Assessing the relationship between depression and obesity using Structural Equation Modeling

Assessing the Relationship Between Depression and Obesity Using Structural Equation Modeling

By

ALINA DRAGAN

A Project

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the degree

Master of Science

McMaster University

June 2005

MASTER OF SCIENCE (2005)
(Statistics)

McMaster University
Hamilton, Ontario

TITLE: Assessing the Relationship Between Depression and Obesity Using Structural Equation Modeling

AUTHOR: Alina Dragan

SUPERVISOR: Dr. Noori Akhtar-Danesh

NUMBER OF PAGES: viii, 123

Abstract

In this project we used structural equation modeling to analyze the data collected for the Canadian Community Health Survey (CCHS) Cycle 1.2 - Mental Health and Well-Being conducted by Statistics Canada. The data are cross-sectional.

We looked at the relation between depression and obesity adjusting for gender, socioeconomic status, gene-environment interactions, eating and physical activity and stress.

We used the AMOS and Mplus softwares to analyze our data. The first one used continuous variables for depression ("persistence of depression", in years) and obesity ("body mass index"-BMI), while the second used categorical variables: lifetime depression, 12 month depression and obesity (normal weight, overweight and obese). We also used two variables to measure different aspects of stress: self-perceived ability to handle an unexpected problem and work stress-social support.

We fitted the models across the entire data, but also across different groups: males versus females and groups based on gender and BMI.

The results indicated that the relationship between depression and obesity is different across gender.

The limitations of the study are also discussed.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Noori Akhtar-Danesh, for his help, guidance, support, encouragement and especially patience throughout the work on this project and not only.

I would also like to thank Dr. Peter Macdonald for giving me the chance to study at this university and the other professors in the department for the interesting courses they taught us.

But none would have been possible without the help and support of my parents and brother. I want to thank them all for helping me make my dreams come true.
And because dreams involve friends, all the best and all my love for the best friends in the world.

Thank you!

Table of Contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Background

Canadians are among the healthiest people in the world and have an average life expectancy that is one of the highest in the world [1]. However, within the Canadian population, there are surprising differences in health status.

Obesity became a widespread problem in Canada with serious public health implications; in fact the World Health Organization (WHO) has recognized the rise in obesity rates as a worldwide epidemic and has been calling for action to address the problem since 2000 [2].

Obesity has a major impact on the burden of disease in Canada. A substantial body of research has linked obesity to major preventable chronic diseases, including Type 2 diabetes, cardiovascular diseases, hypertension, stroke, gallbladder disease and some cancers [2].

Research also showed that obesity is not just a problem for adults in Canada, but is also having an impact on children's health. Obese children and adolescents have a greater occurrence of hypertension and high cholesterol levels, two known risk factors for cardiovascular disease. Previously seen only among adults, Type 2 diabetes is now increasingly found among obese children, particularly adolescents [3].

Canadian Institute for Health Information considers that the obesity problem is mostly due to Canada's decreasing emphasis on physical education, a lack of grassroots sports development, too many fast-food options, and not enough emphasis on alternate transportation like cycling. Also, we live in the new "era of video and computer games", which may explain why obesity among boys is increasing faster than among girls in Canada and many other countries [4].

In 2000-2001, based on self-reported data, more than 6 million adults aged 20 to 64 years, were overweight and almost 3 million obese (Statistics Canada [5]).

Canada does not conduct routine dietary surveillance. The only national nutrition survey in Canada spanning all age groups was conducted 30 years ago (Nutrition Canada, 1970-1972). That is why it is difficult to assess the extent to which energy intake is increasing or decreasing. However, provincial nutrition surveys of adults gathering detailed information on food intake and a small national survey from 1997-1998 found that overall energy intake was lower compared to 30 years ago. This trend is consistent to nutrition surveys from UK and US, which lead to the speculation that what accounts for rising levels of overweight and obesity is the decrease in physical activity rather than increased intakes [6, 7].

"Depression is not just feeling blue for a day, but is the result of actual chemical changes that take place in the brain, causing profound episodes of sadness, crying and loss of energy" [8]. Depression and manic depression (known as affective disorders or mood disorders) are among the most common diseases in our today's society affecting more than 10 out of every 100 people. Because of high prevalence, economic cost, risk of suicide and loss of quality of life, mood disorders present a serious public health concern in Canada [9].

Depression is commonly associated with loss of appetite but in a significant minority, depression can be associated with increased appetite and weight gain [9]. The increase in weight experienced in this situation often leads to further feelings of worthlessness and sometimes even the medications can lead to an even greater weight gain.

Using the Canadian Community Health Survey data (described below), the objective of this study is:

- to describe the relationship between depression and obesity adjusting for gender, socioeconomic status, gene-environment interactions, eating and physical activity and stress.

1.2 CCHS data

In this analysis we will use the data obtained from The Canadian Community Health Survey (CCHS), Statistics Canada. It is a cross-sectional survey that collects information related to health status, health care utilization and health determinants for the Canadian population.

The CCHS operates on a two-year collection cycle. The first year of the survey cycle ".1" is a large sample, general population health survey, designed to provide reliable estimates at the health region level. The second year of the survey, cycle ".2"- "Mental Health and Well-being", is a smaller survey designed to provide provincial level results on specific focused health topics.

The information was collected between May 2002 and December 2002, for the ten provinces, but we will restrict our analysis only to the province of Ontario.

CCHS Cycle 1.2 collects responses from persons aged 15 years or older, living in private occupied dwellings. Excluded from the sampling frame are individuals living on Indian Reserves and on Crown Lands, Health Care institution residents, full-time members of the Canadian Armed Forces and residents of certain remote regions.

The well-being and determinants of health in Cycle 1.2 are based on sources used on surveys such as the National Population Health Survey (NPHS), the CCHS (Cycle 1.1), the Health Promotion Survey (HPS) and others.

The Cycle 1.2 of CCHS based its findings on height and weight measurements that respondents themselves reported. Studies have shown that both men and women who respond to health surveys tend to underestimate their weight and overestimate their height [9a]. This can lead to potentially substantial underestimates of obesity and overweight.

In the autumn of 2005, Statistics Canada will release the results of CCHS Cycle 2.2 on nutrition, which for the first time will include estimates of BMI based on direct measures of height and weight.

CCHS 1.2 collected data from 36,984 persons, with a response rate of 77.0 % at Canada level, from which 12 376 were in Ontario.

Because we deal with a cross-sectional data the method of analysis chose is Structural Equation Modeling (SEM) introduced in Chapter 3. In Chapter 2 a short review of the literature concerning depression, obesity and their relationship is presented. In Chapter 4 we describe the data and the variables we are using and in Chapter 5 we present the analysis for both continuous and categorical variables which measure depression and obesity.

The analyses were conducted using SPSS 13, AMOS 5 and Mplus 3.12.

Chapter 2

Literature review

2.1 Depression

Although almost everybody uses the expression "I'm depressed" from time to time, feeling sad or down is not the same as suffering from depression. Depression does not go away in a few days and it is not caused by a lack of will power.

Major depressive disorder is a recurrent illness with frequent episode relapses and recurrences. The more severe and long lasting the symptoms in the initial episode, due in some cases to a delay in receiving effective treatment, the less likely is the full recovery [9]. Depression also has a major impact on the mental health of family members and caregivers, often with an increased presence of depression symptoms [9].

Mood disorders have no single cause, but several factors such as biochemical imbalance in the brain, psychological factors and socio-economic factors, tend to make some individuals prone to such disorders.

Studies have shown that individuals with depression often find a history of it in immediate family members [10].

One episode of major depression is a strong predictor of future episodes. More than 50% of individuals who have an episode of major depression experience a recurrence [11].

Stress has been viewed as a major risk for depression. Recent studies have indicated, however, that stress may predispose individuals only for an initial episode and not for recurring ones [12].

Because of the high prevalence, mood disorders have a major effect on the Canadian economy, first through the associated loss of productivity in the workplace due to absenteeism and diminished effectiveness; and second with the high health care costs attributable to primary care visits, hospitalizations and medication. At the individual and family level, the loss of income and the cost of medication create a strain on the family financial resources [9].

Epidemiological surveys usually conceptualize depression as a diagnosis based on the criteria of diagnostic systems such as DSM-IV diagnostic classification (the fourth edition of Diagnostic and Statistical Manual of Mental Disorders [13] or ICD-10 (International Classification of Disorders) [14].

According to World Health Organization, major depression is the leading cause of years of life lived with disability [15].

Based on one study carried out by World Health Organization in 14 countries [16] 24% of people worldwide seeking help from a primary care provider received an ICD-10 psychiatric diagnosis. The most frequent co-morbid disorders were depression and anxiety. However, substantial differences are observed in the prevalence of depression in different countries. In a study by Goldberg and Lecrubier [17] the lowest and highest rates were seen in Japan (2.6%) and Chile (29.5%), respectively.

Canadian studies looking at lifetime incidence of major depression found that 7.9% to 8.6% of adults over 18 years of age, living in the community met the criteria for a diagnosis of major depression at some time in their life [18]. During any 12-month period, about 5% of the canadian population will experience major depression [18].

According to the 1994/1995 National Population Health Survey (NPHS), 6% of the Canadian population aged 12 years or over had symptoms consistent with depression at the time of the survey [19].

The most recent study used the CCHS Cycle 1.2 data and found that the prevalence of lifetime depression in the province of Ontario is 11.0 %, while 12-month depression prevalence is 4.8% [20].

Mood disorder affects individuals of all ages, but usually appears in adolescence or young adulthood. However, late diagnosis is common. Studies have also shown higher rates of depression among women than among men [21].

2.2 Obesity

Obesity is an abnormal accumulation of body fat frequently resulting in an impairment of health. The prevalence of obesity has been rapidly increasing in both children and adults in the past 20 years. It has been proved that it contributes to a series of physical illness such as hypertension, coronary heart disease and diabetes mellitus, associations also being reported with osteoarthritis, gallbladder disease and cancer, but less is known about the possible link between obesity and mental disorders [2].

It is believed that obese people experience more psychological distress that may lead to depression. However, the information on the relationship between obesity and depression is inconsistent [22].

Obesity is a very common disorder among Americans: 65% of them are considered overweight or obese [23]. It is believed that it will occur together with major depression, for which the prevalence has been estimated at 10 % [24].

The World Health Organization's MONICA (monitoring of trends and determinants in cardiovascular diseases) study estimated that, internationally, between 50% and 75% of adults are either overweight or obese [25]. With more than 280,000 deaths per year attributable to obesity [26], it is second only to smoking as a cause of death [27]. Weight gain also has a negative impact on social and psychological well-being due to discrimination against overweight and obese people in various areas of life.

The adverse effects of weight gain, in combination with the difficulties and stigma associated with severe mental illness, may profoundly compromise an individual's physical health and quality of life [28], [29], [30].

In Canada, in 1994-1995, the first cycle of the National Population Health Survey found an overall prevalence of obesity of 30.6 % in non-institutionalized Canadians [31] (the definition of obesity was a BMI greater or equal to 27).

According to *Statistics Canada* the rates of obesity (and overweight) have increased slightly during the past three years. In 2000/01, 14.1% of the adult population (aged 18 and over) was considered obese and 32.4% overweight. By 2003, 14.9% of adult Canadians were considered obese and 33.3% of them were overweight. An estimated 46.7% were in the normal range, and about 2.7% were underweight. Being underweight is considered to harbor the same health risks as being overweight. About 15.9% of adult men were considered obese, slightly higher than the rate of 13.9% among adult women. Rates of obesity were highest in the age group 45 to 64 [32].

As noted in the "Joint Canada–US Survey of Health", released June 2, 2004 in "The Daily", Statistics Canada, obesity rates are higher in the United States than in Canada, especially among women.

In a study which compared the prevalence of childhood obesity in Canada, Scotland, England and Spain, while all of the countries had comparable rates in the early 1980s, Canada's rates rose by leaps and bounds by the late 1990s. Today the prevalence of childhood obesity in Canada is double that of those countries, and is similar to rates in the United States [33].

Usually estimates of the prevalence of obesity among Canadians vary widely, depending on whether a study is based on self-reports or objective measures, and on what measure of obesity is used.

2.3 Depression and obesity

Depression is the number one mental disorder and obesity almost reached epidemic proportions in some countries. However, little is known about the relationship between the two disorders and there has been little systematic research on the topic [34]. They were both studied as two largely independent disorders, the investigators from the two fields having little contact.

The existing information on the relationship between obesity and depression is inconsistent. Some studies reported that obese people were at elevated risk for depression [35, 36, 37]. Others found that heavier people were less depressed [38, 39] especially middle-aged men. There were several reports indicating no effect of obesity on the risk of depression [40].

Some surveys revealed that the association between obesity and depression might be sex-specific. Onyike et al (2003) found that obesity was associated with past month depression in women, but not significantly associated in men [41].

When obesity was stratified by severity, heterogeneity in the association with depression was observed [41]. Hence, the association between obesity and depression may be limited primarily to the individuals with extreme obesity.

Severity of obesity may in part account for the failure to find an association between obesity and depression in men in some studies, since the prevalence of severe obesity is much lower in men than in women [42].

There is a relationship between the treatment of depression in the presence of obesity and the other way around. Treatment of obesity often leads to a decrease in depression, for example, it was observed an improvement in the mood that accompanies the large weight losses achieved through gastric bypass surgeries [43].

In contrast to the positive results of the treatment of obesity on depression, the treatment of depression can have a negative effect on obesity. Rarely has treatment of depression had a stronger impact upon another disorder than it does on obesity [19]. Traditional antidepressants have long been known to produce weight gain.

2.4 Body Mass Index (BMI)

BMI is an index of weight-to-height (kg / m^2). Although it is high levels of body fat that are most closely linked to health risks for overweight people and low levels of fat and lean tissue that pose a risk for underweight people, methods to directly measure these tissues are not practical for widespread use [44].

BMI is not a direct measure of body fat or lean tissue, but it is, to date, the most widely investigated and most useful indicator of health problems that are associated with under and overweight [45].

The relationship between BMI and health risk is independent of height as differences in height are accounted for in the BMI formula.

However, BMI as an indicator of risk, may have limitations for individuals or populations who are very tall or very short or who have very long or short limb lengths in relation to trunk measurement [45].

The relationship between BMI and health risk is also independent of sex. Although women, on average, have higher levels of body fat than men at the same BMI

[45], their risk of health problems does not differ appreciably within the same BMI category.

For young adults who have not reached full growth, or for adults who naturally have a very lean body build, a BMI somewhat less than 18.5 is not necessarily an indication of health problems [44].

2.4.1 BMI in adults

The Canadian body weight classification system uses BMI to identify weight-related health risks.

The Canadian Guidelines for Body Weight Classification in Adults provides a scheme for classifying body weight, as measured by the body mass index (BMI), according to the level of health risk [44].

Table 2.1 Body Mass Index

Classification	**BMI Category**	**Health Risk**
Underweight	<18.5	Increased
Normal Weight	18.5-24.9	Least
Overweight	25.0-29.9	Increased
Obese	≥30.0	High

The weight classification system is appropriate for adults age 18 years and older. Although an adult is defined physiologically as one who has reached his/her full growth potential, for practical purposes age 18 has been used as the lower age limit, as full growth will be reached by the majority of individuals by this age. An upper age limit has not been established [44]. However, limitations regarding weight classification for seniors, age 65 years and older, are described later.

In addition, there are limitations to weight classification that arise from the wide variation in body builds and body proportions in individuals and in populations.

Finally, the weight classification system is not appropriate for use with pregnant and lactating women. The weight classification system can be applied to both populations and individuals in order to:

1. Conduct meaningful comparisons of body weight patterns within and between populations;
2. Identify populations and individuals at increased risk of morbidity and mortality;
3. Identify priorities for intervention at the population and individual levels;
4. Evaluate interventions at the population and individual levels.

2.4.2 BMI in adults over 65 years of age

Research suggests that for adults age 65 years and older the *normal* range may begin somewhat above BMI 18.5 and extend into the *overweight* range (BMI 25.0 to 29.9). With regard to underweight, the point at which the health risks begin to increase for adults over age 65 has been shown in some studies to be at BMIs ranging from around 18.5 to the low 20s [46]. It has been suggested that these risks may be linked to unrecognized underlying disease. Nevertheless, BMIs in the low 20s and below should be used with some flexibility. Further assessment procedures will be needed to clarify health risks in individual seniors.

2.4.3 BMI in children

Arbitrary centile cut-offs of selected reference data have been used to identify overweight or obese children.

International comparisons of the prevalence of overweight and obesity have been hampered in the past by the lack of a standard classification system.

The International Obesity Task Force worked on the development of such a system and preliminary reports of this work have been published [47, 48].

The new criteria differ from previous arbitrary centile cut-offs because they are linked to the adult health-related definitions of overweight and obesity. There were identified the centiles corresponding to overweight (BMI 25 kg/m^2) and obesity (BMI 30 kg/m^2) in 18-year-old males and females from six different population samples, and projected and smoothed these percentiles by least mean square (LMS) regression back into childhood and adolescence, pooling the results across countries.

In "Supplement to Health Reports, Statistics Canada", 2003, the young respondents were classified as overweight or obese according to the age- and sex-specific BMI cut-offs defined by Cole et al.[48], so the new classification system was adopted in Canada.

Overweight, therefore, was defined as having a BMI that falls within the 85th to 95th centile curves modeled by Cole et al. for children of the same age and sex; obese, as having a BMI that falls at or over the 95th percentile within the age and sex group.

For example, a 13-year-old girl who was 160 cm (5 feet, 3 inches) tall would be considered obese if she weighed 73 kg (161 pounds) (BMI = 28.5) (see Appendix 1).

Chapter 3

Introduction to SEM

Structural equation modeling, or SEM, is a very general, chiefly linear statistical modeling technique. Factor analysis, path analysis and regression, all represent special cases of SEM.

SEM is a largely confirmatory, rather than exploratory, technique. That is, a researcher is more likely to use SEM to determine whether a certain model is valid, rather than using SEM to "find" a suitable model, although SEM analysis often involves a certain exploratory element.

It is a statistical methodology that takes a confirmatory (hypothesis -testing) approach of a structural theory bearing on some phenomena. This theory represents "causal" processes that generate observations on multiple variables [49].

The term structural equation modeling conveys two important aspects of the procedure:

- that the causal processes under study are represented by a series of structural (regression) equations and
- that these structural relations can be modeled pictorially to enable a clearer conceptualization of the theory under study.

The hypothesized model can then be tested statistically in a simultaneous analysis of the entire system of variables to determine the extent to which it is consistent with the data.

In this project we apply the structural equation modeling method to analyze the relationship between depression and obesity in a sample of residents from the province of Ontario, Canada. The analysis will be performed using the AMOS software [50]. There are many software packages for SEM. LISREL, which is marketed by SPSS Inc., used to be the most recommended program, but it was replaced with AMOS. Many SEM programs, including LISREL, require a strong background in matrix algebra and programming skills. On the contrary, AMOS is GUI-based (Graphical User Interface). You can literally draw a model on the program canvas, enter the data, and then test the model.

3.1 Basics

SEM, also called *causal modeling*, *latent variable structural equation* (LVSE) modeling and *analysis of covariance structures*, is a method for representing, estimating and testing a theoretical network of mostly linear relations between variables where those variables may be either observable or directly unobservable, and may only be measured imperfectly.

SEM is a generalization of both regression and factor analysis, and subsumes most linear modeling methods as "special cases".

The procedure emphasizes covariances rather than cases. The fundamental hypothesis for these structural equation procedures is that the covariance matrix of the observed variables is a function of a set of parameters. If we assume that the model is correct and the parameters are known, then the population covariance matrix would be exactly reproduced by SEM (except for sampling variation).

Multiple regression is used to identify and estimate the amount of variance in a single dependent variable attributed to one or more independent variables. This method basically determines the overall contribution of a set of observed variables to prediction, tests full and restricted models for the significant contribution of a variable in a model, or delineates the best subset of multiple independent predictors.

However, it is not robust to measurement error and model misspecification. Therefore we need path analysis.

Path analysis is a subset of structural equation modeling, the multivariate procedure that, as defined by Ullman [51], "allows examination of a set of relationships between one or more independent variables, either continuous or discrete, and one or more dependent variables, either continuous or discrete".

The general SEM model can be decomposed into two submodels: a measurement model and a structural model.

The measurement model defines relations between the observed and unobserved variables. In other words it provides the link between scores on a measuring instrument (the observed indicator variables) and the underlying constructs they are designed to measure (the unobserved latent variables). It specifies the pattern by which each measure loads on a particular factor.

The structural model defines relations among the unobserved variables. It specifies the pattern by which particular latent variables directly or indirectly influence ("cause") changes in the values of certain other latent variables in the model.

3.2 General characteristics of SEMs

Latent versus Observed Variables

In some fields researchers are often interested in studying theoretical constructs that cannot be observed directly. These abstracts phenomena are termed latent variables, or factors. Because latent variables are not observed directly, it follows that they cannot be measured directly. Thus, the researcher must operationally define the latent variable of interest in terms of behavior believed to represent it. So the unobserved variable is linked to one that is observable, thereby making its measurement possible. These measurements are termed observed or manifest variables. Within the context of SEM terminology, they serve as *indicators* of the underlying constructs that they are presumed to represent.

By convention, when graphically representing the model the observed variables are enclosed by rectangles or squares and latent variables are enclosed by ovals or circles. Residuals are always unobserved, so they are represented by ovals or circles.

Exogenous versus Endogenous Latent Variables

It is helpful in working with SEM models to distinguish between latent variables that are exogenous and those that are endogenous. Exogenous latent variables are synonymous with independent variables; they "cause" fluctuations in the values of other latent variables in the model. Changes in the values of exogenous variables are not explained by the model. Rather, they are considered to be influenced by other factors that are external to the model. Background variables such as gender, age or socioeconomic status are examples of such external factors.

Endogenous latent variables are synonymous with dependent variables and are influenced by the exogenous variables in the model, directly or indirectly. Fluctuation in the values of endogenous variables is said to be explained by the model because all latent variables that influence them are included in the model specification.

Cause versus Effect Indicators

Cause indicators are observed variables that are assumed to "cause" a latent variable. Indicators depend on the latent variable, i.e. the latent variable determines its indicators. Most researchers in the social sciences assume that indicators are effect indicators, since it is not easy to distinguish them sometimes. Causal effects are represented by single-headed arrows in the path diagram.

Standard versus Non-Standard Model

With standard models, all constructs that constitute the structural portion of the model are presented as latent variables with multiple manifest indicators.

With a nonstandard model, at least one of the constructs that constitute the structural portion of the model if represented as a single manifest variable.

Recursive versus Nonrecursive Model

The two major types of structural equations with observed variables are recursive and nonrecursive ones. Recursive models are systems of equations that contain no reciprocal causation or feedback loops.

The structural model written in the matrix equation

In essence, SEM is a multivariate extension of the multiple linear regression model with one dependent (Y) variable:

$$Y = i + XB + e,$$

where Y is a vector containing observed scores on the dependent variable, i is a vector representing the Y-intercept, X is a matrix of continuously distributed or categorical (dummy-coded) independent variables, B is the vector of regression weights, and e represents the vector of residual or error or leftover scoring unexplained by the model.

SEMs consist of series of multiple regression equations - all equations are fitted simultaneously. In fact, you can conduct a multiple regression analysis using SEM software. A structural model specifying relations among the latent variables will be constructed. The structural model letting the latent dependent variables to be regressed on latent explanatory variables is specified in the following matrix equation:

$$\eta = B\eta + \Gamma\xi + \varsigma$$

where

η is a *m*-vector of latent endogenous variables;

ξ is a *n*-vector of latent exogenous variables;

ς is a *m*-vector of latent residual variables;

B is a *m* x *m* matrix of path coefficients between latent endogenous (η) variables;

B has zeros on the diagonal which means that we assume that the variable is not an immediate and instantaneous cause of itself, and also the model assumes that *I-B* is nonsingular so that $(I-B)^{-1}$ exists;

Γ is a *m* x *n* matrix of path coefficients between ξ and η variables;

(ς, ξ) are uncorrelated.

The measurement model written in the matrix equation

The measurement model for the dependent variables written in vector notation is

$$y = \Lambda_y \eta + \varepsilon$$

where

y is a *p*-vector of manifest indicator variables for η variables;

Λ_y is a *p* x *m* matrix of factor loadings of *y* on η;

ε is a *p*-vector of latent error variables.

The measurement model for the explanatory variables can be written as

$$x = \Lambda_x \varepsilon + \delta$$

where

x is a *q*-vector of manifest indicator variables for ξ variables;

Λ_x is a q x n matrix of factor loadings of x on ξ;

δ is a q-vector of latent error variables.

However, in applied work, structural equation models are most often represented graphically.

Lets denote by Σ the population covariance matrix of y and x and by $\Sigma(\theta)$ the covariance matrix written as a function of the free model parameters in Σ. Because the population covariance matrix, Σ, is not known in practice, it is replaced by the sample covariance matrix denoted by S. On average, the sample covariance matrix S equals the population covariance matrix Σ.

The basic hypothesis of the general structural equation model is: $\Sigma = \Sigma(\theta)$. Thus, the essence of SEM is to determine the fit between the restricted covariance matrix implied by the hypothesized model and the sample covariance matrix S.

The implied covariance matrix for the general structural equation model as a function of the model parameters, $\Sigma(\theta)$, is made up by the covariance matrix of the observed y variables, the covariance matrix of x and y and the covariance matrix of x. If we denote by $\Sigma_{yy}(\theta)$ the implied covariance matrix of y written as a function of the unknown model parameters that are stacked in the vector θ we have that:

$$\Sigma_{yy}(\theta) = E(yy') = E[(\Lambda_y \eta + \varepsilon)(\Lambda_y \eta + \varepsilon)'] =$$

$$E[(\Lambda_y \eta + \varepsilon)(\eta' \Lambda'_y + \varepsilon')] =$$

$$E(\Lambda_y \eta \eta' \Lambda'_y + \Lambda_y \eta \varepsilon' + \varepsilon \eta' \Lambda'_y + \varepsilon \varepsilon') =$$

$$\Lambda_y E(\eta \eta') \Lambda'_y + \Lambda_y E(\eta \varepsilon') + E(\varepsilon \eta') \Lambda'_y + E(\varepsilon \varepsilon') =$$

$$\Lambda_y E(\eta\eta')\Lambda'_y + E(\varepsilon\varepsilon') = \Lambda_y E(\eta\eta')\Lambda'_y + \Theta_\varepsilon$$

where Θ_ε is the covariance matrix of ε variables.

But

$$\eta = (I - B)^{-1}(\Gamma\xi + \varsigma),$$

hence we get:

$$\Sigma_{yy}(\theta) = \Lambda_y E[(I - B)^{-1}(\Gamma\xi + \varsigma)((I - B)^{-1}(\Gamma\xi + \varsigma))']\Lambda'_y + \Theta_\varepsilon =$$

$$\Lambda_y (E((I - B)^{-1}(\Gamma\xi + \varsigma)(\xi'\Gamma' + \varsigma')((I - B)^{-1})')\Lambda'_y + \Theta_\varepsilon =$$

$$\Lambda_y (I - B)^{-1} E((\Gamma\xi + \varsigma)(\xi'\Gamma' + \varsigma'))((I - B)^{-1})'\Lambda'_y + \Theta_\varepsilon =$$

$$\Lambda_y (I - B)^{-1} E(\Gamma\xi\xi'\Gamma + \Gamma\xi\varsigma' + \varsigma\xi'\Gamma' + \varsigma\varsigma')((I - B)^{-1})'\Lambda'_y + \Theta_\varepsilon =$$

$$\Lambda_y (I - B)^{-1}(\Gamma E(\xi\xi')\Gamma' + \Gamma E(\xi\varsigma') + E(\varsigma\xi')\Gamma' + E(\varsigma\varsigma'))\Lambda'_y + \Theta_\varepsilon =$$

$$\Lambda_y (I - B)^{-1}(\Gamma\Phi\Gamma' + \Psi)((I - B)^{-1})'\Lambda'_y + \Theta_\varepsilon$$

where $\Psi = E(\varsigma\varsigma')$ is the covariance matrix of ς variables and $\Phi = E(\xi\xi')$ is the covariance matrix of ξ variables.

If we denote now by $\Sigma_{yx}(\theta)$ the implied covariance matrix of y and x written as a function of the unknown model parameters that are stacked in the vector θ we have that:

$$\Sigma_{yx}(\theta) = E(yx') = E[(\Lambda_y\eta + \varepsilon)(\Lambda_x\xi + \delta)'] =$$

$$E[(\Lambda_y\eta + \varepsilon)(\xi'\Lambda'_x + \delta')] =$$

$$E(\Lambda_y\eta\xi'\Lambda'_x + \Lambda_y\eta\delta' + \varepsilon\xi'\Lambda'_x + \varepsilon\delta') =$$

$$\Lambda_y E(\eta\xi')\Lambda'_x + \Lambda_y E(\eta\delta') + E(\varepsilon\xi')\Lambda'_x + E(\varepsilon\delta') =$$

$$= \Lambda_y E(\eta \xi')\Lambda'_x.$$

Because $\eta = (I-B)^{-1}(\Gamma\xi + \varsigma)$, we get:

$$\Lambda_y E((I-B)^{-1}(\Gamma\xi + \varsigma)\xi')\Lambda'_x =$$

$$\Lambda_y (I-B)^{-1} E(\Gamma\xi\xi' + \varsigma\xi')\Lambda'_x =$$

$$\Lambda_y (I-B)^{-1} \Gamma E(\xi\xi')\Lambda'_x =$$

$$= \Lambda_y (I-B)^{-1} \Gamma\Phi\Lambda'_x.$$

$\Sigma_{xy}(\theta)$ is the transpose of $\Sigma_{yx}(\theta)$, hence

$$\Sigma_{xy}(\theta) = \Lambda_x \Phi\Gamma'((I-B)^{-1})'\Lambda'_y.$$

The last one is $\Sigma_{xx}(\theta)$ for which we get:

$$\Sigma_{xx}(\theta) = E(xx') = E[(\Lambda_x\xi + \delta)(\Lambda_x\xi + \delta)'] =$$

$$E[(\Lambda_x\xi + \delta)(\xi'\Lambda'_x + \delta')] =$$

$$E(\Lambda_x\xi\xi'\Lambda'_x + \Lambda_x\xi\delta' + \delta\xi'\Lambda'_x + \delta\delta') =$$

$$\Lambda_x E(\xi\xi')\Lambda'_x + \Lambda_x E(\xi\delta') + E(\delta\xi')\Lambda'_x + E(\delta\delta') =$$

$$= \Lambda_x \Phi\Lambda'_x + \Theta_\delta$$

We put all together into a single matrix:

$$\Sigma(\theta) = \begin{pmatrix} \Lambda_y (I-B)^{-1}(\Gamma\Phi\Gamma' + \Psi)((I-B)^{-1})'\Lambda'_y + \Theta_\varepsilon & \Lambda_y (I-B)^{-1}\Gamma\Phi\Lambda'_x \\ \Lambda_x \Phi\Gamma'((I-B)^{-1})'\Lambda'_y & \Lambda_x\Phi\Lambda'_x + \Theta_\delta \end{pmatrix}.$$

Estimation proceeds by selecting values for the unknown parameters θ so that $\Sigma(\theta)$ matches the covariance matrix of the observed variables. Attempts to find those values introduce the concept of identification.

3.3 Identification and Parameter Estimation

Identification is a structural or mathematical requirement in order for the SEM analysis to take place.

SEM programs require an adequate number of known correlations or covariances as inputs in order to generate a sensible set of results. An additional requirement is that each equation be properly identified.

Identification refers to the idea that there is at least one unique solution for each parameter estimate in a SEM model. Models in which there is only one possible solution for each parameter estimate are said to be *just-identified.* Models for which there are an infinite number of possible parameter estimate values are said to be underidentified. Models that have more than one possible solution (but one best or optimal solution) for each parameter estimate are considered *overidentified.*

A just-identified model is one in which there is a one-to-one correspondence between the data and the structural parameters. That is to say, the number of data variances and covariances equals the number of parameters to be estimated. However, despite the capability of the model to yield a unique solution for all parameters, the just-identified model is not scientifically interesting because it has no degrees of freedom and therefore can never be rejected.

An overidentified model is one in which the number of estimable parameters is less than the number of data points (variances, covariances of the observed variables). This situation results in positive degrees of freedom that allow for rejection of the model, thereby rendering it of scientific use. The aim in SEM, then, it is to specify a model such that it meets the criterion of overidentified.

Finally, an underidentified model is one in which the number of parameters to be estimated exceeds the number of variances and covariances (data points). As such, the model contains insufficient information (from the input data) for the purpose of attaining a determinate solution of parameter estimation; that is an infinite number of solutions are possible for an underidentified model.

A number of rules can be used to assess the identification level of your models, but these rules are not perfect, and they are very difficult (almost impossible, in fact) to evaluate by hand, especially for complex models.

SEM software programs such as AMOS perform identification checks as part of the model fitting process. They usually provide reasonable warnings about underidentification conditions.

An additional complication that can arise is empirical underidentification. Empirical underidentification occurs when a parameter estimate that establishes model identification has a very small (close to zero) estimate. When the SEM program performs its matrix inversion, that parameter estimate may drop from the solution space defined by the list of model parameters, and the program thus suddenly detects what it perceives to be a structural underidentification problem. Due to the iterative nature of SEM estimation, a parameter estimate such as a variance may start out with a positive value and gradually approach zero with each successive iteration.

For example, a path coefficient whose value is estimated as being close to zero may be treated as zero by the SEM program's matrix inversion algorithm. If that path coefficient is necessary to identify the model, the model thus becomes underidentified.

The remedy for all forms of underidentification is to try to locate the source of the identification problem and determine if the source is empirical underidentification or structural underidentification. For structural underidentification, the only remedy is to respecify the model. Empirical underidentification may be correctable by collecting more data or respecifying the model.

In order to determine how many data points we have to work with we count how many variables we have (p) , then the number of data points is $p(p+1)/2$ (variances and covariances). The number of the data points minus the parameters to be estimated gives us the number of *degrees of freedom*.

SEM proceeds by assessing whether a sample covariance or correlation matrix is consistent with a hypothetical matrix implied by the model specified by the user.

The inputs of SEM are either raw data or sample moments computed from the data, and a model to be evaluated. The sample moments will include either variances and covariances or correlations, and may also include means and higher order moments as well. The model consists of a network of proposed equations, with some parameters fixed to particular values and others "free to be estimated".

The output from SEM falls into five general groups:

- estimates of the designated model parameters;
- estimates of the standard errors for the estimated parameters;

- for the dependent variables, estimates of the proportion of variance explained, often called squared multiple correlations (SMCs), which are akin to the R^2 statistic in regression;
- overall goodness of fit statistics, which assess the overall consistency between the specified model and the data;
- diagnostic statistics, which aid in pinpointing the sources of any fit problems.

For estimation, start values of the free parameters are chosen in order to generate an estimated population covariance matrix, $\Sigma(\theta)$, from the model. Start values can be chosen by the researcher from prior information, by computer programs used to build SEM or from multiple regression analysis.

The goal of estimation is to produce a $\Sigma(\theta)$ that converges upon the observed population covariance matrix, *S*, with the residual matrix (the difference between $\Sigma(\theta)$ and *S*) being minimized. Various methods can be used to generate $\Sigma(\theta)$. Choice of method is guided by characteristics of the data including sample size and distribution. Most processes used are iterative.

The general form of the minimization function is:

$$Q = (s - \sigma(\theta))'W(s - \sigma(\theta))$$

where

s is the vector containing the variances and covariances of the observed variables;

$\sigma(\theta)$ is the vector containing corresponding variances and covariances as predicted by the model and W is the weight matrix. Some authors refer to Q as F.

The weight matrix, W, in the function above, corresponds to the estimation method chosen. W is chosen to minimize Q, and $Q(N-1)$ gives the fitting function, in most cases a χ^2 distributed statistic. The performance of the χ^2 is affected by sample size, error distribution, factor distribution, and the assumption that factors and errors are independent [51].

Some of the most commonly used estimation methods are:

Generalized Least Squares (GLS)

$$Q_{GLS} = 1/2tr[([S-\Sigma(\theta)]W^{-1})^2]$$

where

tr is the trace operator, takes sum of elements on main diagonal of matrix;

W^{-1} is the optimal weight matrix, must be selected by researcher (most common choice is S^{-1}).

Maximum Likelihood (ML)

$$Q_{ML} = \log|\Sigma(\theta)| - \log|S| + tr(S\Sigma^{-1}(\theta)) - p$$

in this case, $W = \Sigma^{-1}$ and p is the number of measured variables.

Ullman discusses some advantages and limitations of the above estimators [51]. ML and GLS are useful for normally distributed data when factors and errors are independent. Whatever function is chosen, the desired result of the estimation process is to obtain a fitting function that is close to 0. A fitting function score of 0 implies that the model's estimated covariance matrix and the original sample covariance matrix are equal.

In this project, for the analyses which were conducted using the AMOS software, by default, the estimation of parameters will be based on the maximum likelihood (ML) method. It is important to note that use of the ML estimation assumes that the following conditions have been met:

- The sample is very large;
- The distribution of the observed variables is multivariate normal;
- The hypothesized model is valid;
- The scale of the observed variables is continuous.

This last condition has been the subject of debates over the past years. The debate is on the treatment of the ordinally scaled variables as if they were of a continuous scale. At the present time, AMOS does not yet offer researchers the option of analyzing data with the categorical nature of the variables taken into account. That is why we use Mplus.

3.4 Limitations

Theoretical Issues

SEM is a confirmatory technique in contrast to exploratory factor analysis. It is used most often to test a theory. One cannot do SEM without prior knowledge of potential relationships among variables. Although SEM is a confirmatory technique, there are ways to test a variety of different models (models that test specific hypotheses, or perhaps provide better fit) after a model has been estimated.

Structural equation modeling is sometimes refered to as causal modeling. However, there is nothing causal about the use of SEM, in the sense of inferring causality. Attributing causality is a design issue, not a statistical one. "However convincing, respectable and reasonable a path diagram and its associated model may appear, any causal inferences extracted are rarely more than a form of statistical fantasy. Essentially, the so-called causal models simply provide a parsimonious description of a set of correlations" [52].

Practical Issues

Covariances, like correlations, are less stable when estimated from small samples. SEM is based on covariances. Parameter estimates and chi-square tests of fit are also very sensitive to sample size. Hence SEM is a large-sample technique.

SEM examines only linear relationships among variables. Linearity among latent variables is difficult to assess; however, linear relationships among pairs of measured variables can be assessed through inspection of scatterplots. If nonlinear relationships among measured variables are hypothesized, these relationships are included by raising the measured variables to powers, as in multiple regression.

As with other techniques, matrices need to be inverted in SEM. Therefore, if variables are perfect linear combinations of one another or are extremely highly correlated, the necessary matrices can not be inverted. An extremely small determinant of the covariance matrix may indicate a problem with multicollinearity or singularity. Usually SEM programs abort and provide warning messages if the covariance matrix is singular.

After model estimation, the residuals should be small and centered around zero. The frequency distribution of the residual covariances should be symmetric.

Nonsymmetrically distributed residuals may signal a poor-fitting model; the model is estimating some of the covariances well and others poorly. It sometimes happens that one or two residuals remain quite large although the model fits reasonably well and the residuals appear to be symmetrically distributed and centered to zero. When large residuals are found it is often helpful to examine the Lagrange Multiplier test and consider adding paths to the model.

3.5 Appropriate Handling of Incomplete Data

Typical ad hoc solutions to missing data problems include *listwise* deletion of cases, where an entire case's record is deleted if the case has one or more missing data points, and *pairwise* data deletion, where bivariate correlations are computed only on cases with available data. Pairwise deletion results in different N's for each bivariate covariance or correlation in the database. Another typically used ad hoc missing data handling technique is substitution of the variable's mean for the missing data points on that variable.

But none of these ad hoc missing data handling methods are appealing from a statistical point of view. Listwise deletion can result in a substantial loss of power, particularly if many cases each have a few data points missing on a variety of variables, not to mention limiting statistical inference to individuals who complete all measures in the database. Pairwise deletion is marginally better, but the consequences of using different n's for each covariance or correlation can have profound consequences for model fitting efforts, including impossible solutions in some instances. Finally, mean substitution will shrink the variances of the variables where mean substitution took place, which is not desirable.

The most important problem with these methods is that they assume that the missing data are missing completely at random, which is often not the case.

If the proportion of cases with missing data is small, say five percent or less, listwise deletion may be acceptable [53]. Of course, if the five percent (or fewer) cases are not missing completely at random, inconsistent parameter estimates can result. Otherwise, missing data experts recommend using a maximum likelihood estimation method for analysis, a method that makes use of all available data points [52].

3.6 Assessing the Fit of the Model

After the model has been specified and then estimated, the major concern is if the model is a "good" one.

Evaluation of model fit should derive from a variety of perspectives and be based on several criteria that can assess model fit from a diversity of perspectives. In particular, this focuses on the adequacy of (a) the parameter estimates and (b) the model as a whole.

a) For the fit of individual parameters in the model there are three aspects of concern: (i) the feasibility of the parameter estimates, (ii) the appropriateness of the standard errors and (iii) the statistical significance of the parameter estimates.

(i) Feasibility of Parameter Estimates

The initial step in assessing the fit of individual parameters in a model is to determine the viability of their estimated values. Parameter estimates should exhibit the correct sign and size and be consistent with the underlying theory. Any estimate falling outside the admissible range signals a clear indication that either the model is wrong or the input matrix lacks of sufficient information. Examples of parameters exhibiting unreasonable estimates are correlations >1.00, negative variances, and covariance or correlation matrices that are not positive definite.

(ii) Appropriateness of Standard Errors

Another indicator of poor model fit is the presence of standard errors that are excessively large or small. For example, if a standard error approaches zero, the test statistic for its related parameter cannot be defined. Likewise, standard errors that are extremely large indicate parameters that cannot be determined. Because standard errors are influenced by the units of measurement in observed and/or latent variables, as well as the magnitude of the parameter estimate itself, no definitive criterion of "small" and "large" has been established.

(iii) Statistical Significance of Parameter Estimates

The test statistic here is the critical ratio (c. r.) which represents the parameter estimate divided by its standard error, as such it operates as a *z*-statistic in testing that the estimate is statistically different from zero. Based on a level of 0.05, the test statistic needs to be $\leq -1.96, \geq 1.96$ before the hypothesis (that the estimate equals 0.0) can be rejected. Nonsignificant parameters, with the exception of error variances, can be considered unimportant to the model; in the interest of scientific parsimony, albeit given an adequate sample size, they should be deleted from the model. On the other hand, it is important to note that nonsignificant parameters can be indicative of a sample size that is too small.

b) We look at the fit between the sample covariance matrix and the estimated population covariance matrix. In SEM, the null hypothesis H_0 being tested, that the postulated model holds in the population, is hoped not to be rejected, in contrast to traditional statistical procedures. A good fit is sometimes indicated by a nonsignificant χ^2. Unfortunately, assessment of fit is not always as straightforward as assessment of χ^2. In the case of large samples, trivial differences between sample and estimated population covariance matrices are often significant because the minimum of the function is multiplied by N-1. In the case of small samples, the computed χ^2 may not be distributed as χ^2, leading to inaccurate probability levels.

A "rule of thumb" related to the χ^2 value is that a good-fitting model may be indicated when the ratio of the χ^2 to the degree of freedom is less than 2.

Because of these problems, a number of measures of model fit have been proposed. In AMOS output, for each set of fit statistics we have three rows; the first one focuses on the hypothesized model under test, the second on the saturated model, and the third on the independence model. This last model is one of complete independence of all variables in the model (i.e. in which all correlations among variables are zero), and is the most restricted. The saturated model, on the other hand, is one in which the number of estimated parameters equals the number of data points (i.e. variances and covariances of the observed variables) and is the least restricted.

Comparative Fit Indices

The *normed fit index* (NFI) evaluates the estimated model by comparing the χ^2 value of the model to the χ^2 value of the independence model:

$$\text{NFI} = \frac{\chi^2_{\text{indep}} - \chi^2_{\text{model}}}{\chi^2_{\text{indep}}}$$

This yields a descriptive index with values between 0 and 1. High values (greater than 0.90) are indicative of a good-fitting model. NFI can underestimate the fit of the model in good-fitting models with small samples.

The *comparative fit index* (CFI) also assesses fit relative to other models as the name implies, but uses a different approach. The CFI employs the noncentral χ^2 distribution with noncentrality parameters, τ_i. The larger the value of τ_i, the greater the model misspecification, i.e. if the estimated model is perfect, then $\tau_i = 0$. The CFI is defined as:

$$\text{CFI} = 1 - \frac{\tau_{\text{model}}}{\tau_{\text{indep}}}$$

with

$$\tau_{\text{indep}} = \chi^2_{\text{indep}} - df_{\text{indep}}$$

$$\tau_{\text{model}} = \chi^2_{\text{model}} - df_{\text{model}}$$

CFI values greater than 0.95 are often indicative of good-fitting models. The CFI is normed to the 0-1 range and does a good job in estimating model fit even in small samples. The values of all these indices depend on the estimation method used.

The *root mean square error of approximation* (RMSEA) estimates the lack of fit in a model compared to a saturated model. The equation for the estimated RMSEA is given by:

$$\text{estimated RMSEA} = \sqrt{\frac{\hat{F}_0}{df_{\text{model}}}}$$

where $\hat{F}_0 = \frac{\chi^2_{\text{model}} - df_{\text{model}}}{N}$ or 0 whichever is smaller but positive. When the model is perfect then $\hat{F}_0 = 0$. The greater the model misspecification the larger $\hat{F}_0$. Values of RMSEA of 0.06 or less indicate a good-fitting model relative to the model degrees of freedom.

Indices of Proportion of Variance Accounted

The *goodness-of-fit* index (GFI) can be defined by:

$$\text{GFI} = \frac{tr(\hat{\sigma}' W \hat{\sigma})}{tr(s' W s)}$$

where the numerator is the sum of the weighted variances from the estimated model covariance matrix and the denominator is the sum of the squared weighted variances from the sample covariance. W is the weight matrix that is selected by the choice of estimation method.

This fit index can be adjusted for the number of parameters estimated in the model. The adjusted *fit index* (AGFI) is estimated by,

$$\text{AGFI} = 1 - \frac{1 - \text{GFI}}{1 - \frac{\text{Number of est. parameters}}{\text{Number of data points}}}$$

The fewer the number of estimated parameters relative to the number of data points, the closer the AGFI is to the GFI. The fit improves by estimating lots of parameters in SEM.

However, the second goal of modeling is to develop a parsimonious model with as few parameters as possible.

Degree of Parsimony Fit Indices

An adjustment can be made to the GFI so that it takes into account the degree of parsimony in the model.

$$\text{PGFI} = [1 - (\frac{\text{Number of est. parameters}}{\text{Number of data points}})]\text{GFI}$$

The larger the fit index (values closer to 1) the better. This index will always be smaller than other indices unless the number of parameters estimated is much smaller than the number of data points.

Other methods that include a parsimony adjustment are the *Akaike Information criterion* (AIC) and the *Consistent Akaike Information criterion* (CAIC):

$$\text{AIC} = \chi^2_{\text{model}} - 2df_{\text{model}}$$

$$\text{CAIC} = \chi^2_{\text{model}} - (\ln N + 1)df_{\text{model}}.$$

Small values indicate a good-fitting, parsimonious model. This index is applicable to models estimated with maximum likelihood methods.

Residual-Based Fit Indices

The *root mean square residual* (RMR) and the *standardized root mean square residual* (SRMR) are the average differences between the sample variances and covariances and the estimated population variances and covariances. The root mean square is given by:

$$\text{RMR} = [2\sum_{i=1}^{q}\sum_{j=1}^{i}\frac{(s_{ij}-\hat{\sigma}_{ij})^2}{q(q+1)}]^{1/2}$$

The RMR is the square root of two times the sum, over all of the variables in the covariances matrix, of the averaged squared differences between each of the sample covariances (or variances) and the estimated covariances (or variances).

Good fitting implies small RMR. It is difficult sometimes to interpret an unstandardized residual because the scale of the variables affects the size of the residual; therefore, a standardized root mean square residual (SRMR) is also available.
Again, small values indicate good fit. The SRMR has a range of 0 to 1; values of 0.08 or less are desired.

Absolute Fit Index

An index that is absolute in that it does not depend on comparison with another model such as the independence or saturated models (CFI) or the observed data (GFI) has been proposed:

$$\text{MFI} = \exp[-0.5\frac{(\chi^2_{\text{model}} - df_{\text{model}})}{N}].$$

These are some of the indices that are to be found in the output of AMOS. We do not need to report the entire set of fit indices. Choosing those indices that give a good sense of how well the model fits the sample data is not an easy task. In most cases good-fitting models produce consistent results on many different indices. If all the indices lead to similar conclusions, the issue of choosing the indices to report is a matter of personal preference. The CFI and RMSEA are probably the most frequently reported fit indices. The AIC and CAIC indices are helpful to use when comparing models that are not nested. If the results of the fit indices are inconsistent then the model should be re-examined; if the inconsistency cannot be resolved then reporting multiple indices should be considered.

Since it is rare that an initial theoretical model demonstrates a good fit, we need to modify the model to get a better fit. The basic methods of **model modification** are chi-square difference tests and Lagrange multiplier tests (LM). They are asymptotically equivalent under the null hypothesis but approach model modification differently.

Chi-Square Difference Test

If models are nested then the difference between the χ^2 of the smaller nested model and the χ^2 of the larger model is also a χ^2 with the number of degrees of freedom equal to the difference between the degrees of freedom in the two models.

There are some disadvantages to this test:

- two models need to be estimated to get the χ^2 difference value and estimating two models for each parameters can be time consuming with large models or slow computers;
- because of the relationship between the χ^2 and the sample size, it is hard to detect a difference between models if the sample sizes are small.

Lagrange Multiplier Test

This test also compares nested models but requires estimation of only one model. The LM test asks if the model is improved if one or more of the parameters in the model that are currently fixed are estimated. This method is analogous to the forward stepwise regression.

The LM test can be examined either univariately or multivariately. There is a danger in examining only the results of univariate LM tests because overlapping variance between parameter estimates may make several parameters appear as if their addition would significantly improve the model. All of these parameters are candidates for inclusion by the results of the univariate LM tests but the multivariate LM identifies the single parameter that would lead to the largest drop in model χ^2 and calculates the expected change in χ^2.

Standardized residuals

The residuals are not independent of one another, thus any attempts to test them (in the strict statistical sense) would be inappropriate. In essence, only their magnitude is of interest in alerting the researcher to possible areas of model misfit. Both the matrix of unstandardized residuals and that of standardized residuals are presented in the optional softwares output. However, because the fitted residuals are dependent on the unit of measurement of the observed variables, they can be difficult to interpret, thus their standardized values are typically examined. Standardized residuals are fitted residuals divided by their asymptotically (large sample) standard errors. In essence they represent estimates of the number of standard deviations the observed residuals are from the zero residuals that would exist if model fit were perfect $(\Sigma(\theta) - S = 0)$.

Values > 2.58 are considered to be large.

Modification Indices (MI)

For each fixed parameter specified, the softwares provide an MI, the value of which represents the expected drop in overall χ^2 value if the parameter were to be freely estimated in a subsequent run; all freely estimated parameters automatically have MI values equal to zero. Although this decrease in χ^2 is expected to approximate the MI value, the actual difference can be larger. Associated with each MI is an expected parameter change (EPC) value which is reported in the accompanying column labeled Par Change. This statistic represents the predicted estimated change, in either a positive or negative direction, for each fixed parameter in the model and yields important information regarding the sensitivity of the evaluation of the fit to any reparameterization of the model.

The MIs and the accompanying expected parameter change statistics are presented first for possible covariances, followed by those for variances and regression weights. Parameters which are freely estimated do not appear.

Chapter 4

Data exploration

4.1 Variables

Based on the CCHS 1.2 – Mental Health and Well-being dataset, **major depressive episode** is defined as a period of 2 weeks or more with persistent depressed mood and loss of interest or pleasure in normal activities, accompanied by symptoms such as decreased energy, changes in sleep and appetite, impaired concentration, and feelings of guilt, hopelessness, or suicidal thoughts.

The questions on major depressive episode are based on a recognized World Mental Health version of the Composite International Diagnostic Interview (WMH-CIDI) modified for the needs of CCHS 1.2. The WMH-CIDI instrument, as part of the WMH2000 Project (World Mental Health 2000) is a World Health Organization worldwide initiative to assess the prevalence rates of various mental disorders in multiple countries. The WMH-CIDI is a standardized instrument for assessment of mental disorders and conditions according to the definitions and criteria of DSM-IV and ICD-10.

Both **lifetime** and **past year** diagnoses are assessed. Past year episode refers to the 12 months preceding the interview.

For the purposes of this survey, respondents who experienced the following CCHS 1.2/WMH-CIDI criteria associated with major depressive episode, were classified as being affected by **lifetime depression**:

1. a period of two weeks or more with depressed mood or loss of interest or pleasure AND at least five additional symptoms from the following nine;
    - depressed mood;
    - diminished interest in hobbies or activities;
    - significant weight loss/gain or change in appetite;
    - insomnia or hypersomnia;
    - psychomotor agitation or retardation;
    - fatigue or loss of energy;
    - feelings of worthlessness;
    - diminished ability to think or concentrate;
    - recurrent thoughts of death.
2. clinically significant distress or social or occupational impairment; and
3. the symptoms are not better accounted for by bereavement.

Respondents who experienced the following CCHS 1.2/WMH-CIDI criteria associated with major depressive episode were classified as having a **12 month depression**:

1. meet the criteria for lifetime diagnosis of major depressive episode;
2. report a 12-month episode; and
3. report marked impairment in occupational or social functioning.

The literature suggests that there might be a relationship between depression and obesity only for higher levels of severity of the two diseases.

Persistence of Major Depressive Episode (DEPRESS) is the variable that identifies the longest episode associated with a major depressive episode experienced by the respondent. It is calculated only for respondents who meet the criteria for major depressive episode.

Other variables that we are going to use in the analysis are summarized below:

Physical activity - Energy Expenditure (PACBDEE)

In order to derive a physical activity index, the energy expenditure (EE) of participants in their leisure activities was estimated. EE is calculated using the frequency and time per session of the physical activity as well as its metabolic energy cost (MET) value. The MET is a value of metabolic energy cost expressed as a multiple of the resting metabolic rate. Thus, an activity of four METs requires four times the amount of energy as compared to when the body is at rest. CCHS questions did not ask the respondent to specify the intensity level of their activities; therefore the MET values adopted correspond to the low intensity value of each activity. This approach is adopted from the

Canadian Fitness and Lifestyle Research Institute because individuals tend to overestimate the intensity, frequency and duration of their activities.

Derived Work Stress Scale - Social Support (WSTBDSOC)

This variable summarizes the social support available to the respondent at his/her main job in the past 12 months. Questions are asked about whether or not the supervisor and the people the respondent worked with were helpful in getting the job done, and whether the respondent was exposed to hostility or conflict from the people they worked with. It is a derived variable measured on a scale from 0 to 12 with larger values indicating strong social support.

Stress – ability to handle unexpected and difficult problems (STRB_1)

This variable asks the respondent to rate his/her ability to handle unexpected and difficult problems such as a family or personal crisis, on a 5 point scale from excellent to poor.

SES - socioeconomic status (JOB-INC)

SES is a latent variable (construct) derived from the following variables: job status during the past year, multiple job status, income per household, and income per person using factor analysis technique.

Eating habits – Eating Attitudes Test Index Score (ETABIND)

This variable is a measure of the extent of the symptoms and concerns characteristic of eating disorders. The EAT is usually administered to individuals who have expressed or displayed symptoms or ill concerns associated with eating attitudes and behaviours. CCHS 1.2 has two screener questions to assess past 12 months and lifetime concerns with eating attitudes and behaviours. Those who responded affirmatively to the screener were then asked all of the items/questions used to assess the EAT-26. Higher scores indicate higher risk of eating disorder.

Relatives with depression (DEPB_88)

In this variable the number of close relatives – including biological parents, brothers, sisters and children who ever had one or several episodes of being depressed is counted.

There are 12376 subjects in the raw data in an SPSS file. The AMOS software uses SPSS .sav files which meant we did not have to convert our original file into another type of file in order to perform the SEM.

As we mentioned before, in structural equation modeling the researchers often require a larger sample size (much larger than in multiple regression) to maintain the accuracy of the estimates and to ensure representativeness. The need for larger sample sizes is due in part to the program requirements and the multiple observed indicator variables used to define the latent variables (degree of freedom in a measurement model). There is a rule-of-thumb suggesting that the number of responses should be at least five times the number of parameters in order to attain reasonable results [52], where by the number of parameters we mean the path coefficients, variances and covariances to be estimated. The larger the sample size the better.

The sample size of the data we are going to use in this project is large enough so we can use this method of analysis even if we deal with missing data.

From the initial 12376 observations we have 3440 subjects that have some missing values in the case when we use the continuous variable DEPRESS (persistence of major depressive episode (in years)), 3413 missing values for DEPLIFE (lifetime depression) and 3407 for DEPYEAR (12-month depression).

We use the DEPRESS variable in the model analysis done in AMOS as it is a continuous variable, and the other two variables will be introduced in the models analyzed in Mplus.

We first analyze the data descriptively using SPSS software. In the table in **Appendix 2**, the means, variances, standard deviations, the ranges and missing values are shown. We should also check if the data is normal, however when the sample is large ($n > 2500$) the method works well even for non-normal data.

Chapter 5

Structural Equation Modeling of CCHS 1.2 data

We illustrate the analysis using two possible types of variables: continuous and categorical ones. The first part of the analysis will be done using the AMOS software, which requires continuous variables except for exogenous variables such as gender, while the second part of the analysis will be done in Mplus software which brings the possibility of using categorical endogenous variables in the analysis of SEM.

5.1 Continuous variables

Using AMOS software and continuous variables of BMI (Body Mass Index) for obesity and of DEPRESS (Persistence of depression, in years) for depression we propose the following cross-sectional models, as we do not have any available longitudinal data for now.

5.1.1 Model with 2 latent variables: GENES and ENVIRON

The first model that we are going to discuss was proposed by Stunkard et al in their article entitled "Depression and Obesity" [34], in which they underline that there might be a "genetic correlation" and an "environmental correlation" between depression and obesity. The first one refers to the fact that there might be a set of genes that promote both depression and obesity, while the second one underlines the possibility of existence of "common life experiences" that may promote both diseases. In other words, certain genotypes or environmental factors may create a relationship between obesity and depression.

In the figure below, we have a pictorial representation of the model, which even if does not completely explain the physiological pathways connecting the two diseases under study, it represents the model that had been used as an empirical framework in studies of genetic epidemiology [54].

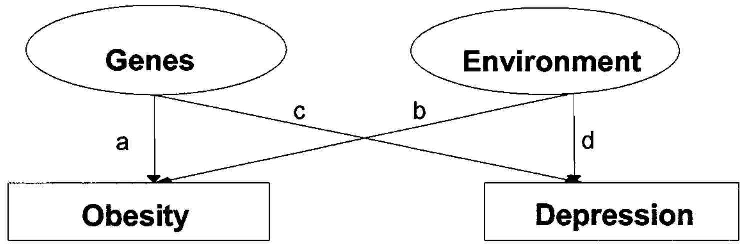

Figure 5.1 Model proposed by Stunkard et al.

Using the information available in the variables contained in our dataset we have proposed the following model (Figure 2).

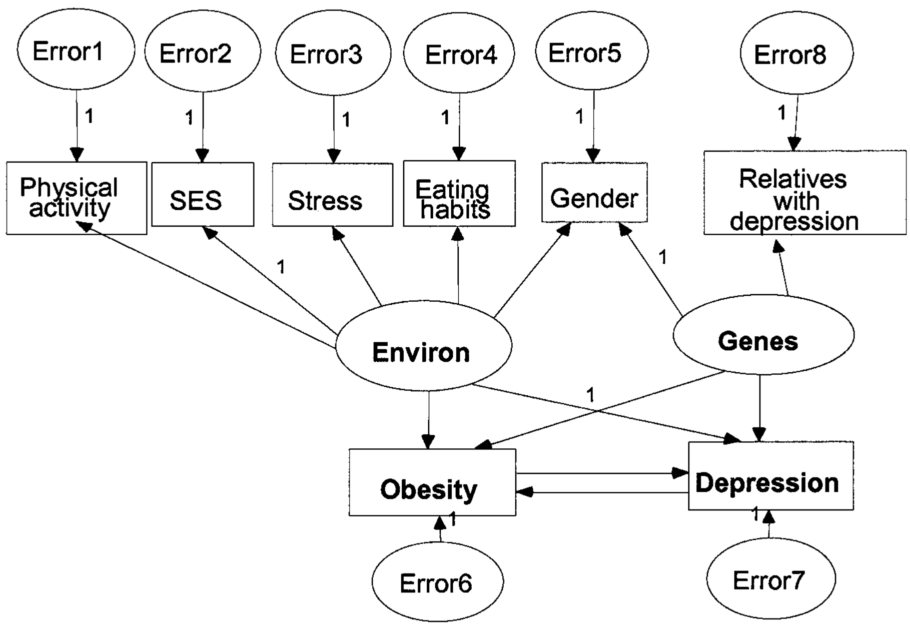

Figure 5.2 Model 1 – two latent variables ENVIRON and GENES

It consists of two unobserved latent variables ENVIRON and GENES as indicated by the ovals (ellipses) and eight observed variables, represented by the rectangles. The observed variables load on the factors in the following pattern:

a) Physical activity/Energy expenditure (PACBDEE), Socioeconomic status (JOB_INC), Stress/Self-perceived ability to handle an unexpected problem (STRB_1), Eating habits (ETABDIND) and Gender (DHHB_SEX) load on the first factor, ENVIRON.

b) Gender (DHHB_SEX) and Relatives with depression (DEPB_88) load on the second factor, namely GENES.

c) both Obesity (HWTBDBMI) and Depression (DEPRESS) regress on the two latent factors.

We also added the errors of measurement associated with each observed variable; and they are uncorrelated.

We notice that in this model the observed variables do not load on only one factor. As we specified above, the gender variable loads on both latent variables.

The model we hypothesize here is a **nonrecursive model**. Namely, it is a model with two structural equations where the dependent variable of each equation appears as a predictor variable in the other equation.

The variables measuring obesity and depression form a feedback loop; meaning that we can follow the path between these two infinitely many times without having to return to the other variables.

Actually, in all the models we will discuss in this paper we hypothesize that obesity is directly influenced by depression and vice versa.

Our models are also **non-standard,** because obesity and depression are two constructs in the structural portion of the model that are represented as a single manifest variable each.

AMOS provides the user with two alternatives of model specification; one within the graphical framework using AMOS Graphics that we have been using until now and the other one within the equation format using AMOS Basic (a more traditional framework; however, we will be not using it in this project).

Provided with the hypothesized model we can now move on to testing it. Descriptive criteria for model fitting were described in more detail in Chapter 3.

After the completion of the analysis, AMOS Graphics will allow us to review the results from three different perspectives: graphical, tabular and textual.

In the graphical output, all estimates are presented in the path diagram.

For Model 1 we cannot show this output because it does not provide an admissible solution. The textual output, selection of which can be seen in Appendix 3, summarizes that the model is nonrecursive and that the analysis was based on 12 376 observations and 18 distinct sample points.

In structural equation modeling of great interest is the extent to which the hypothesized model "fits", describes the data. We look at the adequacy of the parameters estimates and of the model as a whole. The first step in assessing the fit of individual parameters in a model is to determine the viability of their estimated values. Parameters should have the correct sign and size and be consistent with the underlying theory, the standard errors should not be very large nor very small, and the parameter estimates should be statistically significant. The test statistic is the critical ratio (c. r.) and it should be ≤ -1.96 or ≥ 1.96 before the hypothesis can be rejected. Nonsignificant parameters should be deleted from the model.

In the case of our analysis, the results of this first step in assessing the fit of the model show us that we deal with a not very good one, as we have a couple of negative variances (for GENES and ERROR5), some excessively large standard errors , (e.g. the standard error of the regression path from DEPRESS to GENES is 1567.335, with an estimate of 133.188) and nonsignificant parameters (e.g. the regression path between GENES and "(number of) relatives with depression" (DEPB_88), between GENES and DEPRESS, etc). And, of course, the model as a whole does not prove a good fitting either, the indices we discussed in the previous chapters are having values less than the minimum accepted as a marginal fit in the literature.

We decide to drop the second latent variable GENES and have the following Model 2:

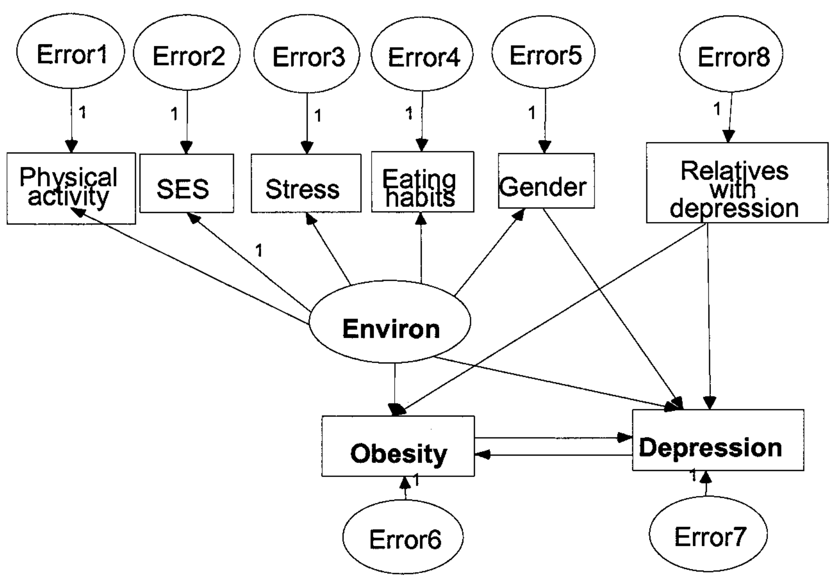

Figure 5.3 Model 2 – one latent variable ENVIRON

However, the model does not improve, as it can be seen in the second part of the Appendix 3. Having these results we decide to drop the second latent variable (GENES) and its indicator, "Relatives with depression" (DEPB_88) as it has a very large number of missing values.

5.1.2 More models with one latent variable: ENVIRON

5.1.2.1 Stress viewed as "Self-perceived ability to handle an unexpected problem" (STRB_1)

The third model that we propose has only one latent variable, namely ENVIRON (Environment) . As indicated by the rectangles, the model also contains 6 measurable, observed variables: Physical activity /Energy expenditure (PACBDEE), Socioeconomic status (JOB_INC), Stress/Self-perceived ability to handle an unexpected problem (STRB_1), Gender (DHHB_SEX), Obesity (HWTBDBMI) and Persistence of depression (DEPRESS) which all load on the ENVIRON factor (Figure 5.4).

Again, we added the errors of measurement associated with each observed variable which are uncorrelated.

As we specified in the previous section all the models that we deal with in the paper are nonrecursive because we assume that obesity is directly influenced by depression and depression is influenced by obesity [34].

Modification indices that help us make the decisions on what can be done with our model in order to improve its fit are not presented in the output of AMOS software for data with missing values like the one we deal with.

In order to get an idea of how to indeed improve the model, we performed the SEM analysis only for the complete data of 8936 participants and with the help of the MI (modification indices) shown in the output we changed the model by adding an arrow from Physical activity /Energy expenditure (PACBDEE) to Obesity (HWTBDBMI) and from Stress (STRB_1) to Depression (DEPRESS).

In other words, we hypothesize that obesity is directly affected by the physical activity (or physical inactivity) and that depression is directly influenced by stress, which makes strong substantive sense and therefore should be included in the model.

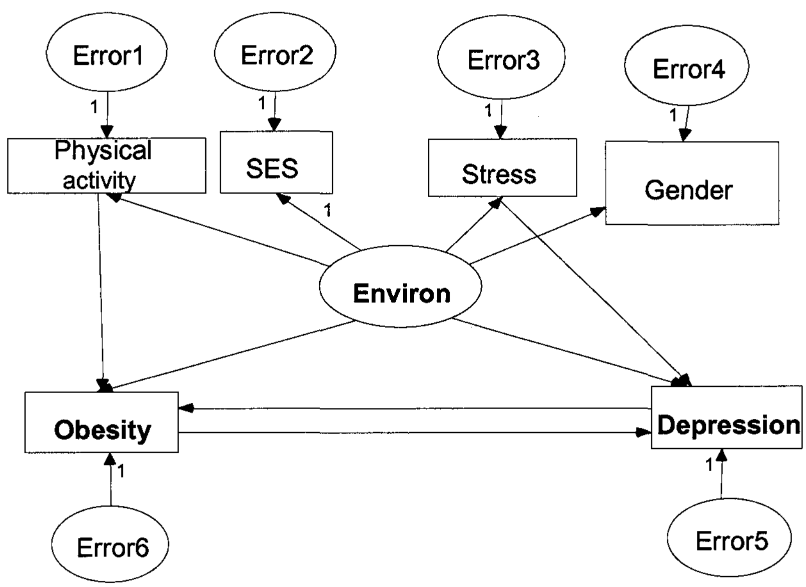

Figure 5.4 Model 3 – one latent variable ENVIRON and 6 measurable ones

We run the program and it seems to perform well. In Appendix 4 the output shows that the analysis was based on 12376 observations and 13 variables (including error terms). The number of distinct sample points was 27 and the parameters to be estimated were 22. The necessary but not sufficient condition for model identification has been met. We have 5 degrees of freedom.

The iteration history indicates that the convergence criterion was satisfied.

We then look at the goodness-of- fit indices and in this case we have a chi-square of 140.238 with 5 degrees of freedom, which is significant. Although a nonsignificant chi-square would have shown support for our model, this significant chi-square does not necessarily indicate a bad fit. We have to look at the other fit indices like CFI, NFI, RMSEA, etc. The values of the fit indices obtained in this analysis are presented in the table below.

Table 5.1 Goodness of Fit Indices for Model 3

MODEL	RMSEA	NFI	CFI
Model 3	0.047	0.895	0.898

The literature suggests that the RMSEA index values ranging from 0.08 to 0.10 indicate mediocre fit, and those greater then 0.10 indicate poor fit. AMOS also reports a 90% confidence interval around the RMSEA value. Presented with a small RMSEA, but a wide confidence interval, a researcher would conclude that the estimated value is quite imprecise, thereby it would be impossible to determine accurately the degree of fit in the population. In contrast, a very narrow confidence interval would argue for good precision of the RMSEA value in reflecting the model fit in the population [55]. In addition to reporting a confidence interval around the RMSEA value, the AMOS program tests for the closeness of fit, meaning that it tests the hypothesis that the RMSEA is "good" in the population (specifically, that it is < 0.05). Joreskog suggested that the p value for this test should be > 0.50 [56].

In our analysis we see that the RMSEA value for our hypothesized model is 0.047, with the 90% confidence interval of (0.040, 0.054) and the p value for the test of closeness of fit is 0.776. The interpretation of the confidence interval indicates that we can be 90% confident that the true RMSEA value in the population will fall within the bounds of 0.040 and 0.054, which represents a good degree of precision.

Given that the RMSEA point estimate is < 0.05 (0.047), that the upper bound of the 90% interval is < 0.05, less than the suggested value of 0.06 in [57] and that the probability value associated with this test of close fit is > 0.50 (0.776), we can conclude that the initially hypothesized model fits the data well. One possible limitation of the RMSEA is that it ignores the complexity of the model [58]. Hence we should have a look at the other indices as well.

The normed fit index (NFI) was the practical criterion of choice for a long period of time but it seemed to underestimate fit in small samples, so in 1990 Bentler [59] revised the NFI to take sample size into consideration and proposed the comparative fit index (CFI). The values for both indices range from 0 to 1 and are derived from comparison of the hypothesized model with the independence model.

A value greater than 0.90 was considered representative of a well-fitting model, Bentler [59] also suggested that the CFI should be the index of choice.
As shown in Table 5.1, the NFI has a value of 0.895 and CFI is 0.898 which do not show a perfect fit, but we can still say that these indices did not show poor fit.

We look now at the factor loadings and path coefficients to check if the standard errors of the estimates are close to zero which would imply that an estimation problem occurred we also check the z- tests to see if we have values greater than 1.96 which would mean that the coefficients are significantly different from zero at the 5% level.

Table 5.2. Regression weights for Model 3

	Estimate	S.E.	C.R.	P
Phys.Activ ← Environ	1.053	.101	10.387	<.001
Stress ← Environ	-.548	.044	-12.420	<.001
Obesity ←Environ	1.738	.256	6.781	<.001
Depression ← Environ	-.963	.141	-6.821	<.001
Obesity ← Phys.Activ	-.252	.021	-11.859	<.001
Depression ← Stress	.274	.034	8.091	<.001
SES ←Environ	1.000			
Gender ← Environ	-.600	.048	-12.385	<.001
Obesity ← Depression	-.282	.116	-2.439	.015
Depression ← Obesity	.150	.045	3.359	<.001

There were no near-zero standard errors for the factor loadings and all of the path coefficients are statistically significant. In other words, the probability of getting a critical ratio (c. r.) by chance as large as the ones observed (in absolute value) is less than .001. The regression weight for our variables in the prediction of another is significantly different from zero.

Of great interest in the analysis are the path coefficients for the "causal" paths that constitute the structural portion of the model. The path coefficients for the path from obesity (HWTBDBMI) to depression (DEPRESS), from depression to obesity, from ENVIRON to obesity and ENVIRON to depression etc. are all significant.

The existence of feedback loops in a nonrecursive model may imply the arising of some certain problems that do not or cannot occur in recursive models. In our model, obesity depends on depression which in turn depends on obesity, and so on; an infinite regress. This infinite sequence of linear dependencies may or may not result in well defined relationships among obesity, depression and the other variables in the model depending on the regression weights [60].

In the case when this infinite sequence of dependencies converges to a set of well-defined relationships, given some values of the regression weights, the system of linear dependencies is called *stable*. Otherwise, it is called unstable.

We cannot tell if a linear system is stable just by looking at the path diagram, but we need the regression weights. AMOS estimates them and also calculates a *stability index*. If this index has a value between -1 and 1 then the system is stable. An unstable one is "impossible"; it implies that our model is wrong or the sample size is too small to provide an accurate estimate of the regression weights. If the hypothesized model has more loops in a path diagram, then AMOS will compute a stability index for each one. If any of the indices is equal or greater to 1 then the system is unstable and should not be modeled in that form. In our example the system is indeed stable, the stability index being equal to 0.042. In the following path diagram the results of the analysis (the regression weights, the intercepts and variances) are displayed (Figure 5.5).

Models can be modified in several ways. For instance, we can either fix paths at zero (eliminating the nonsignificant path from the model) or free paths to be estimated (adding new paths to the model). In our case all the path coefficients are significant which means we do not need to eliminate any. Also we have to keep in mind the parsimony of the model. In this situation, our hypothesized model is the best that can be obtained.

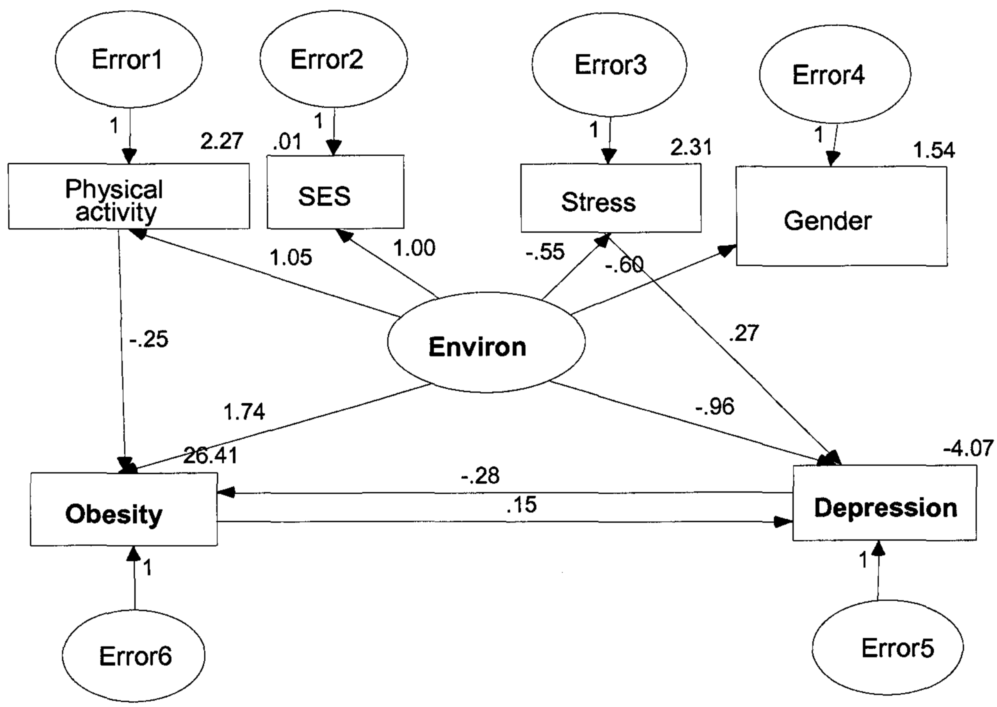

Figure 5.5 Graphical output of Model 3

5.1.2.2 Stress viewed as "Work Stress - social support" (WSTBDSOC)

The fourth model proposed in the present paper will look at the relationship between depression and obesity under the influence of the surrounding environment in which, nowadays, work stress has an important role. The following models and results will only apply to the working population (age between 15 and 75).

The model will look similar to the previous one except the observable variable measuring the work stress- social support, namely WSTBDSOC, is added to the model.

Again we have one latent variable and six observed ones: Physical activity /Energy expenditure (PACBDEE), Socioeconomic status (JOB_INC), Work Stress - social support (WSTBDSOC), Gender (DHHB_SEX) which all load on the ENVIRON factor, Obesity (HWTBDBMI) and Persistence of depression (DEPRESS).

We run the program and start establishing the goodness of fit of the model. The output is shown in the second part of Appendix 4.

The chi-square that was obtained in this case is 80.9 with 5 degrees of freedom, again a significant statistics. As we mentioned before this does not imply the fact that the model does not provide a good fit. We have to look at the other goodness-of-fit indices which might reveal a relatively good fit, even if the chi-square test rejects the model.

In Table 5.3 we present some of those indices: RMSEA, CFI, NFI.

Table 5.3 Goodness of Fit Indices for Model 4

MODEL	RMSEA	NFI	CFI
Model 4	0.035	0.917	0.921

The NFI and CFI should have values over 0.90 in order to have a moderate / acceptable fit and RMSEA value should be less than 0.06 which is the case. We have a better fit than for the previous model, but we still should have a look at the estimates of the regression weights, their statistical significance, the stability index etc.

The path coefficients and factor loadings do not have close to zero standard deviations, the system is stable as the stability index is 0 but this time not all regression weights are statistically significant.

The regression weight for Depression in the prediction of Obesity is not significantly different from zero at the 0.05 level having a critical ratio c. r.= 0 .083 and a p value = 0.934. In Figure 5.6 the graphical output of Model 4 is presented.

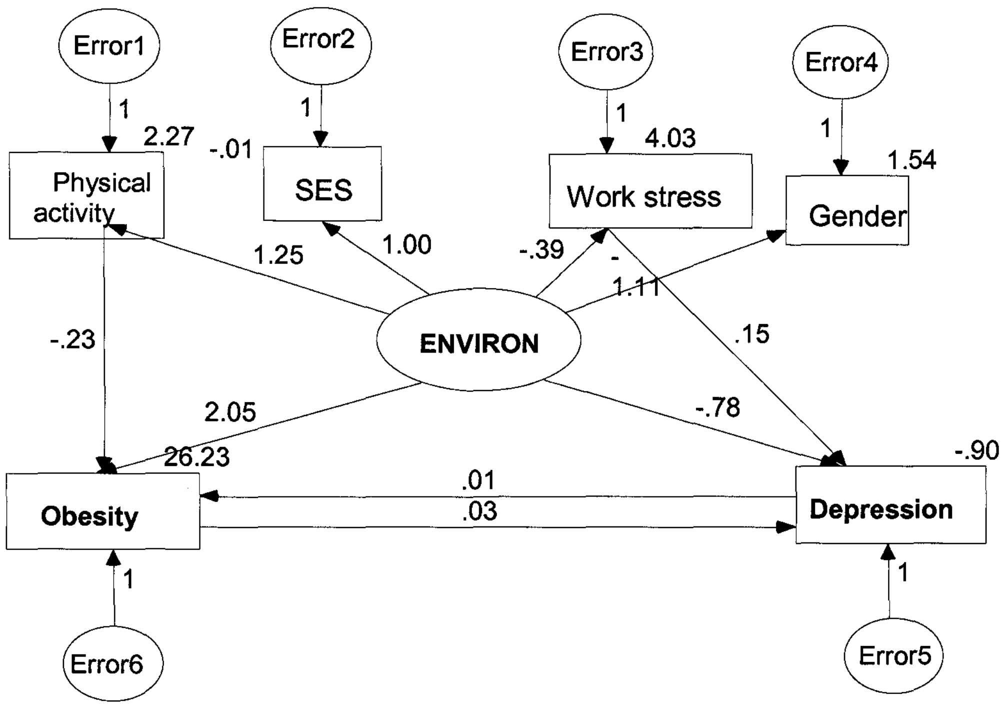

Figure5.6 Graphical output for Model 4

We modify our initial model by deleting this non significant path. The model will look as:

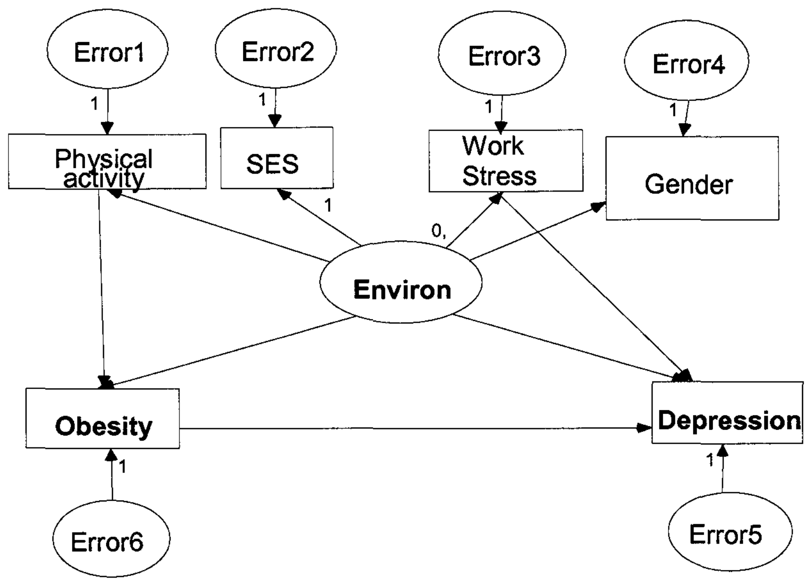

Figure 5.7 Path diagram for Model 5

This time we do not deal with a nonrecursive model. We hypothesize that obesity has a direct effect on depression, and depression is directly influenced by the work stress-social support. We run AMOS software and check the goodness of fit of this new model. The values of the goodness of fit indices are shown in the following table:

Table 5.4 Goodness of Fit Indices for Model 5

MODEL	CHI-SQUARE	RMSEA	NFI	CFI
Model 5	80.939(6 DF)	0.032	0.917	0.922

All regression weights are statistically significant this time and no other problems seem to have arisen (Appendix 4). The goodness of fit indices show a slightly improvement from the previous model, but taking into account the parsimony property of the models, we choose this recursive model as better fitting the population under study. The results are shown in the path diagram below.

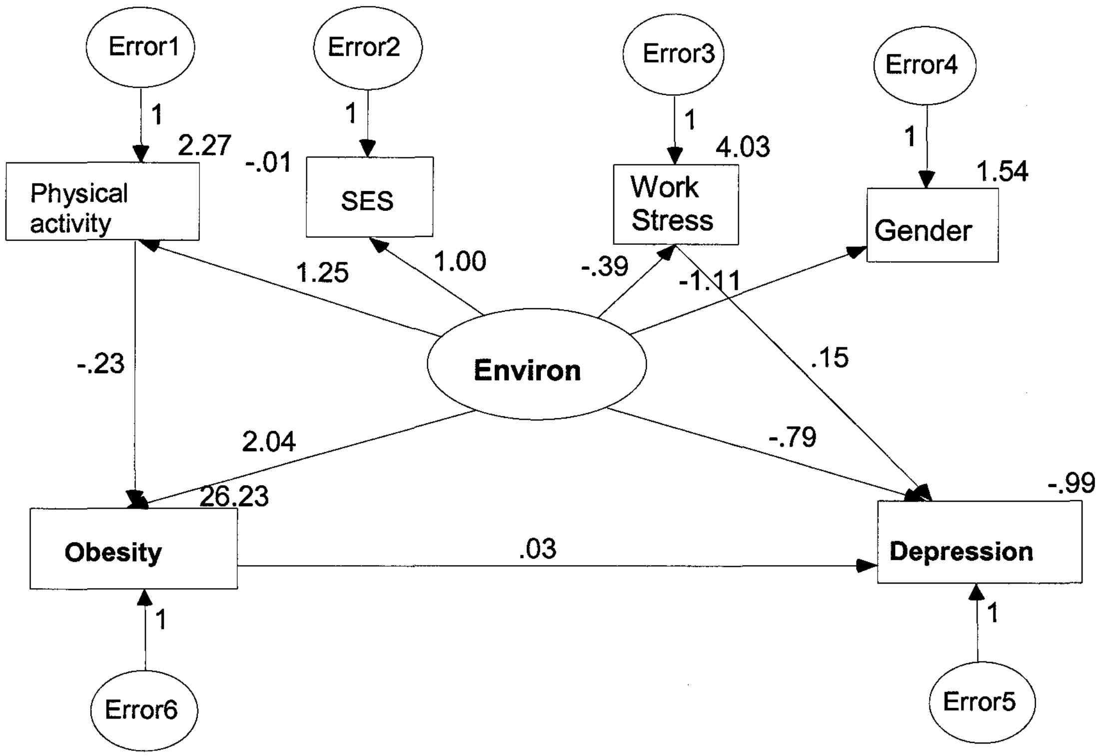

Figure 5.8 Graphical output for Model 5

5.1.3 Models for males and females datasets

Next we will look at the diagrams which model the datasets for males and females separately. Also, we will discuss separately the two cases a) stress viewed as "Self-perceived ability to handle an unexpected problem and b) stress viewed as "Work Stress - social support".

5.1.3.1 Stress - "Self-perceived ability to handle an unexpected problem"

From the total sample of 12376 records we form the two groups, namely "males" and "females", having sizes of 5660 and 6716, respectively.

Because the gender variable was initially in our model we will have to remove/delete it during this part of the analysis.

The model will look as shown below in the path diagram:

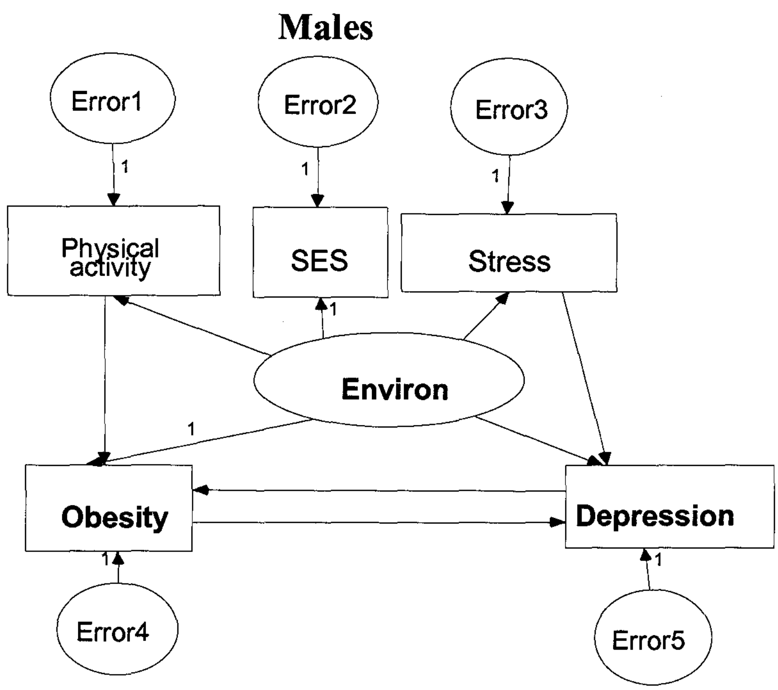

Figure 5.9 Model 6 - Males data

This baseline model fits acceptable well the data for males having the goodness of fit indices RMSEA= 0.056, CFI = 0.890, NFI = 0.889 and statistically significant parameter estimates. We test the same model for the female group, however this time, the model fit is not very good and we deal with non significant estimates as they can be seen in the table below:

Table 5.5 Regression weights: Model 7(a) – Females data

	Estimate	S.E.	C.R.	P
Phys.Activ ← Environ	1.317	.255	5.158	<.001
Stress ← Environ	-2.412	1.160	-2.080	.038
Obesity ←Environ	1.000			
Obesity ← Phys.Activ	-.324	.037	-8.668	<.001
Depression ← Environ	-6.002	4.127	-1.454	.146
Depression ← Stress	-.316	.481	-.658	.511
SES ←Environ	1.000			
Obesity ← Depression	-.009	.135	-.063	.950
Depression ← Obesity	.047	.061	.764	.445

Because we are mainly interested in the relationship between obesity and depression measured by the two variables HWTBDBMI and DEPRESS, we first eliminate the path connecting stress (STRB_1) with depression (DEPRESS) and check if the model fit improves. The fit indices show a moderate fit but some of the regression weights are still nonsignificant.

Table 5.6 Regression weights: Model 7(b) – Females data

	Estimate	S.E.	C.R.	P
Phys.Activ ← Environ	1.347	.264	5.094	<.001
Obesity ←Environ	1.000			
Obesity ← Phys.Activ	-.332	.037	-9.057	<.001
Depression ← Environ	-3.536	.566	-6.252	<.001
SES ←Environ	1.000			
Stress ← Environ	-1.830	.400	-4.579	<.001
Obesity ← Depression	-.082	.129	-.635	.525
Depression ← Obesity	.079	.058	1.360	.174

With a p value of 0.525 or 0.174, one of the path between depression and obesity will have to be eliminated. We exclude the most non significant path from the model and we run the program again. The results are not satisfactory. In the next table representing the regression weights for this third model ran on the female data, very large standard errors are shown:

Table 5.7 Regression weights: Model 7(c) – Females data

	Estimate	S.E.	C.R.	P
Phys.Activ ← Environ	-503.588	21038.901	-.024	.981
Obesity ←Environ	1.000			
Obesity ← Phys.Activ	-.311	.035	-8.809	<.001
SES ←Environ	1.000			
Stress ← Environ	811.194	33893.787	.024	.981
Depression ← Environ	1188.291	49644.559	.024	.981
Depression ← Obesity	.039	.009	4.212	<.001

We decide to eliminate the second non significant path from Model 7(b) and the results are presented below:

Table 5.8 Regression weights: Model 7(d) – Females data

	Estimate	S.E.	C.R.	P
Phys.Activ ← Environ	1.369	.255	5.360	<.001
Depression ← Environ	-3.417	.547	-6.243	<.001
Obesity ←Environ	1.000			
SES ←Environ	1.000			
Stress ← Environ	-1.546	.293	-5.269	<.001
Obesity ← Phys.Activ	-.326	.035	-9.204	<.001
Obesity ← Depression	.092	.020	4.488	<.001

All estimates are indeed significant and the model fit indices show an acceptable fit: RMSEA has a value of 0.038 with a 90% CI (0.028, 0.048), CFI is 0.879 and NFI is 0.872. The output of the two models (for males and females) in its graphical form can be seen below:

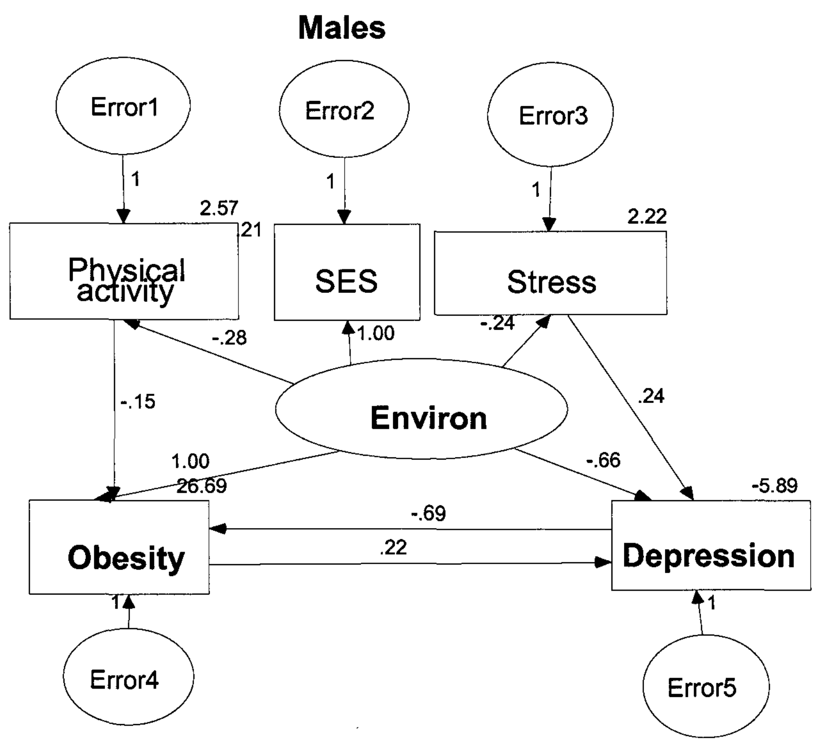

Figure 5.10 Graphical output for Model 6

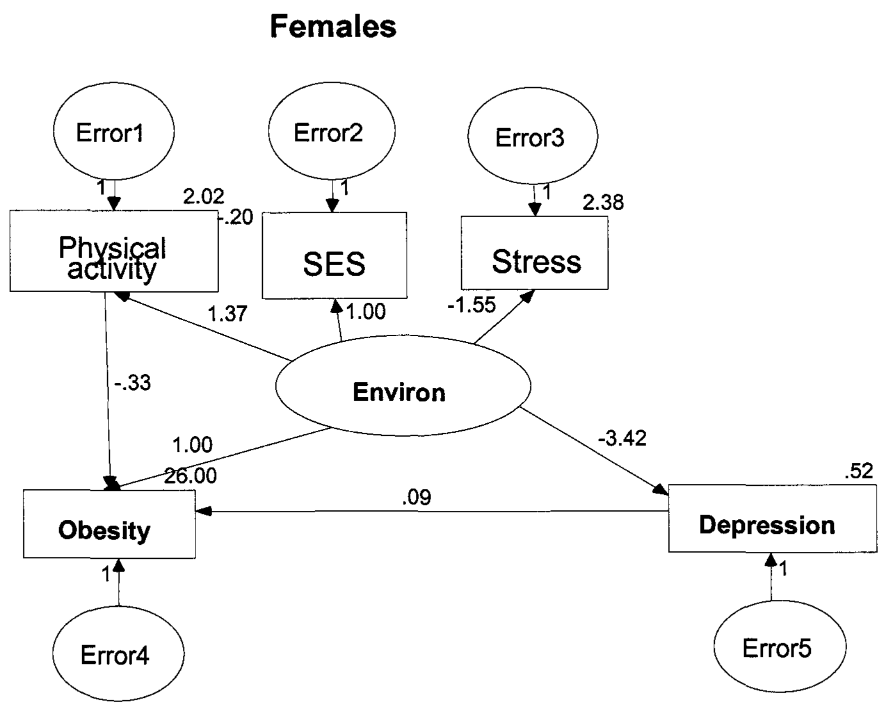

Figure 5.11 Graphical output for Model 7(d)

5.1.3.2 Stress - "Work stress-social support"

We will perform the same analysis on the same two groups, males and females, except the variable measuring stress will now refer to work stress, the social support aspect. The path diagram will look the same as before.

We run the program for the males' sample and obtain the following results:

Table 5.9 Goodness of Fit Indices for Model 5

MODEL	RMSEA	NFI	CFI
Model 8	0.031	0.941	0.947

The output reveals that this model fits the data very well. However the regression weights are not all significant:

Table 5.10 Regression weights: Model 8(a) – Males data

	Estimate	S.E.	C.R.	P
Phys.Activ ← Environ	-.291	.089	-3.260	**.001**
Work Stress ← Environ	.248	.078	3.177	**.001**
Depression ← Environ	-.442	.147	-2.999	**.003**
Obesity ← Phys.Activ	-.130	.022	-5.878	**<.001**
Obesity ←Environ	1.000			
Depression ← Work Stress	.111	.018	6.307	**<.001**
SES ← Environ	1.000			
Obesity ← Depression	-.079	.242	-.326	**.744**
Depression ← Obesity	.046	.068	.669	**.504**

As we did in the previous analysis we will eliminate one of the non-significant path. We start with the least significant one and we obtain a very good fitted model for the males samples with all parameter estimates significant:

Table 5.11 Regression weights: Model 8(b) – Males data

	Estimate	S.E.	C.R.	P
Phys.Activ ← Environ	-.282	.080	-3.545	**<.001**
Work Stress ← Environ	.234	.069	3.382	**<.001**
Obesity ← Phys.Activ	-.127	.022	-5.805	**<.001**
Obesity ←Environ	1.000			
Depression ← Environ	-.405	.081	-5.009	**<.001**
SES ← Environ	1.000			
Depression ← Work Stress	.113	.017	6.503	**<.001**
Depression ← Obesity	.023	.007	3.261	**.001**

For the simplicity of the interpretation of the results we present the output in the graphical form as well:

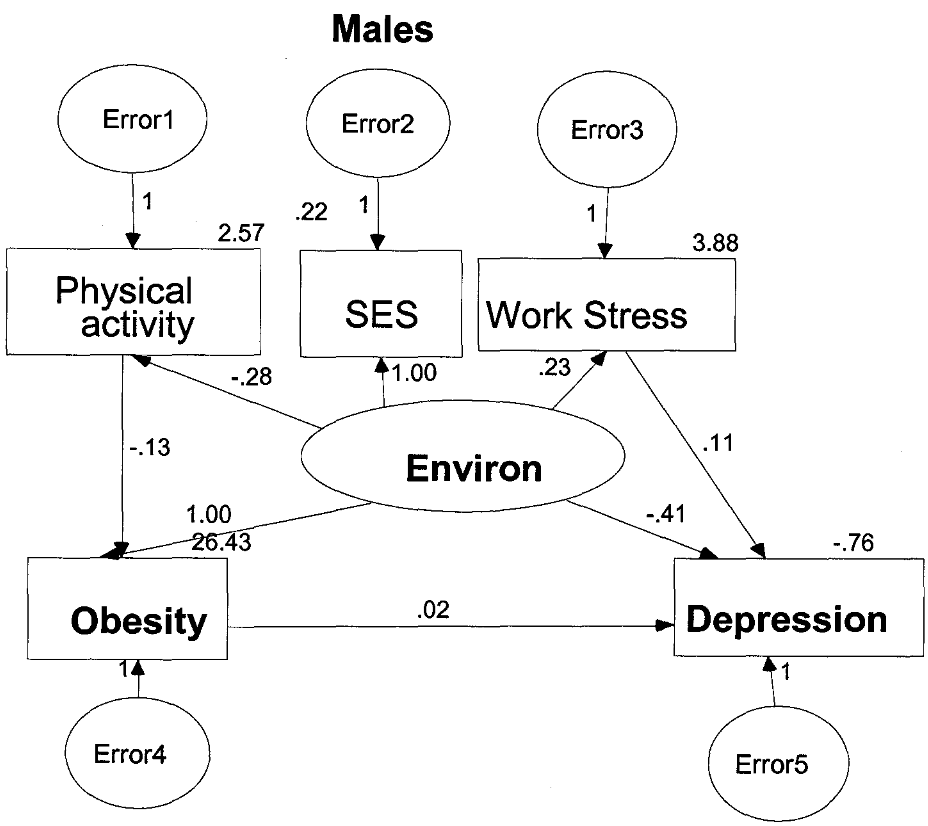

Figure 5.12 Graphical output for Model 8(b)

For the females' sample we start with the same model as in the previous case except this time stress is measured by WSTBDSOC variable (work stress-social support).

The fit indices show a good fit of the model, however some of the parameters estimates are again non-significant:

Table 5.12 Regression weights: Model 9(a) – Females data

	Estimate	S.E.	C.R.	P
Phys.Activ ← Environ	1.974	.477	4.141	**<.001**
Work Stress ← Environ	-.475	.646	-.736	**.462**
Obesity ←Environ	1.000			
Obesity ← Phys.Activ	-.283	.042	-6.798	**<.001**
Depression ← Environ	-8.553	3.337	-2.563	**.010**
Depression ← Work Stress	.128	.054	2.358	**.018**
SES ← Environ	1.000			
Obesity ← Depression	.813	.234	3.475	**<.001**
Depression ← Obesity	-.339	.113	-3.005	**.003**

We delete the path between stress and environment and we obtain our final model for the female data having a very good fit shown by the following indices: non-significant chi-square= 5.956 with 3 degrees of freedom and a p-value of 0.114, RMSEA= 0.012, CFI= 0.981, NFI=0.965 and all significant estimates which appear in the graphical output (Model 9 (b)):

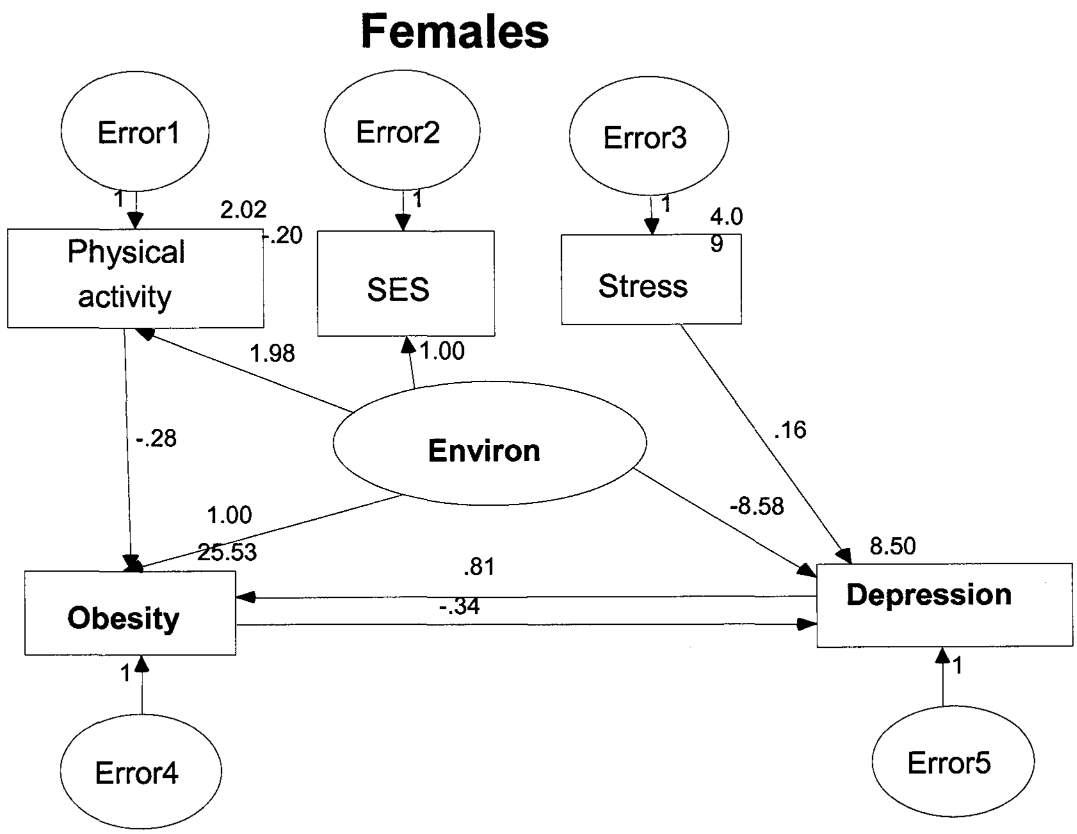

Figure 5.13 Graphical output for Model 9(b)

5.1.4 Multiple group analysis

We can use multiple group analysis in structural equation modeling in order to compare multiple samples across the same measurement instrument or multiple population groups, for instance males versus females, for any identified structural equation model.

AMOS will allow us to do this analysis; hence we test whether the groups meet the assumptions that they are equal by examining if different sets of path coefficients are *invariant*. By invariant we understand that the path coefficients in the model are equal for our groups. With AMOS we can test the equalities of variables' variances, means, intercepts, covariances between variables and the equalities of path coefficients across two or even more groups.

From Joreskog's work [61], a procedure able to test the invariance simultaneously across groups was derived. He recommended that all tests concerning the invariance begin with a global test of the equality of covariance structures across groups. We test the hypothesis $H_0 : \Sigma_1 = \Sigma_2 = ... = \Sigma_m$, where Σ is the population variance-covariance matrix and m represents the number of groups. If the null hypothesis is rejected then we can conclude that the groups are not equivalent and we proceed to further testing in order to identify the source of noninvariance. On the other hand, if the null hypothesis cannot be rejected then the groups are considered to be equivalent and other tests for invariance are unjustified. All subsequent investigative work should be done based on single-group analysis.

Although this test appears reasonable and straightforward it was proved to be leading to contradictory findings with respect to equivalencies across groups, meaning that the null hypothesis may be rejected and yet tests for the invariance hold. Depending on the model and hypotheses to be tested, sets of parameters are put to test. The most commonly parameters of interest in answering questions related to group invariance are factor loading paths, factor variances/covariances and structural regression paths.

The tests of the hypotheses related to group invariance usually start with the investigation of the measurement model; the pattern of factor loadings for the observed variables being tested for its equivalence across groups. Those measures which are found to be group-invariant are then constrained equal while further testing of the structural parameters is conducted. In other words, to determine the nonequivalence of parameters across groups we test a series of increasingly restrictive hypotheses.

In order to test for the factorial invariance we first have to consider a baseline model that is estimated for each group separately. This is the model that best fits the data from the parsimony and substantive meaningfulness points of view.

Because the estimation of the baseline model does not involve between group constraints, we can analyze the data separately for each group, while in the case of the invariance testing, because equality constraints are imposed on some parameters, data has to be analyzed simultaneously to obtain good estimates [62].

We will perform multiple group analysis in order to test for invariance of the proposed model across gender and across level of obesity (groups: males, females and obese, non-obese).

5.1.4.1 Stress -"Self-perceived ability to handle an unexpected problem"

In order to fit the model in AMOS, we first draw the model for each single group and then fit it for the group's sample data to ensure that the model is properly identified and that no minimization or other unexpected problems arise during the model fitting process.

The model will look as shown below in the path diagram:

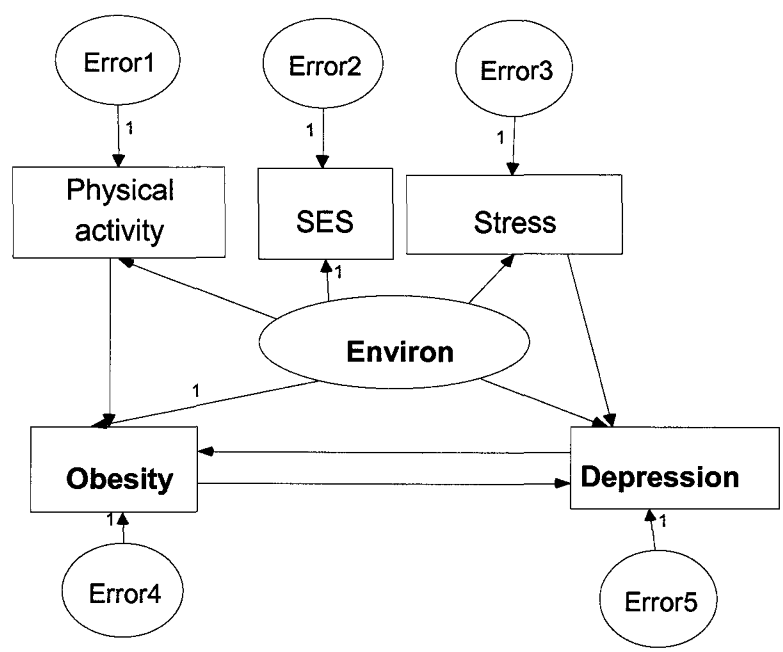

Figure 5.14 Multiple group analysis with Stress-STRB_1

This baseline model fits acceptable well the data for males having the goodness of fit indices RMSEA= 0.056, CFI = 0.890, NFI = 0.889 and statistically significant parameter estimates. We test the same model for the female group and obtain RMSEA= 0.053, CFI = 0.879 and NFI = 0.879.

When testing for multigroup invariance, the researcher often tests one-sample models separately first (i.e. for a male sample and for a female sample), in order to provide an overview of how consistent the model results are, but it does not constitute testing for significant differences in the model's parameters between groups.

First a baseline chi-square value is derived by computing model fit for the pooled sample of all groups. Then we are going to add constraints that various model parameters must be equal across groups and the model is fitted, yielding a chi-square value for the constrained model. A chi-square difference test is then applied to see if the difference is significant.

We can assume now that both models converged correctly and we proceed with the multiple group analysis.

Both groups in the analysis will have identical path diagram structure, hence we only have to draw it for the first group and AMOS will use the same structure for the second one by default. AMOS program will then test invariance simultaneously across groups and the fit of this estimated model provides the baseline value against which all following specified models are compared.

While the same model structure is specified for males and females, there is no restriction that the parameters must have the same values in the two groups. This means that the regression weights, covariance paths and variances may all be different for males and females.

We ran the program and the goodness-of-fit indices indicate that our hypothesized model is moderate, but acceptable fitting across the 2 groups. In this way we obtain a baseline chi-square value which is derived by computing model fit for pooled sample of all groups.

Table 5.13 Goodness of Fit Indices for Model 10

MODEL	RMSEA	NFI	CFI
Model 10	0.038	0.884	0.885

We accept the hypothesis that the proposed model is correct for both males and females and we next start to look at the parameter estimates, being interested in how males' estimates compare to the females' estimates.

In the next tables we will have these two parts of the output.

First, here are the females' parameter estimates:

Table 5.14 Regression weights: (Females)

	Estimate	S.E.	C.R.	P
Phys.Activ ← Environ	**1.317**	.255	5.158	<.001
Stress ← Environ	**-2.412**	1.160	-2.080	.038
Obesity ←Environ	**1.000**			
Depression ← Environ	**-6.002**	4.127	-1.454	.146
Obesity ← Phys.Activ	**-.324**	.037	-8.667	<.001
Depression ← Stress	**-.316**	.481	-.658	.511
SES ← Environ	**1.000**			
Obesity ← Depression	**-.009**	.135	-.063	.950
Depression ← Obesity	**.047**	.061	.764	.445

Next are the estimates for males' sample:

Table 5.15 Regression weights: (Males)

	Estimate	S.E.	C.R.	P
Phys.Activ ← Environ	**-.276**	.100	-2.763	.006
Stress ← Environ	**-.235**	.050	-4.731	<.001
Obesity ←Environ	**1.000**			
Depression ← Environ	**-.656**	.197	-3.330	<.001
Obesity ← Phys.Activ	**-.149**	.024	-6.200	<.001
Depression ← Stress	**.238**	.039	6.079	<.001
SES ← Environ	**1.000**			
Obesity ← Depression	**-.689**	.225	-3.068	.002
Depression ← Obesity	**.216**	.062	3.498	<.001

We now examine the question of whether obesity and depression follow the same dynamics in males as in females.

In structural equation modeling, testing for the invariance of parameters across groups can be done by placing constraints on particular parameters, i.e. we specify that the parameters are equivalent across these groups by giving them a label. Those parameters that are unlabeled will be freely estimated, taking different values across groups, while the labeled ones will be kept equal across same groups.

Supposing that we are mainly interested in the regression weights, we hypothesize that the females and males have the same regression weights. Hence the variances and covariances of the variables would be still allowed to differ between groups while the regression weights are group-invariant.

What we want to do is to evaluate whether a fixed unit change on an exogenous variable will always correspond to the same change in the endogenous one, independent of whether the respondent is male or female. If this modified model is confirmed by the data then the same regression weights can be used for all groups, which would simplify the prediction of the endogenous variables.

More specifically we test for structural invariance across groups. We test if the arrows connecting the latent and observed variables from the structural part of the model are properly drawn in the same way for each group in the analysis. We run AMOS for this model in which the regression weights representing the structural paths are kept equal (for both groups they are labeled with Wi, i=1,3,4,6,7) as it can be seen in the below path diagram:

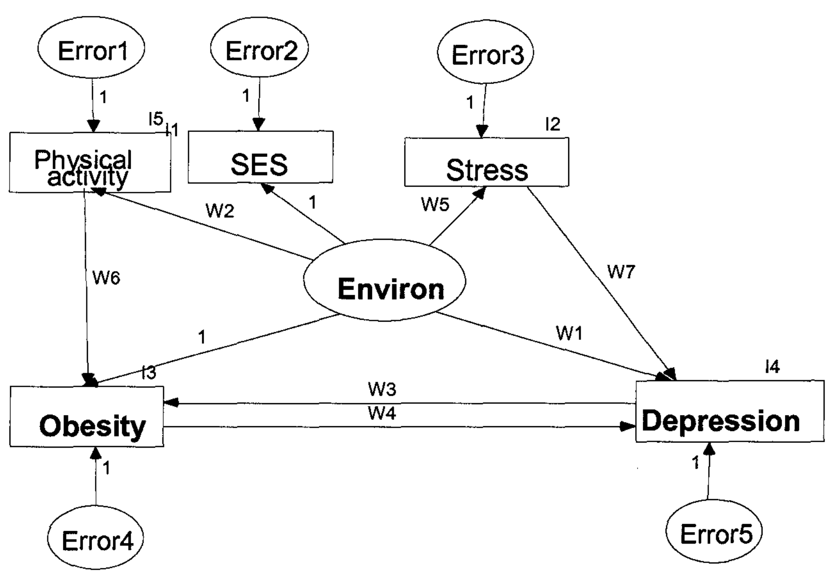

Figure 5.15 Labeled path diagram for multiple group analysis

For this second model we obtain a chi-square equal to 136.6 with 9 degrees of freedom and comparing the two models (unconstrained and constrained) we get a chi-square difference of 59.247 with 5 degrees of freedom which is significant. If the baseline and constrained model are significantly different, one inference is that "there is a moderating effect on causal relationships in the model, and this effect varies by group."

5.1.4.2 Stress - "Work stress-social support"

We will perform the same analysis on the same two groups, males versus females, except the variable measuring stress will now refer to work stress, the social support aspect. The path diagram will look the same as before.

After drawing the model for each single group we fit it for the group's sample data to ensure that the model is properly identified and that no minimization or other unexpected problems arise during the model fitting process.

This baseline model fits very well the data for males having the goodness of fit indices RMSEA = 0.031, CFI = 0.947, NFI = 0.941. For the female group we obtained RMSEA= 0.016, CFI = 0.977 and NFI = 0.968 and even a nonsignificant chi-square of 5.5 with 2 degrees of freedom. We can assume now that both models converged correctly and we proceed with the multiple group analysis.

The model fits the data from both groups very well as it can be seen from the next table:

Table 5.16 Goodness of Fit Indices for Model 11

MODEL	RMSEA	NFI	CFI
Model 11	0.017	0.953	0.960

Because we are interested in how males' estimates compare to the females' estimates we have a look on the following results:

Table 5.17 Regression weights: (Females)

	Estimate	S.E.	C.R.	P
Phys.Activ ← Environ	**1.974**	.477	4.141	<.001
Work Stress ← Environ	**-.475**	.646	-.736	.462
Obesity ←Environ	**1.000**			
Depression ← Environ	**-8.553**	3.337	-2.563	.010
Obesity ← Phys.Activ	**-.283**	.042	-6.798	<.001
Depression ← Work Stress	**.128**	.054	2.358	.018
SES ← Environ	**1.000**			
Obesity ← Depression	**.813**	.234	3.475	<.001
Depression ← Obesity	**-.339**	.113	-3.005	.003

Table 5.18 Regression weights: (Males)

	Estimate	S.E.	C.R.	P
Phys.Activ ← Environ	**-.291**	.089	-3.260	.001
Work Stress ← Environ	**.248**	.078	3.177	.001
Obesity ←Environ	**1.000**			
Depression ← Environ	**-.442**	.147	-2.999	.003
Obesity ← Phys.Activ	**-.130**	.022	-5.878	<.001
Depression ← Work Stress	**.111**	.018	6.307	<.001
SES ← Environ	**1.000**			
Obesity ← Depression	**-.079**	.242	-.326	.744
Depression ← Obesity	**.046**	.068	.669	.504

We test for structural invariance across groups by labeling the structural paths to be equal in both groups and rerun the analysis.

For the constrained model we obtain a chi-square equal to 85.024 with 9 degrees of freedom and comparing the two models (baseline and constrained) we get a chi-square difference of 66.793 with 5 degree of freedom which is significant. Again in the situation when the baseline and constrained model are significantly different, we have to conclude that the moderating effect that exists on the causal relationships in the model varies by group.

Next we will compare 4 groups: obese females, non-obese females, obese males and non-obese males. We make sure that the model fits well each group separately and then we run the multiple group analysis. In the next table we will find the fit indices for the four models. This time not all groups will have the same diagram.

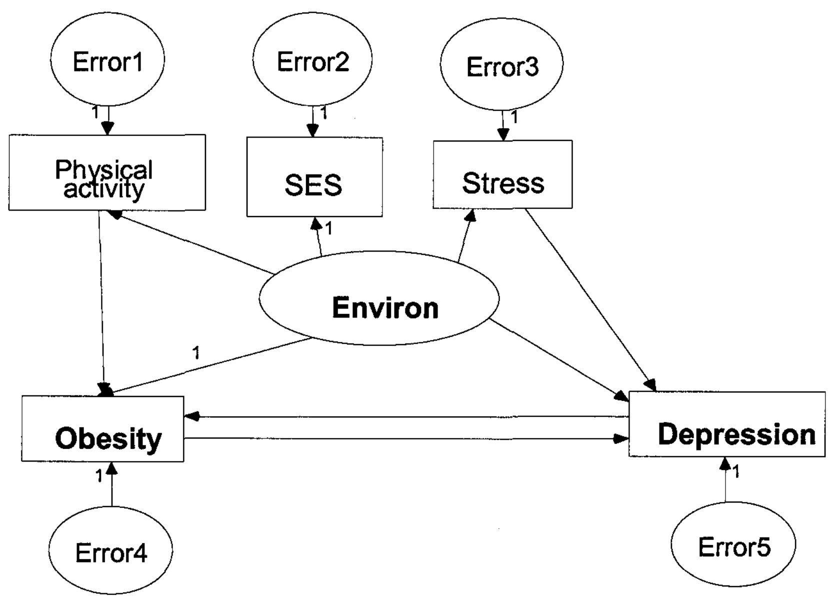

Figure 5.16 Model for non-obese females' data

The difference will consist in that for the groups "non-obese females" and "obese males" the path from stress (ability to handle an unexpected problem) and depression was deleted or set to zero in order to have an acceptable model fit.

Table 5.19 Goodness of Fit Indices for 4 groups

MODEL	RMSEA	NFI	CFI
Non-obese females	0.030	0.879	0.890
Obese females	0.00	0.984	1.00
Non-obese males	0.051	0.912	0.914
Obese males	0.054	0.836	0.847

We perform the multiple group analysis in order to see whether or not this model fits the data across groups.

The indices are below the minimum acceptable and also the estimates for the non-obese males group has some very large standards errors. There are differences among the four groups made up based on gender and obesity status.

If we use the work stress - social support variable the model does not fit the obese males data.

5.2 Categorical variables

In this section we will analyze the data using the Mplus software. Like AMOS, Mplus features Full Information Maximum Likelihood (FIML) handling of missing data. This method enables Mplus to make use of all available data points, including the cases with some missing values. The difference between Mplus and AMOS is that the former has the extra unique ability of generating the modification indices even for data that are not complete. Another useful feature that makes Mplus "better" than the other softwares for SEM is its ability to fit latent variables models to datasets that contain ordinal or dichotomous outcome variables. However, Mplus will not fit models with nominal outcome variables that contain more than two levels.

All Mplus commands are specified using command syntax. Once launched the Mplus we can build a command file using the following 9 headings: TITLE, DATA, VARIABLE, DEFINE, SAVEDATA, ANALYSIS, MODEL, OUTPUT AND MONTECARLO. They may come in any order and DATA and VARIABLE must appear in all analyses. All commands must begin on a new line and are followed by a colon, while semicolons separate command options. In the DATA command we specify where the data set is located, the format of the file and the names of the variables. The accepted types of data files for Mplus are tab-delimited text, space-delimited text and comma-delimited text. So we had to convert our SPSS file into .dat file in order to be able to analyze it using this software.

The next command is the VARIABLE command which names the columns of data that Mplus reads using the previous command.

The dichotomous variables that we are going to use are DEPLIFE (lifetime depression) and DEPYEAR (12-month depression) which we defined in the previous chapter and BMI (the body mass index categorized as follow: 1 acceptable weight, 2 overweight, 3 obese).

No matter what method of handling the missing data we use, we need to tell Mplus how these missing values are represented in the database. This can be done in the MISSING subcommand of VARIABLE. In our case we specified that all missing values are -9.

1. Lifetime Depression

The syntax used to run this analysis is shown in Appendix 5.

The model is similar to the one presented in Figure 5.4 above.

The output reveals a marginal fit:

```
Chi-Square Test of Model Fit

          Value                              272.772

          Degrees of Freedom                       5

          P-Value                             0.0000

RMSEA (Root Mean Square Error Of Approximation)

          Estimate                             0.066

CFI                                            0.815
```

The statistically significant chi-square test of absolute model fit and the poor RMSEA and CFI fit statistics values suggests that this model may need some modification in order to fit better the data. We can use the modification indices provided by Mplus to make this decision. The MI are presented in the second part of Appendix 5.

They suggest that there should be a path between physical activity and stress. The syntax is: PACBDEE ON STRB_1. We run again the program and check the new model's fit. There is an improvement in the model fit as it can be seen as follows:

```
Chi-Square Test of Model Fit

          Value                              185.680

          Degrees of Freedom                       4
```

```
        P-Value                                  0.0000

RMSEA (Root Mean Square Error Of Approximation)

        Estimate                                 0.061

CFI                                              0.874
```

We can say that there is a moderate fit of the data in this case.

2. 12-month Depression

When depression is defined as 12-month depression and measured by the variable DEPYEAR we use the same model and syntax as in the previous case and we obtain:

```
Chi-Square Test of Model Fit

        Value                                  219.017

        Degrees of Freedom                           4

        P-Value                                 0.0000

RMSEA (Root Mean Square Error Of Approximation)

        Estimate                                 0.066

CFI                                              0.868
```

which shows a poor fit of this model for our data.

In this case modification indices do not suggest any improvement that can be made, so we conclude that the hypothesized model does not fit our data.

Chapter 6

Discussion and Conclusion

6.1 Summary of findings

Data were analyzed using the AMOS and Mplus softwares. These analyses used the maximum likelihood methods of parameter estimation.

The dataset was cross-sectional; hence our models were cross-sectional ones. For depression we had three variables, one continuous and 2 categorical (DEPRESS – persistence of depression , in years; DEPLIFE – lifetime depression and DEPYEAR- 12 month depression). We also had information about two types of stress: ability to handle an unexpected problem and work stress-social support (STRB_1 and WSTBDSOC). We looked at models using all types of depression and stress separately using AMOS software for continuous variables and Mplus for models with categorical variables. In Mplus, obesity was considered as a categorical variable, according to the BMI level, as it was explained in Chapter 2.

We also looked at how the models fit the data across different groups: males versus females, and obese males, non-obese males, obese females, non-obese females.

Model 3, having one latent variable, ENVIRON, and stress variable represented by STRB_1 fits our data pretty well, conclusion indicated by the fit indices in Table 5.1 and by the significant regression weights (Table 5.2). If we consider the two variables of main interest for our study, depression and obesity, we can conclude from Model 3 that the higher the level of obesity the more severe the depression of the patient is. Not the same conclusion is drawn from this model for the effect of depression on obesity: the more severe the depression the less weight gain that can end up in obesity problems. An explanation can be that nowadays the new medicines used to treat patients with depression have no weight gain side effect as they used to. However we did not have information regarding the treatment of these patients, if any. Another important result obtained through this model is the "positive" effect that environment exerts on obesity status. A noteworthy feature of this model illustrated by the statistically significant unstandardized regression coefficient includes the negative relationship between Physical activity and Obesity, which translates into "more physical activity less weight problems". Another negative relationship seems to be between Environment and Depression, present in all the models discussed in this project. As we would have expected stress, both viewed as ability to handle an unexpected problem and work stress-social support, has a direct effect on depression (Model 3, Model 5).

Model 5 was similar to the previous model discussed above except that the path from Depression to Obesity was not significant, so it had to be eliminated. The conclusions are the same as before; no changes in the signs of the regression weights, just in their magnitude.

Afterwards we fit our model to the female and male samples separately. We eliminated the necessary paths so we end up with the best model possible. For males sample the output is shown in Figure 5.10 (Model 6) for stress – ability to handle an unexpected problem and in Figure 5.12 (Model 8(b)) for work stress. As before, stress has a direct effect on depression in both cases, environment and depression are in a negative relation and so is physical activity and obesity (BMI). The higher the BMI the more severe the depression, while for higher levels of severity of depression we have lower BMI levels in the males' sample. In the work stress case this path between depression and obesity was not significant, hence eliminated from the model.

For the female sample things are a little different in the sense that stress-ability to handle an unexpected problem does not have a direct effect on depression, the path between these two variables being eliminated in order to improve the fit of the model (Figure 5.11). Nonsignificant path is also the one from obesity to depression. In Model 7(d) for the females sample we have that high levels of severity of depression does have an effect on obesity, different from the males' sample. In Model 9(b), when we used the work stress variable, the only nonsignificant path was the one between stress and environment. Stress does not load on the latent variable, but it is part of the structural part of the model, having a positive relationship with depression. In this model, for the females sample and work stress, we have a different conclusion: the higher the depression level the more increase in the BMI level and the higher the BMI the lower the severity of depression.

The results seem to indicate that the relationship between depression and obesity is different for the two groups, males and females. Because testing the models separately on the two samples does not constitute testing for significant differences in the model's parameters between groups, we perform multiple group analysis using the AMOS software. As we assumed, the moderating effect on the "causal" relationship in the model varies by group.

In the last part of the previous chapter we analyzed the models using Mplus software and the entire data set, as we dealt with categorical variables measuring depression. The model containing lifetime depression had a poor to moderate fit, while the one using 12 month depression variable had an even poorer fit.

6.2 Limitations

SEM requires a sound theoretical perspective. To get a better-fit model a good design guided by a substantive theory and prior research is needed. It is also recommended a comprehensive effort to identify relevant variables and their relationships. This will help determine valid and relevant indicator variables of latent variables, provide a theoretic perspective of the model and help establish latent variables relationships grounded in prior research studies. Also a well fitting model is not necessarily unique. There can be other reasonable models for the same data. We can have different models designs and different outcomes.

In our project we used a data set already collected, and it was not a study designed for the research question we try to answer. That is why a very important limitation of our study is the lack of information/variables about history of depression and obesity in the family, history of debilitating diseases in the family, adverse childhood experience, history of debilitating diseases in the household (excluding the members of blood relative), teasing, etc. Even the information we had about eating habits had a very big numbers of missing values and could not be used.

Goodness of fit measures are sensitive to sample size, method of estimation and model misspecification. The chi-square test for model fit should be used with great caution as it has a number of weaknesses. Our data are not strictly normally distributed which might be one of the reasons of some models not fitting very well.

Another limitation of our models might be the number of indicators per latent variable. The more variables that are used to assess a construct the more reliable the model will be. It is recommended to have at least 3 indicators.

A way of improving the model fit is to respecify the models which we did in this study. We tried different models with different variables and different groups' datasets.

6.3 Conclusion

The data we used in the study is a cross-sectional one. In order to be able to draw conclusions on the causal relationship between obesity and depression longitudinal study with information as stated above needs to be designed.

It is also important to note that even though these models fit the data well and provide a theoretically consistent set of findings, there may be other *equivalent models* that fit the data as well. Or there may also be non-equivalent alternative models that fit the data better than these models.

References:

[1] Health Canada (1996) *Report on Health of Canadians.*

[2] World Health Organization (2000) WHO Technical Report Series no. 894. *Obesity: Preventing and Managing the Global Epidemic.* Geneva: World Health Organization.

[3] Must, A., Strauss, R. S. (1999) "Risks and consequences of childhood and adolescent obesity." *International Journal of Obesity,* **23**(Supplement 2), 2–11.

[4] Canadian Institute for Health Information (2004) "Improving the health of Canadians".

[5] Statistics Canada, "Body mass index (BMI) international standard, by sex, household population aged 20 to 64 excluding pregnant women, Canada, provinces, territories, health regions and peer groups, 2000/2001". Health Indicators (Statistics Canada, Catalogue 82-221-XIE); 2002(1).

[6] Stephen A. M., Wald N. J. (1990) "Trends in individual consumption of dietary fat in the United States, 1920–1984." *American Journal of Clinical Nutrition,* **52**, 457–469.

[7] Stephen A. M., Sieber G. M. (1994). "Trends in individual fat consumption in the UK 1900–1985." *British Journal of Nutrition,* **71**, 775–788.

[8] http://www.weight.com/depression.asp

[9] Public Health Agency of Canada, (2002) "A report on Mental Illnesses in Canada".

[9a] Villanueva E.V., (2001) "The validity of self-reported weight in US adults: a population based cross-sectional study", *BMC Public Health,* **1,** 11.

[10] Horwath E., Weissman M.M.(1995) "Epidemiology of depression and anxiety disorders". In: Tsuang M.T., Tohen M., Zahner G.E.P., eds. *Textbook in Psychiatric Epidemiology. New York: Wiley-Liss*, 317-44.

[11] Spaner, D., Bland, R.C., Newman, S.C. (1994) "Major depressive disorder". *Acta Psychiatr. Scand.*, (Suppl. 376), 7-15.

[12] De Marco, R.R.(2000) The epidemiology of major depression: implications of occurrence, recurrence and stress in a Canadian community sample. *Can. J. Psychiatry*, **45**, 67-74.

[13] American Psychiatric Association, (1994) "Diagnostic and Statistical Manual of Mental Disorders", Fourth Edition.Washington, DC: *American Psychiatric Association.*

[14] World Health Organization, (1992) "The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines". Geneva: *World Health Organization.*

[15] World Health Organization. (2001) "World Health Report 2001. Mental Health: New Understanding, New Hope". Geneva, Switzerland: *World Health Organization.*

[16] Sartorius, N., Ustun, T.B., Lecrubier, Y., et al (1996) Depression comorbid with anxiety: results from the WHO Study on Psychological Disorders in Primary Health Care. *British Journal of Psychiatry*, **168** (suppl.30), 38-43.

[17] Goldberg, D.P., Lecrubier, Y. (1995) Form and frequency of mental disorders accross centres. *Mental Illness in General Health Care: an international study*, 323-

334, John Wiley.

[18] Sagar, V. P., Raymond, W. L. (2001) Canadian clinical practice guidelines for the treatment of depressive disorders. *Can J Psychiatry*, **46**, Suppl.1.

[19] Stephens, T., Joubert, N. (1999) Mental Health of the Canadian population: a comprehensive analysis, *Chronic Diseases in Canada*, **20**:3.

[20] Akhtar-Danesh, N., Landeen , J. (2005) "Depression and its relation with age, sex and marital status in Ontario" (Personal Communication).

[21] Bland, R.C. (1997) Epidemiology of affective disorders: a review. *Can J Psychiatry,* **42**:367-377.

[22] Dong, C., Sanchez, L.E. and Price, R.A. (2004) Relationship of Obesity to Depression: a Family-based Study. *International Journal of Obesity*, **28**, 790-795.

[23] Flegal, K.M., Caroll, M.D., Ogden, C.L., and Johnson, C.L. (2002) Prevalence and trends in obesity among US adults. *JAMA*, **288**, 1723-1727.

[24] Kessler, R.C., et al (1994) Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the US: Results from the National Comorbidity Survey. *Arch Gen Psychiatry*, **51**, 8-19.

[25] WHO Monica Project (1989) Risk factors, *Int J. Epidemiol.*, **18**, S46-55.

[26] Allison, D.B., et al. (1999) Annual deaths attributable to obesity in the US. *JAMA*, **282**, 1530-1538.

[27] McGinnis, J., Foege, W. (1993) Actual causes of deaths in the United States. *JAMA*, **270**, 2207-2212.

[28] Aronne, L.J. (2001) Epidemiology, morbidity and treatment of overweight and

obesity. *Journal of Clinical Psychiatry*, **62**, 13-22.

[29] Kawachi, I. (1999) Physical and psychological consequences of weight gain. *Journal of Clinical Psychiatry*, **60,** (Suppl 21), 5-9.

[30] Hassan, M.K., Joshi, A.V., Madhavan, S.S., Amonkar, M.M. (2003) Obesity and health-related quality of life: a cross-sectional analysis of the US population. *International Journal of Obesity*, **27**, 1227-1232.

[31] Trakas, K., Lawrence, K., Shear, N.(1999) Utilization of health care resources by obese Canadians. *CMAJ*, 160, 1457-1462.

[32] Statistics Canada, *The Daily*, Canadian Community Health Survey, June 15, 2004.

[33] Tremblay, M.S., Katzmarzyk, P.T., Willms, J.D. (2002) Temporal trends in overweight and obesity in Canada 1981-1996. *International Journal of Obesity*, **26**, 538-543.

[34] Stunkard, A.J., Faith, M.S., Allison, K.C. (2003) Depression and obesity. *Society of Biological Psychiatry,* **54**, 330-337.

[35] Roberts, R.E., Strawbridge, W.J., Deleger, S., Kaplan, G.A.(2002) Are the fat more jolly?. *Ann Behav Med*, **24**, 169-180.

[36] Roberts, R.E., Kaplan, G.A., Shema, S.J., Strawbridge, W.J. (2000) Are the obese at greater risk for depression? *Am J Epidemiol*, **152**, 163-170.

[37] Ross, C.E. (1994) Overweight and depression. *J Health Soc Behav*, **35**, 63-79.

[38] Stewart, A.L., Brook, R.H. (1983) Effects of being overweight. *Am J Public Health*, **73**, 171-178.

[39] Palinkas, L.A., Wingard, D.L., Barrett-Connor, E. (1996) Depressive symptoms in

overweight and obese older adults: a test of the "jolly fat" hypothesis. *J Psychosom Res*, **40**, 59-66.

[40] Friedman, M.A., Brownell K.D. (1995) Psychological correlates of obesity: moving to the next research generation, *Psychol Bull*, **117**, 3-20.

[41] Onyike, C.U., Crum, R.M., Lee, H.B., Lyketsos, C.G., Eaton, W.W. (2003) Is obesity associated with major depression? Results from the Third National Health and Nutrition Examination Survey. *American Journal of Epidemiology*, **158**, 1139-1147.

[42] American Obesity Association (2002) Fact sheets, http://www.obesity.org/subs/fastfacts/obesity_global_epidemic.shtml.

[43] Waters, G.S., Pories, W.J., Swanson, M.S. (1991) Long-term studies of mental health after the Greenville gastric bypass operation for morbid obesity. *Am J Surg*, **161**, 154-158.

[44] Health Canada, Office of Nutrition Policy and Promotion, (2003) "Canadian Guidelines for Body Weight Classification in Adults".

[45] World Health Organization (1995) Physical Status: The Use and Interpretation of Anthropometry. *Report of a WHO Expert Committee*. Geneva, WHO.

[46] Helat, A., Vaccarino, V., Krumholz, H.M.(2001) An evidence-based assessment of federal guidelines for overweight and obesity as they apply to elderly persons. *Arch Intern Med*, **16** (9), 1194-203.

[47] Dietz, W.H., Bellizzi, M.C.(1999) Introduction: the use of body mass index to assess obesity in children. *Am J Clin Nutr*, **70**(1), 123S-5S.

[48] Cole, T.J., Bellizzi, M.C., Flegal, K.M., Dietz, W.H. (2000) Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ*, **320**, 1240-3.

[49] Bentler, P.M.(1988) Causal modeling via structural equation systems. In J.R. Nesselroade and R.B. Cattell (Eds.), *Handbook of multivariate experimental psychology*, New York: Plenum.

[50] Byrne, B.M.(2001) *Structural Equation Modeling with Amos. Basic concepts, Applications and Programming*, Lawrence Erlbaum Associates Publishers.

[51] Ullman, J.B.(1996) *Structural Equation Modeling* (In: Using Multivariate Statistics, Third Edition, B.G. Tabachnick and L.S. Fidell, Eds.). HarperCollins College Publishers, New York.

[52] Tabachnick, B.G. and Fidell, L.S. (2001) Using Multivariate Statistics, Fourth Edition,Allyn and Bacon.

[53] Roth, P. (1994) Missing data: A conceptual review for applied psychologists, *Personnel Psychology*, **47**, 537-560.

[54] Kendler, K.S., Walters, E.E., Neale, M.C., Kessler, R.C., Heath, A.C., Eaves, L.J. (1995) The structure of the genetic and environmental risk factors for six major psychiatric disorders in women: Phobia, generalized anxiety disorder, panic disorder, bulimia, major depression, and alcoholism. *Arch Gen Psychiatry,* **52**: 374-383.

[55] MacCallum, R.C., Browne, M. W., Sugawara, H.M. (1996) Power analysis and determination of sample size for covariance structure modeling. *Psychological*

Methods, **1**, 130-149.

[56] Joreskog, K.G., Sorbom, D. (1996) *Structural equation modeling*, Workshop presented for the NORC Social Science Research Professional Development Training Sessions, Chicago.

[57] Hu, L.-T., Bentler, P.M. (1999) Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, **6**, 1-55.

[58] Byrne, B.M. (1994) Testing for the factorial validity, replication, and invariance of a measuring instrument: A paradigmatic application based on the Maslach Burnout Inventory. *Multivariate Behavioral Research,* **29**, 289-311.

[59] Bentler, P.M. (1990) Comparative fit indices in structural models. *Psychological Bulletin*, **107**, 238-246.

[60] Arbuckle, J.L. Wothke, W. (1999) Amos 4.0 User's Guide.

[61] Joreskog, K.G. (1971) Simultaneous factor analysis in several populations, *Psychometrika,* **36,** 409-426.

[62] Bentler, P.M. (1995) EQS: Structural equations program manual, Encino, CA: Multivariate Software, Inc.

Appendix 1

Table A1: International classification of overweight and obesity in children (adapted from [48]) based on BMI.

Age (years)	Overweight		Obese	
	Boys	Girls	Boys	Girls
12.0	21.22	21.68	26.02	26.67
12.5	21.56	22.14	26.43	27.24
13.0	21.91	22.58	26.84	27.76
13.5	22.27	22.98	27.25	28.20
14.0	22.62	23.34	27.63	28.57
14.5	22.96	23.66	27.98	28.87
15.0	23.29	23.94	28.30	29.11
15.5	23.60	24.17	28.60	29.29
16.0	23.90	24.37	28.88	29.43
16.5	24.19	24.54	29.14	29.56
17.0	24.46	24.70	29.41	29.69
17.5	24.73	24.85	29.70	29.84
18+	25.00	25.00	30.00	30.00

Mid-year age points were chosen as the age criteria in our study (for example, 12.5 for 12-year-olds).

Appendix 2

The variable names notations are explained below.

- PACBDEE – Energy expenditure / Physical Activity
- JOB_INC – Socioeconomic status
- HWTBDBMI – Body Mass Index
- DEPRESS – Number of years of major depression persistence (depression lifetime)
- DEPLIFE – Major Depressive Episode – Life
- DEPYEAR – Major Depressive Episode - 12 month
- DHHB_SEX – Gender
- STRB_1 – Self-perceived ability to handle an unexpected problem
- WSTBDSOC – Work Stress Scale - social support

Table A2. Descriptive statistics

VARIABLE	MEAN	MEDIAN	SD	VARIANCE	RANGE	MISSING
Physical activity	2.27	1.6	2.35	5.51	28.70	1
SES	0	.33	1.00	1.00	4.95	2073
STRESS	2.31	2.0	0.93	0.86	4.00	24
Work stress	4.0	4.0	2.11	4.464	12.00	4318
BMI	25.75	25.0	4.83	23.38	48.20	1415
Depression	0.41	0	3.04	9.24	67.00	94

Appendix 3

Output for Model 1:

Analysis Summary

Groups

Group number 1 (Group number 1)

Notes for Group (Group number 1)

```
The model is nonrecursive.
Sample size = 12376
```

Variable Summary (Group number 1)

Variable counts (Group number 1)

```
Number of variables in your model:  18
Number of observed variables:        8
Number of unobserved variables:     10
Number of exogenous variables:      10
Number of endogenous variables:      8
```

Models

Default model (Default model)

Notes for Model (Default model)

Computation of degrees of freedom (Default model)

```
            Number of distinct sample moments:  44
Number of distinct parameters to be estimated:  28
                 Degrees of freedom (44 - 28):  16
```

Result (Default model)

```
Iteration limit reached
The results that follow are therefore incorrect.
Chi-square = 209.134
```

Degrees of freedom = 16
Probability level = .000

Group number 1 (Group number 1 - Default model)

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
depress <--- Environ	-11.523	6.931	-1.663	.096
hwtbdbmi <--- Genes	1.000			
hwtbdbmi <--- Environ	1.976	1.077	1.835	.067
depress <--- Genes	133.188	1576.335	.084	.933
Job_inc <--- Environ	1.000			
wstbdsoc <--- Environ	-.573	.112	-5.115	***
pacbdee <--- Environ	1.427	.118	12.138	***
etabdind <--- Environ	-6.284	.997	-6.305	***
dhhb_sex <--- Genes	1.000			
depb_88 <--- Genes	-.378	.322	-1.174	.240
dhhb_sex <--- Environ	-4.770	2.921	-1.633	.102
depress <--- hwtbdbmi	-.074	.104	-.714	.475
hwtbdbmi <--- depress	.425	.231	1.841	.066

Standardized Regression Weights: (Group number 1 - Default model)

	Estimate
depress <--- Environ	-.549
hwtbdbmi <--- Environ	.059
Job_inc <--- Environ	.145
wstbdsoc <--- Environ	-.039
pacbdee <--- Environ	.088
etabdind <--- Environ	-.103
dhhb_sex <--- Environ	-1.387
depress <--- hwtbdbmi	-.118
hwtbdbmi <--- depress	.267

Intercepts: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
hwtbdbmi	25.531	.107	239.069	***
depress	2.322	2.671	.869	.385
Job_inc	-.004	.010	-.458	.647
wstbdsoc	4.011	.024	170.514	***
etabdind	9.973	.206	48.430	***
dhhb_sex	1.543	.004	344.478	***
pacbdee	2.274	.021	107.753	***
depb_88	1.641	.077	21.288	***

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
Environ	.021	.013	1.610	.107
Genes	-.008	.099	-.084	.933
error6	24.446	1.459	16.750	***
error7	154.405	1751.417	.088	.930
error1	5.471	.074	73.776	***
error8	8.999	.328	27.465	***
error5	-.221	.301	-.735	.463
error4	77.568	2.614	29.675	***
error3	4.457	.070	63.406	***
error2	.978	.019	52.478	***

Notes for Model (Group number 1 - Default model)

The following variances are negative. (Group number 1 - Default model)

	Genes	error5
	-.004	-.222

Notes for Group/Model (Group number 1 - Default model)

This solution is not admissible.
Stability index for the following variables is .033
bmi
depress

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	28	209.072	16	.000	13.067
Saturated model	44	.000	0		
Independence model	8	1102.636	36	.000	30.629

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.810	.573	.822	.593	.819
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.031	.028	.035	1.000
Independence model	.049	.046	.051	.756

Output for Model 2:

Notes for Group (Group number 1)

The model is nonrecursive.
Sample size = 12376

Variable Summary (Group number 1)

Variable counts (Group number 1)

```
Number of variables in your model:  17
Number of observed variables:        8
Number of unobserved variables:      9
Number of exogenous variables:       9
Number of endogenous variables:      8
```

Models

Computation of degrees of freedom (Default model)

```
            Number of distinct sample moments:  44
Number of distinct parameters to be estimated:  28
                 Degrees of freedom (44 - 28):  16
```

Result (Default model)

Iteration limit reached
The results that follow are therefore incorrect.
Chi-square = 210.776
Degrees of freedom = 16
Probability level = .000

Group number 1 (Group number 1 - Default model)

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
dhhb_sex <--- Environ	-2.460	.783	-3.140	.002
depress <--- Environ	-1042.399	57255.697	-.018	.985
hwtbdbmi <--- Environ	2.167	.293	7.393	***
depress <--- dhhb_sex	-421.253	23142.892	-.018	.985
depress <--- depb_88	.046	.028	1.643	.100
hwtbdbmi <--- depb_88	-.025	.048	-.520	.603
Job_inc <--- Environ	1.000			
wstbdsoc <--- Environ	-.586	.112	-5.209	***
pacbdee <--- Environ	1.416	.116	12.166	***
etabdind <--- Environ	-6.252	.998	-6.265	***
depress <--- hwtbdbmi	-.200	.092	-2.173	.030
hwtbdbmi <--- depress	.625	.232	2.698	.007

Standardized Regression Weights: (Group number 1 - Default model)

	Estimate
dhhb_sex <--- Environ	-.997
depress <--- Environ	-69.243
hwtbdbmi <--- Environ	.091
depress <--- dhhb_sex	-69.030
depress <--- depb_88	.046
hwtbdbmi <--- depb_88	-.016
Job_inc <--- Environ	.202
wstbdsoc <--- Environ	-.056
pacbdee <--- Environ	.122
etabdind <--- Environ	-.143
depress <--- hwtbdbmi	-.318
hwtbdbmi <--- depress	.393

Intercepts: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
dhhb_sex	1.543	.004	344.477	***
depb_88	1.641	.077	21.287	***
hwtbdbmi	25.489	.121	211.354	***
depress	655.323	35700.374	.018	.985
Job_inc	-.005	.010	-.476	.634
wstbdsoc	4.017	.024	170.750	***

	Estimate	S.E.	C.R.	P
etabdind	9.981	.206	48.468	***
pacbdee	2.274	.021	107.753	***

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
Environ	.041	.013	3.044	.002
error8	8.995	.327	27.494	***
error5	.001	.078	.018	.985
error6	26.011	2.350	11.067	***
error7	-243.842	13956.164	-.017	.986
error1	5.432	.074	73.748	***
error4	76.786	2.595	29.589	***
error3	4.450	.070	63.367	***
error2	.959	.019	51.800	***

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	28	210.776	16	.000	13.173
Saturated model	44	.000	0		
Independence model	8	1102.636	36	.000	30.629

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.809	.570	.821	.589	.817
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.031	.028	.035	1.000
Independence model	.049	.046	.051	.756

Appendix 4

Output for Model 3:

Analysis Summary

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is nonrecursive.
Sample size = 12376

Variable Summary (Group number 1)

Variable counts (Group number 1)

Number of variables in your model:	13
Number of observed variables:	6
Number of unobserved variables:	7
Number of exogenous variables:	7
Number of endogenous variables:	6

Models

Default model (Default model)

Notes for Model (Default model)

Computation of degrees of freedom (Default model)

Number of distinct sample moments:	27
Number of distinct parameters to be estimated:	22
Degrees of freedom (27 - 22):	5

Result (Default model)

Minimum was achieved
Chi-square = 140.238
Degrees of freedom = 5
Probability level = .000

Group number 1 (Group number 1 - Default model)

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
pacbdee <--- Environ	1.053	.101	10.387	***
strb_1 <--- Environ	-.548	.044	-12.420	***
HWTBDBMI <--- Environ	1.738	.256	6.781	***
DEPRESS <--- Environ	-.963	.141	-6.821	***
HWTBDBMI <--- pacbdee	-.252	.021	-11.859	***
DEPRESS <--- strb_1	.274	.034	8.091	***
Job_inc <--- Environ	1.000			
dhhb_sex <--- Environ	-.600	.048	-12.385	***
HWTBDBMI <--- DEPRESS	-.282	.116	-2.439	.015
DEPRESS <--- HWTBDBMI	.150	.045	3.359	***

Standardized Regression Weights: (Group number 1 - Default model)

	Estimate
pacbdee <--- Environ	.183
strb_1 <--- Environ	-.240
HWTBDBMI <--- Environ	.146
DEPRESS <--- Environ	-.129
HWTBDBMI <--- pacbdee	-.122
DEPRESS <--- strb_1	.084
Job_inc <--- Environ	.407
dhhb_sex <--- Environ	-.490
HWTBDBMI <--- DEPRESS	-.177
DEPRESS <--- HWTBDBMI	.238

Intercepts: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
strb_1	2.309	.008	276.443	***
pacbdee	2.274	.021	107.753	***
HWTBDBMI	26.411	.085	310.527	***
DEPRESS	-4.068	1.165	-3.490	***
Job_inc	-.010	.010	-.982	.326

	Estimate	S.E.	C.R.	P
dhhb_sex	1.543	.004	344.477	***

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
Environ	.166	.016	10.211	***
error1	5.330	.072	73.743	***
error3	.812	.012	69.730	***
error6	23.748	.877	27.083	***
error7	9.313	.263	35.420	***
error5	.189	.006	33.818	***
error2	.835	.018	45.355	***

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	22	140.238	5	.000	28.048
Saturated model	27	.000	0		
Independence model	6	1341.695	21	.000	63.890

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.895	.561	.899	.570	.898
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.047	.040	.054	.776
Independence model	.071	.068	.075	.000

Output for Model 4:

Analysis Summary

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is nonrecursive.
Sample size = 12376

Variable Summary (Group number 1)

Variable counts (Group number 1)

Number of variables in your model:	13
Number of observed variables:	6
Number of unobserved variables:	7
Number of exogenous variables:	7
Number of endogenous variables:	6

Models

Default model (Default model)

Notes for Model (Default model)

Computation of degrees of freedom (Default model)

Number of distinct sample moments:	27
Number of distinct parameters to be estimated:	22
Degrees of freedom (27 - 22):	5

Result (Default model)

Minimum was achieved
Chi-square = 80.933
Degrees of freedom = 5
Probability level = .000

Group number 1 (Group number 1 - Default model)

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
pacbdee <--- Environ	1.249	.115	10.866	***
wstbdsoc <--- Environ	-.391	.112	-3.502	***
HWTBDBMI <--- Environ	2.046	.258	7.944	***
DEPRESS <--- Environ	-.784	.149	-5.279	***
HWTBDBMI <--- pacbdee	-.228	.021	-10.966	***
DEPRESS <--- wstbdsoc	.147	.016	9.260	***
Job_inc <--- Environ	1.000			
dhhb_sex <--- Environ	-1.111	.164	-6.776	***
DEPRESS <--- HWTBDBMI	.028	.043	.646	.518
HWTBDBMI <--- DEPRESS	.009	.110	.083	.934

Standardized Regression Weights: (Group number 1 - Default model)

	Estimate
pacbdee <--- Environ	.162
wstbdsoc <--- Environ	-.056
HWTBDBMI <--- Environ	.129
DEPRESS <--- Environ	-.078
HWTBDBMI <--- pacbdee	-.111
DEPRESS <--- wstbdsoc	.103
Job_inc <--- Environ	.304
dhhb_sex <--- Environ	-.678
DEPRESS <--- HWTBDBMI	.045
HWTBDBMI <--- DEPRESS	.006

Intercepts: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
wstbdsoc	4.028	.024	170.953	***
pacbdee	2.274	.021	107.753	***
HWTBDBMI	26.229	.079	331.017	***
DEPRESS	-.900	1.120	-.804	.421
Job_inc	-.007	.010	-.750	.453
dhhb_sex	1.543	.004	344.477	***

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
Environ	.092	.015	6.215	***
error3	4.478	.071	63.254	***
error1	5.369	.072	74.208	***
error6	22.792	.351	64.889	***
error7	9.062	.117	77.230	***
error5	.134	.017	7.910	***
error2	.907	.019	48.712	***

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	22	80.933	5	.000	16.187
Saturated model	27	.000	0		
Independence model	6	978.662	21	.000	46.603

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.917	.653	.922	.667	.921
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.035	.029	.042	1.000
Independence model	.061	.057	.064	.000

Ouput for Model 5:

Analysis Summary

Group number 1 (Group number 1)

Notes for Group (Group number 1)

The model is recursive.
Sample size = 12376

Variable Summary (Group number 1)

Variable counts (Group number 1)

Number of variables in your model:	13
Number of observed variables:	6
Number of unobserved variables:	7
Number of exogenous variables:	7
Number of endogenous variables:	6

Models

Default model (Default model)

Notes for Model (Default model)

Computation of degrees of freedom (Default model)

Number of distinct sample moments:	27
Number of distinct parameters to be estimated:	21
Degrees of freedom (27 - 21):	6

Result (Default model)

Minimum was achieved
Chi-square = 80.939
Degrees of freedom = 6
Probability level = .000

Group number 1 (Group number 1 - Default model)

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
pacbdee <--- Environ	1.249	.115	10.873	***
HWTBDBMI <--- Environ	2.039	.240	8.481	***
wstbdsoc <--- Environ	-.390	.111	-3.501	***
HWTBDBMI <--- pacbdee	-.228	.021	-10.981	***
DEPRESS <--- Environ	-.789	.135	-5.855	***
Job_inc <--- Environ	1.000			
dhhb_sex <--- Environ	-1.114	.162	-6.870	***
DEPRESS <--- wstbdsoc	.147	.016	9.229	***
DEPRESS <--- HWTBDBMI	.032	.006	5.176	***

Standardized Regression Weights: (Group number 1 - Default model)

	Estimate
pacbdee <--- Environ	.162
HWTBDBMI <--- Environ	.128
wstbdsoc <--- Environ	-.056
HWTBDBMI <--- pacbdee	-.111
DEPRESS <--- Environ	-.079
Job_inc <--- Environ	.304
dhhb_sex <--- Environ	-.679
DEPRESS <--- wstbdsoc	.102
DEPRESS <--- HWTBDBMI	.050

Intercepts: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
pacbdee	2.274	.021	107.753	***
wstbdsoc	4.028	.024	170.951	***
HWTBDBMI	26.233	.066	397.825	***
Job_inc	-.007	.010	-.750	.454
dhhb_sex	1.543	.004	344.477	***
DEPRESS	-.989	.171	-5.780	***

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P
Environ	.092	.015	6.275	***
error1	5.369	.072	74.357	***
error6	22.805	.317	71.963	***
error3	4.478	.071	63.259	***
error5	.134	.017	7.989	***
error2	.908	.018	49.109	***
error7	9.062	.117	77.283	***

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	21	80.939	6	.000	13.490
Saturated model	27	.000	0		
Independence model	6	978.662	21	.000	46.603

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.917	.711	.923	.726	.922
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.032	.026	.038	1.000
Independence model	.061	.057	.064	.000

Appendix 5

The command syntax used in Mplus for depression measured by the categorical variable DEPLIFE:

```
TITLE:    Obesity-Depression: CFA with categorical outcomes
DATA:      FILE IS
C:\\Documents and Settings\Alina\My Documents\PROIECTUL\dataMplus.dat ;
VARIABLE:
  NAMES ARE                 pacbdee
                            job_inc
                            strb_1
                            dhhb_sex
                            BMI
                            DEPLIFE ;
  USEVARIABLES ARE   pacbdee- DEPLIFE ;
  MISSING ARE ALL (-9);
  CATEGORICAL ARE strb_1 - DEPLIFE ;
ANALYSIS:    TYPE = general  missing h1;
             PARAMETERIZATION = THETA;
             ITERATIONS = 10000000;


MODEL:
                 environ by job_inc@1 pacbdee strb_1 dhhb_sex;
                 BMI DEPLIFE on environ;
                 BMI on pacbdee;
                 DEPLIFE on strb_1;
                 BMI on DEPLIFE;
```

```
                  DEPLIFE on BMI;

OUTPUT:      SAMPSTAT MODINDICES RESIDUAL STANDARDIZED CINTERVAL H1SE
             H1TECH3  PATTERNS  TECH1 TECH3 TECH4 TECH5;
```

OUTPUT FROM MPLUS for deplife:

```
                    Estimates     S.E.  Est./S.E.
 ENVIRON  BY
    JOB_INC             1.000    0.000      0.000
    PACBDEE             1.353    0.124     10.925
    STRB_1             -0.462    0.050     -9.294
    DHHB_SEX          143.128  *******      0.009

 BMI       ON
    ENVIRON             0.242    0.064      3.765
 DEPLIFE  ON
    ENVIRON            -0.680    0.078     -8.763
 BMI       ON
    PACBDEE            -0.051    0.005    -10.518
    DEPLIFE            -0.362    0.051     -7.126
 DEPLIFE  ON
    STRB_1              0.250    0.014     17.888
    BMI                 0.398    0.054      7.384
 Intercepts
    PACBDEE             2.274    0.029     78.715
    JOB_INC             0.000    0.011      0.003
 Thresholds
    STRB_1$1           -0.877    0.013    -67.343
    STRB_1$2            0.297    0.011     25.943
    STRB_1$3            1.289    0.015     83.528
    STRB_1$4            2.105    0.027     77.496
    DHHB_SEX$1         -0.003    0.320     -0.009
    BMI$1              -0.107    0.015     -7.021
    BMI$2               0.816    0.021     39.399
    DEPLIFE$1           1.087    0.027     39.722
 Variances
    ENVIRON             0.000    0.011     -0.005

 Residual Variances
    PACBDEE             5.514    0.042    130.057
    JOB_INC             1.000    0.024     42.388
```

MODEL MODIFICATION INDICES

```
Minimum M.I. value for printing the modification index    10.000

                                M.I.     E.P.C.  Std E.P.C.  StdYX E.P.C.
```

```
ON Statements

PACBDEE   ON STRB_1          98.255    -0.213     -0.213      -0.002
PACBDEE   ON BMI             12.878     0.391      0.391       0.161
PACBDEE   ON DEPLIFE         12.860    -0.141     -0.141      -0.058
STRB_1    ON PACBDEE         98.278    -0.039     -0.039      -3.338
STRB_1    ON BMI             12.861     0.126      0.126       4.491
STRB_1    ON DEPLIFE         12.846     0.317      0.317      11.274

WITH Statements

DHHB_SEX WITH JOB_INC       999.000     0.000      0.000       0.000
STRB_1   WITH PACBDEE        98.253    -0.213     -0.213      -3.337
```

The command syntax used in Mplus for depression measured by categorical variable DEPYEAR:

```
TITLE:    Obesity-Depression: CFA with categorical outcomes

DATA:       FILE IS
C:\\Documents and Settings\Alina\My Documents\PROIECTUL\dataMplusY.dat
;

VARIABLE:

  NAMES ARE                  pacbdee
                             job_inc
                             strb_1
                             dhhb_sex
                             BMI
                             DEPyear;

  USEVARIABLES ARE   pacbdee - DEPyear;
  MISSING ARE ALL (-9);
  CATEGORICAL ARE strb_1 - DEPyear;

ANALYSIS:   TYPE = general  missing h1;
            PARAMETERIZATION = THETA;
            ITERATIONS = 10000000;

MODEL:

                environ by job_inc@1 pacbdee strb_1 dhhb_sex;
                BMI depyear on environ;
                BMI on pacbdee;
                DEPyear on strb_1;
                BMI on DEPyear;
                DEPyear on BMI;
```

```
                PACBDEE   ON STRB_1;


OUTPUT:        SAMPSTAT MODINDICES RESIDUAL STANDARDIZED CINTERVAL H1SE
H1TECH3
 PATTERNS  TECH1 TECH3 TECH4 TECH5;
```

OUTPUT FROM MPLUS for depyear:

```
MODEL RESULTS

                       Estimates     S.E.  Est./S.E.
 ENVIRON  BY
    JOB_INC               1.000     0.000      0.000
    PACBDEE               1.266     0.121     10.429
    STRB_1               -0.462     0.050     -9.298
    DHHB_SEX            145.128   *******      0.008

 BMI       ON
    ENVIRON               0.318     0.062      5.142
 DEPYEAR  ON
    ENVIRON              -0.606     0.097     -6.226
 BMI       ON
    PACBDEE              -0.052     0.005    -10.563
    DEPYEAR              -0.259     0.035     -7.430

 DEPYEAR  ON
    STRB_1                0.364     0.018     20.611
    BMI                   0.270     0.042      6.450

 PACBDEE  ON
    STRB_1               -0.190     0.022     -8.730

 Intercepts
    PACBDEE               2.274     0.029     78.714
    JOB_INC               0.000     0.011      0.003

 Thresholds
    STRB_1$1             -0.877     0.013    -67.314
    STRB_1$2              0.297     0.011     25.930
    STRB_1$3              1.289     0.015     83.313
    STRB_1$4              2.105     0.027     77.343
    DHHB_SEX$1           -0.003     0.346     -0.008
    BMI$1                -0.117     0.016     -7.438
    BMI$2                 0.843     0.019     44.382
    DEPYEAR$1             1.626     0.029     56.705
 Variances
    ENVIRON               0.000     0.011     -0.004
 Residual Variances
    PACBDEE               5.478     0.043    127.954
    JOB_INC               1.000     0.023     42.960
```