1-1-2010

Alcohol Use Disorders And Labor Market Outcomes: An Analysis Using 2001-02 National Epidemiology Survey On Alcohol And Related Conditions

Shammima Jesmin
Wayne State University

ALCOHOL USE DISORDERS AND LABOR MARKET OUTCOMES: AN ANALYSIS USING 2001-02 NATIONAL EPIDEMIOLOGY SURVEY ON ALCOHOL AND RELATED CONDITIONS

by

SHAMMIMA JESMIN

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2010

MAJOR: ECONOMICS

Approved by:

____
Advisor Date

____

____

____

____

DEDICATION

I would like to dedicate this piece of my endeavor to my parents, who sacrificed all they could to educate me and encouraged me for higher academic pursuits and excellence. Mother and Father, you will always remain the most important persons in my life.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my academic and dissertation advisor, Dr. Allen Goodman, for his guidance, support, patience, and encouragement throughout this dissertation and all my years at Wayne State University. My thanks will never be enough to the support that I have got from Dr. Allen. Acceptance in the Ph. D. program was a special consideration of him. In this regard, I must mention about the special favor I received from two of my teachers in Masters Program at Eastern Michigan University, Dr. Abdullah Dewan and Dr. Sharon Erenburg. They took personal initiative and spent their valuable time to drive me to Dr. Allen's office for an interview to be accepted in the Ph. D. program. I had to take some time off from the program at the end of fourth year because I had to move to Canada to resolve immigration issue for my entire family, and I could not think I would be able to return to the program with so many hurdles on the way. Once again, the opportunity to return to the program was a special gift of Dr. Allen. He also took special initiatives so many times to help me out from the situations that were almost going to disrupt my ultimate academic goal. I also feel very fortunate to have such an advisor who gave insightful feedback and technical and theoretical assistance that I needed at each stage of the dissertation.

I would also like to gratefully acknowledge my Dissertation Committee Members, Dr. Gail Jensen Summers, Dr. Janet Hankin and Dr. Stephen J. Spurr for their input and suggestions during the dissertation work. This dissertation is as much as a product of their work as is mine, because many of the ideas were hammered out from the discussions with them. I am forever grateful to Dr. Janet for directing me to choose the dissertation topic as well as the source of reliable data while I was finding myself in unsettled territory. I must mention that the idea of my research came from a term paper on alcohol use related issues that I did for Dr Janet's course 'Medical Sociology'. I am especially thankful to Dr. Gail for guiding me to use the accurate estimation strategy. I am also deeply indebted to Dr. Spurr for his suggestions to consider some of the explanatory variables that could have affect on the subject matter of this research.

I am also deeply indebted to Ms. LaVerna Patrick, Ms. Cheri Miller and Ms. Delores Tennille for all their office support during the years I have spent in Department of Economics. I have accumulated many debts in the course of writing this dissertation. It would be difficult to mention everyone who has supported my work. To all my teachers and friends at Department of Economics, Wayne State University, I would like to say 'thank you' for their support throughout the dissertation-writing period.

I would also like to gratefully express my thanks to my family for the prayers, support and patience while I worked on my dissertation. I am grateful to my husband, Mostafiz Khan, for his love, support and encouragement throughout this process, and also for the patience in proof reading my dissertation paper. To my son, Shihab Khan, I appreciate the support that you gave me time to time in solving computer application related problems. To my daughters, Bushra Mehzabin and Mashira Mehzabin, I appreciate your love and support and how you always tried not to disturb when I was busy in my dissertation work. Finally, my special thanks go to one of my family friend, Ms. Shahana Ahmed (University of Toronto), for her time and patience in proof reading the dissertation.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

APA	American Psychiatric Association
CHD	Coronary Heart Disease
DSMALC	Diagnostic and Statistical Manual Criterion
DSM-III	Diagnostic and Statistical Manual of Mental Disorders, III
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, IV
ECA	Epidemiological Catchments Area
FASD	Fetal Alcohol Spectrum Disorders
GMM	Generalized Method of Moments
HDL	High Density Lipoprotein-cholesterol
IV	Instrumental Variables
LFPFULL	Labor Force Participation Full Time
LIML	Limited Information Maximum Likelihood
MLE	Maximum Likelihood Estimation
NLSY	National Longitudinal Survey of Youth
NLAES	National Longitudinal Alcohol Epidemiologic Survey
NESARC	National Epidemiology Survey on Alcohol and Related Conditions
NHSDA	National Household Surveys on Drug Abuse
NIAAA	National Institute of Alcohol Abuse and Alcoholism
NUMC	Number of Cases in the Data Set
OECD	Organization for Economic Co-operation and Development
OLS	Ordinary Least Squares

QES	Quality of Employment Survey
SAMHSA	Substance Abuse and Mental Health Services Administration
2SLS	Two Stage Least Square
WHO	World Health Organization

LIST OF FIGURES

LIST OF TABLES

LIST OF ANNEXES

CHAPTER 1

INTRODUCTION

In American society like many other societies, alcohol consumption [1] is widely accepted as a pleasurable social experience for most adults as long as it causes no harm to themselves or others. Although not intended by any of the alcohol consumers, it is well documented that harmful effects attributed by alcohol use and misuse (alcohol use disorders)[2] may result in short and long-term physical, psychological and cognitive impairment (NIAAA: 1993, 2000, World Health Organization-WHO: 2002, 2004). Due to these multi-dimensional adverse affects (NIAAA-Strategic Plan 2001-2005: 2001) alcohol is identified as one of the leading causes[3] of preventable death, illness, injury and accident, disability and suffering (Heather: 2001, Horgan: 2001, Farrell et al: 2003, WHO: 2004). The recent report of WHO (2004), based on latest available data of 2002, indicated that globally alcohol caused 1.8 million deaths (3.2% of total), a loss of 58.3

---

[1] The National Institute on Alcohol Abuse and Alcoholism (NIAAA: 2000) defined "alcohol consumption" in terms of standard drinks consumed. One standard drink is one absolute ounce of alcohol. This includes: 12 ounces of beer, 5 ounces of wine, or 1.5 ounces (a jigger) of 90−proof liquor.

[2] To date, different phrases have been found in literature to describe alcohol use disorders such as alcohol misuse or excessive drinking or heavy drinking or problem drinking (alcoholism) or alcoholic. Modern studies preferred the phrases "alcohol dependence" and "alcohol abuse" to describe alcohol use disorders. APA (American Psychiatric Association) recognizes two alcohol use disorders, including alcohol abuse and alcohol dependence (alcoholism) based on DSM-IV criterion (Diagnostic and Statistical Manual of Mental disorders, 4th ed, Washington, D.C. 1994). These disorders are the result of excessive use of alcohol.

[3] The WHO (2004) reported that alcohol use is linked to more than 60 types of diseases and injuries, and it is the leading risk factor for disease burden in developing countries and the third largest risk factor in developed countries.

million (4% of total) of Disability-Adjusted Life Years (DALYS)[4], and 76.3 million or 3.8 percent of total alcohol users was diagnosed with alcohol use disorders. Besides imposing large morbidity and mortality costs, alcohol use imposes substantial financial burden[5] (WHO: 2004) to the society for lost productivity due to reduced employment, working hours, job performance and earnings (NIAAA: 1993, 2001, WHO: 2002, 2004).

The broad range and severity of alcohol-related consequences and related substantial costs makes this a major public health, social and economic issue worldwide (NIAAA: 1993, 2000, 2001, WHO: 2002, 2004). It is also a serious cause of concern for the United States because alcohol is the most widely used drug nationwide[6] and also identified as the most "acutely destructive" (Schuckit: 2000) in terms of causing death and disease[7], affecting workforce adversely, and imposing substantial financial burden (WHO: 2004, Mokdad et al: 2004, NIAAA: 2001). The enormous impact of alcohol use on an individual's personal, family and social life has led researchers to explore different alcohol related issues from various perspectives.

---

[4] The WHO (2004) defines "DALYS" as a measure that combines mortality in terms of life years lost due to premature death, and morbidity in terms of life years lived in disability.

[5] The WHO Report (2004) provided more details on country-specific alcohol related costs.

[6] About three-quarters of the adult population drink (Apgar and Burgess: 2001).

[7] Alcohol use and misuse is the third leading cause of preventable death and disease in US (WHO: 2004, Mokdad et al: 2004). It causes approximately 100,000 deaths or 5 percent of all death annually (NIAAA-Strategic Plan 2001-2005: 2001).

The main focus of this research (see more details in section 1.3.2) is to explore to what extent alcohol use disorders affect labor force participation and employment income. I use the largest nationally representative sample[8] of the United States population (between 18 and 98 years of age) collected by the NESARC (National Epidemiologic Survey on Alcohol and Related Conditions) in 2001-2002. With an expectation to contribute in the continuing process (Mullahy and Sindelar: 1996, NIAAA: 2000) of alcohol research and to understand the complex nature of the relationship between alcohol use and labor market outcomes, I proceed by focusing on two issues.

First, I consider the importance of exclusive classification[9] of observations based on individual's past (before 2000) and present (2000-2001) alcohol use record and clinical diagnosis[10] which identified someone as regular or casual drinker or alcohol abuser or alcohol dependent.[11] It is expected to provide more complete comparison in terms of the effect of drinking on labor market outcomes based on the individual's alcohol use status (see Tables 2A and 2B for definition and descriptions).

---

[8] I am very much grateful to my advisor Prof. Allen Goodman (Undergraduate and Ph.D. advisor, Economics Department, Wayne State University, Michigan, USA) for making the full data set available for this research.

[9] Johansson and others (2004), in a recent study using Finnish data (2000), indicated the significance of this issue. For example, they pointed out that empirical results could be misleading if someone considers ex-drinkers (who stopped drinking, named "dry alcoholics") as abstainers.

[10] The NESARC followed clinical-diagnostic criteria described (see Table 2) in the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV).

[11] See details in Chapter 4, section 4.2.1.

Second, I also address[12] the issue of endogeneity[13] to estimate the impact of alcohol use on labor market outcomes. Previous alcohol studies indicated that most often endogeneity arises due to the fact that some common and unobserved individual factors influence both alcohol use and labor market outcomes and because of this problem standard estimation methods such as Ordinary Least Square (OLS) and Maximum Likelihood Method (MLE) would produce biased and inconsistent estimates[14] if someone does not address the problem with appropriate estimation technique (Greene: 2000). I perform formal tests to detect the presence of endogeneity. Traditionally, this problem is addressed by using one of the instrumental variables (IV) estimation techniques[15]. In order to choose appropriate IV method of estimations, I also perform the test (as suggested by Baum et al.: 2003) for the possible presence of heteroscedasticity (which is most common in cross-section data)[16]. As the diagnostic

[12] See details in Chapter 3, section 3.2.

[13] Endogeneity usually arises due to omitted variables, measurement error, simultaneity bias, or a combination of these factors (Wooldridge: 2002).

[14] It violates one of the fundamental assumptions of classical regression model that independent variable included in the model should be uncorrelated with error terms (Greene: 2000).

[15] Baum et al (2003) indicated the following IV methods such as Two Stage Least Square (2SLS), Generalized Method of Moments (GMM), and Limited Information Maximum Likelihood (LIML) should be selected respectively for the existence of “endogeneity”, “endogeneity and heteroscedasticity”, and “weak instruments”. See details in Chapter 3, section 3.2.

[16] One of the classical assumptions of the ordinary regression model is that the disturbance variance is constant or homogeneous across observations. If this assumption is violated, the errors are said to be "heteroscedastic." If heteroscedasticity is present, the estimated parameters are still consistent but they are no longer efficient. Thus, inferences from the standard errors are likely to be misleading (Gujarati: 1995).

tests results (see Chapter 5) confirmed the presence of endogeneity and heteroscedasticity in the current model, I choose Generalized Method of Moments (GMM) estimation method to estimate labor market outcome equations.

The analysis of existing nature of the problem related to alcohol use justifies my attempting further research on the association of alcohol use and labor market variables.

1.1 Nature of the problem.

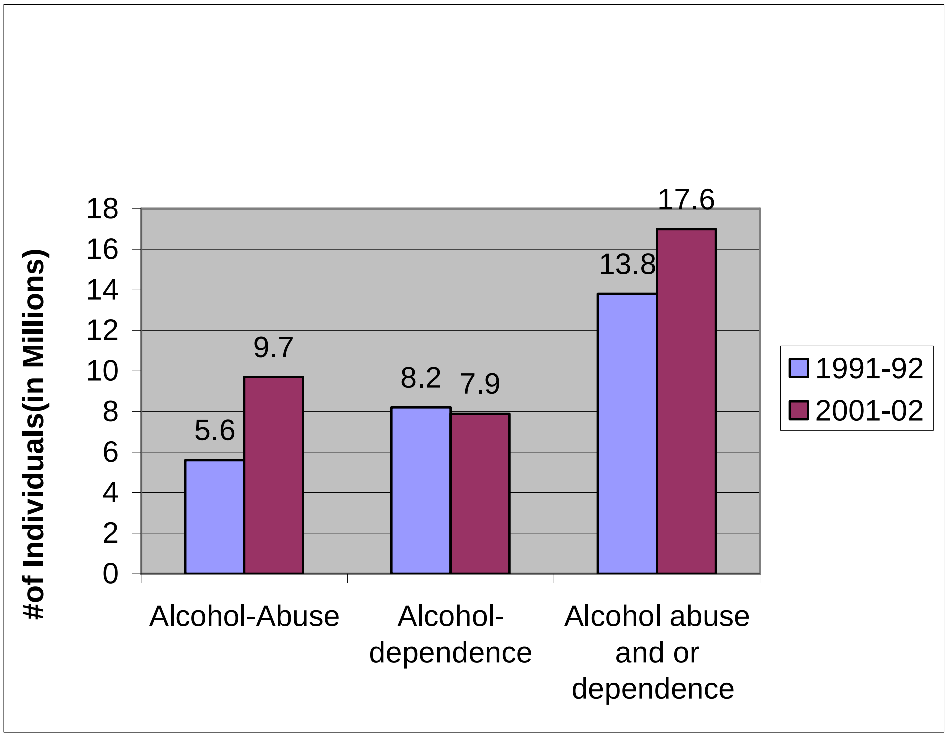
1.1.1 Americans with alcohol use disorders and working population

Alcohol use disorders (alcohol abuse and dependence) are among the most prevalent illnesses in American society (Swift: 2001). Based on recent NESARC survey (2001-02), the NIAAA (2004) indicated (Figure-1) that the prevalence of 12-month alcohol use disorders (alcohol abuse or dependence) among American adults (based on DSM-IV diagnosis criteria) had been increased from 13.8 million in 1991-92 to 17.6 million in 2001-02 (NIAAA: 2004). The prevalence of alcohol abuse alone increased from 5.6 million to 9.7 million across the decade, while the prevalence of alcohol dependence, commonly known as alcoholism, declined from 8.2 to 7.9 million.

Although the alcohol dependence declined, the rate of alcohol abuse and dependence among younger aged population (18-29 and 30-44 years) increased. It is considered as a worrisome fact because substantial number of “working population” may suffer from alcohol use disorders (NIAAA: 2004). In an earlier study, it was reported that 85 percent of heavy drinkers (those who consume five or more drinks at one time each week) in the United States were employed (National Household Survey on Drug Abuse-NHSDA, Substance Abuse and Mental Health Services Administration -

SAMHSA: 1999). A recent study (Frone: 2006) reported that alcohol use related impairment affected about 15 percent of the U.S. workforce.

Figure 1: Prevalence of 12-Month Alcohol Abuse and Dependence in US based on DSM-IV diagnosis criteria: 1991-92 and 2001-02



Source: NIAAA Newsletter, No. 4, Spring/Summer 2004.

1.1.2 The interrelationship between alcohol use, health and work

Poor health status caused by excessive (Farrell et al: 2003) or even occasional[17] (Mullahy and Sindelar: 1996) alcohol use can be attributed to a wide range of adverse health consequences because of the adverse affect on each of the six primary health

---

[17] Mullahy and Sindelar (1996) reported the following adverse consequences of alcohol consumption, such as, reduced hand eye coordination, preoccupation with alcohol, significant time spent seeking alcohol, unusual or unstable behavior, increased accidents and injuries, liver and heart damage are the direct adverse outcomes.

status dimensions (Ware: 1986)[18]. Examples of serious health problems (Ezzati et al: 2002) includes liver disease, heart disease (high blood pressure, cardiovascular disease, certain kinds of stroke), cancer (esophagus, mouth, throat, voice box, colon and rectum), pancreatitis (inflammation of the pancreas), head trauma, depression (NIAAA: 1996, and Weiss and Lonnquist: 1997), and damage to the nervous system (OECD: 2003). Studies have shown that compared to moderate drinkers[19], binge drinkers[20] have an increased risks of angina (heart pain) and a six times higher risk of having a fatal heart attack (Walitzer: 1999). Adverse health impacts of alcohol use have been linked to the reduction of labor market productivity[21], poor job experience and lower educational attainment (Cook and Moore: 1993, Mullahy and Sindelar: 1993, 1994, 1996, Farrell et al: 2003). Besides affecting current workforce, alcohol use of current generation may have affect on the future generation through instable family environment, lower health and economic status and creating brain anomalies in children including FASD- Fetal Alcohol Spectrum Disorders (Russo et al: 2004). For example, in 1996, six million children lived with parents who were alcohol dependents, and

[18] Ware (1986) identified six primary health dimensions from the extensive review on literature on health. Those are: physical functioning, mental health, social well-being, role functioning, general health perceptions, and physical and psycho-physiologic symptoms.

[19] Moderate drinkers use approximately 2 or 2.5 drinks per day on average (French et al: 1995).

[20] Binge drinking: it is seen as drinking that occurs at a hazardous level- five or more drinks for men, and three or more drinks for women (See Annex 1)

[21] The NIAAA (1993, 2001) and the WHO (2002, 2004) reported that it was the result of reduction in employment, working hours (work less hours or frequent absence from the job), job performance and earnings.

approximately one out of ten American children lived in a household where at least one parent met clinical standards for alcohol and/or illicit drug use dependence (Chatterji and Markowitz: 2001 and Horgan: 2001).

1.1.3 Lost productivity and enormous economic costs

Based on the findings of five major studies (Berry et al: 1977, Cruze et al: 1981, Harwood: 1984, Rice et al: 1990, Harwood et al: 1998, 2000), the NIAAA (2001) indicated that in US, the costs of multidimensional health problems and lost productivity due to alcohol use were substantial[22]. For example, according to the most cited and latest available estimate (Harwood: 1998, 2000), the economic costs of alcohol abuse were $184.6 billion in 1998 which was equivalent to 2 percent of gross domestic product (GDP) for that year (Terza: 2002). The breakdown of this cost indicated that 14.3 percent ($26.3 billion) was for health care costs, and more than 70 percent (134.2 billion) was for productivity losses due to lost employment, job absences and reduced earnings.

1.1.4 Non-consensus empirical evidence

Since alcohol use imposed significant financial burden to the society for lost labor productivity by reducing employment, working hours, job performance and earnings (Rice et al. 1990, Kenkel and Ribar: 1994, Mullahy and Sindelar: 1993, 1996, NIAAA: 2001), researchers examined the relationship between alcohol use and labor market outcomes (employment/unemployment, hours of work, wage/earnings) with special

[22] Several other empirical works also provided the evidence of alcohol related economic costs (Farahati: 2003, Feng et al: 2001, Zarkin, et al: 1998, French et al: 1998, Mullahy and Sindelar :1996, and Rice et al: 1991) and non-economic costs (Russo: 2004, Cook and Moore: 1993, NIAAA: 1993, and Mullahy and Sindelar: 1989, 1994).

interests[23]. The magnitude and the direction of impact of alcohol use on labor market outcomes varied (negative, positive, inverted-U shaped, inverted-J shaped) widely across empirical studies. The evidence of negative impact of alcohol use on labor market outcomes were found in several major studies (Fisher[24]: 1927, Mullahy and Sindelar: 1989, 1991, 1993, 1994, 1995, and 1996, Bryant et al: 1992, 1993, and 1996, Cook and Moore: 1993, Kenkel and Ribar: 1994, Yamada, et. al.: 1996, Chatterji: 1998). Berger and Leigh (1988) took the leading role in reporting the first evidence of positive relationship between drinking and earning. Although the work of Berger and Leigh (1988) was highly criticized,[25] later the same work spirit was adopted in some major works (Cook: 1991, French and Zarkin: 1995, Heien: 1996, Hamilton and Hamilton: 1997, Zarkin et al.:1998, MacDonald and Shields: 2001, and Auld: 2005). These studies indicated positive, and even a non-linear (inverted-U shaped, inverted-J shaped) relationship between alcohol consumption and labor market outcomes.

Variation in the results were mainly due to differences in alcohol use measures, model specifications, differences in sample, and use of different econometric methods.

---

[23] The major empirical works on alcohol use and labor market outcomes done by Berger and Leigh (1988), Cook (1991), Kenkel and Ribar (1994), French and Zarkin (1995), Heien (1996), Mullahy and Sindelar (1993, 1996), Zarkin et al. (1998), Hamilton and Hamilton (1997), MacDonald and Shields (1998). See Tekin (2002) for an overview of the literature on alcohol use and labor market outcomes.

[24] Fisher (1927) did the first research in the US context and reported that the drinkers were less productive and extrapolated that the prohibition would increase the nation's productivity by 5 percent.

[25]Berger and Leigh (1988) did not separate the alcohol and non alcohol factors, and did not consider the self selection problem in choosing the work (Heien: 1996, Peters: 2004).

Therefore, much can be explored using better alcohol use measurement, rich data sets, appropriate model and method estimation (Mullahy and Sindelar: 1996, NIAAA: 2000).

1.2 Motivation of the research

The motivation for this research derives from three major considerations.

1.2.1 Alcohol use and misuse[26] is a major social and economic problem

It is still considered an interesting research topic because it is a major social and economic problem, and there is a key role of alcohol use in shaping human capital endowment and health status in terms of complete and full functioning physical, social and mental well being (WHO:1986), and in achieving labor market success (Currie and Madrian: 1999). Also, a recent NIAAA directed survey conducted by NESARC (a representative survey of American adults 18 years of age and older) on alcohol use and related conditions calls for special attention to estimate the impact on labor market because the survey reported that the prevalence of higher rates of alcohol abuse and dependency among younger aged population (18-29 and 30-44 years) in 2001-2002 represents substantial portion of work force.

1.2.2 Ability to test hypotheses using a large and rich data set

As the existence of conflicting findings keep the opportunity open to analyze the same topic based on richer and larger data sets, this research seeks to estimate to what

[26] Please see footnotes 1 and 2.

extent the impact of alcohol use and misuse (alcohol use disorders[27]) affects labor market success in terms of the probability of full time employment and annual earnings from full time work based on newly available data from 2001-02 National Epidemiology Survey on Alcohol and Related Conditions (NESARC). Some unique features of this dataset can be mentioned here.

First, it is the largest survey ever conducted on alcohol use and co-occurring conditions.

Second, the most important advantage of using this data set is that it gives opportunity to use clinically diagnosed or medically verified alcohol use measures directly such as ‘alcohol abuse’ and ‘alcohol dependent’ (based on DSM-IV criterion) instead of self selected criteria[28]. This is very crucial because prior research indicated that the direction and magnitude of the result based on self selected alcohol measures were questionable (Mullahy and Sindelar: 1995 and Heien: 1996). Also, the dataset contains very detailed information on individual drinking behavior indicating number of drinks used in a day, week or a month, ethanol content and drinking frequency such as light, moderate or heavy alcohol consumption. One advantage of using clinically diagnosed alcohol use measures was also indicated in some prior research. This

---

[27] The current research considers two alcohol use disorders such as alcohol abuse and alcohol dependence as defined in NESARC data set based on DSV-IV criteria (see Annex 2).

[28] Some research did not classify alcohol drinker or alcohol measure according to standard clinical diagnosis.

suggested that the estimation problem related to endogeneity[29] of alcohol use variable measure could be less if someone uses clinically diagnosed alcohol use measures. Buchmueller and Zuvekas (1998) provided the argument that the correlation between alcohol use variable and labor market variables is very unlikely where clinical based alcohol measures are included in the analysis, because clinical based alcohol measures are less influenced by income than frequency of use, physiological responses to alcohol use, and genetically determined factors.

Third, a more comprehensive set of explanatory variables as well as instrumental variables[30] are available for the current research as it contains large number of variables with richer information on health, socioeconomic, and other background data.

Fourth, this data set contains information on labor market outcomes such as full time and part time work participation, personal income, household income, and other sources of income. It also contains very detailed information on different occupations and industry types.

1.2.3 The extension of prior work

The main idea and strategy of this research were largely influenced by the pioneer work of Mullahy and Sindelar (1993 and 1996) as they pointed out the way

---

[29] It arises in empirical estimation mainly due to two facts. First, the ‘simultaneity’ arises due to the fact that the estimation of labor market variable includes alcohol use measure as an explanatory variable when actually alcohol use measures and labor market variables jointly determined by individual characteristics. Second, the ‘unobserved heterogeneity’ arises due to the existence of unobserved characteristics that may be correlated with both alcohol consumption and labor market variables (Buchmueller and Zuvekas: 1998, and Peters: 2004).

[30] Please see details in Chapter 3, Section 3.3.

toward further research on related topics (labor market outcomes such as earnings or wages, and labor supply are affected by abusive/problem drinking, problem drinking affects women's use of time in market versus home activities, and indirect affects of alcohol use) as new and richer data are gathered. With the NESARC data (2001-2002), it is possible to estimate the impact of six categories of alcohol users or drinkers[31] (ex-drinkers, ex-abusers/dependents, new-drinkers, ex-new-drinkers, alcohol abusers and dependents) on labor market variables.

The current research extends the main idea of Mullahy and Sindelar's[32] works in two important directions by incorporating some of the ideas from other previous studies (Chatterji et al: 2007, French and Zarkin: 1995, Mullahy and Sindelar: 1993) in the following aspects.

First, this research adopted the estimation model presented by French and Zarkin (1995). They specified the relationship between dependent variable wage, and independent variables alcohol use, demographic variables (age, gender and race), human capital variables (education, health status, and job tenure) and worksite specific variables. This research focuses on estimating personal income equation instead of wage equation as the data on wage is not available. Also, the probability of fulltime work force participation equation is estimated on the same set of independent variables. This research does not include job tenure as an explanatory variable (data are not available).

---

[31] See Chapter 4, section 4.1 for details and Tables 2A and 2B

[32] Mullahy and Sindelar (1993, 1994) supported the conventional view that hypothesized the negative relationship between problem drinking (alcohol abuse or dependence) and labor market variables.

Second, French and Zarkin (1995) used alcohol use (the total number of drinks consumed in last year) as an independent variable to estimate the impact on wage. Instead of including the amount or frequency of alcohol consumption, I include six indicator variables (represent ex-drinkers, ex-abusers/dependents, new-drinkers, ex-new-drinkers, alcohol abusers and dependents) as the key variables of interest.

Third, this research includes indicator variables for different categories of industries. It is indicated from some prior studies that some industries (also some occupations[33]) might have higher incidence of alcohol related problems than others because of the nature of occupation or industry type, availability of alcohol, social pressure to drink, stressful work, and lack of supervision. For example, the British census (1995) indicated that some industries (alcoholic drinks industries, hotel and catering industries, the shipping industry) and occupations (the military, doctors, lawyers, and journalists) had higher than average alcohol consumption related problems.

Fourth, this research addresses the 'endogeneity' (most often arise due to 'simultaneity' or 'unobserved heterogeneity') issue along with the heteroscedasticity problem (see detail in Chapter 3, section 3.2).

1.3 Underlying questions and research objectives

1.3.1 Underlying questions

This research was intended to find out the answers to some key questions based on the current data set.

[33] Initially I included different occupational categories. Finally I had to drop this variable because estimations with current data set did not converge at all.

i) Do individuals who had no alcohol use disorders (lifetime abstainers, ex-drinkers, ex-abuser/dependents, new-drinkers, and ex-new-drinkers)[34] work more and earn more than the individuals who had alcohol use disorders (alcohol abuser and alcohol dependent)?

ii) Do individuals' performances in labor market (work participation and earnings) vary due to the fact that they quit drinking (ex-drinkers, ex-abuser/dependents)[35]?

iii) How much variation in labor market variables (personal income and full-time labor force participation) do arise due to different socio-demographic variables such as age, gender, race, education, location, health and work related variable (industry type)?

1.3.2 Hypotheses.

The core hypothesis of this research is that individuals meeting criteria for a diagnosis of alcohol use disorders (alcohol abuse or dependence) will have lower probabilities of full time employment and earn lower incomes from the full time job than the individuals who did not have alcohol use disorders (abstainers, ex-drinkers, ex-abusers/dependents, new-drinkers, and ex-new-drinkers). The following are the detailed hypotheses.

---

[34] During the year 2000-01, ex-drinkers and ex-abuser/dependents did not drink, new-drinkers started to drink, and ex-new-drinkers continued to drink (see details in Chapter 4, Tables 2A and 2B).

[35] Please see footnote 9.

1. Ex-drinkers and ex-abusers/dependents are expected to have lower probabilities of being employed (full time), and earn less income than abstainers (did not drink in life time), holding other factors constant.

2. Among the drinkers (who were not alcohol abusers or alcohol dependents), ex-new-drinkers (individuals who were drinking for longer periods, i.e., drank both in current (2000-01) and previous years) are expected to have the higher probabilities of being employed (full time), and earn higher income than new-drinkers (who started drinking last year), and vice versa when compares with ex-drinkers holding other factors constant.

3. New drinkers (not abusers or dependents) are expected to have lower probabilities of working, and earning lower income than the ex-drinkers (who did not drink last year) and ex-new-drinkers.

4. Among the problem drinkers (such as alcohol abusers and alcohol dependents), alcohol dependents are expected to have the lower probabilities of being employed full time and earning than alcohol abusers.

1.3.3 Research Objectives

Five research objectives are intended to achieve in order to answer the underlying questions and tests the hypotheses throughout this research.

1. To review previous economic literature, empirical work and findings on the relationship between alcohol use and labor market outcomes to build up the model and to choose estimation method for this research.

2. To estimate the effect of alcohol use on two labor market outcomes (labor force participation and earning or income) including six indicator variables[36] for alcohol use status as key independent variables. Various socioeconomic and demographic variables (see Table 4) are also included as other exogenous variables in each of the labor outcome equations, such as the individual's age, gender, race, marital status, education level, health status, location (whether individual lived in a central city or not), other source of income, and work related characteristic (industry type).
3. To perform formal tests to detect the existence of potential problems in estimations, the endogeneity of alcohol use and heteroscedasticity in the sample.
4. To re-estimate the labor market outcome equations by accounting endogeneity and heteroscedasticity using GMM-IV method as tests results detect the presence of endogeneity and heteroscedasticity in the model.
5. To provide comparative analysis of the estimation results (before and after addressing the endogeneity and heteroscedasticity issues).

1.4 Organization of the paper

The plan for the rest of this research proceeds as follows: Chapter 2 presents an analysis of theoretical background and a review of previous empirical works on the relationship between alcohol use or consumption and labor market variables. Chapter 3

---

[36] Each indicator variable takes the value of one if the individual was identified (based on previous 12 months' alcohol use record) as ex-drinkers, ex-abusers/dependents, new-drinkers, ex-new-drinkers, alcohol abusers and alcohol dependents, and zero otherwise.

presents empirical model, estimation methodology and econometric tests. Chapter 4 provides a discussion on data and variable description. Chapter 5 presents an analysis of empirical results. Finally, Chapter 6 summarizes the findings, provides limitations of the study, implications and offers suggestions for additional research.

CHAPTER 2

THEORETICAL BACKGROUND AND PREVIOUS EMPIRICAL WORKS

2.1 Theoretical background

2.1.1 Alcohol consumption and health status

Most of the previous empirical work on the relationship between alcohol use and different labor market outcomes were based on theoretical model formulated by Becker (1964) known as human capital theory. The human capital theory considers investment in human capital such as education, training, and healthy life style as positive investments, and some activities such as unhealthy eating, lifestyle, smoking, and drinking as negative investments. Positive investments are rewarded by better health status, higher work ability, improved work skill that lead to higher occupational attainment, career progression, productivity and earnings. Negative investments are penalized by poor health status and lower work ability or skill that lead to poor job progression, lower productivity and earnings.

Later, the importance of health status as a predictor of wages was recognized by Grossman (1972) as he first applied the human capital theory to the area of health. Since then, in labor market related analysis, individual health status is considered as an important variable. Original work on the relationship between any type of substance abuse and labor market outcomes was done by Culyer (1973). His thought was that since physical and psychological well being of individuals are negatively affected by any substance/drug use; there can be some obvious and frequent spill over effects on labor market. As a result, alcohol or drug users may end up with lower aggregate level of human capital accumulation and thus lower productivity (Kaestner: 1994). This does not only stay confined to individual's disrupting life (where he/she can not perform expected normal role in family, work and social life) also threatens the future labor force by harming the stable family environment where a child grows up. For example, alcoholic parents usually do not provide appropriate time and care in rearing children, and particularly in cases of alcoholic mothers, the evidence of children's mental and physical disorder are well documented (Chatterji and Markowitz: 2001).

The concept became more popular in the last two decades when Mincer (1974) developed the model based on the relationship between wages and various measures of human capital (include education, marital status, job experience, and other socio-demographic characteristics) that affect labor market. Any type of substance use including alcohol consumption can have an adverse impact on individual and society through its impact on health (mental and physical well being) and criminal activity (MacDonald and Pudney: 2000). Alcohol, drug or any other type of substance users may have lower productivity levels, and so may have lower wage or lower income and lower living standards as economic theory tells that workers receive payment for their marginal product (Folland et al: 2001). Zabel (1993) developed the most general model for labor supply, where the work decision depends on wages, desired hours, and the decision to participate.

2.1.2 Alcohol consumption, work and income

In the analysis of the relationship of labor market variables and alcohol use or any other substance use problem, most empirical studies hypothesized that labor market variables are a function of personal, demographic, economic, human capital, and substance use attributes of an individual (Bryant et al.: 1996).

Several types of theoretical explanations of the causal relationship between alcohol use and labor market outcomes are found.

First, some empirical work relies on the neoclassical explanation of individual's labor supply behavior. It states that rational individual's decision to work or not to work or hours of work depends on the choice between consuming leisure (time not spent at work) and consumption (also includes alcohol) that purchased with income earned at work. Economists consider leisure and alcohol consumption as normal goods. [37] If wage increases (so does income increase), consumer works more and earns more, and spends more on consumption (includes alcohol), vice versa.

Second, some researchers argued that individual's choice to consume alcohol is completely determined by his/her own personality and own preference to time allocation between work and leisure. Becker and Murphy (1988) explained that individuals with a high rate of time preference tend to base current consumption decisions (includes pleasurable alcohol consumption) without considering future adverse health effect, and select jobs with a current high wage but tend not to invest in human capital (i.e., flatter age-earnings). Some studies argued that this type of individual's rate of time preference is the reason of arising unobserved heterogeneity in the estimation process.

Third, the negative relation between alcohol use and work/income can be explained by using the formal setting as considered by Bradley and others (2002). In this explanation, supply of work hours 'H' (equals total time endowment 'T' minus

[37] Normal goods are something (i.e., consumption good) that consumers desire more when income or purchasing power increases.

demand for leisure hour ‘L’) is determined by the utility maximizing assumption, tastes, prices, and endowments of wealth. Combining these assumptions and the basic notion of Grossman (1972) theory, the detrimental health effect of alcohol misuse or excessive drinking can clearly be explained by diminishing tastes for work, raising marginal values of leisure time, and devoting more time to alcohol consuming activities and less to health maintenance activities. All of these jeopardize the individual’s promotion opportunities, occupational advancement, higher productivity and wage/income through reduced work time, frequent sickness leave, absence from the work, and reduced work efficiency. On the other hand, it was also argued that nature of work and certain type of occupation may be responsible for inducing heavy/excessive drinking.

Fourth, the evidence of parabolic relationship between alcohol use and earning such as ‘non-drinkers and heavy drinkers both earn less than moderate drinkers’ is often explained by medical literature supported view. Some researches provided the evidence of the primary health benefit of moderate alcohol consumption that leads to reduced stress and tension levels and lower the incidence of disease such as CHD-coronary heart disease (Dave and Kaestner 2002, Mirowsky et al: 2000, Heien: 1996, Hutcheson et al: 1995). Mirowsky et. al (2000) provided evidence of U-shaped relationship between drinking and illness, and also many others (Shaper: 1990, Doll et al: 1994, Gronbaek et al: 1995 , Fuchs et al: 1995 and Thun et al: 1997) provided the evidence of U-shaped relationship between alcohol consumption and mortality.

Fifth, besides supportive medical benefit from alcohol use, another type of benefit generates from alcohol consumption which is known as ‘networking effect’ (Hutcheson et al.: 1995, MacDonald and Shields: 2001). It usually happens when colleagues and business associates get together, and spend social time and sometimes exchange useful information about work and institutes and businesses. Even senior members obtain important information about their junior colleagues such as job commitment and future job skill improvement plan which eventually help co-workers to achieve future promotion opportunities.

2.2 Previous Empirical Works

Numerous studies analyzed the impact of alcohol use disorders on various aspects of labor market such as labor force participation, hours of work, wage or income and productivity, education, occupation, and job mobility. In order to achieve one of the objectives, this research presents major findings from some major studies first, and then focuses on several important conceptual and computational issues that explain the controversial results.

2.2.1 Major findings

The existing mixed results on the relationship between alcohol consumption and labor market can be divided into four groups based on the directions of impact such as ‘no’ effect, negative, positive, and nonlinear (inverted-U shaped and inverted-J shaped).

2.2.1A. In the first group of studies, only few empirical works were found that reported ‘no effect’. Benham and Benham (1982) utilized the Diagnostic and Statistical Manual (DSMALC) criterion (as proposed by APA). It emphasized on alcohol induced individual’s social behavior instead of amount of alcohol use to define ‘alcoholism’ and reported that alcoholism did not have any significant effect either on income or employment although Mullahy and Sindelar (1989) used the same DSMALC criterion and found that alcoholics earned less than the non-alcoholics. Bryant and others (1992) explored the impact of number of drinks (self-reported) on individual hourly wages using National Longitudinal Survey of Youth (US, early 80s) and reported the evidence of no effect.

2.2.1B. In the second group of studies, the earliest evidence on drinking and productivity was found from the work of prominent economist Irving Fisher in 1927 (Peters: 2004). Fisher reported that drinking made workers less productive, and extrapolated that prohibition would increase the nation's productivity by five percent. The recurrence of similar results on the relationship between alcohol use and labor market outcomes were evidenced in some major empirical works. The works of Mullahy and Sindelar (1989, 1991, 1993, 1994, 1995, and 1996) took the leading role in providing most comprehensive set of alcohol studies with negative evidence. The similar type of evidence was reported in many other major empirical works (Bryant et al: 1992, 1993, 1996, Cook and Moore: 1993, Kenkel and Ribar: 1994, Yamada et al: 1996 and Chatterji: 1998).

In 1989, Mullahy and Sindelar found a relatively larger impact of alcoholism on labor market supply than on changes in workers wages. They used data from the New Haven Site of the Epidemiological Catchment Area (ECA) survey of individuals 18 years and older in the study and investigated the life-cycle effects of alcoholism on human capital formation and labor market outcomes. The reported results were sensitive to different measures of income and different age populations. They found a statistically significant negative impact of alcohol use on education. They also explained that young students with alcohol problem quit from school early and entered into the work force, and earned more than their non-alcoholic peers. Although job experience- accumulated, the lower level of education ultimately limited their wage growth, occupational choice, and earnings at older age.

In 1991, Mullahy and Sindelar investigated gender differences in the effects of alcoholism (alcohol dependence and abuse, according to DSM-III) on personal income of employment and household income by using multiple-site data from the ECA survey (US, 1980-1981). They found consistency with the results of the earlier work (1989) and reported stronger negative effects of alcoholism on labor market participation and household income for women than for men. The negative effect on labor force participation and income, varied across the life cycle, by gender and by labor market variables such as labor force participation, employed/unemployed, and income. They found that the young male alcoholics earned less than their non-alcoholics peers, but the same results were not found for female group. Surprisingly, they found that female alcoholics' incomes increased with age whereas female non-alcoholics' income declined with age. They urged researchers to consider the importance of the differential impact of 'direct' and 'indirect' effects of alcohol dependency, because alcohol dependency not only directly affects labor market variables, but it also affects labor market variables indirectly through educational attainment and marital stability. They argued that this type of differentiation is more important for females. They reported that direct effect of alcoholism on earnings was not significant but indirect effect of alcoholism on variables through educational attainment and marital stability was significant.

In 1993 study, Mullahy and Sindelar investigated the 'alcohol-income' relation by looking at the impact of alcoholism (a binary variable which was obtained by utilizing the DSMALC criterion) on the probability of full time work, productivity and income. They compared the estimated results obtained for individuals with different age groups such as ‘20-29’ ‘30-60’ and ‘60-64’ years old, and reported that in both younger (20-29 years old) and older group (30-60 years old) alcohol dependent individuals had a higher probability of full-time work than non-alcohol dependent individuals whereas in middle age group (30-60 years old) alcohol dependent individuals had a lower probability of full-time work than non-alcohol dependent individuals. They provided the logical explanation that early age school drop out rate might be higher among alcohol dependent individuals and they entered in work force and earned income at younger ages as opposed to the fact that non-alcohol dependent individuals invested more time to finish education and acquire more human capital, and then entered the work field at later age (that is between 30-60). As a result, non-alcohol dependent individuals had a higher probability of full-time work and earnings later in their life, and also accumulated more wealth. Non-alcohol dependent individuals took early retirement based on their accumulated wealth and accordingly they had a lower probability of full-time work at older age, whereas the older alcoholics could not take early retirement and worked more at older age as they saved little for retirement.

In a 1994 study, Mullahy and Sindelar developed a model to investigate the separate impact of direct and indirect effects of alcoholism on productivity and income as they argued that most of the previous works might have underestimated the effects of alcohol abuse and dependence since they only considered the impact of direct effects (such as, the effect of alcoholism on earnings when controlling for factors possibly affected by alcoholism, such as education and marital status).

Mullahy and Sindelar (1995), using the sample of 15,000 men and women drawn from the 1988 Alcohol Survey of the National Health Interview Survey, found greater negative impact of problem drinking on employment. The impact varied across gender and estimation method (Ordinary Least Squares (OLS) approach and the Instrumental Variable (IV) estimation approach). The IV approach produced more negative impact for men than the OLS approach, and for women the IV estimate changed the sign from a positive to a negative. In the IV approach, they used different variables as instruments for alcohol problems such as state-level excise taxes on beer, state-level excise taxes on cigarettes (a complementary good), state-level apparent ethanol consumption (a measure of per-capita state-level sales), the quadratics of the two taxes and apparent ethanol, and three variables describing history of living with alcoholic relatives.

In 1996, Mullahy and Sindelar investigated the impact of alcohol dependence and abuse (according to DSM-III) on number of drinks (self-reported), employment, and unemployment based on data from Alcohol Supplement of the National Health Interviews (US, 1988). They used multinomial IV estimation approach and reported that problem drinking leads to reduced employment and increased unemployment.

Bryant and others (1992, 1993, and 1996) hypothesized the negative consequences of drinking on wage or earnings, and used simultaneous equations systems (similar estimation method used by Berger and Leigh: 1988) estimation method. In their 1992 study, they used control for an income effect on alcohol use, and found that wage premiums disappeared because of alcohol use. In 1993 study, they investigated the importance of drinking patterns over time and found that wage premiums disappeared when an individual’s drinking history was considered, and heavy drinking over an extended period imposed wage penalties. In 1996 study, they found that heavy and prolonged drinking influenced wage negatively for young men.

Cook and Moore (1993) investigated the impact of drinking on the number of years of post-secondary school completed by 1988 based on the NLSY survey (1982) sample of youths who were in high-school. They found that reduction of education year was associated with increased number of drinks (with frequent drinking on more than one occasion per week, and frequent drinking at least four occasions consuming six or more drinks in the last month).

Kenkel and Ribar (1994) used different estimation methods (OLS method, panel data approaches, IV approaches) and data from National Longitudinal Survey of Youth (US, 80s) to estimate impact of alcohol dependence and abuse (according to DSM-III criteria) on number of drinks (self-reported), income, and hours worked. They found the evidence of lower wage for alcohol dependent individuals but no evidence of lower labor supply.

Yamada, Kendix and Yamada (1996) estimated the causal relationship between drinking and schooling (high school graduation) using the National Longitudinal Survey of Youth (NLSY) sample of 1981-82 on twelve grade students, and reported that contemporaneous drinking was associated with reduction in the probability of high-school graduation.

Chatterji (1998) found the evidence of sensitiveness of the relation of drinking and schooling to alternative assumptions about the correlation of unobserved determinants. The test results indicated that alcohol use is not exogenous. Although their estimated Ordinary Least Squares (OLS) results using NLSY data indicated that weekly alcohol use reduced the number of grades completed by the age of twenty-one when exogeneity of alcohol use assumed, but weekly use of alcohol before the age of seventeen did not have significant effect on the years of education attained when instruments for drinking was considered (such as state excise tax on beer, percent of counties in the state that prohibit alcohol sales).

2.2.1C. In the third group of studies with positive association of alcohol consumption and labor market outcomes, the first work done by Berger and Leigh (1988), and later reporting of same type of results continued by many others (Cook:1991, Heien:1996, Hamilton and Hamilton:1997, MacDonald and Shields:2004, and Auld:2005).

Berger and Leigh (1988) investigated the impact of number of drinks (self-reported) on American workers' hourly wages by using the survey data from the Quality of Employment Survey (1972-73). They estimated separate equations for drinkers and non-drinkers by OLS method and reported that drinkers received higher wages than nondrinkers. Most surprising fact of their results was that the difference between drinkers and non-drinkers wages became even greater as the frequency of drinking went up. Although their research method popularly followed by many others (Pittman et al: 1989, Bryant et al: 1992, Heien: 1996) their works were highly criticized because they did not separate the alcohol and non-alcohol related factors (Bryant et al: 1992), and ignored self-selection problems (Gill and Michaels: 1992).

Cook (1991) studied the impact of drinking (measured by the level of drinking per month) on earning based on the Quality of Employment Survey (QES) and reported that drinkers earned more than the abstainers.

Heien (1996) investigated the impact of alcohol consumption at different levels (abstainers: lifetime abstainers and ex-drinkers, moderate drinkers and abusive drinkers) on wage or earnings based on National Household Survey data on alcohol use (US, 1979 and 1984), number of drinks (self-reported) and household income. His findings by applying non-linear 3SLS estimation method confirmed his hypothesis that moderate drinkers earn more than either abstainers or abusive drinkers, but was not statistically significant to the extent of other human capital variables (such as education or age). He adopted the theoretical model of Grossman (1972) with little modification as he took a quadratic functional form.

Hamilton and Hamilton (1997) reported OLS wage regressions results with selectivity correction that moderate drinkers had the highest wages based on General Social Survey (Canada, 1995) data on number of drinks (self-reported), and annual pre-tax Income.

In 2004, McDonald and Shields used the data from Health Survey for England (1997-1998) on number of drinks (self-reported), alcohol dependence and abuse, and employment, and reported the results of Instrumental Variables approach estimation that problem drinking (alcohol dependence and abuse) lead to reduced employment.

Auld (2005) used the Canadian General Social Survey (1985 and 1991) data and studied the ‘alcohol/income puzzle’ and reported maximum simulated likelihood estimates that moderate drinkers earned 10 percent higher income, and heavy drinkers earned 12 percent higher income than abstainers.

2.2.1D A fourth group of studies indicated an inverted-U shaped or inverted-J shaped association between alcohol use and wage or income. The inverted-U shaped or inverted J-shaped association between alcohol use and wage or income is the inverted application of U-shaped or J-shaped relation between drinking intensity and health benefits as found in medically supported literature. For example, Baum and Baicker (1985) first predicted inverted J-shaped and Shaper (1988) first predicted inverted-U shaped relationship after reviewing epidemiological evidence that moderate alcohol consumption could be beneficial to health. Later, major works were done by French and Zarkin (1994 and 1995). The works of Zarkin et al. (1998) and McDonald and Shields (2001) also supported the same view.

French & Zarkin (1995) studied a group of 1,000 workers aged between 30 to 59 at four work sites (US, 1991-1993) around USA and gathered data on number of drinks (self-reported), weekly wages, drinking behavior, and personal characteristics and reported OLS regression results that indicated an inverse U-shaped relationship between alcohol consumption and wages. They found that moderate drinkers earned the highest wages with drinking between 1.7 and 2.4 drinks a day on average, although the evidence of quick drop of wages was found for higher levels of drinking. The inverted-U reached a peak approximately at 1.5 to 2.5 drinks per day on average. This corresponds to a wage premium of around five percent over nondrinkers. However, they did not find any statistically significant results for women in the case of alcohol use premium and nonlinear relationship.

Zarkin and others (1998) replicated French and Zarkin’s (1995) findings by selecting prime age workers at six work-sites from the National Household Surveys on Drug Abuse-NHSDA (1991-1992), a nationally representative cross-sectional surveys of the U.S. that contain detailed information on alcohol and labor market information) and took data on various measures of current and lifetime substance use, number of drinks (self-reported), and hourly wages. Their findings did not confirm uniformly any evidence of an inverse-U-shaped relationship between wages and alcohol consumption for either men or women, and they did not report any evidence of a turning point at a particular alcohol consumption level. They found some evidence of an inverse-U-shaped relationship between alcohol use and wages at low drinking levels for men, but could not reject the hypothesis of a constant 7 percent wage premium over a wide range of alcohol use. For women, they did not find any evidence of a nonlinear relationship between alcohol consumption and wages but indicated that alcohol user women earned approximately 4 percent wage premium than women who did not drink.

McDonald and Shields (2001) estimated the relationship between the number of drinks (self-reported) and hourly wages by using Instrumental Variables approach based on data from Health Survey for England (1992-1996). The results indicated that moderate drinkers earned highest wages, and showed the evidence of an inverted U-shape for the relationship between drinking intensity and mean hourly wages for both males and females. The peak of the turning points occurred in the range of 21-36 units drinking for men and 14-28 units drinking for women.

2.2.2 Conceptual and computational issues

Review of related literature and empirical work on the interrelationship between alcohol use and labor market outcome reveals some noticeable facts and issues. There are various reasons identified for heterogeneous results that mostly accounted for measurement and conceptual issues used in the estimation process. These are: different self selected criteria to measure alcohol use, various causal path and directions (positive, negative and nonlinear (inverted-U shaped and inverted-J shaped) through which alcohol consumption influences labor market variables, different estimation methods to address endogeneity (arise due to simultaneity bias or unobserved heterogeneity) and different labor market variables.

2.2.2A Empirical results in different studies varied substantially due to the choice of a methodological approach to define alcohol use disorders. Some studies used simply binary representation of alcohol use such as whether individuals are alcoholic or not, and some studies used continuous variable such as frequency and amount of alcohol use measurement, for example number of drinks used last week, month or year and medically oriented measurement of alcohol use disorders such as alcohol dependence or alcohol abuse.

Two types of methodologies were used to measure alcohol use disorders. In the first methodology (followed by American Psychiatric Association -APA, 1994), alcohol use measure was defined based on the effect of alcohol use on the failure to fulfill major obligations at work, school or home (American Psychiatric Association, 1994). In second methodology (followed by NIAAA), the link between the alcohol problem, the amount and type of alcohol consumption is considered.

The magnitude of impact on hours of work, productivity and wage/income differs by the extent of alcohol use (types and magnitude) and even behavioral reaction to alcohol use. For example, Berger and Leigh (1988) used self-reported data and binary variable approach on alcoholism and reported that unadjusted mean wages of male and female drinkers were higher by 12.8 percent and 25.2 percent respectively than the non-drinkers male and female. Heien (1996) used NIAAA linked definition proposed by

American Psychiatric Association (APA) and found moderate drinkers earned more than abusive drinkers and abstainers. Mullahy and Sindelar (1989, 1991, 1993, 1994, and 1998) considered alcohol induced individual's social behavior instead of amount of drink as emphasized in DSMALC criterion (proposed by APA) for the most of their studies and observed alcoholics earned less than the non-alcoholics. Finally, Bray et al. (2000) used National Household Survey on Drug Abuse (NHSDA) data that were based on less strict criterion than DSMALC criterion.

The study based on UK (MacDonald & Shields: 2004) and US data (Kenkel & Ribar: 1994, Mullahy and Sindelar: 1991, 1993 and 1996) on alcohol dependence indicated penalties in the form of lower wages, higher unemployment, and lower rates of labor market participation. Also, different levels of drinking measurement used in different studies influenced the result. For example, the level/stages of drinking is defined as 'heavy drinking' for consumption of at least 100 alcoholic beverages per month, 'light drinking' for 1–10 or 1–20 drinks per month and moderate drinking for the consumption of 21–59 alcohol beverages per month (Ruhm and Black: 2002). French and Zarkin(1995) using Substance Abuse and Mental Health Services Administration-SAMHSA 1992 and 1993 defined heavy use as drinking 5 or more drinks per occasion on 5 or more days in the past 30 days. The 60 drink cut-off is frequently used to define 'chronic' drinking (Dee: 2001) which is in between moderate and heavy drinking. Two drinks per day on an average (1.5 to 2.5) were defined as moderate drinking in the study of French and Zarkin (1995). Heien (1996) mentioned that following the NIAAA link, some researchers used self selected approach and defined alcohol abuse as consumption of two of more drinks per day. In that case there is a possibility that moderate alcohol use measurement might be used for estimating heavy alcohol effect of labor market variables in some analysis.

2.2.2B The causal path and directions (positive, negative and nonlinear (inverted-U shaped and inverted-J shaped) through which alcohol consumption influenced labor market outcomes were explained by several mechanisms.

The negative causal direction between alcohol use and work/income based on the belief that alcohol abuse is harmful for health, and lowers human capital accumulation, and eventually reduces productivity through job absence, lesser amounts of work, unemployment, and lower work performance (Rice et al.: 1990, Kenkel and Ribar: 1994, Mullahy and Sindelar: 1993 and 1996).

The positive causal direction between alcohol use and work/income was explained in two ways. First, some studies reported positive wage effects of moderate alcohol use (Berger and Leigh: 1988, French et al: 1995, and Hamilton: 1997) based on health benefits of moderate drinking such as reduction of risk of cardiovascular disease. Better health leads to greater productivity, and therefore higher earnings. Second, in addition to the medical relationship, another informal mechanism referred as ‘networking role of alcohol’ leads to the positive association between alcohol consumption and labor market outcomes. Social gathering or time spent for drinking with co-workers’ may give opportunity to share important job information, co-workers commitment and motivation to the job (MacDonald and Shields: 2001) that eventually help to achieve job progression. Although some studies reported that job stress and nature of job might be responsible for increasing intensity of drinking (Sokejima et al: 1998, Fenwick et al: 1994, Karasek et al: 1990, and Baker: 1985).

Evidence of a quadratic or nonlinear (inverted-U shaped and inverted-J shaped) relationship between drinking intensity and wages or alcohol consumption and wages was found in some major works (Baum and Baicker: 1985, French and Zarkin: 1995, Heien: 1996, Hamilton and Hamilton :1997, and MacDonald and Shields: 2000) using different sources of data. Explanation of these mechanisms grounded on medical literature identifies a U-shaped association between alcohol consumption and the risk of cardiovascular heart disease (Shaper: 1988, Zarkin et al.: 1998, and Sesso: 2001). Zarkin and others (1998) considered the findings of some recent studies based on epidemiological data sets (Marmot and Bruner: 1991; Beaglehole and Jackson: 1992, Shaper: 1990, and Coate: 1993) as consistent and supported the evidence that abstainers and heavy drinkers have higher rates of cardiovascular disease than do light or moderate drinkers. Finally, they argued that since the affect of alcohol use is related to wages through human capital variable, and moderate alcohol users (approximately two drinks per day on average) have a lower risk of coronary heart disease and better health conditions relative to abstainers and heavy drinkers, moderate alcohol users get higher wage than abstainers and heavy drinkers. Alcohol consumption at moderate levels is beneficial for health by relieving stress and reducing the incidence of cardiovascular heart disease. Sesso (2001) argued that alcohol reduces the risk of cardiovascular disease through increases in plasma high density lipoprotein-cholesterol (HDL) levels.

2.2.2C Since the theory of labor supply literature indicates that demographic variables can determine the marginal rate of substitution between hours worked and alcohol/substance consumption (Zarkin et al.: 1998)**,** some researchers tried to explain the causality between alcohol use and labor market variables by considering interactions among demographic variables such as age, race, marital status, number of children living in the household, location (urban/rural), and health status although there always exists some uncertainty (Fuchs: 2004).[38] Some researchers explained non-consensus results by the existence of uncertainty. The evidence of uncertainty was observed in the form of estimation problem such as endogeneity or simultaneity [39]. This type of estimation problem occurred when someone estimated single equation model of wage or income or hours of equation including alcohol use as an explanatory variable without considering the fact that alcohol use and wage or income are jointly determined. As a result, the reported results might be biased. For example, in substance use cases, Kaestner (1991) mentioned that a negative relationship between substance abuse and wage can rarely be observed if someone accounts for endogeneity. However, three reasons explain the causes of occurring endogeneity or simultaneity problem (Ettner: 1996, Ettner et al.: 1997, Marcotte et al.: 2000, Zarkin et al.: 1998, Dee: 2001).

[38] Fuchs (2000): There are numerous reasons for this uncertainty. Many of the socio-economic variables are correlated with each other; sometimes it is difficult to estimate the independent relationship of each one. ……….Interactions among the variables are numerous and varied, as are nonlinearities. One of the biggest problems is establishing causality………. Even when the positive correlation is strong and stable, the interpretations can include causality running from income to health, from health to income, and/or "third variables" that effect health and income in the same direction.

[39] Generally when two variables in a structural simultaneous model (based on theory) are jointly determined (for example, price and quantity) both are considered as endogenous variables (Ramanathan: 1998). Therefore, if someone does not consider the joint relationship between alcohol consumption and income or wage, or alcohol consumption and educational attainment, and estimates single-equation of wage or income on an explanatory variable alcohol use without accounting simultaneous causation (such as between alcohol use and wages or alcohol consumption and educational attainment) then endogeneity problem or simultaneity bias may arise. As a result, the estimation result will be biased.

First, endogeneity occurs because the causality may be bi-directional between labor market outcomes and alcohol use since both employment status and alcohol consumption are jointly determined by unobserved characteristics of individuals (such as individuals time preference for work and leisure, and personality) that influence both employment status and alcohol consumption.

Second, excessive drinking adversely affects health that can harm job progression or occupational advancement in one hand, and on the other hand, heavier drinking can also be induced by work environment such as trouble at work, stressful job, and frustration over not getting promotion. Therefore, it may not be always true that alcohol use affects labor market variables or work place issues may induce drinking.

Third, the affect of unobserved factors[40] such as attitude, motivation, and time preference influence alcohol consumption decision as well as work and wage. For example, individuals with high time preference give priority to current consumption of alcohol without considering its future adverse health consequences and select jobs with a current high wage rather than invest in human capital that choose a flatter age-earnings profile (Becker and Murphy: 1988).

Most of the research addressed the endogeneity issue by using instrumental variable estimation procedure (Hamilton and Hamilton: 1997, Mullahy and Sindelar: 1996, Heien: 1996, Barrett: 2002, McDonald and Shields 2001). In this case,

---

[40] Unobservable heterogeneity may reflect underlying causal factors such as correlations in the direct marginal utilities of health, income, and lifestyle choices which may in turn be related to differences in genetic characteristics, childhood circumstances, and attitudes to risk and the rate of time preference. It may also reflect correlations in the marginal products of lifestyles with respect to health.

instruments selected were uncorrelated with labor market variables but correlated with alcohol use. Taxes or prices were used as instruments in most of these researches. However, different studies used different types of variables as instruments. For example, Mullahy and Sindelar (1996) used alcohol habits or alcohol dependency of the parents as instruments. McDonald and Shields (2001) used long-term, non-acute illnesses, such as asthma or diabetes as instruments since these illnesses inhibit drinking to some extent but do not affect (if not serious enough) labor market outcomes. Barrett (2002) considered alcohol taxes which affect alcohol consumption but not labor market success. Heien (1996) and Hamilton & Hamilton (1997) both used religiosity as instruments considering the fact that religious individuals drink less.

Some researchers had different opinions to address the potential endogeneity problem of alcohol use related variables. First, the estimations using current (rather than past) drug use and predetermined variables (such as family and education variables) with respect to the current period (Burgess et al.: 1998, Kenkel and Ribar et al.: 1994) are able to address the endogeneity issue. Second, if substance or alcohol consumption is considered as a normal consumption good and the level of consumption is determined in response to market wages and non-labor income by the utility maximizing consumer then endogeneity should not be a problem.

2.2.2D Different conclusions that were drawn from these studies might be related to the different labor market variables that were used in the analysis (MacDonald and Pudney: 2000). For example, Kenkel and Ribar (1994) focused on the hours of labor supplied, whereas Mullahy and Sindelar (1989, 1991, 1993, and 1996) considered participation in labor market, wage, income and hours of work in their different studies.

Dave and Kaestner (2002) used reduced form approach (they estimated the relationship between alcohol taxes and labor market outcomes) instead of structural approach (structural relationship between alcohol consumption and labor market out comes as theory described) in order to take care of the endogeneity issue of estimation. In reduced form, Dave and Kaestner (2002) assumed that alcohol taxes are positively related to labor market variables. They argued that most of the literatures suggested that alcohol taxes tended to be negatively related to employment and hours of work, and positively related to wages. They mentioned that previous studies had shown that alcohol taxes were negatively related to alcohol consumption and negatively related to labor supply and wages and thus alcohol taxes were positively related to both labor market outcomes (except Zarkin and others (1998) hypothesized that alcohol taxes had similarly signed effects on labor supply and wages). Dave and Kaestner (2002) did not find the presence of such relation in empirical results (weak and indeterminate association between alcohol taxes and labor market outcomes), and therefore, concluded that either or both sets of previous estimates were misleading and suggested that there existed weak and indeterminate relationship between alcohol taxes and labor market outcomes, and so did between alcohol consumption and labor market outcomes.

CHAPTER 3

EMPIRICAL MODEL, METHODOLOGY AND ECONOMETRIC TESTS

3.1 Empirical Model

Considering the main focus of this research, studying the impact of alcohol use disorders on the labor market outcomes, I choose to estimate the following two labor market outcome (see Table 1 for descriptions) equations. The specification of the equations is guided by the theoretical framework of human capital as described in literature (Becker: 1965, Grossman: 1972, 1991, and Mincer: 1974) and the empirical model specifications of prior studies (Chatterji et al: 2007, French and Zarkin: 1995, Mullahy and Sindelar: 1993).

(3.1.1) $\text{Prob}(\text{LFPFULL}_i) = \alpha^w + X_i \beta^w + \text{ALC}_i \delta^w + \xi_i^w$

(3.1.2) $\text{LN}(\text{INCOME}_i) = \alpha^y + X_i \beta^y + \text{ALC}_i \delta^y + \xi_i^y \quad \text{if } \text{LFPFULL}_i = 1$

where LFPFULL_i is a vector of observations on full time work participation ($\text{LFPFULL}_i = 1$ if $\text{LFPFULL}_i > 0$, otherwise 0) for the *i*th individual. INCOME_i is a vector of observations on annual income for the *i*th individual who worked full time during the preceding year (2001-02). X_i is a vector of personal, socioeconomic and demographic characteristics. ALC_i is a vector of indicator variables for alcohol use status, and ξ_i is the error term. α, β and δ are the unknown parameters to be estimated. The superscripts '*w*' and '*y*' refer to work and earnings. 'Prob' indicates probability and 'LN' indicates natural logarithm. Equation (3.1.1) is a logistic model of full time work participation and equation (3.1.2) is a linear income model, conditional on full time work.

To estimate the effect of alcohol use on two labor market outcomes, this research considers six indicator variables for alcohol use status (ALC_i)[41] takes the value of 'one' if the *i*th individual was identified (based on previous 12 months alcohol use record) as ex-drinkers, ex-abusers/dependents, new-drinkers, ex-new-drinkers only, alcohol abusers and alcohol dependents, and 'zero' otherwise. Various socioeconomic and demographic variables (see Table 4) are also included as other exogenous variables (X) in each of the labor outcome equations, such as the individual's age, gender, race, marital status, education level, health status, location (whether individual lived in a central city or not), other source of income, and work related characteristic (industry type).

The key interest of this research is to obtain unbiased and consistent estimates (δ^w and δ^y) of ALC_i in the estimation of labor market outcomes, holding all other socioeconomic and demographic variables constant. This research intends to use standard Maximum Likelihood Method (MLE) for estimating logistic equation (3.1.1) and Ordinary Least Square (OLS) method for estimating log-linear equation (3.1.2).

Prior to estimating the models using standard procedures, it is necessary to check if there is any estimation problem. Reviews of previous empirical works on the relationship between labor market outcomes and alcohol use measures indicated that generally two main empirical problems such as simultaneity and unobserved heterogeneity most often termed as 'endogeneity'[42] complicated standard estimation

---

[41] See Table 2A and 2B for definition and descriptions.

[42] In econometric terms, endogeneity occurs when independent variables included in the model are correlated with error terms (Greene: 2000).

procedure (DeSimone and Wolaver: 2005, Peters: 2004, Farahati : 2003, Dee and Evans: 2003, and MacDonald and Sheilds: 2001, Dee: 2001, Marcotte et al.: 2000, Zarkin et al:1998, Hamilton and Hamilton: 1997, Mullahy and Sindelar: 1996, Kenkel and Ribar: 1994, Ettner et al :1997, Ettner: 1996, and Heien: 1996). For example, there may be simultaneous causation or bi-directional causality between labor market outcomes and alcohol use measure, that is, work status or earnings influence alcohol use decision and alcohol use also influences work status or earnings.

Also, unobserved heterogeneity is the most common problem in cross sectional data. It arises when unmeasured correlates of alcohol use decision sometimes referred as ‘omitted variables’ affect labor market variables. If alcohol use measure is really endogenous, standard estimation method by including alcohol use measure as a right hand side variable without addressing endogeneity would produce biased and inconsistent estimates. Since it violates one of the fundamental assumptions of classical regression model, the independent variable included in the model should be uncorrelated with error terms (Greene: 2000).

3.2 Methodology

If the endogeneity is detected by the diagnostic tests, it is common practice to use instrumental variable (IV) estimation method instead of standard OLS and MLE method, if appropriate (valid) instruments[43] are found.

[43] Wooldridge (2002) defined that a valid instrument is a variable which must be correlated to a potential endogenous variable (implies “instrument relevance”) and uncorrelated with the error terms of the structural equation (implies “exogeneity”).

The IV estimation can be done in the following two stages (Greene: 2000, Chatterji et al: 2007). Considering the labor market outcome equation (3.1.1) where full time work participation (LFPFULL) is a binary variable, the IV estimation is applied in the following two stages.

(3.2.1) Stage 1: Prob (ALC_i) $= \alpha_1^w + X_i \beta_1^w + Z_i \partial_1^w + \xi_{1i}^w$

$ALC_i = 1$ if $ALC_i > 0$, otherwise 0

Stage 2: Prob ($LFPFULL_i$) $= \alpha_2^w + X_i \beta_2^w + ALC\hat{}_i \delta_2^w + \xi_{2i}^w$

$LFPFULL_i = 1$ if $LFPFULL_i > 0$, otherwise 0

where $ALC\hat{}_i$ is the predicted values obtained from the first stage estimation of alcohol use status indicator variables (ex-drinkers, ex-abusers/dependents, new-drinkers, ex-new-drinkers, alcohol abusers and alcohol dependents) including instrumental variables (Z_i) and the same set of exogenous variables (X_i) included in equation (3.1.1).

Considering the labor market outcome equation (3.1.2), where the labor outcome variable is a continuous variable, the IV estimation is applied in the following two stages.

(3.2.2) If $LFPFULL_i = 1$

Stage 1: Prob (ALC_i) $= \alpha_3^y + X_i \beta_3^y + Z_i \partial_3^y + \xi_{3i}^y$

$ALC_i = 1$ if $ALC_i > 0$, otherwise 0

Stage 2: LN ($INCOME_i$) $= \alpha_4^y + X_i \beta_4^y + ALC\hat{}_i \delta_4^y + \xi_{4i}^y$

where $ALC\hat{}_i$ is the predicted values obtained from the first stage estimation of alcohol use status indicator variables using instrumental variables (Z_i) and the same set of exogenous variables (X_i) included in equation (3.1.2).

Several IV estimation methods are available, for example, Two Stage Least Square (2SLS), Generalized Method of Moments (GMM), and Limited Information

Maximum Likelihood (LIML). Selection of any of the IV estimation processes (2SLS or GMM or LIML) depends on the following three situations (Baum et al: 2003).

First, if there are RHS endogenous regressors and there is no heteroscedasticity in the model then 2SLS is the appropriate option (Baum et al: 2003). Amemiya (1985) and Foster (1997) suggested using GMM-IV method instead of 2SLS-IV for estimating a logistic model even with endogeneity problem alone.

Second, if both endogeneity and heteroscedasticity are present in the model (linear or logistic) then the GMM is the appropriate option (Baum et al: 2003). Since the presence of heteroscedasticity of unknown form is the most common problem in cross-sectional data, it should be tested by appropriate method (Baum et al. 2003) such as the Pagan-Hall (1983) test of heteroscedasticity, and if the heteroscedasticity is detected in the model then the GMM should be used, otherwise IV regression results would be consistent but inefficient.

Third, if selected instruments are not strong (weak instruments in the sense that they are poorly correlated with the endogenous regressors) then the LIML method is the appropriate option. Econometric literatures (Murray: 2005, Staiger and Stock: 1997, Bound, Jaeger and Baker: 1995, Hahn and Hausman: 2002) suggested that in the presence of weak instruments, using inappropriate IV method such as 2SLS may produce less reliable (i.e., more biased and inconsistent) results than the OLS and MLE methods do.

In order to choose appropriate IV method (2SLS or GMM or LIML), this research conducts required diagnostic tests for endogeneity, heteroscedasticity and instrument relevancy.

3.3 Econometric tests

3.3.1 Selection of instruments

Before applying any of the IV estimation methods, it is required to test the presence of endogeneity bias. The test for endogeneity requires selection of proper instruments. Generally researchers use standard econometric tests and existing theories to choose proper instrument(s) for endogenous variable. Econometrically, three criteria are considered in choosing instrument variables: exogeneity (no endogeneity), relevancy (indicates IV strength) and validity of instruments (Wooldridge: 2002). Exogeneity requires that selected instruments must be uncorrelated with the error terms of the structural equation. It also ensures that instruments have affects on dependent variables (labor market outcomes) only through RHS endogenous variables (alcohol use status variable). Relevancy requires that selected instruments must be correlated with RHS endogenous variables. Validity requires the conditions of relevancy and exogeneity to be met.

The diagnostic test for endogeneity requires identifying instruments for the potential endogenous variables. In order to choose appropriate instruments, this research relies on the previous empirical works. These works addressed endogeneity issue and implemented instrumental variable estimation methods to study the relation between alcohol use and labor market outcomes. Different variables were used as identifying instruments. For example, parental alcohol problem, lived with alcoholic, beer tax, apparent ethanol consumption per capita (Mullahy and Sindelar: 1996, Terza: 2002), long-term non-acute illnesses such as asthma or diabetes, the number of dependent children and its square, parents' smoking habit (McDonald and Shields:

2001), alcohol taxes between regions in a country (Barrett: 2002), religiosity (Heien: 1996, Hamilton and Hamilton: 1997), try to stop or cut down drinking, beer tax and its square(Terza: 2002), state-level alcohol and illicit drug policies and prices (Barrett, 2002; DeSimone, 2002), percentage of individuals who were abstainers or heavy drinkers within residing regions (Barrett: 2002) and perceived risks of consuming alcohol (Zarkin et al: 1998). I choose the following identifying instruments based on the current data set[44], review of previous studies and tests results of selection of instruments. These are (see Table 5 for definition and descriptions): Beer tax, Squared Beer Tax, Numbers of Children, Squared Numbers of Children, Parent problem drinker, Try to stop drinking, Spouse problem drinker, Smoker and Ex-smoker.

3.3.2 The test for endogeneity: the Durbin-Wu-Hausman test

Since prior empirical works related to alcohol use indicated that the RHS alcohol use measures might be potentially endogenous, this research conducts endogeneity test for alcohol use measures using the Durbin-Wu-Hausman (DWH) test which is an augmented version of Hausman (1978) test proposed by Davidson and MacKinnon (1993) and Wooldridge (2002). It is performed in two steps. In step 1, it obtains residuals from the regression of six potential endogenous variables (for six alcohol use status such as ex-drinkers, ex-abusers/dependents, new-drinkers, ex-new-drinkers, alcohol abusers, alcohol dependents) on selected instruments (see Table 5) and all other exogenous variables in the system (see Table 4). In step 2, it uses the residuals

[44] The data on beer tax is taken from “Alcohol Epidemiology Program, Alcohol Policies in the United States: Highlights from the 50 States”, 2000 University of Minnesota, School of Public Health.

from the first stage as additional regressors in the auxiliary regressions of LFPFULL and LINCOME equations, and performs (Wooldridge: 2002) the Wald- test[45] for auxiliary regression of LFPFULL and joint F-test[46] for auxiliary regression of LINCOME. The Wald test or the F-test is simply a test of the null hypothesis (Ho) that all of the residuals are jointly equal to zero (also implies “no endogeneity”). If the Wald-test or F-test rejects the null hypothesis that all of the residuals are jointly equal to zero then the endogeneity bias in the model is confirmed.

3.3.3 Tests for IV strength: instruments relevancy

It is essential to check the strength of the included instruments in the first stage regression of IV regression. Estimates of IV method based on weak instruments are biased towards the baseline estimates that are obtained without considering endogeneity (Bound et al: 2003). Weak instruments may also cause identification problem by partially affecting the dependent variable of structural equation (Dollar and Kraay: 2003). Prior studies suggest two simple diagnostic tests. Staiger and Stock (1997) suggested a rule of thumb that the F-test statistic of less than ten on the excluded instruments from the first stage regression is an indication of weak instruments. Baum et al (2003) and many others also suggested the use of Shea’s partial-R^2 in the model with multiple endogenous variables. A smaller value of Shea’s

---

[45] Wooldridge (2002) suggested that this test for the probit or logistic model, where the test can be used for testing the exclusion of single or more variables. The test is simply a Wald test (a test of significance on excluded residuals in the second step). The Wald-test has chi-square distribution with degrees of freedom (# of excluded restrictions) at chosen level of p-value.

[46] The F-test has F-distribution with degrees of freedom in numerator ((# of excluded restrictions) and in denominator (N – K, N=observations, K= # of estimated parameters including intercept) at chosen level of significance.

(1997) partial-R^2 than the standard partial-R^2 indicates weak instruments problem (Baum et al: 2003). This research uses Godfrey's (1999) version[47] to calculate Shea's partial-R^2.

3.3.4 The test of overidentifying restrictions (validity of instruments)

If the number of instruments excluded from the equation exceeds the number of included potential RHS endogenous variables, researchers must test the over-identifying restrictions (Davidson and Mackinon: 1993, Baum et al: 2003). It is a test for the exogeneity of the overall set of instruments. It also indicates if instruments have any direct influence on the dependent variable of the structural equation. Baum and others (2003) pointed out that the Sargan test (Sargan: 1958) is appropriate when there is no heteroscedasticity in the model, and the Hansen-J test[48] (Hansen: 1982) is appropriate in the presence of heteroscedasticity in the model. If the test statistic fails to reject (with high p-value, for example, more than 0.10) the null hypothesis (Ho) of excluding extra instruments, justifies the validity of the model (Baum at al.: 2003). It also confirms that the instruments are uncorrelated with the error process of structural model.

---

[47] Godfrey (1999) developed an easier version to calculate Shea's (1997) partial-R^2 model which equals squared of [(standard error of the coefficient$_{OLS}$/ standard error of the coefficient$_{2SLS}$) $\times$ (residual standard deviation$_{2SLS}$ / residual standard deviation$_{OLS}$)].

[48] Baum et al (2003) defined the Hansen-J statistic which is the minimized value of the corresponding generalized method of moments objective function multiplied by the sample size. It has chi-square distribution with degrees of freedom equals the numbers of extra restrictions and at chosen p-values.

3.3.5 Test of Heteroscedasticity

Since the presence of heteroscedasticity is most common in cross-sectional data, it is appropriate to use the Pagan-Hall (1983) test for checking the existence of heteroscedasticity when someone deals with endogenous regressors in the model and uses IV regression (Baum et al: 2003). The Pagan-Hall test is more appropriate for the IV estimation because it considers the presence of heteroscedasticity in all equations of the system. The Pagan-Hall (1983) pointed out that the standard test procedure developed by White (1980)/ Breusch-Pagan (1979) / Godfrey (1978) will only be valid if heteroscedasticity is present only in one equation and not in the other equation(s) of the system. This research uses one of the versions of Pagan-Hall test (four versions described by Baum et al: 2003). Wallentin et al (2002) and Gujarati (2003) also explained this simple method in the following steps. In step 1, it saves residuals and fitted values of the dependent variable from the second stage of IV regression. In step 2, it estimates auxiliary regression of squared residuals using the fitted values of the dependent variable and squared fitted values of the dependent variable, and then forms the Pagan-Hall test statistic ($N*R^2$, where N= observations, R^2= unadjusted R-square) using step 2 results. Pagan and Hall (1983) pointed out that diagnostic test is invalid if the fitted values of the dependent variable from the IV estimation is not used. The statistic ($N*R^2$) has chi-square distribution with two degrees of freedom at chosen p-value. If the statistic is statistically significant then it rejects the null hypothesis (Ho) that there is no heteroscedasticity (i.e., errors are homoscedastic). This confirms the presence of heteroscedasticity in the sample.

CHAPTER 4

DATA AND VARIABLES DESCRIPTIVE

4.1 DATA

4.1.1 Data Source

This research uses data from the first wave of survey sample collected by National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) in 2001-2002. The survey was conducted by the National Institute on Alcohol Abuse and Alcoholism (NIAAA). It is a large nationally representative sample of the United States population between 18 and 98 years of age including citizens and non-citizens. The major strength of NESARC survey is that it provides detailed information on the comprehensive measures for alcohol use disorders, and its related conditions. The nature of alcohol use disorders is classified based on clinical-diagnostic criteria described (see Appendix 2) in the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV). These clinical measures of alcohol separate casual or regular alcohol use cases from the problematic cases such as alcohol abuse and dependence. This survey also provides extensive information on labor market variables such as full and part time employment, personal and household income, work related variables such as industry, and occupation, basic socio-demographic variables including age, race, sex, and marital status, and human capital variables such as educational attainment and current health status. Total sample size is 43,093.

4.1.2 Categories of alcohol users or drinkers

As mentioned earlier (Chapter 3, Section 3.1), I use six binary variables for representing alcohol use status as the key independent variables. The following categories are considered based on alcohol use record and clinical diagnosis (see details in Tables 2A and 2B).

1. Lifetime abstainers refer to abstainers (did not drink in their lifetime). In the estimations, this category used as base category.
2. Ex-drinkers did not drink during the last year (2000-01).
3. Ex-abusers/dependents did not drink during the last year (2000-01).
4. New-drinkers drank only during the last year (2000-01).
5. Ex-new-drinkers drank during the last year (2000-01) as well as prior to last year.
6. Alcohol abusers were identified by DSM-IV criteria for alcohol abuse during the last year (2000-01).
7. Alcohol dependents were identified by DSM-IV criteria for alcohol dependence during the last year (2000-01).

According to this specification, it can be indicated that ex-drinkers, ex-abusers/dependents, new-drinkers, and ex-new-drinkers are the individuals who did not have any alcohol use disorders and alcohol abusers and dependents are the individuals who had alcohol use disorders. Tables 2A and 2B provide information on different drinkers categories based on current (2000-01) and previous (previous year of 2000-01) period drinking record as well as clinical diagnosis for using alcohol. In the context of LFPFULL (see Table 2A), the percentages of abstainers (base category), ex-drinkers, ex-abusers/dependent, new-drinkers, ex-new-drinkers, alcohol abuses and alcohol dependents were 17.0, 12.0, 5.0, 2.0, 55.0, 5.0 and 4.0 respectively. In the context of LINCOME (see Table 2B), the respective percentages were 12.0, 9.0, 5.0, 2.0, 62.0, 6.0 and 4.0.

4.2 Variables description

4.2.1 Dependent Variables

Two labor market outcomes are considered as dependent variables in this research (see Chapter 3.1, and Table 1). The first dependent variable 'LFPFULL' (full time work participation) is an indicator variable with a value of 1 if the respondent worked full time last year (i.e., worked at least 35 hours a week) and '0' otherwise. The second dependent variable is 'LINCOME' (natural log of annual personal income who worked full time). Annual personal income is used as a proxy measure for wage earnings since the data is not available. Wage rate could be the best variable but the NESARC did not include any information on wage variable. The results may not be sensitive to monthly or annual measures but it seems consistent to use annual income since this research considers another dependent variable (the probability of full time work participation) of labor market outcomes which is based on one year work information. Also, in order to avoid measurement error in the earnings variable and to improve the quality of the proxy earning variable this research sets annual personal income to zero for individuals who did not work full time last year. That is, this research limits the analyses of income to those who participated full time in the workforce (this procedure used by Buchmueller and Zuvekas: 1998, Marcotte: 2000 and Farahati: 2002). For this reason, and for omitting the non-reported income from work, the sample size is further reduced to 21406 for income (see Table 1) where as total sample size used for the logistic regression of the probability of full time work participation is 43093. Table 1 reports that 53 percent respondents worked full-time time in 2000-01, and mean annual income was US$39,337.

Table 3 reports the weighted mean of full-time labor force participation and annual income for abstainers (base category), ex-drinkers, ex-abusers/ dependent, new-drinkers, ex-new-drinkers, alcohol abuses and alcohol dependents. It indicates that 39.0, 39.0, 48.0, 41.0, 60.0, 72.0, and 56.0 percent of abstainers (base category), ex-drinkers, ex-abusers/dependent, new-drinkers and ex-new-drinkers, alcohol abuses and alcohol dependents worked full-time in the year 2000-01. It is seen that average annual income of abstainers (base category), ex-drinkers, ex-abusers/dependent, new-drinkers, ex-new-drinkers, alcohol abuses and alcohol dependents was $32884, $33410, $37304, $21023, $42366, $41193, and $32587 respectively. It can be noted here that lower incomes of new-drinkers are probably explained by the fact that majority of new-drinkers was young adults, and they might have entered in the job market with less experience and lower level of education. For example, by further investigation of the sample (not reported in the Table), I found that among new-drinkers, 1.0, 4.0, 6.0, 72.0, and 14.0 percent respectively had post-graduate, graduate, technical, high school and some high school education.

4.2.2 Independent variables

In the estimation of each labor market model (Chapter 3, section 3.1), besides alcohol use variables, a set of explanatory variables are also included for individual, socioeconomic and demographic variables characteristics (see Table 4). These are individual's age, gender, race, marital status, education level, health status, location (whether individual lived in a central city or not), other source of income, and work related characteristics (industry type).

Gender: Gender is an indicator variable with a value of 1 if respondents are male and 0 otherwise (i.e. females). According to observations that are used to estimate two labor market equations (LFPFULL and LINCOME), it is seen that the male-female ratios are 48:52 and 58:42 respectively. Gender is expected to have positive marginal effect on the probabilities of full time work and earnings. If it is found positive, it will indicate that the probabilities of full time work and earnings by males are higher than those by females (since females indicate by Gender=0), and this will support the gender disparity in terms of labor market variables.

Age groups: Two indicator variables for two age groups are used with a value 1 which indicates individuals belong to a particular age group, '0' otherwise. These variables are: AGE18-29 and AGE30-59 (see Table 4). According to observations that are used to estimate LFPFULL equation (Table 4), it is seen that 22.0 and 57.0 percent were in age groups AGE18-29 and AGE30-59 respectively, whereas these were 22.0 and 73.0 percent respectively in the context of LINCOME equation. The impacts of AGE18-29 and AGE30-59 on the probabilities of full time work participation and earnings are expected to be larger than age group AGE60-98 (base category). Comparatively, the impact of AGE30-59 is expected to be larger than the impact of AGE18-29.

Location: 'Central' is an indicator variable with value 1 indicating individuals who lived in a central city, 0 otherwise. According to observations that are used in labor market models (LFPFULL and LINCOME), it is observed that 30 percent of individuals lived in central city (see Table 4). The impact of 'Central' on labor market outcomes may be negative or positive. The direction and magnitude of the effect depends on the relative strength of living in a central city with more job opportunities and easy to access alcohol use. Usually the central city offers more job opportunities with probable higher payment than the other places (not a central city). Living in a central city may also open more opportunity and easy to access alcohol use, that is expected to impact negatively towards work participation.

Race: Three indicator variables for races are used with a value of 1 which indicates a particular race, '0' otherwise. These variables are: White, Black and Hispanic-Latino (see Table 4). According to observations that are used to estimate labor market models (LFPFULL and LINCOME), it is seen (Table 4) that respective percentages of White, Black, and Hispanic-Latino are (71.0 or 70.0 percent, 11.0 percent, and 12.0 respectively) almost the same in both estimations. No sign for the coefficient of race variables are assigned a priori.

Marital Status: 'Married' is an indicator variable with the value of 1 if a respondent is married or marriage-like relationship and '0' otherwise (widowed, divorced, separated and never married). From Table 4, it is seen, according to observations that are used to estimate labor market models (LFPFULL and LINCOME), that the percentages of married individuals 62.0 and 65.0 respectively. It is expected that marital status has positive impact on the labor market outcomes since married individuals live stable life and carry more responsibilities to support the family.

Other sources of Income: 'Had other source of income' is an indicator variable with a value 1 if individuals reported any positive amount of other household income, 0 otherwise. According to observations that are used to estimate labor market models (LFPFULL and LINCOME), it is seen that 81.0 and 64.0 percent of respondents had other source of income'. This variable is expected to have negative effect on labor market variables since other source of income discourages individuals from working more.

Education: Five indicators for education levels are used with value 1 which indicates a level of education, '0' otherwise. These are: Post-graduate, Graduate, Technical, Highschool, and Some high school (see Table 4). According to observations that are used to estimate LFPFULL equation, it is seen that 8.0, 17.0, 9.0, 51.0, and 9.0 percent of respondents had education level of Post-graduate, Graduate, Technical, Highschool, and Some high school. In the context of LINCOME these are 10.0, 21.0, 10.0, 49.0, and 7.0 respectively. It is generally expected that individuals with higher education will have a higher probability of participation in the labor force, and earn higher income, other things being equal.

Health Status: 'Excellent health status' is an indicator variable with a value of 1 if a respondent reported health status as excellent and very good, and 'Good health status' is an indicator variable with a value of 1 if a respondent reported health status as good or fairly good and 0 otherwise. The positive effects on labor market variables are expected for both of these variables compared to poor health status (base category). According to observations that are used to estimate LFPFULL equation, it is seen (Table 4) that 60.0 and 24.0 percent of respondents had 'Excellent health status' and 'Good health status' respectively. These are respectively 71.0 and 22.0 percent in the context of observations that are used to estimate LINCOME equation.

Industry type: Thirteen industry or business categories are indicated by thirteen indicator variables with a value 1 which indicates a particular industry where an individual worked, ‘0’ otherwise. These are (see Table 4): Agriculture, Mining, Construction, Manufacturing, Transportation, Wholesale Trade, Retail Trade, Finance/Insurance, Business, Personal Service, Entertainment, Professional and Related Services, and Public Administration. According to observations that are used to estimate LFPFULL and LINCOME equations, it is seen that majority of individuals worked in only four industries: Manufacturing, Retail Trade, Personal Service, Professional and Related services. No sign for the coefficient of industry variables are assigned a priori.

CHAPTER 5

EMPIRICAL RESULTS

The main focus of this research is to obtain the consistent estimates of the impact of alcohol use disorders on labor market outcomes. The estimation strategies are as follows. First, using the NESARC 2001-02 survey sample of 43093, I estimate a Logit model by the Maximum Likelihood Estimation (MLE) method where the dependent variable is an indicator variable for full-time labor force participation (Chapter 3, Section 3.1). Second, using a sample of 21406 (the observations for individuals who had job), I estimate a model with the logarithm of annual earnings by the Ordinary Least Square (OLS) method. In these estimations, I include a set of explanatory variables: six binary variables to indicate alcohol use status (see Tables 2A and 2B), and other binary variables for other personal and socio-demographic characteristics (see Table 4), such as individual's age, gender, race, marital status, education level, health status, location (whether individual lived in a central city or not), other source of income, and work related characteristics (industry type). Third, I perform formal tests to detect the existence of potential problems in estimations, the endogeneity and heteroscedasticity in the sample. As mentioned earlier, many prior empirical works indicated that alcohol consumption or alcohol use related right hand side (RHS) variables are potentially endogenous, and estimations without accounting endogeneity are biased and inconsistent. The estimation strategy (appropriate estimation method) to address endogeneity is required to consider the potential presence of heteroscedasticity in the sample since the sample set is cross-sectional (see details in Chapter 3, section 3.2). Fourth, as the statistical tests confirmed the existence of both problems (endogeneity and heteroscedasticity) in current estimation, I re-estimate labor market outcome equations by addressing endogeneity and heteroscedasticity using GMM-IV method (proposed by Baum et al: 2003, Amemiya: 1985 and Foster: 1997).

I begin analysis[49] by presenting the results of formal tests of endogeneity and heteroscedasticity. Though initially I estimated labor market equations by MLE and OLS methods, I present these results as baseline estimation results[50] along with the estimation results of the second stage of GMM-IV regression of the probability of full time labor force participation (See Table 8) and log annual income (See Table 9). The baseline results are used to evaluate relative contribution achieved by GMM-IV for addressing endogeneity and heteroscedasticity problems.

5.1 Results: the test for endogeneity: the Durbin-Wu-Hausman (DWH) test.

The endogeneity of each alcohol use variable (represent by a binary variable) is tested by the DWH test (following the procedure suggested by Davidson and Mackinnon: 1993, Wooldridge: 2002). The DWH test statistic is derived in two steps. In step one, I saved residuals from the regressions of alcohol use variables (each represent by a binary variable) on nine selected instrumental variables (see Table 5 for details) such as Beer Tax (cents per ounce), Squared Beer Tax, Numbers of Children, Squared Numbers of Children, Parental Problem Drink, Spousal Problem Drink, Current-smoker, Ex-smoker, and all others exogenous variables of the model (see Table 4 for details). In step two, I estimate auxiliary regression of each structural equation (equations 3.1.1 and 3.1.2) by including all variables of the structural model as

---

[49] All analyses were conducted in SAS 9.1

[50] These are weighted estimation results.

well as six residual variables obtained from step one. Exogeneity (i.e., no endogeneity) of each alcohol use variable (represent by a binary variable) is then tested by the DWH test statistic. The null hypothesis (Ho) is that the residuals (saved from the first stage regressions of alcohol use variables) are excluded jointly from the auxiliary regression of structural equation.

The DWH test results are presented in Table 6A. For the equation of the probability of full time labor force participation (LFPFULL), the DWH statistic is a chi-square value of Wald-test[51] (25.34 with p-value 0.003), and for the equation of log annual income (LINCOME), the DWH statistic is F-test value (8.04 with p-value less than 0.0001). These results clearly reject the null hypothesis of joint exclusion of residuals from the auxiliary regression of structural equations (equations 3.1.1 and 3.1.2). Though the estimated parameter of each residual variable (to be in perspective, the estimated coefficients of other exogenous variables are not presented in Table 6A) is not statistically significant individually in the auxiliary regressions of labor market outcomes, the rejection of the null hypothesis (i.e., joint exclusion of residuals) by the DWH test establishes joint significance of residuals and the endogeneity for alcohol use variable (represent by a binary variable). It also confirms that there is correlation between the potential endogenous variables and errors of the structural model (i.e. the evidence of endogeneity).

---

[51] Wooldridge (2002) indicated that the Wald chi-square value is appropriate for testing joint exclusion of variables in the probit or logistic model which is similar to the F-test value used for testing joint exclusion of variables in the linear regression model.

These results justify the necessity of using any of the IV regression methods (2SLS or GMM or LIML) instead of using the standard MLE and OLS methods to address endogeneity issue (see detail in Chapter 3, section 3.2).

5.2 Results: the test of heteroscedasticity (the Pagan-Hall test).

Considering the possibility of the presence of unknown form of heteroscedasticity in the sample (See Chapter 3, section 3.2), I conduct the Pagan-Hall (1983) test for heteroscedasticity. This test is chosen as it is described as a more appropriate test (Baum et al: 2003) for checking the existence of heteroscedasticity while dealing with endogenous explanatory variable in the model and choosing appropriate IV estimation method (considering the presence of heteroscedasticity in the system). The Pagan-Hall test (1983) is used for testing the null (Ho) hypothesis that there is no heteroscedasticity in the equation. This test statistic is calculated based on the results obtained from the auxiliary regression of squared residuals on the fitted values of dependent variable and squared fitted values of dependent variable that were saved from the second stage of IV regression.

The Pagan-Hall's test results are presented in Table 6B. The Pagan-Hall test statistic has chi-square distribution with degrees of freedom (equals the numbers of exogenous variables included in auxiliary regression) and p-values. The critical value of chi-square is 10.60 (with p-value of 0.005 and 2 degrees of freedom). It is seen that the Pagan-Hall test statistic for two labor market equations are 37142 and 21407 respectively. These exceed the critical value, and thus strongly reject the null hypothesis (Ho) of no heteroscedasticity. Since the Pagan-Hall test confirms the presence of heteroscedasticity in the model, this research chooses the GMM approach of IV estimation instead of 2SLS (see detail in Chapter 3, section 3.2).

5.3 Results: The GMM-IV estimation results.

As the statistical tests results confirmed that there exist endogeneity and heteroscedasticity problems in current model, I apply GMM-IV method (see Chapter 3, section 3.2) to estimate labor market outcome equations that address both issues. The analyses of the estimated results of GMM-IV are presented in two steps: the analysis of the results of the first stage regressions of GMM-IV and the second stage regressions of GMM-IV.

5.3.1 Results: the first stage regression of GMM-IV

The performance of the estimated results of the first stage regressions of GMM-IV (See Table 7A and 7B) is analyzed by assessing the strength of included identifying instrumental variables (IV strength).

A. Selection of identifying instrumental variables.

I select a set of nine instrumental variables for the GMM-IV estimation (see Table 5 for definition and descriptions): Beer tax, Squared Beer Tax, Numbers of Children, Squared Numbers of Children, Parent problem drinker, Try to stop drinking, Spouse problem drinker, Smoker, and Ex-smoker. These identifying instrumental variables are used in the first stage regression of each potential endogenous alcohol use variable (represent by a binary variable) but not in the second stage labor market equations. I started with a larger set of instrumental variables based on the reviews of previous empirical works, and applied formal testing for these variables using trial and error method (adding or dropping one by one) and finally select the set of instrumental variables that satisfy at least one assessment criterion (of instrumental relevancy) for all of the RHS endogenous variables in the estimation of both labor market equations by GMM-IV. Second, it is also seen (see Table 7A and 7B) that most of the estimated coefficients of individual instrument in the first stage regressions of GMM-IV are statistically significant.

B. Tests for IV Strength (instruments relevancy)

To check the strength of the included instruments (see details in Chapter 3, section 3.3.2) in the first stage regression of GMM-IV method (i.e., each RHS endogenous variable is regressed on instruments, and all other exogenous variables of the model) I apply two simple diagnostic tests (as suggested by Baum et al: 2003): the Wald test[52] for testing exclusion of instruments jointly from the first stage of GMM-IV regression, and the Shea's partial-R^2 (since multiple endogenous variables are included in the model). The tests results for the regression of LFPFULL and LINCOME are presented in Table 7A and 7B respectively. The Wald test statistics indicate strong relevance of included instruments since all of the Wald test statistics (see Table 7A and 7B) are much larger than the comparable F-value of 10 (the F-value of lower than 10 is considered as a flag of weak instruments by Staiger and Stock: 1997). In context of the estimation of LFPFULL equation by GMM-IV, the Shea's partial-R^2 showed higher values than the standard partial-R^2 in almost all of the first stage regressions of RHS endogenous variables except alcohol abuser case (see Table 7A). These results

[52] Wooldridge (2002) and Davidson and Mackinnon (1993) indicated that it is appropriate using the Wald-chi-square value for testing exclusion restrictions or significance of variables in the probit or logistic model. It is comparable to joint F-test for testing exclusion restrictions or significance of variables in the linear equation.

indicate strong relevance of included instruments. In context of the estimation of LINCOME equation by GMM-IV, the Shea's (1997) partial-R^2 is higher than the standard partial-R^2 for two RHS endogenous variables (ex-abusers/dependent and alcohol dependent) in the first stage regressions (see Table 7B), indicating strong relevance of instruments. The Shea's partial-R^2 showed lower value than the standard partial-R^2 for other four RHS endogenous variables (ex-drinkers, new-drinkers, new-ex-drinkers, alcohol-abusers) in the first stage regressions of LINCOME (see Table 7B), indicating weak relevance.

5.3.2 Results: the second stage regressions results of GMM-IV estimation

To analyze the performance of the estimated results of the second stage regressions of GMM-IV (See Tables 8 and 9), first, I present the test results of overidentifying restrictions and second, comparison of the GMM-IV estimates with baseline estimates of MLE and OLS.

5.3.2A Results: the test of overidentifying restrictions (validity of instruments)

This is a way to test the validity of included identifying instruments in the model (see details in Chapter 3, section 3.3.4). It is a test[53] of the null hypothesis (Ho) that the exclusions of extra instruments are valid (it also indicates true exogeneity of instruments i.e., instruments are uncorrelated with the error process of structural model). Considering the presence of heteroscedasticity in the sample (see section 5.2 of this

[53] Davidson and Mackinnon (1993): It is required to apply the test of overidentifying restrictions if the numbers of instruments excluded from the equation (second stage structural equation) exceeds the number of included potential RHS endogenous variables.

Chapter and Table 6B), I choose the Hansen-J test (1982) for testing the overidentifying restrictions (as suggested by Baum et al: 2003).

Table 6B presents the Hansen-J test statistics for overidentifying restrictions which are used to test the validity of nine identifying instruments included in the model. The Hansen-J test statistics[54] from the GMM-IV regressions (before heteroscedasticity correction) of LFPFULL and LINCOME are 5.37 (with p-value of 0.15) and 5.11 (with p-value of 0.16) respectively. Clearly, the null hypothesis is not rejected since the Hansen-J test statistics are lower (with higher p-value) than the critical chi-square value. The critical chi-square value is 6.25 (with 3 degrees of freedom for 3 excluded restrictions) at p-value of 0.10. These results confirmed that all instruments included in the model are truly exogenous. After correcting heteroscedasticity, the Hansen-J test statistics are 5.41 (with p-value of 0.14) and 5.16 (with p-value of 0.16) for respective GMM-IV regressions of LFPFULL and LINCOME. It also failed to reject the null hypothesis and thus verify the validity and exogeneity of included instruments.

5.3.2B The baseline estimations results by MLE and OLS vs. the second stage regressions results of GMM-IV

To allow comparison between the results of estimation before and after addressing estimation problems, I present the baseline estimations results along with the estimation results of the second stage of regression of GMM-IV (See Table 8 and 9).

[54] The Hansen-J test statistics has chi-square distribution with degrees of freedom equals the numbers of excluded restrictions (Baum et al. 2003).

5.3.2B.i. The baseline MLE results vs. the second stage results of GMM-IV of the probability of full time labor force participation (LFPFULL)

The baseline MLE results for LFPFULL

The estimated results of LFPFULL equation by the MLE and the second stage of regression by GMM-IV are presented in Table 8 (column 2-4) and Table 8 (column 5-7) respectively. I report estimated parameters, standard errors and marginal effects. It can be noted here that for an indicator variable in LFPFULL equation, the marginal effect is interpreted as the percentage point change in the probability of full time labor force participation resulting from a discrete change in binary explanatory variable.

Surprisingly, the estimated coefficients of all explanatory variables from the MLE of the probability of full time labor force participation (LFPFULL) are found statistically significant. According to main focus of this research, first I look at the estimated effects of alcohol use status indicating binary variables on LFPFULL. The key result from the baseline MLE is that the marginal effect of being alcohol dependent is -0.06 (significant at 1 percent level), indicating that alcohol dependents have a 6 percent less probability of being fully employed than lifetime abstainers (base category).This result is consistent with expectation. The positive significant marginal effects are found for ex-new- drinkers and alcohol abusers (0.04 and 0.07 respectively). The effect of alcohol abuser is not consistent with prior expectation. The indicator variables of being ex-abuser/dependents and new-drinkers showed significant (at 1 percent level) negative marginal effect on LFPFULL (-0.05 and -0.04 respectively). The marginal effect of being ex-drinkers on the probability of being employed full-time is 0.00, and it indicates that ex-drinkers have the same probability of being employed full-time as life-time abstainers (base category).

The estimated marginal effects of other personal and socio-demographic variables have shown expected impact on LFPFULL (see Table 8, column 4). The positive marginal effect of Male is 0.21 indicating male have a 21 percent higher probability of being employed fulltime than its female counterpart (base category). As expected, both younger age group (age 18-29) and middle age group (age 30-59) are found to have higher probability (in terms of marginal effect of 37 percent and 39 percent respectively) of being employed full-time than senior age group (age 60-98). Individuals who lived in central city have a 1 percent higher probability of being employed fulltime. The marginal effects are found to be positive for belonging to a particular race (White, Black, and Hispanic-Latino). The positive marginal effect of marital status (21 percent) on LFPFULL suggest that married individuals have higher probability of being employed full-time compared to their unmarried counterpart. The estimated marginal effects of different level of education on LFPFULL suggest that higher the level of education, higher the probability of being employed fulltime. Individuals who have other sources of income have a lower probability of being employed full-time. It is seen that the marginal effect of ‘excellent health status’ and ‘good health status’ on LFPFULL are 25 percent and 21 percent respectively. Interestingly, the marginal effects of all indicator variables representing industry type are almost the same (vary from 42 to 45 percent), indicating that the probability of being employed full-time in these industries (Agriculture, Mining, Construction, Manufacturing, Transportation, Whole Sale, Retail Trade, Finance or Insurance, Business, Personal Service, Entertainment, Professional, and Public Administration) is 42 to 45 percent higher compared to being employed full-time by Armed Force and Other Miscellaneous industries.

The second stage estimates of GMM-IV for LFPFULL

The second stage of regression results of GMM-IV for LFPFULL is presented in Table 8 (column 5-7). As compared to the MLE results, the marginal effects of alcohol use variables (represent by binary variables) on LFPFULL are found much larger in magnitude. The key result is that the marginal effects for alcohol abuse and alcohol dependent are -0.38 and -0.04 respectively (though not significant). It indicates that alcohol abusers and alcohol dependents have 38 percent and 4 percent less probability of being employed full-time respectively than life-time abstainers (base category), holding all other variables constant. The marginal effect of being ex-drinker on LFPFULL is negative and statistically significant. Ex-drinkers have 43 percent less probability of being employed full-time than life-time abstainers (base category). The positive marginal effects of being ex-abuser/dependents, new drinkers and ex-new-drinkers on LFPFULL (not significant) are 11 percent, 45 percent and 2 percent respectively though not significant.

The analysis of the estimated impacts of other exogenous variables on LFPFULL by GMM-IV indicates that the marginal effects of other socio-demographic variables on LFPFULL are much smaller in magnitude as opposed to MLE estimates. The estimated effects of other personal and socio-demographic variables on LFPFULL are mostly consistent with expectations and observed in many prior empirical works. The marginal effect of Male is 0.0 (not significant), and it indicates that the probability of being employed fulltime for a male is the same as for a female. As expected, it is found that middle age group (Age 30-59) has slightly higher (2 percent higher, but not significant) and younger age group (age 18-29) has lower (8 percent lower, but not significant) probability of being employed compared to age group 60-98 (base category). No difference was observed in the probability of being employed full-time for living or not in a central city.

The marginal effects of being White and Black are found positive (only significant for Black) whereas it is negative (not significant) for Hispanic-Latino. The positive (significant) marginal effect of marital status on LFPFULL suggests that married individuals have higher (4 percent) probability of being employed full-time than their unmarried counterparts. Surprisingly, binary variables indicating the levels of education are found to have inconsistent marginal effects on LFPFULL whereas the MLE estimations indicated consistent impacts. In MLE estimates, the estimated impacts of education variables indicate that individuals with higher level of education have a higher probability of being employed fulltime. It is found as expected that individuals having other sources of income have 14 percent (significant) less probability of being employed full-time compared to individuals without other sources of income. Health status indicator variables did not show any consistent or expected impact on LFPFULL.

Interestingly, the marginal effects of all indicator variables representing industry types vary from 3 to 8 percent (all are significant). It indicates that the probability of being employed full-time in certain industries (Agriculture, Mining, Construction, Manufacturing, Transportation, Whole Sale, Retail Trade, Finance or Insurance, Business, Personal Service, Entertainment, Professional, and Public Administration) are

3 to 8 percent higher than being employed in Armed Force or other miscellaneous industries (base category).

5.3.2B.ii. The baseline OLS results vs. the second-stage GMM-IV estimation of log of income (LINCOME)

The estimated results for the log annual income equation by the OLS and GMM-IV are presented in Table 9 (column 2-4) and Table 9 (column 5-7) respectively.

The baseline OLS estimates for LINCOME

Similar to the analysis of estimated results of LFPFULL equation, I first focus on the analysis of marginal effects of alcohol use variables (represent by binary variables) on annual income. The marginal effects of these variables (ex-drinkers, ex-abuser/dependents, ex-new drinkers, alcohol abusers, alcohol dependents) on annual income are found positive (statistically significant) except for the new-drinkers. In terms of marginal effect (not significant), new-drinkers have lower incomes than lifetime abstainers. In terms of affects of alcohol use variables on annual income, the estimated results are not consistent with prior expectation (except the impact of new-drinkers). Particularly, the marginal effects of indicators variables of being alcohol abuser and dependents on annual income are positive, indicating that individuals who have alcohol use disorders have higher annual income than life-time abstainers. Considering the reminiscent effect of drinking as mentioned by Johansson and et al. (2004), it was expected (holding all other variables constant) that ex-drinkers and ex-abuser/dependents would have lower probability of being employed full-time and lower income compared to lifetime abstainers (base category). Also, I expected that individuals who drank for longer period (drank in the past years and also a current drinker, i.e., drank in 2000-01) would have lower probability of being employed full-time and lower income than lifetime abstainers (base category).

The marginal effects of most of the other explanatory variables are found as expected. For example, being male is associated with an increase in annual income. Being married is also associated with an increase in income. Other sources of income are found to have negative impacts on income. In terms of education, it is found that higher the level of education, higher the annual income. In terms of health status, it is found that better health status is associated with higher income as compared to poor health status. In terms of race, the marginal effect of being White is positive (significant) whereas the marginal effect of being Black and Hispanic-Latino are negative (but only significant for Hispanic-Latino). The marginal effect of being lived in a central city on income is positive but not significant. In terms of age, the impact of both age groups on income is somewhat inconsistent. The marginal effects of age group 18-29 and age group 30-59 on income are negative (the effect of age group 18-29 is only statistically significant). Positive (significant) effects on income are observed for the binary variables indicating industries of Mining, Manufacturing, Transportation, Finance or Insurance and Public Administration whereas negative (significant) impacts are observed for the industries of Agriculture, Retail Trade and Personal Service.

<u>The second-stage GMM-IV estimates for LINCOME</u>

The second stage of regression results of GMM-IV for LINCOME is presented in Table 9 (column 5-7). Compared to OLS estimates, the estimated parameters, standard errors and marginal effects of all explanatory variables are found much larger in magnitude. The key result is that the marginal effect for alcohol abuse on annual income is -112057 (though not significant). It indicates that alcohol abusers earned $112057 less than life-time abstainers (base category), holding all other variables constant. The marginal effects of all other alcohol use status indicating binary variables (ex-drinker, ex-abuser/dependent, new-drinker, ex-new-drinker and alcohol dependents) are found to have positive impacts on annual income (statistically significant for ex-drinkers and new-ex drinkers). The positive impact on annual income for being alcohol dependent is inconsistent with expectation.

Compared to OLS estimates, the analysis of the estimated impacts of other exogenous variables on LINCOME by GMM-IV indicates that the marginal effects of other socio-demographic variables on LINCOME are mostly similar but much larger in magnitude. Similar to OLS estimates, the estimates of the most socio-demographic variables from the second stage of GMM-IV are found consistent with expectation. Both variables being 'male' (significant effect) and 'married' (not significant) have positive effects on annual income. The indicator variable representing 'other sources of income' has a significant negative impact on income. The following indicator variables such as Post-graduate, Graduate, Technical, High school, and Some High School (represent different level of education) have marginal effects of $39879, $33500, $22927, $16659 and $1425 on annual income (all are statistically significant except 'Some high School'). This result suggests that higher the level of education, higher the annual income. Both health status indicator variables (Excellent health status and Good health status) show positive impacts on annual income (though only significant for 'Excellent health status'). In terms of race, the marginal effects of being White, Black and Hispanic-Latino are negative (but only significant for Black). The marginal effect of being lived in a central city on income is found positive but not significant. In GMM-IV estimations, the impact of both age groups on income is found consistent with expectation as opposed to OLS results. The marginal effects of age group 18-29 and age group 30-59 are -$19530 and $1983 respectively on income though not statistically significant. None of the binary variables indicating industry types has significant effect on annual income.

In sum, differences in the estimated results mainly arise due to the differences of estimation strategy between two procedures. In the estimation of LFPFULL by GMM-IV, the estimated coefficients of indicator variables represent alcohol use status become larger but the estimated coefficients of other variables become smaller compared to the MLE estimates. The estimated coefficients of all variables in LINCOME equation from the GMM-IV estimation are also found larger in magnitude than the OLS estimation. The standard errors of the estimated coefficients of all variables are larger in the GMM-IV estimations of both labor market equations compared to the results obtained by the baseline MLE and OLS. This result is also consistent with the findings of previous empirical works that found larger effects from the estimation accounting endogeneity compared to the baseline estimates that did not account endogeneity.

CHAPTER 6

CONCLUSIONS

6.1 Summary

A. A couple of important existing facts motivated this research to re-examine to what extent performance of individuals who had alcohol use disorders (alcohol abuse and alcohol dependence) differ from individuals who had no alcohol use disorders (lifetime abstainers, ex-drinkers, ex-abuser/dependents, new-drinkers, and ex-new-drinkers) in terms two labor market outcomes: the probability of full time work participation and annual personal income. First, based on most recent survey on alcohol and alcohol related conditions the NIAAA (2004) reported that alcohol use disorders (alcohol abuse and dependency) became more prevalent in the working age population and its effects on employment and productivity were likely to become more detrimental for the society. Second, findings of a large literature on the nature and extent of relationship between alcohol use and labor market outcomes remained debatable, heterogeneous and often counterintuitive. Explanations for differences have included, types of alcohol use measures (self reported or clinical), various labor market variables (employment, hours of work, productivity, wage or income), different statistical method of estimations (particularly to address the endogeneity problem), and quality of data sets.

B. The main focus of this research is to explore to what extent alcohol use disorders affects labor market outcomes. This research uses a typical specification (see details in Chapter 3, section 3.1) of labor market outcomes (French and Zarkin: 1995) with some modifications guided by the theoretical framework of human capital described in literature (Becker: 1965, Grossman: 1972, 1991, and Mincer: 1974) and the empirical model specifications of prior studies (Chatterji et al: 2007, French and Zarkin: 1995, Mullahy and Sindelar: 1993). The estimation strategies are as follows.

First, using the NESARC 2001-02 survey sample of 43093, I estimate a Logit model by the maximum likelihood estimation (MLE) method where the dependent variable is an indicator variable for full-time labor force participation (Chapter 3, Section 3.1). Using a sample of 21406 (the observations for individuals who had full-time job), I estimate a model with the logarithm of annual earnings by the ordinary least square (OLS) method. In these estimations, I include a set of explanatory variables: six binary variables to indicate alcohol use status (see Tables 2A and 2B), and other binary variables for other personal and socio-demographic characteristics (see Table 4), such as individual’s age, gender, race, marital status, education level, health status, location (whether individual lived in a central city or not), other sources of income, and work related characteristic (industry type).

Second, since reviews of previous empirical work on the relationship between labor market outcomes and alcohol use measures indicated that the ‘endogeneity’ problem (due to simultaneity and unobserved heterogeneity) complicated standard estimation procedure and the estimation without addressing this issue resulted in biased and inconsistent estimates, I apply the formal tests to detect the presence of endogeneity. After detecting the presence of endogeneity, I consider to use the IV method of estimation to address this issue. In order to choose appropriate IV method of estimations (2SLS or GMM or LIML, see details in Chapter 3, Section 3.2), I also perform the required test (as suggested by Baum et al.: 2003) for possible presence of heteroscedasticity (which is most common in cross-sectional data).

Third, as the diagnostic tests results (see Chapter 5, sections 5.1 and 5.2) confirmed the presence of endogeneity and heteroscedasticity in the current model, I re-estimate labor market outcome equations by addressing endogeneity and heteroscedasticity using GMM-IV method (as proposed by Baum et al: 2003, Amemiya: 1985 and Foster: 1997).

C. This research contributes to the existing literature in several ways.

First, the results can be considered as nationally representative since it used a large nationally representative survey data set of the United States population (NESARC: the first wave data sets, 2001-2002) .This data set not only provided rich information of socioeconomic, demographic and work specific variables but also provided more reliable measures of alcohol use disorders (diagnosed by clinical criteria as described in DSM-IV).

Second, besides using the rich and reliable data sets, this research indicates the importance of exclusive classification of observations based on individual's past (before 2000) and present (2000-2001) alcohol use record and clinical diagnosis to identify someone as regular or casual drinker or alcohol abuser or alcohol dependent. Individuals were either abstainers (who never drank) or alcohol drinkers. Based on individual's drinking record of current (2000-01) and prior years, three categories of alcohol drinkers are considered, such as ex-drinkers (did not drink during the year 2000-01, but drank prior to that), new-drinkers (drank only during 2000-01) and ex-new-drinkers (drank during 2000-01, and also drank prior to 2000-01). Each category again classified for the diagnosis of alcohol abuse or dependents. Ex-drinkers, ex-abusers/dependents, new-drinkers, and ex-new-drinkers are the individuals who did not have any alcohol use disorders whereas alcohol abusers and dependents are the individuals who had alcohol use disorders.

Observed difference in the estimated results of labor market variables, the probabilities of full time work participation (LFPFULL) and annual income (INCOME), largely relates to the following questions: i) were the individuals abstainers or alcohol drinkers? ii) did individuals use to drink last year or years before that or in both periods? iii) were the individuals' alcohol abusers or dependents or neither?

Third, differences in the estimated results also arise due to different estimation strategies that are used before and after addressing endogeneity and heteroscedasticity. Though the estimated results by MLE and OLS methods are biased and not reliable (because these do not address endogeneity and heteroscedasticity), these baseline results are used to evaluate relative contribution achieved by GMM-IV for addressing endogeneity and heteroscedasticity problems. The followings are the key results.

1. The baseline MLE and OLS estimations of labor outcome models treat observed alcohol use related indicator variables as exogenous. After detecting the presence of endogeneity and heteroscedasticity in the model by the statistical tests (Chapter 5, sections 5.1 and 5.2), labor market equations are estimated by GMM-IV method (as described Chapter 3, section 3.2). The estimations of GMM-IV are performed in two stages. In the first stage, the predicted or fitted values are derived from the estimation of each alcohol use variable (represented by a binary variable, see

Tables 2A and 2B) using identifying instruments and all other exogenous variables of the structural model (Chapter 3, section 3.1). The predicted values of alcohol use related indicator variables are used (instead of observed values of alcohol use related indicator variables) in the second stage of GMM-IV estimations of labor outcome models.

2. The results of the first stage regressions of GMM-IV are assessed by the strength of the included instrumental variables (see details in Chapter 3, section 3.3.2) since each RHS endogenous variable (each alcohol use status indicating variable) is regressed on instruments, and all other exogenous variables of the model. To check the strength of the included instruments, two diagnostic tests (as suggested by Baum et al: 2003) are used: the Wald test for testing exclusion of instruments jointly from the first stage of GMM-IV regression, and the Shea's partial-R^2 (since multiple endogenous variables are included in the model). For both labor market equations, the Wald test statistics indicate strong relevance of included instruments (see Table 7A and 7B). According to the Shea's partial-R^2, strong relevance of included instruments is found in the first stage GMM-IV of LFPFULL whereas weak relevance is found in the first stage GMM-IV of LINCOME (see details in Chapter 5, section 5.3.1).

3. The estimated results of the second stage regressions of GMM-IV (See Tables 8 and 9) are assessed by the test results of overidentifying restrictions and the performance of explanatory variables (as compared to expectation and prior empirical findings) in second stage regressions of GMM-IV. To perform the test of overidentifying restrictions (for establishing the validity of instruments), I apply the Hansen-J test (1982), considering the presence of heteroscedasticity in the sample (as suggested by

Baum et al: 2003). After correcting heteroscedasticity, the Hansen-J test statistics (Table 6B) are 5.41 (with p-value of 0.14) and 5.16 (with p-value of 0.16) for respective GMM-IV regressions of LFPFULL and LINCOME. Since the Hansen-J test statistics are lower (with higher p-value) than the critical chi-square value of 6.25 (with 3 degrees of freedom and p-value of 0.10), they failed to reject the null hypothesis of the validity of the exclusions of extra instruments. This result suggests the true exogeneity of included instruments i.e., instruments are uncorrelated with the error process of structural model.

4. According to main focus of this research, first I look at the estimated effects of alcohol use variables (each represent by a binary variable) on the probability of full time labor force participation (LFPFULL). Surprisingly, the estimated coefficients of all explanatory variables from the MLE of LFPFULL are found statistically significant. The key result from the baseline MLE is that the marginal effect of being alcohol dependent is -0.06 (significant at 1 percent level), indicating that alcohol dependents have a 6 percent less probability of being fully employed than lifetime abstainers (base category).This result is consistent with expectation. The positive significant marginal effects are found for ex-new- drinkers and alcohol abusers (0.04 and 0.07 respectively). The effect of alcohol abuser is not consistent with prior expectation though the effect of ex-new- drinkers is consistent with prior empirical finding. The indicator variables of being ex-abuser/dependents and new-drinkers showed significant (at 1 percent level) negative marginal effect on LFPFULL (-0.05 and -0.04 respectively). The impact of ex-abuser/dependents indicates the reminiscent effect of prior drinking as described in previous empirical work (Johansson and et al.: 2004) The marginal effect of being ex-drinkers on the probability of being employed full-time is 0.00, indicating that ex-drinkers

have the same probability of being employed full-time as life-time abstainers (base category).

In second stage GMM-IV regression of LFPFULL, the marginal effects of alcohol use variables (represent by binary variables) on LFPFULL are found much larger in magnitude as compared to the MLE results. The key result of GMM-IV regression of LFPFULL is that the marginal effects for alcohol abuser and alcohol dependent are -0.38 and -0.04 respectively (though not significant). indicating that alcohol abusers and alcohol dependents have 38 percent and 4 percent less probability of being employed full-time respectively than life-time abstainers (base category), holding all other variables constant. The marginal effect of being ex-drinker on LFPFULL is negative (statistically significant) and consistent with expectation. Ex-drinkers have 43 percent less probability of being employed full-time than life-time abstainers (base category). The positive marginal effects of being ex-abuser/dependents, new drinkers and ex-new-drinkers on LFPFULL (not significant) are 11 percent, 45 percent and 2 percent respectively though not significant. The results for ex-abuser/dependents and new drinkers are not consistent with prior expectation though the result for ex-new-drinkers in consistent with previous empirical findings.

5. According to the estimated results of log annual income (LINCOME) by the OLS, the marginal effects of being ex-drinkers, ex-abuser/dependents, ex-new drinkers, alcohol abusers, alcohol dependents on annual income are found positive (statistically significant) except for the new-drinkers. In terms of marginal effects (not significant), new-drinkers have lower earning than lifetime abstainers. In terms of effects of alcohol use variables (represent by binary variables as described in Tables 2A and 2B) on

annual income, the estimated results are mostly inconsistent with prior expectation and previous empirical findings. This could be the result of estimation without addressing endogeneity and heteroscedasticity.

In second stage GMM-IV regression results of LINCOME compared to OLS estimates, the estimated parameters, standard errors and marginal effects of all explanatory variables are found much larger in magnitude. The key result of the GMM-IV regression of LINCOME is that the marginal effect for alcohol abuse on annual income is -112057 (though not significant). It indicates that alcohol abusers earned $112057 less than life-time abstainers (base category), holding all other variables constant. The marginal effects of all other alcohol use status indicating binary variables (ex-drinker, ex-abuser/dependent, new-drinker, ex-new-drinker and alcohol dependents) are found to have positive impacts on annual income (statistically significant for ex-drinkers and new-ex drinkers). The positive impact on annual income for being alcohol dependent is inconsistent with expectation.

6. To be concise in reporting the reliable (addressed endogeneity and heteroscedasticity) estimates of the effects of other socio-demographic variables on labor market equations, I focus mainly on the analysis of estimated effects in the GMM-IV regressions of LFPFULL and LINCOME as compared to MLE and OLS regression.

6a. In the GMM-IV regressions of LFPFULL, the marginal effects of other socio-demographic variables on LFPFULL are found much smaller in magnitude as opposed to MLE estimates. The marginal effect of Gender (Male) is 0.0 (not significant) indicating that the probability of being employed fulltime for a male is the same as for a female. As expected, it is found that middle age group (Age 30-59) has slightly higher (2 percent higher, but not significant) and younger age group (age 18-29) has lower (8 percent lower, but not significant) probability of being employed compared to age group 60-98 (base category). No difference was observed in the probability of being employed full-time for living or not in a central city. The marginal effects of being White and Black are found positive (only significant for Black) whereas it is negative (not significant) for Hispanic-Latino. The positive (significant) marginal effect of marital status ('Married') on LFPFULL suggests that married individuals have higher (4 percent) probability of being employed full-time than their unmarried counterparts. Surprisingly, in GMM-IV regression of LFPFULL, the estimated effects of education variables and health status variables did not show any consistent or expected impact on LFPFULL. In this context, the MLE estimates show consistent impacts: individuals with higher level of education have a higher probability of being employed fulltime, and individuals with better health have a higher probability of being employed full-time. Interestingly, the marginal effects of all indicator variables representing industry types vary from 3 to 8 percent (all are significant), and indicates that the probability of being employed full-time in certain industries (Agriculture, Mining, Construction, Manufacturing, Transportation, Whole Sale, Retail Trade, Finance or Insurance, Business, Personal Service, Entertainment, Professional, and Public Administration) are 3 to 8 percent higher than being employed in Armed Force or other miscellaneous industries (base category). In MLE estimates, it is seen that the impacts of all indicator variables representing industry type on LFPFULL are much larger (vary from 42 to 45 percent) in magnitude.

6b. Comparison of OLS estimates of LINCOME and GMM-IV estimates of LINCOME in terms of the effects of other exogenous variables (besides alcohol use variables) on LINCOME indicates that in GMM-IV regression of LINCOME, the marginal effects of other socio-demographic variables on LINCOME are mostly similar but much larger in magnitude. As expected, both variables, Gender (significant) and 'Married' (not significant) are found to have positive effects on annual income. The indicator variable representing 'other sources of income' has a significant negative impact on income. Different levels of education such as Post-graduate, Graduate, Technical, High school, and Some High School (represent different level of education) have marginal effects of \$39879, \$33500, \$22927, \$16659 and \$1425 on annual income (all are statistically significant except 'Some high School'). This result suggests that higher the level of education, higher the annual income. Both health status indicator variables (Excellent health status and Good health status) show positive impacts on annual income (though only significant for 'Excellent Health Status'). In terms of race, the marginal effects of being White, Black and Hispanic-Latino are negative (but only significant for Black). The marginal effect of being lived in a central city on income is found positive but not significant. Compared to OLS results, the marginal effects of age group 18-29 and age group 30-59 are -\$19530 and \$1983 respectively on income (though not statistically significant) as expected. None of the binary variables (indicating industry type) has significant effect on annual income.

7. In brief, the estimated results of LFPFULL by GMM-IV compared to the MLE estimation results indicates that the estimated parameters and marginal effects of alcohol use variable (represent by binary variables) are larger whereas the estimated parameters and marginal effects of other variables are smaller. The estimated result of LINCOME by GMM-IV compared to the OLS estimation indicates that the estimated

parameters and marginal effects of all variables became larger in magnitude. The standard errors of the estimated parameters of all variables are found to be larger in GMM-IV estimations of both labor market equations compared to the results obtained by the baseline MLE and OLS. This result is also consistent with the findings of previous empirical works that found larger effects from the estimation accounting endogeneity as compared to the baseline estimations without addressing endogeneity. Compared to GMM-IV estimates, the MLE (without addressing endogeneity and heteroscedasticity) underestimates the effects of alcohol use variables (represent by binary variables) and overestimates the effects of other socio-demographic variables on labor market variables, and the OLS (without addressing endogeneity and heteroscedasticity) underestimates the effects of all explanatory variables on labor market outcomes.

8. As the statistical tests confirmed the endogeneity of alcohol use related variables and the presence of heteroscedasticity in the sample, the estimation by applying GMM-IV are expected to produce consistent and efficient estimates compared to baseline estimates without addressing the endogeneity and heteroscedasticity. Based on present data set, the efficiency gain by GMM-IV estimations is observed only in the following contexts. First, alcohol abusers and dependents have lower probability of being employed fulltime (though not significant) compared to lifetime abstainers. Alcohol abusers earn $112057 less annually than lifetime abstainers. According to current data set, the similar evidence was not found for alcohol dependents though it was expected. Second, the GMM-IV results of LFPFULL also indicates that individuals with alcohol disorders (abusers and dependents) perform worse in labor market (in terms of the probability of full-time employment) than individuals who with no alcohol

use disorders (abstainers, ex-drinker, ex-abuser/dependents, new-drinker, and ex-new-drinker). The evidence found from the estimation of log of annual income equation by GMM-IV is not similar.

6.2 Limitations of the research

The current research is restricted by the following limitations.

First, although the NESARC provided the largest data sets ever available on alcohol use and related conditions, and this research used weighted estimation method to maintain that national representation of the original data sets, it is possible that results might be underestimated or overestimated since only 13 percent of individuals had alcohol use disorders (among the ex-new-drinkers) compared to 87 percent of individuals who had no alcohol use disorders.

Second, this research used lifetime measures only for abstainers since the data set did not provide lifetime measures for the drinker categories. Lifetime alcohol measures could further minimize potential endogeneity bias in estimation.

Third, as mentioned earlier, wage rate or hours of work could be the best labor market variable to reflect earning capacity or productivity but the NESARC did not include any information on wage or hours of work. For this reason, this research used annual personal income (if respondents worked full time) as a proxy measure for earnings (this procedure used by Buchmueller and Zuvekas: 1998, Marcotte: 2000, and Farahati: 2002). Still, there exists a possibility of arising measurement errors for using the proxy measure.

6.3 **Implications**

Besides the limitations, the estimated results are largely consistent with results found in parallel health and labor economics literature. The results can be considered representative because this research used the rich and nationally representative NESARC data source. This research suggests that results can be useful for further research by addressing the limitations of this study and for the employers and respective authorities to face the challenges of having and maintaining productive and healthy work force. Since it is observed that major proportions of the alcohol abusers or dependents were in the full time job in the current year (2000-2001), this might imply that alcohol use disorders contribute to work loss through absence from work or fewer hours of work and health-related work limitations (though it did not provide any explicit measure because of the lack of data on absence from work or hours of work). To minimize the adverse impacts of alcohol use and misuse (alcohol use disorders), it requires clear and well communicated policies concerning recruitment, monitoring, early prevention, access to effective treatment, and maintaining positive work environment.

6.4 Future Research Questions

The current research used the data from the first wave (2001-2002) of the NESARC study. Further research or a comparative study on same issue can be carried out using upcoming second wave data. Besides this, more research can be conducted to address the following highlighted issues.

First, this research estimated the probability of work participation and income equation using the data of full time work (individuals who worked at least 35 hours a week in last year). As alcohol use disorders can limit the working hours to the extent of drop out or job loss. Obviously, there exists scope to investigate the relation of drop out or job loss with alcohol use disorders.

Second, prior research indicated that female reaction to the amount of alcohol consumption is different than that of males. Further research on the relationship of labor market variables and different amount of alcohol consumption (light, moderate and heavy) are expected to clarify more.

Third, alcohol use disorders may contribute to work loss and productivity through absent from the job, or less working hours, force to choose part-time job, and health-related work limitations. These can be explored in future research in order to quantify the actual loss due to alcohol use.

Table 1: Definition and Descriptive Statistics of Dependent Variables (Labor Market Outcomes)

(Weighted Mean)

Dependent Variables (Labor Market Outcomes)	Definition	N	Mean	SD
LFPFULL	=1 if respondents worked full time (worked at least 35 hours in a week) during the year 2000-2001 [a/], '0' otherwise.	43093	53.0	34.64
INCOME	Annual Income (US$) during the year 2000-2001 [a/], if LFPFULL=1	21406	39337	1354784.1

Notes: N= the number of observations. SD=Standard Deviation.

a/ The preceding year of the sample collection year 2001-2002.

Table 2A: Definition and Descriptive Statistics of Different Drinkers Categories used as Key Independent Variables in Full-time Work Participation (LFPFULL) Equation

			(Weighted Mean)	
Variable Name	Definition	N	Percent	SD
Lifetime Abstainers (used as base category)	=1 if respondent did not drink in lifetime	8266	17.0	26.26
Ex-drinkers [d/]	=1 if respondent drank prior to last years, but did not drink during the last year (2000-01) [a/]	5645	12.0	22.54
Ex-abuser/dependent [d/]	=1 if respondent drank prior to last year, but did not drink during the last year (2000-01) [a/]	2236	5.0	15.58
New-drinkers [d/]	=1 if respondent did not drink prior to last year, but drank during the last year (2000-01) [a/]	973	2.0	10.51
Ex-New drinkers [d/]	=1 if respondent drank both during the last year (2000-01) [a/] and prior to last years	22646	55.0	34.58
Alcohol Abuser	=1 if respondent diagnosed for alcohol abuse (who drank in last year and also prior to last year) [b/]	1843	5.0	14.63
Alcohol Dependent	=1 if respondent drank diagnosed for alcohol dependence (who drank in last year and also prior to last year) [c/]	1484	4.0	13.29
Total Observations:		43093	100.0	

Notes:

a/ The preceding year of the sample collection year 2001-2002.
b/ Includes observations of 28 alcohol abusers who started drinking only in last year (2001-02).
c/ Includes observations of 62 alcohol dependents who started drinking only in last year (2001-02).
d/ Individuals who did not have any alcohol use disorder for alcohol abuse or dependence during the last year (2000-01).

Table 2B: Definition and Descriptive Statistics of Different Drinkers Categories used as Key Independent Variables in Income Equation (if LFPFULL=1)

		(Weighted Mean)		
Variable Name	Definition	N	Percent	SD
Lifetime Abstainers (used as base category)	=1 if respondent did not drink in lifetime	2901	12.0	16.41
Ex-drinkers [d/]	=1 if respondent drank prior to last years, but did not drink during the last year (2000-01) [a/]	1990	9.0	14.06
Ex-abuser/dependent [d/]	=1 if respondent drank prior to last year, but did not drink during the last year (2000-01) [a/]	998	5.0	10.71
New-drinkers [d/]	=1 if respondent did not drink prior to last year, but drank during the last year (2000-01) [a/]	376	2.0	6.39
Ex-New drinkers [d/]	=1 if respondent drank both during the last year (2000-01) [a/] and prior to last years	13055	62.0	24.15
Alcohol Abuser	=1 if respondent diagnosed for alcohol abuse (who drank in last year and also prior to last year) [b/]	1271	6.0	12.13
Alcohol Dependent	=1 if respondent drank diagnosed for alcohol dependence (who drank in last year and also prior to last year) [c/]	815	4.0	9.71
Total Observations:		21406	100.0	

Notes:

a/ The preceding year of the sample collection year 2001-2002.
b/ Includes observations of 28 alcohol abusers who started drinking only in last year (2001-02).
c/ Includes observations of 62 alcohol dependents who started drinking only in last year (2001-02).
d/ Individuals who did not have any alcohol use disorder for alcohol abuse or dependence during the last year (2000-01).

Table 3: Mean Measures of Labor Market Outcomes by Drinkers Category

(Weighted Mean)

Drinkers Category [b/]	Labor Market Outcomes [a/] LFPFULL N= 43093		INCOME (US$) N=21406	
	Percent	SD	Mean	SD
Lifetime Abstainers (used as reference category)	39.0	32.20	32884	1119850
Ex-drinkers [c/]	39.0	32.39	33410	1058600
Ex-abuser/dependent [c/]	48.0	35.13	37304	126865
New-drinkers [c/]	41.0	34.83	21023	726053
Ex-New drinkers [c/]	60.0	34.63	42366	1431442
Alcohol Abuser	72.0	32.41	41193	1425084
Alcohol dependent	56.0	36.28	32587	1182442
Total:	53.0	34.64	39337	1354784

Note:

a/ Please see Table 1 for details.

b/ Please see Table 2A or 2B for details.

c/ Individuals who did not have any alcohol use disorder for alcohol abuse or dependence during the last year (2000-01).

Table 4: Definitions and Descriptive Statistics of Other Control Variables used in Two Labor Market Outcome Equations

(Weighted Mean)

Control Variables (base category)	Definition	Labor Market Outcomes LFPFULL N=43093 %	SD	INCOME N=21406 %	SD
Age-group (age 60-98):					
AGE18-29	=1 if respondents are in age group 18-29	22.0	28.68	22.0	20.47
AGE30-59	=1 if respondents are in age group 30-59	57.0	34.43	73.0	22.04
Location (not in central city):					
Central	=1 if respondents from the central city	30.0	31.69	30.0	22.76
Gender(Female):					
Male	=1 if respondents are Male	48.0	34.70	58.0	24.53
Race (other races):					
White	=1 if respondents are White, not Hispanic/ Latino	71.0	31.55	70.0	22.77
Black	=1 if respondents are Black, not Hispanic or Latino	11.0	21.79	11.0	15.76
Hispanic-Latino	=1 if respondents are Hispanic or Latino	12.0	22.21	12.0	16.42
Marital status (other):					
Married	=1 if respondents reported married	62.0	33.78	65.0	23.67
Other Income (none):					
Had other source of income	=1 if respondents reported any source of other income	81.0	27.19	64.0	23.81
Education level (below 9th grade):					
Post-graduate	= 1 if respondents reported Master's / Higher	8.0	18.92	10.0	14.95
Graduate	= 1 if respondents reported Bachelor degree	17.0	25.97	21.0	20.15
Technical	= 1 if respondents reported Asst. technical degree	9.0	19.84	10.0	15.17
High School	= 1 if respondents reported Completed High School or GED	51.0	34.73	49.0	24.87
Some High School	= 1 if respondents reported Some High School	9.0	20.32	7.0	12.37
Health Status (Poor):					
Excellent health status	=1 if respondents reported excellent or very good health status	60.0	33.95	71.0	22.62
Good health status	=1 if respondents reported good or fair health status	24.0	29.60	22.0	20.64

Table 4 continues…….

Control Variables (base category)	Definition	LFPFULL N=43093		INCOME N=21406	
Industry type (Armed Services or other)		%	SD	%	SD
Agriculture	= 1 if respondents industry type was Agriculture	3.0	11.80	3.0	8.54
Mining	= 1 if respondents industry type was Mining	0.1	4.43	1.0	3.70
Construction	= 1 if respondents industry type was Construction	6.0	16.01	8.0	13.27
Manufacturing	= 1 if respondents industry type was Manufacturing	10.0	20.64	14.0	17.06
Transportation	= 1 if respondents industry type was Transportation/Communication	6.0	16.18	8.0	13.62
Wholesale	= 1 if respondents industry type was Wholesale Trade	2.0	9.00	2.0	7.35
Retail Trade	= 1 if respondents industry type was Retail Trade	12.0	22.23	11.0	15.64
Finance/Insurance	= 1 if respondents industry type was Finance/ Insurance/ Real Estate	5.0	14.60	6.0	12.09
Business	= 1 if respondents industry type was Business repair service	3.0	12.11	4.0	10.01
Personal service	= 1 if respondents industry type was Personal service	9.0	19.79	9.0	13.97
Entertainment	= 1 if respondents industry type was Entertainment/ Rec. service	2.0	10.65	2.0	7.48
Professional	= 1 if respondents industry type was Professional service	19.0	27.43	25.0	21.47
Public Administration	= 1 if respondents industry type was Public Administration	4.0	13.96	6.0	11.77

Table 5: Definitions and Descriptive Statistics of Identified Variables that are Instrumenting RHS endogenous variables (drinkers' categories)[a/] used in Two Labor Market Outcome Equations

(Weighted mean)

		Labor Market Outcomes[b/]			
Identified Instruments	Definition	LFPFULL N=43093 %	SD	INCOME N=21406 %	SD
Beer Tax	State level Excise Tax on Beer (Cents per Ounce)	19.0	9.34	18.0	6.54
Squared Beer Tax	$(\text{Beer tax})^2$	5.0	6.74	5.0	4.65
Numbers of Children	Numbers of Children	74.0	77.81	84.0	55.83
Squared Numbers of Children	$(\text{Numbers of Children})^2$	180.0	299.89	197.0	203.66
Parent problem drinker	=1 if respondent's Biological Mother/ Father was problem drinker	22.0	28.71	25.0	21.55
Tried to stop drinking	=1 if respondents tried to cut down or stop drinking	21.0	28.46	22.0	20.69
Spouse problem drinker	=1 if respondents reported married or lived with problem drinker	10.0	21.17	10.0	14.95
Current Smoker	=1 if respondents were smoker in last year	28.0	31.07	31.0	22.94
Ex-smoker	=1 if respondents were smoker prior to last year	0.19	27.35	16.0	18.41

Notes:

a/ See Table 2A or 2B for definition of drinkers' categories.

b/ See Table 1 for definition of labor market outcomes.

Table 6A: Results of Test of Endogeneity: the Durbin-Wu-Hausman (DWH) Test

Independent variables [c/]: Residuals saved from separate regression of each potential RHS endogenous variables (different drinker categories[a/])	Auxiliary Regression of Labor Market Outcomes	
	Maximum Likelihood Estimates[b/] of LFPFULL	Least Square Estimates[b/] of Log of INCOME
	Estimated Parameter (Wald-Chi Square value)[c/]	Estimated Parameter (t-value)[d/]
Residual_Ex-drinkers	0.49***(3.47)	-0.23* (-4.12)
Residual_Ex-abuser/dependent	0.20***(3.38)	-0.07* (-2.80)
Residual_New-drinkers	0.66* (13.57)	-0.001 (-0.02)
Residual_ Ex-New drinkers	0.48* (9.28)	-0.14* (-3.32)
Residual_Alcohol Abuser	-0.41 (1.84)	0.05 (0.75)
Residual_Alcohol dependent	0.39* (7.43)	0.05 (1.28)
Test of Endogeneity (The Durbin-Wu-Hausman Test)	25.34* [e/] (p-value: 0.003)	8.04*[f/] (p-value: <0.0001)
Observations:	43093	21406

Notes: '*' and '**' indicate statistical significance at 1% and 5% level.

[a/] See Table 2A or 2B for detail definition.

[b/] The estimated coefficients of all other exogenous variables are not reported here since the key interest is to test the statistical significance of residuals in the auxiliary regression, and also to obtain the Durbin-Wu-Hausman test statistic.

[c/] The critical value of Wald chi-square (with 1 degrees of freedom for excluding individual coefficient) is 6.63 and 2.70 at 1% and 10% level respectively.

[d/] The critical t-value (with N-K=21363 degrees of freedom, where N=Observations, K=# of estimated parameters including intercept) is 2.58 (approximate) at 1% level.

[e/] The critical value of Wald chi-square (with 6 degrees of freedom for excluding 6 residuals variables) is 16.81 at p-value of 0.01.

[f/] The critical F-value (with degrees of freedom in numerator=6 and in denominator=21364) at 1% level is 2.80 (approximate).

Table 6B: Results of Test of Heteroscedasticity (the Pagan-Hall test), and Results of Overidentifying Restrictions (the Hansen-J test)

	Tests results obtained from the IV regressions of two Labor Market Outcomes	
Tests	LFPFULL	Log of INCOME
Test of Heteroscedasticity :		
The Pagan-Hall's Test Statistic[a/]	37142*	21407*
Test of Overidentifying Restrictions:		
The Hansen-J Test Statistic[b/]		
Before Heteroscedasticity correction:	5.37 (p-value=0.15)	5.11(p-value= 0.16)
After Heteroscedasticity correction:	5.41 (p-value=0.14)	5.16 (p-value=0.16)
Observations:	43093	21406

Notes: Figures in parentheses are the p-values. '*' indicates statistical significance at 1% level with given p-value.

[a/]. The critical value of chi-square (with p-value of 0.005 and 2 degrees of freedom) is 10.60

[b/] The critical chi-square value is 6.25 (with 3 degrees of freedom for 3 excluded restrictions) at p-value of 0.10.

Table 7A: First Stage Regression Result to Predict Six Drinker Categories [a/] that are used in the Second Stage GMM Estimation of LFPFULL Equation

First Stage Regression [a/] Results of RHS Endogenous Variables (Drinkers Categories) by GMM-IV												
	Ex-drinkers		Ex-abuser/ dependent		New-drinkers		Ex-New drinkers		Alcohol Abuser		Alcohol dependent	
Instruments	Coefficient	Std err	Coefficient	Std err	Coefficient	Std err	Coefficient	Std err	Coefficient	Std err	Coefficient	Std err
Beer Tax	-0.065**	0.031	0.068*	0.021	-0.040*	0.013	-0.270*	0.045	-0.043**	0.019	-0.029***	0.016
Squared Beer Tax	0.085**	0.040	-0.070*	0.028	0.053*	0.018	0.139**	0.061	0.029	0.024	0.039***	0.022
Numbers of Children	-0.005	0.003	0.002	0.002	0.003***	0.002	0.016*	0.006	-0.001	0.002	-0.010*	0.002
Squared Numbers of Children	0.002**	0.001	-0.001	0.000	-0.001*	0.000	-0.003***	0.001	0.000	0.000	0.002*	0.000
Parent problem drinker	-0.045*	0.004	0.126*	0.004	-0.019*	0.001	-0.057*	0.006	0.032*	0.003	0.114*	0.004
Tried to stop drinking	-0.008*	0.003	0.026*	0.003	-0.003***	0.002	0.004	0.006	0.011*	0.003	0.014*	0.003
Spouse problem drinker	-0.002	0.004	0.046*	0.004	-0.004*	0.001	-0.002	0.007	0.004	0.003	0.008*	0.003
Current Smoker	-0.040*	0.004	0.010*	0.003	-0.011*	0.002	0.138*	0.006	0.033*	0.003	0.029*	0.003
Ex-smoker	0.011**	0.005	0.054*	0.003	-0.006*	0.001	0.127*	0.006	0.006**	0.003	-0.011*	0.002
Observations	**43093**		**43093**		**43093**		**43093**		**43093**		**43093**	
Test for IV Strength: Instrument Relevance [b/]	436.9*(<.0001)		1635.4*(<.0001)		379.3*(<.0001)		908.6*(<.0001)		377.94*(<.0001)		1296.0*(<.0001)	
Shea's Partial-R^2 [c/]	0.004		0.164		0.003		0.02		0.001		0.014	
Standard Partial-R^2	0.001		0.0002		0.002		0.004		0.002		2.102749E-7	

Note: '*' '**' and '***'indicate statistical significance level at 1%, 5% and 10% respectively. Figures in parentheses are p-values.

a/ Each drinker category is regressed on nine instruments and other exogenous variables (different education levels, marital status, gender, age groups, health status, race, other income, and industry types). The coefficients of other variables are not reported here.

b/ The critical value of Wald Chi-square (with 9 degrees of freedom for excluding 9 instruments) is 23.59 at p-value =0.005

c/ Following the procedures suggested by Godfrey (1999) and Baum et al (2003): see footnote 47 (in Chapter 3, section 3.3.3).

Table 7B: First Stage Regression Result to Predict Six Drinker Categories [a/] that are used in the Second Stage GMM Estimation of log of INCOME (LINCOME) Equation

First Stage Regression Results of RHS Endogenous Variables (Drinkers Categories) [a/] by GMM-IV												
	Ex-drinkers		Ex-abuser/ dependent		New-drinkers		Ex-New drinkers		Alcohol Abuser		Alcohol dependent	
Instruments[b/]	Coefficient	Std err	Coefficient	Std err	Coefficient	Std err	Coefficient	Std err	Coefficient	Std err	Coefficient	Std err
Beer Tax	0.004	0.041	0.102*	0.029	-0.018	0.018	-0.313*	0.067	-0.104*	0.030	-0.026	0.025
Squared Beer Tax	0.009	0.057	-0.108*	0.039	0.016	0.025	0.159***	0.092	0.111*	0.039	0.042	0.034
Numbers of Children	-0.005	0.004	0.001	0.003	0.004***	0.002	0.026*	0.007	-0.0001	0.003	-0.011*	0.003
Squared Numbers of Children	0.001	0.001	0.000	0.001	-0.001***	0.001	-0.006*	0.002	-0.0001	0.001	0.002**	0.001
Parent problem drinker	-0.034*	0.004	0.102*	0.005	-0.015*	0.002	-0.083*	0.008	0.042*	0.005	0.111*	0.005
Tried to stop drinking	-0.007	0.005	0.027*	0.004	-0.003	0.002	-0.004	0.008	0.010*	0.004	0.014*	0.004
Spouse problem drinker	-0.009	0.006	0.050*	0.006	-0.003*	0.002	-0.005	0.011	0.004	0.005	0.010**	0.005
Current Smoker	-0.049*	0.004	0.004	0.003	-0.009*	0.002	0.111*	0.008	0.037*	0.004	0.029*	0.004
Ex-smoker	-0.010***	0.006	0.050*	0.005	-0.007	0.002	0.096*	0.009	0.006	0.004	-0.010*	0.003
Observations	**21406**		**21406**		**21406**		**21406**		**21406**		**21406**	
Tests for IV Strength: Instrument Relevance[b/]	265.20* (<.0001)		700.60* (<.0001)		171.04* (<.0001)		383.97* (<.0001)		233.47* (<.0001)		700.39* (<.0001)	
Shea's Partial-R^2 [c/]		0.0007		0.016		0.0002		0.001		9.21E-05		0.0003
Standard Partial-R^2		0.004		0.0002		0.014		0.018		0.009		0.00007

Notes: '*' '**' and '***'indicate statistical significance level at 1%, 5% and 10% respectively. Figures in parentheses are p-values.

a/ Each drinker category is regressed on nine instruments and other exogenous variables (different education levels, marital status, gender, age groups, health status, race, other income, and industry types). The coefficients of other variables are not reported here.

b/ The critical value of Wald Chi-square (with 9 degrees of freedom for excluding 9 instruments) is 23.59 at p-value =0.005.

c/ Following the procedures suggested by Godfrey (1999) and Baum et al (2003): see footnote 47 (in Chapter 3, section 3.3.3).

Table 8: Estimated Results of Full Time Labor Force Participation (LFPFULL)

	Baseline MLE Estimates (Weighted)			GMM-IV Estimates (Heteroscedasticity Corrected)		
Variable (Base category)	Parameter	Std. Error	Marginal Effect a/	Parameter	Approx Std Err	Marginal Effect a/
Intercept	11.05	1.31	0.47	0.918*	0.35	0.21
Drinker category **<u>(Life-time abstainers):</u>**						
Ex-drinkers	0.003	0.001	0.00	-2.360**	1.35	-0.43
Ex-abuser/dependent	-0.19*	0.001	-0.05	0.461	0.30	0.11
New-drinkers	-0.16*	0.001	-0.04	4.063	2.62	0.45
Ex-New drinkers	0.15*	0.001	0.04	0.098	0.42	0.02
Alcohol Abusers	0.30*	0.001	0.07	-1.885	3.28	-0.38
Alcohol dependent	-0.23*	0.001	-0.06	-0.171	1.03	-0.04
<u>Gender (Female):</u>						
Male	0.92*	0.0005	0.21	0.016	0.04	0.00
Age-group (age 60~98)						
AGE18~29	2.08*	0.001	0.37	-0.309	0.29	-0.08
AGE30~59	2.37*	0.001	0.39	0.078	0.09	0.02
Location (not in central city):						
Central	0.03*	0.001	0.01	-0.002	0.01	0.00
Race (other races):						
White	0.04*	0.001	0.01	-0.001	0.04	0.00
Black	0.30*	0.001	0.07	0.131*	0.05	0.03
Hispanic-Latino	0.32*	0.001	0.08	-0.035	0.07	-0.01
Marital status (other):						
Married	0.95*	0.0004	0.21	0.174*	0.05	0.04
Other Income (none):						
Had other source of income	-18.61*	1.307	-0.53	-0.562*	0.06	-0.14
Education level (below 9th grade):			`			
Post-graduate	0.55*	0.001	0.13	-0.072	0.15	-0.02
Graduate	0.44*	0.001	0.11	0.011	0.08	0.00
Technical	0.35*	0.001	0.09	0.015	0.06	0.00
High School	0.26*	0.001	0.06	-0.042	0.06	-0.01
Some High School	-0.12*	0.001	-0.03	-0.033	0.04	-0.01

(Continue in next page………..)

(Table 8 continues……….)

Variable (Base category)	Baseline MLE Estimates (Weighted) Parameter	Std. Error	Marginal Effect a/	GMM-IV Estimates (Heteroscedasticity Corrected) Parameter	Approx Std Err	Marginal Effect a/
Health Status (Poor)						
Excellent Health status	1.15*	0.001	0.25	-0.027	0.10	-0.01
Good Health status	0.92*	0.001	0.21	-0.009	0.07	0.00
Industry type (Armed Services or other)						
Agriculture	3.06*	0.002	0.43	0.228*	0.04	0.06
Mining	3.12*	0.003	0.43	0.268	0.18	0.07
Construction	3.26*	0.001	0.44	0.288*	0.03	0.07
Manufacturing	3.76*	0.001	0.45	0.332*	0.03	0.08
Transportation	3.49*	0.001	0.44	0.279*	0.05	0.07
Wholesale	3.46*	0.002	0.44	0.275*	0.07	0.07
Retail Trade	2.96*	0.001	0.43	0.152*	0.08	0.04
Finance/Insurance	3.73*	0.001	0.45	0.309*	0.06	0.08
Business	3.40*	0.001	0.44	0.242*	0.08	0.06
Personal service	2.95*	0.001	0.43	0.227*	0.05	0.06
Entertainment	2.83*	0.002	0.42	0.134*	0.08	0.03
Professional	3.34*	0.001	0.44	0.289*	0.04	0.07
Public Administration	3.69*	0.001	0.45	0.329*	0.06	0.08

Notes: '*', and '**' indicates statistical significance level at 1% and 5% respectively.

a/ Since the dependent variable (LFPFULL) and explanatory variables all are binary variables, the marginal effect is interpreted as the percentage point change in the probability of full time labor force participation resulting from a discrete change in binary explanatory variable.

Table 9: Estimated Results of Income Equation

Variable (Base category)	Dependent Variable: LINCOME (log of Annual Income) Baseline OLS Estimates (Weighted) Parameter	Std Err	Marginal Effect [a]/	GMM-IV Estimates (Heteroscedasticity Corrected) Parameter	Std Err	Marginal Effect [a]/
Intercept	9.53	0.04	-	8.00	0.75	-
Drinker category (Life-time abstainers):						
Ex-drinkers	0.01	0.02	411.47	3.42**	1.71	134546.82
Ex-abuser/dependent	0.04**	0.02	1538.08	0.41	0.48	16037.26
New-drinkers	-0.01	0.03	-583.76	5.48	6.37	215675.09
Ex-New drinkers	0.12*	0.01	4810.52	1.75***	1.04	68715.21
Alcohol Abusers	0.13*	0.02	5094.14	-2.85	5.75	-112056.95
Alcohol dependent	0.06**	0.02	2306.33	3.79	3.54	148976.58
Gender (Female):						
Male	0.27*	0.01	9558.50	0.41*	0.16	16173.80
Age-group (age 60~98)						
AGE18~29	-0.37*	0.02	-14482.70	-0.50	0.42	-19529.64
AGE30~59	-0.03	0.02	-1064.07	0.05	0.07	1983.73
Location (not in central city):						
Central	-0.01	0.01	439.77	0.01	0.03	321.03
Race (other races):						
White	0.07*	0.02	2795.68	-0.03	0.06	-1071.93
Black	-0.01	0.02	-510.20	-0.19**	0.09	-7625.08
Hispanic-Latino	-0.09*	0.02	-3489.59	-0.14	0.10	-5355.73
Marital status (other):						
Married	0.17*	0.01	6670.77	0.11	0.10	4369.28
Other Income (none):						
Had other source of income	-0.14*	0.01	-5862.79	-0.21***	0.12	-8136.47
Education level (below 9th grade):						
Post-graduate	0.97*	0.02	38339.81	1.01*	0.12	39879.18
Graduate	0.78*	0.02	30704.89	0.85*	0.10	33499.82
Technical	0.53*	0.02	20797.87	0.58*	0.09	22926.94
High School	0.38*	0.02	15123.11	0.42*	0.07	16659.22
Some High School	0.16*	0.02	6271.50	0.04	0.10	1424.71

(Table 9 Continue in next page……….)

(Table 9 continues……….)

Variable (Base category)	Baseline OLS Estimates			GMM-IV Estimates (Heteroscedasticity Corrected)		
	Parameter	Std Err	Marginal Effect [a]/	Parameter	Std Err	Marginal Effect [a]/
Health Status (Poor)						
Excellent Health status	0.14*	0.01	5518.19	0.19*	0.05	7284.11
Good Health status	0.05*	0.02	2056.93	0.07	0.05	2646.55
Industry type (Armed Services or other)						
Agriculture	-0.15*	0.04	-6044.13	-0.03	0.26	-1246.59
Mining	0.21*	0.06	8164.39	0.73	0.68	28858.33
Construction	0.004	0.03	-154.59	0.20	0.30	7826.73
Manufacturing	0.07**	0.03	2766.18	0.25	0.33	9857.46
Transportation	0.14*	0.03	5350.23	0.28	0.28	11035.52
Wholesale	0.04	0.04	1637.99	0.17	0.29	6754.16
Retail Trade	-0.12*	0.03	-4833.73	-0.03	0.23	-1168.70
Finance/Insurance	0.16*	0.03	6443.79	0.27	0.22	10497.43
Business	0.03	0.03	1248.56	0.20	0.26	7777.28
Personal service	-0.14*	0.03	-5552.02	0.01	0.28	284.68
Entertainment	-0.05	0.04	-2142.69	0.08	0.26	2983.08
Professional	0.04	0.03	1482.22	0.17	0.28	6868.12
Public Administration	0.08	0.03	3036.03	0.23	0.27	8991.14

Note: '*', '**' and '***' indicate statistical significance level at 1% and 5% respectively.
a/ Marginal effects are evaluated at dependent mean.

Annex 1: Alcohol Use Measures

Drinking measures	Definition
Level/stages of drinking	1. Heavy drinking' for consumption of at least 100 alcoholic beverages per month, 2. 'Light drinking' for 1–10 or 1–20 drinks per month and 3. Moderate drinking for the consumption of 21–59 alcohol beverages per month. (Ruhm and Black: 2002).
Moderate Drinking	1. Moderate drinking, particularly red wine appears to offer health benefits. Moderate drinking is defined as equal to or less than two drinks a day for men and equal to or less than one drink a day for women. @ 2. Moderate drinkers use approximately 2 or 2.5 drinks per day on average. French et al (1995)
Heavy drinking @	Hazardous drinking puts people at risk for adverse health events. People who are heavy drinkers consume the following: -More than 14 drinks per week or four to five drinks at one sitting, for men. -More than seven drinks per week or three drinks at one sitting, for women. -Frequent intoxication.
Binge drinking@	Binge drinking is seen as drinking that occurs at a hazardous level- five or more drinks for men, and three or more drinks for women. It is defined as at least three times the rate of drinking that would keep an individual within a 0.05 blood alcohol content. Two drinks for the first hour and one drink thereafter for males, and one drink per hour for females, would roughly maintain a blood alcohol content of 0.05.
Alcohol abuse and dependence @ (See Clinical diagnostic Criteria in Annex 2).	American Psychiatric Association (APA),1994, considered alcohol abuse and dependence as two severe alcohol-use disorders.

Source: @ http://www.healthandage.com/html/well_connected/pdf/doc56.pdf

Annex 2: Diagnostic Criteria for Substance (includes Alcohol) Abuse and Dependence

Abuse	Dependence
(A) A maladaptive pattern of substance use leading to clinically significant impairment or distress, as manifested by *one* (or more) of the following, occurring within a 12-month-period (1) Recurrent substance use resulting in a failure to fulfill major role obligations at work, school, or home (e.g. repeated absences or poor work performance related to substance use; substance-related absences, suspensions, or expulsions from school; neglect of children or household) (2) Recurrent substance use in situations in which it is physically hazardous (e.g. driving an automobile or operating a machine when impaired by substance use) (3) Recurrent substance-related legal problems (e.g. arrests for substance-related disorderly conduct) (4) Continued substance use despite having persistent or recurrent social or interpersonal problems caused or exacerbated by the effects of the substance (e.g. arguments with spouse about consequences of intoxication, physical fights) (B) The symptoms have never met the criteria for Substance Dependence for (the particular substance of concern, e.g. alcohol)	A maladaptive pattern of substance use, leading to clinically significant impairment or distress, as manifested by *three* (or more) of the following, occurring at any time in the same 12-month-period (1) Tolerance, as defined by either of the following: (a) A need for markedly increased amounts of the substance to achieve intoxication or desired effect, (b) Markedly diminished effect with continued use of the same amount of the substance. (2) Withdrawal, as manifested by either of the following: (a) The characteristic withdrawal syndrome for the substance (defined elsewhere in DSM-IV) (b) The same (or a closely related) substance is taken to relieve or avoid withdrawal symptoms. (3) The substance is often taken in larger amounts or over a longer period than was intended. (4) There is a persistent desire or unsuccessful efforts to cut down or control substance use. (5) A great deal of time is spent in activities necessary to obtain the substance, use the substance, or recover from its effects. (6) Important social, occupational, or recreational activities are given up or reduced because of substance use. (7) The substance use is continued despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by the substance (continued drinking despite recognition that an ulcer was made worse by alcohol consumption)

Source: APA (1994).

i) A diagnosis of *current* alcohol abuse requires meeting at least one of the four abuse criteria in the past 12 months (and never having met the criteria for alcohol dependence).

ii) A diagnosis of *current* alcohol dependence requires meeting at least three of the seven dependence criteria in the past 12 months.

REFERENCES

Amemiya, T., 1985. Advanced Econometrics. Cambridge, MA: Harvard University Press.

American Psychiatric Association, 1994. Diagnostic and statistical manual of mental disorders, Fourth Edition. Washington, DC: American Psychiatric Association.

Alcohol Epidemiology Program, 2000. Alcohol policies in the United States: highlights from the 50 States. University of Minnesota, School of Public Health.

Apgar, K.R., Burgess, A.G., 2001. Alcohol use to alcohol abuse: the economic burden, health consequences and impact on the workplace. The Washington Business Group.

Auld, M.C., 2005. Smoking, drinking, and income. The Journal of Human Resources XL 2, pp. 505-517.

Baker, D.E., 1985. The study of stress at work. Annual Review of Public Health 6, pp. 367-381.

Barrett, G.F., 2002. The effect of alcohol consumption on earnings. Economic Record, Vol. 78, pp. 79-96.

Baum-Baicker, C.,1985. The psychological benefits of moderate alcohol consumption: A review of the literature. Drug and Alcohol Dependence 15, pp. 305-322.

Baum, C.F., Schaffer, M.E., Stillman, S., 2003. Instrumental variables and GMM: estimation and testing. Working Paper No. 545, Boston, MA: Boston College.

Beaglehole, R., Jackson, R., 1992. Alcohol, cardiovascular diseases and all causes of mortality: A review of the epidemiological evidence. Drug and Alcohol Review II (1992), pp. 275-290.

Becker, G., 1965. A theory of the allocation of time. Economic Journal 75, pp. 493-517.

Becker, G. S., 1964., Human Capital*.* Chicago: University of Chicago Press,

Becker, G.S., Murphy K.M., 1988. A theory of rational addiction. Journal of Political Economy 96:675-700.

Benham, L., Benham, A., 1982. Employment, earnings, and psychiatric diagnosis, Economics Aspects of Health, Fuchs V. (Ed.) University of Chicago Press, Chicago.

Berger, M., Leigh, J.,1988. The Effect of alcohol use on wages. Applied Economics, 20, pp. 1343-1351.

Berry Jr., R.E., Boland, J.P., and Smart, C.N., 1977. The economic cost of alcohol abuse-1975. Report prepared for the National Institute on Alcohol Abuse and Alcoholism, U.S. Department of Health, Education, and Welfare. Brookline, MA: Policy Analysis, Inc., 1977.

Bound, J., Jaegar, D. A., and Baker, R. M., 1995. "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak." Journal of the American Statistical Association 90 (June 1995), pp. 443-50.

Bound, J., Waidmann, T.A., Schoenbaum, M., Bingenheimer, J., 2003. The labor market consequences of race differences in health*.* Milbank Quarterly, 81(3), pp. 441-473.

Bray, J.W., Zarkin G.A., Dennis, M.L., French, M.T. 2000. Symptoms of dependence, multiple substance use, and labor market outcomes. The American Journal of Drug and Alcohol Abuse, 2000 Vol. 26 No. 1, pp. 77-95.

Breusch, T. S., and Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. Econometrica, 47, 1287-1293.

Bryant, R., Samaranayake, V., Wilhite, A., 1992. Alcohol use and the wages of young men: whites vs. non-whites. International Review of Applied Economics 1*, 6*(2), pp. 184-202.

Bryant, R., Samaranayake, V., Wilhite, A.,1993. The influence of current and past alcohol use on earnings: three approaches to estimation. Journal of Applied Behavioral Science, 29, pp. 9–31.

Bryant, R., Samaranayake, V., Wilhite, A.,1996. The Impact of alcohol and drug use on employment: A Labor Market Study Using the National Longitudinal Survey of Youth. Institute for Research on Poverty**,** Discussion Paper no. 1092-1096.

Buchmueller, T. C., Zuvekas., S. H., 1998. "Drug use, drug abuse, and labour market outcomes." Health Economics 7: 229-245.

Burgess, S.M., Propper, C., 1998. Early health-related behaviours and their impact on later life chances: evidence from the US. Health Economics 7**,** pp. 381-399.

Chatterji, P., 1998. Does use of alcohol and illicit drugs during adolescence affect educational attainment?: results from three estimation methods," Johns Hopkins University School of Hygiene and Public Health Working paper.

Chatterji P., Markowitz S., 2001. The impact of maternal alcohol and illicit drug use on children's behavior problems: evidence from the children of the National Longitudinal Survey of Youth, Journal of Health Economics 12, pp. 411-429.

Chatterji, P., Alegria, M., Lu, M., Takeuchi, D., 2007. Psychiatric disorders and labor market outcomes: evidence from the national Latino and Asian American Study. Health Economics. (in press, Accepted 2006), John and Wiley Sons Ltd. 2007.

Coate, D.,1993. Moderate drinking and coronary heart disease mortality: evidence from NHANES I and the NHANES I follow-up. American Journal of Public Health 83 (1993), pp. 888-890.

Cook, P.J., 1991. The social costs of drinking, in expert meeting on negative social consequences of alcohol use, Oslo: Norwegian Ministry of Health and Social Affairs.

Cook P.J., Moore, M.J., 1993. Drinking and schooling. Journal of Health Economics 12 (1993), pp. 411-429.

Cruze, A., Harwood, H., Kristiansen, P., Collins, J., Jones, D., 1981. Economic costs to society of alcohol and drug abuse and mental illness, 1977. Report prepared for the Alcohol, Drug Abuse, and Mental Health Administration, U.S. Department of Health and Human Services. DHHS Pub. No. (ADM)81-1179. Rockville, MD: Alcohol, Drug Abuse, and Mental Health Administration, 1981

Culyer A, 1973. Should social policy concern itself with drug abuse? Public Finance Quarterly 1, pp. 449-456.

Currie, J., Madrian, B.C., 1999. Health, health insurance and the labor market, Handbook of Labor Economics, Vol. 3c (edited by Ashenfelter, O., Card, D.), pp. 3309-3416.

Dave, D., and Kaestner, R., 2002. Alcohol taxes and labor market outcomes, Journal of Health Economics 21, pp. 357-371.

Davidson, R., Mackinnon, J.G., 1993. Estimation and inference in econometrics, New York: Oxford University Press.

Dee, T.S., 2001. Alcohol abuse and economic conditions: evidence from repeated cross sections of individual-level data. Health Economics 103 (2001), pp. 257-270.

Dee, T.S., Evans, W.N., 2003. Teen drinking and educational attainment: evidence from Two-Sample Instrumental Variables (TSIV) estimates. *Journal of Labor Economics* 21(1), January 2003, pp. 178-209.

DeSimone, J., 2002. Illegal drug use and employment, Journal of Labor Economics, University of Chaicago Press, Vol 20(4), pp. 952-977.

DeSimone, J., Wolaver A.M., 2005. Drinking and academic performance in high school. NBER Working Papers 11035, National Bureau of Economic Research, Inc.

Doll R, Peto, E., Hall, W. K., and Gray, R., 1994. Mortality in relationship to consumption of alcohol: 13 years' observations on male British doctors. British Medical Journal 309, pp. 911-918.

Dollar, D., Kraay A., 2003. Institutions, trade and growth, Journal of Monetary Economics, 50 (1), pp. 133-162.

Ettner, S.L.,1996. New evidence on the relationship between income and health. Journal of Health Economics 15(1), pp. 67-85.

Ettner, S.L.,1997. Measuring the human cost of a weak economy: does unemployment lead to alcohol abuse? Social Science Medicine 44 2 (1997), pp. 251-260.

Ezzati, M., A.D. Lopez, A., Rodgers, S. Vanderhoorn, and. Murray, C.J.L., 2002. Selected major risk factors and global and regional burden of disease", The Lancet, Vol. 360, pp.1347-1360.

Farahati, F., Marcotte, D.E., Wilcox-Gök, V., 2003. Gender differences of psychiatric disorders: effects on labor market outcomes. Working paper Centers for Mental Healthcare and Outcomes Research, University of Arkansas for Medical Sciences.

Farrell *S.,* Manning W., and Finch M., 2003. Alcohol dependence and the price of alcoholic beverages. Journal of Health Economics 22, pp. 117-147.

Feng, W., Zhou, W., Butler, J., Booth, B., and French, M., 2001. The impact of problem drinking on employment. Health Economics 10(6), 509-21.

Fenwick, R., and Tausig M.,1994. The macroeconomic context of job stress. Journal of Health and Social Behavior, 3S3, pp. 266-282.

Fisher, I.,1927. Prohibition at its worst, Alcohol Information Committee. New York.

Folland, S., Goodman, A.C., Stano, M., 2001. The economics of health and health Care, Third edition, Prentice-Hall, New Jersey.

Foster, E.M., 1997. Instrumental variables for logistic regression: an illustration. Social Science Research 26, pp. 487-504.

French, M.T., Zarkin, G.A., Bray, J.W., Hartwell, T.D., 1994. Costs of Employee Assistance Programs: Findings from a National Survey. Research Triangle Institute. Research Triangle Park, NC.

French, M., and Zarkin, G., 1995. Is moderate alcohol use related to wages? Evidence from four worksites. Journal of Health Economics,14(3), pp. 319-344.

French, M.T., Zarkin, G.A., & Dunlap, J.L., 1998. Illicit drug use, absenteeism, and earnings at six U.S. worksites. Contemporary Economic Policy 16(3), 334-346.

Frone, M., R., 2006. Prevalence and distribution of alcohol use and impairment in the workplace: a U.S. National Survey. Journal of Studies of Alcohol 76, pp.147-156

Fuchs, C. S., Stampfer, M. J., Colditz, G. A., 1995. Alcohol consumption and mortality among women. New England Journal of Medicine 332, pp. 1245–1250.

Fuchs, V.R., 2004. The future of health economics. Journal of Health Economics 19(2), pp.141–157.

Fuchs, V.R., 2004. Reflections on the socio-economic correlates of health, Journal of Health Economics 23, pp. 653-661.

Godfrey, L.G., 1978. Testing for multiplicative heteroscedasticity, Journal of Econometrics 8, pp. 227-236.

Godfrey, L.G., 1999. Instrument relevance in multivariate linear models, Review of Economics and Statistics 81(3), pp. 550-552.

Gill, A., Michaels, R., 1992. Does drug use lower wages? Industrial and Labor Relations Review 45, pp. 419-434.

Gujarati, Damodar N. *Basic Econometrics*. New York: McGraw-Hill, 1995.

Greene, W. H., 2000: Econometric Analysis; Prentice Hall, (4th edition).

Gronbaek, M., Deis, A, Sorensen, T.I.A., Becker, U., Schnohr, P., and Jensen G., 1995. Mortality associated with moderate intakes of wine, beer and spirits. British Medical Journal 310, pp.1165-1169.

Grossman, M., 1972. The demand for health: A theoretical and empirical investigation. Columbia University Press for the National Bureau of Economic Research, New York.

Grossman, M., 1991.The demand for health: A theoretical and empirical investigation, The National Bureau of Economic Research, Reproduced in "The Economics of Health", edited by Culyer, AJ. (1991), Vol.1.

Hamilton, V and Hamilton, B., 1997. Alcohol and earnings: does drinking yield a wage premium? Canadian Journal of Economics, 30, pp. 135-151.

Hahn, J., Hausman, J., 2002. Notes on bias in estimators for simultaneous equation models. Economics Letters 75 (2), pp. 237-41.

Hansen, L. P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, pp.1029-1084.

Harwood, H.J., Napolitano, D.M., Kristiansen, P.L., and Collins, J.J., 1984. Economic Costs to Society of Alcohol and Drug Abuse and Mental Illness: 1980. Report prepared for the Alcohol, Drug Abuse, and Mental Health Administration, U.S. Department of Health and Human Services.

Harwood, H., Fountain, D., Livennore, G., 1998. The economic costs of alcohol and drug abuse in the United States 1992. NIH Publication No. 98-4327, The Lewin Group, Fairfax ,V A.

Harwood, H., 2000 Updating Estimates of the Economic Costs of Alcohol Abuse in the United States: Estimates, Update Methods and Data. Report prepared by the Lewin Group for the National Institute on Alcohol Abuse and Alcoholism, 2000.

Hausman, J., 1978. Specification tests in econometrics, Econometrica 46, pp.1251-1271.

Heather N, 2001. Pleasures and pains of our favourite drug in Heather N, Peters T and Stockwell T (eds), International Handbook of Alcohol Dependence and Problems , John Wiley and Sons, Chichester, p 6.

Heien, D., 1996. Do drinkers earn less? Southern Economic Journal, 63, pp. 60-60.

Horgan, C., Skwara, K.C., Strickler G., 2001. Substance abuse: The nation's number one health problem, Schneider Institute for Health Policy, Brandeis University.

Hutcheson, G.D., Hemderson, M.M., and Davies, J. B., 1995. Alcohol in the workplace: costs and responses, Department for Education and Employment Research Series No 59, London.

Johansson, E., Alho, H., Kiiskinen, U., and Poikolainen, K. (2004). Abstaining from alcohol and labor market underperformance – have we forgotten the "dry" alcoholics? The Research Institute of the Finnish Economy, Discussion Papers, ISSN 0781-6847; No. 931.

Karasek, R. A., Theorell, T., 1990. Healthy work: stress, productivity, and the reconstruction of working life. Basic Books, New York.

Kaestner, R., 1991. The effect of illicit drug use on the wages of young adults. Journal of Labor Economics, 9(4): 381–412.

Kaestner, R., 1994. The effect of illicit drug Use on the labour supply of young adults. Journal of Human Resources 29**,** pp. 126-155.

Kenkel, D. and Ribar, D. 1994. Alcohol consumption and young adults' socio-economic status. Brookings Papers on Economic Activity, June 1994 (Micro), pp. 119-161.

Marcotte, D.E., Wilcox-Gök, V., and Redmon, D.P., 2000. The labor market effects of mental illness: The case of affective disorders. In the Economics of Disability, D. Salkever & A. Sorkin, eds. (Greenwich, CT: JAI Press), 181-210.

McDonald, Z. and Pudney, S., 2000. Illicit drug use, unemployment, and occupational attainment. Journal of Health Economics 19, pp. 1089-1115.

McDonald, Z., and Pudney, S., Shields, M., 2001. The impact of alcohol consumption on occupational attainment in England. Economica 68, pp. 427-454.

MacDonald, Z and Shileds, M. A., 2004. Does problem drinking affect employment? Evidence from England, Health Economics.

MacDonald, Z and Shileds, M. A., 2004. Does drinking affect employment? Evidence from England. Health Economics, 13, pp. 139-155.

Mincer, J., 1974 Schooling, experience, and earnings. Columbia University Press.

Marmot M., and Brunner E., 1991. Alcohol and cardiovascular disease: The status of the U-shaped curve. British Medical Journal 303, pp. 565-568.

Mirowsky, J., Ross, *C.B.,* Reynolds, J., 2000. Links between social status and health status, in Handbook of Medical Sociology (ed. By Bird, C.E., Conrad, P., and Fremont, A.M.), Prentice-Hall Inc., New Jersey, USA.

Mokdad, A.H., Marks, J.S., Stroup, D.F. Gerberding, J.L., 2004. Actual causes of death in the United States, 2000, Journal of the American Medical Association 2004:291(10), pp. 1238-1245.

Murray, M.P., 2005. Econometrics: A Modern Introduction, Publisher: Prentice Hall

Mullahy, J., Sindelar, J.L., 1989. Lifecycle effects of alcoholism on education, earnings, and occupation. Inquiry 26 (Summer 1989), pp. 272-282.

Mullahy, J., Sindelar, J.L., 1991. Gender differences in labor market effects of alcoholism. American Economic Review 81, pp. 161-165.

Mullahy, J., Sindelar, J.L.,1993. Alcoholism, work, and income. Journal of Labor Economics 11 (1993), pp. 494-520.

Mullahy, J., Sindelar, J.L.,1993. Alcohol and the labor market. In: M. Hilton and G.

Bloss, Editors, Economic and socioeconomic issues in the prevention of alcohol-related problems, National Institute on Alcohol Abuse and Alcoholism, Washington, DC.

Mullahy, J., Sindelar, J.L.,1994. Alcoholism and income: The role of indirect effects. The Milbank Quarterly 72, pp. 359-375.

Mullahy, J., Sindelar, J.L., 1995. Employment, unemployment, and problem drinking. The NBER Working Paper No. 5123.

Mullahy, J., Sindelar, J.L.,1996. Employment, unemployment, and problem drinking. Journal of Health Economics 15, pp. 409-434.

National Household Survey on Drug Abuse, Substance Abuse and Mental Health Services Administration (SAMHSA), 1999. Worker Drug Use and Workplace Policies and Programs.

National Institute on Alcohol Abuse and Alcoholism (NIAAA), 1993. Eight Special Report to Congress on Alcohol and Health, U.S. Department of Health and Human Services, Washington, DC.

National Institute on Alcohol Addiction and Abuse (NIAAA), 2000. In: Proceedings of the 10th Special Report to Congress on Alcohol and Health, NIH Publication No. 00-1583.

National Institute on Alcohol Abuse and Alcoholism (NIAAA), 2001. Alcohol Alert No. 51, Reports on NIAAA Strategic Plan, 2001-2005.

National Institute on Alcohol Addiction and Abuse (NIAAA) Newsletter, 2004. Spring/Summer No. 4, Department of Health and Human Service, National Institute of Health.

Organization for Economic Co-operation and Development (OECD) Health at a glance: OECD indicators 2003, Volume 434.

Pagan. A.R., Hall, D., 1983. Diagnostic tests as residual analysis. Econometric Reviews 2(2), pp. 159-281.

Peters, B. l., 2004. Is there a wage bonus from drinking? unobserved heterogeneity examined. Applied Economics, November vol. 36(20), pp 2299-2315,

Pittman, D. J., Heien, D. M., 1989 "The economic costs of alcohol abuse: an assessment of current methods and estimates." J. Stud. Alcohol, 1989, 50(4): pp.567-579.

Ramanathan, R., 1998. Introductory econometrics with applications, 4th edition. 1998.

Rice, D.P. , Kelman, S., Miller L.S., Dunmeyer, S., 1990. The economic costs of alcohol and drug abuse and mental illness: 1985. , Institute for Health and Aging, University of California, San Francisco, California.

Ruhm, C. J., and W. Black, W., 2002. Does drinking really decrease in bad times?, Journal of Health Economics 21, pp. 659-678.

Russo, D., Purohit, V., Foudin, L., Salin, M. 2004. Workshop on alcohol use and health disparities 2002: a call to arms, Alcohol 32, pp.37-43.

Sargan, J., 1958. The estimation of economic relationships using instrumental variables, Econometrica 26(3), pp. 393-415.

Schuckit, M.A., 2000. Drug and Alcohol Abuse: A clinical guide to diagnosis and treatment, Fifth Edition, New York. Kluwer Academic/Plenum Publisher.

Sesso, H.D., 2001. Alcohol and cardiovascular health: recent findings. American Journal of Cardiovascular Drugs.1(3), pp. 167-72.

Shaper, A., G., 1988. Alcohol and mortality in British men: explaining the U-Shaped Curve. Lancet (December 3), pp. 267-273.

Shaper, A., G., 1990. Alcohol and mortality: A review of prospective studies. British Journal of Addiction 85 (1990), pp. 837-847.

Shea, J. 1997. Instrument relevance in multivariate linear models: a simple measure. Review of Economics and Statistics, 79 (2). pp. 348-352.

Sokejima, S., Kagamimori, S., 1998. Working hours as a risk factor for acute myocardial infarction in Japan: a case-control study. The British Medical Journal 317 19 (1998), pp.775-780.

Staiger, D., Stock, J.,1997. Instrumental variables regression with weak instruments." Econometrica, Vol. 65(3), pp. 557-586.

Swift, R.M., 2001. The pharmacotherapy of alcohol dependence: clinical and economic aspects. TEN, 2001 3(12), pp. 62-66.

Terza, J. V., 2002. Alcohol abuse and employment: a second look. Journal of Applied Econometrics, 17, pp. 393-404.

Thun, M.J., Peto, R., Lopez, A.D., Monaco, J.H., Henley, S.J., Heath, C.W., and Doll , R., 1997. Alcohol consumption and mortality among middle aged and elderly US adults. New England Journal of Medicine 337 (1997), pp. 705-714.

Walitzer, K, Conners, G, 1999. Treating problem drinking, Alcohol Research & Health, 23: PP.138-143.

Wallentin, B., Agren, A., 2002. Test of heteroscedasticity in a regression model in the presence of measurement errors. Economics Letters, 76, pp. 205-211.

Ware, J. E., 1986. The assessment of health status, in Applications of social science to clinical medicine and health policy, edited by Aiken, L.H., and Mechanic, D., New Brunswick, NJ: Rutgers University Press.

Weiss, G.L., Lonnquist, L.E., 1997, The Sociology of Health, Healing and Illness, 2nd edition, Prentice Hall, New Jersey, USA.

White, H., 1980. A heteroskedasticity consistent covariance matrix and a direct test for heteroskedasticity. Econometrica 48, pp. 817-838.

Wooldridge, J.M., 2002. Econometric analysis of cross section and panel data. Cambridge, Massachusetts.

World Health Organization (WHO), 1986. Ottawa Charter for Health Promotion, Health Promotion, pp iii-v.

World Health Organization (WHO), 2002. The World Health Report 2002 - Reducing Risks, Promoting Healthy Life. Geneva, World Health Organization.

World Health Organization (WHO), 2004. Global status report on alcohol. Geneva, Switzerland: World Health Organization, Department of Mental Health and Substance Use.

Yamada, T., Kendix, M., Yamada, T. 1996. “The impact of alcohol consumption and marijuana use on high school graduation,” Health Economics, 5: 77-92

Zabel, J. E. 1993. “The relationship between hours of work and labor force participation in four models of labor supply behavior.” Journal of Labor Economics, 11(2): 387–416.

Zarkin, G., French, M., Mroz, T., Bray, J., 1998. Alcohol use and wages: new results from the National Household Survey on Drug Abuse. Journal of Health Economics 17, pp. 53-68.

ABSTRACT

ALCOHOL USE DISORDERS AND LABOR MARKET OUTCOMES: AN ANALYSIS USING 2001-02 NATIONAL EPIDEMIOLOGY SURVEY ON ALCOHOL AND RELATED CONDITIONS

by

SHAMMIMA JESMIN

May 2010

Advisor: Dr. Allen Goodman

Major: Economics

Degree: Doctor of Philosophy

BACKGROUND

Attempt of this research to explore the impact of alcohol use disorders on labor market outcomes is justified on the following grounds. First, alcohol use disorders (alcohol abuse and dependence) became more prevalent among the working age population over the decades (NIAAA: 2004). Second, existence of a large body of research on the nature and extent of relationship between alcohol use and labor market outcomes remained debatable, heterogeneous and counterintuitive with various explanations although it is generally agreed that harmful effects attributed by alcohol consumption may results in short-run and long-run physical and mental impairments, and it may entail enormous economic and non-economic costs to the society.

OBJECTIVE

The main focus of this research is to obtain the consistent estimates of the impact of alcohol use disorders on labor market outcomes. This also examines to what extent performance of individuals with alcohol use disorders (alcohol abuse and

dependence) differ from abstainers and individuals who had no alcohol use disorders (ex-drinkers, ex-abuser/dependents, new-drinkers, and ex-new-drinkers) in terms two labor market outcomes: the probability of full time work participation and annual earnings (annual personal income).

RESEARCH DESIGN AND METHODS

The estimation strategies are as follows. First, using the NESARC 2001-02 survey sample of 43,093, I estimate a Logit model by the Maximum Likelihood Estimation (MLE) method where the dependent variable is an indicator variable (LFPFULL) for full-time labor force participation (Chapter 3, Section 3.1). Second, using a sample of 21406 (the observations for individuals who had job), I estimate a model with the logarithm of annual earnings (LINCOME) by the Ordinary Least Square (OLS) method. In these estimations, I include a set of explanatory variables: six binary variables to indicate alcohol use status (see Tables 2A and 2B), and other binary variables for other personal and socio-demographic characteristics (see Table 4), such as individual’s age, gender, race, marital status, education level, health status, location (whether individual lived in a central city or not), other source of income, and work related characteristics (industry type). Third, I perform formal tests to detect the existence of potential problems in estimations, the endogeneity and heteroscedasticity in the sample. The estimation strategy (appropriate estimation method) to address endogeneity is required to consider the potential presence of heteroscedasticity in the sample since the sample set is cross-sectional (see details in Chapter 3, section 3.2). Fourth, as the statistical tests confirmed the existence of both problems (endogeneity and heteroscedasticity) in current estimation, I re-estimate labor market outcome equations by addressing endogeneity and heteroscedasticity using GMM-IV method (proposed by Baum et al: 2003, Amemiya: 1985 and Foster: 1997).

RESULTS

As the statistical tests confirmed the endogeneity of alcohol use related variables and the presence of heteroscedasticity in the sample, the estimation by applying GMM-IV are expected to produce consistent and efficient estimate compared to baseline estimates (from the MLE and OLS) which do not address the endogeneity and heteroscedasticity problems. Thus, the estimated results by MLE and OLS methods could be biased and not reliable. The followings are the key results of GMM-IV estimations.

The key result of GMM-IV regression of LFPFULL is that the marginal effects for alcohol abuse and alcohol dependent are -0.38 and -0.04 respectively (though not significant). indicating that alcohol abusers and alcohol dependents have 38 percent and 4 percent less probability of being employed full-time respectively than life-time abstainers (base category), holding all other variables constant. The marginal effect of being ex-drinker on LFPFULL is negative (statistically significant) and consistent with expectation. Ex-drinkers have 43 percent less probability of being employed full-time than life-time abstainers (base category). The positive marginal effects of being ex-abuser/dependents, new drinkers and ex-new-drinkers on LFPFULL (not significant) are 11 percent, 45 percent and 2 percent respectively though not significant.

The key result of the GMM-IV regression of LINCOME is that the marginal effect for alcohol abuse on annual income is -112,057 (though not significant). It indicates that alcohol abusers earned $112,057 less than life-time abstainers (base category), holding all other variables constant. The marginal effects of all other alcohol use status indicating binary variables (ex-drinker, ex-abuser/dependent, new-drinker, ex-new-drinker and alcohol dependents) are found to have positive impacts on annual income (statistically significant for ex-drinkers and new-ex drinkers). The positive impact on annual income for being alcohol dependent is inconsistent with expectation.

Compared to GMM-IV estimates, the MLE (without addressing endogeneity and heteroscedasticity) underestimate the effects of alcohol use variables (represent by binary variables) and overestimate the effects of other socio-demographic variables on labor market variables, and the OLS (without addressing endogeneity and heteroscedasticity) underestimate the effects of all explanatory variables on labor market outcomes.

CONCLUSION

Some observed unexpected results should be treated with cautions considering the limitations of this research: there might be measurement errors in proxy earnings variable, there were data limitations on previous drinking record of individuals who had alcohol use disorders and some labor market information such as hours of work and loss of working hours (absence from the job) due to alcohol use. Besides these limitations, overall results are largely consistent with the results that observed in parallel labor and health economics literature and can be considered representative since this research used rich and nationally representative NESARC data source. The results of this study can be useful for policy and management research to face the challenges of having and maintaining productive and healthy work force. The results imply the necessity of adopting clear and well communicated policies concerning recruitment, monitoring, early prevention and access to effective treatment, and maintaining positive work environment. The results also imply that public or private policies addressing related issues of alcohol use and employment should take into account the fact that women react differently to the amount of alcohol consumption and their work decision also different than men.

AUTOBIOGRAPHICAL STATEMENT
SHAMMIMA JESMIN

EDUCATION

2001	Master of Arts (Applied Economics) Eastern Michigan University, Michigan, USA.
1990	Master in Social Science (Economics) University of Dhaka, Dhaka, Bangladesh.
1988	Bachelor in Social Science (Economics), University of Dhaka, Dhaka, Bangladesh.

AWARDS/ACHIEVEMENTS

- Graduate Teaching Assistantship: Wayne State University, Michigan, USA.

- 1999 and 2000 special academic recognition award, 1999 and 2000 graduate assistant recognition award, and 1999 the tutor recognition award: Eastern Michigan University, Michigan, USA.

- USAID scholarship (1999-2000): Nominated by the Central Bank of Bangladesh.

PROFESSIONAL EXPERIENCES

2000-2005	GRADUATE TEACHING ASSISTANT Department of Economics Wayne State University Detroit, Michigan, USA.
1998-2000	GRADUATE TEACHING ASSISTANT Department of Economics Eastern Michigan University Ypsilanti, Michigan, USA.
1992-1998	DIRECTOR (RESEARCH) Research Department, Bangladesh Bank The Central Bank of Bangladesh, Dhaka, Bangladesh.