* and D_2^* and use a probit model for their parts of the likelihood function. The likelihood function for individual i is expressed as follows, where $f(.)$ is the multivariate normal probability density function (PDF):[6]

$$L_i = \iiiint f(\varepsilon_{Yi}, \varepsilon_{D_2 i}, \varepsilon_{Ei}, \varepsilon_{D_1 i})\, d\varepsilon_{Yi}\, d\varepsilon_{D_2 i}\, d\varepsilon_{Ei}\, d\varepsilon_{D_1 i} \qquad (26)$$

When I restrict all covariance terms to zero, the PDF in equation 26 simplifies to $f(\varepsilon_{Yi})f(\varepsilon_{D_2 i})f(\varepsilon_{Ei})f(\varepsilon_{D_1 i})$ and produces estimates equivalent to estimating each equation separately with maximum likelihood. This is how I estimate the system for a baseline set of results. When I add instruments, I estimate the covariance terms between the equations with instruments and the equations for endogenous variables, which allows for the instruments to inform relevant parameter estimates. For example, when I use cohort religiosity to instrument for adolescent depression and identify β_1 and γ_1, I estimate σ_{Y,D_1} and σ_{D_2,D_1}. When I add instruments to the adult depression and education equations to identify β_2 and β_3, I also estimate σ_{Y,D_2} and $\sigma_{Y,E}$. Estimated covariance terms represent endogeneity in a recursive system. If the covariance is different than zero, this is evidence that the conditional mean assumption is violated. For example, the consistency of β_1 relies on the assumption that $E(D_1\varepsilon_Y) = 0$. Expanding and substituting for D_1, we get:

$$E(D_1\varepsilon_Y) = E(D_1)E(\varepsilon_Y) + Cov(C_{D_1}\varphi_1, \varepsilon_Y) + Cov(Z_{D_1}\varphi_2, \varepsilon_Y) + Cov(\varepsilon_{D_1}, \varepsilon_Y) \qquad (27)$$

The first three terms converge to zero by assumption. The term $Cov(\varepsilon_{D_1}, \varepsilon_Y)$ is estimated by σ_{Y,D_1} in the system. If the estimate of $\sigma_{Y,D_1} \neq 0$, then we have evidence to reject the null

[6] Let θ_1 through θ_4 be be the parameters and variables on the right-hand sides of equations 21-24. Let $Q_2 = (2D_{2i} - 1)\theta_2$ and $Q_4 = (2D_{1i} - 1)\theta_4$. The lower bound of each integral is $-\infty$. The upper bounds are as follows, from left to right: $-\theta_1$, Q_2, $-\theta_3$, Q_4.

hypothesis that our estimate of β_1 is consistent. Since identification based on parametric assumptions could be weak and parameters likely suffer from omitted variables bias, I use instrumental variables to better identify parameters. Note that instruments do not address endogeneity by bringing the covariance terms to zero. Instead, instruments provide identifying variation in the endogenous variable that is conditionally uncorrelated with confounding unobservables in the error term. In fact, testing whether a covariance term is unchanged after adding instruments is a test of the validity of the instruments, similar to a Sargan test.

I always restrict $\sigma_{E,D_1} = 0$ to prevent Z_{D_1} from identifying the effect of adolescent depression on years of education. This is done because Z_{D_1} may be related to the error term in the years of education equation, violating A1. While this assumes away a potentially important endogeneity issue for the effect of adolescent depression on education, it avoids another type of bias that could arise from using an invalid instrument. It also avoids using variation from one instrumental variable to identify three separate parameters. When $\sigma_{E,D_1} = 0$, we can rewrite the PDF as:

$$f(\varepsilon_Y, \varepsilon_{D_2}, \varepsilon_E, \varepsilon_{D_1}) = f(\varepsilon_Y, \varepsilon_{D_2} | \varepsilon_E, \varepsilon_{D_1}) f(\varepsilon_E \mid \varepsilon_{D_1}) f(\varepsilon_{D_1})$$

$$= f(\varepsilon_Y, \varepsilon_{D_2} | \varepsilon_E, \varepsilon_{D_1}) f(\varepsilon_E) f(\varepsilon_{D_1}) \quad (28)$$

Since ε_E is no longer conditional on ε_{D_1}, the estimate of α_1 (contained in ε_E) is no longer informed by the parameter on the instrument in the adolescent depression equation (contained in ε_{D_1}). This prevents the instrument from being used to identify the effect of adolescent depression on education (α_1), while still being used to identify its effect on earnings and adult depression (β_1 and γ_1, respectively).

Across all equations, I control for measures of gender, race/ethnicity, grade, adolescent health behaviors (cigarette use, marijuana use, alcohol use), adolescent anxiety, age-standardized

vocabulary test score, log of family income in adolescence, mothers' highest education, adolescent BMI, and indicators for imputed values. In equations 21-23, I also control for whether the respondent has been to jail. In equations 21 and 22, I also control for whether the respondent is enrolled in school in wave 4, the number of children they have, rural/urban status, region, and adult religiosity.[7] The total derivative of earnings with respect to adolescent depression can be written as the following, where γ_1 is an average marginal effect:

$$dE(Y_i)/dD_{1i} = \beta_1 + \beta_2\gamma_1 + \beta_3\alpha_1 \qquad (29)$$

I use a parametric bootstrap routine with 500 replications to estimate the standard errors of the direct, indirect, and total effects in equation 29.

When I use hourly wages as the labor market outcome, I add a sample selection equation to model whether each observation has observed wages. All notation is identical to the sample selection model described in the previous section. The updated system of equations and variance/covariance matrix is below:

$$Y_i = \beta_0 + D_{1i}\beta_1 + D_{2i}\beta_2 + E_i\beta_3 + C_{Yi}\beta_4 + S_i + \varepsilon_{Yi} \qquad (30)$$

$$W_i = \delta_0 + D_{1i}\delta_1 + D_{2i}\delta_2 + E_i\delta_3 + C_{Yi}\delta_4 + Z_{wi}\delta_5 + S_i + \varepsilon_{wi} \qquad (31)$$

$$D_{2i}^* = \gamma_0 + D_{1i}\gamma_1 + C_{D_2i}\gamma_2 + Z_{D_2i}\gamma_3 + S_i + \varepsilon_{D_2i} \qquad (32)$$

$$Y_i = \beta_0 + D_{1i}\beta_1 + D_{2i}\beta_2 + E_i\beta_3 + C_{Yi}\beta_4 + S_i + \varepsilon_{Ei} \qquad (33)$$

$$D_{1i}^* = \varphi_0 + C_{D_1i}\varphi_1 + Z_{D_1i}\varphi_2 + S_i + \varepsilon_{D_1i} \qquad (34)$$

[7] I control for adult religiosity to shut down any backdoor pathway connecting adolescent cohort religiosity to adult earnings when it is used as an instrument.

Figure 11: Error Distribution for Wages System

$$\begin{matrix}\epsilon_Y \\ \epsilon_W \\ \epsilon_{D_2} \\ \epsilon_E \\ \epsilon_{D_1}\end{matrix} \sim N \begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}\sigma_Y^2 & \sigma_{Y,W} & \sigma_{Y,D_2} & \sigma_{Y,E} & \sigma_{Y,D_1} \\ & \sigma_W^2 & 0 & 0 & 0 \\ & & 1 & \sigma_{D_2,E} & \sigma_{D_2,D_1} \\ & & & \sigma_E^2 & \sigma_{E,D_1} \\ & & & & 1\end{bmatrix}$$

Across all specifications with wages, I estimate $\sigma_{Y,W}$ and restrict all other covariances involving the selection equation to zero. The covariance between wages and being in the wage sample ($\sigma_{Y,W}$) represents the degree to which sample selection is an issue. If $\sigma_{Y,W} \neq 0$ then adjusting for sample selection addresses bias. Identical to the difference method, I use interview date as an instrumental variable in the observed wages equation. I use the same instruments, the same estimation method, and the same bootstrapping procedure for hourly wages as I do for yearly earnings.

A Test of Instrument Excludability

This subsection briefly outlines how I test for instrument validity when estimating a system of equations (e.g., equations 21-24). First, consider the covariance term $\sigma_{D_1,Y}$ when no instrument is included in the adolescent depression equation:

$$Cov(\varepsilon_{D_1}, \varepsilon_Y)' = Cov(D_1, \varepsilon_Y) + Cov(C_{D_1}\varphi_1, \varepsilon_Y) + Cov(S, \varepsilon_Y) \quad (36)$$

Second, consider the covariance term $\sigma_{D_1,Y}$ when Z_{D_1} is included as an instrument in the adolescent depression equation:

$$Cov(\varepsilon_{D_1}, \varepsilon_Y)'' = Cov(D_1, \varepsilon_Y) + Cov(C_{D_1}\varphi_1, \varepsilon_Y) + Cov(S, \varepsilon_Y) + Cov(Z_{D_1}, \varepsilon_Y) \quad (37)$$

Taking the difference between the two terms, we get:

$$Cov(\varepsilon_{D_1}, \varepsilon_Y)'' - Cov(\varepsilon_{D_1}, \varepsilon_Y)' = Cov(Z_{D_1}, \varepsilon_Y) \quad (38)$$

The validity assumption (A1) for Z_{D_1} is that equation 38 equals zero. Therefore, we can test the validity of an instrument by testing whether the difference between $Cov(\varepsilon_{D_1}, \varepsilon_Y)'$ and $Cov(\varepsilon_{D_1}, \varepsilon_Y)''$ is zero. I conduct a Z-test for the difference between parameters, where the null hypothesis is that $Cov(\varepsilon_{D_1}, \varepsilon_Y)'' - Cov(\varepsilon_{D_1}, \varepsilon_Y)' = 0$. If the test statistic is sufficiently large, then we can reject the null hypothesis that the instrument is valid.

CHAPTER V: DATA

This dissertation requires data that: (1) tracks a large sample of individuals from adolescence through adulthood, (2) contains reliable measures of depression and labor market outcomes, and (3) contains measures of important mediating and confounding factors. Several data sets satisfy one or more of these requirements. The National Comorbidity Survey (NCS) is perhaps the most comprehensive survey of mental disorders. However, the 10-year follow-up period provides a short window of future outcomes to explore and a limited sample size of adolescents. The National Longitudinal Survey of Youth 97 (NLSY97) contains a wealth of data on youth surveyed every year through adulthood, but their measures of mental health are not specific enough to distinguish mental illnesses from one another.

No data set satisfies these requirements as well as the National Longitudinal Study of Adolescent to Adult Health (Add Health). Add Health is a school-based study that tracks a nationally representative sample of youth in the United States through adulthood. An in-school questionnaire is administered to students in grades 7-12 in 1994-1995; a representative subsample participates in a follow-up in-home interview, which comprises of wave 1 of the survey (mean age of 15.59). Follow-up in-home interviews are given again in 1996 (wave 2), 2000 (wave 3), and 2008 (wave 4, mean age 28.48). Add Health collects a wide array of data from the respondents themselves, family members, peers, and school administrators. Add Health was designed to explore 'the causes of health and health-related behaviors of adolescents and their outcomes in young adulthood,' making it a great match for this dissertation (Harris, 2009).

In addition to the public-use data, I also gained access to the restricted-use data, which allows me to link youth to their school, family, peer group, and community. This enables important within-group comparisons. A rich set of data at the census-block, census-tract, school,

county, and state levels is also available. I use data from wave 1 for adolescent measures and wave 4 for adult measures.

Measuring Adolescent Depression

Wave 1 contains 19 of the 20 items of the Center for Epidemiologic Studies Depression Scale (CES-D), a reliable measure of depressive symptoms (Hann et al., 1999; Radloff, 1977). The CES-D was originally introduced by Radloff (1977) as a self-report scale designed to measure depressive symptoms. In the decades since, the CES-D has been used as a screening tool for depression; respondents who score above a threshold are referred to additional evaluation. The CES-D is also known to be useful for tracking symptoms of depression over time. The CES-D has been widely validated across several characteristics and populations (Blodgett et al., 2021).

Table 1 presents all 20 items of the CES-D and which 19 items are included in the Add Health. To complete the CES-D, the respondent is asked, "How often was each of the following things true during the past week?" and then presented with each item. Respondents choose from the responses of "never or rarely," "sometimes," "a lot of the time," and "most of the time or all of the time." Values of 0-3 are assigned to each response and are summed to calculate a "CES-D Score" with a possible range of 0-57. A higher CES-D score indicates more frequent and/or severe symptoms of depression. Following the recommendation of the Value Options (*Value Options Depression Screening*, n.d.), youth with five or more missing responses to CES-D questions are given a missing CES-D score.

Following the recommendation of the author of the CES-D and subsequent research, I use the cutoff score of 16 to create a discrete measure of depression, where those with a CES-D score of 16 or greater are characterized as having adolescent depression (Radloff, 1977; Weissman et

al., 1977). This is not a diagnosis of depression. Instead, this measure compares youth with a high level of symptoms of depression to youth with a low level of symptoms. The cut-off score of 16 is often used to categorize youth as either having moderate to severe risk of depression or having a low risk of depression. Using the terms 'adolescent depression' and 'not depressed in adolescence' to describe this variable is done for convenience for the rest of this paper. For my main results, I use the dichotomous measure of adolescent depression for convenience – it allows comparison between two groups, one with high risk and one with low risk of depression. I use the CES-D score in sensitivity checks to test robustness of results to measure of depression.

Measuring Adult Depression

Adult depression is measured using the abbreviated 10-item CES-D included in wave 4, where a cutoff score of 10 is used to create a discrete measure of depression (Kilburn et al., 2016; Oppong Asante & Andoh-Arthur, 2015). Table 1 includes details on the items included in each measure of depression. Similar to the measure of adolescent depression, the abbreviated CES-D measures the level of symptoms of depression present in the past week. The discrete measure of depression does not indicate a diagnosis, but a significantly higher risk of depression. Similar to adolescent depression, my main results use the dichotomous measure of adult depression. I use the 10-item CES-D to test robustness of results to an alternative measure of depression.

Measuring Educational Attainment

In wave 4, the respondent is asked to report their highest level of educational attainment. I use this response to create a categorical variable for educational attainment with the following categories: less than high school, high school graduate, some college, bachelor's degree, and more than a bachelor's degree. In my preferred specifications, I use a measure of years of

education that I create by assigning years to each of the reported categories. The following is a list of the categories respondents choose from with the imputed years of education is in parentheses: 8th grade or less (8), some high school (10), high school graduate (12), some vocational/technical training after high school (13), completed vocational/technical training after high school (14), some college (14), completed bachelor's degree (16), some graduate school (17), completed a master's degree (18), some graduate training beyond a master's degree (19), completed a doctoral degree (20), some post baccalaureate professional education (e.g., law school, med school, nurse) (18), and completed post baccalaureate professional education (e.g. law school, med school, nurse) (20). In robustness checks, I use the categorical measure of educational attainment to show the robustness of results to different measures of education.

Earnings, Wages, and Hours Worked

I measure adult earnings using the natural log of the respondent's gross yearly earnings from the previous calendar year, as reported in wave 4. If the respondent does not know their income, then they are asked to give their "best guess" of their gross income by picking 1 of 12 categories listed (less than $5,000, $5,000-$10,000, etc.). I assign these respondents the midpoint of the category they pick. Respondents with zero reported income (about 7% of the sample) are excluded from earnings analyses. I use the natural log of earnings instead of earnings for several reasons. First, the value of an additional dollar varies depending on level of income. A $1,000 drop in earnings could be an important result to take note of for a worker making $10,000 per year, while a $1,000 drop in earnings is more or less noise for someone making $150,000. Taking the natural log of earnings allows me to think about income in relative terms, not in dollar terms. I can estimate effects in percentage terms, which are easily generalizable across the earnings distribution. Second, earnings is right skewed. Taking the natural log brings the

Table 1: Measures of Depreession in the Add Health

	CES-D	Add Health	Feelings Scale	CES-D 20-item	CES-D 10-item
1	I was bothered by things that usually don't bother me	You were bothered by things that usually don't bother you	Y	Y	Y
2	I did not feel like eating; my appetite was poor	You didn't feel like eating, your appetite was poor	Y	Y	N
3	I had trouble keeping my mind on what I was doing	You had trouble keeping your mind on what you were doing	Y	Y	Y
4	I felt that everything I did was an effort	You felt that you were too tired to do things	Y	Y	Y
5	I talked less than usual	You talked less than usual	Y	Y	N
6	I could not get "going"	It was hard to get started doing things	Y	Y	N
7	My sleep was restless	Trouble falling asleep or staying asleep	N	Y	N
8	I felt that I could not shake off the blues even with help from my family or friends	You felt that you could not shake off the blues, even with help from your family and your friends	Y	Y	Y
9	I felt depressed	You felt depressed	Y	Y	Y
10	I thought my life had been a failure	You thought your life had been a failure	Y	Y	N
11	I felt fearful	You felt fearful	Y	Y	N
12	I felt lonely	You felt lonely	Y	Y	N
13	n/a	You felt life was not worth living	Y	N	N
14	I felt sad	You felt sad	Y	Y	Y
15	I had crying spells	Frequent crying	N	Y	N
16	I felt that people dislike me	You felt that people disliked you	Y	Y	Y
17	People were unfriendly	People were unfriendly to you	Y	Y	N
18	I felt that I was just as good as other people	You felt that you were just as good as other people	Y	Y	Y
19	I felt hopeful about the future	You felt hopeful about the future	Y	Y	N
20	I was happy	You were happy	Y	Y	Y
21	I enjoyed life	You enjoyed life	Y	Y	Y

Note: The Feelings Scale is administered in wave 1. The 10-item CES-D is administered in wave 4. Respondents are asked, "How often was each of the following things true during the past week?" and then presented with each item. Respondents choose from the responses of "never or rarely," "sometimes," "a lot of the time," and "most of the time or all of the time." Values of 0-3 are assigned and summed to calculate a "CES-D Score".

distribution closer to normal and makes estimating a conditional mean function easier and less influenced by outliers. Third, using log earnings can avoid potential issues of heteroskedasticity.

Respondents are then asked about their current job (if employed) or most recent job (if unemployed), including the job's start date, end date, and the typical number of hours worked at that job every week. Information about the respondent's current/most recent job and their yearly income may not be related. Each respondent's recent employment history determines whether we have enough information to calculate their hourly wages. Figure 12 details possible labor market histories and whether they result in observed or unobserved wages.

I observe wages for respondents who are either (1) currently employed and have worked at their current job since January of the income year or before, (2) currently unemployed and have worked at their most recent job from January of the income year or before to December of the income year or later, or (3) currently unemployed, started their most recent job in January of the income year or before, and left this job during the income year. Assuming that hours worked per week is constant, average hourly wage is calculated by the following:

$$\textit{Average Hourly Wage} = (\textit{Yearly Income})/((\textit{Weeks Worked})(\textit{Hours Worked per Week})) \quad (39)$$

Weeks worked is assumed to be 52 in the first two cases and is estimated using the end date of the job in the third case. Every other scenario described in Figure 12 prevents the calculation of an hourly wage. For example, someone who started their current job during June of the income year reports weeks worked and hours worked for the latter half of the income year; however, their reported income may also include income from their previous job, for which we do not observe labor supply. As another example, someone who started their current job during the interview year does not provide any labor supply information that overlaps with their income from the income year. Note that since the Add Health only collects data on the respondent’s

primary job, average hourly wage could be overestimated for those who work several jobs at once.

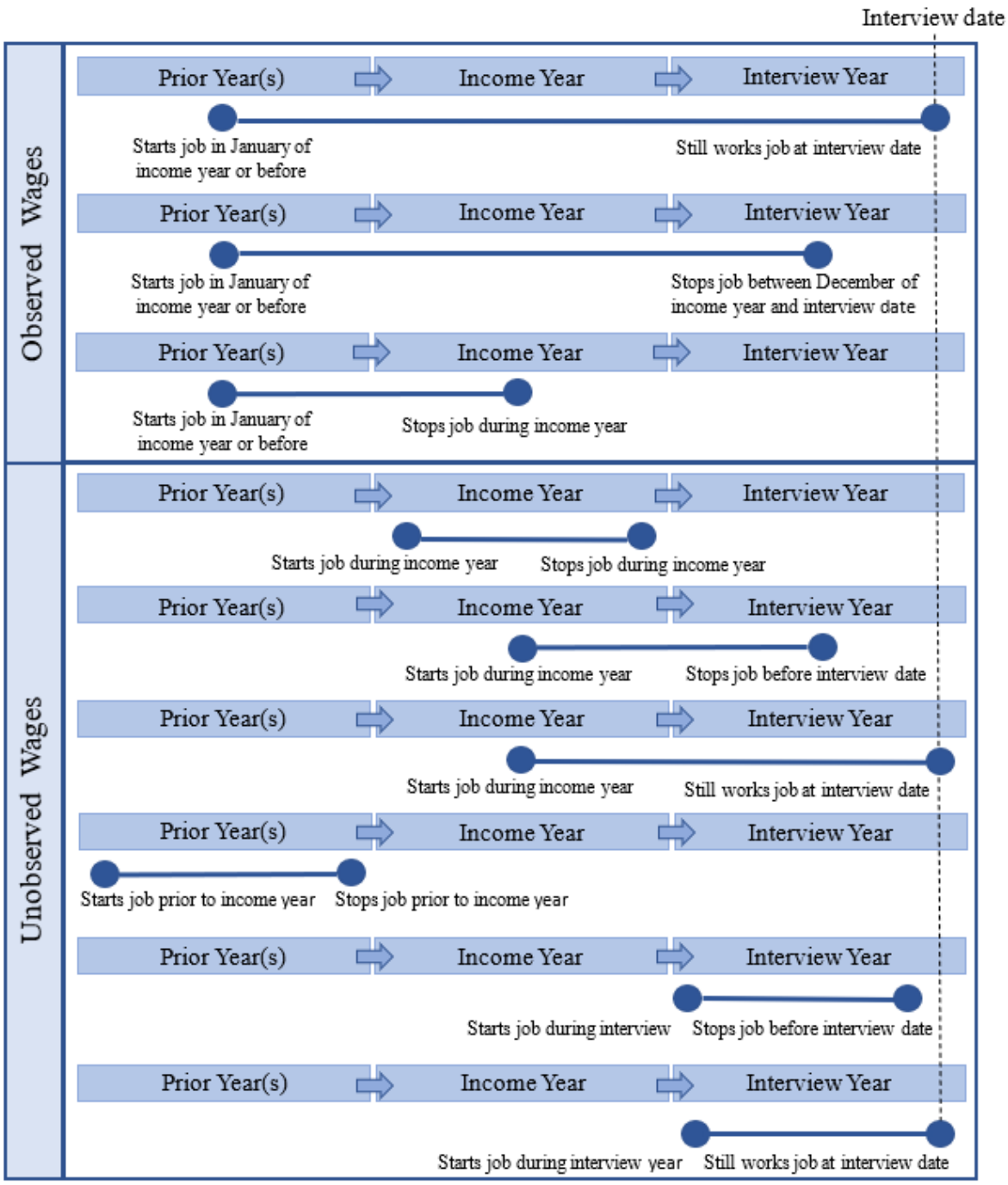
Confounding Variables

I use data in the Add Health to create several measures of variables that may be related to adolescent depression, education, adult depression, and/or earnings. These variables are included in regression analyses in the following sections.

I create binary measures of demographic characteristics, such as grade, gender, race (white, black, Asian, other), and ethnicity (Hispanic or not Hispanic). Using wave 1 data, I create measures of two important family characteristics: log family income and mothers' highest education (less than high school, high school degree, college degree). I categorize adolescent alcohol use into four categories: no alcohol use, drink in the past year, binge in the past year, and 'problem drinking' in the past year. Measures of marijuana and cigarette use in the past month are also created. I measure adolescent health using adolescent body mass index (BMI) and a binary measure of adolescent anxiety based on the Add Health anxiety index. I use an age-standardized score on the Peabody Picture Vocabulary Test to measure the respondent's vocabulary ability, which is a rough proxy for ability.

Using wave 4 data, I create several measures of potentially important characteristics in adulthood. Two family characteristics are measured: a binary measure of marital status (equaling 1 if the respondent is married and 0 otherwise) and a variable equaling the number of children the respondent has. Current school enrollment is also measured using a binary measure. Geographic variation is measured using a categorical variable for region (northeast, southeast, west, Midwest) and categorical variables for Rural-Urban Commuting Area (RUCA) codes. I also measure criminal record with a binary variable equal to 1 if the respondent has been to jail.

Figure 12: Observed Wages by Labor Market History



Samples

There are 15,690 respondents who participated in both the wave 1 and wave 4 interviews of the Add Health. To avoid significant attrition, I mean impute missing values for adolescent BMI and family income. I median impute maternal education. I use a probit model to impute missing values for marital status – race, ethnicity, gender, and a quadratic for age are used as right-hand side variables. I remove respondents who have any missing data for measures of yearly earnings, education, depression,[8] and the list of confounding variables described in the previous subsection. In line with the Value Options Depression Screening, I mark respondents with more than 4 missing items from the CES-D as missing for the adolescent depression measure. In total, attrition of these factors brings the final sample size to 14,561. All respondents have observed data on their school, allowing me to match respondents based on their school.

The Add Health surveyed 5,512 youth from 2,633 families. The family sample in this paper consists of all respondents in the full sample who have at least one other family member in the full sample, resulting in 2,593 youth from 1,261 families.

[8] Those with more than 4 missing items from the CES-D are marked as missing for the adolescent depression measure (Value Options Depression Screening).

CHAPTER VI: DESCRIPTIVE RESULTS

This section provides a descriptive overview of how adolescent depression impacts adult earnings and wages. First, I use a set of tables and graphs to survey relationships. Next, I use the difference method and linear regression to estimate baseline results for direct, indirect, and total effects.

Graphs and Tables

Columns 1-2 of Table 2 presents means and standard deviations of the overall sample. The average respondent is 28.5 years old in wave 4, earned more than $35,000 in the previous year, has more than 14 years of education, and has a 1 in 5 chance of being depressed in adulthood. About 1 in 4 respondents had depression in adolescence, with an average CES-D score of 12.29. Nearly half of all adolescents drank alcohol in the past year, 14% used marijuana in the past month, and almost 1 in 5 smoked cigarettes. Most respondents have a mother with a high school degree, but only about 1 in 4 have a mother with a college degree.

Columns 3-4 of Table 2 present means by adolescent depression. Column 5 presents p-values from a t-test of equality in means between the two groups.[9] Adolescent depression is related to a variety of demographic characteristics and adverse outcomes. First, youth with depression earn about 20% less than those without depression. Youth with depression also receive about 0.75 fewer years of education and are more than twice as likely to have depression in adulthood. Adolescent depression is associated with a statistically significant increased likelihood of alcohol use, cigarette use, marijuana use, and involvement with the criminal justice system. Youth with depression are disproportionately older, female, black, and Hispanic.

[9] The t-test assumes unequal variances between the samples.

Table 2: Descriptive Statistics

	Full Sample		Adolescent Depression		T-Test
			No	Yes	
	Mean	Std. Dev.	Mean	Mean	P-Value
Wave 4 - Adulthood					
Yearly Earnings	35,132	(44,583)	37,060	29,661	(0.00)
Hourly Wages	20.42	(20.85)	21.07	18.40	(0.00)
Currently Employed	0.78	(0.41)	0.80	0.73	(0.00)
Years of Education	14.27	(2.20)	14.45	13.74	(0.00)
Depressed in Adulthood	0.20	(0.40)	0.15	0.34	(0.00)
Adult CES-D Score	6.07	(4.67)	5.33	8.17	(0.00)
Been to Jail or Prison	0.15	(0.35)	0.14	0.16	(0.00)
Age	28.48	(1.76)	28.37	28.78	(0.00)
Wave 1 - Adolescence					
Adolescent Depression	0.26	(0.44)			
CES-D Score	12.29	(6.67)	9.12	21.28	(0.00)
Drank Alcohol Past Year	0.47	(0.50)	0.43	0.58	(0.00)
Binged Past Year	0.26	(0.44)	0.23	0.35	(0.00)
Used Marijuana Past Month	0.14	(0.35)	0.11	0.21	(0.00)
Smokes Cigarettes	0.19	(0.39)	0.16	0.28	(0.00)
Age	15.58	(1.72)	15.48	15.88	(0.00)
Grade	9.65	(1.63)	9.58	9.84	(0.00)
Family Income (Thousands)	44.22	(45.12)	45.36	41.01	(0.00)
Mother has High School Degree	0.86	(0.35)	0.88	0.80	(0.00)
Mother has College Degree	0.26	(0.44)	0.28	0.20	(0.00)
Demographics					
Female	0.53	(0.50)	0.50	0.63	(0.00)
White	0.60	(0.49)	0.63	0.52	(0.00)
Black	0.21	(0.40)	0.20	0.24	(0.02)
Asian	0.09	(0.28)	0.08	0.11	(0.00)
Non-White Hispanic	0.11	(0.31)	0.10	0.13	(0.02)
Observations	14,561		10,760	3,801	14,561

Note: Columns 1 and 2 report the means and standard deviations for the full sample. Columns 3 and 4 report means by adolescent depression status. Column 5 reports the p-value of a two-sample t-test of the equality of means in columns 3 and 4.

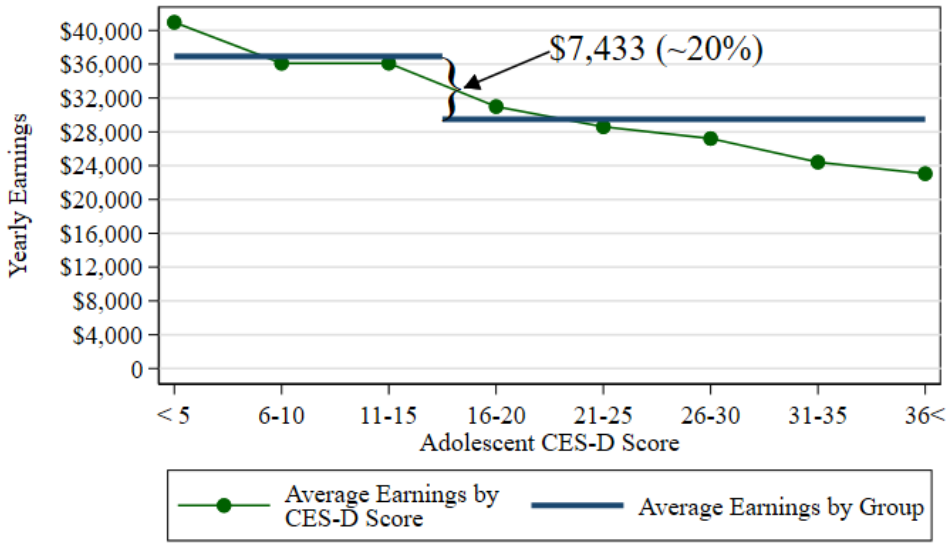
Important differences at the family, school, and community-level are also related to adolescent depression – for example, youth with depression come from families with significantly lower income and lower educated mothers. Figure 13 takes a closer look at the

earnings gap by plotting average adult earnings by adolescent CES-D score. As symptoms of adolescent depression increase, average earnings steadily decrease. When the cut-off score of 16 is used, a 20% earnings gap shows up, although the exact point of the cutoff does not appear responsible for an especially large drop in earnings. This is an economically significant disparity in earnings – from the depressed individual's standpoint, $7,500 per year could be the difference between financial security and economic hardship. Lower socioeconomic status is associated with long-term health problems and a reliance on government subsidies and programs. When viewed through a broader lens of labor economics, a drop in earnings represents a drop in both productivity and overall economic output. To put this gap into context, its magnitude falls between the black/white earnings gap (17.7%) and the gender earnings gap (29.9%) found in this sample. Of course, both gender and race likely confound the relationship between depression and earnings.

Figure 14 plots average adult hourly wages by adolescent CES-D score. This figure only includes respondents who had an observed hourly wage in wave 4. We see a similar pattern in wages as we do yearly earnings. As symptoms of depression increase, average wages decrease. Note that the drop in wages, while notable, is not as dramatic as the drop in earnings in Figure 13. On average, youth with depression have hourly wages about 12.7% lower than youth without depression. Although there are potential measurement issues with this measure of wages, this implies that a significant portion of the relationship between adolescent depression and earnings shows up in both labor supply and productivity.

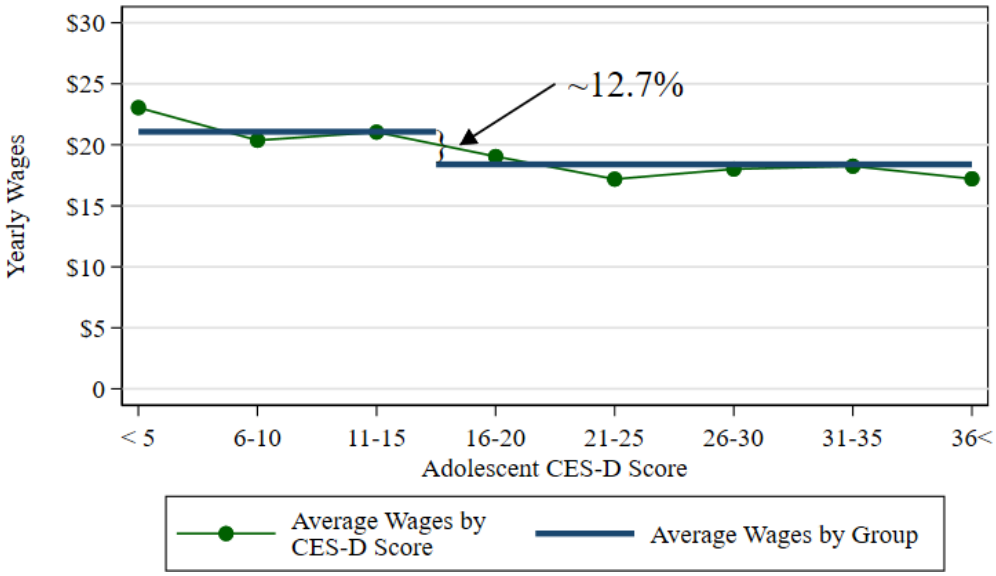
Several individual, environmental, and family characteristics may confound the earnings gap between youth with and without depression. These characteristics must be considered to understand the magnitude of the effect of adolescent depression on earnings. However,

Figure 13: Average Adult Earnings by Adolescent Depression



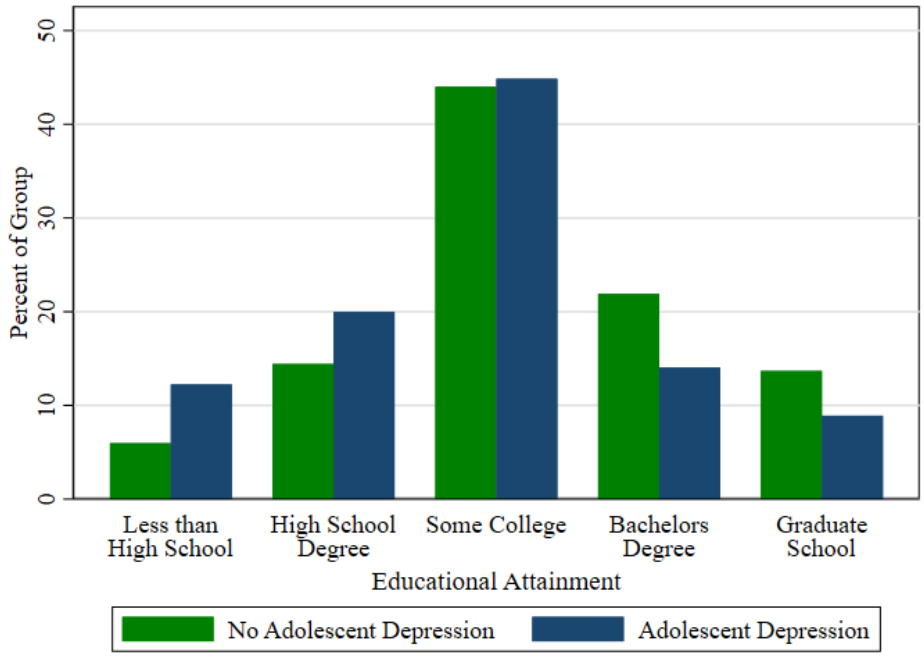
Note: Average yearly earnings in adulthood is plotted by adolescent CES-D score, a self-reported measure of symptoms of depression in the past two weeks. Respondents have an average age of 15.6 in adolescence and 28.5 in adulthood.

Figure 14: Average Adult Hourly Wages by Adolescent Depression



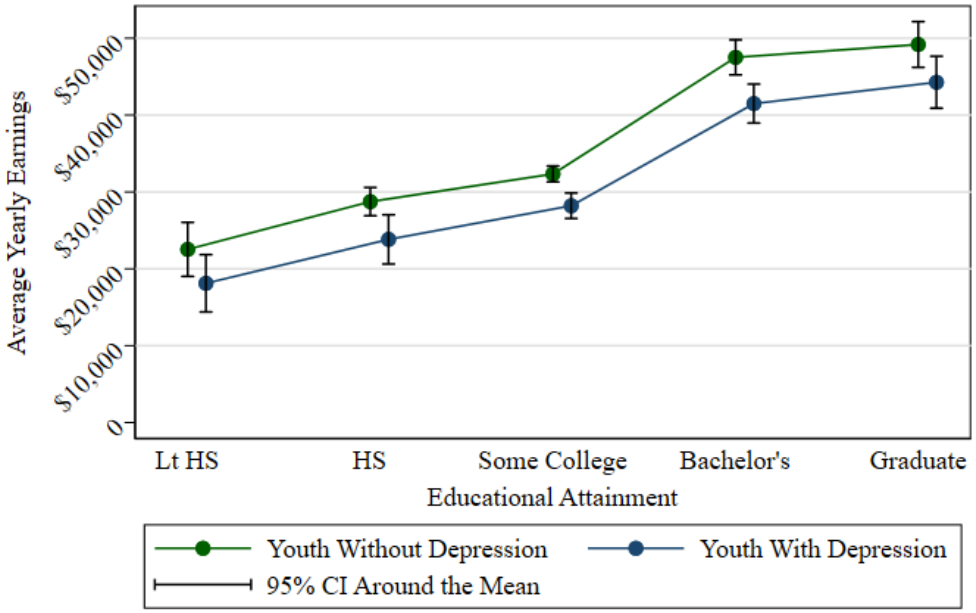
Note: Average hourly wages in adulthood is plotted by adolescent CES-D score, a self-reported measure of symptoms of depression in the past two weeks.

Figure 15: Educational Attainment and Adolescent Depression



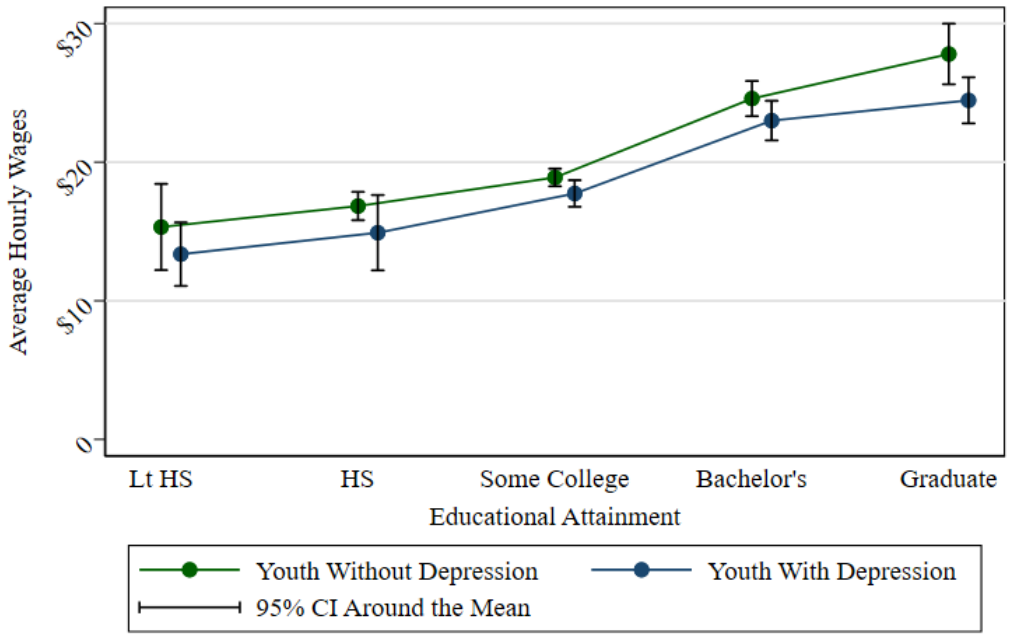
Note: Highest level of educational attainment is plotted by adolescent depression.

Figure 16: Average Adult Earnings by Educational Attainment and Adolescent Depression



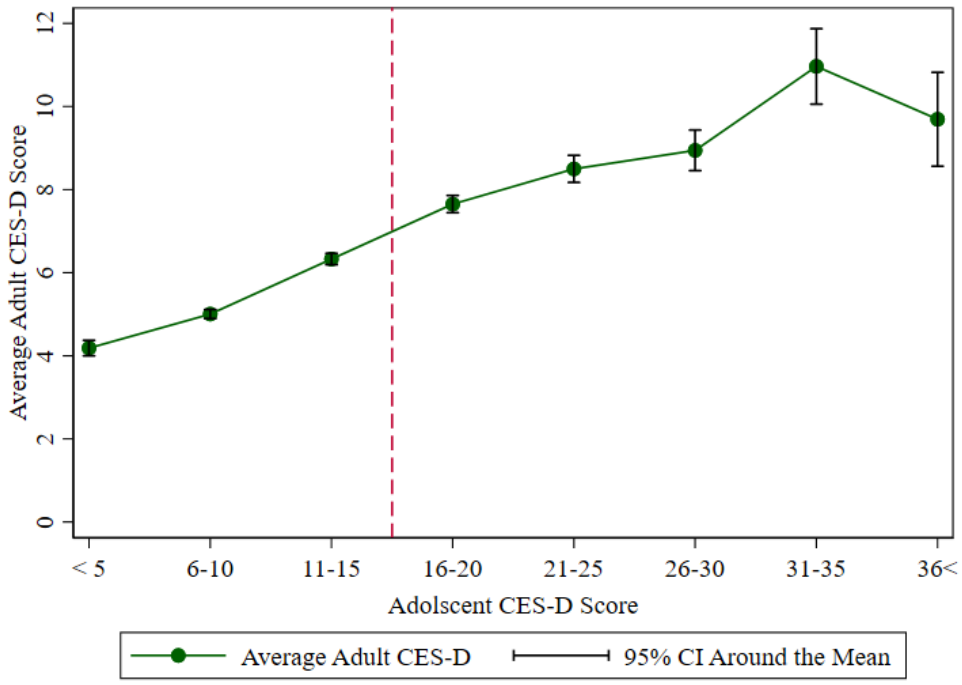
Note: Average yearly earnings are plotted by educational attainment and adolescent depression.

Figure 17: Average Hourly Wages by Educational Attainment and Adolescent Depression



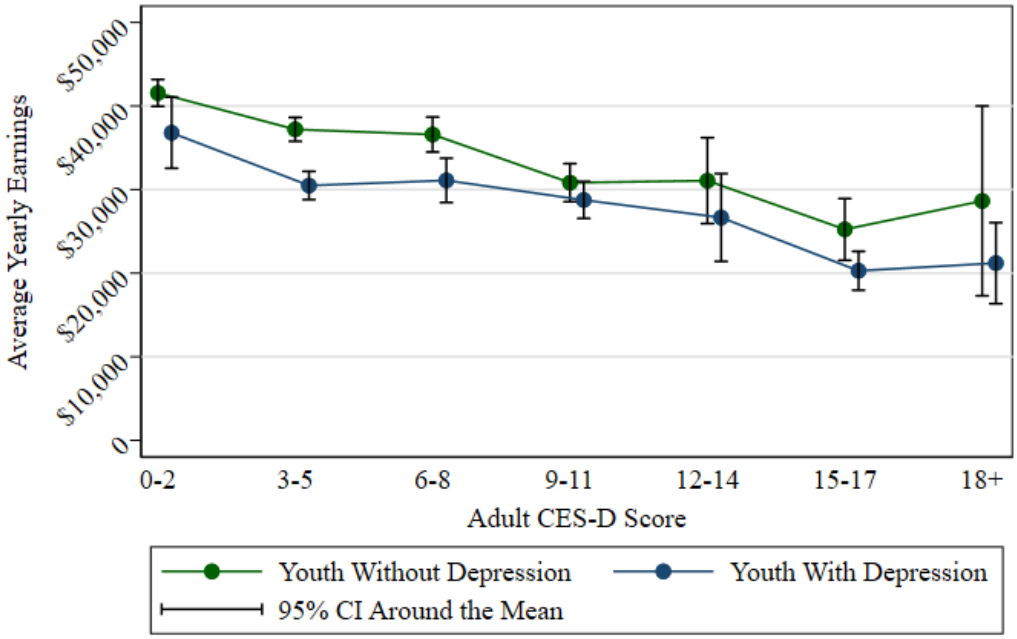
Note: Average hourly wages are plotted by educational attainment and adolescent depression.

Figure 18: Adolescent and Adult CES-D Scores



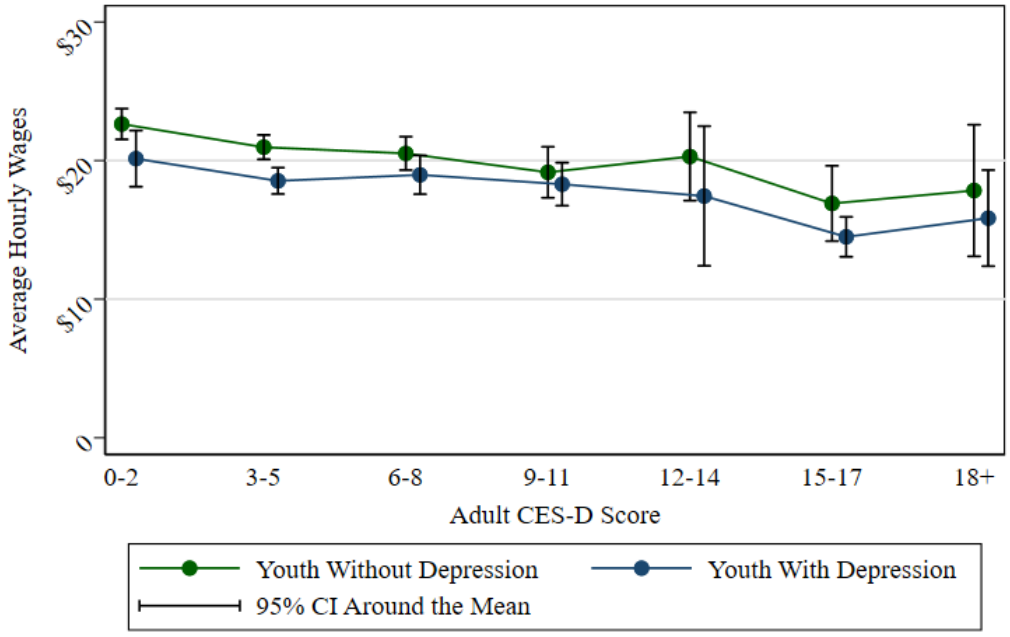
Note: Average adult abbreviated CES-D score is plotted by adolescent CES-D score.

Figure 19: Average Adult Earnings by Adult CES-D Score and Adolescent Depression



Note: Average yearly earnings are plotted by adult CES-D score and adolescent depression.

Figure 20: Average Hourly Wages by Adult CES-D Score and Adolescent Depression



Note: Average hourly wages are plotted by educational attainment and adolescent depression.

accounting for important differences could crowd out parts of the effect if the differences themselves are driven by adolescent depression. For example, youth with symptoms of depression have a more difficult time focusing on school and receive less education in the long run (J. M. Fletcher, 2008). Lower education then leads to lower average earnings in adulthood. Treating educational attainment as a confounding variable would waive away an important part of the relationship between adolescent depression and earnings. Similarly, symptoms of depression in adolescence strongly predict the likelihood of depression in adulthood, which can lead to lower labor supply and earnings. Ignoring the intertemporal relationship of depression would ignore a potentially relevant part of the effect of adolescent depression one earnings.

Figure 15 splits educational attainment apart by adolescent depression. Youth with depression are more likely to drop out of high school or stop pursuing education after high school than their counterparts. Conversely, youth without depression are more likely to have a bachelor's degree or go to graduate school afterwards. If this figure is indicative of a causal relationship between depression and education, then it could have significant earnings implications. Descriptively, it is consistent with evidence from previous literature on the relationship between adolescent depression and educational attainment.

Figure 16 plots average earnings by educational attainment and adolescent depression. Three points jump out. First, the returns to education are positive. Combined with the negative relationship described in Figure 13, this implies a negative effect of adolescent depression on earnings through education. Second, the average returns to education are approximately constant for youth with and without depression. This suggests that the incentive to pursue a higher education is similar across groups and that youth with depression receive a similar return on investment in education. If education mediates an important effect, it is because youth with

depression receive less of it, not because it is less beneficial. Third, for every level of educational attainment, youth with depression earn less than their counterparts without depression. Although, differences are not always statistically significant, this suggests that education does not fully account for the gap in earnings between the two groups. Figure 17 plots average hourly wages by educational attainment and adolescent depression. The aforementioned three points also hold up in Figure 17, suggesting that features of the relationship may not vary much by productivity and labor supply.

Figure 18 takes a close look at the relationship between adolescent and adult symptoms of depression. The dashed redline splits the graph into two groups – youth classified as having depression and youth without depression. As adolescent CES-D score increases, adult CES-D score increases significantly across most values. Youth with the highest-level symptoms of depression have slightly fewer symptoms on average than youth with CES-D score of 31-35. This could be due to an increased likelihood of receiving treatment or it could be noise, as the means are statistically indistinguishable.

Figure 19 plots average earnings by adult CES-D score and adolescent depression. Symptoms of depression in adulthood are negatively related to earnings across the board. Although depression could be a determinant of earnings, lower earnings could also lead to symptoms of depression. Across all levels of symptoms, the change in earnings is similar between groups with and without adolescent depression. This suggests that the average adverse effect of adult depression on earnings may be independent of symptoms in adolescence (i.e., symptoms do not have a compounding effect). Finally, across all levels of adult symptoms, the gap in earnings between groups remains relatively constant, suggesting that differences in adult depression do not account for the entirety of the earnings gap either. Figure 20 plots average

hourly wages by adult CES-D score and adolescent depression. While the above points hold for hourly wages as well, there is a smaller drop in wages as adult CES-D increases. This is in line with the idea that adult depression primarily impacts earnings via its effect on labor supply, not wages.

The tables and figures in this section provide descriptive evidence of three points. First, youth with depression earn significantly less as adults. Second, part of this relationship is driven by the negative impact of adolescent depression on educational attainment. Third, another part of this relationship is driven by a higher likelihood of depression in adulthood. These results are consistent with both the conceptual framework outlining why adolescent depression is related to adult labor market outcomes and the peer-reviewed literature on the subject. In the following subsection, I take a deeper dive into these relationships using the difference method to estimate exact magnitudes of relationships.

Regression Results: The Difference Method

This subsection outlines the results from single equation regressions, including the difference method. First, I build up earnings and wages equations using the difference method to calculate initial estimates of direct and indirect effects. Next, I build up specifications for adult depression and years of education equations one at a time. The conditional independence assumption needed for results to be interpreted as causal likely fails. Therefore, these results provide an in-depth descriptive view of relationships. These results also mimic the methods used in previous literature on this topic and they act as a launching-off point for the causal methods used in the next chapter.

Earnings

Table 3 presents coefficients and standard errors from a log earnings regression estimated with OLS. I convert coefficients to percentage change effects using the following equation for each coefficient: $\%\Delta Y = e^{\Delta x \beta} - 1$. Column 1 regresses log earnings on adolescent depression, finding that those with depression in adolescence earn about 22.5% less in adulthood than their counterparts without depression. This gap is similar in size to the gap in means found in Table 2. Column 2 adds adult depression, which lowers the coefficient on adolescent depression to -0.195 or about -17.7%. Adult depression is associated with a statistically significant average earnings drop of 27.1%. When years of education is added in column 3, the coefficient on adolescent depression drops even further to -0.131 but remains statistically significant at the 0.1% level. This reveals that about half of the initial association between adolescent depression and earnings in column 1 was driven by the omission of adult depression and years of education.

Several other factors are important determinants of earnings. Column 4 adds measures of gender, race, ethnicity, and grade to the right-hand side. Females earn about 30.4% less than males, all else constant – responsible for a larger gap than any other factor. Respondents in higher grades earn less than those in lower grades due to their additional average experience in the labor market at the time of interview. Differences in earnings by race and ethnicity are also economically and statistically significant. Controlling for these demographic characteristics drops the coefficient on adolescent depression further to -0.099, or about -9.4%.

Health and health behaviors are also correlated with adolescent depression and several factors related to long-term labor market success. Column 5 adds measures of alcohol use, binge drinking, problem drinking, cigarette use, marijuana use, and BMI to the right-hand side. Marijuana use and cigarette use are associated with lower long-term earnings, while all alcohol

use measures are associated with higher earnings, consistent with literature on the drinker's bonus (e.g., Bray, 2005). Adding these characteristics to the regression leads the effect of adolescent depression to drop slightly to about -8.2%.

Column 6 adds measures of log family income, mother's highest education, and the age-standardized Peabody Picture Vocabulary Score to control for measures of human capital. Family income is a statically significant positive predictor of earnings, consistent with literature on intergenerational transfer of wealth. While the vocab test score is not statistically significant, it is important to control for a proxy for ability. Adding these characteristics drops the coefficient on adolescent depression further to -0.077.

Finally, column 7 adds marital status, number of kids, school enrollment status, whether the respondent has been to jail, region indicators, and RUCA codes to the right-hand side. Those who are married earn more, while having additional children is associated with lower earnings. Respondents currently enrolled in school earn about 20% less than their counterparts. Those who have been to jail or prison face an earnings penalty of about 18%. Region and rural/urban status also accounts for some variation in adult earnings. Accounting for these adult characteristics drops the coefficient on adolescent depression even further to -0.067, or nearly one-fourth the size of its initial size in column 1.

While a lengthy list of relevant characteristics are included in the earnings regression in column 7, there could be unobserved characteristics related to both adolescent depression and earnings that are biasing the coefficient on adolescent depression. Some of these characteristics may be attributes of the school, community, state, or environment that youth find themselves in. In column 8, I add a school fixed effect, effectively controlling for shared characteristics at the school level. Estimates remain largely unchanged, suggesting that the direct effect of adolescent

depression is not being biased by school-level heterogeneity, conditional on right-hand side variables. Note that while school-level heterogeneity does not appear to play a significant role, I still use a school fixed effect in the next chapter due to its role in my preferred identification strategy (i.e., within school variation in cohort religiosity).

Columns 9 and 10 add community and state fixed effects to the regression, respectively. The coefficient on adolescent depression is relatively unchanged in either case, suggesting that community-level and state-level characteristics are not biasing the direct effect of adolescent depression. The explanatory roles of right-hand side variables do not qualitatively change either.

A large majority of the correlation between adolescent depression and earnings is driven by confounding and mediating variables – the coefficient on adolescent depression drops from -0.255 to -0.068. Coefficients on adult depression and years of education confirm expected signs and magnitudes, suggesting that they could be mediating variables. Adult depression is associated with about a 13.1% drop in earnings, while a year of education has about a 9.6% average earnings premium. This sets up the regression specification needed to estimate direct and indirect effects.

Table 4 uses the difference method to calculate indirect and total effects. Column 1 presents the fully specified model without any fixed effect (i.e., column 7 of Table 3) and acts as the starting specification. Column 1 estimates that the direct effect of adolescent depression on adult earnings is about -6.5%, which is equivalent to about a $2,200 annual drop in earnings in this sample.

Column 2 removes adult depression from the regression and the estimated coefficient on adolescent depression is about -0.087. This coefficient is slightly larger than the average coefficient on adolescent depression estimated by Johar & Truong (2014), which excluded adult

depression from their wage equation. Taking the difference in coefficients between columns 1 and 2 implies that adult depression mediates a 2% average drop in earnings.

Column 3 removes years of education from the regression and produces an estimated coefficient of -0.085. This is approximately the size of the coefficient on adolescent depression found by Philipson et al. (2020), which omitted educational attainment from the earning equation. Taking the difference between estimates in columns 1 and 3, years of education mediates a 1.8% average drop in earnings. Note that while Johar & Truong (2014) and Philipson et al. (2020) referred to their estimates as direct effects, these coefficients are combinations of the direct effect and the indirect effect through the omitted mediator.

Both mediated effects are economically significant but relatively modest compared to the direct effect of -6.5%. Column 4 estimates the total effect by removing both adult depression and years of education from the regression. The estimated coefficient is -0.109, implying a total effect of -10.3% or about $3,688 annually for this sample. This effect is driven primarily by the direct effect. Results from the difference method appear internally consistent, as this estimate of the total effect is consistent with taking the sum of the individual direct and indirect effects (i.e., $6.5 + 2.0 + 1.8 = 10.3$).

Columns 5-8 present difference method estimates for specifications with a school fixed effect. Results are fairly similar to columns 1-4. Column 5 estimates a coefficient of -0.068, implying a direct effect of -6.6%. When adult depression is removed in column 6, the coefficient rises to -0.087. In column 7, years of education is removed and the coefficient is also estimated at -0.087. This implies that adult depression and years of education each mediate a 2.1% drop in earnings, similar in magnitude to estimates from columns 2 and 3. Finally, column 8 estimates a coefficient of -0.111, or a total effect of -10.5%. Results are relatively internally consistent –

Table 3: Building an Earnings Regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Adolescent Depression	-0.255***	-0.195***	-0.131***	-0.099***	-0.086***	-0.077***	-0.067**	-0.068**	-0.068**	-0.067**
	(0.032)	(0.030)	(0.026)	(0.021)	(0.022)	(0.022)	(0.021)	(0.021)	(0.020)	(0.020)
Adult Depression		-0.316***	-0.244***	-0.196***	-0.192***	-0.185***	-0.142***	-0.140***	-0.137***	-0.139***
		(0.024)	(0.022)	(0.022)	(0.021)	(0.021)	(0.022)	(0.022)	(0.022)	(0.019)
Years of Education			0.118***	0.125***	0.122***	0.111***	0.094***	0.092***	0.091***	0.091***
			(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)
Female				-0.362***	-0.357***	-0.348***	-0.346***	-0.345***	-0.348***	-0.344***
				(0.028)	(0.028)	(0.027)	(0.026)	(0.027)	(0.027)	(0.032)
Smokes Cigarettes					-0.104***	-0.105***	-0.070**	-0.063**	-0.070**	-0.065**
(Wave 1)					(0.023)	(0.023)	(0.022)	(0.022)	(0.024)	(0.020)
Used Marijuana Past					-0.077**	-0.085**	-0.078**	-0.080**	-0.081**	-0.071**
Month (Wave 1)					(0.029)	(0.029)	(0.028)	(0.028)	(0.028)	(0.026)
Alcohol Use (Wave 1)					0.050*	0.043*	0.045*	0.037	0.038	0.041*
					(0.021)	(0.021)	(0.020)	(0.020)	(0.022)	(0.018)
Log Family Income						0.075***	0.057***	0.040**	0.041**	0.049***
						(0.013)	(0.013)	(0.013)	(0.014)	(0.012)
Vocab Test Score						0.014	0.015	0.018	0.016	0.018*
						(0.011)	(0.011)	(0.011)	(0.010)	(0.008)
Married							0.123***	0.125***	0.128***	0.127***
							(0.016)	(0.016)	(0.016)	(0.021)
Been to Jail or Prison							-0.199***	-0.191***	-0.191***	-0.199***
							(0.029)	(0.029)	(0.030)	(0.027)
Fixed Effect	None	None	None	None	None	None	None	School	Comm.	State
Observations	13,298	13,298	13,298	13,298	13,298	13,298	13,174	13,174	12,911	13,174

Note: Coefficients are reported with standard errors in parentheses. Standard errors are clustered at the school level in columns 1-7 and at the level of the fixed effect in columns 8-10.

**** for p < 0.001, ** for p < 0.01, * for p < 0.05*

Table 4: Earnings Difference Methods Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Adolescent Depression	-0.067**	-0.087***	-0.085***	-0.109***	-0.068**	-0.087***	-0.087***	-0.111***
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)
Adult Depression	Yes	No	Yes	No	Yes	No	Yes	No
Years of Education	Yes	Yes	No	No	Yes	Yes	No	No
Fixed Effect	None	None	None	None	School	School	School	School
Observations	13,174	13,174	13,174	13,174	13,174	13,174	13,174	13,174

Note: Coefficients are reported with standard errors in parentheses. Standard errors are clustered at the school level. Columns 1-4 present difference method results without a fixed effect, while columns 5-8 present results with a school fixed effect.

**** for p < 0.001, ** for p < 0.01, * for p < 0.05*

when direct and indirect effects are summed, they amount to -10.8%. Results from using a school fixed effect are qualitatively similar to those from without a school fixed effect – the total effect of adolescent depression on earnings is about -10-11%, the direct effect accounts for a majority of this effect, and both adult depression and years of education mediate some of the effect.

Wages

Table 5 presents coefficients and standard errors from a log wages regression estimated with OLS. Wages are only observed for a subset of the population, as described in Chapter V: Data. I convert coefficients to percentage change effects using the following equation for each coefficient: $\%\Delta Y = e^{\Delta x \beta} - 1$. Column 1 regresses log wages on adolescent depression, finding a coefficient of -0.138 – those with depression in adolescence earn about 12.9% less in adulthood than their counterparts without depression. This gap is almost identical in size to the gap in means found in Table 2. Column 2 adds adult depression, which lowers the coefficient on adolescent depression to -0.111 or about -10.5%. Adult depression is associated with a statistically significant average drop in hourly wages of about 13.9%. When years of education is added in column 3, the coefficient on adolescent depression drops to -0.060 but remains statistically significant at the 0.1% level. Similar to results for earnings, about half of the initial association between adolescent depression and wages in column 1 of Table 5 is driven by the omission of adult depression and years of education. Education appears to play an especially important role in driving this coefficient downward; an additional year of education leads to about a 9.6% wage premium.

Column 4 adds measures of gender, race, ethnicity, and grade to the right-hand side. Females have hourly wages that are about 17.6% less than males, all else constant. The gender gap in hourly wages remains the largest of any factor throughout all specifications. Respondents

in higher grades earn lower wages than those in lower grades due to their additional average experience in the labor market at the time of interview. Differences in earnings by race and ethnicity are also economically and statistically significant. Controlling for demographic characteristics drops the coefficient on adolescent depression further to -0.055, or about a -5.4% average drop in wages.

Column 5 adds measures of alcohol use, binge drinking, problem drinking, cigarette use, marijuana use, and BMI to the right-hand side. Cigarette use is associated with lower wages at the 5% level, while neither marijuana use nor alcohol use are significant predictors of hourly wages. Adding these characteristics to the regression leads to a small drop in the effect of adolescent depression on wages to about -5.3%. Column 6 adds measures of log family income, mother's highest education, and the age-standardized Peabody Picture Vocabulary Score to control for measures of human capital. Family income is a statically significant positive predictor of wages – a 1% increase in family income leads to a 0.07% increase in average hourly wages. Vocabulary test score is a statistically significant predictor of higher wages, confirming that a measure of ability or productivity is necessary when modeling wages. Adding these characteristics drops the coefficient on adolescent depression further to -0.044.

Finally, column 7 adds marital status, number of kids, school enrollment status, whether the respondent has been to jail, region indicators, and RUCA codes to the right-hand side. Those who are married have higher wages, while having additional children is associated with lower wages. Those who have been to jail or prison face a wages penalty of about 9%. Region and rural/urban status also accounts for some important variation in adult wages. Accounting for these adult characteristics actually raises the coefficient on adolescent depression slightly to -0.046, or about one-third the size of its initial size in column 1.

Columns 1-7 uses the sample of respondents who have observed wages in wave 4, which is likely an unrepresentative sample of the full sample. To address potential sample selection bias, column 8 presents results from a Heckman sample selection model. In the first-stage, interview date is strong predictor of whether we observe wages. Being interviewed 10 days later is associated with a 0.9% decrease in the probability of observing wages. Put another way, someone interviewed on January 1 is nearly 28% more likely to have observed wages than someone interviewed on December 31 of the same year. The covariance between the error terms is estimated to be 0.191 and is significant at the 1% level, indicating that there is positive sample selection bias with respect to wages. We are more likely to observe wages for high wage earners than low wage earners, which could bias estimates. Despite this potential bias, the coefficient on adolescent depression only rises slightly to -0.049, or about -4.8%. The coefficient on adult depression rises from -0.067 to -0.080, while the return to a year of education stays at its level in column 7. Sample selection does not appear to be driving most results.

Important unobserved characteristics at the school or environmental level could impact the productivity or skill development of youth, ultimately influencing hourly wages. In column 9, I estimate column 7 but with a school fixed effect, which controls for shared characteristics at the school level. The coefficient on adolescent depression remains qualitatively unchanged, rising slightly to -0.046. However, estimated coefficients on years of education and log family income drop, indicating that a school fixed effect accounts for some important variation related to socioeconomic status and adult wages.

In column 10, I estimate the school fixed effect model with a Heckman sample selection model. Overall, results are similar to column 8. Interview date is a strong predictor of whether we observe hourly wages and the covariance between error terms is positive, suggesting positive

sample selection bias in results. The coefficient on adolescent drops from -0.048 in column 9 to -0.050 in column 10, while the coefficient on adult depression drops from -0.068 to -0.078 across the same columns. The explanatory roles of most right-hand side variables do not qualitatively change.

A majority of the correlation between adolescent depression and hourly wages is driven by confounding variables, mediating variables, and sample selection bias – the coefficient on adolescent depression drops from -0.138 to -0.049. Coefficients on adult depression and years of education confirm expected signs and magnitudes, suggesting that they could be mediating relationships between adolescent depression and wages. Adult depression is associated with about a 7.7% drop in wages, while a year of education has about a 7.6% average wages premium.

Table 6 uses the difference method to calculate indirect and total effects of adolescent depression on wages. Columns 1-4 present results with school fixed effects but without any sample selection model. Column 1 presents the fully specified model (i.e., column 7 of Table 5) and acts as the starting specification. Column 1 estimates that the direct effect of adolescent depression on wages is about -4.7%. Column 2 removes adult depression from the regression and the estimated coefficient on adolescent depression is about -0.057. This coefficient is a little smaller than the average coefficient on adolescent depression estimated by Johar & Truong (2014), which excluded adult depression from their wage equation. Taking the difference in coefficients between columns 1 and 2 implies that adult depression mediates a 0.9% average drop in earnings.

Column 3 removes years of education from the regression and puts adult depression back in, producing an estimated coefficient of -0.061. This is slightly smaller than the size of the coefficient on adolescent depression found by Philipson et al. (2020), which omitted educational

attainment from the earning equation. Taking the difference between estimates in columns 1 and 3, years of education mediates a 1.3% average drop in earnings.

Column 4 estimates the total effect by removing both adult depression and years of education from the regression. The estimated coefficient is -0.072, implying a total effect of -6.9%, or about a $1.40 loss per hour. This effect is driven primarily by the direct effect of -4.7%, as the mediated effects add up to only -2.2% combined. Results from the difference method are internally consistent, as this estimate of the total effect is consistent with taking the sum of the individual direct and indirect effects (i.e., $4.7 + 0.9 + 1.3 = 6.9$).

Columns 5-8 present difference method estimates using a Heckman sample selection model. for specifications with a school fixed effect. Results are fairly similar to columns 1-4. Interview date is a strong predictor of observing wages and there is positive sample selection bias across all columns. Column 5 estimates a coefficient of -0.050, implying a direct effect of -4.9%. When adult depression is removed from both equations in column 6, the coefficient rises to -0.061. In column 7, years of education is removed instead, and the coefficient is also estimated at -0.063. This implies that adult depression mediates about a -1.1% effect on wages, while years of education mediates about a -1.3% effect. Finally, column 8 estimates a coefficient of -0.075, or a total effect of -7.2%. Results are relatively internally consistent – when direct and indirect effects are summed, they amount to -7.3%, similar to the column 8 estimate of -7.2%. Accounting for sample selection leads reveals that each effect is slightly larger than indicated in columns 1-4. The total effect rises from -6.9% to about -7.2%, driven by small increases in each direct and indirect effect. Similar to columns 1-4 and to earnings results, the direct effect makes up a majority of the total effect.

Table 5: Building a Wages Regression

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ln(Wages)										
Adolescent Depression	-0.138***	-0.111***	-0.062**	-0.055**	-0.054**	-0.044**	-0.046**	-0.049**	-0.048**	-0.050**
	(0.025)	(0.024)	(0.021)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Adult Depression		-0.150***	-0.112***	-0.087***	-0.087***	-0.077***	-0.067**	-0.080***	-0.068**	-0.078***
		(0.023)	(0.022)	(0.021)	(0.022)	(0.022)	(0.023)	(0.022)	(0.022)	(0.022)
Years of Education			0.092***	0.095***	0.094***	0.081***	0.073***	0.073***	0.069***	0.070***
			(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Female				-0.194***	-0.189***	-0.183***	-0.188***	-0.194***	-0.189***	-0.193***
				(0.015)	(0.014)	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)
Smokes Cigarettes					-0.053*	-0.051*	-0.030	-0.034	-0.024	-0.027
(Wave 1)					(0.024)	(0.024)	(0.025)	(0.025)	(0.025)	(0.025)
Alcohol Use (Wave 1)					0.026	0.020	0.023	0.027	0.020	0.023
					(0.021)	(0.020)	(0.020)	(0.019)	(0.020)	(0.020)
Log Family Income						0.073***	0.060***	0.059***	0.040***	0.040***
						(0.012)	(0.011)	(0.012)	(0.011)	(0.011)
Vocab Test Score						0.016*	0.021**	0.020**	0.023**	0.023**
						(0.007)	(0.007)	(0.007)	(0.008)	(0.008)
Married							0.076***	0.082***	0.076***	0.081***
							(0.015)	(0.014)	(0.015)	(0.014)
Been to Jail or Prison							-0.098***	-0.115***	-0.099***	-0.112***
							(0.020)	(0.022)	(0.020)	(0.021)
Observed Wages										
Interview Date (/10)								-0.009***		-0.009***
								(0.001)		(0.001)
$\sigma_{Y,W}$ (Or λ)								0.191**		0.157***
								(0.070)		(0.049)
Fixed Effect	None	None	None	None	None	None	None	None	School	School
Observations	7,762	7,762	7,762	7,762	7,762	7,762	7,760	14,349	7,760	14,349

Note: Coefficients are reported with standard errors in parentheses. Standard errors are clustered at the school level. When estimated, the observed wages equation includes all right-hand side variables in the wages equation, as well as the interview date as an instrument. Sample selection bias is being addressed if $\sigma_{Y,W}$ is not equal to zero.

**** for p < 0.001, ** for p < 0.01, * for p < 0.05*

Table 6: Hourly Wages Difference Method Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(Wages)								
Adolescent Depression	-0.048**	-0.057**	-0.061***	-0.072***	-0.050**	-0.061***	-0.063***	-0.075***
	(0.017)	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)	(0.017)	(0.017)
Observed Wages								
Interview Date (/10)					-0.009***	-0.009***	-0.009***	-0.009***
					(0.001)	(0.001)	(0.001)	(0.001)
$\sigma_{Y,W}$ (Or λ)					0.157**	0.162**	0.120**	0.127**
					(0.049)	(0.049)	(0.043)	(0.043)
Adult Depression	Yes	No	Yes	No	Yes	No	Yes	No
Years of Education	Yes	Yes	No	No	Yes	Yes	No	No
Fixed Effect	School	School	School	School	School	School	School	School
Observations	7,760	7,760	7,760	7,760	14,349	14,349	14,349	14,349

Note: Coefficients are reported with standard errors in parentheses. Standard errors are clustered at the school level. When estimated, the observed wages equation includes all right-hand side variables in the wages equation, as well as the interview date as an instrument. Sample selection bias is being addressed if $\sigma_{Y,W}$ is not equal to zero.

**** for p < 0.001, ** for p < 0.01, * for p < 0.05*

Years of Education

This subsection presents results from estimating a years of education equation, with special attention placed on the association between adolescent depression and education. Table 7 presents results from a regression with years of education as the dependent variable, estimated with OLS. In column 1, years of education is regressed only on adolescent depression. The coefficient suggests youth with depression receive about -0.72 fewer years of education on average than youth without depression, similar to the difference in means found in Table 2. Column 2 adds measures for gender, race, ethnicity, grade, and vocabulary test score. Females earn about two-thirds of a year of education more than males, on average. A standard deviation increase in the vocabulary test score is associated with nearly a 0.80 increase in years of education, representing the tie between ability and educational attainment. The coefficient on adolescent depression drops to -0.584.

Health behaviors are also correlated with outcomes in school and mental health. Column 3 controls for cigarette use, marijuana use, measures of alcohol use, and adolescent BMI. Youth who use cigarettes receive nearly a year less of education than their non-smoking counterparts, a gap larger than nearly any other explanatory factor. Marijuana use is associated with a statistically significant drop of more than a quarter of a year of education. BMI is also related to lower educational attainment. Accounting for these factors drops the coefficient on adolescent depression to -0.419.

Column 4 adds measures of family characteristics and controls for whether the respondent has been to jail. Higher family income is associated with a small increase in educational attainment – a 10% increase in family income leads to an average increase in 0.04 years of education. Mother's highest level of education is much more significant. Youth with a

Table 7: Years of Education Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Adolescent Depression	-0.716***	-0.584***	-0.419***	-0.312***	-0.320***	-0.320***	-0.317***
	(0.056)	(0.054)	(0.050)	(0.040)	(0.041)	(0.041)	(0.041)
Female		0.657***	0.622***	0.495***	0.497***	0.497***	0.495***
		(0.036)	(0.034)	(0.032)	(0.031)	(0.031)	(0.031)
Vocab Test Score		0.792***	0.769***	0.536***	0.477***	0.477***	0.478***
		(0.046)	(0.040)	(0.032)	(0.024)	(0.024)	(0.024)
Smokes Cigarettes			-0.945***	-0.777***	-0.752***	-0.752***	-0.747***
(Wave 1)			(0.072)	(0.057)	(0.057)	(0.057)	(0.057)
Used Marijuana Past			-0.286***	-0.215***	-0.215***	-0.215***	-0.217***
Month			(0.059)	(0.053)	(0.050)	(0.050)	(0.051)
Alcohol Use (Wave 1)			0.018	-0.008	-0.024	-0.024	-0.023
			(0.046)	(0.044)	(0.041)	(0.041)	(0.041)
BMI in Adolescence			-0.042***	-0.030***	-0.025***	-0.025***	-0.025***
			(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Log Family Income				0.431***	0.351***	0.351***	0.352***
				(0.035)	(0.029)	(0.029)	(0.029)
Mother has High				0.319***	0.340***	0.340***	0.348***
School Degree				(0.073)	(0.073)	(0.073)	(0.073)
Mother has Bachelor's				1.187***	1.067***	1.067***	1.071***
Degree				(0.104)	(0.097)	(0.097)	(0.097)
Been to Jail or Prison				-0.879***	-0.826***	-0.826***	-0.821***
				(0.044)	(0.042)	(0.042)	(0.042)
Observations	14,561	14,561	14,561	14,561	14,561	14,561	14,420
Fixed Effect	None	None	None	None	School	Comm.	State

Note: Coefficients are reported with standard errors in parentheses. Heteroskedasticity robust standard errors are reported in columns 1-4. Standard errors are clustered at the level of the fixed effect in columns 5-7.

**** for p < 0.001, ** for p < 0.01, * for p < 0.05*

mother who graduated high school receive about 0.32 more years of education compared to youth with a mother who dropped out of high school, on average. For youth with a mother with a bachelor's degree, the difference is 1.18 years. Youth who have been to jail can also expect receive nearly 0.9 fewer years of education. Controlling for family characteristics and criminal justice involvement drops the coefficient on adolescent depression to -0.312, nearly two-fifths the size of the association in column 1. Demographic, health, and family characteristics are all

important to account for when estimating adolescent depression's impact on educational attainment.

Column 5 adds a school fixed effect to control for time-invariant, school-level heterogeneities. The coefficient on adolescent depression rises slightly to -0.320, but the result is qualitatively similar. Some coefficients drop in size, such as those on vocabulary test score and family income, suggesting that their explanatory power was due in part to their correlation with school-level differences. Columns 6 and 7 add community and state fixed effects, respectively. Adding these fixed effects does not have any significant implications for how adolescent depression impacts years of education.

Overall, there is consistent evidence that adolescent leads to lower educational attainment, even after controlling for the average impacts of a variety of confounding variables.

Adult Depression

This subsection focuses on the predictors of adult depression, with a focus on how well adolescent depression predicts adult depression. Table 8 presents results from an adult depression equation; columns 1-7 are estimated with a probit and present average marginal effects, while columns 8-10 are estimated with OLS and present coefficients.

Column 1 regresses adult depression on adolescent depression – youth with depression are 19.5 percentage points more likely to have depression in adulthood. In column 2, measures of gender, race, ethnicity, and grade are added, dropping the average marginal effect of adolescent depression to 18.7 percentage points. Females are about 4.4 percentage points more likely to have depression in adulthood than males.

Since health and health behaviors are strongly correlated to adolescent depression, they may be correlated with long-term depression and influence our marginal effects. Column 3 adds

measures of cigarette use, marijuana use, alcohol use, and adolescent BMI. Youth who use cigarettes, use marijuana, or have a higher BMI are all more likely to have depression in adulthood, although the marginal effects are relatively small in magnitude. Accounting for health behaviors, youth with depression are about 16.7 percentage points more likely to have depression in adulthood. Column 4 adds measures of adolescent family characteristics and youth ability. Youth from families with higher incomes or higher educated mothers are less likely to have depression in adulthood, while vocabulary test score is negatively related to adult depression. The average marginal effect of adolescent depression drops slightly to -0.151.

Column 5 adds several adult characteristics, including region, RUCA codes, marital status, number of kids, BMI, and criminal justice involvement. Respondents who are married are nearly 5 percentage points less likely to be depressed, while those who have been to jail are 4.3 more less likely to have depression. Youth with depression are now only 13.7 percentage points more likely to experience depression in adulthood, all else constant. This effect is only 30% smaller than the original effect size in column 1, suggesting a majority of the association cannot be explained away by observables.

Columns 6 and 7 add school and state fixed effects, respectively. None of these fixed effects qualitatively change results – adolescents with depression are 13.7 percentage points more likely to have depression in adulthood. This is consistent with the idea that the persistence of mental illness over the life course is an individual-level phenomena. Between-group variation in time-invariant characteristics does not appear to account for the persistence of depression into adulthood.

One concern with using a discrete measure for an outcome is that the cut-off point used to create the discrete measure is driving results. In columns 8-10, I use adult CES-D score as the

Table 8: Adult Depression Results

	Adult Depression							Adult CES-D Score		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Adolescent Depression	0.195***	0.187***	0.162***	0.151***	0.137***	0.137***	0.137***	2.106***	2.103***	2.106***
	(0.009)	(0.009)	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.112)	(0.113)	(0.112)
Female		0.044***	0.043***	0.042***	0.043***	0.043***	0.043***	0.593***	0.605***	0.593***
		(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.070)	(0.072)	(0.071)
Smokes Cigarettes			0.034**	0.029**	0.018	0.015	0.015	0.240*	0.187	0.206
(Wave 1)			(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.108)	(0.110)	(0.107)
Alcohol Use (Wave 1)			-0.014	-0.008	-0.010	-0.010	-0.010	-0.195*	-0.190*	-0.192*
			(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.094)	(0.094)	(0.094)
BMI in Adolescence			0.003***	0.002**	0.001	0.001	0.001	0.015	0.015	0.015
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.012)	(0.012)	(0.012)
Vocab Test Score				-0.013**	-0.012**	-0.014**	-0.014***	-0.165**	-0.193***	-0.195***
				(0.004)	(0.004)	(0.004)	(0.004)	(0.053)	(0.051)	(0.050)
Log Family Income				-0.022***	-0.015**	-0.016**	-0.015**	-0.206**	-0.232***	-0.198**
				(0.005)	(0.005)	(0.005)	(0.005)	(0.064)	(0.061)	(0.061)
Mother has Bachelor's				-0.030*	-0.017	-0.017	-0.016	-0.333	-0.346	-0.333
Degree				(0.013)	(0.013)	(0.013)	(0.013)	(0.177)	(0.176)	(0.173)
Married					-0.049***	-0.047***	-0.048***	-0.745***	-0.723***	-0.740***
					(0.006)	(0.006)	(0.006)	(0.083)	(0.083)	(0.081)
Been to Jail or Prison					0.043***	0.044***	0.043***	0.804***	0.813***	0.812***
					(0.010)	(0.010)	(0.010)	(0.122)	(0.124)	(0.124)
Observations	14,561	14,561	14,561	14,561	14,349	14,349	14,319	14,349	14,349	14,349
Fixed Effect	None	None	None	None	None	School	State	None	School	State

Note: Columns 1-7 present coefficients from probit regression. Columns 8-10 present coefficients from an OLS regression with Adult CES-D score as the dependent variable. When there is no fixed effect, standard errors are heteroskedasticity robust. When there is a fixed effect, standard errors are clustered at the level of the fixed effect

**** for p < 0.001, ** for p < 0.01, * for p < 0.05*

dependent variable. In contrast to columns 1-7 where marginal effects are with respect to crossing a threshold of symptoms, estimated effects in columns 8-10 are the average marginal effects on additional symptoms (or severity of symptoms) of depression in adulthood.

In column 8, youth with depression have CES-D scores about 2.1 higher than youth without depression, all else constant. When a school and state fixed effects are added in columns 9 and 10, respectively, this marginal effect does not significantly change. The standard deviation in adult CES-D score is 4.67 (see Table 2). Results in columns 8-10 imply that youth depression see nearly a one-half of a standard deviation increase in adult CES-D score. This confirms the important role of adolescent depression in determining adult mental health status.

CHAPTER VII: ESTIMATION RESULTS

Instrumental Variable Results

This section presents a summary of estimation results using instrumental variables and a system of equations. First, I give an overview of results for log earnings. Next, I give an overview of results for log wages. In contrast to difference method results, the results in this section aim to identify causal effects.

Earnings

Table 9 presents system of equations results for earnings. For convenience, average marginal effects are presented in lieu of coefficients for the adult depression and adolescent depression equations. Column 1 estimates each equation independently, which is identical to column 4 in Table 3, and provides initial evidence that both adult depression and years of education are mediators. Youth with depression are about 13.7 percentage points more likely to be depressed in adulthood, all else constant. Those with depression in adulthood make about 13.1% less than those without depression, which is expected to come through lower labor supply. This estimate is similar in magnitude to Ettner (1997) but a bit larger than some other estimates in the literature (e.g., Baldwin & Marcus, 2007; Cseh, 2008). Youth with depression also receive about one-third fewer years of education than those without depression, all else constant. The average return to a year of education is about 9.6%, which is on the higher end of estimates in the literature.[10] Identical to column 4 of Table 3, column 1 estimates an average direct effect of about -6.6%, suggesting that there is a an economically significant direct effect on earnings.

[10] The literature generally finds an average wage premium of 6-8% (Harmon et al., 2003). The high return found in this paper could be driven by three factors. First, I use yearly earnings instead of hourly wages, so the return to a year of education includes effects on both wages and labor supply. Second, zero-earners are excluded

Column 2 is identical to column 1 but estimates the σ_{Y,D_1} and σ_{D_2,D_1} covariance terms. This opens up the covariance terms needed to identify β_1 and γ_1 with instruments in column 3. While the estimates of β_1 and γ_1 are identified in column 2, they are not well identified – both parameters drop in magnitude and are noisily estimated.

Column 3 adds cohort religiosity to the adolescent depression equation, which identifies β_1 and γ_1. Cohort religiosity is a strong predictor of adolescent depression – a one standard deviation increase in cohort religiosity leads to about a 1.7 percentage point decrease in the probability of having depression in adolescence. Instrumenting for adolescent depression drops the direct effect to -0.036, or about -3.5%. This effect is statistically insignificant and about half the size of the effect in column 1. This reveals that about half of the direct effect in column 1 was driven by the correlation between the error terms in the earnings and adolescent depression equations – in other words, endogeneity. The Z-Test Statistic indicates that we fail to reject the null hypothesis that σ_{Y,D_1} is unchanged between columns 2 and 3, failing to find evidence that the instrument is invalid. Instrumenting for adolescent depression also slightly drops its average marginal effect on adult depression– youth with depression are about 12.8 percentage points more likely to suffer from depression in adulthood. In contrast to column 2, this estimate is significant at the 0.1% level. We also fail to reject the null hypothesis that σ_{D_2,D_1} is unchanged between columns 2 and 3.

Column 4 re-estimates column 3 but unrestricts σ_{Y,D_2} and $\sigma_{D,E}$. The estimate of a return to a year of education is unchanged, while the effect of adult depression on adult earnings grows in

---

from the earnings regression, plausibly excluding those who have a lower return to education, on average. Third, unobserved ability could be biasing the return to a year of education upward.

magnitude slightly to about -14.4%. In contrast to column 2, estimates remain precisely estimated.

Column 5 adds instrumental variables to the adult depression and years of education equations. Using instruments for identification grows the effect of adult depression on earnings to about -16.8% in column 5, larger than any previous specification. Whether a friend or family member attempted suicide in the past year is a relatively strong instrument. Having a death in the family is a weaker instrument only statistically significant at the 5% level. The Z-Test Statistic suggests that we fail to reject the null hypothesis that σ_{Y,D_2} is unchanged between columns 4 and 5, failing to find evidence that the included instruments are invalid. Taken in context, results suggest that adult depression may play a larger mediating role than implied by column 1.

Using instruments for identification also has a small effect on the mediating role of years of education. Youth with depression receive about one third fewer years of education – this estimate stays relatively constant across all specifications because instruments are not used to identify the effect of adolescent depression on years of education. In column 5, the return to a year of education is about 11.5%, a noticeable increase compared to previous estimates. The tract-level proportion of adults with a bachelor's degree is a strong and meaningful instrument – a 0.10 increase in the proportion increases years of education by about 0.16. However, tract-level school enrollment is only significant at the 5% level. Finally, the direct effect of adolescent depression in column 5 drops slightly to about -2.3% and remains statistically insignificant. This maintains the result that column 1 significantly overstates the importance of the direct effect.

Two of the instruments used in column 5 are only statistically significant at the 5% level, which is relatively weak and could be biasing estimates. In column 6, I remove having a death in the family and the proportion of those enrolled in school as instruments. The coefficient on adult

depression drops to -0.164, closer to its initial level in column 4. The coefficient on years of education drops back to 0.091, unchanged compared to column 4. This suggests that the weaker instruments were biasing the coefficients on adult depression and years of education away from zero. When only strong instruments are used, the marginal effects of adult depression and years of education are relatively unchanged compared to specifications without instruments. The direct effect in column 6 is about -3.1% and imprecisely estimated, giving a result that is qualitatively similar to column 5.

Figure 21 plots the direct, indirect, and total effects of adolescent depression on earnings as percent changes. All effects are calculated using the estimates from columns 1, 3, and 6 of Table 9 and are presented in Table 10.

When each equation in the system is estimated separately, the average direct effect is -6.56% and is clearly the pathway with the most significant magnitude. Adult depression mediates a -1.90% effect, while years of education mediates about a -2.88% effect. Mediated effects make up a notable portion of the total effect of adolescent depression on earnings – treating adult depression and education as confounders would underestimate the total effect about 42%.

When cohort religiosity is used as an instrumental variable in the adolescent depression equation, the direct effect falls to -3.21% and is imprecisely estimated. The effect mediated by adult depression drops slightly to -1.88% due to a drop in the average marginal effect of adolescent depression on adult depression. The drop in the direct effect drives a large drop in the total effect as well from -11.33% to -7.06%.

When instruments are also added to the adult depression and years of education equations, the direct effect moves to -3.23% and remains statistically insignificant. An increase

Table 9: System of Equations Results, Yearly Earnings

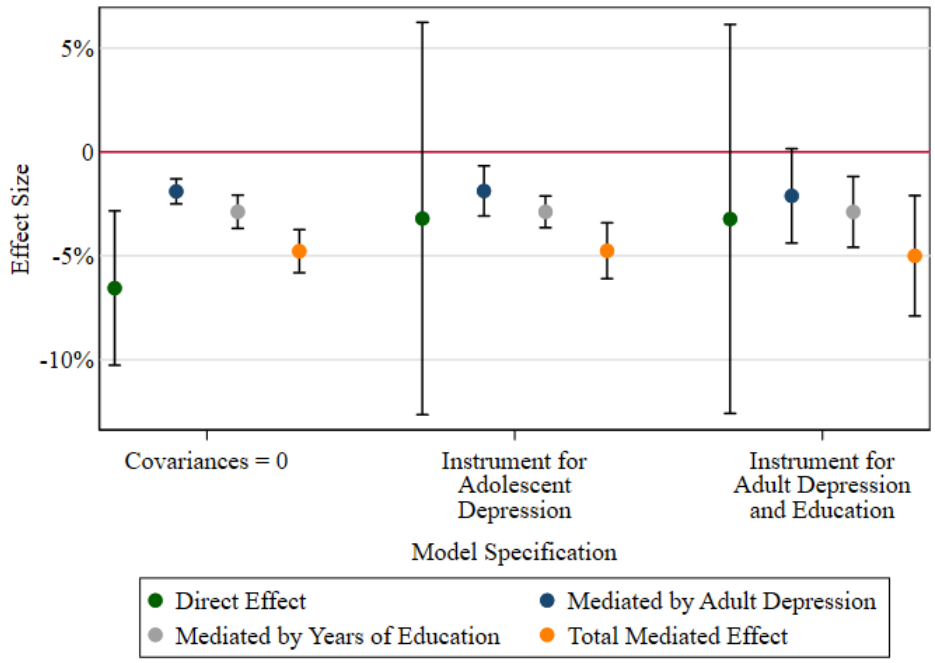
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(earnings)						
Adolescent Depression	-0.068**	-0.017	-0.036	-0.033	-0.023	-0.032
	(0.021)	(0.039)	(0.052)	(0.052)	(0.053)	(0.052)
Adult Depression	-0.140***	-0.143***	-0.140***	-0.155**	-0.184*	-0.164**
	(0.022)	(0.022)	(0.022)	(0.052)	(0.090)	(0.062)
Years of Education	0.092***	0.092***	0.092***	0.092***	0.110*	0.091***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.051)	(0.026)
Adult Depression						
Adolescent Depression	0.137***	0.080	0.128**	0.128**	0.126**	0.126**
	(0.009)	(0.086)	(0.042)	(0.042)	(0.042)	(0.042)
Family Member Passed Away in Past 12 Months					0.049	
					(0.042)	
Family/Friend Attempted Suicide in Past 12 Months					0.051***	0.051***
					(0.013)	(0.013)
Years of Education						
Adolescent Depression	-0.320***	-0.320***	-0.320***	-0.320***	-0.313***	-0.315***
	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.042)
Proportion 25+ with Bachelor's Degree					1.285***	1.575***
					(0.235)	(0.206)
Proportion 16-19 Enrolled in School					0.478*	
					(0.214)	
Adolescent Depression						
Religiosity Score			-0.005***	-0.005***	-0.005***	-0.005***
			(0.001)	(0.001)	(0.001)	(0.001)
$\sigma_{Y,E}$	= 0	= 0	= 0	0.000	-0.036	0.002
				(0.001)	(0.100)	(0.050)

σ_{Y,D_4}	= 0	= 0	= 0	0.009	0.027	0.014
				(0.030)	(0.055)	(0.036)
σ_{Y,D_1}	= 0	-0.031	-0.021	-0.021	-0.021	-0.021
		(0.020)	(0.028)	(0.028)	(0.028)	(0.028)
σ_{D_2,D_1}	= 0	0.114	0.018	0.017	0.022	0.021
		(0.180)	(0.080)	(0.080)	(0.081)	(0.082)
Z-Test Statistic: σ_{Y,D_1}			-0.291		0.000	0.000
Z-Test Statistic: σ_{D_2,D_1}			0.487		0.044	0.035
Z-Test Statistic: σ_{Y,D_2}					0.287	0.107
Z-Test Statistic: $\sigma_{Y,E}$					-0.036	0.002
Observations	14,561	14,561	14,561	14,561	14,561	14,561
Fixed Effects	School	School	School	School	School	School

Note: This table estimates the system of equations with earnings as the dependent variable using limited information maximum likelihood. For the adolescent and adult depression equations, average marginal effects are presented with standard errors in parentheses. For the earnings and years of education equations, coefficients are reported. Standard errors are clustered at the school level. In column 1, all covariance terms are restricted to zero. In column 2, σ_{Y,D_1} and σ_{D_2,D_1} are estimated. In column 3, I instrument for adolescent depression. Column 4 estimates σ_{Y,D_2} and $\sigma_{Y,E}$. Column 5 adds instrumental variables to the adult depression and years of education equations. Column 6 removes two weak instruments from the system. The Z-Test Statistics are the test statistics for the Z-test of the difference between the covariance terms with and without an instrument added.

**** for p < 0.001, ** for p < 0.01, * for p < 0.05*

Figure 21: Direct, Indirect, and Total Mediated Effects for Yearly Earnings



Note: Effect sizes are reported as percentage terms with 95% confidence intervals calculated using a parametric bootstrap with 500 replications. The results match columns 1, 3, and 6 of Table 9.

Table 10: Direct, Indirect, and Total Effects (% Change) for Yearly Earnings

	(1)	(2)	(3)
Direct	-6.56	-3.21	-3.23
	(1.90)	(4.82)	(4.78)
Mediated by Adult Depression	-1.90	-1.88	-2.12
	(0.31)	(0.62)	(1.16)
Mediated by Years of Education	-2.88	-2.88	-2.89
	(0.41)	(0.39)	(0.87)
Total Mediated Effect	-4.78	-4.76	-5.00
	(0.53)	(0.69)	(1.48)
Total Effect	-11.33	-7.96	-8.23
	(1.95)	(5.10)	(5.13)
Instrument for Adolescent Depression	No	Yes	Yes
Instrument for Adult Depression and	No	No	Yes

Note: Effect sizes are reported as percentage terms with standard errors in parentheses. The results match columns 1, 3, and 6 of Table 9.

in the marginal effect of adult depression on earnings leads the indirect effect through adult depression to grow to -2.12%. Little changes for the indirect effect through education – lower educational attainment resulting from adolescent depression results in a 2.89% drop in average earnings. Compared to estimating equations separately, the total effect drops from -11.33% to -8.23% and is statistically insignificant due to the drop in the direct effect and its noisy estimation. As displayed in Figure 21, the total *mediated* effect is about -5% and statistically significant at the 1% level.

In sum, system of equation results find that adolescent depression lowers adult earnings primarily through its impact on years of education and adult depression, rather than through a direct mechanism. The magnitudes of the indirect effects are robust to identification strategy and similar to results from the difference method.

Hourly Wages

Table 12 presents system of equation results for hourly wages. Across all specifications, interview date is relatively strong predictor of having observed wages. Someone who is interviewed on December 31 is about 32 percentage points less likely to have observed wages than someone who is interviewed on January 1 of the same year. The estimate of $\sigma_{Y,W}$ is about 0.15 and significantly different than zero, implying that positive sample selection bias is being addressed with the sample selection model. I interpret results with respect to the whole sample, not just the sample with observed wages.

Column 1 estimates each equation independently – the earnings equation is identical to column 4 in Table 5. Similar to earnings, there is evidence that both adult depression and years of education are mediators. Youth with depression receive about one third fewer years of education than youth without depression, and an additional year of education has a wage

premium of about 7%. Both estimates are statistically significant at the 0.1% level, implying a strong mediated effect for wages. Youth with depression are about 13.7 percentage points more likely to be depressed in adulthood. Those with depression in adulthood have hourly wages about 7.5% lower than those without depression, which accounts for a majority of the earnings gap in Table 9. Column 1 also estimates a direct effect of adolescent depression on wages of -4.8%, which is statistically significant at the 1% level. Similar to earnings results, there appears to be an economically significant direct effect on wages.

Column 2 also estimates the σ_{Y,D_1} and σ_{D_2,D_1} covariance terms, the terms needed to identify β_1 and γ_1 with instruments in column 3. The estimate of the direct effect flips signs and drops to 0.013 but is noisily estimated. The effect of adolescent depression on adult depression remains similar in magnitude but the standard errors become large. Both parameters are not well identified without instrumental variables.

Column 3 adds cohort religiosity to the adolescent depression equation to better identify β_1 and γ_1. Cohort religiosity is a strong predictor of adolescent depression – a one standard deviation increase in cohort religiosity leads to about a 1.7 percentage point decrease in the probability of having depression in adolescence. The Z-Test Statistic suggests that we fail to reject the null hypothesis that σ_{Y,D_1} is unchanged between columns 2 and 3, failing to find evidence that the included instruments are invalid. In contrast to the earnings results, instrumenting for adolescent depression leads to an increase in the direct effect to -0.062, or about -6%. While slightly larger in magnitude, this effect has large standard errors and is statistically indistinguishable from zero or other estimates of the direct effect. Instrumenting for adolescent depression also slightly drops its average marginal effect on adult depression, but

adolescent depression remains an important positive predictor of adult depression. The estimate is statistically significant at the 1% level.

Column 4 re-estimates column 3 but unrestricts σ_{Y,D_2} and $\sigma_{Y,E}$. The estimate of a return to a year of education is unchanged, while the effect of adult depression on adult earnings grows in magnitude to about -8.5%. In contrast to column 2, estimates remain precisely estimated.

Column 5 adds instrumental variables to the adult depression and years of education equations. The marginal effect of adult depression hourly wages grows to about -11.4% and remains significant at the 5% level – a magnitude noticeably larger than previous columns. Whether a friend or family member attempted suicide in the past year is a strong instrument significant at the 0.1% level. Having a death in the family is a weaker instrument only statistically significant at the 5% level. The Z-Test Statistic suggests that we fail to reject the null hypothesis that σ_{Y,D_2} is unchanged between columns 4 and 5, failing to find evidence that the included instruments are invalid. Similar to earnings results, instrumenting for adult depression reveals that it may play a larger mediating role than implied by column 1.

The mediating role of years of education is unchanged when instruments are used for identification. Youth with depression receive about one third fewer years of education – this estimate stays constant across all specifications because instruments are not used to identify the effect of adolescent depression on years of education. In column 5, the return to a year of education stays at about 7% and statistically significant at the 0.1% level. The tract-level proportion of adults with a bachelor's degree is a strong and meaningful instrument – a 0.10 increase in the proportion increases years of education by about 0.13. Tract-level school enrollment is only significant at the 5% level and explains less variation in years of education.

Finally, the direct effect of adolescent depression in column 5 drops slightly to -5.4% but is qualitatively similar to column 4.

Two of the instruments used in column 5 are only statistically significant at the 5% level, which is relatively weak and could be biasing estimates. In column 6, I remove having a death in the family and the proportion of those enrolled in school as instruments. The coefficient on adult depression drop slightly to -0.114, while the coefficient on years of education remains largely unchanged but is less precisely estimated. The direct effect is a similar magnitude as well. Whether the weaker instruments were driving results one way or the other is unclear – estimates in columns 5 and 6 remain qualitatively similar.

Figure 22 plots the direct, indirect, and total effects of adolescent depression on hourly wages as percent changes. All effects are calculated using the estimates from columns 1, 3, and 6 of Table 11 and are presented as point estimates in Table 12. When each equation in the system is estimated separately, the average direct effect on wages is about -4.77%. Adult depression mediates a -1.06% effect, while years of education mediates about a -2.21% effect. Similar to results for earnings, the direct effect makes up a majority of the total effect of adolescent depression on wages. Mediated effects make up a notable portion as well. Years of education plays a larger role relative to adult depression, consistent with the idea that education has a stronger impact on wages than adult depression. Treating adult depression or years of education as confounders would underestimate the total effect of adolescent depression on wages by about -3.25 percentage points.

When cohort religiosity is used as an instrumental variable in the adolescent depression equation, the direct effect rises slightly to -6% and is imprecisely estimated, more so than estimates for earnings. The effect mediated by adult depression drops slightly to -0.97% due to a

drop in the average marginal effect of adolescent depression on adult depression. The indirect effect through years of education is unchanged.

When instruments are also added to the adult depression and years of education equations, the direct effect drops slightly but remains similar to previous specifications. An increase in the marginal effect of adult depression on wages leads the indirect effect through adult depression to grow to -1.5%. The effect mediated by years of education remains constant at -2.22%. Across all specifications, results are similar. There is a direct effect of adolescent depression on wages of about -5-6%, which is imprecisely estimated when instruments are used. The effect mediated by years of education is about -2.21% and the effect mediated by adult depression is between -1-1.5%. The total mediated effect remains between -3.25-3.5%. In contrast to circumstantial evidence for earnings, it is unclear whether adolescent depression has a meaningful direct effect on hourly wages.

Sensitivity of Results

In this chapter I test the sensitivity of results to several potentially influential characteristics. First, I examine how different measures of depression impact results. Second, I use dichotomous measures of educational attainment. Finally, I conduct a brief family-level analysis. I find that my main results are not sensitive to changes in these factors.

Measure of Depression

To assess whether my findings are driven by the measures of depression that I have chosen, I consider three alternative measures of adolescent depression and one alternative measure of adult depression. First, I use the adolescent CES-D score. When used in linear regression, the coefficient on the CES-D score provides an average marginal effect of an increase in symptoms. Second, I use a dichotomous measure with a lower CES-D cutoff of 14+. Third, I

Table 11: System of Equations Results, Hourly Wages

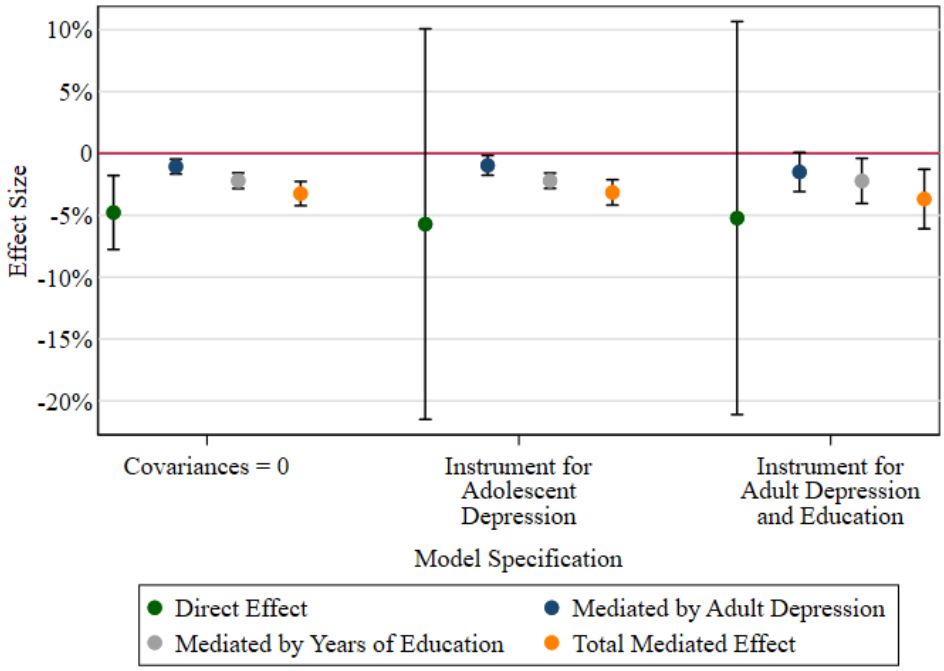
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Wages)						
Adolescent Depression	-0.049**	0.013	-0.062	-0.061	-0.055	-0.059
	(0.017)	(0.074)	(0.085)	(0.085)	(0.085)	(0.087)
Adult Depression	-0.078***	-0.079***	-0.077***	-0.087*	-0.121*	-0.114*
	(0.022)	(0.022)	(0.022)	(0.043)	(0.051)	(0.048)
Years of Education	0.070***	0.070***	0.070***	0.070***	0.069***	0.068*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.016)	(0.028)
Observed Wages						
Interview Date (/10)	-0.009***	-0.009***	-0.009***	-0.009***	-0.009***	-0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Adult Depression						
Adolescent Depression	0.492***	0.435	0.447**	0.446**	0.455***	0.497***
	(0.031)	(0.224)	(0.136)	(0.136)	(0.133)	(0.129)
Friend/Family Attempted					0.193***	0.194***
Suicide Past 12 Months					(0.046)	(0.046)
Family Member Passed					0.181*	
Away Past 12 Months					(0.083)	
Years of Education						
Adolescent Depression	-0.322***	-0.322***	-0.322***	-0.322***	-0.314***	-0.317***
	(0.041)	(0.041)	(0.041)	(0.041)	(0.042)	(0.042)
Proportion 25+ with					1.275***	1.571***
College Degree					(0.234)	(0.205)
Proportion in School					0.489*	
					(0.212)	
Adolescent Depression						
Cohort Religiosity Score			-0.018***	-0.018***	-0.018***	-0.018***
			(0.005)	(0.005)	(0.005)	(0.005)
$\sigma_{Y,W}$	0.152	0.151	0.151	0.151	0.152	0.152
	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)
$\sigma_{Y,E}$	=0	=0	=0	0.000	0.002	0.006
				(0.000)	(0.052)	(0.092)
σ_{Y,D_2}	=0	=0	=0	0.010	0.045	0.038

				(0.034)	(0.042)	(0.040)
σ_{Y,D_1}	=0	-0.066	0.014	0.014	0.013	0.015
		(0.078)	(0.088)	(0.088)	(0.088)	(0.088)
σ_{D_4,D_1}	=0	0.033	0.027	0.027	0.023	-0.003
		(0.130)	(0.079)	(0.079)	(0.076)	(0.074)
Observations	14,561	14,561	14,561	14,561	14,561	14,561

Note: This table estimates the system of equations with hourly ages as the dependent variable using limited information maximum likelihood. For the adolescent depression, adult depression, and selection equations, average marginal effects are presented with standard errors in parentheses. For the earnings and years of education equations, coefficients are reported. Standard errors are clustered at the school level. In column 1, all covariance terms are restricted to zero except for the covariance between the wages and the selection equations. In column 2, σ_{Y,D_1} and σ_{D_2,D_1} are estimated. In column 3, I instrument for adolescent depression. Column 4 estimates σ_{Y,D_2} and $\sigma_{Y,E}$. Column 5 adds instrumental variables to the adult depression and years of education equations. Column 6 removes two weak instruments from the system

**** for p < 0.001, ** for p < 0.01, * for p < 0.05*

Figure 22: Direct, Indirect, and Total Mediated Effects for Hourly Wages



Note: Effect sizes are reported as percentage terms with 95% confidence intervals calculated using a parametric bootstrap with 500 replications. The results match columns 1, 3, and 6 of Table 11.

Table 12: Direct, Indirect, and Total Effects (% Change) for Hourly Wages

	(1)	(2)	(3)
Direct	-4.77	-6.01	-5.72
	(1.53)	(8.05)	(8.10)
Mediated by Adult Depression	-1.06	-0.97	-1.50
	(0.31)	(0.41)	(0.81)
Mediated by Years of Education	-2.21	-2.21	-2.22
	(0.32)	(0.32)	(0.93)
Total Mediated Effect	-3.25	-3.15	-3.69
	(0.50)	(0.52)	(1.23)
Total Effect	-7.86	-8.68	-8.73
	(1.54)	(7.87)	(7.77)
Instrument for Adolescent Depression	No	Yes	Yes
Instrument for Adult Depression / Education	No	No	Yes

Note: Effect sizes are reported as percentage terms with standard errors in parentheses. The results match columns 1, 3, and 6 of Table 9.

use a dichotomous measure with a higher CES-D cutoff of 18+. For adult depression, I use the CES-D score as an alternative to the usual dichotomous measure. While each of these measures depression in a different way, similar results across measures would suggest that depressive symptoms are driving results, not an arbitrary cutoff or measure.

Table 13 presents system of equations results for earnings with each alternative measure. All specifications include school fixed effects. Each pair of columns (e.g., columns 1-2) compares results with covariances restricted to zero to results when instruments are used. Table 13 reveals three points about robustness to the measures of depression.

First, coefficient estimates without instruments (odd columns) are not dependent on the measures of depression used. In column 3, a lower cutoff for adolescent depression results in an estimate of the direct effect coefficient of -0.059; in column 5, a higher cutoff results in a coefficient of -0.048. Both are statistically different than zero but indistinguishable from the estimate of -0.068 in column 1. In column 7, using CES-D score produces a coefficient of -0.004. When scaled by the average difference in CES-D scores between youth with and without adolescent depression (see Table 1), this coefficient is about -0.049, comparable to estimates in columns 1, 3, and 5.

Second, using instruments for identification in each specification results in similar changes in results despite the measure of depression. The takeaways are the same. Compared to column 1, the direct effect in column 2 drops in magnitude and is imprecisely estimated. Using lower or higher cutoff scores in columns 4 and 6 leads to similar changes in the direct effect – a large drop in magnitude and statistical insignificance. When CES-D score is used for adolescent and adult depression in column 8, the direct effect grows slightly but the standard errors are

inflated. Overall, the impact that instrumental variables has on the results does not appear to be driven by the measures of depression used.

Third, the measures of depression used do not affect instrument relevance. Peer religiosity remains a strong instrument across all even columns, and the instruments used in the adult depression equation have similar levels of relevance. Altogether, Table 13 suggests that the measures of adolescent and adult depression do not have a significant impact on results, which remain qualitatively similar.

Table 14 presents identical sensitivity checks for a system of equations with hourly wages as the dependent variable. Similar to sensitivity checks for earnings, results are relatively robust.

First, coefficient estimates without instruments are similar across measures of depression. In column 1, adolescent depression has a direct effect coefficient of -0.049. A lower cutoff results in a coefficient of -0.033 and a higher cutoff results in a coefficient of -0.032. Adolescent CES-D score presents a coefficient of -0.003. While different cutoffs present slightly smaller direct effects, confidence intervals overlap and effects are qualitatively similar.

Second, adding instruments has a similar effect on direct effect coefficients. In column 2, using instruments for identification raises the direct effect to -0.059. With the lower cutoff in column 4, the direct effect falls slightly to -0.028. With a higher cutoff in column 6, the direct effect rises to -0.08. Using CES-D score presents an effect of 0.003 that is statistically insignificant. These results are consistent with the idea that instrumenting for adolescent depression produces a noisy estimate of the direct effect. In contrast to yearly earnings, the direct effect of adolescent depression on wages is not driven to zero with an alternative identification strategy.

Table 13: Earnings Results Robustness to Measures of Depression

Adolescent Measure	Normal		Low Cutoff		High Cutoff		CESD	
Adult Measure	Normal		Normal		Normal		CESD	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(Earnings)								
Adolescent Depression	-0.068**	-0.032	-0.059**	-0.039	-0.048*	-0.008	-0.004**	-0.008
	(0.021)	(0.052)	(0.019)	(0.086)	(0.024)	(0.031)	(0.002)	(0.010)
Adult Depression	-0.140***	-0.164**	-0.141***	-0.176*	-0.144***	-0.177*	-0.018***	-0.033
	(0.022)	(0.062)	(0.022)	(0.070)	(0.022)	(0.072)	(0.002)	(0.034)
Years of Education								
Adolescent Depression	-0.320***	-0.315***	-0.329***	-0.326***	-0.323***	-0.314***	-0.029***	-0.029***
	(0.041)	(0.042)	(0.037)	(0.037)	(0.041)	(0.042)	(0.003)	(0.003)
Proportion 25+ with Bachelor's Degree		1.575***		1.577***		1.573***		1.555***
		(0.206)		(0.205)		(0.206)		(0.208)
Adult Depression								
Adolescent Depression	0.492***	0.458**	0.463***	0.342*	0.513***	0.306***	0.185***	0.142***
	(0.031)	(0.142)	(0.029)	(0.158)	(0.032)	(0.038)	(0.007)	(0.032)
Family/Friend Attempted Suicide in Past 12 Months		0.192***		0.186***		0.186***		0.658***
		(0.045)		(0.045)		(0.045)		(0.144)
Adolescent Depression								
Religiosity Score		-0.018***		-0.018***		-0.017***		-0.101***
		(0.005)		(0.005)		(0.005)		(0.021)
$\sigma_{Y,E}$	=0	0.002	=0	0.000	=0	0.001	=0	-0.002
		(0.050)		(0.050)		(0.050)		(0.050)
σ_{Y,D_4}	=0	0.014	=0	0.021	=0	0.019	=0	0.072
		(0.036)		(0.041)		(0.042)		(0.152)
σ_{Y,D_1}	=0	-0.021	=0	-0.010	=0	-0.033	=0	0.044

		(0.028)		(0.053)		(0.015)		(0.055)
σ_{D_2,D_1}	=0	0.021	=0	0.073	=0	0.184	=0	0.059
		(0.082)		(0.095)		(0.021)		(0.043)
Observations	14,561	14,561	14,561	14,561	14,561	14,561	14,561	14,561

Note: Coefficients are reported with standard errors in parentheses. Every column includes a school fixed effect in each equation, and standard errors are clustered at the school level. Each pair of columns uses a different measure of depression and compares estimation with covariances restricted to zero to estimation with instruments. Columns 1-2 use the regular depression measures. Columns 3-4 use a low cut-off (14+) for adolescent depression, while columns 5-6 use a high cutoff (18+). Columns 7-8 use CES-D score for both adolescent and adult depression. When the CES-D score is used, the variable is modeled as continuous; when a cut-off is used, the variable is treated as binary and a probit part is used in the likelihood function.

**** for p < 0.001, ** for p < 0.01, * for p < 0.05*

Table 14: Hourly Wages Results Robustness to Measures of Depression

Adolescent Measure	Normal		Low Cutoff		High Cutoff		CESD	
Adult Measure	Normal		Normal		Normal		CESD	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hourly Wages								
Adolescent Depression	-0.049**	-0.059	-0.033*	0.028	-0.032	-0.080	-0.003*	0.003
	(0.017)	(0.087)	(0.015)	(0.088)	(0.017)	(0.062)	(0.001)	(0.009)
Adult Depression	-0.078***	-0.114*	-0.081***	-0.120*	-0.081***	-0.108*	-0.010***	-0.058*
	(0.022)	(0.048)	(0.022)	(0.048)	(0.022)	(0.044)	(0.002)	(0.026)
Adult Depression								
Adolescent Depression	0.492***	0.497***	0.463***	0.362**	0.513***	0.448***	0.185***	0.143***
	(0.031)	(0.129)	(0.029)	(0.138)	(0.032)	(0.134)	(0.007)	(0.032)

Friend or Family Attempt Suicide		0.194***		0.188***		0.193***		0.682***
		(0.046)		(0.046)		(0.046)		(0.144)
Years of Education								
Adolescent Depression	-0.322***	-0.317***	-0.329***	-0.326***	-0.323***	-0.314***	-0.029***	-0.029***
	(0.041)	(0.042)	(0.037)	(0.037)	(0.041)	(0.042)	(0.003)	(0.003)
Tract Proportion with College Degree		1.571***		1.578***		1.573***		1.555***
		(0.205)		(0.204)		(0.206)		(0.207)
Adolescent Depression								
Cohort Religiosity		-0.018***		-0.018***		-0.020***		-0.100***
		(0.005)		(0.005)		(0.006)		(0.021)
$\sigma_{Y,W}$	0.152	0.152	0.156	0.156	0.156	0.155	0.156	0.150
	(0.050)	(0.050)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.047)
$\sigma_{Y,E}$		0.006		-0.008		0.000		0.001
		(0.092)		(0.093)		(0.093)		(0.088)
σ_{Y,D_2}		0.038		0.039		0.029		0.346
		(0.040)		(0.039)		(0.036)		(0.170)
σ_{Y,D_1}		0.015		-0.061		0.055		0.034
		(0.088)		(0.094)		(0.063)		(0.083)
σ_{D_4,D_1}		-0.003		0.061		0.038		0.059
		(0.074)		(0.083)		(0.073)		(0.044)
Observations	14,561	14,561	14,561	14,561	14,561	14,561	14,561	14,561

Note: Coefficients are reported with standard errors in parentheses. Every column includes a school fixed effect in each equation, and standard errors are clustered at the school level. Each pair of columns uses a different measure of depression and compares estimation with covariances restricted to zero to estimation with instruments. Columns 1-2 use the regular depression measures. Columns 3-4 use a low cut-off (14+) for adolescent depression, while columns 5-6 use a high cutoff (18+). Columns 7-8 use CES-D score for both adolescent and adult depression. When

the CES-D score is used, the variable is modeled as continuous; when a cut-off is used, the variable is treated as binary and a probit part is used in the likelihood function.

**** for $p < 0.001$, ** for $p < 0.01$, * for $p < 0.05$*

Third, consistent with results for earnings, the measures of depression used do not affect instrument relevance. Peer religiosity remains a strong instrument across all even columns, and the instruments used in the adult depression equation have similar levels of relevance.

Measure of Education

Results suggest that adolescent depression has an important indirect effect on earnings through years of education. Understanding whether depression affects each level of education differently could provide insight into where this indirect effect is the most important and have important policy implications.

I split reported educational attainment into four categories: less than high school, high school diploma, some college, and bachelor's degree or higher. I use three linear probability models[11] to estimate the effect of adolescent depression on the probability of moving from one group to another. Each model only uses respondents who report to be in at least the lower of the two categories (e.g., the model estimating the effect on college graduation uses those with some college or greater). I then estimate the earnings equation where I replace years of education with the dummies for each level of educational attainment. Indirect effects are calculated by multiplying the marginal effect of adolescent depression on the probability of an educational outcome by the change in earnings with respect to having that level of education.

Table 15 presents marginal effects of the education and earnings models and finds evidence that adolescent depression is important across several margins of educational attainment. Column 1 finds that adolescent depression leads to about a 3.3 percentage point drop in the probability of graduating high school, all else constant. Among those who have a high

[11] A linear probability model is used instead of a probit so that observations from small schools without much variation in educational attainment can be used in estimation.

school degree, column 2 suggests that youth with adolescent depression are about 5.5 percentage points less likely to enroll in college. Finally, among youth who enroll in college, column 3 shows that those with adolescent depression are about 3.5 percentage points less likely to finish college. Column 4 presents marginal effects of educational attainment on log earnings, where less than a high school diploma is the base category. Each level of education is an important determinant of earnings, with those with a bachelor's degree earning nearly double than those without a high school degree.

Table 16 presents estimates of the indirect effects through each margin of education. When scaled by the returns to education, adolescent depression leads to a 0.90% drop in earnings through the margin of graduating high school, a 0.66% drop through the probability of enrolling in college, and a 1.07% drop via whether they finish college. This highlights that an additional year of education that leads to a degree is more meaningful than one that does not. However, the total indirect effect of -2.63% is similar and statistically indistinguishable from the estimated effect of -2.88% in column 1 of Table 10. When years of education is used. This suggests that accounting for the nonlinear returns to education does not qualitatively change results.

Although I do not explore the several unobserved mechanisms that could be driving differences in earnings (e.g., quality of education, occupation), results are suggestive that adolescent depression has broad negative effects on earnings throughout the educational career of an adolescent. These effects are more likely to be direct in the case of high school completion, as symptoms of depression in adolescence may immediately affect the youth's ability to perform well in school. However, the effects on enrolling in college and/or finishing college could be mediated by several pathways, including changing expectations of the youth's career path, their social network, or their symptoms of depression in early adulthood.

Table 15: Robustness to Discrete Measures of Education

	(1) *High School*	(2) *Some College*	(3) *Bachelor's*	(4) *Ln(Earnings)*
Adolescent Depression	-0.034***	-0.057***	-0.036**	-0.069**
	(0.006	(0.009)	(0.012)	(0.021)
High School Diploma				0.262***
				(0.049)
Some College				0.379***
				(0.051)
Bachelor's Degree				0.678***
				(0.055)
Observations	14,561	13,505	9,775	13,174

Note: Coefficients are reported with standard errors in parentheses. Standard errors are clustered at the school level. Columns 1-3 estimate a linear probability model with high school completion, some college, and bachelor's degree or more as dependent variables, respectively. Youth on or above the margin of the outcome are included in each regression. Column 4 presents earnings results with the ordered measure of education. All specifications include a school fixed effect.

**** p < 0.001, ** p < 0.01, * p < 0.05*

Table 16: Indirect Effects of Adolescent Depression on Earnings by Margin of Education

	(1) High School	(2) Some College	(3) Bachelor's	(4) Total
Adolescent Depression	-0.90%**	-0.66%***	-1.07%***	-2.60%***
	(0.19)	(0.19)	(0.34)	(0.46)

Note: Indirect effects are presented as percent changes with standard errors in parentheses. Standard errors are non-parametrically bootstrapped with 100 replications. Effect sizes are calculated using results from Table 15.

**** p < 0.001, ** p < 0.01, * p < 0.05*

Family-Level Heterogeneity

To examine the robustness of results to family-level heterogeneity, I estimate the earnings, adult depression, and years of education equations using the family sample and assess whether including a family fixed effect significantly alters coefficient estimates. The Add Health surveyed 5,512 youth from 2,633 families. The family sample in this paper consists of all respondents in the full sample who have at least one other family member in the full sample, resulting in 2,593 youth from 1,261 families. To appropriately assess the impact of a family fixed effect, I focus on the 'identifying sample' for each coefficient (Miller et al., 2019), which is the sample of respondents with within-family variation in the variable of interest. I compare the coefficient produced by OLS with this sample to the coefficient produced when a family fixed effect is included. Since there are potential sample selection issues into the family sample and the identifying sample, I focus on the change in estimates when family fixed effects are included, not how they differ from estimates using the full sample.

Table 17 presents earnings results for the family sample. In column 1, adolescent depression has about a -5.6% direct effect on earnings that is statistically insignificant. Column 2 estimates the same model for the identifying sample for adolescent depression, where the direct effect drops slightly to -5.4%. When a family fixed effect is added in column 5, the direct effect grows to -7.6% but remains noisy and statistically insignificant. Column 3 estimates the model using the identifying sample for adult depression, finding about a -20% average drop in earnings. This result is unchanged when a family fixed effect is added in column 5. Column 4 estimates the model using the identifying sample for years of education, which increases only slightly in column 5.

Table 18 presents regression results for adult CES-D score and years of education with the family sample. In columns 1-3, I estimate an equation for adult CES-D score using OLS. I use adult CES-D score as the dependent variable to avoid issues of incidental parameters bias in a probit with a large number of fixed effects. Column 1 runs OLS on the whole family sample. When only the identifying sample is used in column 2, the marginal effect of adolescent depression drops notably. Compared to column 2, adding a family fixed effect in column 3 has a marginal impact on the estimated coefficient, bringing it from 2.047 to 1.907. In columns 4-6, I estimate the years of education equation with OLS. When the whole family sample is used, the coefficient is -0.30 and statistically significant, similar to main results in Table 10. When only the identifying sample is used, the coefficient drops to -0.046 and is statistically insignificant, indicating that the coefficient is sensitive to sample selection. When a family fixed effect is added, the effect of adolescent depression on years of education decreases slightly from -0.046 to -0.044, suggesting that family-level heterogeneity has little effect on the coefficient estimate. If we had not used the identifying sample results as a baseline, we would incorrectly infer that family-level heterogeneity plays a large role.

Time-invariant, family-level heterogeneity appears to have minimal impacts on coefficient estimates. While estimates of adolescent depression's impacts on adult depression, years of education, and earnings change slightly, they are statistically indistinguishable from estimates without fixed effects and are qualitatively similar.

Table 17: Yearly Earnings Family Sample Results

	(1)	(2)	(3)	(4)	(5)
Adolescent Depression	-0.058	**-0.055**	-0.059	-0.013	**-0.079**
	(0.045)	**(0.071)**	(0.079)	(0.058)	**(0.065)**
Adult Depression	-0.224***	-0.382***	**-0.224****	-0.263***	**-0.229*****
	(0.051)	(0.085)	**(0.070)**	(0.070)	**(0.066)**
Years of Education	0.103***	0.113***	0.115***	**0.091*****	**0.096*****
	(0.010)	(0.018)	(0.020)	**(0.012)**	**(0.016)**
Sample	Family	ID	ID	ID	Family
Observations	2,593	931	828	1,640	2,593
Fixed Effect	None	None	None	None	Family

Note: Coefficients are reported with standard errors in parentheses. Standard errors are clustered at the family level. Ln(Earnings) is the dependent variable in all columns. Column 1 runs the full model on the family sample using OLS. Columns 2-4 run the full model on their respective 'identifying samples' – the samples which have within-family variation in the variable of interest. This is done so that the presented coefficients in columns 2-4 are identified by the same sample as the coefficients in column 5. In column 5, a family fixed effect is added, addressing within-family heterogeneity. All confounding variables are included in all columns, except for family characteristics that are excluded from column 5. I also control for whether a youth was the first born in their family.

**** for p < 0.001, ** for p < 0.01, * for p < 0.05*

Table 18: Adult Depression and Years of Education Family Sample Results

	Adult CES-D Score			Years of Education		
	(1)	(2)	(3)	(4)	(5)	(6)
Adolescent Depression	2.341***	2.047***	1.907***	-0.300***	-0.046	-0.044
	(0.234)	(0.317)	(0.318)	(0.088)	(0.113)	(0.108)
Sample	Family	ID	Family	Family	ID	Family
Observations	2,593	931	2,593	2,593	931	2,593
Fixed Effect	None	None	Family	None	None	Family

Note: Coefficients are reported with standard errors in parentheses. Standard errors are clustered at the family level. In column 1, adult CES-D score is regressed on adolescent depression and a full set of

covariates with OLS. Column 2 runs column one with the identifying sample, and column 3 adds a family fixed effect. Column 4 regresses years of education on adolescent depression and a full set of covariates using OLS. Column 5 runs column 4 with the identifying sample and column 6 adds a family fixed effect. In columns 3 and 6, only the 'identifying sample' is used – the sample of youth from a family with within-family variation in adolescent depression. While the whole family sample is being used in columns 2 and 4, only the identifying sample is identifying the coefficient on adolescent depression. Therefore, the same sample identifies the coefficient on adolescent depression across columns 2-3 and 5-6.

**** for p < 0.001, ** for p < 0.01, * for p < 0.05*

CHAPTER VIII: DISCUSSION AND CONCLUSION

This dissertation investigates why adolescents with depression do worse in the labor market as adults. I use a nationally representative survey of youth in the United States and a mediation analysis framework to estimate the direct and indirect effects of adolescent depression on adult earnings and wages. I present detailed descriptive and causal evidence outlining the relationship between adolescent depression and adult labor market outcomes.

The descriptive results teach us three things. First, we find consistent evidence that educational attainment and adult depression play a mediating role between adolescent depression and adult earnings. In Table 4, using the difference method finds that years of education mediates about a -1.9% effect on adult earnings, all else constant. Adult depression also mediates about a -1.9% average effect on adult earnings. A combination of single-equation results confirm these findings by finding strong relationships between adolescent depression, adult depression, educational attainment, and earnings. The estimated mediated effects are both less than -2%. While smaller in magnitude than some estimates in the literature, a combined mediated effect of -3.8% equates to about a $1,335 loss for the average respondent in this sample. The mediating roles of adult depression and education confirm the conceptual framework outlined at the beginning of this dissertation. These results also confirm the graphical relationships between adolescent depression, education, adult depression, and earnings/wages found in Chapter V.

Second, I find evidence that part of adolescent depression's effects on earnings may be via its effects on hourly wages. After accounting for sample selection issues, Table 6 finds that adolescent depression has about a -7.2% total effect on hourly wages, compared to a -10.5% total effect on yearly earnings. This implies that a majority of the total effect on earnings is driven by adolescent depression's effects on hourly wages. Difference method results imply that about 1.1

percentage points of the total effect is mediated by adult depression, and 1.3 percentage points of the effect is mediated by years of education. This is consistent with the conceptual framework behind how changes in productivity show up in wages, which in turn lower earnings. Youth with depression in adolescence achieve a lower level of schooling, which leads to lower productivity or a worse signal in the labor market, ultimately winding up with lower wages in adulthood. Early onset depression increases the likelihood of depression in adulthood, which can dampen productivity at work and make it difficult to advance. The slightly larger effect mediated by education could be indicative of education playing a stronger role when determining wages.

Third, throughout all specifications there is a negative and statistically significant direct effect of adolescent depression on adult earnings and wages. Using a mediation analysis approach reveals that some of the previous literature has overestimated the direct effect by omitting important mediators or confounding variables. Some papers have omitted educational attainment, adult depression, and other important factors, artificially biasing the direct effect away from zero. When I approximately replicate these specifications, results from some previous literature are relatively consistent. This is a central finding – clearly defining mediating and confounding variables, as well as the assumptions required to interpret effects, is crucial when researching this topic. A conceptual framework that outlines why variables are excluded and what effect that has on estimated effects is a bare minimum to interpreting estimated effects in a meaningful way.

The statistical and economic significance of the direct effect stands in stark contrast to the idea that an adolescent condition should not have a direct effect on an adult outcome. Several things could be driving the result of a significant direct effect. First, there could be omitted mediating variables. If an important mediating variable were omitted from the regression, then

the direct effect would pick up the mediated effect, similar to how it does in the difference method. Second, measurement error could lead to a similar result. Measurement error of a mediator is a common cause of an overstated direct effect in mediation analysis (T. J. VanderWeele, 2016). Omitted confounding variables could also drive the direct effect away from zero, as well as bias other parameter estimates. Finally, it is possible that the model is misspecified and that there is a conceptual reason for there to be a large direct effect of adolescent depression on adult labor market outcomes.

Most of these possibilities imply that the estimates in the earnings and wages regressions could be biased. Additionally, coefficient estimates on adult depression and years of education could also be biased. Adult labor market outcomes could influence depressive symptoms, resulting in two-way effects. Years of education is commonly thought to be endogenous in a wage equation due to its correlation with ability and family characteristics. It could also be that the impact of adolescent depression on adult depression is biased if common unobserved factors impact the likelihood of the onset of depression. In light of the single equation results, the system of equation results address several potential issues of endogeneity and reveal new insights about the relationship between adolescent depression and adult earnings and wages.

First, I find that the mediating pathways of adult depression and educational attainment are robust to identification strategy, which confirms the conceptual framework of this dissertation and peer-reviewed literature. When the product method is used to estimate direct and indirect effects, estimated indirect effects are slightly larger than those from the difference method. Youth with depression earn nearly 5% less as adults due to the impact of adolescent depression on adult depression and years of education. When instruments are used for identification, the indirect effects via years of education and adult depression stay relatively

constant, and there is little evidence that effects are biased. Adult depression mediates more than a 2% drop in earnings, while years of education mediates nearly a 3% drop in earnings. Sensitivity checks suggest that these pathways are robust to the measure of depression used, the measure of education used, and family-level heterogeneity.

Second, in contrast to mediated effects, the direct effect is not robust to identification strategy. When peer religiosity is used to instrument for adolescent depression, the direct effect of adolescent depression on earnings drops closer to zero and is imprecisely estimated. This evidence is robust across specifications for yearly earnings, suggesting that the initial large and statistically significant direct effect is at least in part driven by omitted variables bias. This is a key result of this dissertation. In line with the conceptual framework of why adolescent depression affects adult labor market outcomes, it is not the direct effect, but mediated effects that are most important. A large and statistically significant direct effect in an earnings equation is likely driven by the omission of important confounding or mediating variables and should not be interpreted as a causal effect without further reasoning.

I also find that adolescent depression's impact on hourly wages accounts for a majority of its impact on yearly earnings, suggesting that the effect is driven by changes in productivity. The total mediated effect of adolescent depression on wages is about -3.7%, compared to -5% for yearly earnings. About -1.5 percentage points are mediated by adult depression and -2.2 percentage points are mediated by years of education. The differences in the magnitudes of indirect effects are not as pronounced as we would expect from the conceptual framework. While educational attainment is expected to play an especially strong role in determining wages, adult depression mediates a meaningful effect. This could be picking up on how the persistence of depressive symptoms in adulthood affect career choices that impact occupation, the likelihood of

getting a promotion, and movement between jobs. In contrast, this could also be picking up on a how hourly wages are measured. Improved measures of hourly wages and labor supply would provide more comprehensive estimates of how the mediating roles of educational attainment and adult depression differ.

The results of this dissertation suggest a large added economic benefit to improving prevention and treatment of adolescent depression - the total mediated effect of -5% is equivalent to about a $1,750 drop in annual earnings. However, this effect is likely an underestimate of the total effect of adolescent depression on earnings if there are more complex mediating pathways. For example, the potential mediating role of alcohol use could be responsible for additional drops in earnings later in life. Since I do not account for differences in receiving treatment for depression, estimated effects are in spite of rates of treatment over the life course. The effect of untreated adolescent depression on earnings is likely much larger. This highlights the potentially large remaining economic benefits that could be achieved from better identifying and treating adolescent depression. All else constant, decreasing symptoms of depression in adolescence will increase educational attainment, decrease the likelihood of depression in adulthood, and lead to increased earnings and wages in the long run.

The prominence of the mediating roles of adult depression and education suggest there may also be additional ways of avoiding lost earnings due to adolescent depression. Holding depressive symptoms in adolescence constant, interventions that prevent those with depression from dropping out of school prematurely may improve educational attainment and avoid a significant portion of the earnings gap. For example, Supported Education interventions hire staff at educational institutions to identify and work with young adults with psychological disorders to help achieve their educational goals (Ringeisen et al., 2017). Interventions at the school level that

simultaneously support education and refer participants to mental health services could address several indirect pathways at once.

Part of the earnings gap could also be addressed by identifying and treating depression in adulthood, weakening the link between adolescent and adult depression. This confirms the idea that it is better late than never – even if depression is not identified or treated in adolescence, some of its labor market consequences can be avoided by treatment later in life. Interventions that help adults with mental illness find and maintain gainful employment could be worthwhile. One such type of policies are Supported Employment policies (Marshall et al., 2014). If youth suffer from depression for an extended period before receiving treatment, it is possible that some of these policies would be the most effective at preventing lower earnings in the long run. Identifying depression in early adulthood could also have implications for earnings via changes in occupation choice.

Results of this dissertation also complement several findings and predictions in previous literature that adult depression and educational attainment may be mediators in this context (Fergusson et al., 2007; J. Fletcher, 2013; Johar & Truong, 2014; Smith & Smith, 2010). For example, results are consistent with Philipson et al. (2020), which finds that about half of the total effect is mediated by adult depression. I also corroborate the evidence presented by Johar & Truong (2014) that educational attainment plays an important mediating role, as well as broad evidence in Lundborg et al. (2014) that adolescent health and mental health problems have labor market consequences.

There are several limitations to the results presented in this dissertation. First, establishing causality without experimental data is difficult. Although the baseline results of this paper are consistent with previous literature, small violations in identifying assumptions could

result in bias that alters results. Future research that uses data from randomized, experimental contexts would create the strongest case for identifying causal effects of adolescent depression on outcomes. Second, self-reported measures of mental health and labor market outcomes, while generally reliable, often suffer from measurement error that could bias effects towards zero. While the CES-D is a tested measure of symptoms of depression, it does not provide the same level of detail that would be obtained from a diagnosis or in-person assessment. Longitudinal surveys that include interactions with healthcare professionals would allow researchers to pinpoint exactly what mental health disorders are responsible for long-term consequences.

Finally, this dissertation is a snapshot connecting data from one point in adolescence to data from one point in adulthood. I am unable to construct long-term histories of labor market outcomes and mental health status that could provide a more detailed look at long-term effects of adolescent depression. The evidence presented in this paper should be taken in context of other work done with longer histories of outcomes (e.g., Philipson et al., 2020).

The long-term effects of adolescent depression are only one part of a much larger area of research: understanding the labor market consequences of adolescent mental health, substance use, and their comorbidities. Our understanding of the long-term impacts of these conditions and behaviors has grown significantly over the past 15 years. However, this dissertation highlights that there are many more methodological, theoretical, and empirical contributions left to be made. In particular, improved measures of mental illness (e.g., anxiety, ADHD) in longitudinal surveys will enable researchers to better track the correlates and consequences of understudied disorders. The unique role of comorbidities is also a topic that is understudied relative to its potential impact on life trajectories and socioeconomic outcomes.

The relationship between adolescent mental illness and criminal justice involvement may also play an important role in shaping long term labor market outcomes. While this dissertation was not able to identify important effects potentially mediated by criminal justice involvement, future research should consider a more detailed view of its role in determining the economic consequences of mental illness. The interconnectedness of criminal justice involvement, educational attainment, and adult mental health makes identifying appropriate mediating pathways especially difficult.

In sum, this dissertation provides important new estimates of the various effects of adolescent depression on adult earnings and wages. These estimates improve our understanding of the consequences of adolescent depression, highlight pathways that can be targeted by policies to avoid its adverse impacts, and provide a framework for future work on the topic.

REFERENCES

Ainsworth, J. W. (2002). Why Does It Take a Village? The Mediation of Neighborhood Effects on Educational Achievement. *Social Forces*, *81*(1), 117–152. https://doi.org/10.1353/sof.2002.0038

Baldwin, M. L., & Marcus, S. C. (2007). Labor Market Outcomes of Persons with Mental Disorders. *Industrial Relations*, *46*(3), 481–510. https://doi.org/10.1111/j.1468-232X.2007.00478.x

Banerjee, S., Chatterji, P., & Lahiri, K. (2017). Effects of Psychiatric Disorders on Labor Market Outcomes: A Latent Variable Approach Using Multiple Clinical Indicators: Psychiatric Disorders and Labor Market Outcomes. *Health Economics*, *26*(2), 184–205. https://doi.org/10.1002/hec.3286

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, *70*(5, Part 2), 9–49. https://doi.org/10.1086/258724

Belmaker, R. H., & Agam, G. (2008). Major Depressive Disorder. *The New England Journal of Medicine*, *358*(1), 55–68.

Bishu, S. G., & Alkadry, M. G. (2017). A Systematic Review of the Gender Pay Gap and Factors That Predict It. *Administration & Society*, *49*(1), 65–104. https://doi.org/10.1177/0095399716636928

Blodgett, J. M., Lachance, C. C., Stubbs, B., Co, M., Wu, Y.-T., Prina, M., Tsang, V. W. L., & Cosco, T. D. (2021). A systematic review of the latent structure of the Center for Epidemiologic Studies Depression Scale (CES-D) amongst adolescents. *BMC Psychiatry*, *21*(1), 197. https://doi.org/10.1186/s12888-021-03206-1

Bobonis, G. J., & Finan, F. (2009). Neighborhood Peer Effects in Secondary School Enrollment Decisions. *Review of Economics and Statistics*, *91*(4), 695–716. https://doi.org/10.1162/rest.91.4.695

Bohman, H., Jonsson, U., Päären, A., Knorring, A.-L. V., Olsson, G., & Knorring, L. V. (2010). Long-term follow-up of adolescent depression. A population-based study. *Upsala Journal of Medical Sciences*, *115*(1), 21–29. https://doi.org/10.3109/03009730903572057

Bray, J. W. (2005). Alcohol Use, Human Capital, and Wages. *Journal of Labor Economics*, *23*(2), 279–312. https://doi.org/10.1086/428025

Cheadle, J. E., & Schwadel, P. (2012). The ‘friendship dynamics of religion,’ or the ‘religious dynamics of friendship’? A social network analysis of adolescents who attend small schools. *Social Science Research*, *41*(5), 1198–1212. https://doi.org/10.1016/j.ssresearch.2012.03.014

Conway, C. C., Rancourt, D., Adelman, C. B., Burk, W. J., & Prinstein, M. J. (2011). Depression socialization within friendship groups at the transition to adolescence: The roles of gender and group centrality as moderators of peer influence. *Journal of Abnormal Psychology*, *120*(4), 857–867. https://doi.org/10.1037/a0024779

Cseh, A. (2008). The Effects of Depressive Symptoms on Earnings. *Southern Economic Journal*, *75*(2), 383–409. eoh.

Eisenberg, D., Golberstein, E., & Hunt, J. B. (2009). Mental Health and Academic Success in College. *The B.E. Journal of Economic Analysis & Policy*, *9*(1). https://doi.org/10.2202/1935-1682.2191

Ettner, S. L., Frank, R. G., & Kessler, R. C. (1997). *The Impact of Psychiatric Disorders on Labor Market Outcomes*. 19.

Fergusson, D. M., Boden, J. M., & Horwood, L. J. (2007). Recurrence of major depression in adolescence and early adulthood, and later mental health, educational and economic outcomes. *British Journal of Psychiatry*, *191*(4), 335–342. https://doi.org/10.1192/bjp.bp.107.036079

Fletcher, J. (2013). Adolescent Depression and Adult Labor Market Outcomes. *Southern Economic Journal*, *80*(1), 26–49. https://doi.org/10.4284/0038-4038-2011.193

Fletcher, J. M. (2008). Adolescent depression: Diagnosis, treatment, and educational attainment. *Health Economics*, *17*(11), 1215–1235. https://doi.org/10.1002/hec.1319

Fox, L., Torche, F., & Waldfogel, J. (2016). *The Oxford Handbook of the Social Science Policy: Intergenerational Mobility*. 10.1093/oxfordhb/9780199914050.013.24

Frijters, P., Johnston, D. W., & Shields, M. A. (2014). The Effect of Mental Health on Employment: Evidence from Australian Panel Data. *Health Economics*, *23*(9), 1058–1071. eoh.

Fruehwirth, J. C., Iyer, S., & Zhang, A. (2019). Religion and Depression in Adolescence. *Journal of Political Economy*, *127*(3), 1178–1209. https://doi.org/10.1086/701425

Giletta, M., Scholte, R. H. J., Prinstein, M. J., Engels, R. C. M. E., Rabaglietti, E., & Burk, W. J. (2012). Friendship Context Matters: Examining the Domain Specificity of Alcohol and Depression Socialization Among Adolescents. *Journal of Abnormal Child Psychology*, *40*(7), 1027–1043. https://doi.org/10.1007/s10802-012-9625-8

Ginther, D., Haveman, R., & Wolfe, B. (2000). Neighborhood Attributes as Determinants of Children's Outcomes: How Robust Are the Relationships? *The Journal of Human Resources*, *35*(4), 603. https://doi.org/10.2307/146365

Gollan, J., Raffety, B., Gortner, E., & Dobson, K. (2005). Course profiles of early- and adult-onset depression. *Journal of Affective Disorders*, *86*(1), 81–86. https://doi.org/10.1016/j.jad.2004.12.009

Hammen, C. (2005). Stress and Depression. *Annual Review of Clinical Psychology*, *1*(1), 293–319. https://doi.org/10.1146/annurev.clinpsy.1.102803.143938

Hann, D., Winter, K., & Jacobsen, P. (1999). *MEASUREMENT OF DEPRESSIVE SYMPTOMS IN CANCER PATIENTS: EVALUATION OF THE CENTER FOR EPIDEMIOLOGICAL STUDIES DEPRESSION SCALE (CES-D)*. 7.

Harmon, C., Oosterbeek, H., & Walker, I. (2003). The Returns to Education: Microeconomics. *Journal of Economic Surveys*, *17*(2), 115–156. https://doi.org/10.1111/1467-6419.00191

Harris, K. M. (2009). *The National Longitudinal Study of Adolescent to Adult Health (Add Health), Waves I & II, 1994–1996; Wave III, 2001–2002; Wave IV, 2007-2009*. Carolina Population Center, University of North Carolina at Chapel Hill.

Hoffman, F., Petermann, F., Glaeske, G., & Bachmann, C. (2012). *Prevalence and comorbidities of adolescent depression in Germany. An analysis of Health Insurance Data*. *40*(6), 399–404.

Johar, M., & Truong, J. (2014). Direct and indirect effect of depression in adolescence on adult wages. *Applied Economics*, *46*(36), 4431–4444. https://doi.org/10.1080/00036846.2014.962227

Kandel, D. B. (2021). Homophily, Selection, and Socialization in Adolescent Friendships. *American Journal of Sociology*, 11.

Kessler, R. C., Angermeyer, M., Anthony, J. C., Graaf, R. D., Gasquet, I., Girolamo, G. D., Gluzman, S., Gureje, O., Haro, J. M., Kawakami, N., Karam, A., Levinson, D., Medina, M. E., Browne, M. A. O., Posada-Villa, J., Stein, D. J., Tsang, C. H. A., Aguilar-Gaxiola, S., Alonso, J., … Üstün, T. B. (2007). Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry*, 9.

Kilburn, K., Thirumurthy, H., Halpern, C. T., Pettifor, A., & Handa, S. (2016). Effects of a Large-Scale Unconditional Cash Transfer Program on Mental Health Outcomes of Young People in Kenya. *Journal of Adolescent Health*, *58*(2), 223–229. https://doi.org/10.1016/j.jadohealth.2015.09.023

Levin, J. (2010). Religion and mental health: Theory and research. *International Journal of Applied Psychoanalytic Studies*, n/a-n/a. https://doi.org/10.1002/aps.240

Maddala, G. S., & Lee, L.-F. (1976). Recursive Models with Qualitative Endogenous Variables. *NBER*, *5*(4), 525–545.

Marshall, T., Goldberg, R. W., Braude, L., Dougherty, R. H., Daniels, A. S., Ghose, S. S., George, P., & Delphin-Rittmon, M. E. (2014). Supported Employment: Assessing the Evidence. *Psychiatric Services*, *65*(1), 16–23. https://doi.org/10.1176/appi.ps.201300262

McCullough, M. E., & Larson, D. B. (1999). Religion and depression: A review of the literature. *Twin Research*, *2*(2), 126–136. https://doi.org/10.1375/twin.2.2.126

McLeod, J. D., Uemura, R., & Rohrman, S. (2012). Adolescent Mental Health, Behavior Problems, and Academic Achievement. *Journal of Health and Social Behavior*, *53*(4), 482–497. https://doi.org/10.1177/0022146512462888

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, *27*(1), 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Miller, D., Shenhav, N., & Grosz, M. (2019). *Selection into Identification in Fixed Effects Models, with Application to Head Start* (No. w26174; p. w26174). National Bureau of Economic Research. https://doi.org/10.3386/w26174

Oppong Asante, K., & Andoh-Arthur, J. (2015). Prevalence and determinants of depressive symptoms among university students in Ghana. *Journal of Affective Disorders*, *171*, 161–166. https://doi.org/10.1016/j.jad.2014.09.025

Pine, D. S., Cohen, E., Cohen, P., & Brook, J. (1999). Adolescent Depressive Symptoms as Predictors of Adult Depression: Moodiness or Mood Disorder? *American Journal of Psychiatry*, *156*(1), 133–135. https://doi.org/10.1176/ajp.156.1.133

Prinstein, M. J. (2007). Moderators of Peer Contagion: A Longitudinal Examination of Depression Socialization Between Adolescents and Their Best Friends. *Journal of Clinical Child & Adolescent Psychology*, *36*(2), 159–170. https://doi.org/10.1080/15374410701274934

Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement*, *1*(3), 385–401. https://doi.org/10.1177/014662167700100306

Ringeisen, H., Langer Ellison, M., Ryder-Burge, A., Biebel, K., Alikhan, S., & Jones, E. (2017). Supported education for individuals with psychiatric disabilities: State of the practice and policy implications. *Psychiatric Rehabilitation Journal*, *40*(2), 197–206. https://doi.org/10.1037/prj0000233

Ringel, J. S., Ellickson, P. L., & Collins, R. L. (2006). The Relationship between High School Marijuana Use and Annual Earnings among Young Adult Males. *Contemporary Economic Policy*, *24*(1), 52–63. eoh.

Roodman, D. (2011). Fitting Fully Observed Recursive Mixed-process Models with cmp. *The Stata Journal: Promoting Communications on Statistics and Stata*, *11*(2), 159–206. https://doi.org/10.1177/1536867X1101100202

Rosenqvist, E. (2018). Two Functions of Peer Influence on Upper-secondary Education Application Behavior. *Sociology of Education*, *91*(1), 72–89. https://doi.org/10.1177/0038040717746113

Santos, M., & Ribeiro, A. E. (2011). Poverty and mental illness: An unbreakable cycle? *European Psychiatry*, *26*(S2), 577.

Shrum, W., Cheek, N. H., & Hunter, S. MacD. (1988). Friendship in School: Gender and Racial Homophily. *Sociology of Education*, *61*(4), 227. https://doi.org/10.2307/2112441

Smith, J. P., & Smith, G. C. (2010). Long-term economic costs of psychological problems during childhood. *Social Science & Medicine*, *71*(1), 110–115. https://doi.org/10.1016/j.socscimed.2010.02.046

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, *87*(3), 355. https://doi.org/10.2307/1882010

Substance Abuse and Mental Health Services Administration. (2020). *Key substance use and mental health indicators in the United States: Results from the 2019 National Survey on Drug Use and Health* (HHS Publication No. PEP20-07-01-001; NSDUH Series H-55). Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administratio. https://www.samhsa.gov/data/

Thompson, O. (2021). *Human Capital and Black-White Earnings Gaps, 1996–2017*. W.E. Upjohn Institute. https://doi.org/10.17848/wp21-343

Tylee, A. (2005). The Importance of Somatic Symptoms in Depression in Primary Care. *Prim Care Companion J Clin Psychiatry*, 10.

Value Options Depression Screening. (n.d.). Value Options. http://www.valueoptions.com/providers/Education_Center/Provider_Tools/Depression_Screening.pdf

Van de Velde, S., Bracke, P., & Levecque, K. (2010). Gender differences in depression in 23 European countries. Cross-national variation in the gender gap in depression. *Social Science & Medicine*, *71*(2), 305–313. https://doi.org/10.1016/j.socscimed.2010.03.035

VanderWeele, T. J. (2016). Mediation Analysis: A Practitioner's Guide. *Annual Review of Public Health*, *37*(1), 17–32. https://doi.org/10.1146/annurev-publhealth-032315-021402

VanderWeele, T., & Vansteelandt, S. (2014). Mediation Analysis with Multiple Mediators. *Epidemiologic Methods*, *2*(1). https://doi.org/10.1515/em-2012-0010

Weersing, V. R., Jeffreys, M., Do, M.-C. T., Schwartz, K. T. G., & Bolano, C. (2017). Evidence Base Update of Psychosocial Treatments for Child and Adolescent Depression. *Journal of Clinical Child & Adolescent Psychology*, *46*(1), 11–43. https://doi.org/10.1080/15374416.2016.1220310

Weissman, M. M., Sholomskas, D., Pottenger, M., Prusoff, B. A., & Locke, B. Z. (1977). Assessing depressive symptoms in five psychiatric populations: A validation study. *American Journal of Epidemiology*, *106*(3). https://doi.org/10.1093/oxfordjournals.aje.a112455

Weller, N. F., Cooper, S. P., Tortolero, S. R., Kelder, S. H., & Hassan, S. (2004). School-Year Employment Among South Texas Middle School Students: Effects on Academic, Social, Mental Health, and Physical Functioning. *Journal of Children's Health*, *2*(1), 87–106. fgh.

Wickrama, K. A. S., Noh, S., & Bryant, C. M. (2005). Racial differences in adolescent distress: Differential effects of the family and community for blacks and whites. *Journal of Community Psychology*, *33*(3), 261–282. https://doi.org/10.1002/jcop.20053