TRAJECTORIES OF ADOLESCENT DEPRESSION AND GENDER/RACIAL DISPARITY

Lin Wang

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Sociology.

Chapel Hill
2005

Approved by
Advisor: Kenneth Bollen
Reader: Peggy Thoits
Reader: Michael Shanahan

ABSTRACT

LIN WANG: Trajectories of Adolescent Depression and Gender/Racial Disparity
(Under the direction of Kenneth Bollen)

This paper attempts to understand the trajectories of adolescent depression and gender/racial disparities. I conduct data analysis using data from the National Longitudinal Survey of Adolescent Health. I use growth curve models that trace individual depression trajectories over different ages employing a measurement that is comparable across gender/race. I find that depression increases at early adolescence and decreases at late adolescence. This pattern is consistently found among both males and females. Females are found to be significantly disadvantaged in depression when compared to males for all races. Racial disparity, unclear in previous literature, is clarified in this study. Minority groups do encounter greater levels of depression, ad compared to their white counterparts, female minorities face greater disadvantages compared to male minorities. However, in spite of gender and racial gaps, after the respondents enter their 20s, the disparities are greatly reduced, as the depression overall has decreased for all groups.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

LIST OF ABBREVIATIONS

BIC: Bayesian Information Criterion

CES-D: Center for Epidemiological Studies Depression Scale

DF: Degrees of Freedom

LCM: Latent Curve Model

LRT: Log-likelihood Ratio Test

MAR: Missing at Random

MGA: Multiple Group Analysis

ML: Maximum Likelihood

SEM: Structural Equation Model

Background

Adolescence is an important developmental stage for understanding the nature, course, and treatment of depression (Petersen, Compas, Brook-Gunn, Stemmler, Ey, and Grant 1993). Elevated rates of depression in adolescence relative to adulthood have been reported in many studies (e.g., Allgood-Merten, Lewinsohn, and Hops 1990; Larsson and Melin 1990). Although depressive symptoms are commonly observed phenomena in adolescence, such symptoms are not uniformly distributed across the population. Two issues were raised over the years in sociological studies of depression. First, depression level changes over time. Adolescents at certain ages are more prone to depression than others, therefore it should be examined in a developmental perspective. Cross-sectional data analysis can no longer fulfill such a task. A longitudinal study is necessary. Second, certain gender/racial groups are thought to be more likely to experience depressive symptoms. To compare the group differences, we need to employ multiple group analysis employing a measurement that is comparable across groups. Considering the dynamic nature of such disparities, they should also be studied longitudinally.

Starting in the eighties, the number of sociological studies of adolescent depression has been steadily increasing. However, we still do not have an accurate description of the trajectories of adolescent depression of different gender/racial groups over a relatively long period of time. This is due to limitations such as the availability of national representative longitudinal panel data and a lack of advanced statistical techniques to enhance measurement, model panel data, and deal with missing values. This is reflected in the conflicting findings in

the literature regarding the timing of change and existence of disparities. This study is aimed at bridging such a gap. To do this, I first briefly review the literature of the conceptualization of depression and its classification and measurement, followed by an overview of its trajectories and distribution in the population. Then I describe their possible causes. Afterwards I develop my hypotheses regarding the trajectories and disparities of depression and statistically test these hypotheses in the following sections. The statistical analysis uses data from a national representative sample and is done with growth curve modeling, which traces individual depression trajectories over different ages employing a measure that is comparable across gender/race.

What Is Depression and How Is It Measured

The term depression covers a spectrum of mood disorders that can range from being mild and transitory to a persistent state of incapacitation. One end of the spectrum can be difficult to distinguish from normal reaction and at the other end there is an overlap into severe psychotic disorders. Three approaches to the assessment and classification of adolescent psychopathology have been reflected in the literature on adolescent depression: (a) depressed mood, (b) depressive syndromes, and (c) clinical depression (Peterson et al. 1993). Each approach reflects different assumptions about the nature of psychopathology, serves different purposes, and reflects a different severity of depressive phenomena. Considerable confusion has arisen in the literature from the use of the term depression to refer to all three levels of depressive phenomena (Petersen et al. 1993).

This study will focus on depressed mood, a less intense but relatively common phenomenon in the general population. Research on depressed mood has been concerned with depression as a symptom and refers to the presence of sadness, unhappiness, or blue

feelings for an unspecified period of time (Petersen et al. 1993). No assumptions are made about the presence or absence of other symptoms. The study of depressed mood is especially important for its connection with greater level of depression in later life. Some investigators of depressed mood have identified a threshold above which a score is thought to be predictive of clinical depression. In the paper, unless otherwise specified, depression refers only to depressed mood.

Depression is typically measured through adolescents' self-reported emotions, either through measures specifically concerned with mood or though items included in checklists of depressive symptoms. The Center for Epidemiological Studies Depression Scale (CES-D) is one of these key measurement instruments. Developed in 1976 for use in general adult population (aged 18 or older), the standard CES-D is a 20-item self-report scale that measures depression (Radloff 1977).

Since its introduction, the CES-D, in either short (i.e., 4-10 questions) or long forms (i.e., 20-questions), has been used to assess depression risks in several populations (e.g., adolescent, elderly, ethnic, and clinical populations) for whom it was not originally designed. In order to be used on different populations or conduct group comparison tests, invariance of both form and parameters in the measurement model has to be established (Bollen, 1989).

A recent development on evaluating CES-D on adolescents by Perreira, Deeb-Sossa, Harris, and Bollen (2005) categorizes the long form (19 items) CES-D items used in the National Longitudinal Study of Adolescent Health into three dimensions: effects, causes, and outcomes. Effect indicators are *determined by* the latent variable or factor that they are presumed to be measuring and should be positively correlated. In other words, we expect that a change in the latent depression variable leads to a change in responses to these items.

Causal indicators, in contrast, are indicators that *determine* the latent variable and may be either positively or negatively correlated with each other. Outcomes are not indicators at all but are *consequences* of the latent variable (Bollen and Lennox 1991).

Techniques for the evaluation of multi-item scales (e.g., reliability estimates) rest on the assumption that all scale items are effect indicators. The multiple-factor model they tested (with all CES-D measurement items), which acknowledges the different dimensions of depression such as effects, causes, and outcomes, seem to behave well for Whites, while not so well for other groups. A possible reason for this is the long form CES-D measurement consists of depressive symptoms that are causes or outcomes. These symptoms (such as appetite) might not be related to depressed mood. This can bring in larger measurement error, which is probably why the multiple dimension measurement is not invariant across groups.

Alternatively, they tested a unidimensional model consisting of only effects (depressed, life not worth living, happy, sad, and blues). These indicators are direct measurements of mood related variables. They conclude that this unidimensional model measures adolescent depression of different racial/ethnical groups more consistently. This is by far the most sophisticated evaluation and solution for CES-D on accessing adolescent depression for different racial/gender groups.

Trajectories and Disparities of Adolescent Depression

First of all, it is clear that children and adolescents do experience the entire range of stressful life events, except those linked to older age or particular adult status (e.g., marriage). Children can report accurately on their own depressed mood and symptoms (Kazdin 1994) and can recognize readily various different emotions (positive/negative valence and self/other perspectives) after age 9 (Harter 1999).

Adolescence is a life-stage of transition from being dependent (childhood) to independent (young adulthood), where they expand their roles into more complex social environments that expose them to a widening array of stressors and life-shaping choices. Physical and psychological maturation, as well as alterations in social roles and environments, often lead to increased stress levels and depression. The majority of adolescents of both genders successfully negotiate this developmental period without any major psychological or emotional disorders, developing a positive personal identity and adaptive relationships with both peers and their families (Powers, Hauser, and Kilner 1989). However, many adolescents do experience various degrees of depressive symptoms. The level of depressed phenomena is linked to the varying intensity of such psychological stresses they experience at different ages of adolescence.

Longitudinal studies concerning the development of depression during adolescence have not been widely conducted. Here I summarize the findings from some of the existing research. Radloff (1991) found dramatic increases in depressed moods between the ages of 13 and 15 years, a peak at approximately 17-18 years, and subsequent decline to adult levels. Hankin et al. (1998) also showed that both boys and girls became increasingly more depressed between the age of 15 and 18 (from 3% to 17%). Holsen, Kraft, and Vitterso (2000) found among boys there were no substantial changes in depressed mood mean scores, while among girls there was a slight tendency of a curvilinear trend, with a peak level reached in mid-adolescence. They also found there was a tendency for adolescents to retain their relative level in depressed mood, most pronounced for a period of 4 years, from age 15 to age 19 years. Overall, elevated rates of depressed mood in adolescence relative to adulthood have been repeatedly reported in previous studies (e.g., Allgood-Merten et al. 1990; Larsson and

Melin 1990). These studies suggest the depression trajectory follows a pattern of rising and then falling during adolescence. Thus, middle adolescence may represent a critical time for increased vulnerability to depressive mood and disorders. I therefore develop my first hypothesis regarding how depression changes during adolescence as following.

Hypothesis 1: The trajectories for all races and genders follow a general pattern of increase in early to mid teenage and then decrease after reaching a peak in late teenage.

However, many studies have shown that this elevated level of depression during adolescence is not equally distributed among the population. The literature suggested mainly two type of disparity: gender disparity (that women are disadvantaged) and racial/ethnical disparity (that minorities experience more depression). The gender difference is very consistent in the literature, while the findings about racial disparity have long been mixed. A review of the literature of group comparisons on depression during adolescence reveals that established measurement theories for group comparison studies of depression have not been incorporated into most of these comparison studies, partly due to the diversity of the assessment methods of depression. So the findings in some of these studies might not be valid due to being unable to employ consistent measurements for different gender or racial groups.

Furthermore, the nature and causes of these disparities can be very different. An understanding of the processes of depression provides us with insights on how the disparities among groups are formed. In general, depression disparities come from two major sources. (1) Exposure to greater levels of stressors; researchers believe that disadvantaged locations in the social structure, as well as certain cultural values and biological factors are the major sources of differential levels of stress outcomes (Gove and Tudor 1973; Radloff 1975;

Hankin and Abramson 1999; Ge, Conger, and Elder 2001). (2) Differential vulnerability to the negative consequences of stress; this differential vulnerability to stress can be due to certain coping styles which are believed to be more likely to lead to unfavorable results of stress (Nolen-Hoeksema 1987; Petersen et al. 1993) and/or their disadvantaged social position and to lack of social and personal resources for combating the deleterious effect of stress on health (Schoenbach, Kaplan, Fredman, and Kleinbaum 1986).

Disparity by Gender

That females face greater challenge from depression has been consistently found in the literature. However, most of them did not establish statistical measurement invariance among groups. The invariance of measurement between males and females has been widely assumed. At best Cronbach's alpha[1] is employed as an indirect way to gauge the comparability of the CES-D or Children's Depression Inventory (CDI). Yet this is an inadequate approach that does not directly compare the equality of the measurement parameters.

According to the literature, gender difference during the adolescence is eventually caused by the difference in the rate of elevation of depression during early adolescence. Starting in early adolescence, more girls than boys begin to become depressed, and this gender difference in depression persists throughout adulthood (Nolen-Hoeksema 1990; Nolen-Hoeksema and Girgus 1994; Holsen et al. 2000) across many countries and cultures (Weissman et al., 1990). More specifically, after the age of 13, the rate of depression rises dramatically for girls while boys' symptoms and mood remained constant (Cole, Martin,

[1] In order to accurately measure the latent constructs, interrelated items may be summed to obtain an overall score for each participant. Cronbach's coefficient alpha estimates the reliability of this type of scale by determining the internal consistency of the test or the average correlation of items within the test (Cronbach 1951). However, Cronbach's alpha is limited compared to confirmatory factor analysis in accessing measurements. Also, Listwise deletion of observations with missing values is necessary to correctly calculate Cronbach's alpha.

Peeke, Seroczynski, and Fier 1999; Ge, Lorenz, Conger, Elder, and Simons 1994; Petersen, Sarigiani, and Kennedy 1991). Hankin et al. (1998) quantified these changes and found this increase for girls was from 4% to 23%, much greater than for boys (from 1% to 11%).

Hormonal changes are considered to be acting behind the gender differences in increased depression level after puberty. Research has also shown social and psychological factors play important roles in causing gender differences. Possible explanations for the disadvantages for girls include dissatisfaction with body-image (Hankin and Abramson 1999), greater negative events in the family (Compas, Davis, and Forsythe 1985), and more susceptibility to network or peer-related disruption (Hankin and Abramson 2001). Having more challenges (Petersen et al. 1991) but less effective coping styles, such as ruminating on their depressed mood and therefore amplifying it (Nolen-Hoeksema 1987; Butler and Nolen-Hoeksema 1994), may increase the likelihood of depression among girls (Petersen et al. 1993).

Another important issue is the possibility that in reality, males and females do not differ in the prevalence rates of depression. However, studies have shown that the greater preponderance of depressed adolescent girls and adult women relative to boys and men appears not to be explained by factors such as response bias on questionnaires, greater openness to acknowledging psychological difficulties, and other attributes apart from actual depression experienced by the individual (e.g., Nolen-Hoeksema 1987; Weissman and Klerman 1977; Gove and Tudor 1973; Nolen-Hoeksema 1990; Nolen-Hoeksema, Girgus, and Seligman 1991). These examinations have concluded that the gender differences appear to be a true difference in the experience of depression. However, these claims still need further statistical examination and support.

I then develop my second hypothesis regarding the unequal distribution of depression by gender and race. This hypothesis can be further broken down to two components. Each can be tested separately. For gender, the literature has been clear and consistent. Therefore the hypothesis can be easily formed.

Hypothesis 2: there is significant gender difference in depression trajectories and females are more likely to be depressed during adolescence.

Disparity by Race/Ethnicity

Other than gender, rates of depression may be higher among adolescents in some ethnic groups or their subgroups (Rushton, Forcier, and Schectman 2002). However, very different from gender, some scholars believe the study of racial disparity faces a greater challenge from the comparability of the measurement due to differential language capabilities and various cultural backgrounds of respondents of different races. Assessment of the comparability of the measurements appears slightly more frequently in the studies of racial disparity compared to gender disparity. Cronbach's alpha is still the most popular assessments in the studies using sum score type of measurement, despite its inability to directly compare measurement parameters. Measurement models using confirmatory factor analysis (CFA) are occasionally employed (e.g., Whitley and Gridley 1993; Crockett, Randall, Shen, Russell, and Driscoll 2005). However, few of them were supported with invariance tests. Studies dedicated to the assessment of the measurement comparability are available. But as I mentioned earlier, the findings of these studies on measurements are not incorporated in the substantive study of depression disparities.

Overall, racial/ethnic disparities on depressive symptoms are not consistent in the literature. On one hand, available work leads us to suspect that minority status is associated

with poorer mental health. Siegel, Aneshensel, and Taub (1998) found that compared to Whites, African Americans or Asian Americans, Latinos adolescents (age 12-17 years) reported more symptoms of depressed mood, a finding that was independent of socioeconomic status. Similar results regarding Hispanics were also found by Gore and Aseltine (2003), Iwata, Turner, and Lloyd (2002), Twenge and Nolen-Hoeksema (2002). In a review of community studies of adolescent depression, Fleming and Offord (1990) reported that in two of five studies where race was examined, African-American adolescents had higher rates of depression and depressed mood than Whites.

Many other studies also find that African-American adolescents (Garrison, Jackson, Marsteller, McKeown, and Addy 1990) and young adults (Gore and Aseltine 2003) have higher depression levels than either Whites or Asian Americans. Greenberger and Chen (1996) found that ethnic differences in depressed mood, not evident in the early adolescent sample, emerged in the college sample, with Asian Americans reporting more symptoms compared to European Americans. Other studies examining certain Asian ethnic groups also found that these minority groups experience greater levels of depression compared to Whites (e.g., Lam, Pepper, and Ryabchenko 2004; Greenberger, Chen, Tally, and Dong 2000).

Racial minorities are believed to be exposed to more stressors given their disadvantaged position in society, and being more vulnerable to its negative effects on mental health. Some research has examined the question of stress exposure by race in adolescents. Youth in low SES neighborhoods perceive greater ambient hazards such as crime, violence, drug use, and graffiti than those in high SES neighborhoods (Aneshensel and Sucoff 1996; Radziszewska, Richardson, Dent, and Flay 1996). The perception of the neighborhood as dangerous, in turn, influences the mental health of the adolescents: the more threatening the

neighborhood, the more common the symptoms of depression, anxiety, oppositional defiant disorder, and conduct disorder.

Gore and Aseltine (2003) showed that the heightened depressed mood among Hispanic and Black high school students are the result of increasingly disadvantaged pathways into adulthood, characterized by poorer prospects for educational advancement and more problematic relationships. Wight, Sepúlveda, and Aneshensel (2004) also argue that the risk of depressive symptoms may be especially pronounced among economically disadvantaged ethnic minority adolescents.

On the other hand, Nettles and Pleck (in press) reviewed several studies and concluded that although African-American youth are at greater risk for many negative behavioral and health outcomes, rates of depressive symptoms in African-American samples are typically lower than in Caucasian youth. In a study of one of the largest multiethnic samples of adolescents, Dornbush, Mont-Reynand, Ritter, Chen, and Steinberg (1991) reported that Caucasian and Asian-American youth reported more depressive symptoms than African-American or Hispanic-American adolescents, even after controlling for level of stressful life events. Given other findings (e.g., Fitzpatrick, Fujii, Shragg, Rice, Morgan, and Felice 1990), it is probably wise to note that a literature review on this subject by Hammen (1991) did not find evidence for Black-White differences in depression among adults either. Overall, the incidence of depression in various ethnic groups and social classes is still understudied. At this point, no solid conclusion on racial/ethnic disparity of depression can be drawn.

Last, it is important to examine how ethnicity may play a role in the development of gender differences in depression. Two studies (Schraedley, Gotlib, and Hayward 1999; Siegel et al. 1998) using CDI found that Hispanic adolescents reported the greatest level of

depressed mood compared with Caucasians or African Americans. Neither study found significant interactions between ethnicity and gender in predicting depressed mood.

As we have seen, the literature on racial disparities consists of conflicting results. The statistical evidences supporting the existence of racial disparity and those against it are both found in the literature. In terms of theoretical development, there has been stronger evidence regarding the existence of racial disparity and this evidence is more consistent with theories regarding other types of disparities such as gender. Those who found there are no significant racial differences in depression have not been able to provide convincing arguments why there are no differences in outcomes while minorities are perceived to normally have greater exposure to stress. In addition, considering that recent studies with more attention on comparability of the measurements are more likely finding racial differences, I also hypothesize my study towards that direction.

Hypothesis 3: there are significant racial differences in trajectories for each gender and minorities experience a greater level of depression.

However, given the past mixed findings, it might not be surprising to find out racial disparities cannot be found significant. Or, although this disparity exists, it might be only on a very small scale.

Data

The National Longitudinal Study of Adolescent Health (Add Health) is a nationally representative study of adolescents who were in grades 7 through 12 in the United States in 1995. The study was designed to examine the causes of adolescent health and health behavior, focusing on the multiple contexts in which young people live. Some minority ethnic groups were sampled in proportion to their size within the U.S. population; smaller

ethnic groups, including Chinese, Puerto Rican, and Cuban youth, were over-sampled. This aspect of Add Health makes it possible to conduct analysis by race-ethnic group (Harris 1999).

Add Health involves three waves of data collection and several data collection components. I use data from all three waves of in-home interviews. The three waves of Add Health data contain 20774 cases in total. The number of cases in each wave and the duration of data collection are shown in Table 1. The survey time and sample size information of each wave is shown in Appendix A. The age distributions at each wave are shown in Table 1.

The reduction of the number of cases in Wave II[2] and Wave III[3] is the result of both change of sample and non-response. There has not been empirical or theoretical evidence that depression is directly liked to the probability of being missing (theoretically, this is impossible to prove because we don't have the values of depression when it is missing). However, there are possible indirect associations between the two. For example, one of the popular reasons for non-response is allocation. Allocation has been found in the literature to be connected with social-economical status, which is linked to social support/resources and then depression. However, these associations identified in the literature are usually minimal.

[2] The sample for the Wave II in-home interview comprised the respondents to the Wave I in-home interview, excluding respondents (4039) who were in the 12th grade at Wave I and who were not part of the genetic sample, and also excluding respondents who were in only the Wave I disabled sample. Of those selected for Wave II, the non-response was only 12%. The top three reasons for non-response are "final unlocatable" (583), "final unavailable" (629), and "final refusal by adolescent" (391).

[3] The sample for the Wave III in-home interview consisted of Wave I in-home respondents, plus 45 Wave II only genetic respondents (their data in wave II never been released), minus 687 Wave I cases without a weight and without a genetic sample flag. Of those selected for Wave III, the non-response was 24%. At wave III, the top three reasons for disposition are "respondent not located" (2125), "refusal by respondent" (1012), and "respondent unavailable after repeated attempts" (658). The frequencies for other reasons are significantly smaller (all less than 200 and most of them under 50).

I assume the data points are missing at random (MAR)[4] (Little and Rubin 1987). That is, the probability of missing on the dependent variable (depression) is not related to depression, after controlled for independent variables such as age, gender, and race. Among all the cases in the sample, 9 cases are deleted due to lack of information on race; 334 Indians are deleted (20431 left) because of their unique cultural background and their different response pattern to CES-D (Chapleski, Lamphere, Kaczynski, Lichtenberg, and Dwyer 1997); 1667 first generation immigrants are deleted (18764 left as second+ generation) due to their diverse and unique cultural backgrounds and sometimes even language capabilities[5] (Harker 2001); 8 cases are deleted due to lack of information on depression at any wave. There is no case that contains missing values for only a particular item of the three items. The sample I eventually use for my data analysis contains 18756 cases. Since the sample does not include Native Americans and first generation immigrants, this study is only representative of native born Whites, Blacks, Asians, and Hispanics.

Measurements

For this analysis, we evaluate the measurement of the CES-D across 8 groups constructed by interacting each adolescent's primary race-ethnic identification with his/her gender. Race/ethnicity was defined using the respondent's self-reported ethnic identity (i.e., Hispanic/Latino) and race (i.e., white, black/African American, Asian/Pacific Islander). The respondents could choose multiple racial categories to describe their racial background and 1,038 out of 20,745 respondents did so.

[4] Gender and race do not appear to affect attrition in this sample. Age obviously does (such as in Wave II). However, it is included in the model as a predictor. Therefore the MAR assumption is still holds. The parameter estimates should not be affected by such selection.

[5] These background differences can reduce the consistency of the measurement. This can further cause unnecessarily greater deviation/complication. Therefore, Native Americans and First generation immigrants are excluded from the sample to reduce the complexity of this study.

Because of the employment of multiple group analysis in this study, I use mutually exclusive race categories[6]. A single race is assigned to those reported multiple racial backgrounds. Hispanic is treated as a racial group, although theoretically they could be subdivided into other groups. The reason for defining Hispanic as an independent group is that they are unique in many of the characteristics causing depression and/or its disparity such as culture or SES suggested by the literature. The generic rule of assigning race is: if the respondent reported single race, he/she will be coded as is; if the respondent reported more than one race, only one race will be selected from the races the respondent reported in the following order: Hispanic, Black, Asian, Indian, and White. This coding method is employed by the add health data manager as a way to obtain exclusive race category. This method is also a popular measurement method among population related studies. Therefore, the findings of this study are more likely to be comparable with other studies.

Add Health uses the CES-D to measure adolescents' depressive symptoms. However, CES-D, although widely used, is not a perfect measure. It was designed originally for adults and White European descendants. When it comes to minorities and adolescents, it might not function as well as it does with its original targets. Minority status could add complications to the measurement of depression and endanger its comparability (Perreira et al. 2005).

Demonstrated in Riddle's 2002 study, the items of the CES-D measurement can be categorized up to four dimensions: depressed affect, positive affect, somatic complains, and interpersonal relations. Various studies have shown that including all four dimensions, although providing a full coverage of the depressive symptoms, adds complications and

[6] The multiple racial backgrounds will cause certain level of measurement error for race. However, considering only 4 percent of the respondents reported mixed background, I assume the validity of the study is not jeopardized. The ordering of recoding race is also effective in reducing the impact the measurement error on race.

decreases the consistency of the measurement when CES-D is used for multi-ethnic study. Researchers started to use short forms of CES-D in order to obtain consistency when doing group comparison analysis. The idea is that the shortened scale included items more conceptually valid for minority groups (Liang, Tran, Krause, and Markides 1989; Chapleski, Lamphere, Kaczynski, Lichtenberg, and Dwyer 1997; Harker 2001; Iwata, Turner, and Lloyd 2002).

In a recent study of depression Perreira et al. (2005) identified a five-indicator measurement of depression that is comparable across groups using the first-wave data of Add Health. The multiple group analysis is conducted across 12 ethnocultural groups constructed by interacting each adolescent's primary race-ethnic identification with his/her immigrant generation. They included 4 ethnicity-race groups: (1) Hispanic of any race, (2) Non-Hispanic Asian, (3) Non-Hispanic Black, and (4) Non-Hispanic White, and 3 immigrant generation: foreign-born to foreign-born parents (1^{st} generation), U.S.-born to foreign-born parents (2^{nd} generation) and U.S.-born adolescents to U.S-born parents (3^{rd+}generation or native) (Harris 1999).

The technique they used is confirmatory factor analysis (CFA). Those five indicators are feeling depressed, feeling sad, feeling blue, being happy, and feeling life is not worth living during the past week, which are all effect indicators. That is, they are indicators that are influenced by the latent depression variable. These effect indicators are direct measures of mood-related variables, which are highly correlated to depression. I further reduced the five indicator measurement to 3 indicators (depressed, sad, and blues) to form my measurement of depressed mood.

The two excluded indicators (happy and life not worth living) are actually not as highly correlated with the rest. Each item is measured on a 0-3 scale, from never or rarely (0) to most of the time or all of the time (3). (The details of the questions are shown in Appendix B.) I then used a simple sum of the three indicators to form an index of depression. Therefore the measurement for each of the three waves has a scale of 0-9, and is treated as approximating a continuous outcome. The measurement has a consistent form for all three waves. A rough assessment of the reliability of the measurement is done by computing the Pearson correlations among the items and their sum at each wave. The tests produce correlations between the sum and each individual item and they are well above 0.8 at each wave. Based on this evidence, I will make the simplifying assumption that the sum of the items forms an index of depression mood that can be studied across groups while recognizing that future research should examine individual item invariance[7].

Analytic Strategy

The measurement of depression is based on the feelings that the respondents had in the past week. The measurement of depression based on such a short period of time will very likely to be disrupted by random factors, which will increase the variance of the trajectory model. I nonetheless assume that although the measurement might be disrupted by short-term fluctuations, the collective information provided by the large sample size can average out the disruption and reflect the overall depression pattern. Short-term patterns are largely determined by the long-term characteristics of mental health.

[7] According to Bollen (1989), in order for a measurement to be comparable across groups, the invariance in both model form and model parameter needs to be established. I conducted invariance test using a measurement model in SEM to access the measurements. Constraining the factor loadings and intercepts to be equal across racial and gender groups did not reduce the model fit much, considering the large sample size. That means this measurement model is an invariant measurement for these groups. Therefore it can be used to compare groups.

In order to capture the trajectory of the mental health development during adolescence, I employed the latent curve model, or LCM, based on structural equation models (SEM). A two-level data structure arises from the repeated observations of a set of individuals over time (Curran 2000). LCM considers the observed repeated measures over time to be fallible indicators of an unobserved true growth trajectory. Using the SEM framework, we can use the observed repeated measures to define one or more underlying latent growth factors (Willet and Sayer 1994; Curran 2000).

Usually two latent factors are defined to represent the intercept and slope of the depression growth trajectory. The intercept factor represents the trajectory of depression that does not change. Therefore the factor loadings relating the observed depression measures to the intercept factor are all fixed to equal one. The slope factor describes the change of depression over time. Since I expect to see a curvilinear growth curve, here the factor loadings relating the observed repeated measures to the slope factor are set to be freely estimated in order to capture the functional form of the growth trajectory over time.

A number of (at least three in order to identify the model) observed measures of depression can be used to estimate a single underlying true trajectory for each case. I estimate a series of parameters that best describes these observed measures and use them to define the individual's true growth trajectory of depression. The trajectory is characterized by a starting point (or the intercept) and a rate of change (or the slope). The individual level model is usually referred as the level 1 model. α_i and β_i are random variables and variation of each variable across individuals can be expressed as the mean plus a disturbance term for each individual. This is commonly called the level 2 model.

Model Equations for LCM:

Level 1 model:

$$y_{it} = \alpha_i + \beta_i\lambda_t + \varepsilon_{it} \quad (1.1)$$

Level 2 model:

$$\alpha_i = \mu_\alpha + \zeta_{\alpha i} \quad (1.2a)$$

$$\beta_i = \mu_\beta + \zeta_{\beta i} \quad (1.2b)$$

Combined model:

$$y_{it} = (\mu_\alpha + \lambda_t\mu_\beta) + (\zeta_{\alpha i} + \lambda_t\zeta_{\beta i} + \varepsilon_{it}) \quad (1.3)$$

In the level 1 model, y_{it} represents the depression measure for person i at time point t; α_i represents the true intercept of the growth trajectory for person i; β_i represents the true slope of the growth trajectory for person i; λ_t represents the value of time at time point t (e.g., 0, 1, or 2), and ε_{it} represents the time specific residual for person i at time t. In the level two model, μ_α represents the mean intercept of the growth trajectory and μ_β represents the mean slope of the growth trajectory; $\zeta_{\alpha i}$ represents the residual of intercept for person i and $\zeta_{\beta i}$ represents the residual of slope for person i. Equation 1.3 clarifies that the observed repeated measures of y can be expressed as an additive combination of a fixed component of growth (e.g., $\mu_\alpha + \lambda_t\mu_\beta$) and a random component of growth (e.g., $\zeta_{\alpha i} + \lambda_t\zeta_{\beta i} + \varepsilon_{it}$).

From the literature, we know that depression is related to age. So I will reorganize the data that are based on wave to use chronological age as the time metric. The model using age as the time metric is illustrated by Figure 1. But in this study, as shown in Appendix A and Table 1, the unequal spacing among the assessments, large amount of missing data in wave II and III, and the respondents' chronological age differences at each wave of assessment created a great deal of missing data.

Fortunately, this problem can be addressed given the recent developments of LCM methods. Bollen and Curran (forthcoming) present these developments on the choices of time metric based on the techniques of handling missing data. Among the different treatments they discussed, I choose the direct Maximum Likelihood (ML). In this approach, the likelihood function is computed for each case using only those variables that are available for that case. The total likelihood is the sum of the values of the likelihood for each case. Therefore the direct ML method makes use of all available information in the data with no need to impute values. One of the most important properties of direct ML is it can maintain the asymptotic properties of ML estimators under the more relaxed assumption of MAR (Bollen and Curran, forthcoming).

Direct ML, like many other methods, relies on the multivariate normality assumption. In this study since distribution of depression is highly skewed towards the right, this assumption is unlikely to hold. However, Bollen and Curran in their overview of literature on this issue concluded that although further examination needs to be done, early studies show that the direct ML is relatively robust.

Instead of organizing the data by wave (shown in Equation 2.1), I use chronological age to restructure the data (shown in Equation 2.2). Note in the restructured data, the unit of age ranged from one year to 5 years. This is done to reduce the number of variables and improve data coverage because there are not many cases having measures at earlier (age 11-12) or later (age 25-27) ages of the study period. I use the mean age of the span to represent age.

$$\begin{pmatrix} dep_{1,1} & dep_{1,2} & dep_{1,3} \\ dep_{2,1} & dep_{2,2} & dep_{2,3} \\ dep_{3,1} & dep_{3,2} & dep_{3,3} \\ \cdots & \cdots & \cdots \\ dep_{N,1} & dep_{N,2} & dep_{N,3} \end{pmatrix}$$

(2.1)

$$\begin{pmatrix} dep_{1,12} & dep_{1,14} & . & . & . & dep_{1,22} & . & . \\ . & dep_{2,14} & dep_{2,16} & . & . & . & dep_{2,24} & . \\ . & . & dep_{3,16} & dep_{3,18} & . & . & . & dep_{3,26} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & .... \\ . & . & dep_{N,16} & dep_{N,18} & . & . & . & dep_{N,26} \end{pmatrix}$$

(2.2)

Table 2 shows us the descriptive statistics of the constructed dependent variables (using age as time metric) that will be used in the LCMs. The "variable" column represents the longitudinal indicators for the trajectory. The "range of age" column represents the ages that are consolidated together to form that indicator variable. The "linear slope factor loadings" for that indicator variable are formed based on the mean age of the interval. N represents the number of non-missing values each variable contains. And the "Mean of depression" represents the means of the variables, or substantively, the mean level of depression at that age range. These means clearly suggest a curvilinear shape.

The data analysis is conducted in eight separated groups with every race/gender combination. Multiple group analysis, or MGA, will be employed in order to formally test my hypotheses regarding group disparities. MGA allows us to constrain certain model parameters to be equal across groups and formally test whether these constraints influence the model fit. If the model fit is not affected by such constraints, we can conclude that there

are no statistically significant group differences in those parameters. On the other hand, if the model fit is significantly worse when such constraints are imposed, we can conclude there are group differences and the magnitude depends on the actual values of the parameters. Since these models are nested, they can be compared through model fit indices such as log-likelihood ratio test (LRT) and the Bayesian Information Criterion (BIC) (Schwartz 1978; Raftery 1995). Of the two, BIC assigns a greater penalty to model complexity, and so has a greater tendency to pick parsimonious models. According to Raftery (1995), the model with smaller BIC is preferred and the strength of the preference depends on the magnitude of the difference.

Results

Testing Hypothesis 1 – Nonlinearity

I first explore the trajectories of the depression with whites, the majority group of this study. I run a series of models separately for White Males and White Females to examine the depression trajectories of each group. This allows me to select the models suitable to describe the depression trajectories and to use it for group comparison analysis. According to the literature, the depression level of adolescents increases during the early and mid-adolescence and decreases in late adolescence and young adulthood. Therefore, I expect a model that can capture this curvilinear trajectory will fit the data the best. I choose to start with a two factor (intercept and slope) LCM[8] with freed slope factor loadings to describe the depression trajectory. This model allows great flexibility on the shape of trajectory. By using different combination of constraints on this model, SEM allows me to test various hypotheses such as

[8] I also tried LCM with a quadratic term. But I encountered problems such as non-convergence and significant non-positive definite variance/covariance matrix. With only up to 3 repeated measures for each case, the data do not provide adequate information to fit a polynomial model. So I will not choose the polynomial model for this study. However, the polynomial model also showed a curvilinear pattern in the trajectories, similar to what I found in the freed loading model.

whether or not depression changes over time and whether it follows a linear change. By imposing constraints, the models are nested within each other. So I can use model fit comparison tests to select the model that fits the data best. The LCMs are estimated in the following configurations[9].

1. "Intercept Only Model" - random Intercept only model,
2. "Linear Model" - random Intercept and Slope model with linear slope factor loadings,
3. "Freed Loading Model" - random Intercept and Slope model with all but the first (fixed to 0) and the fifth (fixed to 1)[10] factor loadings unconstrained,
4. "Fixed Slope Model" - random intercept and fixed slope model (variance of slope set to zero) with all but the first and the fifth factor loadings unconstrained.

I compare the fit of these models using LRT and the BIC. The least restrictive model is the Freed Loading Model and the other models are nested within it. Table 3 shows the results of LRT between the Freed Loading Model and each other model. The LRT results between the Intercept Only Models and Freed Loading Models are highly significant for both White males and females. This suggests that depression changes over time. So a slope factor is necessary for the model to capture this change. The Freed Loading Model also fits the data much better than the Fixed Slope Models. This indicates that random slopes are necessary for the model to fit the data well. The significant LRT results also suggest that the Linear Models

[9] With some models, I encountered non-positive definite covariance matrices for the latent variables. But the problematic values (e.g.: a correlation that is greater than 1 or smaller than -1, or a variance that is negative) in these matrices are not significant, which means those values are not significantly different from zero. They appear to be a result of sampling fluctuations rather than a symptom of specification error.

[10] The fifth factor loading is the one with highest value for most racial/gender groups. The first factor loading is constrained to 0 and the fifth to 1. These two loadings define the unit of change for the data. This can ensure the change between the two scaled factor loadings is large enough to be estimated. Then this change is a reliable standard for estimating other factor loadings.

do not fit the data well, as I expected, indicating a curvilinear trajectory for depression. However, the large sample sizes of white males and females, both having over 5,000 cases, create huge statistical power.

With this statistical power, the models are more likely to produce significant statistical tests. I further examined the model fits with the BIC. The Freed Loading Model produces the smallest BIC and Sample-size Adjusted BIC compared to the other models; the values of BICs for the Freed Loading Model are 100 to 350 times smaller than that of the other models for both groups. The BIC indicates that constraining the model appears to cause real a decline model fit. Therefore, model fit comparison tests favor the Freed Loading Model for both White male and female groups. I then present the parameter estimates of the Freed Loading Model in Table 4.

As we can see in Table 4 for both White males and females, the Freed Loading Models produce significant parameter estimates, including most factor loadings (except a few measurement indicators at later ages[11]), factor means, and factor variances. The slope factor loadings from this model follow a curvilinear pattern. These results suggest that for both groups, depression changes following a curvilinear trajectory during adolescence. The significance of the parameter estimates explains why fixing certain model parameters to be 0 in the more constrained models significantly reduced the model fit. The results of these models strongly support Hypothesis 1 that for adolescents, the trajectories of depression are not linear and follow a reversed U-shape pattern. Therefore, the freed loading model is the best model for both White males and females. Since White females and males share the same model form, next, I can use the Freed Loading Model to conduct group comparison analysis.

[11] This suggests that the Factor loadings are not significantly different from 0. In other words, the depression levels at those years are not significantly different from that of the beginning of the study period.

Testing Hypothesis 2 – Gender Disparity

In order to test for Hypothesis 2a and 2b, I conducted a series of group comparison tests. I use white males and white females to examine gender disparities with the freed loading model chosen in the previous section. In order to compare groups, the MGA is conducted by constraining certain parameter(s) to be equal across groups. The models estimated include "Freed Loading Model", in which factor loadings are freely estimated except for the first and fifth ones, "Equal Factor Loadings Model", constraining the freely estimated factor loadings in the Freed Loading Model to be equal across groups, and "Equal Factor Means Model," further constraining factor means to be equal across groups.

Among the three models estimated, the Freed Loadings Model is the least constrained and all other two models are nested within the Freed Loadings Model. I conduct nested model comparison tests using model fit indices such as LRT and BIC. The results of the model comparison tests (LRTs), BICs, and Sample-size Adjusted BICs are shown in the first panel of Table 5. We can see from the test results that constraining the factor loadings to be equal across groups result in lost of model fit (39 with 7 degrees of freedom). This lost is moderate considering such a large sample size will generate huge statistical power and can produce a significant statistical test with a slight change on the model. So BIC would be a more suitable measure of model fit.

As we can see, the Equal Factor Loadings Model produced smaller BIC and Sample-size Adjusted BIC. This suggests that the more parsimonious model, which constrains the factor loadings to be equal across groups, has a better fit as indicated by the BIC than does having separate trajectory parameter estimates for each group. In other words, the depression trajectories for White males and females share a similar shape. The model which further

constrains the factor means to be equal across groups, on the other hand, resulted in serious lost of model fit (447 with 9 degrees of freedom) and has much larger BICs. Therefore constraining the factor means to be equal across groups is inappropriate and the average starting point and rate of change for White male and female respondents are statistically different. I present the parameter estimates of the "Equal Factor Loadings Model" in Table 6 to further examine the gender differences in terms of depression trajectory.

From Table 6, we can see that the slope factor loadings are significant except for the last two, which are not significantly different from zero. The model produced highly significant intercept factor means for both groups. The intercept means represent the average level of depression at the starting point, which is around age 13.5 in this study. Females have much higher starting point than males (1.286 VS 0.906). At the age of 13.5, white females are more depressed compared to white males. The significant latent intercept factor variance indicates there is variability in the initial level of depression within each group. We can see that males (0.359) are more homogeneous than females (0.949) in terms of their starting points of depression according to their variance value.

The estimates of the slope factor means are also significant. The slope factor means need to be explained together with the factor loadings. From table 6 we can see the factor loadings follow a pattern of rise and then fall back to the initial value[12]. The slope factor mean for White females is substantially higher than that of White males (0.581 VS 0.287). This means greater changes for White females than White males. During early to mid-adolescence, white females, although having higher initial level of depression already,

[12] The last two factor loadings are not significant. That means the factor loadings are not significantly different from zero.

experience a faster rate of increase in depression, as shown by their higher slope factor mean (0.581 VS 0.287). This results in an even greater gender gap.

Alternatively, during late adolescence and young adulthood, the depression level decreases for both groups. The gender gap is reduced because white females' depression level decreases faster, also due to the same slope factor mean. The variances of the slope factor are also significant for both males and females, indicating there are significant individual differences in the rate of change within each gender group.

I also conducted similar statistical analysis to test the between-gender differences of Blacks, Asians, and Hispanics. Both Blacks and Asians showed very similar results to that of Whites. The results are shown in Table 5 and Table 6. However, the analysis for Hispanics was not carried out sucessfully because the restricted models (equal factor loading and equal factor means) did not converge. This is probably due to the greater gender differences between Hispanics, which makes it impossible to find parameter estimates that can be reasonable for both males and females[13].

The model implied trajectories of both white males and females are illustrated in the first section of Figure 2. From the figure, we can see that the trajectories of depression for white males and white females follow very similar patterns but the one for females is much higher and curvier. These results strongly support Hypothesis 2a that gender differences are truly present in the developmental trajectories of mental health at this stage. Besides that, the gender gap increases during early to mid-adolescence, while start to decline after mid-

[13] This speculation is supported by the later study conducted separately for male racial groups and female groups. We can see from the section b and c of figure 2, Hispanics, being the highest among both males and females, have the greatest gender gap. Especially since the peak for Hispanic females (2.3) is considerably higher than the peak for males (1.5).

adolescence. After the adolescents enter their 20s, the gender difference is reduced to about the same level in early adolescence.

Testing Hypothesis 3 – Racial Disparity

I then conducted racial disparity tests separately for each gender. The racial disparity tests are done in the same way as the gender disparity tests. Three models (Freed Loading, Equal Factor Loadings, and Equal Factor Means) are estimated for model comparison tests.

The results of the group comparison tests for males of different races are shown in the second panel of Table 5. Constraining the factor loadings across groups to be equal reduces model fit. But this reduction (51 with 27 degrees of freedom) is also moderate considering the large sample size (over 9,000). The BIC, with a much smaller value for the Equal Factor Loadings Model, clearly favors the model constraining the factor loadings to be equal across groups. However, further constraining the factor means to be equal across groups did not affect the model fit as much; the Equal Factor Means Model produces slightly smaller (1.3) BIC values compared to the Equal Factor Loadings Model. Only when the BIC are adjusted for sample size did the model fit comparison favor the less constrained Equal Factor Loadings Model. In other words, the differences of factor means for each male racial group are hardly significant. This is probably because of the greater similarity among male groups.

The parameter estimates of the Equal Factor Loadings Model for males groups are shown in the first section of table 7. Their model implied growth trajectory is presented in the second section of Figure 2. As reflected in the model fit comparison tests, constraining factor means to be equal across groups does not reduce the model fit dramatically. The estimated factor means for male racial groups are very close. Intercept factor means range from 0.8 to 1.0 and slope factor means range from 0.36 to 0.47, indicating different groups have similar

starting point and rate of change. The magnitude of racial differences among males is substantively small. The variances are not significant except for White males, suggesting there is not much variation within each male racial group except for White males. Overall, males are much more homogenous regardless of race/ethnicity and their depression trajectories are relatively flat compared to females.

For female racial groups, the results are presented in similar fashion in the corresponding tables and figure as male groups. Constraining factor loadings to be equal across groups does not affect the model fit at all according the model fit comparison tests, indicating strong similarity in the shape of the trajectory. However, unlike male groups, further constraining the factor means to be equal across groups results in serious model fit reduction according to the BICs. Reflected in parameter estimates is that female groups have noticeable between-group differences on factor means; the estimated intercept factor means range from 1.35 for White females to 1.62 for Hispanic females and slope factor means range from 2 for Black female to lofty 3.8 for Asian females.

These differences appear to be significant according to the model fit comparison tests. And the magnitude of these differences is also substantial. Reflected in the third section of Figure 2, the trajectory for white females, being the lowest, is much lower than that of the Hispanic females. Asian and Hispanic females experienced a large amount of changes in their depression levels during the study period. We can also see from Table 7 that 3 out of 4 female groups (except Asian females) have a significant amount of within-group variation for starting values. And both White females and Hispanic group have significant within-group variation for rate of change. Overall, the depression trajectories for female groups are higher

and have quite a high amount of change over the life course. They also have much more within and between group variations compared to male groups.

Discussion

Although depressive symptoms are not devastating for most adolescents as indicated by this research, there are strong evidences indicating these symptoms are linked to more profound disorders later. As early as in the 1970s research has demonstrated that those youth who experienced psychological difficulties in adolescence are more likely to develop serious psychiatric disorders in adulthood (Rutter, Gradham, Chadwich, and Yule 1976; Weiner and DelGaudio 1976). As depressive symptoms become a reality for an increasing number of adolescents, it has become critical to examine the nature and course of these symptoms in this stage, as well as to identify the implications of this increase in depression for later life. The present study represents a further step towards understanding the longitudinal trajectory of depressive symptoms for each gender/race in a nationally representative sample of adolescents.

In terms of trajectory of depression during adolescence, the findings of this study are consistent with those in the previous literature. During early adolescence, depression quietly emerges and keeps increasing through mid-adolescence. At the age of 15 to 18, the respondents experienced the highest level of depression in both adolescence and young adulthood. Notice that for male groups, the peak is reached at the age of 18 and the peak is sharp or that it does not sustain (section b, Figure 2). While for female groups, the peak is reached earlier at 15 and the peak lasts for almost 4 years and has a plateau shape (section c, Figure 2). From late adolescence, the trajectory for depression appears to be steadily decreasing and stabilizes at a relatively low level during young adulthood.

Consistent with former studies, females are again found to be significantly disadvantaged in mental health when compared to males for all races and the disadvantage persists through out adolescence and into young adulthood. Longitudinally, the gender gap widens before mid-adolescence and is reduced after mid-adolescence. Racial disparity, unclear in the previous literature, is clarified in this study. Overall, minority groups do encounter greater level of depression. However, the extent of this disadvantage for minorities differs substantially for males and females.

On one hand, the magnitude of the racial differences for males is small and its significance is arguable. On the other hand, compared to their white counterparts, female minorities face significantly greater disadvantage. Knowing who is most susceptible to depression and when people are most susceptible can help us more efficiently direct limited resources for prevention and treatment to those who need them most. Understanding the higher level of depression for female minorities during mid-adolescence is substantively important in that they are caught in cross-fire and demand the most attention. However, in spite of gender and racial gaps, after the respondents enter their 20s, the disparities are greatly reduced as the depression overall start to decrease for all groups. From the results, we can see the gender/racial gap decreased to about the same level as it was in early adolescences.

The greatest strength of this study lies in its employing advanced statistical techniques in the treatment of longitudinal data. Add health data are excellent for studying adolescent depression, because of advantages such as its large sample size and high quality of data collection (e.g., low non-response rate). However, it was limited in that it had only three waves of data collection. However, this study overcomes this problem by conducting the

analysis by age instead of wave, so depression is measured at 9 different age-intervals instead of at only three waves.

This greatly expanded the number of indicators in the growth curve model. However, this is achieved at the cost of high percentage of missing values with larger number of missing patterns. In this case, more waves of data collection and relatively equal spacing between waves will help to reduce the number of missing values created by reorganizing the data using age as the time metric. A greater number of repeated measures for each respondent is also benifical in examining higher order growth models (such as a growth model with quadratic or even cubic terms) as an alternative to the freed loading model employed in this study to capture non-linear growth.

Another improvement of this study over previous ones is that it utilizes a better measurement of depression, which is recommended by the literature dedicated to assessment of measurement for group comparisons. The three indictor measurement employed by the present study focuses on the effect of depressed mood and reduces complications and disruptions introduced by including causes and outcomes. The short forms of measuremnt have been shown to have better construct validity for multiple-group comparison study.

Exclusion of first-generation immigrants and Native Americans further reduced the complixity of measurement such as cultural and language differences. Therefore we have greater confidence on reducing measurement error. However, the measurement error can be further reduced with a true measurement model based on confirmatory factor analysis (CFA). CFA was not employed in this study because it over-complicates the data analysis.

Historically adolescents' experiences with depression did not receive as much attention as adult depression. The study of adolescent depression tends to be less mature compared to

studies of adults due to limitations such as lack of longitudinal national representive data and a lack of developments in measurements that are more suitable for adolescents. With the utilization of the latest development of LCM, better measurement for group comparison, as well as many other improvements over the previous studies (such as longitudinal national representative sample) I provided by far the most precise description of the depression trajectories and disparities of different races/genders in adolescence. This study can greatly further our understanding of adolescent depression. With detailed description of depression in gender/race specific and developmental perspective, this study lays the groundwork for further examination of the risk factors and correlates of depressive symptoms in adolescence.

Table 1
Frequency Distribution of Age, by Wave (Counts and Percentages)

Race	Wave I (1995)		Wave II (1996)		Wave III (2001-2002)	
Age	Freq	Pct	Freq	Pct	Freq	Pct
11	10	0.05				
12	482	2.57	9	0.07		
13	2122	11.33	577	4.3		
14	2602	13.89	1800	13.41		
15	3357	17.92	2189	16.31		
16	3656	19.52	2797	20.84		
17	3461	18.48	2921	21.76		
18	2572	13.73	2172	16.18	121	0.87
19	402	2.15	796	5.93	1331	9.59
20	55	0.29	141	1.05	1888	13.61
21	10	0.05	19	0.14	2297	16.56
22					2586	18.64
23					2630	18.96
24					2185	15.75
25					713	5.14
26					102	0.74
27					17	0.12
Total	18729	100	13421	100	13872	100

Note:
The respondents of Add Health in Wave I contain a wide range of age, with majority of them in Middle school or High school.

Table 2
Descriptive Statistics of the Data Rearranged Using Age as the Time Metric

Variables	Range of Age	Appr .Medium of Age Range	Linear Factor Loadings[14]	Means of Depression	N	Min	Max	Std Dev
Dep 1	11-13	13.5	0	1.112	2610	0	9	1.505
Dep 2	14	14.5	1	1.289	4162	0	9	1.721
Dep 3	15	15.5	2	1.479	5271	0	9	1.827
Dep 4	16	16.5	3	1.587	6116	0	9	1.852
Dep 5	17	17.5	4	1.637	6075	0	9	1.889
Dep 6	18	18.5	5	1.633	4721	0	9	1.910
Dep 7	19-20	20	7.5	1.380	4440	0	9	1.810
Dep 8	21-22	22	9.5	1.223	4895	0	9	1.739
Dep 9	23-27	24.5	12	1.135	5636	0	6	1.673

[14] This is for Slope Factor.

Table 3
Nested Model Comparisons of Intercept Only, Linear, and Fixed Slope Models to Freed Loading Model for White Males and for White Females

	White Male (N=5340)					White Female (N=5448)				
Model	DF	LRT	P-value	Change of BIC	Change of BIC'	DF	LRT	P-value	Change of BIC	Change of BIC'
Intercept Only	10	264.588	0	+179	+211	10	407.496	0	+322	+354
Linear	7	208.364	0	+148	+170	7	162.998	0	+103	+125
Fixed Slope	2	127.212	0	+110	+117	2	115.777	0	+99	+105

Note:
LRT stands for Log Likelihood Ratio Test.
BIC' is Sample-Size Adjusted BIC

Table 4
Parameter Estimates for Freed Loading Models for White Males and for White Females

	Freed Loading Model			
	White/Male		White/Female	
	Estimates	P	Estimates	P
S BY				
DEP1	0		0	
DEP2	0.099	0.277	0.375	***
DEP3	0.481	***	0.895	***
DEP4	0.786	***	1.061	***
DEP5	1		1	
DEP6	0.941	***	0.723	***
DEP7	0.639	***	0.153	0.186
DEP8	0.199	0.03	-0.15	0.293
DEP9	0.101	0.259	-0.346	0.031
Means				
I	0.844	***	1.344	***
S	0.404	***	0.478	***
Variances				
I	0.488	***	0.914	***
S	1.031	***	0.612	0.006
I with S	-0.048	0.711	0.131	0.312
DF	23		23	
Loglikelihood	-22255		-26315	
Sample Size	5340		5448	

Note:
*** Stands for significant P-vlaue

Table 5
Nested Model Comparisons of MGA with Equal Factor Loading and MGA with Equal Factor Means to MGA with Freed Loading

Model	DF	LRT	P-value	Change of BIC	Change of BIC'
MGA for Whites, by Gender (N=10788)					
Equal Factor Loading Model	7	38.881	0	-26.109	-3.864
Equal Factor Means Model	9	447.072	0	+363.579	+392.18
MGA for Blacks, by Gender (N=4519)					
Equal Factor Loading Model	7	22.207	.002	-35.438	-13.195
Equal Factor Means Model	9	204.803	.000	+129.150	+157.748
MGA for Asians, by Gender (N=789)					
Equal Factor Loading Model	7	6.609	.471	-40.069	-17.841
Equal Factor Means Model	9	37.455	.000	-22.497	+6.093
MGA for Hispanics, by Gender (N=2660)					
Equal Factor Loading Model	7	-		-	-
Equal Factor Means Model	9	-		-	-
MGA for Males, by Race (N=9249)					
Equal Factor Loading Model	21	51	0	-140.752	-74.018
Equal Factor Means Model	27	104	0	-142.072	-56.271
MGA for Females, by Race (N=9507)					
Equal Factor Loading Model	21	30.005	0.092	-162.326	-95.591
Equal Factor Means Model	27	114.135	0	-133.126	-47.324

Note:
LRT stands for Log Likelihood Ratio Test.
BIC' is Sample-Size Adjusted BIC
the similar tests for Hispanics are not carried out sucessfully because the restrained models (equal factor loading and equal factor means) didn't converge.

Table 6
Parameter Estimates for MGA with Equal Factor Loading for Whites, Blacks, and Asians, by Gender

	White/Male		White/Female		Black/Male		Black/Female		Asian/Male		Asian/Female	
	Est.	P	Est.	P	Est.	P	Est.	P	Est.	P	Est.	P
S BY												
DEP1	0		-	-	0		-	-	0		-	-
DEP2	0.294	***	-	-	0.495	***	-	-	0.887	***	-	-
DEP3	0.726	***	-	-	0.897	***	-	-	1.129	***	-	-
DEP4	0.895	***	-	-	1.032	***	-	-	1.031	***	-	-
DEP5	1		-	-	1		-	-	1		-	-
DEP6	0.837	***	-	-	1.094	***	-	-	0.889	***	-	-
DEP7	0.309	***	-	-	0.864	***	-	-	0.223	0.256	-	-
DEP8	-0.014	0.869	-	-	0.407	***	-	-	-0.004	0.986	-	-
DEP9	-0.128	0.165	-	-	0.316	***	-	-	-0.132	0.5	-	-
Means												
I	0.906	***	1.286	***	0.963	***	1.33	***	1.008	***	1.306	***
S	0.287	***	0.581	***	0.337	***	0.612	***	0.347	0.001	0.783	***
Variances												
I	0.359	***	0.949	***	0.503	0.073	2.533	***	-0.12	0.659	0.385	0.392
S	0.472	0.01	1.011	***	0.871	0.059	3.493	***	0.04	0.908	0.891	0.206
I with S	0.264	0.014	-0.018	0.917	-0.119	0.731	-2.19	0.003	0.516	0.067	0.7	0.15
Sample Size	5340		5448		2131		2388		438		351	
DF	73				73				73			
Log Likelihood	61122				26248				4464			

Note:
*** Stands for significant P-vlaue
The parameter estimates for Equal Factor Loading for Hispanics are not presented because the model did not converge.

Table 7
Parameter Estimates for MGA with Equal Factor Loading for Males/Females, by Race
a. Male

	White/Male		Black/Male		Asian/Male		Hispanic/Male	
	Estimates	P	Estimates	P	Estimates	P	Estimates	P
S BY								
DEP1	0		-	-	-	-	-	-
DEP2	0.302	***	-	-	-	-	-	-
DEP3	0.699	***	-	-	-	-	-	-
DEP4	0.802	***	-	-	-	-	-	-
DEP5	1		-	-	-	-	-	-
DEP6	1.029	***	-	-	-	-	-	-
DEP7	0.584	***	-	-	-	-	-	-
DEP8	0.221	0.003	-	-	-	-	-	-
DEP9	0.092	0.262	-	-	-	-	-	-
Means								
I	0.833	***	1.002	***	0.943	***	1.011	***
S	0.378	***	0.363	***	0.459	***	0.471	***
Variances								
I	0.375	***	0.208	0.252	0.033	0.922	0.238	0.327
S	0.829	0.001	0.129	0.713	0.415	0.53	0.539	0.255
I with S	0.07	0.646	0.476	0.046	0.281	0.528	0.394	0.22
Sample Size	5340		2131		438		1340	

DF: 109 Loglikelihood: -38878.481

b. Female

	White/Female		Black/Female		Asian/Female		Hispanic/Female	
	Estimates	P	Estimates	P	Estimates	P	Estimates	P
S BY								
DEP1	0		-	-	-	-	-	-
DEP2	0.528	***	-	-	-	-	-	-
DEP3	0.976	***	-	-	-	-	-	-
DEP4	1.041	***	-	-	-	-	-	-
DEP5	1		-	-	-	-	-	-
DEP6	0.833	***	-	-	-	-	-	-
DEP7	0.207	0.018	-	-	-	-	-	-
DEP8	-0.211	0.064	-	-	-	-	-	-
DEP9	-0.423	***	-	-	-	-	-	-
Means								
I	1.351	***	1.587	***	1.422	***	1.623	***
S	0.438	***	0.379	***	0.719	***	0.629	***
Variances								
I	0.901	***	0.703	***	0.578	0.127	1.584	***
S	0.504	0.005	0.288	0.264	0.477	0.402	1.201	0.002
I with S	0.158	0.212	0.397	0.052	0.84	0.057	-0.231	0.422
Sample Size	5448		2388		351		1320	

DF: 107 Loglikelihood: -45981.162

Note:

*** Stands for significant P-vlaue

Figure 1
Path Diagram of the Linear LCM Using Age as Time Metric

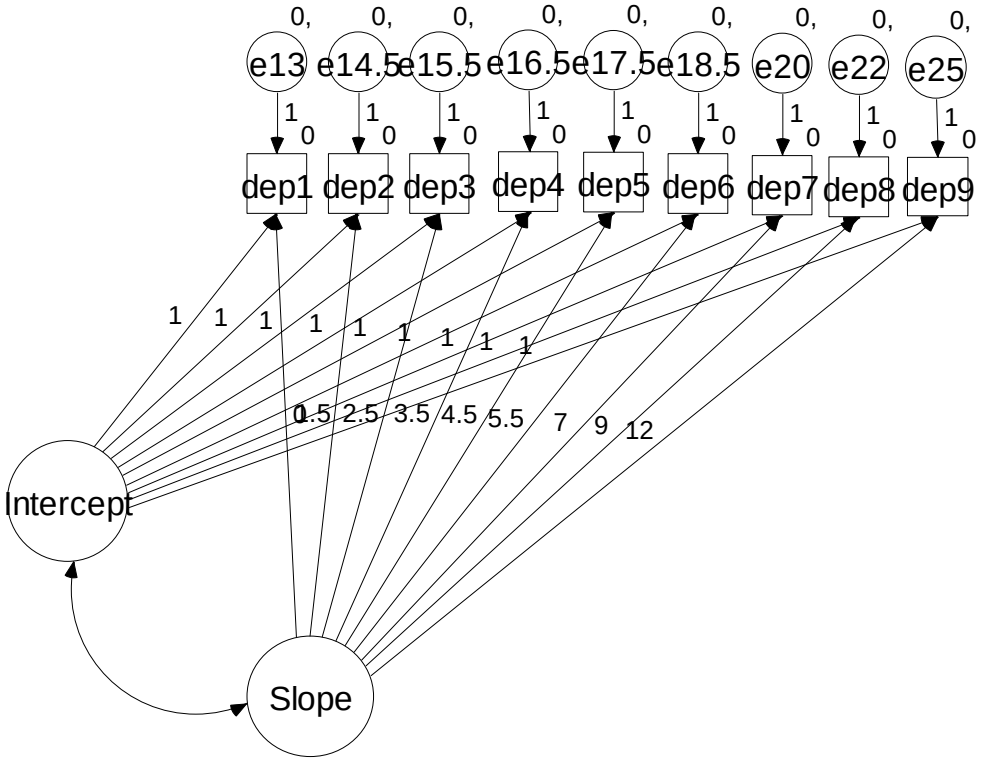
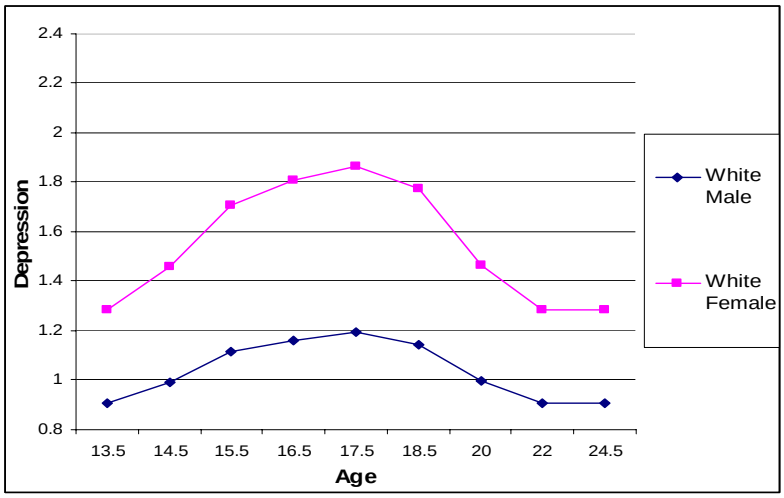
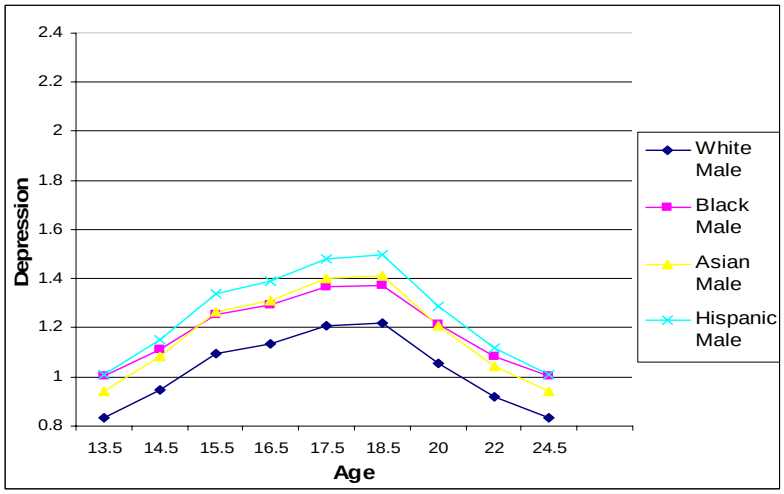


Figure 2
Model Implied Trajectories of Adolescent Depression, by Gender/Race

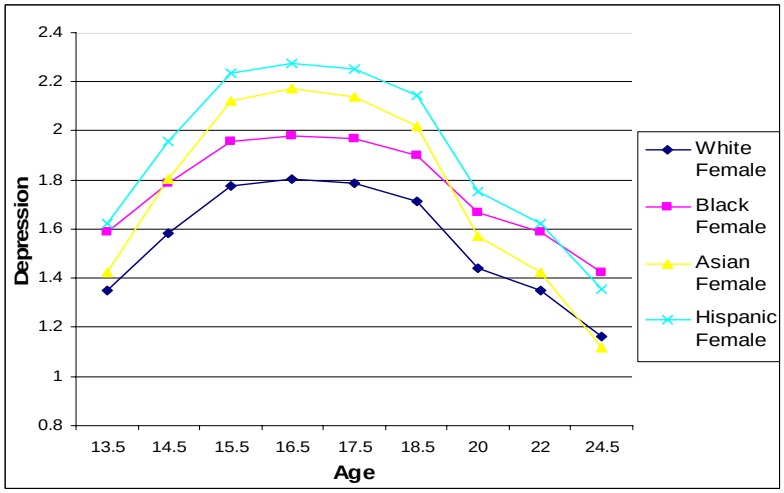
a.



b.



c.



Appendix A
Add Health Survey Schedule and Number of Respondents

Variable	Collection Dates	N
Wave I: In-Home Interview	April 1995 to December 1995	20,745
Wave II: In-Home Interview	April 1996 to August 1996	14,738
Wave III: In-Home Interview	August 2001 to April 2002	15,197
Total		20,774

Note:
The samples for each wave are slightly different.

Appendix B
Sub-scale for Selected CES-D items

QUESTIONS

These questions will ask about how you feel emotionally and about how you feel in general. How often was each of the following things true during the past week?

You felt that you could not shake off the blues, even with help from your family and your friends.

0 never or rarely
1 sometimes
2 a lot of the time
3 most of the time or all of the time

You felt depressed.

0 never or rarely
1 sometimes
2 a lot of the time
3 most of the time or all of the time

You felt sad.

0 never or rarely
1 sometimes
2 a lot of the time
3 most of the time or all of the time

REFERENCES

Allgood-Merten, B., P. M. Lewinsohn, and H. Hops. 1990. "Sex-differences and Adolescent Depression." *Journal of Abnormal Psychology* 99(1):55-63.

Aneshensel, C. S. and C. A. Sucoff. 1996. "The Neighborhood Context of Adolescent Mental Health." *Journal of Health and Social Behavior* 37(4):293-310.

Bollen, Ken. A. 1989. *Structural Equations with Latent Variables*. New York: John Wiley & Sons.

Bollen, Ken. A. and R. Lennox. 1991. "Conventional Wisdom on Measurement: A Structural Equation Perspective." *Psychological Bulletin* 110:305-14.

Bollen, K. A. and P. J. Curran. Forthcoming. *Latent Curve Models: A Structural Equation Approach*. New York: Wiley.

Butler, L. D. and Susan Nolenhoeksema. 1994. "Gender Differences in Responses to Depressed Mood in a College Sample." *Sex Roles* 30 (5-6):331-346.

Chapleski, E. E., J. K. Lamphere, R. Kaczynski, P. A. Lichtenberg, and J. W. Dwyer. 1997. "Structure of a Depression Measure among American Indian Elders: Confirmatory Factor Analysis of the CES-D Scale" *Research on Aging* 19(4):462-485.

Cole, D. A., J. M. Martin, L. A. Peeke, A. D. Seroczynski, and J. Fier. 1999. "Children's Over- and Underestimation of Academic Competence: A Longitudinal Study of Gender Differences, Depression, and Anxiety." Child Development 70(2):459-473.

Compas, B. E., G. E. Davis, and C. J. Forsythe. 1985. "Characteristics of Life Events during Adolescence." *American Journal of Community Psychology* 13 (6):677-691.

Crockett, L. J., B. A. Randall, Y. L. Shen, S. T. Russell, and A. K. Driscoll. 2005. "Measurement Equivalence of the Center for Epidemiological Studies Depression Scale for Latino and Anglo Adolescents: A National Study." *Journal of Consulting and Clinical Psychology* 73(1):47-58.

Curran, P. J. 2000. "A Latent Curve Framework for Studying Developmental Trajectories of Adolescent Substance Use." Pp. 1-42 in *Multivariate Applications in Substance Use Research* edited J. Rose, L. Chassin, C. Presson, and J. Sherman. Hillsdale, NJ: Erlbaum.

Curran, P. J. 2003. "Have Multilevel Models Been Structural Equation Models All Along?" *Multivariate Behavioral Research* 38:529-569.

Dornbusch, S. M., R. Mont-Reynand, P. L. Ritter, Z. Chen, and Steinberg. 1991. "Stressful Events and Their Correlates among Adolescents of Diverse Backgrounds." Pp. 111-130 in *Adolescent Stress: Causes and Consequences* edited by M. E. Colten and S. Gore. New York: Aldine de Gruyter.

Fitzpatrick, S. B., C. Fujii, G. P. Shragg, L. Rice, M. Morgan, and M. E. Felice. 1990. "Do Health-care Needs of Indigent Mexican-American, Black, and White Adolescents Differ?" *Journal of Adolescent Health* 11(2):128-132.

Fleming, J. E. and D. R. Offord. 1990. "Epidemiology of Childhood Depressive-Disorders - A Critical Review." *Journal of the American Academy of Child and Adolescent Psychiatry* 29 (4):571-580.

Garrison, C. Z., K. L. Jackson, F. Marsteller, R. McKeown, and C. Addy. 1990. "A Longitudinal Study of Depressive Symptomatology in Young Adolescence." *Journal of the American Academy of Child and Adolescent Psychiatry* 29(4):581-585.

Ge, X. J., F. O. Lorenz, R. D. Conger, G. H. Elder, and R. L. Simons. 1994. "Trajectories of Stressful Life Events and Depressive Symptoms during Adolescence." *Developmental Psychology* 30(4):467-483.

Ge, Xiaojia, R. D. Conger, and Glen H. Elder Jr. 2001. "Pubertal Transition, Stressful Life Events, and the Emergence of Gender Differences in Adolescent Depressive Symptoms." *Developmental Psychology* 37: 404-17.

Gore, S and R. H. Aseltine. 2003. "Race and Ethnic Differences in Depressed Mood Following the Transition from High School." *Journal of Health and Social Behavior* 44(3):370-389.

Gove, W. R. and J. F. Tudor. 1973. "Adult Sex Roles and Mental Illness." *American Journal of Sociology* 78(4):812-35.

Greenberger, E., C. S. Chen. 1996. "Perceived Family Relationships and Depressed Mood in Early and Late Adolescence: A Comparison of European and Asian Americans." *Developmental Psychology* 32(4):707-716.

Greenberger, E., C. S. Chen, S. R. Tally, and Q. Dong. 2000. "Family, Peer, and Individual Correlates of Depressive Symptomatology among US and Chinese Adolescents." *Journal of Consulting and Clinical Psychology* 68(2):209-219.

Hammen, C. 1991. *Depression Runs in Families*. New York: Springer Verlag.

Hankin, B. L., L. Y. Abramson, Phil A. Silva, Rob McGee, Terrie E. Moffitt, and Kathryn E. Angell. 1998. "Development of Depression from Preadolescence to Young Adulthood: Emerging Gender Differences in a 10-year Longitudinal Study." *Journal of Abnormal Psychology* 107(1):128-140.

Hankin, B. L. and L.Y. Abramson. 1999. "Development of Gender Differences in Depression: Description and Possible Explanations." *Annals of Medicine* 31(6):372-379.

Hankin, B. L. and L.Y. Abramson. 2001. "Development of Gender Differences in Depression: An Elaborated Cognitive Vulnerability-Transactional Stress Theory." *Psychological Bulletin* 127(*6*):773-796.

Harker, K.. 2001. “Immigrant Generation, Assimilation, and Adolescent Psychological Well-being.” *Social Forces* 79(3):969-1004.

Harris, Kathleen Mullan. 1999. “The Health Status and Risk Behavior of Adolescents in Immigrant Families.” Pp. 286-347 in *Children of Immigrants: Health, Adjustment, and Public Assistance* edited by Donald J. Hernandez. Washington, D.C.: National Academy Press.

Harter, S. 1999. *The Construction of the Self: A Developmental Perspective*. New York: Guilford Press.

Holsen, I., P. Kraft, and J. Vitterso. 2000. "Stability in Depressed Mood in Adolescence: Results from a 6-year Longitudinal Panel Study." *Journal of Youth and Adolescence* 29(1):61-78.

Iwata, Noboro, R. Jay Turner, and Donald A. Lloyd. 2002. “Race/Ethnicity and Depressive Symptoms in Community-Dwelling Young Adults: A Differential Item Functioning Analysis.” *Psychiatry Research* 110: 281-289.

Kazdin, A. E. 1994. “Informant Variability in the Assessment of Childhood Depression.” Pp. 249-274 in *Handbook of Depression in Children and Adolescents: Issues in Clinical Psychology* edited by W. M. Reynolds and H. F. Johnston. New York: Plenum Press.

Lam, C. Y., C. M. Pepper, and K. A. Ryabchenko. 2004. “Case Identification of Mood Disorders in Asian American and Caucasian American College Students.” *Psychiatric Quarterly* 75(4):361-373.

Larsson, B. and L. Melin. 1990. “Depressive Symptoms in Swedish Adolescents.” *Journal of Abnormal Child Psychology* 18(1):91-103.

Liang, J., T. V. Tran, N. Krause, and K. S. Markides. 1989. "Generation-Differences in the Structure of the SEC-D Scale in Mexican-Americans." *Journals of Gerontology* 44(3):S110-S120.

Little, R. J. A. and D. A. Rubin. 1987. *Statistical Analysis with Missing Data*. New York: John Wiley and Sons.

Nettles, S. M. and J. H. Pleck. In Press. “Risk, Resilience, and Development: the Multiple Ecologies and Black Adolescents.” in *Risk and Resilience in Children: Developmental Approaches* edited by R. J. Haggerty, N. Garmezy, M. Rutter, and L. Sherrod. New York: Cambridge University Press.

Nolen-Hoeksema, Susan. 1987. “Sex Differences in Unipolar Depression Evidence and Theory.” *Psychological Bulletin* 101(2):259-282.

Nolenhoeksema, Susan. 1990. *Sex Differences in Depression*. Stanford, CA: Stanford University Press.

Nolenhoeksema, Susan, J. S. Girgus, and M. E. P. Seligman. 1991. "Sex-differences in Depression and Explanatory Style in Children." *Journal of Youth and Adolescence* 20(2):233-245.

Nolenhoeksema, Susan and J. S. Girgus. 1994. "The Emergence of Gender Differences in Depression During Adolescence" *Psychological Bulletin* 115(3):424-443.

Perreira, K. M., N. Deeb-Sossa, K. M. Harris, and K. Bollen. 2005. "What Are We Measuring? An Evaluation of the CES-D across Race/Ethnicity and Immigrant Generation." *Social Forces* 83 (4):1567-1601.

Petersen, Anne C., Pamela A. Sarigiani, and Robert E. Kennedy. 1991. "Adolescent Depression: Why More Girls?" *Journal of Youth and Adolescence* 20(2): 247-272.

Petersen, A. C., B. E. Compas, J. Brooksgunn, M. Stemmler, S. EY, and K. E. Grant. 1993. "Depression in Adolescence." *American Psychologist* 48(2):155-168.

Powers, S. I., S. T. Hauser, and L. A. Kilner. 1989. "Adolescent Mental-Health." *American Psychologist* 44(2):200-208.

Radloff, L. S. 1975. "Sex Differences in Depression: The Effects of Occupation and Marital Status." *Sex Roles* 1:249-267.

Radloff, L. S. 1977. "The Center for Epidemiologic Studies Depression Scale: A Self Report Depression Scale for Research in the General Population." *Applied Psychological Measurement* 1(3):385-401.

Radloff, L. S. 1991. "The Use of the Center for Epidemiologic Studies Depression Scale in Adolescents and Young-Adults." *Journal of Youth and Adolescence* 20(2):149-166.

Radziszewska, B., J. L. Richardson, C. W. Dent, and B. R. Flay. 1996. "Parenting Style and Adolescent Depressive Symptoms, Smoking, and Academic Achievement: Ethnic, Gender, and SES Differences." *Journal of Behavioral Medicine* 19(3):289-305.

Raftery, Adrian E. 1995. "Bayesian Model Selection in Social Research." *Sociological Methodology* 25:111-163.

Riddle, A.S., M.R. Blais and U. Hess. 2002. "A Multi-Group Investigation of the CES-D's Measurement Structure Across Adolescents, Young Adults and Middle-Aged Adults." Working Paper.

Rushton, Jerry L., Michelle Forcier, and Robin M. Schectman. 2002. "Epidemiology of Depressive Symptoms in the National Longitudinal Study of Adolescent Health." *Journal of the American Academy of Child and Adolescent Psychiatry* 41: 199-205.

Rutter, M., P. Graham, O. F. D. Chadwich, and W. Yule. 1976. "Adolescent Turmoil – Fact or Fiction?" *Journal of Child Psychology and Psychiatry and Allied Disciplines* 17:35.

Schoenbach, Victor J., Berton H. Kaplan, Lisa Fredman, and David G. Kleinbaum. 1986. "Social Ties and Mortality in Evans County, Georgia." *American Journal of Epidemiology* 123: 577-91.

Schraedley, P. K., I. H. Gotlib, and C. Hayward. 1999. "Gender Differences in Correlates of Depressive Symptoms in Adolescents." *Journal of Adolescent Health* 25(2):98-108.

Schwartz. 1978. "Estimating the Dimension of a Model." *The Annals of Statistics* 6: 461-464.

Siegel, J. M., C. S. Aneshensel, and B. Taub. 1998. "Adolescent Depressed Mood in a Multiethnic Sample." *Journal of Youth and Adolescence* 27(4):413-427.

Twenge, Jean M. and Susan Nolen-Hoeksema. 2002. "Age, Gender, Race, Socioeconomic Status, and Birth Cohort Differences on the Children's Depression Inventory: A Meta-Analysis." *Journal of Abnormal Psychology* 111:578-588.

Weiner, I. B. and A. DelGaudio. 1976. "Psychopathology in Adolescence." *Archives of General Psychiatry* 34:98.

Weissman, M. M. and G. L. Klerman. 1977. "Sex Differences and the Epidemiology of Depression." *Archives of General Psychiatry* 35:1304-1978

Weissman, M. M., R. C. Bland, G. J. Canino, C. Faravelli, S. Greenwald, H. G. Hwu, P. R. Joyce, E. G. Karam, C. K. Lee, J. Lellouch, J. P. Lepine, S. C. Newman, M. RubioStipec, J. E. Wells, P. J. Wickramaratne, H. U. Wittchen, and E. K. Yeh. 1996. "Cross-national Epidemiology of Major Depression and Bipolar Disorder" *Journal of the American Medical Association* 276(4):293-299.

Whitley, B. E. and B. E. Gridley. 1993. "Sex-role Orientation, Self-esteem, and Depression – a Latent-Variables Analysis." *Personality and Social Psychology Bulletin* 19(4):363-369.

Wight, R. G., J. E. Sepúlveda, and C. S. Aneshensel. 2004. "Depressive Symptoms: How Do Adolescents Compare to Adults." *Journal of Adolescent Health* 34:314-323.

Willett, J. B. and A. G. Sayer. 1994. "Using Covariance Structure Analysis to Detect Correlates and Predictors of Individual Change over Time." *Psychological Bulletin* 116:363-381.