Trajectories of Adolescent Depression and Gender/Racial Disparity

Victor Wang
University of North Carolina at Chapel Hill

Abstract

This paper attempts to understand the trajectories of adolescent depression and gender/racial disparities. I conduct data analysis using Add Health with growth curve models that trace individual depression trajectories over different ages employing a measurement that is comparable across gender/race. This study found that depression increases at early adolescence and decreases at late adolescence. While this pattern is consistently found among females, it is less clear for males. Females are again 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 level of depression. And compared to their white counterparts, female minorities face greater disadvantage 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.

Background

Adolescence is an important developmental stage for understanding the nature, course, and treatment of depression. Previous studies have shown that adolescence is a highly stressful part of the life span. Elevated rates of depression in adolescence relative to adulthood have been reported in many studies (e.g., Allgood-Merten, Lewinsohn, & Hops, 1990; Larsson & Melin, 1990). Physical and psychological maturation, as well as alterations in social roles and environments, often lead to increased stress levels and depression. Adolescence is a life-stage of transition from being dependent (childhood) to independent (young adulthood), where adolescents are expanding their roles into more complex social environments that expose them to a widening array of stressors and life-shaping choices.

The majority of adolescents of both genders successfully negotiate this developmental period without any major psychological or emotional disorders, develop a positive sense of personal identity, and manage to forge adaptive peer relationships at the same time they maintain close relationships with their families (Poers, Hauser, & Kilner, 1989). However, not everyone is fortunate. Many adolescents do experience various depressive symptoms. Some of them even had their lives disrupted by severe clinical depression. Some people neglect depression and regard it as natural for this stage of life. Although these depressive symptoms are often not devastating in the adolescence, there is strong evidence 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, & Yule, 1976; Weiner & DelGaudio, 1976). Looking at depression in a developmental perspective is a most important way to conduct research on this subject.

Although depressive symptoms are commonly observed phenomena in adolescence, such moods are not uniformly distributed across the population. Research on depression is especially important considering the high rates of their occurrence during adolescence. It is clear that treatment efforts will not be sufficient to meet the needs of the population in the near future (Petersen et al., 1993). So it is important to know who are most susceptible to depression and when people are most susceptible.

Various literatures documented depression among adolescents. Some discussed the group disparities of depression. Some researchers have sought to understand this from a longitudinal perspective. However, most previous literature falls short on certain

important issues, three of which are most significant. First, measurements employed by some group comparison studies did not provide enough evidence on the comparability of the measures of depression across different groups. Second, most research on this topic consists of cross sectional group comparisons or longitudinal studies of a single group over a short period. With that, we can only see small pieces of this complex subject. Last but not least, the methods employed by many of the previous longitudinal studies of the depressed are considered problematic by current standards. One popular problem is using wave to represent time while it is age that correlates with depression. Another problem is ignoring the clustering in the data. The clustering is usually produced by repeated measurements. Research with the above problems can produce false outcome, which is bad for theoretical development and policy making. To sum up, there are various improvements need to be made in order to have a rigorous analysis and more accurate understanding of the adolescent depression.

To bridge the gap, I do not attempt to develop or test the theories regarding the reasons for adolescent depression, but I will try to accurately describe the trajectories of adolescent/young adult depression of different gender/racial group over a relatively long period of time. This will help us to find out whether or not, and in what degree, disparities exist, as well as to understand the momentum of depression among adolescence/young adults such as when and how it rises and falls. This study is especially important in clarifying some of the conflicting results in previous studies. These clarifications will eventually benefit the theoretical development of the causes of the depression and its disparities. For example, we do not have to waste efforts on developing theories to understand the causes of certain group disparities which in fact do not exist.

To do this, I first consider the conceptualization of depressive phenomena and describe its process, followed by an overview of the literature on trajectories of depression and gender/racial disparities. Then I conduct data analysis with growth curve models that trace individual depression trajectories over different ages employing a measurement that is comparable across gender/race. Finally, I will link the findings of this study to the literature.

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. Officially, as defined in the Diagnostic and Statistical Manual of Mental Disorders, depression is a disorder of mood, characterized by sadness and loss of interest in usually satisfying activities, a negative view of the self and hopelessness, passivity, low energy, poor concentration, indecisiveness, suicidal intentions, loss of appetite, weight loss, sleep disturbances, and other physical symptoms. Some or all of these symptoms may be present in people suffering from depression. These problems can become chronic or recurrent and lead to substantial impairments in an individual's ability to take care of his or her everyday responsibilities.

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, Compas, Brook-Gunn, Stemmler, Ey, and Grant, 1993). Each approach reflects different assumptions about the nature of psychopathology, serves different purposes, and reflects

a different severity of depressive phenomena. This study will focus on depressed mood, a less intense but relatively common phenomenon in the general population. Briefly, depressed mood refers to depression as an affective quality, such as feeling sad or unhappy, for an unspecified time period, but it does not signify whether other symptoms are present or absent. The study of depressed mood during adolescence has emerged from developmental research in which depressive emotions are studied along with other features of adolescent development. 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, if I do not specify, 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). It may be used to identify populations at risk of developing clinical depression or anxiety disorders. 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.

A recent development on evaluating CES-D on adolescents by Perreira, Deeb-Sossa, Harris, and Bollen (2005) categorizes the long form CES-D items into three dimensions: effects, causes, and outcomes. The multiple factor model with all CES-D measurement items seem to behave well for Whites, while not so well for other groups. This measurement consists of depressive symptoms that might not be related to mood, such as appetite. This can bring in larger measurement error, which is probably why it is not invariant across groups. Alternatively, they tested a unidimensional model consisting of only effects (depressed, life, 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 group more consistently. This is by far the most advanced evaluation and solution for CES-D on accessing adolescent depression for different racial/gender groups.

Where Is Depression Coming from?

Depression is a result of psychosocial stressors that include both biochemical and social components (American Psychiatric Association (APA), 1994; Beck, 1967; Ingram & Holle, 1992). The relationships between depression and other factors have been illustrated through schematic diagrams in many previous studies (e.g., Burke 1991, Park and Folkman 1997). Here I integrate some of the existing illustrations and present the entire process of depression in the following diagram. The solid boxes and arrows represent the main process of depression. The dashed boxes and arrows represent factors that influence the process and cause the disparity among groups.

[Figure 1 about here]

Stress is a relationship between external conditions and the current state of the person; and distress or anxiety is the internal, subjective response to that relationship. (Burke, 1991) Anxiety (distress) results from stress (Endler and Edwards 1982). Biochemical components of the stressor are usually disruptions of hormones, including puberty, pregnancy, and menopause, etc. Research on social stress(or) shows that social stress come from two major resources: the excessive demands and pressures arising from the many roles and identities that people maintain (Holroyd and Lazarus 1982; House 1974); and interruptions of the continual "identity process" (Burke 1991). Although stress was often thought of as a set of demands on individuals that tax or exceed their resources for managing them, more recent cognitive views of stress focuses on interruption ("autonomic activity resulting whenever some organized action or thought process is interrupted") and subsumes the idea of overload (Mandler 1982). Adolescence is a stage of life when people start to adopt multiple roles and their identities are highly dynamic. Identity conflict and identity change are common among adolescents. Both the role overload and interruption explanation suggest that adolescence is at more risk of depressive symptoms.

Disparities and Trajectories of 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). 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%). Elevated rates of depressed mood in adolescence relative to adulthood have been reported in other studies as well (e.g., Allgood-Merten, Lewinsohn, & Hops, 1990; Larsson & Melin, 1990). These studies suggest the depress 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. While due to a lack of studies on the transition from adolescence to young adulthood, our knowledge about relative mental health disparities between genders and races during this period is limited.

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 relatively consistent in the literature, while the findings about racial disparity have long been controversial. A review of the literature of group comparisons on depression during adolescence lead us to conclude 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 (Gove and Tudor, 1973; Radloff, 1975; Hankin and Abramson 1999; Ge, Conger, & 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; Cramer 1979; 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 et al. 1986).

Gender Disparity

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.

[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.

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. In early adolescence, depression emerges and increases. The group disparities appear as a result of differential rate of increase. Starting in early adolescence, more girls than boys begin to become depressed, and this gender difference in depression persists throughout adulthood (Nolen-Hoeksema, 1990) across many countries and cultures (Weissman et al., 1990). The emerging gender difference in depressed mood and depressive disorders appear after the age of 13 or mid-puberty, where the rate of depression rises dramatically for girls while boys' symptoms and mood remained constant (Cole, Martin, Peeke, Seroczynski, & fier, 1999; Ge, Lorenz, Conger, Elder, & Simons, 1994; Petersen, Sarigiani, & Kennedy, 1991). More precisely, this increase for girls was (from 4% to 23%), much greater than for boys (from 1% to 11%), according to Hankin et al. (1998).

Hormonal changes are considered to be acting behind the gender differences in increased depression level after puberty. Other than biological reasons, research has shown social and psychological factors play important roles in causing gender differences. Pre-puberty children are sometimes seen as "no-gender" (gender neutral). After the biological changes of puberty, social and psychological (e.g., gender role, anxiety on sexuality, physical appearance associated with sexuality) factors start to make greater influences on adolescents. One of the explanations for increased experience of depression among girls is that girls experience more challenges than do boys (Petersen et al. 1991). Since the 70s, studies have linked the feminine role or stereotype with more depression (Gove and Tudor, 1973; Radloff, 1975). Girls identify more strongly with a feminine stereotype of needing to appear thin and consequently become more dissatisfied with

their body shape and physical appearance, which in turn is associated with increased depression (Hankin and Abramson 1999). Girls are also more likely to experience negative events in the family than boys (Compas et al., 1985), and these adversities are in turn associated with elevated depression. Research has found the adolescent girls are more susceptible to network or peer related disruption than boys (Hankin & Abramson 2001). Research on adult women has also shown that they are more vulnerable to their negative effects on mental health (Turner & Lloyd 1995). And this might apply to adolescent girls too.

In terms of coping strategy, men and women may have different response styles to depression in which men distract themselves, whereas women ruminate on their depressed mood and therefore amplify it (Nolen-Hoeksema, 1987). Cramer (1979) found that adolescent girls reported using "internalizing" defense mechanisms-turning against the self, intellectualization, and rationalization-more than boys. In contrast, boys reported using "externalizing" mechanism-projection, turning against others-more than girls. Having more challenges but less effective coping styles 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 and Klerman, 1977; Gove & Tudor, 1973; Nolen-Hoeksema, 1990; Nolen-Hoeksema,

Girgus, & Seligman, 1991). These examinations have concluded that the gender differences appear to be a true difference in the experience of depression. Theories about the existence and causes of gender disparities have been well established. However, these studies were done only in theoretical level and were not carried out with formal statistical tests.

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 model using confirmatory factor analysis (CFA) is occasionally employed. However, few of them were supported with invariance test. Studies dedicated to the assessment of the measurement comparability are available. But as I mentioned earlier, the findings of these studies 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. Hispanics and Blacks are the groups most likely to be identified as having greater levels of depressive symptoms in the literature. Siegel, et

al. (1998) conducted research on adolescents (age 12-17 years) where depressed mood was assessed by the Children's Depression Inventory (CDI). They found that compared to Whites, African Americans or Asian Americans, Latinos reported more symptoms of depressed mood, a finding that was independent of socioeconomic status. Many other studies also report that Hispanic Americans have higher depression levels than other ethnic groups (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. Gore and Aseltine (2003) also find that African-American young adults have higher depression levels than either Whites or Asian Americans.

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. 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 racial differences in depression are the result of disadvantaged pathways into adulthood. Wight et al. (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 et al., 1990), it is probably wise to note Hammen's (1991) conclusion that there is no evidence for Black-White differences in depression among adults. 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.

Interactions between Gender and Race

Last, it is important to examine how ethnicity may play a role in the development of gender differences in depression. Two studies (Schraedley, Gotlib, & 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.

Hypothesis

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

2. The trajectories for different race and gender are significantly different from each other. This hypothesis can be further broken down to two components. Each can be tested separately.

2a. there is significant gender difference in depression trajectories.

2b. there is significant racial difference in trajectories for each gender.

Data, Measurement, and Methods

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 (see Bearman, Jones & Udry 1997 for more details on the Add Health design). The survey time and sample size information of each wave is shown in Appendix Table A. The age distributions at each wave are shown in the following Table 1.

[Table 1 about here]

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. I assume the

data points are missing at random (MAR)[2] (Little and Rubin, 1987). This assumption will allow me to use direct maximum likelihood (DML) method to handle the missing data (Bollen and Curran, forthcoming). When selecting the sample, I employed a series of identifiers of race/ethnicity and generation of immigration created by Harris. According to the identifiers, 9 cases are deleted due to lack of information on race. 334 Indians are deleted (20431 left) because some of the environments in which they grow up are very unique compared to other respondents. 1667 first generation immigrants are deleted (18764 left as second+ generation) due to their cultural backgrounds and sometimes even language capabilities[3]. Finally, 8 more 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 Whites, Blacks, Asians, and Hispanics of the second+ generations.

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

[2] Cases with incomplete data differ from cases with complete data, but the pattern of data missingness is traceable or predictable from other variables in the database rather than being due to the specific variable on which the data are missing. The actual variables where data are missing are not the cause of the incomplete data. Instead, the cause of the missing data is due to some other external influence.

[3] These background differences can reduce the consistency of the measurement. This can further cause unnecessarily greater deviation/complication which sometimes can not be explained by mainstream theory or captured by general information collection.

describe their racial background and 1,038 out of 20,745 respondents did. Since the employment of multiple group analysis in this study, I use mutually exclusive race categories[4]. 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 management as a way to obtain exclusive race category. This method is also a popular measurement 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) I used a substitute (three variables) of the CES-D to measure depression at each wave.

[4] 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.

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 decreases the consistency of the measurement when CES-D is used for multi-ethnic study. Researchers started to use substitutes of CES-D in order to obtain consistency when doing group comparison type of analysis. In the 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 measurement 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 in this study because the other two

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 Table 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[5].

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.

[5] 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. The constraining the factor loadings and intercepts to be equal across racial and gender groups do not seem to reduce the model fit much considering the large sample size. That means this measurement model is invariant measurement for these groups. Therefore it can be used to compare groups.

In order to capture the trajectory of the mental health development in the transition from adolescents to young adults, I employed the latent curve model, or LCM, based on structural equation models (SEM). Handling the 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 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., $\zeta_{\alpha i} + \lambda_t\zeta_{\beta i} + \varepsilon_{it}$).

[Figure 2 about here]

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 2. But in this study, as shown in Table 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 lots 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. The method they suggested is Direct Maximum Likelihood (DML). With the assumption of MAR, the direct ML technique makes use of all available information in the data. No cases are discarded and all the values of the variables available for a case go into the calculations. Unlike multiple imputation, the direct ML does not involve imputing the values of any variables that are missing and no values are "filled in." (Bollen and Curran, forthcoming). 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 is two years. This is done to reduce the number of variables and improve data coverage. I use the mean age of the two-year span to represent age.

$$\begin{pmatrix} dep_{1,1} & dep_{1,2} & dep_{1,3} \\ dep_{2,1} & dep_{2,2} & dep_{2,3} \\ dep_{3,1} & dep_{3,2} & dep_{3,3} \\ \cdots & \cdots & \cdots \\ dep_{N,1} & dep_{N,2} & dep_{N,3} \end{pmatrix} \tag{2.1}$$

$$\begin{pmatrix} dep_{1,12} & dep_{1,14} & . & . & . & dep_{1,22} & . & . \\ . & dep_{2,14} & dep_{2,16} & . & . & . & dep_{2,24} & . \\ . & . & dep_{3,16} & dep_{3,18} & . & . & . & dep_{3,26} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ . & . & dep_{N,16} & dep_{N,18} & . & . & . & dep_{N,26} \end{pmatrix} \tag{2.2}$$

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 substantially, the mean level of depression at that age range.

[Table 2 about here]

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 the disparities suggested by the literature. MGA allows us to constrain certain parameters to be equal across groups and formally test whether these constrain influence the model fit. Therefore it is a great tool to examine group differences.

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[6] 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 whether or not depression changes over time and whether it follows a linear change. And 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[7].

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)[8] 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 log-likelihood ratio (LR) tests and the Bayesian Information Criterion (BIC) (Schwartz, 1978; Raftery, 1995). The least

[6] 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.

[7] 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.

[8] 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.

restrictive model is the Freed Loading Model and the other models are nested within it. Table 3 shows the results of LR tests between the Freed Loading Model and each other model. The LR test 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 LR test results also suggest that the Linear Models 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 also looked at the Bayesian Information Criteria (BIC). BIC assigns a greater penalty to model complexity, and so has a greater tendency to pick parsimonious models. According to Raftery, the model with smaller BIC is preferred and how strong the preference depends on the magnitude of the difference. 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 smaller than that of the other models for both groups. The BIC indicates constraining the model appears to cause real declination of the 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.

[Table 3 about here]

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[9]), 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 adolescence, 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.

[Table 4 about here]

Testing Hypothesis 2 – Gender and Racial 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,

[9] 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.

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 LR test and BIC. The results of the model comparison (LR) tests, BICs, and Sample-size Adjusted BICs are shown in the first section 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.

[Table 5 about here]

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.

[Table 6 about here]

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[10]. The slope factor mean for White females is substantially higher that 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, experience faster increase on depression due to their higher slope factor mean (0.581 VS 0.287). This results in even greater gender gap. On the other hand, 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

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

both males and females, indicating there are significant individual differences in the rate of change within each gender group.

The model implied trajectories of both white males and females are illustrated in the first section of figure 3 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.

[Figure 3 about here]

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 section 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). And the BIC, with much smaller value for the Equal Factor Loadings Model, clearly favors the model constraining the factor loadings to be equal across. However, further constraining the factor means to be equal across groups do not seem to affect the model fit as much; the Equal Factor Means Model produces smaller BIC values compared to the Equal Factor Loadings Model. Only when the BIC is adjusted for sample size, the model fit comparison favors the less constrained Equal Factor Loadings Model. 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 3 As we can see, the 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 similar starting point and rate of change. 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 very homogenous among racial groups. Their depression trajectories are relatively flat.

[Table 7 about here]

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, compared to male groups, further constraining the factor means to be equal across groups results in serious model fit reduction according to the BICs. Female groups have noticeable between-group differences; 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. Reflected in the third section of Figure 5, the trajectory for white females, being the lowest, is greatly lower than that of the Hispanic females. And Asian and Hispanic females experienced 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 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 amount of change over the course. They also have much more within and between group variations compared to male groups.

Discussion

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 adolescence, as well as to identify the implications of this increase in depression for later life. The present study represents a first step in understanding the longitudinal trajectory of depressive symptoms in a national 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. 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 when compared to males for all races. This supported the theory that puberty is the stage when gender gap emerges. The gender difference is believed to be due to a combined effect of hormonal levels (Ge, Conger, & Elder 2001), cognitive explanatory style (Broderick & Korteland 2002), and self-concept (Facio & Batistura 2001) factors. Sex role socialization in early adolescence related to the biological changes of puberty that heighten an identity with one's gender is also thought to produce the observed change in these gender differences by mid-adolescence.

Racial disparity, unsure in the previous literature, is clarified in this study. Minority groups do encounter greater level of depression. And compared to their white counterparts, female minorities face much greater disadvantage compared to male minorities. Female minorities, caught in cross-fire, 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 disparity test, I found that the gender disparity has a much greater magnitude compared to racial disparities. Therefore, it is more substantively important to study gender disparity and its causes.

A limitation of this study is its measurement is still less than ideal. But it is already a great improvement over the measurement employed by most previous studies. The three indictor measurement employed by the present study focuses on the core concept of depression and reduces complications and disruptions. Exclusion of first-generation immigrants further improves the consistency of the measurement. The measurement of this paper is well based on the literature dedicated to assessment of measurement for group comparisons. This study is also limited by the fact there are only three repeated measurement over 18 years of life span. This problem plus the unequal spacing between waves provide only moderate amount of information.

It is only recently that the researchers start to pay attention to adolescents' experiences with depression. We are still in the early stages of exploration in this area. With the utilization of the latest development of LCM, as well as many other improvements over the previous studies such as longitudinal national representative sample and improved measurement for group comparison, I provided by far the most

precise description of the depression trajectories and disparities of different races/genders in adolescence and young adulthood. With detailed description of depression in gender/race specific and developmental perspective, this study lays groundwork for further examination of the risk factors and correlates of depressive symptoms in adolescence.

Reference

Block J.H., Block J., Gjerde P.F. 1986. "The personality of children prior to divorce – a prospective-study" *Child Development* 57 (4): 827-840

Broderick, Patricia C. and Constance Korteland. 2002. "Coping Style and Depression in Early Adolescence: Relationships to Gender, Gender Role, and Implicit Beliefs." *Sex Roles* 46: 201-13.

Burke, Peter J. (1991). "Identity Process and Social Stress" *American Sociology Review* 56 (6): 836-849

Curran, P.J. (2003). "Have multilevel models been structural equation models all along?" Multivariate Behavioral Research, 38, 529-569.

Curran, P.J. 2000. "A latent curve framework for studying developmental trajectories of adolescent substance use." In J. Rose, L. Chassin, C. Presson, & J. Sherman (Eds.), *Multivariate Applications in Substance Use Research*, (pp. 1-42). Hillsdale, NJ: Erlbaum.

Ge, Xiaojia, Rand 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., R. H. Aseltine Jr., and M.E. Coten. 1993. "Gender, Social-Relational Involvement, and Depression." *Journal of Research on Adolescents*(3):101-25.

Guo, Guang and John R. Hipp. 2004. "Analysis of Linear Longitudinal Data." Pp. 347-368 in New Handbook on Data Analysis, edited by M. A. Hardy. London: Sage.

Hankin, BL, Abramson, LY, Moffitt, TE, et al. 1998 "Development of depression from preadolescence to young adulthood: Emerging gender differences in a 10-year longitudinal study" *J ABNORM PSYCHOL* 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 Dec 1999

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.

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*. Donald J. Hernandez, ed. Washington, D.C.: National Academy Press.

Harter, S. (1999). *The construction of the self: A developmental perspective*. New York:

Guilford Press

Herman-Stahl, Mindy A., Mark Stemmler, and Anne C. Pertersen. 1995. "Approach and Avoidant Coping: Implications for Adolescent Mental Health." *Journal of Youth and Adolescence* 24(6): 649-665.

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.

Joyner, Kara, and J.R. Udry. 2000. "You Don't Bring Me Anything but Down: Adolescent Romance and Depression." *Journal of Health and Social Behavior* 41(4): 369-391.

Kazdin, A. E. 1994. "Informant variability in the assessment of childhood depression." in W. M. Reynolds & H. F. Johnston (Eds.), *Handbook of depression in children and adolescents: Issues in clinical psychology* (pp. 249-274). New York: Plenum Press.

Little, R.J.A. & Rubin, D.A. 1987. Statistical analysis with missing data. New York: John Wiley and Sons.

Mizell, C. Andre. 1999. "Life Course Influences on African American Men's Depression: Adolescent Parental Composition, Self-Concept, and Adult Earnings." *Journal of Black Studies* 29(4): 467-490.

NOLEN-HOEKSEMA, S. 1987. "Sex Differences in Unipolar Depression Evidence and Theory" *Psychological Bulletin* 101 (2): 259-282

NOLEN-HOEKSEMA, S. 1991. "Responses to Depression and Their Effects on the Duration of Depressive Episodes" *Journal of Abnormal Psychology* 100 (4): 569-582

Park, Crystal L., Susan Foldman. 1997. "Meaning in the Context of Stress and Coping" *Review of General Psychology* 1 (2): 115-144

Petersen, Anne C., Pamela A. Sarigiani, Robert E. Kennedy. 1991. "Adolescent depression: why more girls?" *Journal of Youth and Adolescence* v20 n2 p247(25)

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

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.

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, PK, Gotlib, IH, Hayward, C. 1999. "Gender differences in correlates of depressive symptoms in adolescents" J ADOLESCENT HEALTH 25 (2): 98-108

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

Siegel, JM, Aneshensel, CS, Taub, B. 1998. "Adolescent depressed mood in a multiethnic sample" J YOUTH ADOLESCENCE 27 (4): 413-427

Turner, R. Jay, Blair Wheaton, and Donald A. Lloyd. 1995. "The Epidemiology of Social Stress." *American Sociological Review* 60: 104-25.

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.

Willett, J. B., & Sayer, A. G. 1994. "Using covariance structure analysis to detect correlates and predictors of individual change over time." *Psychological Bulletin* 116, 363-381.

Table 1
Distribution of Age at Each Wave

Race	Wave I		Wave II		Wave III	
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 Constructed Data for LCM Using Age as the Time Metric

Variables	Range of Age	Appr .Medium of Age Range	Linear Factor Loadings[11]	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

[11] 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 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:
(1). LRT stands for Log Likelihood Ratio Test.
(2). BIC' is Sample-Size Adjusted BIC

Table 4
MGA Results of Freed Loading Model Estimated Separately for White Males and 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 Equal Factor Loading Model and Equal Factor Means Model to Freed Loading Model for Different MGAs

Model	DF	LRT	P-value	Change of BIC	Change of BIC'
MGA for White Groups (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 Male Groups (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 Female Groups (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:
(1). LRT stands for Log Likelihood Ratio Test.
(2). BIC' is Sample-Size Adjusted BIC

Table 6
MGA Results of Equal Factor Loading Model for White Male and White Female

	White/Male		White/Female	
	Estimates	P-value	Estimates	P-value
S BY				
DEP1	0		-	-
DEP2	0.294	***	-	-
DEP3	0.726	***	-	-
DEP4	0.895	***	-	-
DEP5	1		-	-
DEP6	0.837	***	-	-
DEP7	0.309	***	-	-
DEP8	-0.014	0.869	-	-
DEP9	-0.128	0.165	-	-
Means				
I	0.906	***	1.286	***
S	0.287	***	0.581	***
Variances				
I	0.359	***	0.949	***
S	0.472	0.01	1.011	***
I	0.264	0.014	-0.018	0.917
Sample Size	5340		5448	

DF: 55 Loglikelihood: -48793.637

Note:
*** Stands for significant P-vlaue

Table 7

a. MGA Results of Equal Factor Loading Model for Males of Different Races

	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. MGA Results of Equal Factor Loading Model for Females of Different Races

	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

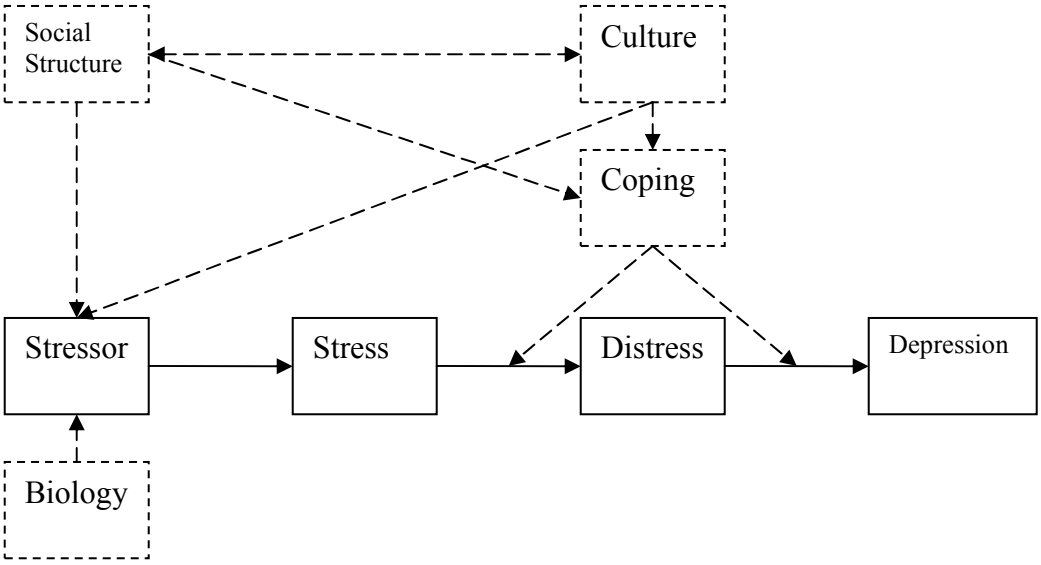


Figure 1
Depression Process

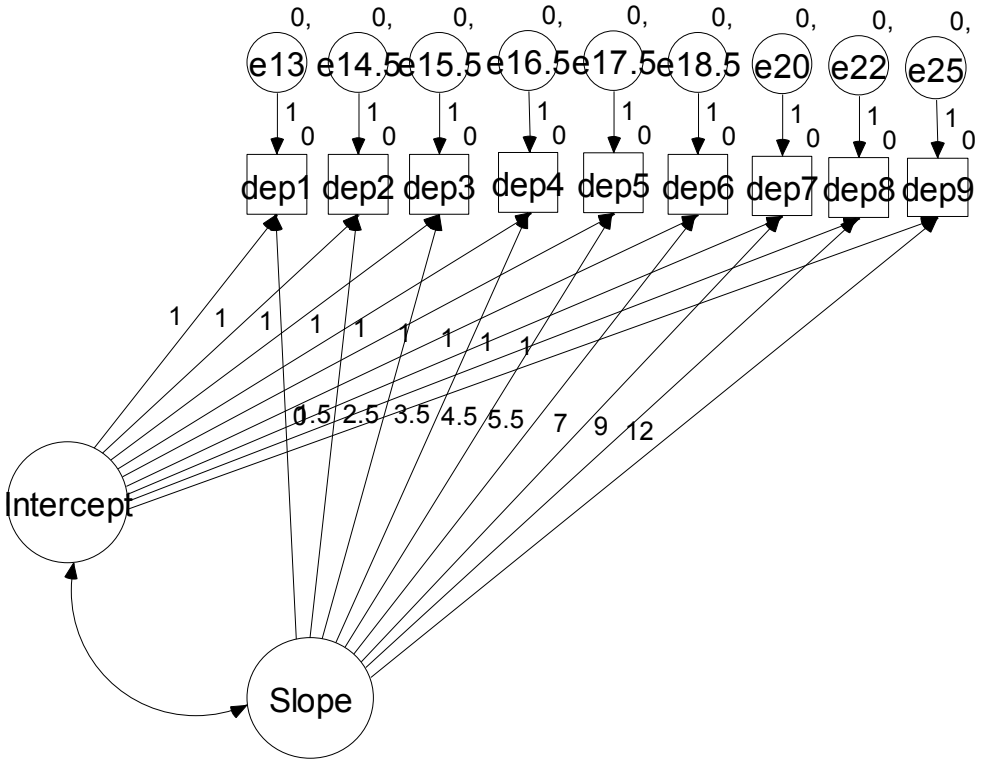
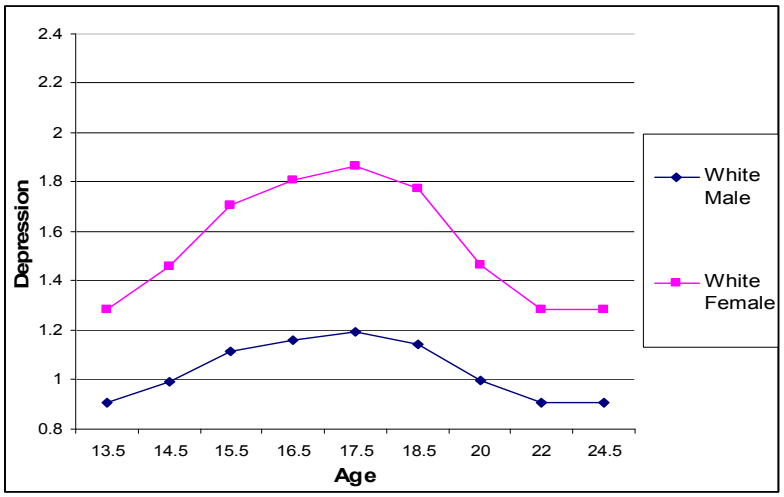
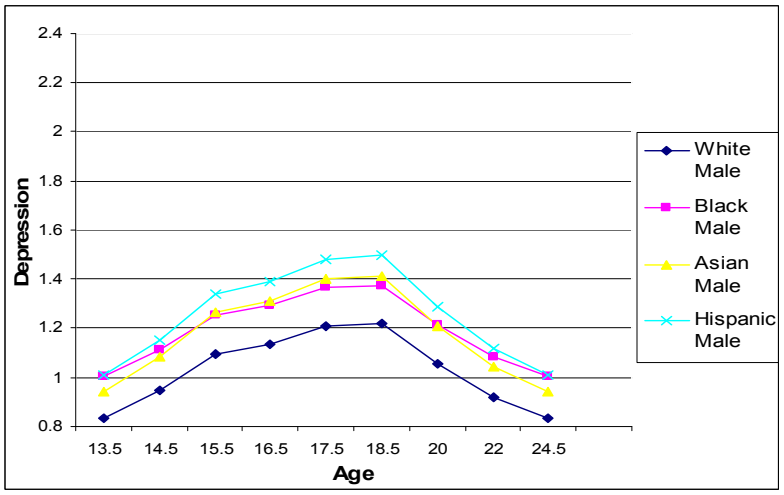


Figure 2
Path Diagram of the Linear LCM of Depression Using Age as Time Metric

a.



b.



c.

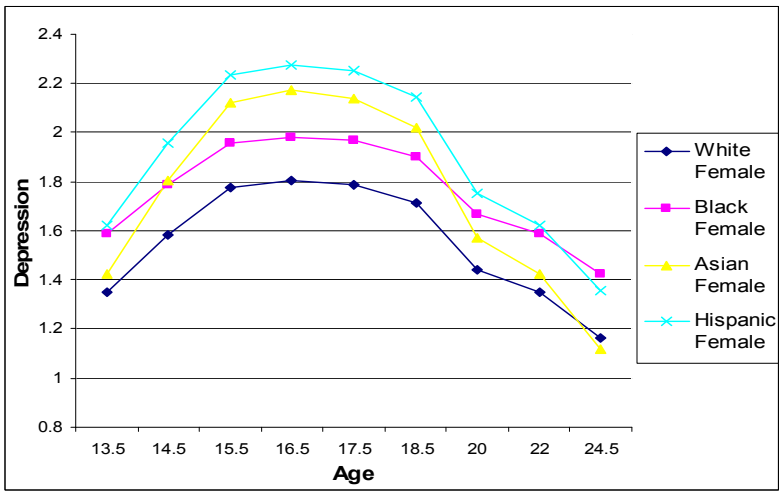


Figure 3
Model Implied Trajectories of Depression for Each Race/Gender group

Appendix

Table A
Add Health Survey Schedule

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

Table B
CES-D 3-item Sub-scale

QUESTION
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