ESSAYS ON THE EDUCATION GRADIENT OF HEALTH IN TURKEY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

HANİFE DENİZ KARAOĞLAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF ECONOMICS

JULY 2015

Approval of the Graduate School of Social Sciences

Prof. Dr. Meliha Altunışık
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Doctor of Philosophy.

Prof. Dr. Nadir Öcal
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Doctor of Philosophy.

Prof. Dr. Aysıt Tansel
Supervisor

Examining Committee Members

Prof. Dr. Meltem Dayıoğlu-Tayfur (METU,ECON)

Prof. Dr. Aysıt Tansel (METU, ECON)

Assist. Prof. Dr. İlhan Can Özen (METU, ECON)

Assoc. Prof. Dr. Çağla Ökten (Bilkent U., ECON)

Assoc. Prof. Dr. Nur Asena Caner (TOBB ETU, ECON)

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: HANİFE DENİZ KARAOĞLAN

Signature :

ABSTRACT

ESSAYS ON THE EDUCATION GRADIENT OF HEALTH IN TURKEY

Karaoğlan, Hanife Deniz
Ph.D., Department of Economics
Supervisor: Prof. Dr. Aysıt Tansel

July 2015, 285 pages

This thesis investigates the association between health and education for Turkey by using Turkish Health Survey (THS) data set. In the analysis, we focus on the individuals who are 25 years old or older. In the thesis, first, we examine the factors that possibly have impact on the individual's health such as education, gender, age, region, marital status, labor market indicators and household income. Second, we examine the effect of education on individual's Self-Assessed Health Status (SAH) empirically by considering first education is exogenous to the individual's health by implementing ordered probability analyses (both ordered probit and ordered logit). Afterwards, we consider the possible endogeneity problem between health and education. In order to overcome the endogeneity problem between health and education, we implement Instrumental Variable (IV) methodology. We introduce a new instrument for education to the literature: The education reforms that took place in the early 1960s, after the 1960 constitution had been implemented. Third, we test the effect of education on different health behaviors, such as smoking, alcohol consumption, fruits and vegetable consumption, exercising and BMI. Last, we specifically focus on the rising health problem both in the World and in Turkey:

Obesity, and we examine how the health behavior determining factors vary across different quantile distribution of the individual's BMI level.

Our results suggest that the individual's education level is positively correlated with his/her SAH. In general, our results also indicate that education an important factor that could reduce the probability of risky health behaviors in Turkey.

Keywords: Self-Assessed Health, Health Behaviors, Education, Instrumental Variable, Quantile Regression

ÖZ

TÜRKİYE'DE EĞİTİM VE SAĞLIK GRADYANTI ÜZERİNE MAKALELER

Karaoğlan, Hanife Deniz
Doktora, İktisat Bölümü
Tez Yöneticisi: Prof. Dr. Aysıt Tansel

Temmuz 2015, 285 sayfa

Bu çalışma, eğitim ve sağlık arasındaki ilişkiyi Türkiye için incelemektedir. Bu tezdeki analizler Türkiye Sağlık Araştırması (TSA) veri seti kullanılarak yapılmıştır. Çalışmada 25 yaş ve üstü kişiler üzerine odaklanılmıştır. Bu tezde öncelikle kişinin sağlık durumunu belirleyici olası etkenler belirlenmiş ve bu etkenlerin insan sağlıyla ne şekilde ilişkili olabileceği tartışılmıştır. Bu faktörler eğitimle beraber, yaş, cinsiyet, yaşanılan bölge (kentsel/kırsal), medeni durum, istihdam durumu ve hane halkı geliri olarak sıralanabilir. Tezde ikinci olarak kişinin kendi rapor ettiği sağlık durumu ve eğitim seviyesi arasındaki ilişki amprik olarak incelenmiştir. Amprik çalışmada öncelikle eğitimin sağlıkla egzojen bir ilişkisi olduğu varsayılmış ve sıralı olasılık modelleri kullanılarak ilişki test edilmiştir. Daha sonra eğitim ve sağlık arasındaki içsellik problemi ele alınmış ve bu problemi çözmek için araç değişken yöntemleri kullanılarak tahminler tekrar edilmiştir. Çalışmada eğitimin araç değişkeni olarak Türkiye'de 1960'lı yılların başında yapılan eğitim reformları kullanılmıştır. Tez de üçüncü olarak sigara tüketimi, alkol tüketimi, meyve ve sebze tüketimi, düzenli egzersiz yapma ve Vücut Kitle Endeksi (VKİ) gibi sağlık davranışları ve eğitim seviyesi arasındaki ilişki incelenmiştir. Tezde son olarak VKİ ve eğitim ilişkisi arasında özellikle durulmuş ve kantil regresyon analizleri kullanılarak VKİ'nin değişik yüzdelik dilimlerinde eğitim ve diğer sosyo-ekononomik faktörlerin nasıl bir rol oynadığı tartışılmıştır.

Tezin sonucunda genel olarak kişinin eğitim seviyesinin sağlığı üzerinde olumlu bir etkisi olduğu tespit edilmiştir. Bulduğumuz sonuçlara göre eğitim seviyesi yüksek olan bir kişinin sağlık durumu düşük eğitim seviyesindekilere göre daha iyidir. Ayrıca, yüksek eğitim seviyesine sahip kişilerin sigara gibi sağlığa zararlı davranışların doğuracağı sonuçlar hakkında daha farkındalık sahibi olduğu gözlenmiştir.

Anahtar Kelimeler: Kişinin kendini hissettiği sağlık durumu, sağlık davranışları, eğitim, araç değişken modelleri, kantil regresyon

In the memory of my grandmother

Ayten Özdem

ACKNOWLEDGMENTS

I would like to express my deepest gratitudes to Prof. Dr. Aysıt Tansel for her supervision and guidance through the development of this thesis. I also would like to thank to Prof. Dr. Meltem Dayıoğlu-Tayfur, Assist.Prof.Dr. İlhan Can Özen, Assoc. Prof. Dr. Çağla Ökten and Assoc. Prof. Dr. Nur Asena Caner for their valuable comments on my thesis.

My thanks should also go for Tuna Kemali for his helps with education and census statistics. I will not be able to convince the jury that my results are satisfying without this statistics. I would also thank to Assoc. Prof. Dr. Murat Kırdar, and Assist. Prof. Dr. İpek Eren for their helps with the empirical model and answering all of my questions with patience. I would also want to thank to TURKSTAT for providing me Turkish Health Survey Data Set.

I would like to give my special thanks to Assit. Prof. Dr. Serkan Küçükşenel, Assoc. Prof. Dr. Şirin Saraçoğlu, Assist. Prof. Dr. Seven Ağır, Assist. Prof. Dr. Hasan Cömert, Assoc. Prof. Dr. Elif Akbostancı and Prof. Dr. Nadir Öcal for their continous support, encouragement and motivation during my Ph.d degree and during the job market process.

My thanks should also go for the participants of several conferences that I attend between 2012 and 2015 and of course the participants of several seminars in different universities that I give in 2014 and 2015. Without their suggestions, I could not improve the academic quality of my thesis.

I am grateful to my colleagues in METU during my assistantship between 2011 and 2015. Each of them are very important for me. But, I should give my special thanks to Burcu Özgün and Ceren Akbıyık for their endless helps in my last semester. My special thanks should also go for Başak Dalgıç, Burcu Fazlıoğlu, Ceren Baran, Gizem Kurt, Selin Gençtürk and Şaban Çelik, my co-authors and my dearest collegues in Hacettepe and Yasar Universities.

Of course, I cannot forget the administrative board of BILMED. My thanks should especially go for Şeref Özer and Sinan Öncel for giving me the chance of being a member of their family in such a stressful period.

Zumba and yoga make me a person who is full of energy in this hard period. I am indebted to Çağla Ateşalp for meeting me with the yoga. Yoga becomes more beautiful with her, Selen Gül, Fikret Çirkin and Pınar Karakuşoğlu. I send my special thanks to Özlem Alataş, Boran Alataş and Elif Karaöz, my patient yoga instructors. I will never forget my lovely Zumba instructors: Gizem Saygılı and Şule Akdoğan: You become my energy machine in the last difficult 3 years. I love you so much!

Although they are far away in this period, I always feel the support and love of ten very special friends: Duygu Pekşen, Kübra Güler, Beliz Osman, Eda Pekalp, Zeynep Çağlayan, Merve Şen, Sedef Şimşir, Derya Aksaray, Merve Doğan and Sel Özcan.

Finally, and most importantly, I would like to thank to my family for their endless love and support. I am proud of being the daughter of Nilgün and Muharrem Karaoğlan, and I am very lucky to be the sister of Pınar Karaoğlan.

TABLE OF CONTENTS

PLAGIARISM iii
ABSTRACT iv
ÖZ vi
DEDICATION..viii
ACKNOWLEDGMENTS ix
TABLE OF CONTENTS xi
LIST OF TABLES xiv
LIST OF FIGURES xvii
LIST OF ABBREVIATIONS xix
CHAPTER
1. INTRODUCTION 1
1.1. BACKGROUND AND SCOPE OF THE THESİS 1
1.2. SUMMARY OF THE CHAPTERS 3
2. THEORETICAL MODELS ON HEALTH PRODUCTION FUNCTION 12
2.1. INTRODUCTION 12
2.2. HEALTH DEMAND THEORY 13
2.3. EXTENSIONS ON HEALTH DEMAND THEORY 18
2.4. EMPIRICAL SPECIFICATION BASED ON THEORETICAL FRAMEWORK 21
2.5. CONCLUSION 23
3. DATA 25
4. SELF –ASSESSED HEALTH STATUS (SAH) AND ITS DETERMINING FACTORS 36
4.1. INTRODUCTION 36
4.2. LITERATURE SURVEY 388
4.2.1. SAH as Health Outcome 388
4.2.2. Studies on Developing Countries 43
4.3. FREQUENCY DISTRIBUTION ANALYSIS 45
4.3.1. SAH and Education Level 46

4.3.2. SAH and Gender 53
4.3.3. SAH and Age 55
4.3.4. SAH and Marital Status 57
4.3.5. SAH and Region 58
4.3.6. SAH and Household Income 59
4.3.7. SAH and Employment Status 64
4.4. CONCLUSION 65

5. METHODOLOGY 67

5.1. INTRODUCTION 67
5.2. ORDERED PROBABILITY MODELS 69
5.3. PROBIT MODEL 71
5.4. INSTRUMENTAL VARIABLE APPROACH 72
5.4.1. Literature Survey on IV Methodology 72
5.4.2. IV Methodology 76
5.4.2.1. IV Estimation Process 76
5.4.2.2. Methodology for the Endogeneity Tests 77
5.4.3. Discussion on the Instrument 78

6. SELF-ASSESSED HEALTH AND EDUCATION 102

6.1. INTRODUCTION 102
6.2. EMPIRICAL SPECIFICATION FOR THE ORDERED PROBABILITY MODELS 104
6.2.1. Ordered Probit Analysis 104
6.2.2. Ordered Logit Analysis 112
6.2.3. EXTENSION OF THE ORDERED MODELS 116
6.2.3.1. Ordered Probit Analysis 116
6.2.3.2. Ordered Logit Analysis 126
6.3. IV ESTIMATION PROCESS 129
6.3.1. Empirical Specification 129
6.3.2. Empirical Results 133
6.3.2.1. Descriptive Statistics 133
6.3.2.2. First Stage Results 135
6.3.2.3. Second Stage Results 140
6.4. CONCLUSION 147
APPENDIX 6.A.1: 2SLS ESTIMATION RESULTS 1499
APPENDIX 6.A.2. IV-ESTIMATION REVISITED 151

7. HEALTH BEHAVIORS AND EDUCATION 153

7.1. INTRODUCTION 153
7.2. LITERATURE REVIEW ON HEALTH BEHAVIORS 155
7.3. DESCRIPTIVE STATISTICS 180
7.3.1. Smoking 180
7.3.2. Alcohol Consumption 182

7.3.3. Fruit and Vegetable Consumption 184
7.3.4. Exercising 185
7.3.5. BMI 187
7.4. Empirical Results 189
7.4.1. Smoking 191
7.4.2. Alcohol Consumption 195
7.4.3. Fruit and Vegetable Consumption 199
7.4.4. Exercise 202
7.4.5. Body Mass Index (BMI) 206
7.5. INSTRUMENTAL VARIABLE (IV) APPROACH 209
7.6. CONCLUSION 216
APPENDIX 7.A.1. IV ESTIMATION REVISITED 218
8. QUANTILE REGRESSION ANALYSIS: THE RELATIONSHIP BETWEEN BMI AND EDUCATION IS REVISITED 220
8.1. INTRODUCTION 220
8.2. LITERATURE SURVEY 222
8.3. QR METHODOLOGY AND EMPIRICAL RESULTS 225
8.4. QUANTILE REGRESSION ANALYSIS WITH IV METHODOLOGY 231
8.5. CONCLUSION 234
9. CONCLUSION 236
9.1. SUMMARY OF THE MAIN FINDINGS AND CONCLUSIONS 236
9.2. POLICY IMPLICATIONS 242
9.3. FURTHER RESEARCH 246
REFERENCES 248
APPENDICES 263
A. ORIGINAL LAWS 263
B. TURKISH SUMMARY 265
C. CURRICULUM VITAE 283
D. TEZ FOTOKOPİSİ İZİN FORMU 285

LIST OF TABLES

TABLES

Table 3.1. The Distribution of Diseases/Health Problems diagnosed by a Physician by reported SAH………………………………………………………....………....29

Table 4.3.1. SAH and Education Level…………………………………………..52

Table 4.3.2. Distribution of Reported SAH Measures………………………………54

Table 4.3.3. Percentage Distribution of SAH by Age Group and Gender…......…....56

Table 4.3.4. Percentage Distribution of SAH by Marital Status……………….........58

Table 4.3.5 Percentage Distribution of SAH by Region…………………………........59

Table 4.3.6. Percentage Distribution of SAH by Household Income……….........62

Table 4.3.7. Percentage Distribution of SAH by Employment Status……….........65

Table 6.2.1.1. Ordered Probit Results (Education Control: Years of Schooling).....108

Table 6.2.1.2. Ordered Probit Results (Education Control: Education Dummies)...110

Table 6.2.2.1. Ordered Logit Results: (Education Control: Years of Schooling).....113

Table 6.2.2.2. Ordered Logit Results (Education Control: Education Dummies)…114

Table 6.2.3.1.1. Ordered Probit Results for the Extended Models ………………..117

Table 6.2.3.1.2 Marginal Effects for the Extended Models …………………….....119

Table 6.2.3.1.3. Ordered Probit Results for the extended models…………………121

Table 6.2.3.1.4. Marginal Effects for the Extended Models……………….……....123

Table 6.2.3.2.1 Ordered Logit Results for the Extended Models …………………126

Table 6.2.3.2.2. Ordered Logit Results for the Extended Models ………………...127

Table 6.3.2.1.1. Summary Statistics…………………………………………….....134

Table 6.3.2.2.1. First- Stage Estimation Results (The individuals who were born in and after 1952 are affected from the change in education policies…......................137

Table 6.3.2.2.2. First- Stage Estimation Results (The individuals who were born in and after 1951 are affected from the change in education policies)……………….138

Table 6.3.2.2.3. First- Stage Estimation Results (Birth Cohort 1951 is excluded)...139

Table 6.3.2.3.1. Estimates of SAH Returns to Schooling (The individuals who were born in 1952 or later are affected from the educational reforms)………………….141

Table 6.3.2.3.2. Estimates of SAH Returns to Schooling (The individuals who were born in 1951 or later are affected from the educational reforms)………………….143

Table 6.3.4.3.3. Estimates of SAH Returns to Schooling (Birth Cohort 1951 is excluded)…………………………………………………………………………...145

Table 6.A.1. OLS and Second Stage IV Estimation Results of SAH Returns to Schooling………………………………………………………………………...149

Table 6.A.2. 2SLS Estimation Results (With Trend Controls)…………………….152

Table 7.3.1. Descriptive Statistics for Smoking……………………….……..180

Table 7.3.2. Descriptive Statistics for Alcohol Consumption……………….…….182

Table 7.3.3. Descriptive Statistics for Fruit and Vegetables (FV) Consumption….184

Table 7.3.4. Descriptive Statistics for Exercising………………………………….186

Table 7.3.5. Descriptive Statistics for BMI……………………………….……….188

Table 7.4.1.1. Marginal Effects from Probit Estimation Results for Smoking (%) with Years of Schooling……………………………………………………………….192

Table 7.4.1.2. Marginal Effects from Probit Estimation Results for Smoking (%) with Education Levels…………………………………………………………...193

Table 7.4.2.1. Marginal Effects from Probit Estimation Results for Alcohol Consumption (%) with Years of Schooling…...…………………………………...196

Table 7.4.2.2. Marginal Effects from Probit Estimation Results for Alcohol Consumption (%) with Education Levels…………………………………………197

Table 7.4.3.1. Marginal Effects from Probit Estimation Results for Fruit and Vegetable Consumption (%) with Years of Schooling…………...………………..200

Table 7.4.3.2. Marginal Effects from Probit Estimation Results for Fruit and Vegetable Consumption (%) with Education Levels………………………………201

Table 7.4.4.1. Marginal Effects from Probit Estimation Results for Exercise (%) with Years of Schooling……………………………………………………………203

Table 7.4.4.2. Marginal Effects from Probit Estimation Results for Exercise (%) with Education Levels…………………………………………………………...204

Table 7.4.5.1. OLS Estimation Results for BMI with Years of Schooling………...207

Table 7.4.5.2. OLS Estimation Results for BMI with Education Levels…………..208

Table 7.5.1. First Stage Results…………………………………………….....212

Table 7.5.2. Second Stage Results………………………………………….....214

Table 7.A.1. IV-Probit (2SLS for BMI) Estimation Results………………………218

Table 8.3.1. OLS and QR Estimation Results……………………………………..227

Table 8.4.1. 2SLS and IV-QR Results (Dependent Variable: BMI)………………233

LIST OF FIGURES

FIGURES

Figure 4.3.1. SAH by Education Level…………………………………...………52

Figure 5.3.2.1. The number of the individuals who have at least primary school degree………………………………………………………..……………...83

Figure 5.3.2.2. City and Village Population in Turkey Over Time…………….......84

Figure 5.3.2.3. Ratio of City and Village Populations to the Total Population……..84

Figure 5.3.2.4. Number of Primary School Graduates in Villages……………........85

Figure 5.3.2.5. Number of Students in Primary Schools in Turkey Over Time…….86

Figure 5.3.2.6. Number of Students in Middle Schools in Turkey over Time…..…..87

Figure 5.3.2.7. Number of Students in High Schools in Turkey Over Time……..…88

Figure 5.3.2.8. Number of Primary School Teachers………………………….......89

Figure 5.3.2.9. Number of Middle School Teachers……………………..…........90

Figure 5.3.2.10. Number of High School Teachers…………………………….…….91

Figure 5.3.2.11. Number of Primary Schools in Turkey Over Time………….........92

Figure 5.3.2.12. Number of Middle Schools in Turkey Over Time………………...93

Figure 5.3.2.13. Number of High Schools in Turkey Over Time…………………...94

Figure 5.3.2.14. Number of Primary School Teachers per School-Aged Children in Turkey Over Time…………………………………………………………...95

Figure 5.3.2.15. Number of Middle School Teachers per School-Aged Children in Turkey Over Time…………………………………………………………...95

Figure 5.3.2.16. Number of High School Teachers per School-Aged Children in Turkey Over Time…………………………………………………………...96

Figure 5.3.2.17. Primary School Enrollment Rate in Turkey Over Time………..…97

Figure 5.3.2.18. Middle School Enrollment Rate in Turkey Over Time……………97

Figure 5.3.2.19. High School Enrollment Rate in Turkey Over time……….............98

Figure 5.3.2.20. Primary School Graduates to Primary School Graduate Aged Population………………………………………………………..………….99

Figure 5.3.2.21. Middle School Graduates to Middle School Graduate Aged Population………………………………………………………………………100

Figure 5.3.2.22. High School Graduates to High School Graduate Aged Population………………………………………………………………………100

Figure 7.4.1.Marginal Effects for Smoking (x100) by Education Level……….....194

Figure 7.4.2. Marginal Effects for Alcohol Consumption (x100) by Education Level……………………………………………………………………………198

Figure 7.4.3. Marginal Effects for Fruit and Vegetable Consumption (x100) by Education Level………………………………………………………………202

Figure 7.4.4. Marginal Effects for Exercising (x100) by Education Level………..205

Figure 7.4.5. Coefficients for BMI by Education Level…………………………...209

LIST OF ABBREVIATIONS

THS	Turkish Health Survey
SAH	Self-Assessed Health Status
BMI	Body Mass Index
IV	Instrumental Variable
QR	Quantile Regression
IV-QR	Instrumental Variable Quantile Regression
TSA	Türkiye Sağlık Araştırması
VKİ	Vücut Kitle İndeksi

CHAPTER 1

INTRODUCTION

1.1. Background and Scope of the Thesis

Human capital is an accelerating factor for the improvement of productivity, welfare and earnings in a country. Schultz (1961) is one of the first studies that points out the importance of investment in the human capital. In that study, Schultz refers to the following as the investments in human capital: Direct expenditures on education, health, internal migration to get better job opportunities, earnings forgone by mature students attending school and earnings forgone workers who participate in job training activities as human capital investment. According to him, such investments are the major factors leading to a rise in the real earnings per worker. Becker (1993) argues that human capital may also be in the form of bad habits, such as addictions to heavy drinking. He states that this form of human capital has a negative impact on productivity in the market and nonmarket activities.

Health and education are the two most important components of human capital. The positive impact of investment in these two forms of human capital on the growth of an economy has been discussed and shown in several studies. For instance, Mankiw et al. (1992) show that the augmented Solow Model which includes education as well as physical capital explains the income differences across the countries better explains than the former model where education is not included. Later on, Weil (2007) includes health to the augmented Solow model and conducts a study that mainly focuses on the direct effect of health on the economic growth. He states that if the individuals are healthier, they become more productive workers, and this leads to a rise in the country's growth rate. Weil suggests that health also has an economically significant effect on determining income differences across countries.

This thesis deals with the relationship between the two forms of human capital, health and education in Turkey, a middle income, developing country. The relationship between health and education is important though not new. The link between health and education has been examined widely in the literature. The association between education and health is simply called as the "education gradient of health". In the health economics literature, the relationship between education and health is generally tested empirically by using several appropriate econometric methods. In general, the empirical results suggest that there is a positive association between health and education. Cutler and Lleras-Muney (2010) explain the possible mechanisms for the positive linkage between education and health as follows: First, higher levels of education bring about greater access to health care. Second, educated people are provided with better health insurance because they have better jobs, thus they work in safer environments. Third, higher level of education leads to better future, therefore, higher educated individuals tend to invest in their health more to protect that future. Finally, more educated people are better informed, hence, they utilize new health related information first.

Most of the previous studies that examined the association between health and education are conducted for developed countries. In general, those studies suggest a strong and positive link between health and education. There is less evidence for developing countries and there are not many studies on this issue for developing countries. To our knowledge, there is no previous study, which specifically examines the association between health and education in Turkey for adults by using the rich micro-data set in one study. In this thesis, we examine the relationship between health and education for the individuals who are 25 years of age or above. The first empirical part of this thesis investigates the association between individual's Self-Assessed Health (SAH) and education level, and the second empirical part is devoted to examining the relationship between education and different health behaviors, such as smoking, alcohol consumption, fruit and vegetable consumption, exercise and Body Mass Index (BMI). This study is new in the sense that both empirical parts are conducted for adults for the first time in Turkey.

There are only a few studies that investigates the association between health and education in Turkey. Erdogan et al. (2012) examines the relationship between gender inequality in education and selected health outcomes such as life expectation at birth, under five mortality rate and infant mortality rate by using Turkish time-series data. Tansel (1993) uses cigarette consumption as health outcome. In that study, she examines the demand characteristics of cigarette consumption and finds that if people are educated about the adverse effects of smoking, the demand for cigarettes may decrease by a larger amount than the decrease in the demand that may result from an increase in cigarette prices. Therefore, her study supports the positive effect of health knowledge on reduction in cigarette consumption. In a recent work, Cesur et al. (2014) examine the association between education and different health behaviors in Turkey for the individuals who are between 18 and 30 years of age. Their results do not indicate a significant relationship between smoking and education level for young men in Turkey. However, they find that young men with high levels of education are more likely to be obese or overweight than young men with low levels of education.

1.2. Summary of the Chapters

The positive association between health and education is exposed theoretically for the first time by Grossman (1972). He is one of the first economists who provide formal explanations of the observed differences in health outcomes by education. In his model, individual's education level is one of the most important determinants of the demand for health. By forming a discrete time optimization problem, Grossman concludes that optimal choice of investment in health depends on individual's wage rate, marginal product of the stock of health in the production of healthy days, marginal cost of gross investment in health, interest rate and depreciation. In addition, Grossman concludes that education is positively associated with health capital and negatively associated with expenditures on health care. Following Grossman, several health demand models are formulated with slight extensions or with new assumptions. Chapter Two reviews this theoretical framework. The first part of Chapter Two examines the Grossman's 1972 health-demand model. The latter part of chapter two examines the other health demand models which extend the assumptions of Grossman's 1972 model. The last part of chapter two introduces the covariates that we use to explain the variations in health outcomes. In that part, we explain why we choose those variables as the determinants of various health outcomes that we consider in this thesis.

For empirical purposes, Turkish Health Survey (THS) data set for the years 2008, 2010 and 2012 are used in this thesis. The THS is a cross sectional data set over individuals. The THS is prepared and conducted by Turkish Statistical Institute (TURKSTAT). The observations for the three waves of this survey are pooled in empirical analysis. For this purpose we tested and did not find significant variations in covariates for different time periods. Chapter Three describes the THS data set, how it is formed, what kind of questions are asked in the questionnaire with a focus on the specific health related questions relevant for our analysis. The advantages and shortcomings of the data set are both discussed in Chapter Three.

The first health outcome considered in this thesis is the individual's self-assessed health (SAH) status. Many authors indicate that SAH is a good subjective measure of health in the sense that it is a good predictor of mortality (Idler and Benyamini, 1997). In addition, SAH is easier to measure than clinical measures of health and standardized self-reported surveys of health status are inexpensive to process (Miilunpalo et al.,1997). In general, the literature suggests that it is advantageous to use SAH as a health outcome. However, there are some studies which suggests that using SAH as a health outcome may be misleading because it may include measurement errors (Butler et al., 1987) or there may be misspecification of health (Baker et al., 2004). Such issues may lead to biases in the estimation results. The first part of Chapter Four examines the literature on the use of SAH as the health outcome and discusses the advantages and disadvantages of using SAH as the indicator for the individual's general health status. In this part, we also review the limited literature on developing countries that examines the variations in SAH empirically.

Individual's SAH is asked frequently in health surveys all around the world. According to OECD 2010 data set, in Turkey, 66 per cent of working age population state that they are in good health. This percentage is low relative to that of the developed countries such as Canada, USA and most of the EU countries. USA and Canada have the highest percentages of population with good SAH. Individual's SAH tends to be poorer in Asian Countries such as Korea and Japan[1]. The second part of Chapter Four provides a descriptive analysis of the variations of SAH by education, gender, age, region, marital status, labor market indicators and household income in Turkey.

In Chapter Five, we describe the econometric techniques that we use in chapters six and seven. We introduce the IV methodology and discuss the validity of the instrument. We use the educational expansion in the early 1960s in Turkey as the instrument for years of schooling. We think this is a good IV as it is implemented in the all around the Turkey in a short time period, therefore, we think it leads to an increase in the individual's years of schooling. This is a unique instrumental variable that is being used for the first time in any study on Turkey. We establish validity of this instrument for years of schooling with extensive investigations of the several educational outcomes in the 1950s and 1960s. Chapter Five includes an elaborate discussion of these results. We also investigate health reforms in the 1960s in Chapter Five. We conclude that health reforms are not as effective as the education reforms, therefore, they do not coincide with each other.

Several studies indicate that although the positive association between health and education is valid, the direction of the causality between them is inconclusive. The difficulty in reaching a conclusion about the direction of the causality occurs because of the fact that education may be endogeneous to health. The endogeneity can arise because of two facts: First, investment in health in earlier ages can affect both education and health in later ages. Second, unobserved variables such as time preference, genetic factors and family background can affect both education and

[1] Source: OECD Health Data Set (http://stats.oecd.org)

health simultaneously (Eide and Showalter 2011, Bolin 2011). In order to detect the causality between education and health correctly, it is necessary to solve the endogeneity problem between these variables.

In this thesis, the endogeneity problem between individual's SAH and education is resolved by using Instrumental Variables (IV) methodology. To our knowledge, this is the first study that examines the validity of the education gradient of health in Turkey by considering the endogeneity problem between individual's health and education level for adults. The educational reforms that took place in October 1960 and January 1961 in Turkey are used as the instrument for the individual's education level. This thesis is the first study which uses the 1960s educational reforms in Turkey as instrument for the individual's years of schooling. Next, in Chapter Five we present a brief literature review on instrumental variable methodology. Then we describe the educational reforms as well as the health reforms of the early 1960s. We discuss the validity of the instrument by graphically examining the relevant data from the Turkish Education Statistics and Census. We show that health reforms do not coincide with education reforms in those years. Therefore, education reforms are purely exogenous factors to individual's health.

The regression analysis of the relationship between individual's SAH and education level is conducted in Chapter Six. Since individual's SAH is a categorical variable, which takes values from 1 to 5, ordered probability analyses are the most appropriate models for this study. We first assume that error terms are normally distributed and conduct ordered probit analysis where the health outcome is SAH and explanatory variables are individual's education level, age, gender, marital status and the region where the individual is located. However, we cannot obtain odds ratios with ordered probit regression. It is possible to obtain odds ratios with ordered logit regression. Hence, ordered logit regression, in which the error terms are assumed to have logistic distribution, is implemented so that we can obtain the odds ratios of the covariates. The dependent and explanatory variables are the same as in ordered probit regression. The results of the two ordered probability models do not significantly differ from each other. Our results suggest that there is positive relationship between health and education in Turkey. In the final part of Chapter Six, the two models are extended by including labor market outcomes and household income. However, such variables can be endogenous to individual’s SAH. Here, we include them in order to check the robustness of our main results.

The last part of Chapter Six reexamines the relationship between individual’s SAH and education by taking the endogeneity problem into account. For continuous dependent variables, the STATA “ivregress” command is used to implement the instrumental variable technique. The dependent variable in our case is categorical. The predicted values of SAH may be out of specific intervals (1 to 5) if linear regression is implemented in the second stage. Therefore, following Roodman (2011), we implement the “Conditional Mixed Process” (CMP) estimation technique. This procedure allows us to run the ordered probit regression in the second stage. Chapter Six introduces the CMP methodology, discusses how it fits with our analysis and reports the findings that result from CMP analysis. In the appendix to Chapter Six, we also carry out the two-stage least Squares (2SLS) regression techniques assuming that SAH is continuous variable. Our results indicate that the education reforms in the early 1960s cause between 0.99 and 1.27 years increase in the adults’ years of schooling. Hence, we can say that educational reforms are effective for increasing the individuals’ years of schooling.

Chapter Seven is devoted to an examination of the relationship between education and five different health behaviors, namely, smoking, alcohol consumption, fruits and vegetable consumption, exercising and body mass index (BMI). We examine the association between different health behaviors and years of schooling along with demographic factors (such as age, gender and region), labor market indicators and household income. The differentials in health behaviors by education are studied mostly for developed countries. There is less evidence on this issue for developing countries. Therefore, the purpose of Chapter Seven is to examine the health behaviors by education in Turkey which is a developing country. This is the first study that investigates the relationship between the main health behaviors and education for the individuals who are 25 years old or over in Turkey in one study.

The first part of Chapter Seven examines the literature on main health behaviors and, argues why they are beneficial or harmful for the individual's overall health. We further discuss the education gradient of health where the health outcomes are different health behaviors.

Smoking is one of the major health behaviors that is used as a health outcome in this thesis. OECD (2010) Health Data set reports that 25.4 per cent of the adult population in Turkey is a regular smoker. This is the second highest rate after Estonia (26.2 per cent) among the OECD countries. Smoking is one of the most harmful health behaviors. Regular smokers are at great risk for cardiovascular disease, chronic lung disease and several types of cancer (Stewart et al., 2009; Chaloupka and Warner, 2000). There is evidence that smoking is less prevalent among the educated individuals compared to less educated individuals in the case of developed countries (Chaloupka and Warner, 2000).

Regarding the other health behaviors, we also examine the relationship between alcohol consumption and education in Turkey. According to the OECD (2010) Health Data set, only 1.5 per cent of adult population in Turkey consumes alcohol. The percentage is very low compared to the alcohol consumption in other OECD countries. The low percentage of alcohol consumption in Turkey is most probably due to the religious traditions which prohibit alcohol consumption.

Fruit and vegetable consumption and exercising are necessary for a healthy life. Sufficient daily consumption of fruit and vegetables could prevent several kinds of diseases, such as cardiovascular diseases, certain cancers and diabetes (Hung et al. 2004). Hence, it is important to investigate the variation in fruit and vegetable consumption. Similarly, regular exercising may prevent several chronic diseases such as heart disease, stroke, high blood pressure, diabetes, obesity, back pain and osteoporosis. It can also improve the psychological condition of the individual (Fletcher et al., 1996).

The final health outcome that is considered in this thesis is the individual's Body Mass Index (BMI). BMI is used as a tool for determining if an individual is overweight or obese. According to the World Health Organization (WHO) criteria, an individual is considered as obese if his/her BMI is greater than 30, overweight if his/her BMI is greater than 25 and underweight if his/her BMI is under 18.5. Obesity is an increasing health problem in Turkey. It is important to analyze the determinants of obesity as it is a major source of certain diseases such as cardiovascular diseases, diabetes, and joint problems (Stewart et al., 2009). OECD (2010) Health Data indicate that 21 percent of females and 13.2 percent of males in Turkey were obese in 2010. World Bank (2008) reports that the adoption of Western diets high in refined carbohydrates, saturated fats and sugars and a more sedentary lifestyle are major contributors to the increase in overweight and chronic diseases in Turkey.

Chapter Seven first presents the descriptive statistics separately for each health behavior. This gives an idea about the variations in health behaviors in Turkey by education and their other determinants. Next, we present the results of a probit estimation where the dependent health behavior is bivariate except that for case of BMI we present OLS estimation results. However, probit (OLS for BMI) estimation may give inconsistent results due to the endogeneity problem between health behaviors and education. Therefore, for smoking, alcohol consumption, fruit and vegetable consumption and exercising, where the dependent variables are bivariate, we establish an ivprobit model in order to obtain consistent estimates. Since BMI is a continuous variable, we implement the 2SLS estimation techniques. The last section of Chapter Seven reports the results from the IV approach. In general, our results suggest the protective effect of education on the health behaviors considered in this thesis with slight deviations from this conclusion. The reasons for the deviations in the empirical results are studied in detail in chapter six.

Chapter Eight is devoted to an examination of the variations in BMI across different quantiles of the BMI distribution. The determinants of the BMI along the distribution of BMI are analyzed by using quantile regression techniques, following Koenker and Basset (1978). We first discuss the advantages of using the quantile regression approach compared to OLS for investigating the variations in BMI. We present and compare the results from quantile regression and from OLS. Next, the quantile regression technique and IV methodology are implemented together by using a user-written STATA command called "ivqreg". This procedure addresses the possible endogeneity problem between individual's BMI and education level.

The main findings of the empirical analysis can be summarized as follows: In general, we find a positive association between education and health in Turkey as in the case of developed countries. We find that additional years of schooling positively affect individual's SAH. We conclude that education is an important factor that could also reduce the probability of main health behaviors in Turkey. Unlike the previous studies in developed countries, we find that the probability of smoking increases with education. However, the effect of university or higher education is smaller than the effect of lower levels of schooling. In addition, IV estimation results suggest that there is a negative relationship between the individual's years of schooling and the prevalence of smoking. Our results also indicate that higher educated individuals tend to consume more alcohol in Turkey than the less educated, however this result differs in IV estimation setting.

Probit model estimation results indicate that higher educated individuals clearly eat more fruits and vegetables compared to the less educated and the illiterate. We also find that higher educated individuals exercise more. Next, we observe that higher educated individuals clearly have BMI levels in the normal range compared to the less educated. Although the protective effect of education on individual's BMI level is not seen in IV estimation settings, both the quantile and the IV quantile regression results suggest that there is negative association between years of schooling and BMI at the top quantiles of the BMI distribution. Therefore we can conclude that higher educated individuals are more aware of the risks of high BMI levels.

Therefore, in general, we can conclude that educated people are better informed about the benefits of good health, and they are more aware of the importance of using the preventative health care services in Turkey. Chapter Nine presents the concluding remarks in detail and discusses the policy implications.

CHAPTER 2

THEORETICAL MODELS ON HEALTH PRODUCTION FUNCTION

2.1. Introduction

This chapter investigates the theoretical models for health demand theory which is first proposed by Grossman (1972). This thesis is mainly an empirical study, however it is important to investigate the theoretical framework in order to understand why we use other covariates except education such as marital status or labor market indicators in order to explain the variations in individuals' health outcomes. Therefore, in this chapter we investigate the basic health demand theories that are relevant for the empirical relationship.

In that chapter, first Grossman's famous health demand model which is conducted in 1972 is briefly examined[2]. We first explain the setting of the individual's utility maximization problem. Then we show how the health investment and consumption functions are formed, and which covariates explain the health demand theory except medical care expenditures. Finally, we focus on the implications of the model.

Next, we consider on other models which can be considered as the extensions of Grossman's model. Indeed, Grossman's 1972 model is still the most famous model in health economics literature. The theoretical models following Grossman, only make slight changes to the original model. Bolin (2011) establishes the most recent health demand model and extends the Grossman's model to continuous time setting. In section two, the major points of Bolin's model is investigated. Then, we refer to the recent literature such as Grossman (2000) and Cutler et al. (2011) in order to

[2] Grossman (1972) study is in fact Micheal Grossman's Ph.d Thesis. Later on Grossman published his Ph.d Thesis in Journal of Political economy.

point out the evaluation of Grossman's model over time (i.e., other covariates rather than what Grossman considers as inputs for health demand are added into the model).

Section 2.4 explains how our empirical models are formed based on the theory. More specifically, we try to explain why we believe that factors such as household income or labor market indicators can explain the variations in health outcomes.Some covariates such as gender, marital status are important determinants of health outcomes. They are examined as the determinants of health outcomes empirically several times, but they are not referred as inputs for individual's health in theoretical models. We examine the significance of these covariates for the individual's health referring to famous existing empirical studies such as Case and Paxson (2005), Fuchs (2004). Last, section 2.5 provides concluding remarks for chapter two.

2.2. Health Demand Theory

The theoretical model for demand for health is first formed by Grossman (1972). In that model, Grossman looks for how individuals allocate their resources to produce health. In Grossman's model, individual is not only health consumer but also health producer. In the model, Grossman introduces the idea of investing in human capital (i.e. investing in both health and education) to improve outcomes in both market and non- market sectors. Throughout the whole model, health is analyzed as a capital good because health depreciates over time. Grossman says that health can be considered as both consumption and investment good. On the one hand, health is consumption good since people get disutility if they are sick because of spending their whole time in bed. On the other hand, health is an investment good since health determines the individual's allocation of time for both market and non-market activities. Grossman concludes that health is not determined by the price of medical care only. The other determinants of health include individual's education level, age and wage rate.

Three important assumptions of Grossman's model are as follows: First, individuals want to maximize their life time utility. Second, individuals have perfect knowledge.

Third, individuals are able to allocate their time between different activities, such as labor versus leisure, health producing time versus non-health producing time. In the light of these three assumptions, Grossman establishes a discrete time optimization model.

In Grossman's model the individual's intertemporal utility function is defined as follows[3]:

$$U=U\,(\Phi_0 H_0,\ \Phi_1 H_1, \ldots\ldots, \Phi_n H_n, Z_0, Z_n) \qquad (2.2.1)$$

In equation (2.2.1), H_0 is the stock of health which the individual inherits. For every i ϵ {1,……,n}, H_i shows the individual's health stock at time i. Grossman defined h_i as $\Phi_i H_i$, which shows individual's total amount of health services consumption and Z_i shows the individual's total amount of consumption for other goods in period i. Grossman states that the individual dies when $H_i=H_{min}$. Hence, he concludes that the individual's length of life depends on the quantity of H_i.

Grossman defines the net investment in health stock by the following equation:

$$H_{i+1}-H_i=I_i-\delta_i H_i \qquad (2.2.2)$$

In equation 2.2.2, I_i refers to the gross investment and δ_i refers to the depreciation rate in period i. In that model the author argues that the rate of depreciation is exogenous but he also states that it may change with individual's age. The author also notes that the rate of production of H depends on the efficiency of investment in H.

Grossman's model is based on human capital model. In other words, it is a model that shows how the individuals invest in themselves, e.g., through education in order to improve their productivity. In the model, the individuals are the producers of H in the sense that they buy market inputs for their health such as medical care, food and clothing and combine them with their own time to produce services that increase their utilities. Grossman assumes that the individuals produce both gross investments in health and the other goods according to the following functions:

[3] The notations are taken from Grossman (1972) "On the Concept of Health Capital and Demand for Health"

$$I_i=I_i\ (M_i,\ TH_i;\ E_i) \qquad (2.2.3)$$

$$Z_i=Z_i\ (X_i,\ T_i;\ E_i)$$

In equation (2.2.3), I_i refers to the investment in health. M_i refers to the medical care, X_i refers to inputs that are used for producing the good Z_i. TH_i refers to the time spent on improving health and T_i refers to the time spent on composite consumption good. and E_i is the individual's education level. From equation (2.2.3) we can infer that in Grossman's model education plays an important role in both health investment and production of other goods. Hence, one can conclude that education influences consumption patterns of the households.

Next, in the model, the individual's budget constraint is defined as:

$$\sum \frac{P_i M_i + V_i X_i}{(1+r)^i} = \sum \frac{W_i TW_i}{(1+r)^i} + A_0 \qquad (2.2.4)$$

Equation (2.2.4) indicates that the present value of spending on health stock and on other goods is equal to the present value of earnings over the whole life plus the initial asset holdings. In equation (2.2.4) P_i is the price of M_i, V_i is the price of X_i. W_i refers to the individual's wage rate at time i. TW_i shows the total hours of work. A_0 is the initial asset holdings and r refers to the interest rate. In addition, individual's time constraint is defined by the following equation.

$$TW_i+TL_i+TH_i+T_i=\Omega \qquad (2.2.5)$$

In equation (2.2.5) TL_i shows the individual's time lost due to sickness or injury and Ω refers to the total available time in any retro. By nature, TL and h have negative relationship. By substituting the equation (2.2.5) into the equation (2.2.4) Grossman left with the full wealth constraint:

$$\sum \frac{P_i M_i + V_i X_i + W_i(TL_i + TH_i + T_i)}{(1+r)^i} = \sum \frac{W_i \Omega_i}{(1+r)^i} + A_0 = R \qquad (2.2.6)$$

Then the optimum levels of H and Z are determined by maximizing the intertemporal utility function subject to the constraints which are given by equations (2.2.4), (2.2.5) and (2.2.6). The first order conditions are as follows:

$$\frac{\pi_{i-1}}{(1+r)^{i-1}} = \frac{W_i G_i}{(1+r)^i} + \frac{(1-\delta_i) W_{i+1} G_{i+1}}{(1+r)^{i+1}} + \cdots + \frac{(1-\delta_i) \ldots (1-\delta_{n-1}) W_n G_n}{(1+r)^n} + \frac{Uh_i}{\lambda} G_i + \cdots + (1-\delta_i) \ldots (1-\delta_{n-1}) \frac{Uh_n}{\lambda} G_{ni} \quad (2.2.7)$$

In equation (2.2.7) $Uh_i = \frac{\partial U}{\partial h_i}$ is the marginal utility of healthy days and λ shows the marginal utility of wealth. $G_i = \frac{\partial h_i}{\partial H_i} = -(\frac{\partial TL_i}{\partial H_i})$ refers to the marginal product of the stock of health in the healthy days production. Finally, π_{i-1} shows the marginal cost of gross investment in health in period i-1. Therefore, the maximization condition implies that the present value of marginal cost of gross investment is equal to the present value of the marginal benefits.

Iterating the equation (2.2.7) for one period results with:

$$\frac{\pi_i}{(1+r)^i} = \frac{W_{i+1} G_{i+1}}{(1+r)^{i+1}} + \frac{(1-\delta_{i+1}) W_{i+2} G_{i+2}}{(1+r)^{i+2}} + \cdots + \frac{(1-\delta_{i+1}) \ldots (1-\delta_{n-1}) W_n G_n}{(1+r)^n} + \frac{Uh_{i+1}}{\lambda} G_{i+1} + \cdots + (1-\delta_{i+1}) \ldots (1-\delta_{n-1}) \frac{Uh_n}{\lambda} G_{ni} \quad (2.2.8)$$

Then equations (2.2.7) and (2.2.8) together imply:

$$\frac{\pi_{i-1}}{(1+r)^{i-1}} = \frac{W_i G_i}{(1+r)^i} + \frac{Uh_i}{\lambda} G_i + \cdots + (1-\delta_i) \frac{\pi_i}{(1+r)^i} \quad (2.2.9)$$

Multiplying the both sides of equation (2.2.) with $(1+r)^i$ gives:

$$G_i \left[W_i + (\frac{Uh_i}{\lambda})(1+r)^i \right] = \pi_{i-1}(r - \tilde{\pi}_{i-1} + \delta_i) \quad (2.2.10)$$

where $\tilde{\pi}_{i-1}$ is the percentage rate of change in marginal cost between the periods i and i-1.

Equation (2.2.10) reveals that marginal benefit from the optimum amount of investment in health should be equal to the marginal cost of investing in health. In the model both r and δ are determined exogenously. Moreover, if we assume that marginal utility of healthy days is equal to 0, equation (2.2.10) becomes:

$$G_i W_i = \pi_{i-1}(r - \tilde{\pi}_{i-1} + \delta_i) \quad (2.2.11)$$

Dividing both sides of equation (2.2.11) by π_{i-1} leads to:

$$\frac{W_i G_i}{\pi_{i-1}} = (r - \tilde{\pi}_{i-1} + \delta_i) \qquad (2.2.12)$$

Therefore, with zero marginal utility of healthy days assumption, Grossman determines the optimal choice of investment in health by equation (2.2.12). Thus, optimal choice of investment in health depends on individual's wage rate, marginal product of the stock of health in the production of healthy days, marginal cost of gross investment in health, interest rate and depreciation.

After explaining how the optimal amount of investment in health is determined, we now turn to the implications of the Grossman's model. First, he argues that health capital is different from education in the sense that while education determines the individual's time productivity (such as wages), the stock of health determines the total amount of productive time an individual uses freely. In Grossman's framework, health is determined within the model (endogeneous) whereas education is taken as given (exogeneous). He concludes that education is positively associated with health capital and negatively associated with expenditures on health care.

Another important covariate that explains the demand for health in Grossman's model is the rate of depreciation on the stock of health. Although the rate of depreciation is assumed to be exogeneous, it is important to note that rate of depreciation varies with age and education. Grossman argues that depreciation rate rises with age and falls with higher levels of education. He asserts that as long as the rate of depreciation increases with age, demand for good health decreases and expenditure on medical care increases as people get older.

The third important covariate in Grossman's model is the wage rate. He finds a positive correlation between wage rate and demand for health as well as demand for health care. In his model, wage rate acts as intermediary between health and education: Higher levels of education enhance the wage rate and wage rate improves the quality of individual's health capital.

2.3. Extensions on Health Demand Theory

Bolin (2011) is one of the recent papers that extends Grossman's model. He forms the life-time utility maximization problem in a continuous time setting. Bolin forms individual's utility as a function of H (health capital) and Z(consumption) according to the utility function:

$$U(H, Z) \tag{2.3.1}$$

where the utility function shows a positive and diminishing returns in both health and consumption. The commodity Z is a market good. Bolin also assumes that in household production market goods and own time is converted to commodities. For example, variety of foods are market goods whereas a meal is a commodity.

Moreover, Bolin formulate the cost function related with the health investment production as:

$$\mathrm{C} = \mathrm{C}(\mathrm{w}(\mathrm{E}), \bar{p}; \mathrm{E}) = \pi(\mathrm{w}(\mathrm{E}), \bar{p}; \mathrm{E}).\mathrm{I} \tag{2.3.2}$$

In equation (2.3.2) E refers to the education level, w is the wage rate which is determined by education level, $\bar{p}$ is the vector of prices of market goods used in health production, and π refers to the one-unit cost of producing I. In the model wage rate is assumed to be increasing in educational capital. In addition, Bolin assumes that the individual's cost of production of health capital decreases with education. Formally this implies that:

$$\frac{\partial w}{\partial E} \geq 0 \; and \; \frac{\partial C}{\partial E} = \frac{\partial \pi}{\partial E}.\mathrm{I} \leq 0 \tag{2.3.3}$$

In addition, the stock of health depreciates over time. δ(t) refers to the rate of depreciation. It is assumed that the rate of depreciation increases with age, which implies that δ'(t)>0. That is, if the health stock is equal to H_t at time t, then the amount of health which depreciates at time t is equal to δ(t).H_t

In the model, the individual's objective is to maximize the sum of utilities over the life-cycle. The individual values current utility different from future utility, which is called as "time preferences", which specify the relative value of current utility in relation to future utility. When time preferences are taken into account, the current value of utility accruing at time t, is:

$$e^{-\rho t}.\mathrm{U}(H_t, Z_t) \qquad (2.3.4)$$

where ρ refers to the rate of time preferences. Individual's life-time optimization problem is constrained with the changes in the stocks of health and wealth, which are defined by the following equations:

$$\dot{H}_t = I_t - \delta_t . H_t \qquad (2.3.5)$$

$$\dot{W}_t = rW_t + w_t.(T - \tau_t{}^s) - C_t - p.Z_t \qquad (2.3.6)$$

Equation (2.3.5) is the equation of motion of health capital and equation (2.3.6) is the equation of motion of wealth. In equation (2.3.6), p refers to the price of the consumption of the commodity. $\tau_t{}^s$ refers to the individual's sick time, r is the opportunity cost of investing in health rather than in the capital market and W_t is the wealth of the individual.

By solving the individual's utility maximization problem Bolin derives predictions from the model and explains how education, age and wage rate affect the individual's health level. He argues that education influences the demand for health in two ways. Firstly, education enhances household production efficiency (the efficiency effect), secondly education increases the cost of each unit of own time used in household production, because education causes rise in market productivity and hence the wage rate (the time-price effect). Efficiency effect decreases the marginal cost of producing health capital since fewer resources are used to produce a certain quantity of gross health investments. Therefore, efficiency effect increases the demand for health. On the other hand, the time price effect causes a decrease in health demand because marginal cost of health capital increases due to a higher unit cost of own time. However, the time price effect cannot outweigh the efficiency effect, or the two effects cannot completely offset each other, since individual's own time is not the only input to health production. In other words, the positive effect of education always dominates its negative effect.

For wage rate, Bolin states that a higher wage rate increases the value of available healthy time, therefore as wage increases, the incentives for being healthy strengthen. He also points out that higher wage rate makes own time used for producing gross investments in health more expensive. A higher unit cost of own time will increase the marginal cost of health capital which leads to a decrease in the demand for health. As in the education case, positive effect of wage rate on health always dominates the negative effect of wage rate on health since individual's own time is not the only input that produces health.

Regarding with the age, Bolin argues that as age increases, the possibility of better health decreases. As the rate of depreciation increases over time, Bolin's model also predicts that health decreases with age. Because the rate of depreciation increases with age, Bolin finds that the equilibrium amount of health for the old individual is lower than the health for the young individual.

Cutler et al. (2011) extend Grossman's model's predictions by considering the relationship between socioeconomic status (SES) and health. They state that different measures of SES may function through different mechanisms. For instance, short-term positive fluctuations in income reduce health, whereas long-term measures of income and wealth are positively associated with health. They argue that these fluctuations in health due to socioeconomic status will not be observed if education is uses as the only measure of SES since education is likely to remain stable throughout adulthood. They argue that components of SES should also include financial resources, rank, race and ethnicity as well as education4. They observe that mortality

[4] In the study, Cutler et al. (2011) state that "Rank" includes 4 different occupation groups. These are: Managerial, Professional and Executive Occupations, Technical, Sales and Administrative Occupations, Service Occupations, which consist of protective services, household services, and precision production, among others; Manual labor, which includes machine operators, fabricators, and various other laborers.

rates increase among individuals with lower ranks and with lower financial resources.

The covariates, which explain the variations in health outcomes, discussed in previous paragraphs are not health related behaviors. In the theory, health related behaviors are also used as input for health production. In Grossman's both 1972 and 2000 health demand models medical care is the only health input. However, as Grossman (2000) suggests, it is oversimplification because other market goods and services such as housing, diet, recreation, cigarette smoking and excessive alcohol use also influence health. Grossman states that smoking and excessive alcohol consumption have negative marginal products in the production of health. However, they are purchased since these risky behaviors may have positive marginal products in producing some commodities such as "smoking pleasure". Nevertheless, in the empirical analysis we do not prefer to use health related behaviors as inputs for other health outcomes due to two-way causality between health behaviors and health outcomes, which may create endogeneity problem.

2.4. Empirical Specification based on Theoretical Framework

In this thesis, we form our health production function as follows:

$$H=f(E,A,G,P,M,L,HI) \qquad (2.4.1)$$

In equation (2.4.1), H refers to health outcome. H is a function of education (E), age (A), gender (G) the place where the individual lives (P), marital status of the individual (M), labor market status (L) and the household Income (HI). For simplicity, education is taken as exogenous in our health production function[5]. Grossman (2004) states that formal school completed is the most important determinant of good health. This finding is valid for different health outcomes such as mortality rates, morbidity rates, self-evaluation of health status, or psychological indicators of good health.

[5] We make our empirical analysis for all health outcomes by assuming education both exogenous and endogenous, which will be discussed in detail in the following chapters. In that chapter we assume that education is exogenously affect individual's health as in Grossman (1972).

As the theory supports that education is positively associated with good health, we assume that health behaviors that improve the quality of individual's health (Fruit and Vegetable Consumption, Exercising, Normal Ranges of BMI) are increasing in education, whereas health behaviors that weakens the individual's health quality (Smoking, Alcohol Consumption, High Ranges of BMI) are decreasing in education. We also assume that level of the individual's Self-Assessed Health (SAH) increases with education.

We include age in our health production function as the theory suggests that health is a capital that depreciates over time. Both Grossman (1972) and Bolin (2011) indicate that rate of depreciation decreases with education level and increases with age. Therefore, other things being equal we assume that H is increasing in A if the H refers to risky health behavior ($\frac{\partial H}{\partial A} > 0$). Similarly, we assume that H is decreasing in A if H refers to SAH or health behaviors that are beneficial for the individual's health when other factors remain constant

We do not observe individual's wage rate from the data set we use in empirical analysis. Therefore, we include individual's household income as a proxy for the individual's income. We also include individual's employment status in our model as a proxy for the socioeconomic status (SES) of the individual. As the theory suggests that higher income levels lead to better health status, we assume H is decreasing in HI if the health behavior is risky $\left(\frac{\partial H}{\partial HI} < 0\right)$, similarly H is increasing in HI if H refers to SAH or health behaviors that are beneficial for the individual's health when other factors remain constant. We assume that the SES of the employed individuals is higher since employed individuals earn their own money and have more social networks than unemployed or inactive individuals. Cutler et al. (2011) argue that low SES in occupation leads to psychosocial stress because of subordination feelings and lack of control. They state that this stress causes deterioration in health. Hence, we expect that the occurrence of risky health behaviors decreases with higher employment status. If we assume L is a number that increments if the individual finds a job or finds a better job if he/she is already employed then H is decreasing in L if H corresponds to risky health behavior $\left(\frac{\partial H}{\partial L} < 0\right)$.

Finally, we include the gender of the individual, the place where the individual lives, and individual's marital status as covariates that explain the variation in health outcome and health related behaviors in our empirical analysis. These covariates are not examined theoretically, but they are examined empirically several times. Regarding gender, the literature suggest that females have better health status than males (Case and Paxson (2005), Fuchs (2004)). Hence, if we define "G" as 1 if the individual is male, and "0" if the individual is female, then we assume that H is increasing in G if H corresponds to risky health behavior $\left(\frac{\partial H}{\partial G} > 0\right)$.

In addition, empirical studies conclude that people living in more-favored places have better health statuses than people living in less-favored places (Reijneveld (2002)). In THS data set we observe whether the individual lives in urban or rural areas. We assume that living in urban areas are more preferred by individuals in Turkey. Hence, if we define "P" as 1 if the individual lives in urban area, and "0" if the individual lives in rural area, then we assume that H is decreasing in P if H corresponds to risky health behavior $\left(\frac{\partial H}{\partial P} < 0\right)$.

For marriage, existing literature suggests that in general, married individuals are healthier than those who are not married since having a spouse is assumed to make positive contribution to an individual' health (Fuchs 2004). Therefore, we assume health outcome is a function of marital status (M). If we assume M is a number that raises if the individual is married then we can say that H is decreasing in M if H corresponds to risky health behavior $\left(\frac{\partial H}{\partial M} < 0\right)$.

2.5. Conclusion

This chapter investigates the theoretical models for health demand theory in order to understand why the covariates that we use in the empirical analysis are crucial to explain the variations in health outcomes. We first review the findings of the Grossman (1972) model, and then we focus on the Bolin's (2011) model, which is the most recent health demand model. The implications of both models are more or less the same. Both models imply that health deteriorates with age and improves with education and wage rate. In addition, Cutler et al. (2011) include the individual's socioeconomic status (SES), which depends on the individual's race, ethnicity and financial resources, as the determinant of good health. The authors argue that higher SES leads to better health.

In Section 2.4., we tell how we establish our empirical framework based on the theory. In the empirical analysis, we consider education, gender, age, region, marital status, labor market indicators and household income as the determinants of individual's health outcomes. In this setting, all the covariates are taken as exogenous. However, it is important to note that there can be strong endogeneity problem between health and education as well as household income and labor market indicators. In fact, in the literature, the main criticism to Grossman's model is the exogeneity assumption for education. The main concern of this thesis is to investigate the relationship between health and education and detect the endogeneity problem between the two and solve it by using an exogenous instrument for education. However, we cannot find instrument for other endogenous covariates due to the limitations of the data set. Therefore, in some empirical specifications, we omit the covariates that may possibly be endogenous to health outcomes except education. In those specifications, we only concentrate on the impact of education as well definitely exogenous covariates, such as gender and age, on health outcomes.

CHAPTER 3

DATA

In this thesis, Turkish Health Survey (THS) data sets for the years 2008, 2010 and 2012 are used for empirical purposes. The THS is cross sectional data set over individuals. The THS is prepared and conducted by Turkish Statistical Institute (TURKSTAT). We pool the observations for three years for the empirical analysis since we do not observe significant variation in the determinants of health outcomes for three years. In THS questionnaire, questions related with individual's demographic factors (such as gender, region, marital status), education level, labor market indicators, household income and health related questions are asked separately for 3 different age groups, 0-6, 7-14 and 15 or above. In the empirical chapters, we examine the validity of education gradient of health in Turkey for the individuals who are 25 or above, since approximately at age 25 individuals complete their schooling in Turkey. We assume that with the restriction of sample to individuals who are 25 or above, the understanding of the exact effect of education on individual's health outcomes will be more appropriate. THS is a rich micro data set which consists of 46473 observations for the individuals who are 25 years old or older. The pooled data set includes 21015 observations for men and 25458 observations for women.

First, in THS we are able to observe the demographic factors of the respondents such as age, gender, marital status, region (urban/rural) as well as education, labor market indicators (employed, unemployed, inactive) and household income. In THS age is divided into nine groups: "0-6, 7-14, 15-24, 25-34, 35-44, 45-54, 55-64, 65-74 and 75+". In addition to these age categories, we observe the individual's self-reported age in the survey year. In the empirical analysis, we use the individual's reported age as the age variable instead of taking the mid-points of each age category so that we can see the age effect on health outcomes more accurately. In addition, single age allows us to calculate the birth year of the individual by subtracting the reported age from the survey year. To have information about the individual's birth year is important for the empirical analysis, since, we form the instrumental variable based on the individual's birth year, which will be discussed in detail in chapter five.

Regarding marital status, in THS, we observe whether the individual is married, single, widowed or divorced. In the analysis, we categorize the marital status of the respondents into 3 groups: Married, single and widowed/divorced. The respondent is referred as widowed if his/her wife/husband is dead and he/she is considered as divorced if he/she is legally separated from his/her wife/husband. Since widowed and divorced individuals have similar history (are married before but now live alone) we combine these two subgroups and form one group.

In THS, the respondents are asked whether they are working in a job or not. If the respondent does not work in a job, then he/she is asked why he/she is not working. There are eight different answers to this question, they are: "*Looking for work, Seasonal Worker, Enrolled in Education, Housewife, Pensioner, Has other Income Sources, Unable to Work* and *Other Reasons*". Based on these answers in this thesis, the employment statuses of the individuals are classified into 3 groups: Employed, Unemployed and Inactive. The respondents who are working in a regular job are referred as "employed", whereas the individuals who are not working but seeking for job are considered as "unemployed". The respondents who are considered as inactive are seasonal workers, students, housewives, pensioners, the individuals who have other income resources and the individuals who are unable to work.

Next, in THS the respondents are asked to tell or (if they do not know) guess their monthly net household income. The income categories are divided into 11 groups. These are: "*Less than 350 Turkish Liras (TL), 351-500 TL, 501-620 TL, 621-750 TL, 751-900 TL, 901-1100 TL, 1101-1300 TL, 1301-1700 TL, 1701-2300 TL, More than 2301 TL, I do not want to answer this question*". We take the mid-points of each household income group and determine the individual's household income accordingly. In the empirical analysis we take the logarithm of the household income. The individuals who do not disclose their household income constitutes 1 per cent of the pooled sample. Nevertheless, the shortcoming of this variable is that the stated monthly household incomes are really low. In addition, the individuals who have monthly household income greater than 1700 TL belongs to highest income groups in our data set although it is not a very high income level in Turkey. Last, the individuals whose household income is 2301 TL and, for instance, 10000 TL are included in the same income group although there is a huge gap between the two. Therefore, we may not correctly determine the effect of very high levels of household income on individual's health outcomes in the empirical analysis. Nevertheless, the variations in health outcomes with these levels of reported incomes give some idea about the relationship between the individual's health and household income.

In THS data set, education is classified into 6 groups: Illiterate, Not illiterate but do not finish school(we simply define it as "Non-Graduate"), primary school graduate, middle school graduate, high school graduate and has university or higher degree. Therefore, although we do not observe the years of schooling completed by the individuals, we are able to calculate an average level of years of schooling. The average years of schooling for a certain degree is approximated by the minimum years of schooling for that degree. Hence, it is equal to 5 for primary school graduate, 8 for middle school graduate, 11 for high school graduate and 15 for the individuals who have university or higher degree. The average years of schooling is taken as 0 for illiterate individuals and 2 for non-graduates.

In this thesis, the health outcome of the first empirical analysis is the individual's Self-Assessed Health (SAH) which, in THS, is determined by the answer of the question:

"In general, how is your health status?"

with respondents choosing one of the outcomes: "*Very Good, Good, Average, Poor, Very Poor, Does not know, Does not want to answer this question*". We coded the SAH of the individual in the following manner: If the individual reports very good health, then his/her SAH is coded as 5, if he/she states good health then his/her SAH is coded as 4. If he/she reports poor and very poor health, then the individual's SAH is coded as 2 and 1 respectively.

The respondents who state that their health is on average are taken as they are in "Average" health (Bobak et al., 1998). The SAH of these individuals are coded as 3 in the data set. Idler and Benyamini (1997) define the average health with the following quote:

"It's right in the middle. It's like those shades of grays between the black and white. So I can't say everything about me is poor and not everything about me is excellent."

Chapter four discusses both advantages and disadvantages of using SAH as health outcome. One of the shortcomings of using SAH is the fact that the individuals may misreport their health conditions. In order to show in THS data set the individual's SAH is consistent with the individual's objective health criteria, we analyze the correlation between SAH and individual's health problem by looking at the frequency distribution analysis of certain chronic illnesses by the individual's reported SAH.

In THS, objective health of the individual is measured in response to answer of the question:

> *"Do you have any disease/health problems stated below, if yes, is it diagnosed by a physician?"*

If the answer of the question is *"Yes",* then the respondent is considered to have that illness. In 2008 data set, the listed diseases/health problems in the survey are: *Asthma, Chronic Obstructive Pulmonary Disease (Chronic Bronchits, Emphysema), Heart Attack/Coronary Heart Disease/Chronic Heart Failure, High Blood Pressure (Hypertension), Stroke (Serebral Heamorrhage, Cerebral Thrombosis), Osteoartrit/Romatoid Artrit/Neck Musculoskeletal System Disorders/Back Musculoskeletal System Disorders, Diabetes, Allergy, such as rhinitis, eye inflammation, dermatitis, food allergy or other (allergic asthma excluded), Ulcer, Cirhosis of the Liver/Liver Dysfunction, Cancer, Migraine and Severe Headeche,* *Urinary Incontinance, Problems in Controlling the Bladder, Chronic Anxiety/ Chronic Depression/Other Psychological Health Problems, Aneamia (Iron Deficiency Aneamia), Permanent Injury or Defect Caused by an Accident, Sinusitis*. In 2010 and 2012 data sets, *Thyroid Disease and Hepatitis* are added into the list. Table 3.1. below reports the association between SAH and chronic illness for individuals who are 25 or above in Turkey.

Table 3.1. The Distribution of Diseases/Health Problems diagnosed by a Physician by reported SAH

SAH	Very Good	Good	Average	Poor	Very Poor	Total
Diseases/ Health Problems						
Asthma	36 (1.30)	709 (25.53)	1153 (41.52)	757 (27.26)	122 (4.39)	2777 (100.0)
Chronic Obstructive Pulmonary Disease (Chronic Bronchits, Emphysema)	45 (1.81)	598 (24.00)	1029 (41.29)	692 (27.77)	128 (5.14)	2492 (100.0)
Heart Attack/ Coronary Heart Disease/ Chronic Heart Failure	27 (0.72)	706 (18.77)	1687 (44.86)	1148 (30.52)	193 (5.13)	3761 (100.0)
High Blood Pressure (Hypertensi on)	92 (1.07)	2072 (24.07)	4054 (47.09)	2053 (23.85)	338 (3.93)	8609 (100.0)
Stroke (Serebral Heamorrhag	4 (0.61)	69 (10.44)	234 (35.40)	264 (39.94)	90 (13.62)	661 (100.0)

e, Cereb
Thrombo

Table 3.1. (Continued)

Osteoartrit/ Romatoid Artrit/Neck Musculoskeletal System Disorders/ Back Musculoskeletal System Disorders	265 (1.96)	4104 (30.34)	6009 (44.43)	2719 (20.10)	429 (3.17)	13526 (100.0)
Diabetes	47 (1.13)	859 (20.71)	1970 (47.49)	1103 (26.59)	169 (4.07)	4148 (100.0)
Allergy, such as rhinitis, eye inflammation, dermatitis, food allergy or other (allergic asthma excluded)	91 (4.12)	807 (36.57)	875 (39.65)	370 (16.76)	64 (2.90)	2207 (100.0)
Ulcer	111 (2.04)	1688 (30.96)	2380 (43.65)	1097 (20.12)	177 (3.25)	5453 (100.0)
Cirhosis of the Liver/Liver Dysfunction	3 (0.90)	56 (16.72)	126 (37.61)	122 (36.42)	28 (8.36)	335 (100.0)
Cancer	4 (1.00)	76 (19.05)	139 (34.84)	139 (34.84)	41 (10.28)	399 (100.0)

Table 3.1. (Continued)

Migraine and Severe Headeche	81 (2.38)	1106 (32.48)	1416 (41.59)	681 (20.00)	121 (3.55)	3405 (100.0)
Urinary Incontinance, Problems in Controlling the Bladder	10 (0.75)	173 (13.01)	560 (42.11)	461 (34.66)	126 (9.47)	1330 (100.0)
Chronic Anxiety/ Chronic Depression/ Other Psychological Health Problems	32 (1.33)	576 (23.92)	1103 (45.81)	579 (24.04)	118 (4.90)	2408 (100.0)
Aneamia (Iron Deficiency Aneamia)	67 (1.88)	1228 (31.73)	1512 (42.53)	730 (20.53)	118 (3.32)	3555 (100.0)
Permanent Injury or Defect Caused by an Accident	21 (3.36)	199 (31.84)	212 (33.92)	163 (26.08)	30 (4.80)	625 (100.0)
Sinusitis	138 (4.94)	1121 (40.11)	1082 (38.71)	396 (14.17)	58 (2.08)	2795 (100.0)
Thyroid Disease	35 (2.27)	520 (33.70)	662 (42.90)	273 (17.69)	53 (3.43)	1543 (100.0)
Hepatitis	21 (5.33)	181 (45.94)	126 (31.98)	54 (13.71)	12 (3.05)	394 (100.0)

Source: Turkish Health Research Survey, 2008, 2010,2012 pooled data

Notes: 1. Percentages are in parenthesis.

2. The data on Thyroid Disease and Hepatitis are not available for 2008.

Table 3.1. reports that in general, the respondents tend to report average health although they have chronic health problems, which are considered as serious, such as thyroid, anemia, depression, ulcer, diabetes, etc…The percentage of reporting good health compared to average health are higher for the individuals who have sinusitis

or hepatitis problems. The prevalence of reporting poor health is seen higher among individuals with serious health problems, such as stroke, cancer and cirrhosis of the liver. For the other diseases, the respondents report that their health is on average. The percentage of reporting very good or very poor health is very low for all diseases.

The respondents who declare that their health is on average are considered to have health which is something between good and bad (Burstörm and Fredlund, 2001 and Bobak et al. 1998). In fact, the perception of health depends on the respondent's psychology, the frequency of feeling the pain, and whether the disease/health problems limit the daily activities or not. Therefore, on the one hand, we can state that most of the detected illnesses do not cause individuals to perceive their health is poor, on the other hand we observe that most of the individuals do not state that their health is good or very good when they have diagnosed health problem. Thus, although we cannot confidently say that there is strong positive association between SAH and objective health measures, the distribution of detected disease/health problems by SAH suggest that the linkage is not so weak.

In the second empirical part of this thesis, the health outcomes are different health behaviors such as smoking, alcohol consumption, fruit and vegetable consumption, exercise and the body mass index (BMI). First, whether the individual is regular smoker or not is determined by the respondent's answer to the following question:

> *"Were you a regular tobacco consumer and do you still consume tobacco?"*

If the response to this question is, *"Yes, everyday"* or *"Yes, sometimes"*, the respondent is considered as smoker. In the empirical analysis, for smoking behavior, we define a discrete variable which is equal to 1 if the individual is currently a regular smoker and 0 if he/she does not smoke. Similarly, for alcohol consumption behavior, the following question is asked:

> *"Did you consume alcohol regularly or occasionally and do you still consume alcohol?"*

If the response to this question is *"Yes, everyday"* or *"Yes, sometimes"*, the individual is considered as an alcohol consumer. In the empirical model, for alcohol consumption behavior, we define a discrete variable which is equal to 1 if the individual consumes alcohol regularly or occasionally and 0 if he/she does not consume alcohol. In fact, alcohol consumption is a risky health behavior if the individual is heavy-drinker, which is defined as consuming 5 or more drinks when an individual drinks (Cutler and Lleras-Muney, 2010). However, in THS data set we cannot observe how many drinks the respondent has when he/she drinks. Rather, we can observe how often an individual consumes alcohol in the last 12 months. The responses for this are: "*Never, I had used to but now give up, once a month or less, twice or four times a month, twice or three times a week, four times or six times a week, everyday*". We consider the respondents who say that they consume alcohol everyday as heavy drinkers. However, in THS the proportion of daily alcohol drinkers is very low, less than one percent (0.5 per cent, 0.4 per cent and 0.2 per cent in 2008, 2010 and 2012 respectively)[6]. Therefore, for alcohol consumption behavior, we combine the daily and occasional alcohol drinkers and call them as "alcohol drinkers" in the empirical analysis. Since some alcohol consumption is even good for individual's health (Kenkel, 1991) we consider alcohol consumption as health behavior, rather than risky health behavior throughout the analysis.

Regarding fruit and vegetable consumption, in THS there are three different questions. For fruit consumption behavior the following question is asked:

"How often do you consume fruit (Except Fruit Juice)?"

Similarly, for vegetable consumption, the following question is asked:

"How often do you consume vegetable or salad (Except Vegetable Juice or Potato)?"

Finally, for fruit or vegetable juice consumption, the following question is asked:

"How often do you consume fruit or vegetable juice?"

[6] It is noteworthy to state that the respondents may also misreport their alcohol consumption related behavior due to religious beliefs in Turkey, which prohibit alcohol consumption.

The answers for these questions are all the same. They are: "*Twice a day or more, Once a day, Less than once a day but at least four times a week, Less than four times a week but at least once a week, Less than once a week, Never, Does not know, Does not want to answer this question*"

The individual is referred as regular fruit and vegetable consumer if he/she states that he/she consumes fruits and/or vegetables and/or fruit and vegetable juice at least once a week. In the empirical analysis, for fruit and vegetable consumption behavior, we define a discrete variable which is equal to 1 if the individual consumes fruits, vegetables and/or their juice at least once a week and 0 otherwise.

In THS we also observe the respondent's frequency of exercise. There are three different questions for exercising. For heavy exercise, the question is as follows:

"How many days did you make heavy body exercise during the last 7 days?"

Similarly, for medium level of exercise, the question is:

"How many days did you make medium level of exercise during the last 7 days?"

Finally, for walking, the question is:

"How many days did you walk at least 10 minutes during the last 7 days?"

The answers for these questions are all the same. If the individual makes one of the exercises above, he/she reports the number of days that he/she exercises. Otherwise, he/she gives the following answers: "*Never, Does not know, Does not want to answer this question*".

In the empirical analysis, we define a discrete exercise variable which is equal to 1 if the individual exercises at a high level or medium level or low level (walking), for at least 10 minutes in a week and 0 if he/she gives the answer "Never" to all of the three questions. In the data set, both physical activities and working in construction sector are referred as heavy exercise. Similarly, both houseworks (such as cleaning the house) and playing tennis are included in the same category (medium level of exercise). Making sports activities directly lead to healthy life and generally the respondents who have higher education levels or who belong to higher income groups tend to enroll in these activities more than the others. On the contrary, people who are illiterate or less educated or who belong to lower income families work in construction sector or clean someone else's houses to earn income. Unfortunately, we cannot observe what kind of activity the individuals are making from the data set. Therefore, the variation in exercising behavior with education may be biased.

Finally, in THS data set the respondent's self-reported height (in centimeters) and weight (in kilograms) are available. Therefore, we are able to calculate the body mass index (BMI) of the individual. We calculate the individual's BMI in the following manner: First, we convert the height into meters by dividing the reported height by 100. Then we divide the reported weight of respondent (in kilograms) to the square of the height in meters. The resulting number gives the BMI of the individual.

In short, THS is a rich data set which consists of several health related questions that give idea about the respondent's general health conditions. Apart from the variables that we use for the empirical analysis, there are more health related questions such as the questions about the psychology of the respondents, how often a respondent makes use of preventative health care services, or whether the respondents have accident recently, etc…These questions can be used for further health researches in Turkey. Health related questions for the children below 15 years of age are not as gorgeous as the questions for adults and they are answered by the parents of children. The most important deficiency of THS data is the fact that the answers to all the questions are self-reported, even for the objective health criteria questions. Therefore, some of the results may be biased. However, the investigations of the variations in health outcomes are usually done with this kind of data sets in the literature and THS is the only richest data set in Turkey that allows for conducting health related empirical studies.

CHAPTER 4

SELF –ASSESSED HEALTH STATUS (SAH) AND ITS DETERMINING FACTORS

4.1. Introduction

Chapter four provides a descriptive analysis that shows how the individual's Self-Assessed Health Status (SAH) varies with the covariates that we discuss in chapter two. These covariates are education, demographic factors such as age, gender, region, marital status; individual's labor market indicators and household income. We explain why we use these covariates in order to explain the variations in SAH in detail in chapter two[7].

There are several studies which support that it is advantageous to use SAH as health outcome. These studies are, Idler and Benyamini (1997); Miilunpalo et al.(1997); Burstörm and Fredlund (2001); Benyamini et al. (1999); Farmer and Ferraro (1997) and Ferraro et al. (1997). However, there are also some studies which indicate SAH may be a misleading health outcome especially due to measurement errors. These studies which discuss the disadvantages of using SAH as health outcome are Butler et al. (1987), Baker et al., (2004) and Crossley and Kennedy (2001). Since the answers to the questions in THS data set is self-reported, there may also be likelihood of misreporting the SAH in THS data set. Nevertheless, In Chapter three, we show that the linkage between SAH and detected disease/health problems is not so weak. Therefore, we assume that SAH is a good proxy for morbidity in our study.

[7]We cannot observe individual's life expectancy or mortality in THS data set. Therefore, following Idler and Benyamini (1997), we assume SAH is a good predictor of mortality and morbidity, therefore we assume SAH is assumed to be the main health outcome throughout this thesis. Hence, we devote a whole chapter only to the descriptive statistics in variation in SAH. The descriptive statistics for the other health behaviors will be presented in Chapter 6 preceding the empirical results on them.

This chapter is important in the sense that it is the first study that examines the SAH and its variations in Turkey, a middle income, developing country. The determinants of SAH are mostly investigated for developed countries, especially for USA in the literature. There are few studies done for developing countries on this subject. These studies include Yen et al. (2010) for China, Vaillant and Wolff (2012) for Albania and Zimmer et al. (2000) for Taiwan, Thailand and Phillipines.

In this chapter, we provide descriptive analyses of SAH for individuals who are 25 years old or older in Turkey by using Turkish Health Survey data set for years 2008, 2010 and 2012, together with a detailed literature review. We report on how SAH varies by education, demographic factors such as gender, age, region, marital status, labor market indicators and household income. Descriptive statistics suggest 59.28 per cent of the individuals who are 25 years old or older report that their health statuses are good or very good in Turkey. We also observe that males report that their SAH is good more often than females although life expectancy of females is higher than that of males. In addition, we find out that age is negatively associated with SAH; older individuals are more likely to report poor SAH. We also observe that people living in urban areas report good SAH more often than people living in rural areas.

For marital status, our results suggest that the percentage of reporting poor SAH is the highest among widowed/divorced females, and lowest among single males. Moreover, we find that education and household income level is positively associated with SAH as the theory implies. For employment status, frequency distributions indicate that employed and unemployed individuals state good SAH more often than inactive individuals.

The rest of the chapter is organized as follows: Section 4.2. makes a brief literature survey on SAH. Section 4.2.1 examines the literature that discusses the advantages and disadvantages of using SAH as health outcome. Then in Section 4.2.2, we provide brief summary of studies that are on developing countries. Next, in Section 4.3 we report the frequency distribution analyses that show the variations in SAH by each determinant of SAH together with a detailed literature review. Finally, Section 4.4 provides the concluding remarks.

4.2. Literature Survey

4.2.1. SAH as Health Outcome

SAH is the health outcome that is most frequently used in health economics literature. SAH has various advantages over the other health outcomes. First, SAH is an indicator of future coming disease of the individual. For instance, a person may not have a diagnosed illness but has symptoms of it, hence the person may report that his/her health is not good although he/she does not have a detected disease. Second, SAH gives idea about the severity of the disease, because, the symptoms of the disease vary over time and between individuals. For example, two individuals may have the same disease, but one of them is at the early stages of the disease, then his/her reported SAH will differ from the other, although they have the same detected illness. Third, SAH reflects family history, as the individual knows the genetic risk factors of his/her family. Next, SAH reflects the dynamic evaluation of health of the individual, not only the current level of health. It is also argued that SAH influences the individual's health related behaviors as people with poor SAH engage in less preventative health care activities. Last, SAH reflects the person's psychological condition as people with depression tend to report poorer SAH. (Idler and Benyamini, 1997).

The literature on SAH mostly focuses on the positive association between SAH and mortality and morbidity. One of the most important study is done by Idler and Benyamini (1997). In that study, the authors review 27 different studies which suggest positive association between SAH and mortality. These studies are done for different developed countries. They conclude that the strong positive relationship between SAH and mortality is universal and it is still valid when other covariates of health are controlled.

Miilunpalo et al. (1997) conduct another study that examines the relationship between SAH and mortality for Finland, a developed country. Miilunpalo and his friends find that SAH is a valid health status indicator and it is a good predictor of mortality. They use longitudinal census data for Finland in their analysis. Thus, they are able to examine the stability of SAH about predicting the mortality. Their results suggest that SAH is a good predictor of mortality throughout the follow-up periods. Therefore, they conclude that SAH is a stable measure of mortality. Miilunpalo et al. also suggest that SAH is easier to measure than clinical measures of health and to conduct standardized self-reported surveys of health status is less expensive for the researcher.

For another developed country, Sweden, Burstörm and Fredlund (2001) examine the relationship between SAH and mortality by using Swedish Survey of Living Conditions data set between the years 1975 and 1997. The objective of this study is to analyze the predictive power of SAH for mortality in different socioeconomic groups. They look at the association between reporting poor SAH and ratio of mortality for the individuals above 16 years old. They find that the relationship between poor SAH and mortality is very strong and it is even greater at younger ages. Their results suggest that the relationship is similar among men and women as well as among individuals with and without a chronic illness. They suggest that the linkage between SAH and subsequent death is stronger in higher socioeconomic groups.

Benyamini et al. (1999) also confirm the relationship between SAH and mortality. Their results indicate that relevant factors that affect SAH, such as function and lack of energy, predict five-year mortality. However, they find that if these factors are controlled, the association between SAH and mortality decrease. They also find that psychological factors that affect SAH, such as anxiety and depression do not predict mortality.

Ferraro et al. (1997) investigate the relationship between negative SAH and morbidity as well as physical disability by using National Health and Nutrition Examination Survey I (NHANES I) data set for the USA for both black and white adults who are above 15. They also investigate whether negative SAH lead to a cycle of health decline over time. Their results confirm the negative association between

SAH and health problems. The negative association is seen in blacks more frequent than whites. Later on, Farmer and Ferraro (1997) examine the negative relationship between SAH and functional disability by considering stress factors. The authors also use NHANES I survey data set in this study. They find that distress causes more negative health perceptions. They also conclude that SAH influences the distress levels of the individual. Their results suggest that negative SAH is certainly related to morbidity even if stress factors are controlled.

There are some opponents of using SAH as health outcome due to different reasons. For instance, Quevedo et al. (2005) suggests that the individuals' reported SAH may change if the questionnaires change. They called this bias as "Reporting Bias". They show how the researchers can face with conditions of reporting bias by using panel data ordered probit and generalized ordered probit models. In addition, they point out the fact that representation of actual health into SAH categories may vary with respondent's characteristics. For instance, the person can evaluate his/her health in an optimistic way. Therefore, an old and unhealthy individual may compare his/her health status with other individuals who belong to the same age group and who are in poorer health condition. In this case, this old individual may report very good SAH although he/she is not.

Crossley and Kennedy (2001) do not support the idea that SAH is a proxy for actual health either. In their paper, they show that the responses of the individuals may change with different type of surveys. Crossley and Kennedy conduct an experimental study and apply both "form-based" and "face to face" questionnaires to randomly selected groups. They find that the same individual may give different answer about his/her health in different types of surveys. The authors state that this could be happen because of three reasons: First, the individuals may report their health with some error in the form based survey. Second, the individuals may answer more honestly to sensitive questions when they fill the forms compared to being personally interviewed. Third, the respondents may learn about their SAH in the period between the form-based questionnaire and personal interview.

Butler et al. (1987) state that the SAH of the individuals may be misleading due to employment conditions. They find out that that people who are out of labor force tend to report misleading results about their SAH. Therefore, they conclude that the individuals who are not working tend to report their health incorrectly owing to social pressure to justify not having job. In contrast, Bound (1991) conducts another study that describes the labor market decisions of males older than 50 years old. His findings suggest that objective health measures do not explain the differences in labor force participation decisions of elderly men better than the self-reported health measures. That is, he claims that objective health measures do not have priority over self-reported health measures.

A very famous study that discuss the measurement problems related to SAH is conducted by Baker et al. (2004). In that study, Baker and his friends look at the relationship between SAH and more objective status of health by using Canadian National Population Health (NPHS) Survey as well as the data taken from Ontario Health Insurance Plan (OHIP) for 1996-1997 period. NPHS includes self-reported health measures such as: SAH, specific work and activity restrictions, the self-reported occurrence of specific illnesses, whereas OHIP includes the individuals' objective health status measures. For the objective health status measures, the authors use the data taken from OHIP, which is a panel data set that reveals the health statuses of the individuals in survey year and in the preceding 5 years. In their study, the authors focus on the individuals who are 12 years old or over. For the analysis where they examine the labor supply decision of the individuals, they exclude the respondents who are 16 years old or younger; since they are enrolled in education. The authors' findings support the measurement errors in self-reported surveys. In the study these measurements are referred as "false negatives" and "false positives". For instance, 74 per cent of individuals who have cancer according to the OHIP data set, do not report that they have cancer, this is called as "False Negative", or 0.5 per cent of individuals report that they have cancer, although actually they do not have cancer, and this is called as "False Positive". Both false positives and false negatives are likely to appear if the individual do not want to reveal their current health conditions. After analyzing the descriptive statistics, the authors estimate the measurement errors in self-reported health statuses. They find nearly 50 per cent of NPHS data does not coincide with OHIP data. They calculate the proportion of false positives and false negatives by using Bayes rules. Next, like Butler et al. (1987) their results support the validity of the “justification hypothesis”, i. e some people may give misleading information on their health in order to justify why they do not participate in labor force. Hence, the authors conclude that false positives are highly seen among the individuals who do not work.

While Baker et al. (2004) conclude that being a non-worker results with misleading self-reported health statuses, Johnston et al. (2009) conclude that it is the income level that leads to misrepresentative self-reported health statuses. In their study, Johnston et al. examine the validity of self-reported outcomes by comparing the individual’s self-reported hypertension with clinical measures of hypertension by using Health Survey for England (HSE). The authors focus on the individuals who are 26 years old or older, since the incidence of hypertension is lower among younger individuals. They pool the HSE data set for the years between 1998 and 2003. They find that the incidence of false positives are very high, as it varies from 79 per cent to 87 per cent in different samples. They regress individual’s both subjective and objective hypertensions on income, education, age, gender, marital status, location and genetic predisposition. The authors also include employment status, obesity and measures of life style as the covariates that explain the differences in hypertension levels. They implement probit regression analysis for both self-reported hypertension (i.e. the individual is asked whether he/she has hypertension problem or not) and clinically detected hypertension levels (i.e. a threshold level of blood pressure level is detected). Their results suggest that as the individual’s income increases, his/her probability of reporting false negative decreases. Therefore, the authors conclude that income is an important factor that may lead bias in reporting the health statuses, most probably due to the fact that people with lower income levels are not able to measure their hypertensions correctly and they are not able to perceive their health conditions accurately.

4.2.2. Studies on Developing Countries

Zimmer et al. (2000) examine the determinants of SAH for three developing countries, Taiwan, Thailand and Philippines. By examining three different countries, their objective is to investigate the cross country variations in determinants of SAH. In their study, they focus on the individuals who are 50 years old or above. They use the data that come from “The Study of Rapid Demographic Change and Welfare of the Elderly”. The data contain series of surveys for Taiwan, Thailand and Philippines, which are conducted in 1996. In the empirical analysis, the dependent variable is SAH which ranges from 1 to 5. The authors use several independent variables to explain the variations in SAH. Some of the covariates they use are strongly endogenous to the individual’s SAH. Examples of those variables are the behaviors that measure the individual’s functional abilities such as walking, bathing and dressing; and chronic health conditions such as high blood pressure, diabetes and heart disease. The authors also include the socioeconomic measures and demographic characteristics such as gender, age and urban/rural residence. Socioeconomic measures they use are health care access, education and employment, and possession of the households. They do not refer household income as socioeconomic factor, since it may differ across three different countries. Next, as a measure of network, the authors include marital status, household size and living with offspring or not. Finally, the authors also include smoking and alcohol consumption (i.e. being current smoker and/or being current alcohol drinker) as health behaviors. The authors conduct ordered probit regression analysis in order to examine the determinants of SAH. The authors first control for country effects and then include the independent variables gradually. Zimmer et al. find that education is positively related to probability of reporting very good health. Their results also suggest that chronic illnesses and risky health behaviors negatively associated with reporting very good health. However, the effect of the rich set of determinants of SAH vary from country to country. Hence, the authors conclude that cultural distinction should also be taken into account in order to get more accurate measure of SAH.

Yen et al. (2010) explore the effect of being regular smoker on individual's SAH in China. Although their main objective is to test the effect of smoking on SAH of the individual, they also control for other factors such as age, province, employment status, education level, and marital status, the variables that indicate the alcohol consumption behavior of the individual, and having hypertension problem or not. They use China Health and Nutrition Survey data set for the year 2006 in their study. The authors conduct their analysis for randomly selected adult males since smoking is highly seen among males rather than females in China (48.9 per cent of males and 3.2 per cent of females are smokers by 2003). The authors first estimate a Gaussian binary treatment effect model. In that model, the authors assume that SAH is a continuous variable, rather than a categorical variable, therefore they implement an OLS regression and then, they conduct a bivariate model such that in the first regression the dependent variable is smoking (i.e. a variable that indicates whether the individual is non-smoker, heavy smoker and moderate smoker), and in the second regression the dependent variable is individual's SAH. Both of the outcomes are categorical variable, hence the authors implement "bivariate ordered probit" regression. Their results suggest that heavy smokers are more likely to report excellent health than non-smokers or moderate smokers. The authors underline the importance of the other factors that determine SAH in explaining this relationship.

Vaillant and Wolff (2012) examine the determinants of SAH for Albania, a developing country. In this study, They use a panel data set collected in Albania. The data set is collected by World Bank in between the years 2002 and 2004 as a part of Living Standard Measurements Study project. In addition to standard question for individual's SAH, different from the previous studies, there is another question like "*Compared with health one year ago, would you say that your health now is much better, somewhat better, about the same, somewhat worse, much worse?*" Therefore, the authors are able to observe whether time also plays an important role in explaining the variations in SAH. The authors conduct a random effect ordered probit model. In the empirical analysis they control several factors such as age, gender, religion, marital status, education and income. Their findings suggest that living in urban areas, having higher education degree and having higher income level improves the individual’s SAH. Since, the authors can also observe the individual’s chronic illnesses, they divide the whole sample into smaller samples according to the chronic illnesses of the individuals, and observe how SAH varies over time if the person has chronic diseases. They observe that if the person has chronic illness, she/he reports poorer SAH. They also note that the SAH of the individual may change over time. The authors find that temporary shocks do not affect the individual’s SAH, whereas permanent illnesses change the individual’s reported SAH.

4.3. Frequency Distribution Analysis

In this section we explain the variations in SAH by various covariates using frequency distribution analysis. Different empirical studies control for several factors in order to explain the variations in SAH, and they investigate how these covariates affect individual’s SAH. Although, in general, the main control variable that explains the variation in SAH is education in the literature, other control variables, especially household income and employment status also attract the attentions of many authors. Since, education is also the most relevant factor that explains the variations in SAH in this thesis, in this section we mainly focus on the literature on the relationship between SAH and education[8]. It is important to note that SAH is not the only health outcome in most of these studies, and the name of the SAH may change as self-reported health or self-rated health or self-reported poor health in different studies. In each section, first we provide a literature review on the relationship between the determinant considered and the health outcome, then we present how SAH varies by that factor in Turkey using the frequency distribution analysis. Since it is the main explanatory variable for this thesis, we first examine the variation in SAH by education level. Next, we look at the variations in SAH by gender, age, marital status, region, household income and labor market status respectively.

[8] This chapter is devoted to descriptive analysis, however in the literature survey part, we discuss the methodology that the studies use to conduct their analysis. We will use similar techniques in later chapters in order to examine the validity of the education gradient. Hence this section provides benchmark for our analysis in the following chapters.

4.3.1. SAH and Education Level

The relationship between health and education is called as “Education Gradient of Health” in the literature. Eide and Showalter (2011) examine the recent literature on the validity of the education gradient of health. In that study, the authors mainly concentrate on the direction of causality between health and education. Detecting the causality between the two is difficult due to the endogeneity problem. The studies that Eide and Showalter focus on generally implement Instrumental Variable (IV) techniques to circumvent the problems due to endogeneity.

In the first part of the paper, they review the literature which supports the direction of causality is from education to health. In general, the literature finds a positive and significant association between health and education by using IV estimation techniques. The instruments for education used in different studies are; the changes in compulsory schooling laws, the twins effect, Vietnam draft, variations in education policy (such as state-level tuition prices and per-student state expenditures on education) and family background variables, (such as parents’ education and occupation).

In that part of the paper, Eide and Showalter also discuss different methodologies apart from IV methods in order to overcome the endogeneity problem between education and health. They stress that another possible solution to the endogeneity problem, is to use propensity score matching (PSM) methods. Therefore, they also focus on the literature that uses PSM methods for investigating the relationship between health and education. The studies that use PSM methods also support the positive association between health and education[9].

In the second part of the paper, Eide and Showalter examine the literature that state the direction of the causality is from health to education. In that part of the paper, first, the authors examine the literature which investigates the correlation between

[9] The positive relationship between the college attendance and the preventative health care use is obtained by Fletcher and Frisvold (2009).

birth-weight and educational outcomes of the individual in later ages. In general, the previous studies related with this issue suggest that increase in the birth weight of an individual will lead to higher level of education in later ages relative to the other individuals who have lower birth weights.

Eide and Showalter also underline the studies which conclude that if an individual face with unpleasant events in utero, such as radioactive fallout, the probability of high school completion of this individual decreases relative to their contemporaries who lead a healthy period in utero. This finding also shows the importance of health in utero for educational outcomes in the future. In summary, Eide and Showalter look at the causality between education and health in both directions by examining various studies. Different methodologies used in these studies give idea for future researches.

Brunello et al. (2015) conduct the first study that relates education as well as risky health behaviors to self-reported poor health. In the study the authors form a health production function where the output is the self-reported poor health and inputs are the risky health behaviors and education (measured by the years of schooling). They estimate the effects of health behaviors that are most probably endogenous to the health outcome as well as education on self reported poor health for 12 European countries[10] by using IV methodology. In their study they take change in the compulsory schooling laws across countries and birth cohorts as instrument for the individual's years of schooling. In order to look at the both long-term and short-term health behaviors of the individuals, they estimate the relationship between health and health related behaviors along with education for the individuals whose ages are above 50 by using two longitudinal data sets, Survey of Health, Ageing and Retirement in Europe (SHARE) and the English Longitudinal Study of Aging (ELSA). They find that education has protective effect on self reported poor health for males and females and their results suggest that the effect for females is higher than for males. They also conclude that in the long-run health behaviors have more

[10] The countries used in the study are: Austria, Belgium, Denmark, England, France, Germany, Greece, Italy, the Netherlands, Spain, Sweden and Switzerland.

significant effects on self-reported poor health than the short-run health behavior effects.

Fonseca and Zheng (2011) conduct a similar study to Brunello et al. (2015). They also examine education gradient of health and find positive effect for the individuals who are above 50 from the 15 OECD countries[11]. The data sources in that study are SHARE, ELSA and the Health and Retirement Study (HRS) in the USA. To examine the effect of education gradient of health in these countries, the authors firstly runs variety of probit regression where the health outcomes differ. The examples of different health outcomes are self-reported poor health, and indicator variables such as whether the individual has chronic illness, cancer, heart disease, lung disease, psychiatric illness, etc. The independent variables are the years of schooling, gender, birth cohort dummies for different age groups, and country. The authors also use IV methods and take the change in the compulsory schooling laws as the instruments for education. Fonseca and Zheng conclude that there is negative association between years of education and self-reported poor health. Their results also suggest that more years of education lead to lower frequency for diabetes and hypertension. Nevertheless, the authors do not find a significant relationship between education and other chronic health conditions, such as heart disease, lung disease, stroke, arthritis and psychiatric illness. Last, both probit and IV-probit estimates suggest that there is positive association between education and probability of being cancer from both IV probit and probit results. Overall, this study is very similar to Brunello et al. (2015) with the data and the methodology they use. However, the study differs from Brunello et al.'s study in the sense that they do not include endogenous health behaviors to health production function.

Silles (2009) conducts an important study for United Kingdom (UK). In the study the author investigates whether the correlation between health and education represents a causal relation for the UK by using the data from the General Household Survey for England, Scotland and Wales. The sample that the authors use includes men and

[11] The countries that they use are USA, England, Austria, Sweden, the Netherlands, Italy, France, Denmark, Greece, Switzerland, Belgium, Germany, Czech Republic, Poland and Spain.

women who are between 25-60 years of ages. The sample is obtained from pooling cross-sections between 1980 and 2003/2004. The author uses both IV and Ordinary Least Squares (OLS) methodology to estimate the causal relationship between health and education[12]. There are various health outcomes for different models. These outcomes are: Self-reported good health, whether the individual has long-term illness or not, whether the individual has activity-limiting illness or not and whether the individual has work-preventing illness or not. Like in the previous studies, Silles uses IV methods due to the possible endogeneity problem between health and education. She also takes the change in the compulsory schooling laws in England as instrument for education. Both OLS and IV result support the positive association between education and good-health. She points out that the effect obtained from the IV results is higher than the effect obtained from the OLS results. Therefore, the author concludes that standard measures for education may underestimate the impact of education on health.

Arendt (2005) investigates the relationship between health and education for Denmark. He uses the Danish National Work Environment Cohort Study (WECS) data set of Danish Workers who have been interviewed in 1990 and 1995. The data set consists of individuals who are above 25 in 1990. He uses both a quantal response panel model and a linear regression model in order to test the effects of education on SAH of individual. In that study, SAH is graded as poor, fair, good, very good and excellent. In order to overcome the possible endogeneity problem, Arendt uses two school reforms in Denmark as instruments for years of schooling. These school reforms were established in 1958 and 1975 respectively. They form two dummy instrumental variables to address whether the individual is affected by these two reforms. Arendt's results provide evidence that education and good health has positive relationship.

Berger and Leigh (1989) is one of the earliest papers that considers the possible endogeneity problem between health and education. Although, the health outcome is not the individual's SAH level in this study, Berger and Leigh (1989) is one of the

[12] Logit and Linear Probability models are also conducted for the sake of convenience.

important papers written on education gradient of health by considering the endogeneity problem between health and education. They use both the Health and Nutrition Examination Survey and National Longitudinal Survey of Young Men data sets in their study. The health outcomes they use are the individual's disability, functional limitations, systolic and diastolic blood pressures. Berger and Leigh propose Two Stage Least Squares (2SLS) estimation techniques in order to detect the relationship between health and education consistently. In the first stage regression the dependent variable is the individual's completed years of schooling and the independent variable is the individual's initial health status as well as other demographic factors such as gender, age, household size, whereas in the second stage regression, current health status is regressed on individual's years of schooling that is estimated from the first stage equation as well as the initial health status and other factors that are assumed to affect the individual's current health status. Their results suggest that the effect of years of schooling is higher than the effect of unobservables[13] on individual's health outcomes.

Likewise, Llears-Muney (2005) does not examine the relationship between SAH and education, rather she investigates the relationship between mortality rates and education level for the USA. Since SAH is believed to be a good predictor of mortality, and since Lleras-Muney's this study is one of the pioneers in education gradient of health literature, we summarize this study briefly in this section as well. In her study, Lleras-Muney uses change in compulsory schooling laws and child labor laws between the years 1915 and 1939 in different 30 states as the instruments for years of schooling. She uses the 1960, 1970 and 1980 censuses of USA. Her concentration is the individuals who are 14 years old between 1915 and 1939. She chooses this age as threshold value since it is the lowest common school-leaving age across the states. In addition to the instrumental variable approach, Lleras-Muney also implements the regression discontinuity (RD) analysis in order to test the effects of increasing level of education on mortality rates in the USA. RD results suggest that additional years of schooling reduces the mortality, however the effect is found to be insignificant. On the other hand, IV results indicate a negative and significant

[13] In the study rate of time discount is referred as unobservables.

association between additional level of education and mortality rates. She concludes that 1 year of additional years of schooling decreases the 10- year mortality rate by approximately 3.6 percentage points.

Adams (2002) tests the effect of education on health for adults above 50 by using US Health Retirement Survey (HRS) for 1992. He implements both OLS and IV techniques in his study. In his analysis, Hausman test statistic results suggest that education is not endogeneous, hence it is not necessary to IV methodology. In spite of the results of the endogeneity tests, he reports the IV results but he notes that IV results can be biased towards OLS. He uses parental and sibling characteristics as instruments for education. Quarter of birth of the individual is another instrument for education level of the individual. SAH and an indicator variable that shows whether the individual has functional limitation or not are the two health outcomes in his model. Both OLS and IV results provide evidence that education is positively associated with good health.

Oreopoulos (2006) indeed examine the effect of additional years of schooling on individual's earnings. However, he has some conclusions about the impact of education on individual's health outcomes in this study. He uses UK General Health Survey data sets for the years between 1983 and 1998 together with Northern Ireland Continuous Household Surveys for the years between 1985 and 1998. He implement IV methodology due to the possible endogeneity between education and earnings (as well as health). In the study he uses compulsory schooling laws change in 1947, which increases the minimum school leaving age from 14 to 15, as instrument for education. He also implements OLS estimation techniques in his study. Both OLS and IV estimation results suggest that additional years of schooling decreases the prevalence of limitations resulting from physical health and mental health problems. In addition, his findings suggest that additional years of schooling decreases the probability of reporting poor health and it increases the probability of reporting good health accordingly.

Table 4.3.1 below shows the frequency distribution of SAH by education levels in Turkey and figure 4.3.1 provides a histogram analysis for the association between individual's SAH and education level.

Table 4.3.1. SAH and Education Level

SAH	Illiterate	Non-Graduate	Primary School Graduate	Middle School Graduate	High School Graduate	University+
Very Poor	296 (4.82)	95 (3.34)	223 (1.08)	28 (0.67)	28 (0.40)	8 (0.14)
Poor	1650 (26.88)	509 (17.90)	1874 (9.06)	200 (4.77)	222 (3.17)	125 (2.23)
Average	2517 (41.00)	1133 (39.84)	6692 (32.35)	1046 (24.96)	1421 (20.28)	855 (15.25)
Good	1574 (25.64)	1013 (35.62)	10530 (50.90)	2520 (60.13)	4449 (63.50)	3741 (66.53)
Very Good	102 (1.66)	94 (3.31)	1368 (6.61)	397 (9.47)	886 (12.65)	877 (15.64)
Total	6139	2844	20687	4191	7006	5606

Source: 2008, 2010 and 2012 pooled Turkish Health Survey

Note: Percentages are shown in parenthesis.

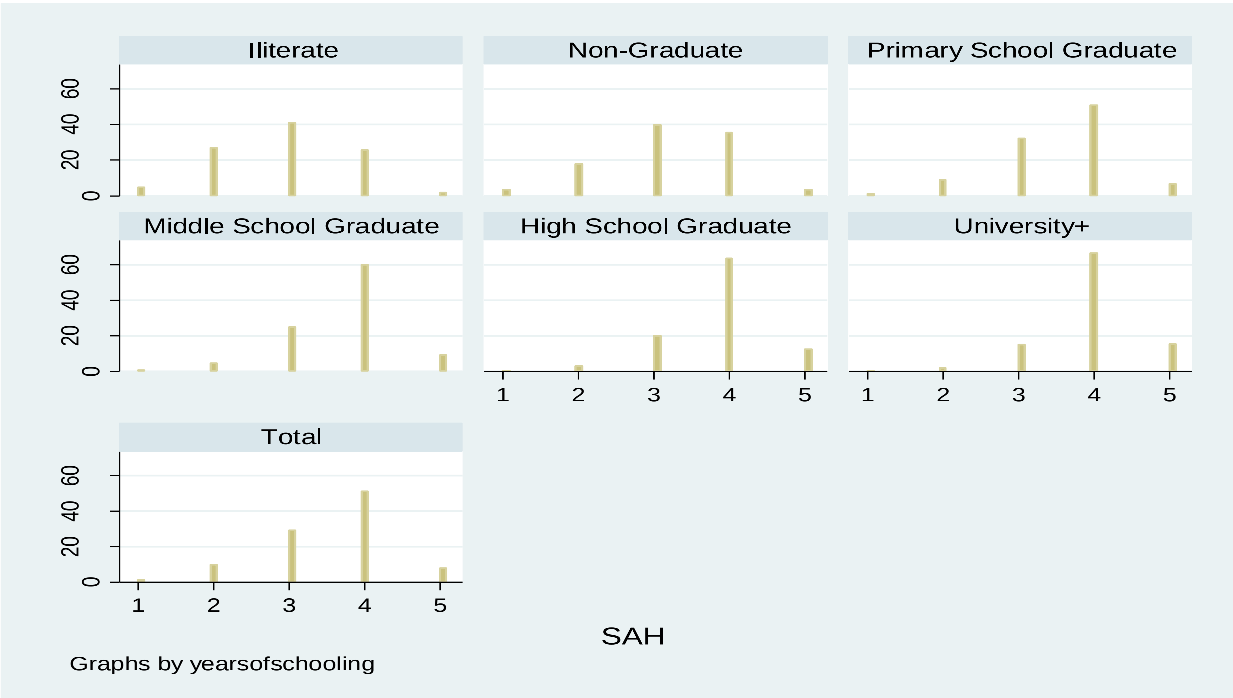


Figure 4.3.1. SAH by Education Level

The frequency distribution suggests that in Turkey the prevalence of reporting very good and good health is highest among the individuals who have university or higher degree. We observe that the frequency of reporting very poor and poor health is highest among the illiterate individuals. The histogram analysis clearly reflects the fact that higher levels of education lead to increase in the frequency of reporting very good and good health and they lead to decrease in the occurrence of reporting very poor and poor health. Therefore, the frequency distribution indicates the positive association between SAH and education level in Turkey.

4.3.2. SAH and Gender

Previous studies on gender differences in SAH suggest that although women have worse SAH than men at every stage of life, they die later than men(Case and Paxson, 2005). Case and Paxson find that that this paradox mainly results from the differences in the distribution of chronic conditions that both genders face. They conduct an analysis for gender differences by using National Health Interview Survey (NHIS) data for USA. The authors report that men experience more suffer forms of diseases, such as cardiovascular or lung diseases, hence they have more tendency to die. Case and Paxson also note that adverse effects of chronic conditions have more impact on males than females. For a developing country, Albania, Vaillant and Wolff (2012) also observe that women are more likely to state they are in poor or average health relative to men. The authors attribute this result to the differences in the observable characteristics of men and women. In fact, three different hypotheses explain why females’ life expectancies are greater than that of males. First, immune systems of females are more active. Next, estrogen protects females. Finally, since males do not have the second X chromosome, they tend to live shorter than females (Austad, 2006).

Table 4.3.2 below shows the frequency of reported SAHs in total and by gender in Turkey.

Table 4.3.2. Distribution of Reported SAH Measures

SAH	Total	Male	Female
Very Poor	678 (1.46)	219 (1.04)	459 (1.80)
Poor	4,580 (9.86)	1,449 (6.90)	3,131 (12.30)
Average	13,664 (29.40)	5,016 (23.87)	8,648 (33.97)
Good	23,827 (51.27)	12,112 (57.64)	11,715 (46.02)
Very Good	3,724 (8.01)	2,219 (10.56)	1,505 (5.91)

Source: 2008, 2010 and 2012 pooled Turkish Health Survey

Note:Percentages are shown in parenthesis.

The frequency distribution indicates that more than half of the respondents tend to report good health in Turkey. We observe that the frequency of reporting very poor health is very low in Turkey (1.46 per cent).

We observe that women are more likely to state that their health conditions are very poor or poor. The statistics suggest that 46.02 per cent of females report good SAH and 5.91 per cent of females report very good SAH. These percentages amount to 57.64 per cent and 10.56 per cent for males respectively. The percentage of females with stated poor health is nearly 12.30, while the proportion of males with poor health is nearly 7 per cent on average. These findings suggest that SAH of males are better than that of females. However, in Turkey, life expectancy at birth for females is 75 years in 2008 and 76 years in 2010. The number of years amounts to 71 in both years for males. In addition adult female mortality rate is 81 per 1000 female adults in 2008 and 77 per 1000 female adults in 2010. Adult male mortality rate is 142 per

1000 male adults in 2008 and 136 per 1000 male adults in 2010[14]. These findings suggest that although perceived health status of males is better than that of females, females live longer than males in Turkey. These findings show that the suggestions of Case and Paxson (2005) may also be valid for Turkey. Although women report that they are in poor health more often than men, men suffer from harder forms of chronic conditions, hence they tend to die earlier than women.

4.3.3. SAH and Age

Age is considered as another important determinant of SAH. Liang et al. (2005) examine the determinants of SAH among elderly Japanese adults by using five-wave panel study of national adults between the years 1987 and 1999. The authors believe that SAH should deteriorate with health as Grossman (1972) suggests since serious health problems are more likely to be seen in older ages, therefore the prevalence of poor SAH should also increase in older ages. They test this hypothesis by using multinomial logistic techniques. In the study, gender, education, marital status and employment indicators are the other determinants of SAH. The authors find that individual’s SAH decreases between ages 60 and 85, whereas it improves a bit between the ages 85 and 95. The improving SAH of elderly individuals is seen most probably due to the fact that those individuals who report better SAH compare their health statuses with their contemporaries whose health is very poor. The authors also refer those individuals as “hardy survivors” Their results also indicate that the individuals who have higher education levels have better reported SAHs. In addition, both employed and married individuals have better SAHs, which implies higher social networks improve the health statuses of the individuals.

Like Liang et. al. (2005), Zimmer et al. (2000) also conclude that although older people are more likely to be in poor health, it is also possible that at very old ages SAH of the individual can be improved. Because, old people are more likely to compare their health status with those of their peers. Table 4.3.3 below shows the distribution of SAH by age group and gender in Turkey.

[14] Source: World Bank (data.worldbank.org/indicator)

Table 4.3.3. Percentage Distribution of SAH by Age Group and Gender

Age Group	Very Poor		Poor		Average		Good		Very Good	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
25-34	0.47	0.49	2.50	4.92	15.47	24.92	66.28	60.00	15.27	9.67
35-44	0.70	0.68	4.62	10.50	21.52	35.55	61.66	47.86	11.50	5.41
45-54	0.77	1.66	7.06	16.26	27.45	41.12	57.40	38.83	7.31	2.13
55-64	2.08	4.05	11.70	21.55	37.06	44.40	44.73	28.33	4.42	1.67
65-74	2.76	6.47	18.62	30.00	43.22	45.10	31.49	17.45	3.91	0.98
75+	4.31	8.55	24.71	36.87	40.78	38.94	27.84	14.75	2.35	0.88

Source: Turkish Health Survey, 2008,2010 and 2012 pooled data

The frequency distribution suggests that age can be an important factor that explains the variations in SAH. The most frequent response among the 25-34 years old individuals is good SAH for both genders, while the most frequent response among the 75+ age group individuals is average SAH. The results reveal that the prevalence of good and very good health decreases with age. Table 4.3.3 suggests that the older the individual, the higher the probability that he/she perceives his/her SAH as poor or very poor. Nevertheless, older people have more health problems than younger. Hence, these results confirm that health problems lead to lower SAH.

4.3.4. SAH and Marital Status

Marital status is included in almost every empirical study as a control variable that explain the variations in SAH. However, to our knowledge, there is no specific study that considers on the importance of marital status as a control variable in these kinds of studies. Liang et. al. (2005) and Zimmer et al. (2000) refer marital status as a network tool of individual, and both of them conclude that married individuals have more improved health conditions relative to the others. Likewise, Fuchs (2004) suggests that in general, married individuals are healthier than those who are not married since having a spouse is assumed to make positive contribution to an individual' health. Table 4.3.4 presents the variation in SAH by marital status in Turkey.

Table 4.3.4. Percentage Distribution of SAH by Marital Status

Marital Status						
SAH	**Married**		**Single**		**Widowed/Divorced**	
	Male	Female	Male	Female	Male	Female
Very Poor	0.99	1.38	0.69	0.76	3.19	4.39
Poor	6.89	10.64	3.24	4.27	16.18	24.13
Average	24.97	33.90	11.28	16.86	30.88	41.74
Good	57.54	48.39	64.49	62.24	42.65	27.05
Very Good	9.62	5.69	20.30	15.87	7.11	2.69
Total	100.0	100.0	100.0	100.0	100.0	100.0

Source: Turkish Health Survey, 2008,2010 and 2012 pooled data

For marital status, Table 4.3.4 reports that single respondents are in better health conditions than married or widowed/divorced respondents for both genders. For single individuals, the percentage of reporting very good health is 20.3 per cent for males and 15.87 per cent for females. The effect is reversed for widowed/divorced respondents. The distribution results show that only 2.69 per cent of widowed/divorced females report that their SAH is very good and 27.05 per cent of females report that their SAH is good. In contrast, 42.65 per cent of the widowed/divorced men report that they are in good health. Widowed/divorced people report their SAH is poor more often than other marital status groups. The reported SAH of widowed/divorced females are even poorer than SAH of widowed/divorced males. The poor SAH of widowed/divorced individuals may be explained by the negative psychological effects of separation. From our results, we can conclude that the negative effect of separation has more impact on females relative to males.

4.3.5. SAH and Region

Almost all the studies that control for the region find that people live in urban areas have better health conditions. Among the other studies, Reijneveld (2002) specifically shows that individuals from less-favored places report poor SAH more often than the other respondents who live in more-favored places in his study. Reijneveld conducts his study for the Netherlands, a developed country. Table 4.3.5 reports that the result is similar for Turkey, a developing country.

Table 4.3.5 Percentage Distribution of SAH by Region

	Region			
SAH	**Urban**		**Rural**	
	Male	Female	Male	Female
Very Poor	0.85	1.34	1.51	2.93
Poor	5.80	10.45	9.62	16.81
Average	21.87	32.62	28.85	37.25
Good	60.34	49.14	50.89	38.43
Very Good	11.13	6.46	9.13	4.59
Total	100.0	100.0	100.0	100.0

Source: Turkish Health Survey, 2008,2010 and 2012 pooled data

From Table 4.3.5 we observe that almost half of the females who live in urban areas report that they are in good health. The percentage amounts to 60.34 for males who live in urban areas. In contrast, 38.43 per cent of females, who live in rural areas, report that they are in good health, whereas 50.89 per cent of males who live in rural areas perceive their SAH as good. Similarly, the percentage of respondents who report poor and very poor health is higher in rural areas for both genders.

4.3.6. SAH and Household Income

To our knowledge, the literature considers income as the second important covariate that explains the variation in health outcomes following education. For instance, Ettner (1996) focuses on the relationship between income and average daily alcohol consumption as well as SAH and other alcoholic variables by using National Survey of Families and Households data for the year 1987, the Survey of Income and Program Participation by using 1986 and 1987 panels, and the National Health Interview Survey for the year 1988. In addition to SAH and alcohol consumption, the author also investigates the relationship between income and work and functional limitations, bed days and scales of depressive symptoms. The authors implement both OLS and IV estimation techniques in their study. They use state unemployment rate, work experience, parental education and spouse characteristics as instruments for family income. For individual's SAH, the author uses 5- point scale (Very Good, Good, Fair, Poor, Very Poor) and for alcohol, the author uses average daily alcohol consumption during the previous two weeks and the number of alcoholic behaviors shown by the respondent during the previous year as health outcomes. In the empirical analysis, the authors also control for several demographic factors such as gender, marital status, the number of individuals living in the household, a dummy variable which indicates whether the individual lives in metropolitan area, race, age, education and the environment that the individual grows up (for instance, they define a dummy variable which is equal to 1 if the individual grows up with an alcoholic).

Ettner's results indicate that household income and SAH are strongly and positively related to each other, whereas household income is negatively related to depressive symptoms, work limitations, functional limitations and the bed days. Their results also suggest that if income increases, physical and mental health improves, however alcohol consumption also improves.

Lantz et al. (2001) examine the relationship between the individual's SAH and income and education as well as risky health behaviors such as smoking, alcohol consumption, physical activity and higher levels of BMI. For their study, Lantz and his friends use American's Changing Lives Study for the year 1986. By implementing multinomial logistic regression analysis, they find that both income and education strongly affect the individual's SAH, whereas the risky behaviors only explain a small portion of the differences in individual's SAH.

Mackenbach et al. (2004) look at the relationship between SAH and income for seven European countries[15]. They conduct the study for the individuals who are 25 years old and older. The data come from socioeconomic determinants of healthy

[15] The countries they examine are Belgium, Denmark, England, Finland, France, The Netherlands, and Norway.

aging study. In their empirical analysis, the health outcome, SAH ranges from 1 to 5. Different from the previous studies, they assume SAH as a continuous variable and run simple OLS regressions. They also apply several non-parametric analysis and finally select LOESS-function for presentation. LOESS-function is referred as locally weighted regression smoother. In LOESS-function presentation income axis is divided into very small parts and linear regression lines are fitted to each small part. For the non-parametric analysis, the authors use OLS regressions for fitting. The LOESS-function results suggest that for middle income groups, the association between SAH and income are linear, and they have positive relationship. In fact, in general, the results suggest that higher income levels improve the SAH of the individual, but that relationship is steeper for the individuals who belong to middle income group. A concave relationship between SAH and income occurs at higher income levels in all countries.

Except from these studies, there are other stuıdies which stress the positive association between SAH and income level include Van Doorslaer and Jones (2003), Quevedo et al.(2005), and Crossley and Kennedy (2002)[16]. Table 4.3.6 shows the percentage distribution of SAH by household income in Turkey.

[16] Van Doorsaler and Jones' (2003) main focus is the inequality of health outcomes, whereas Quevedo et al.'s (2005) and Crossley and Kennedy's (2002) primary fouses are measurement errors in SAH.

Table 4.3.6. Percentage Distribution of SAH by Household Income[17]

Household Income Level						
SAH	**Very Poor**	**Poor**	**Average**	**Good**	**Very Good**	**Total**
Income Level						
<350 TL	4.11	22.10	34.23	35.93	3.64	100.0
351-500 TL	2.50	17.17	35.61	40.43	4.29	100.0
501-620 TL	2.05	13.63	34.28	44.93	5.12	100.0
621-750 TL	1.64	12.29	31.54	47.88	6.64	100.0
751-900 TL	1.38	9.82	32.44	49.52	6.84	100.0
901-1100 TL	1.15	8.20	29.71	52.71	8.22	100.0
1101-1300 TL	1.13	7.95	30.22	52.28	8.43	100.0

[17] Source: Turkish Health Survey, 2008-2010 Pooled Data

Table 4.3.6. (Continued)

1301-1700 TL	0.97	6.66	28.28	55.71	8.37	100.0
1701-2300 TL	0.49	4.98	25.63	57.83	11.07	100.0
>2300 TL	0.47	3.86	19.97	62.23	13.47	100.0

Table 4.3.6 reports that there exists a positive relationship between income level and reported SAH. The distribution suggests that almost 76 per cent of respondents who belong the highest income level report good or very good SAH, whereas nearly 39 per cent of respondents from lowest income group report that they are in good health. The distribution results show that people from higher income groups more frequently report very good SAH.

4.3.7. SAH and Employment Status

Labor market indicators are also examined in the literature as the determinants of SAH. For instance, Böckerman and Ilmakunnas (2009) look at the relationship between individual's SAH and unemployment for Finland, a developed country. They use European Community Panel for Finland for the period between 1996 and 2001. The authors apply matching methodologies together with differences-in-differences method. The matching method overcomes the selection bias (i.e. the individuals with poor SAH can already be unemployed). Their results suggest that the individuals who report poor SAH are continually unemployed over time, whereas the individuals who report better health statuses are continually employed over time. Therefore, the authors support the idea that being unemployed leads to poor SAH, which is a valid finding in studies that use cross-section data, is no longer valid in panel studies.

Butler et al. (1987) show that people who are out of labor force tend to report misleading results regarding with their SAH. While Bound (1991) does not identify the impact of health on labor force participation of older men, Haan and Myck (2009) argue that poor SAH increases the probability of job loss and it significantly lowers the employment probability for men whose ages are between 30-59. They conclude that the effect is even stronger for the men whose ages are 59. Table 4.3.7 below shows the distribution of SAH by the individual's employment status and by gender in Turkey.

Table 4.3.7. Percentage Distribution of SAH by Employment Status

	Employment Status					
SAH	**Employed**		**Unemployed**		**Inactive**	
	Male	Female	Male	Female	Male	Female
Very Poor	0.23	0.76	0.47	0.87	2.96	2.08
Poor	3.30	7.36	5.80	3.79	15.07	13.67
Average	19.06	27.01	20.67	25.36	35.09	35.84
Good	64.48	55.02	61.09	56.27	41.84	43.62
Very Good	12.94	9.85	11.97	13.70	5.04	4.80
Total	100.0	100.0	100.0	100.0	100.0	100.0

Source: Turkish Health Survey, 2008, 2010 and 2012 pooled data

For employment status, Table 4.3.7 shows that inactive people more frequently state that they are in poor or very poor health relative to employed and unemployed people. This may occur due to age effect since pensioners, people who are above 65, are considered as inactive. In fact, distribution results of age group by employment status indicate that 92.09 per cent of individuals who are above 75 years old report that they are inactive. Another reason of why inactive individuals report poorer SAH relative to other employment groups can be explained as Butler et al. (1987) suggest. Butler and his friends support that individuals who are not working tend to report their health incorrectly owing to social pressure to justify not having job. Finally, we observe that the frequency of reporting very good or good SAH are almost the same for employed and unemployed individuals for both genders.

4.4. Conclusion

This chapter provides a descriptive analysis of variations in SAH in Turkey. In addition, the chapter provides a detailed literature review on the determinants of SAH, especially on education, which is the primary concern of this thesis. We examine the variations in SAH in Turkey by education, gender, age, marital status, region, household income and employment status. We use the data from 2008, 2010 and 2012 Turkish Health Survey for the frequency distribution analysis.

The most important results of this chapter can be summarized as follows: Education and household income level are positively associated with individual's SAH level. More than half of the people state they are in good health in Turkey. Males report that their SAH is good more often than females although life expectancy of females is higher than that of males. Age is negatively associated with SAH; young individuals are more likely to report good SAH. People living in urban areas report good SAH more often than people living in rural areas. Single adults state that they are in good health more often than married adults, most probably due to age effect. The percentage of reporting poor SAH is the highest among widowed/divorced females. People out of labor force state poor health more often than the employed and unemployed individuals.

Although the descriptive analysis give idea about the variations in SAH in Turkey, we need to establish an appropriate econometric model in order to have correct inference about the directions and significance of these covariates. Next, chapter five provides an empirical analysis for the relationship between SAH and education as well as the other factors.

CHAPTER 5

METHODOLOGY

5.1. Introduction

In this chapter, we focus on the important estimation methodologies that we implement throughout this thesis. In Chapter Six, we first investigate the relationship between self-assessed health (SAH) and education by implementing ordered probability models (ordered probit and ordered logit). Here, we assume that the individual's education level is exogenous to his/her reported SAH. The setting of the ordered probability models is discussed in Section 5.2.

In Chapter Seven, we examine the relationship between health behaviors and education. We form binary dependent variables for each health behavior, except Body-Mass Index (BMI) (BMI is continuous variable). In the first empirical part of Chapter Seven, we test the effect of education together with other controls on each health behavior by implementing probit models (for BMI Ordinary Least Squares (OLS)) estimation techniques. The theoretical formation of the probit estimation methodology is discussed in section 5.3.

Both ordered probability models (for SAH) and probit models (for health behaviors) give idea about the relationship between health and education, however the results we find may be inconsistent due to possible endogeneity problem between health and education. The endogeneity between health and education may occur because of two reasons: First, investment in health in earlier ages can affect both health and education in older ages. Second, unobserved variables such as time preference, genetic factors and family background can affect both health and education simultaneously (Eide and Showalter (2011), Bolin (2011)).

We investigate the relationship between health and education by addressing the problem of endogeneity via using Instrumental Variables (IV) methodology throughout this thesis. This methodology provides consistent estimates and as a result we obtain proper information on the association between health and education in Turkey. Several studies use IV approach to solve the endogeneity problem between health and education. Some of the examples of these studies include Eide and Showalter (2011), Kemptner et al. (2011), Brunello et al. (2015), Fonseca and Zheng (2013), Silles (2009), Arendt (2005), Llears-Muney (2005) and Adams (2002), Grossman (2008) and Cutler and Lleras-Muney (2008). All of these studies agree that endogeneity exists between health and education, and suggest that a variable, which is directly correlated with the years of schooling, and which is not related to the individual's health outcome, should be used as instrument for education in order to detect the relationship between health and education correctly. Most of these studies use change in compulsory schooling laws or educational reforms as the instruments for education[18]. In this thesis, we use the education expansion that took place in October 1960 and January 1961 in Turkey as the instrument for years of schooling. To our knowledge, we are the first who use 1960s educational reforms in Turkey as instrument for individual's years of schooling[19]. In Section 5.4, we first

[18] Adams (2002) is the exception. He uses parental and sibling characteristics as instruments for education. In his study, quarter of birth of the individual is another instrument for education level of the individual.

[19] There are several studies that examine the returns of change in compulsory schooling laws in Turkey. For instance, Aydemir and Kırdar (2013) examine the relationship between years of schooling and individual's earnings by using change in compulsory schooling laws as the instrument for education. They find a positive association between individual's education level and earnings. Kırdar et al. (2012) examine the impact of the raise in compulsory schooling years on teenage marriage and births. Finally, Kırdar et al. (2014) investigate whether the increase in compulsory schooling years narrow the education gap between men and women, as well as the gap between the individuals who live in urban and rural areas. In all these studies, the authors consider the compulsory schooling reforms that take place in 1997, which increases compulsory schooling from 5 to 8 years. We do not use this change as instrument for years of schooling in this study. Because, first, most of the individuals who are bound to this law are very young and most of them can be still enrolled in education. Second, we do not expect a young individual to report poor health status.

discuss the formation of IV methodology, then we discuss the validity of our instrument by showing pictures from Turkish Education and Census Data. Last, we also discuss the reform that promotes the socialization of health care services in Turkey. The law is also enacted in the early 1960s, however it cannot be implemented effectively. Therefore, we conclude that health reforms do not coincide with the education reforms in the early 1960s. Hence, we can confidently state that the educational expansion in the 1960s is purely exogenous factors to the individuals' health.

5.2. Ordered Probability Models

In the empirical model in Chapter Six, the dependent variable SAH is a categorical variable which ranks from 1 to 5. Applying ordinary least squares (OLS) estimation methodology is not appropriate since, the linear model assumes that the distance between the five categories are all equal. Therefore, we use ordered probability models in order to investigate the impact of education on health. First, we apply ordered probit model. Next, we apply ordered logit model. Both ordered probit and ordered logit models are originated from a latent variable model. The difference between the two models is while the error terms are assumed to have normal distribution in the ordered probit models, the error terms are assumed to have logistic distribution in the ordered logit models. However, we can obtain odds ratios of the covariates only from the ordered logit model. Obtaining the odds ratios of the covariates is important in the sense that it enables the researcher to interpret the coefficients in terms of marginal effects on odds ratio. Therefore, we both apply ordered probit and ordered logit models in this section.

One can derive the log-likelihood function of the ordered probit model as follows[20]:

[20] The notation in this part comes from Wooldridge (2002).

Let y be a categorical variable such that y ϵ {0,1,2,…….,J} where J is a known integer. Suppose that a latent variable y* is determined by:

$y^* = x\beta + \varepsilon$ where we assume ε has standard normal distribution and it has 0 mean and its variance is equal to 1.

The vector β is Kx1 matrix, and x does not contain a constant. Let $\alpha_1 < \alpha_2 < \ldots\ldots < \alpha_J$ be threshold parameters and define:

y=0 if $y^* \leq \alpha_1$

y=1 if $\alpha_1 < y* \leq \alpha_2$

.

. (5.2.1)

.

y=J if $y* \geq \alpha_J$

Equation (5.2.1) implies that y ϵ {0,1,2} if there are two threshold parameters α_1 and α_2.

As ε has standard normal distribution, we can directly derive the conditional distribution of y given x and we can easily calculate the each response probability by:

P(y=0|x)= P($y^* \leq \alpha_1|x) = P(x\beta + \varepsilon \leq \alpha_1|x) = \Phi(\alpha_1 - x\beta)$

P(y=1|x)= P $(\alpha_1 < y* \leq \alpha_2|x) = \Phi(\alpha_2 - x\beta) - \Phi(\alpha_1 - x\beta)$

.

.

.

P(y=J-1|x)= P $(\alpha_{J-1} < y* \leq \alpha_J|x) = \Phi\left(\alpha_J - x\beta\right) - \Phi(\alpha_{J-1} - x\beta)$

P(y=J|x)= P $(y* \geq \alpha_J|x) = 1 - \Phi(\alpha_J - x\beta)$

Then, the parameters will be estimated by using maximum likelihood (MLE) estimation techniques. For each i, the log-likelihood function will be as follows:

$$l_i(\alpha,\beta) = 1[y_i = 0]\ log[\Phi(\alpha_1 - x_i\beta)] + 1[y_i = 1]\ log[\Phi(\alpha_2 - x_i\beta) - \Phi(\alpha_1 - x_i\beta)] + \cdots + 1[y_i = J]\log\left[1 - \Phi(\alpha_J - x_i\beta)\right] \qquad (5.2.2)$$

For the ordered logit model, the derivation of the log-likelihood function is exactly the same. The only difference is Φ, normal distribution function, is replaced with Λ, logit function.

5.3. Probit Model

In Chapter Seven we test the effects of education on health behaviors (smoking, alcohol consumption, fruits and vegetable consumption and exercising) by using probit models of the form:

$$P(y = 1|x) = G(x\beta) \qquad (5.3.1)$$

Here x is 1xK vector of explanatory variables whose first element is unity, and β is Kx1 vector, the vector of coefficients. In addition, G (.) takes the values between 0 and 1. In equation (5.3.1) y refers to different health behaviors which are defined as binary variables. For instance, if y refers to smoking, it is equal to 1 if the individual reports that he/she is current smoker, 0 otherwise. The variable is defined in the similar manner for the other health behaviors. The probit model can be written in the following form:

$$G(z) \equiv \Phi\ (z) = \int_{-\infty}^{z} \Phi(v)dv \qquad (5.3.2)$$

Here Φ(z) refers to the standard normal density function which is equal to $(2\pi)^{-1/2}$ $\exp(-z^2/2)$.

In probit model, the parameters are estimated by maximum likelihood estimation (MLE) techniques. For each x_i, we can write the density of y_i as follows:

$$f(y|x_i;\beta) = [G(x_i\beta)]^y[1 - G(x_i\beta)]^{1-y} \qquad (5.3.3)$$

where y takes only values 0 and 1. Then the log-likelihood function can be written as:

$$l_i(\beta) = y_i log[G(x_i\beta)] + (1 - y_i) log[1 - G(x_i\beta)] \quad (5.3.4)$$

If the sample size is N, the log-likelihood can be written as:

$$L(\beta) = \sum_{i=1}^{N} l_i(\beta) \quad (5.3.5)$$

Then the MLE of β ($\hat{\beta}$) is the value which maximizes this log-likelihood.

5.4. Instrumental Variable Approach

In both Chapter Six and Chapter Seven, we investigate the relationship between health and education by addressing the problem of endogeneity via using Instrumental Variables (IV) methodology. In those empirical specifications, we instrument education by using the change in education reforms in the early 1960s in Turkey. By implementing IV methodology our objective is to obtain consistent estimates so that we can make a valid inference for the association between health and education. In this section we first provide a literature survey on IV methodology. Next we provide the theory of IV estimation techniques. Then we introduce our instrument and argue that it is a valid instrument by discussing both the education reforms and health reforms in Turkey in the early 1960s.

5.4.1. Literature Survey on IV Methodology

In this subsection, we focus on the studies that introduce the IV methodology, and how can one infer that an instrument is valid or not. These studies mainly discuss the IV methodology by applications on the relationship between earnings and education and they generally use the change in compulsory schooling laws as the instrument for education. The positive relationship between increase in the years of compulsory education and individual’s years of education is first proposed by Angrist and

Krueger (1991). They, in fact use the individual's quarter of birth as the instrument for education since they think that individuals who were born in later quarters attend to school more than the others since compulsory schooling laws force individuals to enroll in education up to a certain age. Therefore, increase in years of compulsory schooling force individuals to attend to school for more years.

Later on, Bound et al. (1995) discuss the potential estimation problems if the instrument is weak, and then they criticize Angrist and Krueger's (1991) study as they believe that quarter of birth is a weak instrument. Bound et al. discuss on the two main problems with IV Estimation process. First, they state that if there is a weak relationship between the instrument and the endogenous explanatory variable, then IV estimation will also give inconsistent results as in the case in OLS estimation. Second, they argue that if the R^2 between the endogenous explanatory variable and the instrument approaches to 0, then the instrument is considered as weak, and IV estimation results approaches to OLS Estimation results. Bound et al. also conclude that F-statistics for the excluded instruments is good indicator for the quality of the instrument and they should be reported at the end of the first-stage estimation results in order to persuade someone that the instrument is valid. According to Bound and his friends, Angrist and Krueger (1991) obtain reasonable standard errors by using quarter of birth as instrument for education, due to the large samples they used. Because, they find out that quarter of birth is weakly related to educational attainment as the R^2 between them nearly approaches to zero (i.e. R^2 is in ranges between 0.0001 and 0.0002).

Staiger and Stock (1997) also use Angrist and Krueger's data set and use quarter of birth as instrument for education and replicate their analysis. They suggest that the first stage F-statistics of the excluded instruments should be greater than 10 in order to claim that the instruments are valid. From their results, they observe that the F-statistics is greater than 10 if three quarters are used as the instruments for the individual's years of schooling. However, they state that the F statistics decreases below 10 if interaction terms such as interaction of state variables with quarter of birth, are added into the regression.

Angrist et al. (1996) bring a new term to the econometrics literature: "Local Average Treatment Effect (LATE)". In their study, they test the effect of being veteran during the Vietnam War on the individual's health outcomes[21]. They define a treatment group which includes individuals who serve army. Similarly, the control group they define consists of the individuals who do not serve the army. Individual's treatment status is determined by lottery number. If the individual's lottery number is high, then that individual is not chosen for military service. However, there cannot be perfect compliance to the lottery. For instance, a person who has high lottery number, may be volunteer to serve in the military, or the person who has low lottery number may not serve in the military since he has medical problems. In that study, the number comes from the draft lottery serves as instrument for veteran status. Angrist and his friends suggest that IV approach can only consistently estimate the average treatment effect for those who have low lottery number and serve in the army as well as those who have high lottery number and do not serve in the military. Because, the act of serving in the army, mainly comes from the exogenous factor, namely, the draft lottery, the instrument. Angrist and his friends call this IV estimand as LATE.

In general, the empirical results suggest that IV estimates are larger than the OLS estimates. Card (1995) is one of the examples of these studies. In his study Card investigates the relationship between earnings and education and use college proximity as the instrument for education. Both OLS and IV estimation results suggest positive relationship between earnings and education. However, Card find out that IV estimates are 50-60 per cent larger than the OLS estimates. Card suggests that this can be indicator of three reasons. First, it confirms the claim that schooling is strongly endogenous to earnings (most probably due to individual's unobserved ability). Second, OLS estimates may be downward bias due to measurement errors in the years of schooling. Third, he suggests that college proximity may only affect the schooling decisions of individuals who have poor family backgrounds. In other words, these individuals may not attend to college if they do not live nearby college.

---

[21] During Vietnam War period, the individuals are chosen to serve army by draft lottery.

Hence, the IV estimate may reflect the results for this subpopulation, which is consistent with the Angrist et al. (1996)'s LATE concept.

The papers discussed are pioneers in the recent IV theory literature. There are very famous studies as the followers of these studies such as Harmon and Walker (1995), Pischke and von Wachter (2008), Ichino and Winter-Ebmer (1999), Kane and Rouse (1995) and Oreopoulos (2006). Except Ichino and Winter-Ebmer, all of these studies use change in compulsory schooling laws as instrument for education[22]. Angrist and Krueger (1991) state that as the change in compulsory schooling laws are determined by the government, it provides exogenous variation in individual's years of schooling. In addition, compulsory schooling reforms are unrelated to individual's health outcomes. Therefore, it is a good instrument for individuals' years of schooling for examining the validity of education gradient of health. Likewise, other educational reforms such as change in education testing system, increase in number of courses in village schools and increase in number of schools are established by government and they also provide exogenous variation in individual's years of schooling[23]. The studies on the endogeneity between health and education become popular in recent decades. The studies that focus on the endogeneity problem between health and education are discussed in chapter four. Change in compulsory schooling laws or education reforms are used as instruments for education almost in all these studies.

[22] Ichino and Winter-Ebmer use father's education background as instrument for the individual's education.

[23] Arendt (2005) is one of the study that uses the educational reforms together with change in compulsory schooling laws as instrument for education.

5.4.2. IV Methodology

5.4.2.1. IV Estimation Process

First, consider the linear population model below:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \beta_k x_k + u \qquad (5.4.2.1)$$

The assumptions of the model are as follows:

$$E(u) = 0, Cov\left(x_j, u\right) = 0, where\ j = 1,2, \ldots, K - 1 \qquad (5.4.2.2)$$

However, x_k may be correlated with u, so it is potentially endogenous variable in equation (5.4.2.1). Therefore estimation of the equation (5.4.2.1) may give inconsistent results for each βs.

In order to solve this problem we need an extra variable (z) that satisfies the two conditions below:

(1) $Cov\ (z, u) = 0, that\ is\ z\ should\ be\ uncorrelated\ with\ the\ error\ term.$

(2) $z\ should\ be\ partially\ correlated\ with\ x_k\ if\ the\ other\ exogenous\ covariates\ are\ excluded.$ $That\ is\ if\ we\ run\ the\ model\ below$:

$$x_k = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \cdots \delta_k x_k + \theta_1 z + r_k \quad (5.4.2.3)$$

$$we\ should\ have\ \theta_1 \neq 0\ with\ the\ assumptions$$

$$E(\ r_k) = 0\ and\ r_k\ \ is\ not\ correted\ with\ any\ covariates.$$

If both conditions (1) and (2) are satisfied, then z can be considered as the instrumental variable candidate for x_k.

The reduced form equation for y can be obtained by substituting equation (5.4.2.3) into the equation (5.4.2.1). That is, we have:

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \cdots \alpha_{k-1} x_k + \lambda_1 z + v \qquad (5.4.2.4)$$

In equation (5.4.2.4) v refers to the reduced form error which is equal to $u + \beta_k x_k$ and for each j∈ $\{1,2, \ldots k - 1\}, \alpha_j = \beta_j + \beta_k \delta_j$, and $\lambda = \beta_k \theta_1$ since by assumption,

v is not correlated with the other explanatory variables, OLS gives us consistent estimates[24].

5.4.2.2. Methodology for the Endogeneity Tests

In this section we discuss the test for endogeneity which is first proposed by Hausman (1978). Hausman proposes that if 2SLS and OLS results for the coefficient of endogenous variable differ by only sampling error, then the possible endogenous variable is said to be exogenous. For notational simplicity we refer the dependent variable as y_1 and the possible endogenous explanatory variable by y_2, and the set of exogenous explanatory variables as x.

The population model is written as:

$$y_1 = x\delta_1 + \alpha_1 y_2 + u_1 \qquad (5.4.2.2.1)$$

Assuming that $E(x'u_1) = 0$ (5.4.2.2.2)

Whether y_2 is really endogenous to y_1 will be tested with the following procedure:

First, we write the linear projection of y_2 on x:

$$y_2 = x\pi_2 + v_2 \; where \; E(x'v_2) = 0 \qquad (5.4.2.2.3)$$

Then y_2 is endogenous if and only if $E(u_1 v_2) \neq 0$. Therefore we should test whether the error terms of equations (5.4.2.2.1) and (5.4.2.2.3) are correlated or not by writing the linear projection of u_1 on v_2:

$$u_1 = \rho_1 v_2 + e_1 \qquad (5.4.2.2.4)$$

By nature, $\rho_1 = \frac{E(v_2 u_1)}{E(v^2{}_2)} \; and \; E(v_2 e_1) = 0 \; and \; E(x'e_1) = 0$.

Therefore y_2 is exogenous if and only if ρ_1=0. We can infer whether ρ_1 is different than 0 or not, first substituting the equation (5.4.2.2.4) into the equation (5.4.2.2.1):

$$y_1 = x\delta_1 + \alpha_1 y_2 + \rho_1 v_2 + e_1 \qquad (5.4.2.2.5)$$

[24] If y is binary variable, the second stage regression is designed as probit regression, and the model is called as IV-Probit Model. We establish IV-Probit models in Chapter Seven.

Whether ρ_1 is different than 0 or not can be detected by standard t-test from the above OLS regression. However, v_2 is not observed. Therefore, first the reduced form parameters π_2 should be estimated from the equation (5.4.2.2.3), then the estimated residuals, denoted by $\widehat{v_2}$, can be easily calculated. If we replace v_2 with $\widehat{v_2}$ in equation (5.4.2.2.5), we have:

$$y_1 = x\delta_1 + \alpha_1 y_2 + \rho_1\widehat{v_2} + error \qquad (5.4.2.2.6)$$

Now we can test if ρ_1 is equal to 0 or not by applying standard t-test. Hausman suggests that if we reject $H_0\colon \rho_1 = 0$, then equation (5.4.2.2.1) should be estimated by applying 2SLS estimation methods as there is endogeneity between y_1 and y_2.

5.4.3. Discussion on the Instrument

This section devotes to the discussion of health and education reforms in the early 1960s in order to show the validity of the instrument that we use in this thesis. We use the educational reforms in October 1960 and January 1961 in Turkey as instrument for years of schooling. This thesis is the first that takes the educational reforms in 1960s as instrument for years of schooling in Turkey. We believe that the policy dummy, which is established based on the individual's birth year by considering the timing of education reforms, is valid instrument for individual's years of schooling. Because, first, all of these reforms are uncorrelated with individual's SAH and health behaviors since those reforms are the results of political developments in the early 1960s. Second, the individual's birth year is independent of unobservable characteristics such as time preferences, genetic factors and parental endowment. In addition, we observe that in THS data set, the correlation between individual's years of schooling and policy dummy is around 0.32. The magnitude of the correlation coefficient is high enough to state that the reforms positively affect the individual's years of schooling. It is also important to note that in the early 1960s the government attempts to make some health reforms, however they do not have significant impact on the improvement in society's health status. Therefore, we confidently state that our instrument is a good instrument to test the effects of years of schooling on individual's health outcomes. Throughout this section, first we discuss the health reforms in the early 1960s, then we discuss the education reforms and present relevant figures from Turkish Education and Census Data sets as well as THS data sets in order to convince the readers that the education reforms in the early 1960s are managed to succeed in increasing the number of students, number of teachers and number of schools in Turkey.

The law that promotes the socialization in health care services is enacted in 1961. This law is considered as the first step for the establishment of the national health care services (Yıldırım, 2013). In general, the law states that health care services should be objective, continuous and should be provided according to the public's needs. In addition, with this law, the government's objective is to provide free or partially free health services to the public. The socialization of health care services is also included in the first five year development plan. Moreover, the discussions on General Health Security (GHS) system, which is now applied in Turkey, also start in 1960s (Yıldırım, 2013).

The 1961 law provides a path for the integrated health care services since it supports the establishment of health centers (Sağlık Ocakları) and health posts (Sağlık Evleri) as primary health care services. Health centers are mainly responsible for the integration of health care services. The law states that each health center should consist of one general practitioner and other health personnels. The number of health personnels will be determined by the population of the region where the health center is established (Bulut, 2007).

Large-scale scale hospitals, which include specialist physicians and more beds are considered as secondary health care services. The objective of 1961 law is making the primary and secondary health care services work together as follows: First, the patient is expected to see the general practitioner in the health center in the region where he/she lives. Then, if it is necessary, the patient is directed to the region hospital by the general practitioner. Finally, the specialist physician in the region hospital evaluates the patient's general health condition and region hospital provides feedback to the health center (Bulut, 2007).

The establishment of the health centers is the major health reform in 1961 law. There are additional minor reforms stated in that law. These reforms are also important factors that may accelerate the socialization of the health care services. For instance, in the law, preventative health care services for children and women living in rural areas are structured with the following manner: The health condition of children and women will be monitored by a midwife with selected household visits and preventative health care services will be taken into account for those women and children.

In addition, the law states that there should be a register card for each individual for monitoring the general health status of the individual. The law also gives to the priority to the education of the health personnels. According to the law, every health personnel working in health centers should be educated regularly on the related issues. Moreover, the law also promotes the creation of health group presidencies. The law states that health centers in close regions (between 5 and 10) should come together and form one health group presidency. Then, these health group presidencies will establish the health group hospitals with 50-100 beds and 4 specialist physicians. Last, for every region with 200,000-400,000 population new hospitals should be built. Medical schools are considered as supporting factor of regional hospitals. Last, the law asserts that the health care services should be free. But everyone has the chance of choosing the physician if he/she pays the money for it. The physicians can work either in the public hospitals or they can work privately (Bulut, 2007).

The socialization of health services is planned to be pervaded all around the country in 15 years' time. However, the reforms are not successful either in the early 1960s or in the planned time (around the mid 1970s). In 1983, the government declares that the socialization of health services is valid in all around the country (Bulut, 2007).

Later on, the health reforms in worldwide, which are occurred due to rapid globalization and liberalization, also lead to acceleration in the health reforms in Turkey in the early 1990s.

The 1961 law cannot be implemented effectively in the planned time period due to several reasons[25]. Bulut (2007) argues that the most important problems are the lack of resources and political obstacles. The necessary investment for the infrastructure of the system cannot be achieved. Several fiscal crises also avoid the implementation of the system. As a result of these kinds of difficulties, the financial model for the system can never be improved and the cooperation between different sectors cannot be achieved. Some governors suggest that the salaries for the physicians should be higher than before to maintain the integrated health care services, however Turkey does not have enough resources to spent on high physician salaries. In addition, the regular education of health personnels and shipment system are also ignored in the related time period, although they do not need considerable capital investments. Bulut (2007) notes that now there are more than 6000 health centers and approximately 12000 health houses in Turkey. However, some of them still do not work effectively and there is still lack of health personnel in the health centers.

Yıldırım (1994) discusses other reasons for the failure of the 1961 health reforms in the planned time period. First, he points out that the resources in hand cannot be used effectively. Because, the government cannot prevent that the resources are used for general, large-scale hospitals. Second, Yıldırım states that the part of the budget, which the government plans to expend for social health care services, decreases due to other economic reasons. Third, it is important to note that health care personnels prefer to work in large-scale hospitals rather than working in health centers. In order to overcome this problem, health care personnels are forced to work in rural areas with a law which is approved in 1981. Yıldırım also notes that, despite this new law, the health centers remain to be ineffective. Therefore, the patients continue to go to

[25] Prof. Dr. Nusret Fişek devotes considerable effort on this issue. He signs some protocols with Ministry of Health in those years for certain regions in Ankara. And the planned socialization in the health system is implemented effectively in Çubuk and Etimesgut regions in Ankara.

large-scaled hospitals. Tansel (2012) also points out that individuals generally ignore the primary health care services and go directly to general hospitals.

In one study Koray (2000) states that in Turkey, in general, the expenditures on health care services are less than the expenditures on education. Therefore, education reforms are implemented immediately in Turkey, therefore their effects become more apparent in a short-time period. There are several educational reforms in Turkey in the early 1960s. First, a law is enacted in January 1961 which increases compulsory schooling to 5 years in all around the country[26]. Compulsory schooling is used to be 3 years in villages before this law is implemented. (Erdogan, 2003; Şen, 2013). Second, in 1960 a law has been implemented such that men can serve army as primary school teachers in villages[27]. Third, the law that passes in 1961 allows the volunteers who have been graduated from middle school to be assigned as primary school teachers and the volunteers who have been graduated from high school or higher to be assigned as middle school teachers after passing a certain course[28]. With these two laws the number of teachers in Turkey has increased (Akyüz, 2007)[29].

We collect all these three main laws and call it as “educational expansion” in the 1960s. Making the compulsory schooling 5 years in villages is the most important component of the educational expansion. First, the figure 5.4.3.1 below shows the fraction of individuals who have at least primary school degree. We show the figure for the individuals who were born between 1946 and 1970 in order to see the relevant impact of education reforms on narrower groups of birth cohorts. From the figure, we observe the relevant jump in the number of graduates between the 1950-51 and 1952-53 birth cohorts. The fraction of individuals who have at least primary school degree

[26] Source:Resmi Gazete dated 12th January 1961

[27] These soldiers continue to be teacher if they would like to. Source: Resmi Gazete dated 15th October 1960

[28] Source: mevzuat.meb.gov.tr/html/24.html

[29] You can find the original primary sources in the appendix part of this thesis.

increases from 0.65 to 0.7 per cent between 1949-1950 and 1950-1951 birth cohorts. The fraction even increases to 0.75 for 1951-1952 birth cohort.

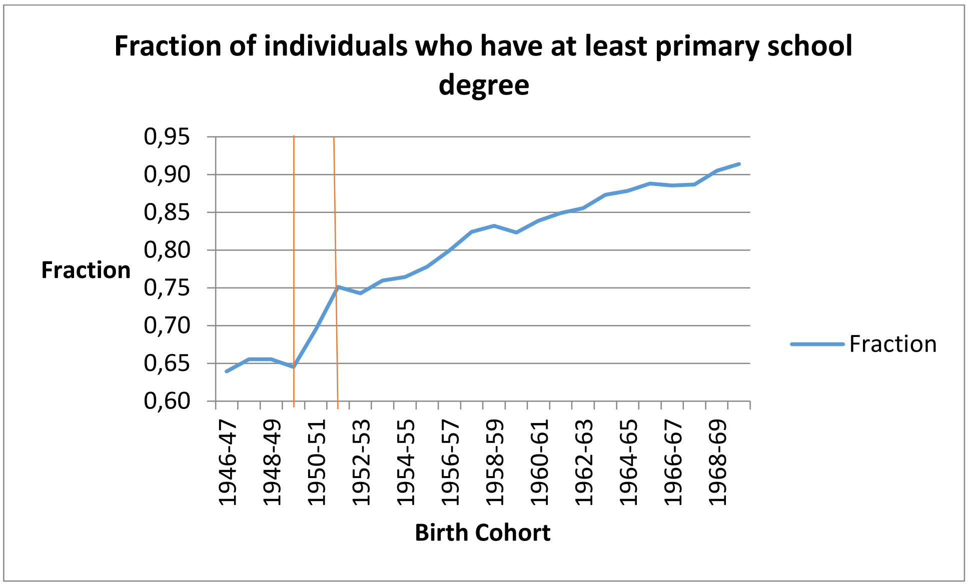


Figure 5.4.3.1. The number of the individuals who have at least primary school degree

Source: THS Data Set

The figures 5.4.3.2 and 5.4.3.3 show that higher amount of students benefit from this law than we expect, since the population in villages are higher than the population in cities in Turkey until mid 1980s. According to 1927 General Census Survey Results nearly 80 per cent of the population lives in villages, while it decreases to 35 per cent in the beginning of 2000s. Both the figures 5.4.3.2 and 5.4.3.3 show that the decrease in the village population becomes rapid after 1955. However, village population is still higher than the city population in 1961, when the new compulsory schooling law passes. Therefore, we can confidently state that the change in compulsory schooling laws affect a huge number of students.

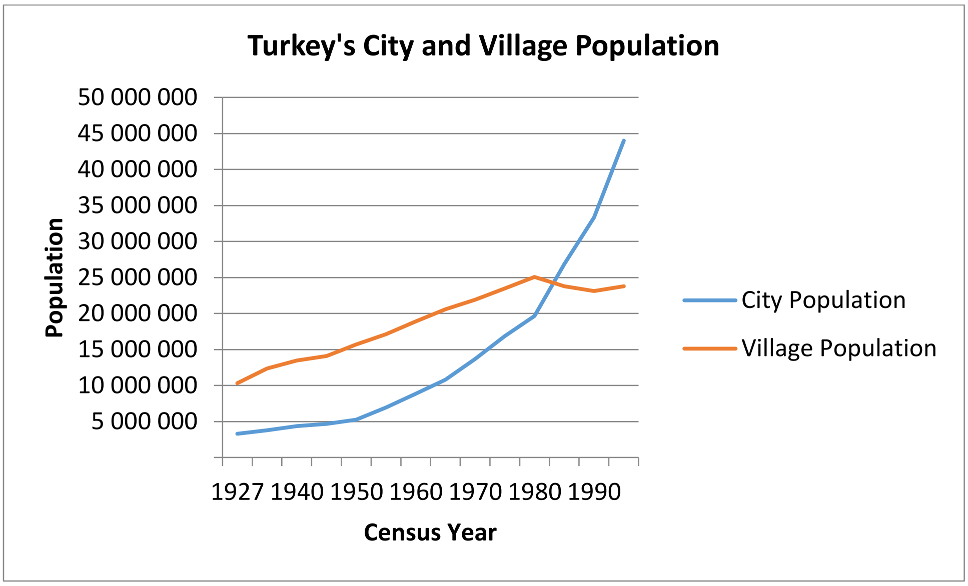


Figure 5.4.3.2.. City and Village Population in Turkey Over Time

Source: TURKSTAT, General Census Results

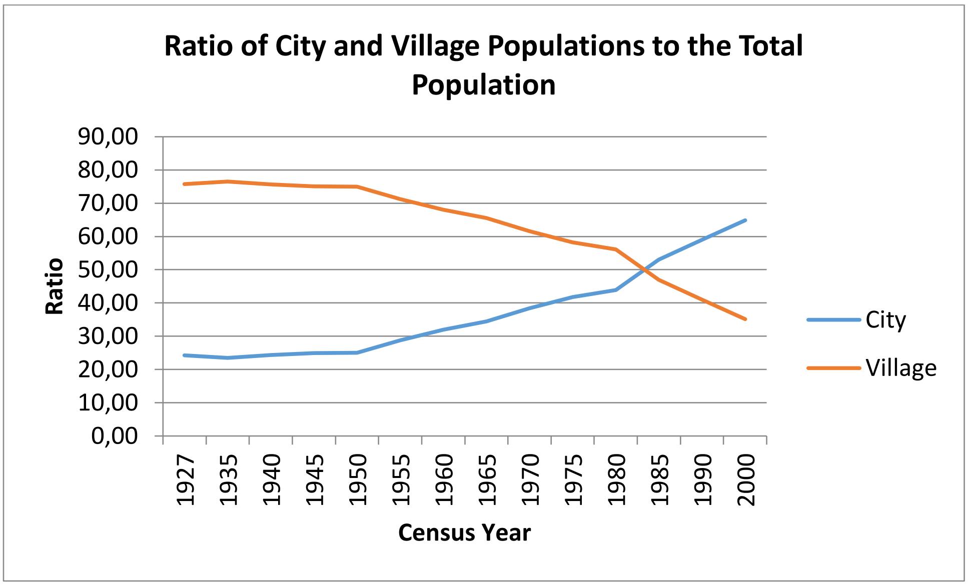


Figure 5.4.3.3. Ratio of City and Village Populations to the Total Population

Source: TURKSTAT, General Census Results

Figure 5.4.3.4 shows the number of primary school graduates in villages between 1950-51 and 1975-76 academic semesters. We see that there is a continuous increasing trend in the number of graduates between 1950-51 and 1959-60 academic semesters. The rising trend is more apparent and becomes steeper after the reforms are implemented starting from the 1960-61 academic semester. Ministry of Education statistics indicate that number of primary school graduates in villages grows by 9.74 per cent in 1960-61 academic year. This steep growth continues over time and it reaches to 19.55 per cent at the end of the 1964-65 academic semester, after five years the reform takes place.

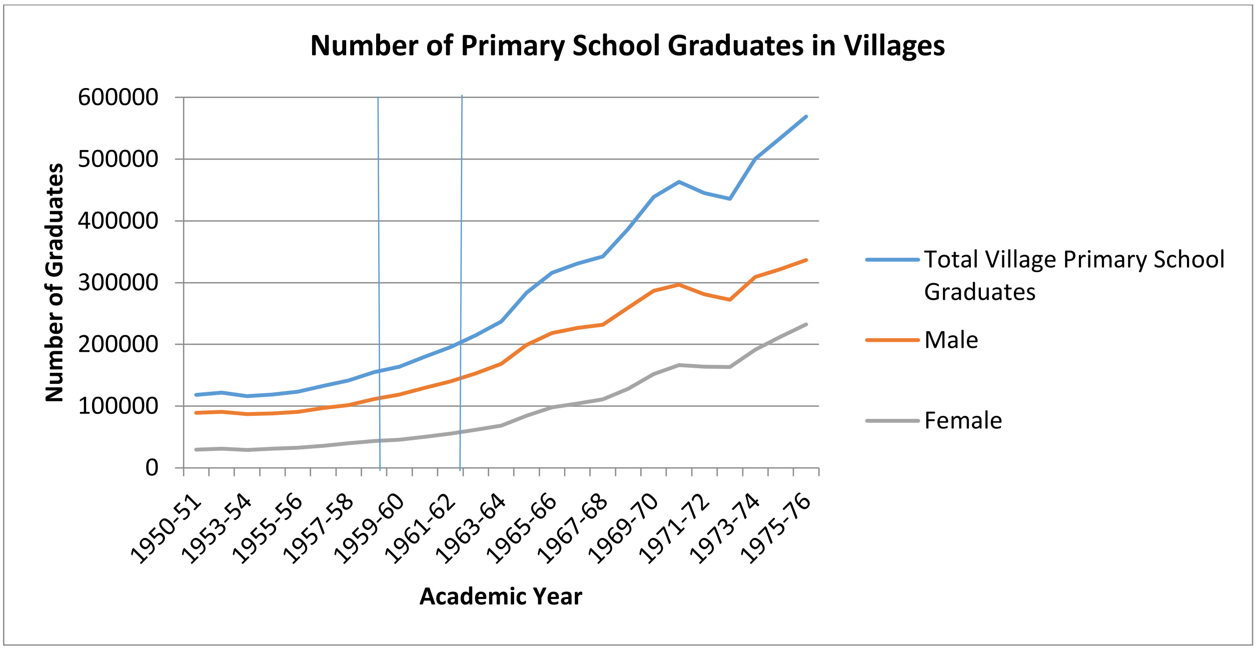


Figure 5.4.3.4. Number of Primary School Graduates in Villages

Source: Turkish Ministry of Education Statistics

Next, we continue to discuss on Turkish Education Data in order to show that the education reforms in 1960s have positive impacts on the school enrollment rates, number of teachers, number of schools, and number of graduates in Turkey.

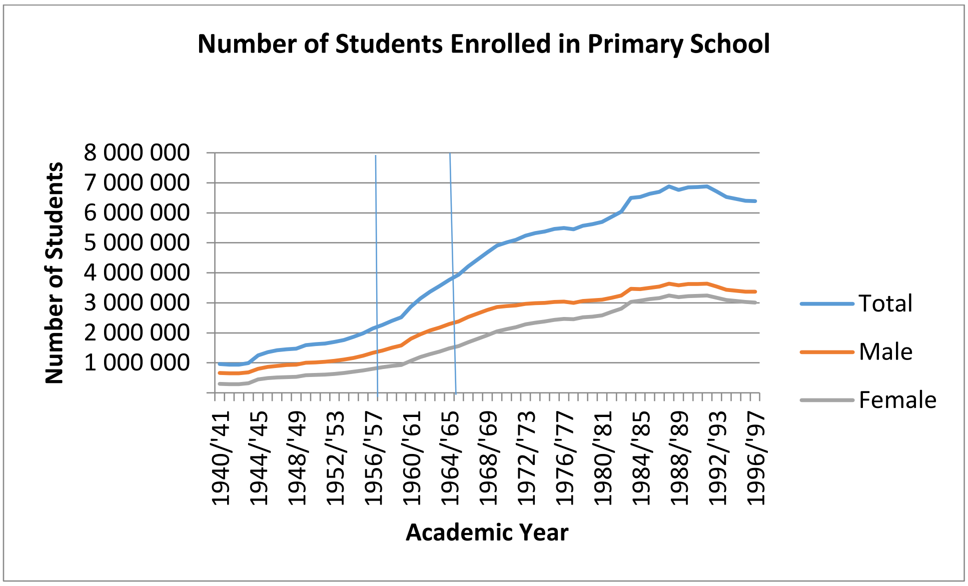


Figure 5.4.3.5. Number of Students in Primary Schools in Turkey Over Time

Source:TURKSTAT

Figure 5.4.3.5 shows the number of students who are enrolled in primary schooling in Turkey over time. We see that there is a continuous increasing trend in number of primary school students. The trend becomes even steeper after the 1960s education reforms take place. Although, the number of female students is lower than the number of male students, the continuous increase is also valid for female students. In 1959-1960 academic year, before the education reforms, the number of students in primary school is around 2.5 million, whereas in 1965-1966 academic year this number has risen to around 3.95 million, which implies approximately 50 per cent increase in the number of primary school students five years after the education reforms take place. In addition, in 1959-1960 academic year, the number of male students who attend to primary school is equal to 1,582,798 whereas it is equal to 931,794 for female students. However, in 1965-1966 academic year these numbers increase to 2,377,666 for males and 1,555,585 for females. This implies a 50 per cent and a 40 per cent growth in number of male and female students respectively.

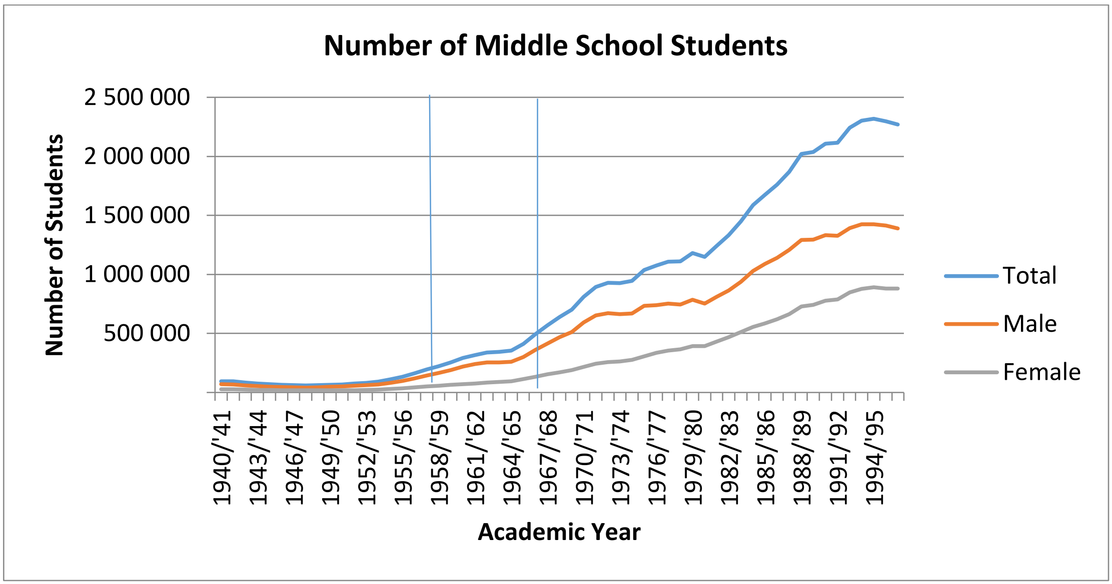


Figure 5.4.3.6. Number of Students in Middle Schools in Turkey over Time

Source:TURKSTAT

Figure 5.4.3.6 shows the number of middle school students in Turkey over time. We observe that starting from early 1950s there is a continuing increase in the number of male students in Turkey, and the increase becomes more apparent after the early 1960s. Turkish Education Statistics suggest that in 1959-1960 academic year, before the new education reforms take place, the number of male students who are enrolled in middle schooling is equal to 190,865 while this number is equal to 64,101 for female students. In 1965-1966 academic year, after 5 years that the reforms take place, the numbers rise to 299,673 for males and 112, 780 for females. This implies a 57 per cent increase for males and 76 per cent increase for females. Therefore, the statistics reveal that the new education reforms benefit the female students even more. In addition, we observe a small decline in the number of middle school students in the early 1980s, most probably due to the terrorist attacks and the military coup afterwards[30].

[30] It is important to stress that, both in figures 5.4.3.3 and 5.4.3.4 above, we do not show the continuing increasing trend in number of middle school and primary school students after 1996-1997 academic year. Since in 1997 compulsory schooling laws change in Turkey (it raises to 5 to 8 years). Therefore the students who enroll in primary school and middle school, are supposed to be enrolled in Basic Education (Ilkogretim). Hence, if we consider the students who are enrolled in Basic Education as primary school students and middle school students,

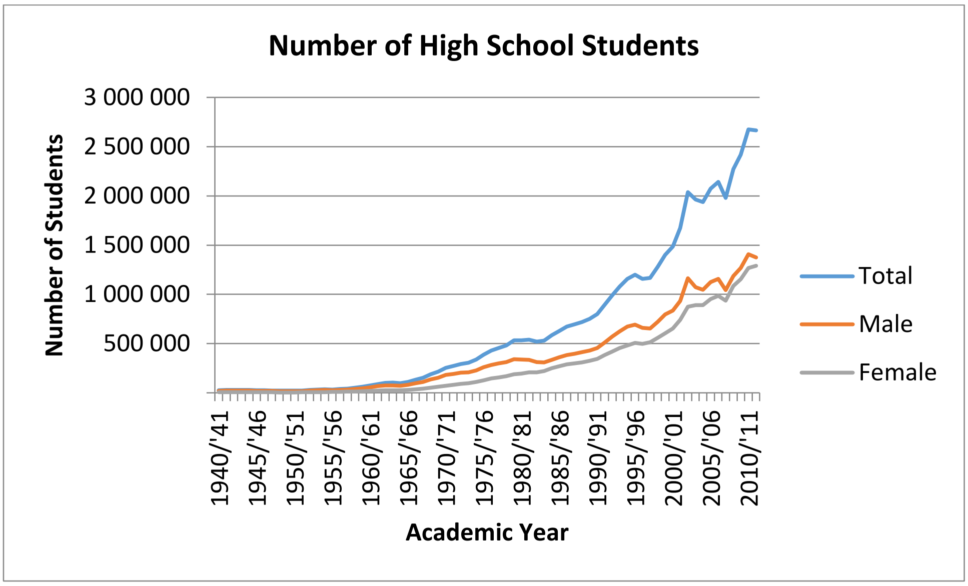


Figure 5.4.3.7. Number of Students in High Schools in Turkey Over Time

Source:TURKSTAT

Although the reforms that take place in the early 1960s are relevant for primary and middle school education, the reforms have spillover effect on high schools. Figure 5.4.3.7 shows that the total number of students who attend to high school in Turkey is around 4500 before 1960s. After 1960-1961 academic year we observe a small rise in number of high school students. This rise continues over time. For females, attendance to high schools is very low until 1964-1965 academic year. From figure 5.4.3.7 we infer the attendance to high schools for females increase afterwards. We also see that the high school attendance gap between males and females gets narrower over time.

there will be a huge jump in both figures after 1996-1997 academic year, and the increasing effects of 1960s education reforms on number of both primary school and middle school students cannot be seen clearly.

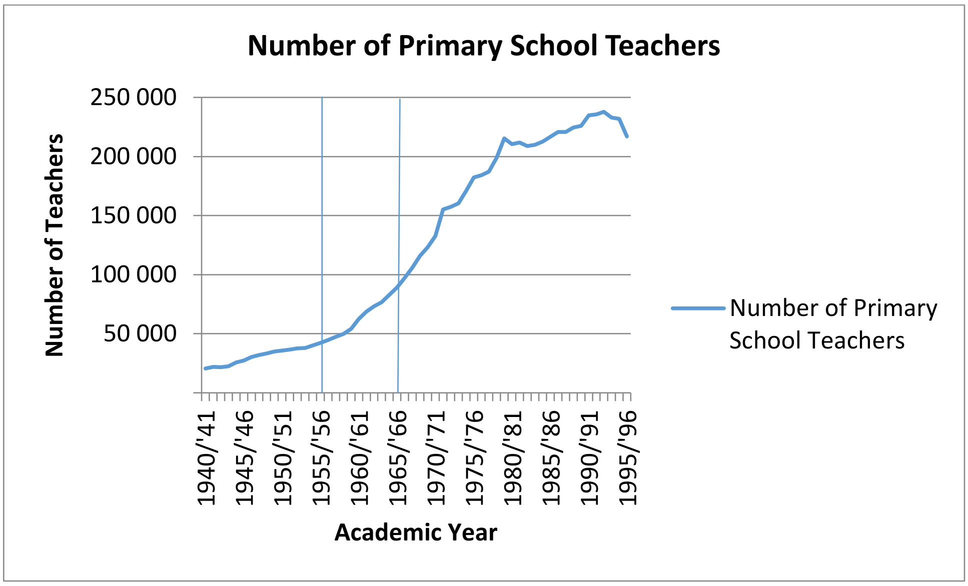


Figure 5.4.3.8. Number of Primary School Teachers

Source:TURKSTAT

One of the important objective of early 1960s education reforms is to increase the number of primary school teachers. Figure 5.4.3.8 reveals that the ongoing increase in number of primary school teachers becomes steeper after 1960-1961 academic year. For instance, the number of primary school teachers increase from 54,000 to 89,000 between 1959 and 1965. This implies a 65 per cent increase in the number of primary school teachers in 6 years time. This time period is short enough to say that the educational reforms are successful in increasing the number of primary school teachers.

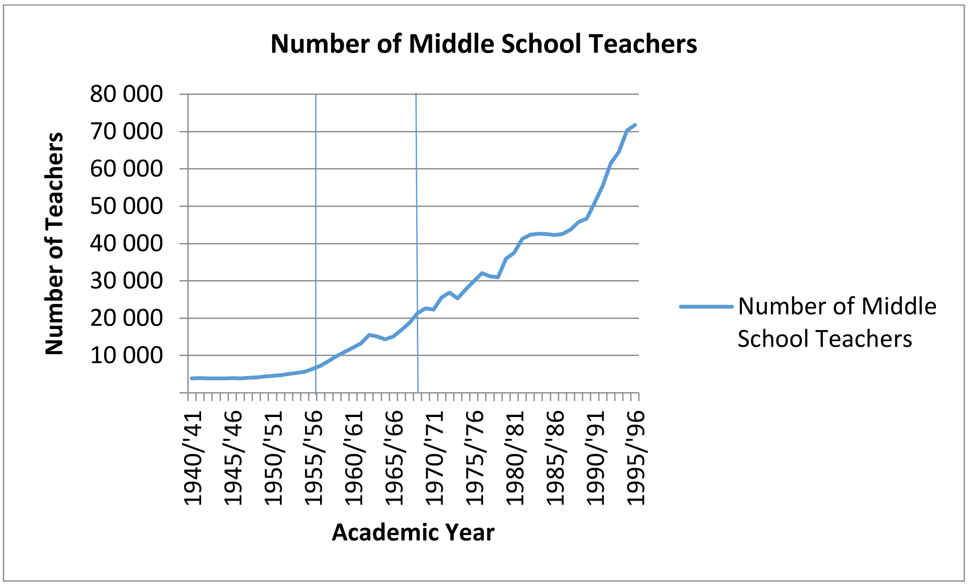


Figure 5.4.3.9. Number of Middle School Teachers

Source:TURKSTAT

The education reforms in 1960s also aim to increase the number of middle school teachers as well by implementing a law that high school graduates may serve as middle school teachers after passing a certain course. Turkish Education Statistics indicate that in 1959-1960 academic year the number of middle school teachers is approximately equal to 11,000, while in 1961-1962 academic year it raises approximately to 13,300. This implies that number of number of middle school teachers raise by around 21 per cent in 2 years time. This implies that the new education reforms also succeed in increasing the number of middle school teachers in a very short time period.

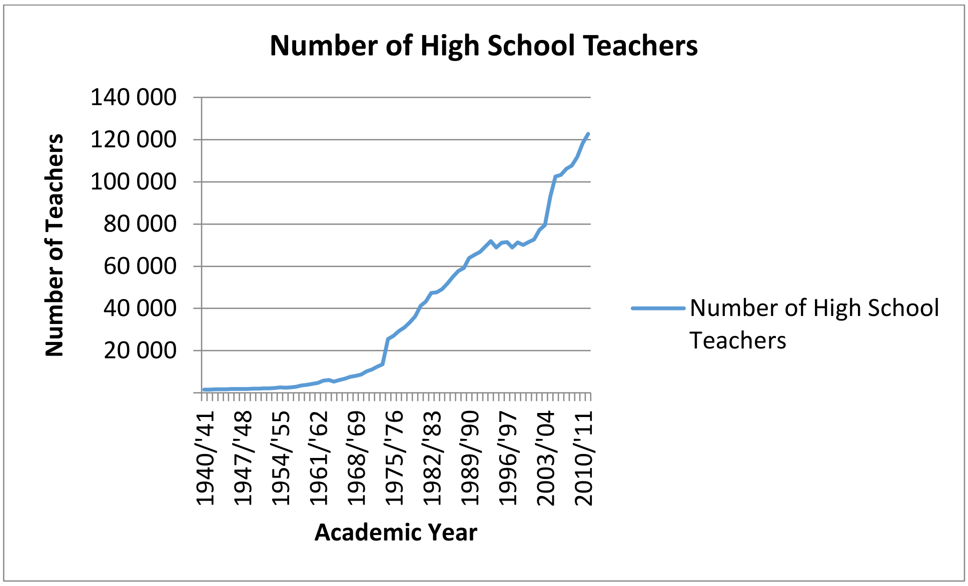


Figure 5.4.3.10. Number of High School Teachers

Source:TURKSTAT

After the education reforms are implemented, we observe a spillover effect in number of high school teachers as well. Figure 5.4.3.10 reveals that beginning from 1959, number of high school teachers start to increase. Moreover, the increase becomes more apparent in the early 1970s, approximately after 10 years the reforms take place.

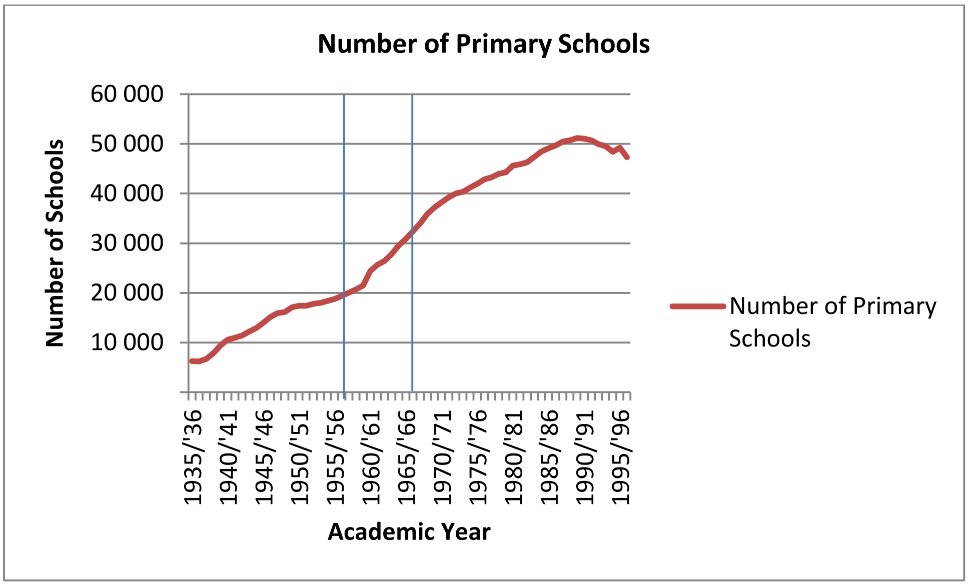


Figure 5.4.3.11. Number of Primary Schools in Turkey Over Time

Source: TURKSTAT

From figure 5.4.3.11, we observe that while the number of primary schools is 21,429 in 1959-1960 academic year, one year before the education reforms take place, it increases to 25,677 in 1961-1962 academic year, following the education reforms, which implies around 20 per cent increase in the number of primary schools in one year. The growth in number of primary schools is permanent in Turkey until 1996-1997 academic year. Most of the primary schools and middle schools are united at the beginning of 1997-1998 academic year when compulsory schooling raises from 5 years to 8 years in Turkey. That's why from figure 5.4.3.11 we observe a reduction in number of primary schools starting this period.

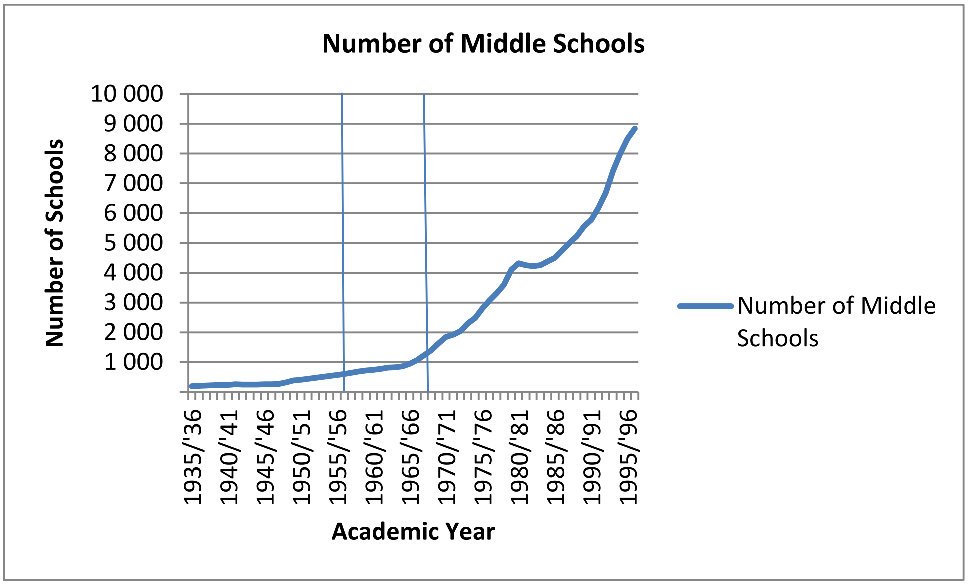


Figure 5.4.3.12. Number of Middle Schools in Turkey Over Time

Source: TURKSTAT

Turkish Education Statistics indicate that while the number of middle schools is equal to 715 in 1959-1960 academic year before the policy, it raises to 1405 in 1968-1969 academic year, after 7 years the education reforms are implemented. This implies nearly 100 per cent increase in number of middle schools in such a short time period. Figure 5.4.3.12 reveals that the growth in number of middle schools is permanent in Turkey over time.

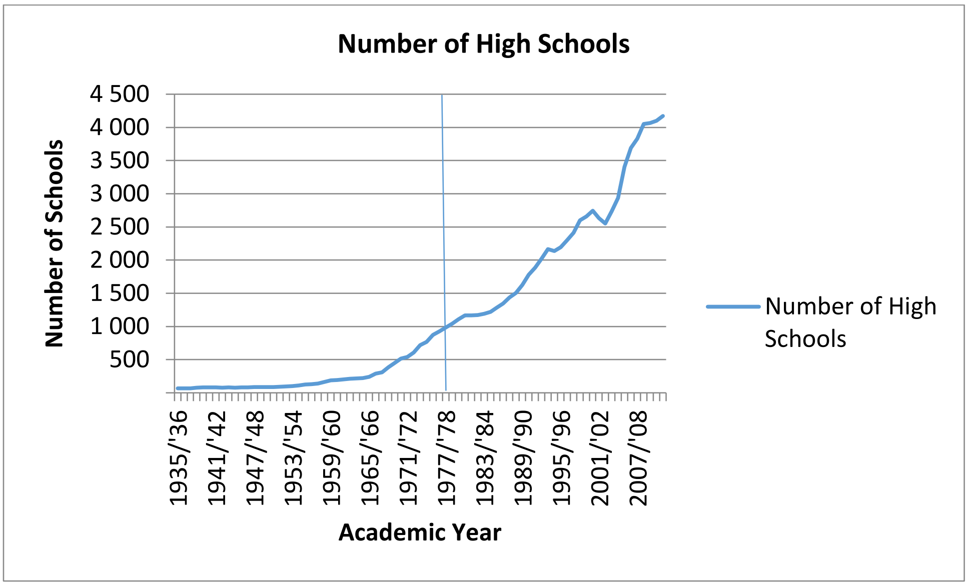


Figure 5.4.3.13. Number of High Schools in Turkey Over Time

Source: TURKSTAT

In this section, we also look at number of high schools in Turkey in order to see how it varies over time as a result of increasing demand to primary and middle schools over time. Figure 5.4.3.13 reveals that there is a significant increase in the number of high schools after the early 1960s. The increase in number of high schools becomes more rapid in the beginning of 1970s, 10 years after the education reforms take place.

It is important to note that Turkey's population has raisen sharply after 1960s. The rise in the population growth may give biased information on how the education reforms affect the number of teachers. Therefore, the figures 5.4.3.14, 5.4.3.15 and 5.4.3.16 below show the increase in number of primary, middle and high school teachers per school aged children in Turkey over time. The figures reveal that there has been a significant increase in number of teachers per school aged children over time in Turkey after 1960s. Figure 5.4.3.14 reveals that while the primary school teachers per school aged children ratio is around 1.5 in 1960-1961 academic year, it raises to 3 per cent in 1980-81 academic year, which implies 100% per cent raise in teachers per primary school aged children ratio in 20 years time. Figure 5.4.3.15 shows us that there is an increasing trend in the number of middle school teachers per middle school aged children. In figure 5.4.3.16. we again observe the spillover effect: The number of high school teachers per high school aged children has also risen over time.

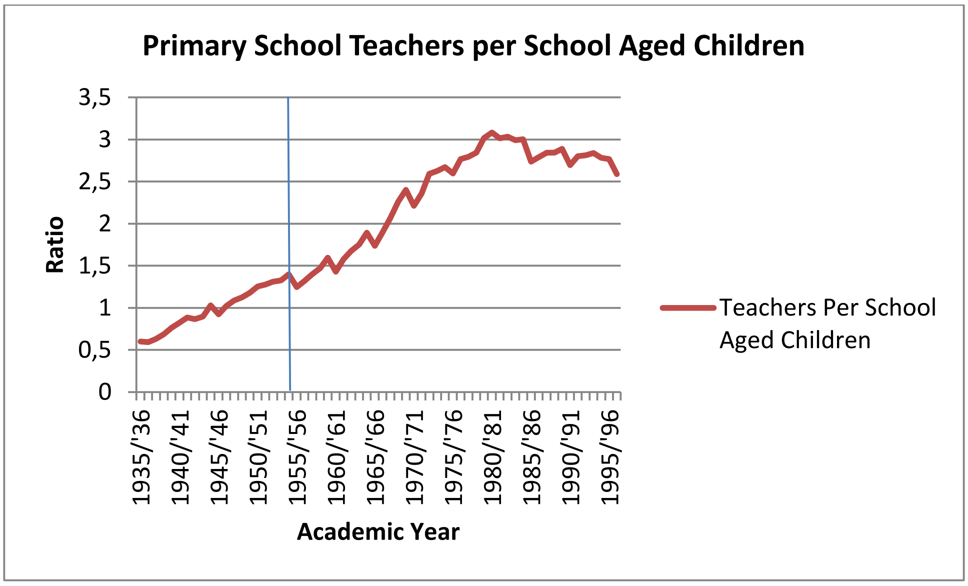


Figure 5.4.3.14. Number of Primary School Teachers per School-Aged Children in Turkey Over Time

Source: TURKSTAT

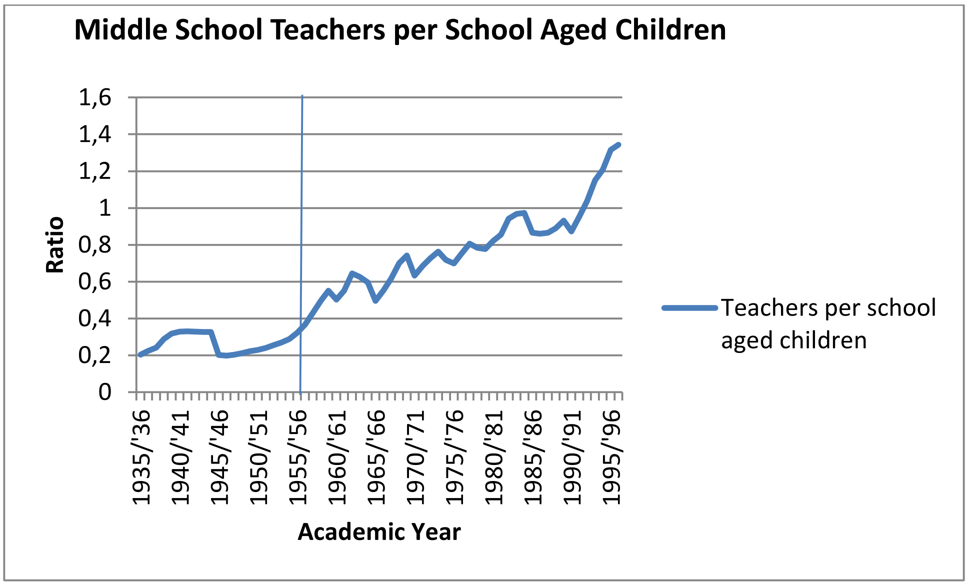


Figure 5.4.3.15. Number of Middle School Teachers per School-Aged Children in Turkey Over Time

Source: TURKSTAT

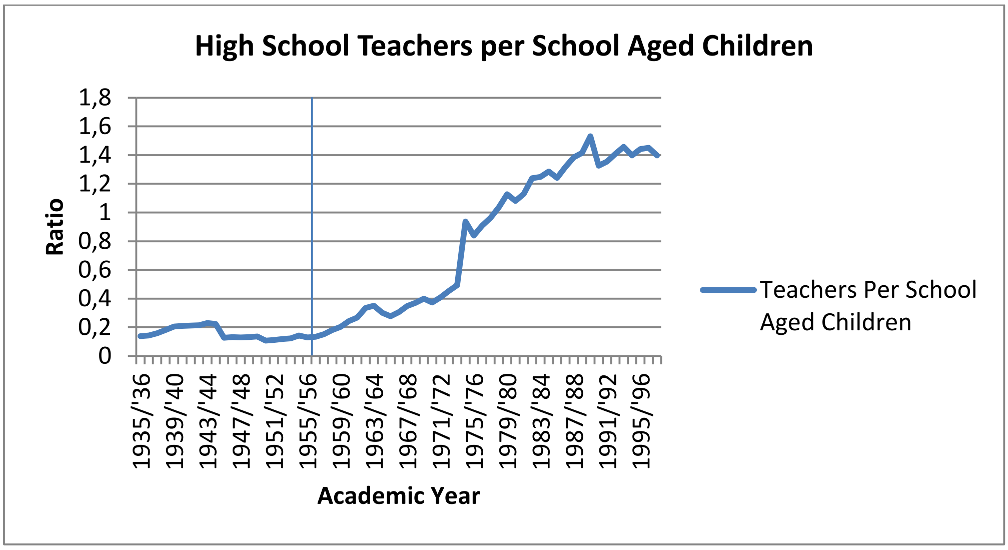


Figure 5.4.3.16. Number of High School Teachers per School-Aged Children in Turkey Over Time

Source: TURKSTAT

As a result of noteworthy increase in number of schools and number of teachers after the education reforms, the number of students who are enrolled in primary and middle schools have also increased. The figures 5.4.3.17 and 5.4.3.18 reveal the growth in the attendance to primary and middle schools. From figure 5.4.3.17 we observe that while the enrollment in primary school rate varies between approximately 20 per cent and 60 per cent before 1960s, it varies between 60 per cent and 95 per cent after the education reforms take place. Similarly, figure 5.4.3.18 reveals that increase in the attendance to middle schools become even larger after the early 1960s. Figure 5.4.3.19 presents us the spillover effect. The figure clearly reflects that the increase in enrollment to high schools is mostly seen after the late 1960s.

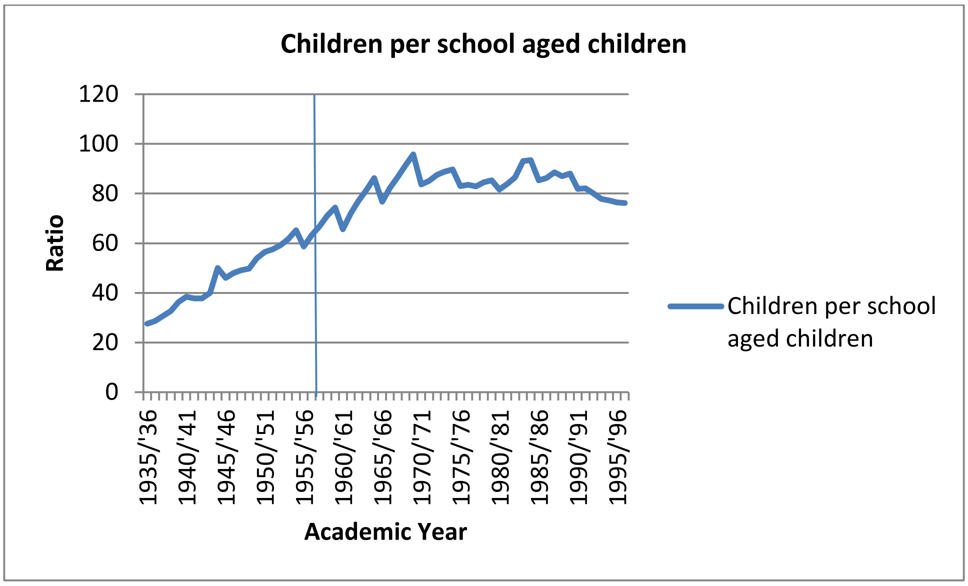


Figure 5.4.3.17. Primary School Enrollment Rate in Turkey Over Time

Source: TURKSTAT

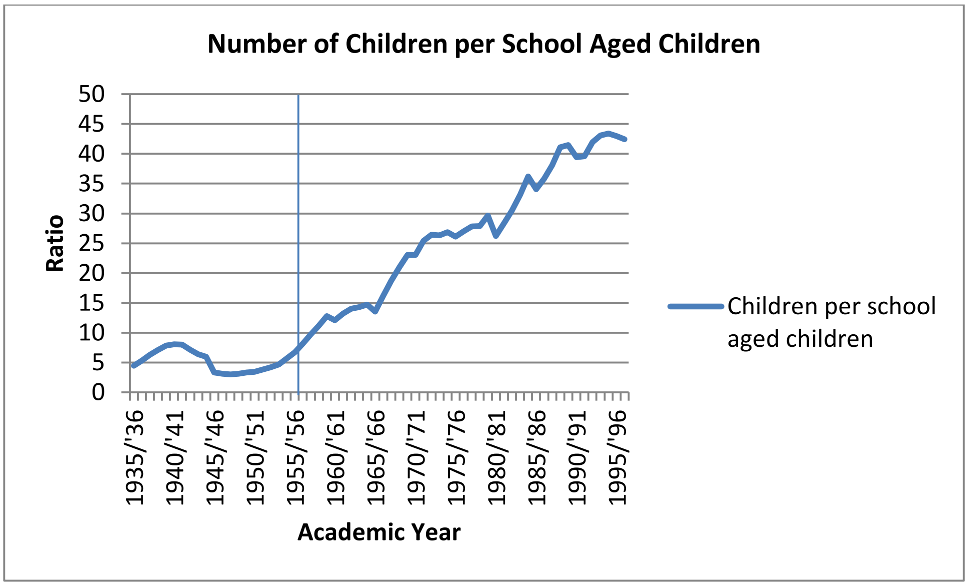


Figure 5.4.3.18. Middle School Enrollment Rate in Turkey Over Time

Source: TURKSTAT

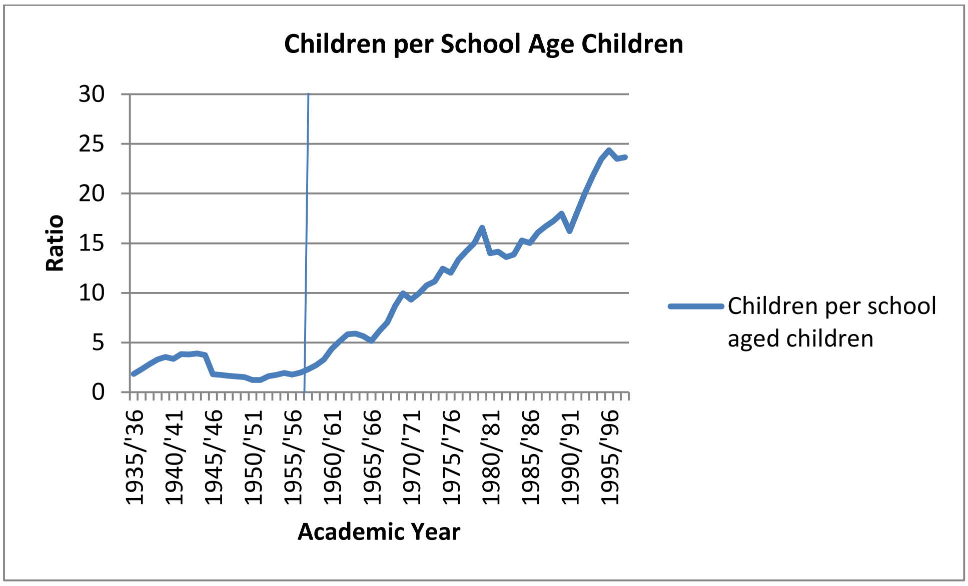


Figure 5.4.3.19. High School Enrollment Rate in Turkey over time

Source: TURKSTAT

In figures 5.4.3.20, 5.4.3.21 and 5.4.3.22 below we look at the fraction of the number of graduates to the number of graduate aged population for primary school, middle school and high school graduates. In other words, these figures show us the fraction completing the 5th, 8th, and 11th grades respectively. For primary school, figure 5.4.3.20 tells us that in 1964-1965 academic year, approximately, after 5 years that the reforms have passed, there is a huge increase in the fraction of both male and female graduates. The growth in the primary school graduates amounts to 14.27 per cent in 1964-1965 academic year, while it is equal to 7.51 per cent in 1959-1960 academic year. We also observe that especially, after the mid 1980s, the grade completion rate gap between males and females is narrowed. For middle and high school graduates, we see this gap is nearly closed after the late 1990s, when the new compulsory schooling law is implemented. Figure 5.4.3.21 shows us that after 1960-1961 academic year, an increasing trend appears in the fraction of middle school graduates, and it continues over time, just one exception: Between 1979-1980 and 1980-1981 academic year the number of middle school graduates have decreased by approximately 22 per cent, most probably due to the high terrorist attacks and the military coup afterwards. For high school graduates the result is the same as in the middle school graduates case: While there is a continuous increasing trend after 1960-1961 academic year, in 1980-1981 academic year number of high school graduates decreased by approximately 21 per cent.

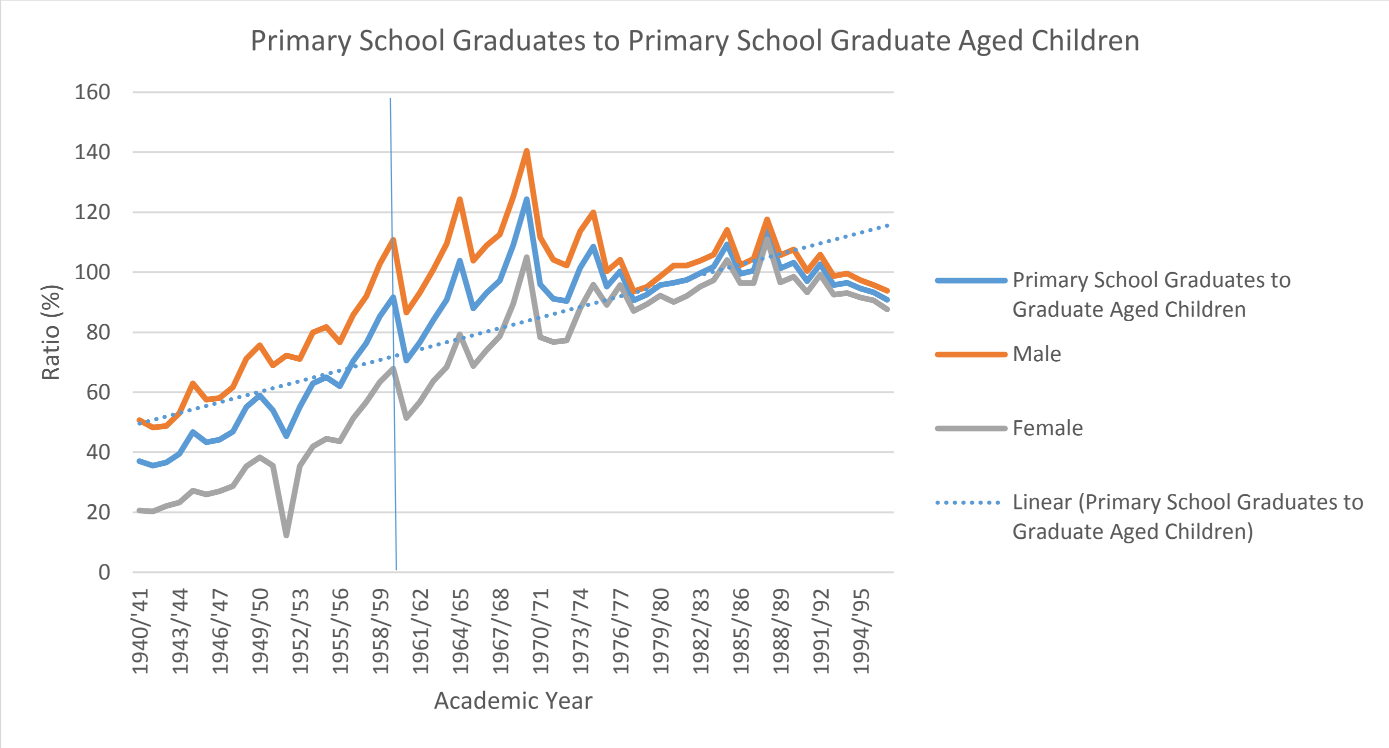


Figure 5.4.3.20. Primary School Graduates to Primary School Graduate Aged Population

Source: TURKSTAT

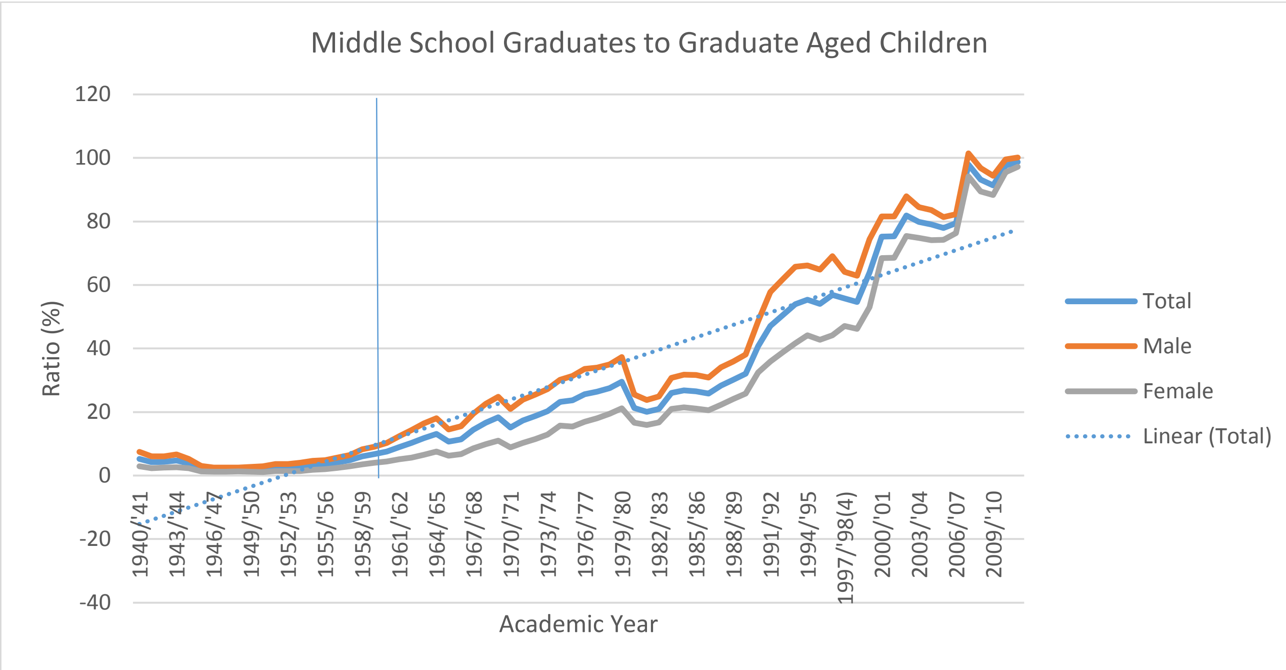


Figure 5.4.3.21. Middle School Graduates to Middle School Graduate Aged Population

Source: TURKSTAT

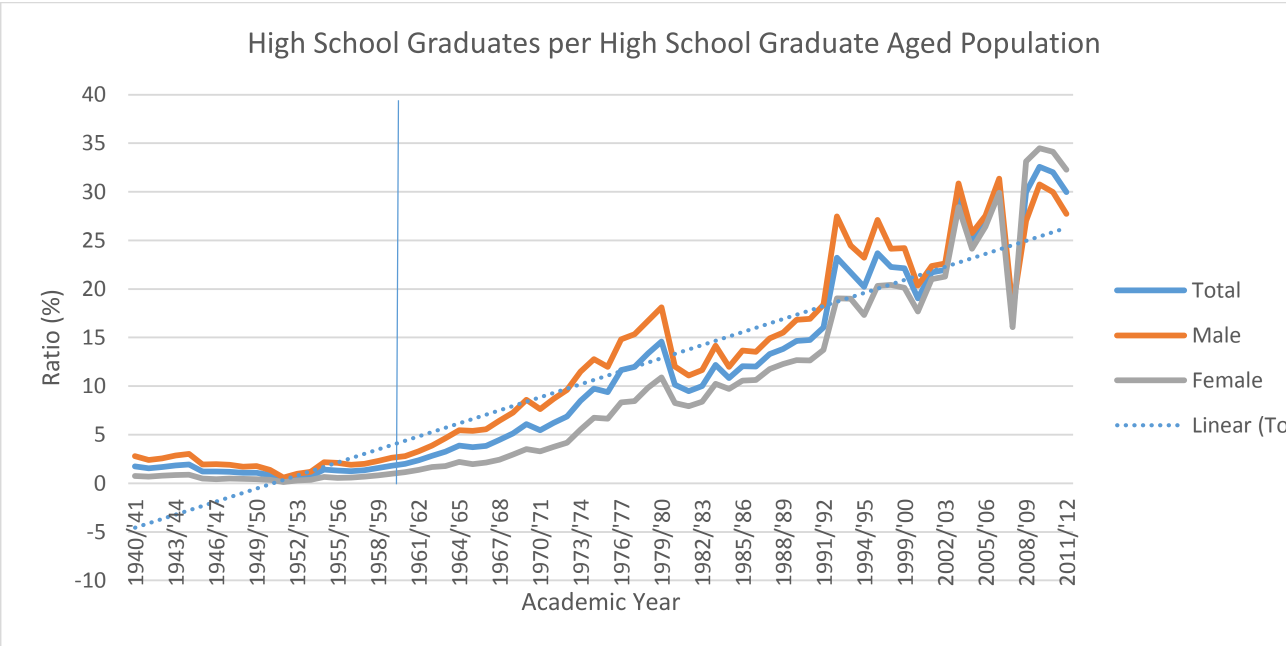


Figure 5.4.3.22. High School Graduates to High School Graduate Aged Population

In this section, we look at the evidence on both the health reforms and the education in the early 1960s in Turkey. It is clear that health reforms are not as successful as education reforms until the beginning of 1990s. In other words, health reforms do not coincide with the education reforms in the early 1960s. Therefore, we cannot infer that health reforms in the early 1960s cause improvement in the individuals' health. In contrast, we observe that the education reforms are implemented more rapidly than the health reforms. We observe that the improvements in the number of students, number of schools and number of teachers become more rapid beginning from the 1960s. In almost all the figures, we see that flatness of the lines disappear beginning from the early 1960s. In short, we conclude that educational expansion in the early 1960s has considerable impact on the improvement in the individual's education level, however we cannot say that health reforms have significant impact on the individual's health condition.

CHAPTER 6

SELF-ASSESSED HEALTH AND EDUCATION

6.1. Introduction

In this empirical part of the thesis, our first objective is to examine the validity of the education gradient of health by considering self-assessed health status (SAH) as the health outcome. This empirical part is new in the sense that it investigates the relationship between SAH and education for the first time in Turkey, a developing country, by using micro-data for adult men and women in Turkey.

In this chapter, first, we examine the association between health and education by using ordered probability models (ordered probit and ordered logit). We test the validity of the education gradient of health in Turkey for the individuals who are above 25 years of age since approximately at age 25 individuals complete their schooling in Turkey. By restricting the sample to individuals who are 25 or over we circumvent the problem of individuals who have not yet completed their education. Both ordered probability models support the positive association between health and education in Turkey as in the case of developed countries.

While investigating the linkage between individual's SAH and education level, we first control for gender, age, the region where the individual lives, marital status and the year dummy variables. In other words, we first examine the effect of exogenous factors on individual's SAH[31]. Next, we extend the two ordered probability models by controlling other variables such as the individual's employment status and the household income (in logarithms). We aim to see whether the positive effect of education on SAH still remains after controlling for the factors that may be endogenous to individual's SAH. We observe that although the magnitude of education coefficient gets smaller, the positive effect of education on SAH still remains when we control for labor market indicators and household income.

Ordered probability models suggest that in Turkey, males report better health conditions compared to females. Age is negatively associated with SAH; old individuals are more likely to report poor SAH. However, the region that the individual lives does not significantly affect his/her reported SAH. The results also indicate that household income level is positively associated with good SAH. In addition, we observe that both employed and unemployed individuals are more likely to state good and very good health relative to inactive individuals.

Next, we move to Instrumental Variable (IV) estimation process in order to overcome the endogeneity problem between health and education as discussed in Chapter Five. IV estimation results also support the positive association between the individual's SAH and years of schooling. Increase in years of education provides better SAH for both men and women in Turkey. We find that our instrument is strongly valid and that change in education policies in 1960 and 1961 lead to increase in the individual's years of schooling. Similar to ordered probit results, IV results indicate that men tend to report better SAH than women and SAH deteriorates with age. Finally, IV estimates suggest that increase in years of schooling have higher returns to SAH for men than for women.

The rest of the chapter is organized as follows: Section 6.2 presents the empirical results for ordered probability models. In Sections 6.2.1 and 6.2.2, we present the results of ordered probit and ordered logit models respectively. Next, in section 6.2.3

[31] According to 2006 Family Structure Survey, approximately 96 per cent of Turkish people marry at least once in their life. Therefore, we confidently assume that there is no two way causality between SAH and marital status. Hence, we refer it as exogenous variable.

we extend both models by adding different control variables. In section 6.3. we report the empirical results of IV estimation. Finally, Section 6.4 presents concluding remarks.

6.2. Empirical Specification for the Ordered Probability Models

6.2.1. Ordered Probit Analysis

In this section, the determinants of variation in SAH is analyzed using an ordered probit model. More specifically;

$$SAH_i = E_i \beta_1 + X_i\beta_2 + m + \varepsilon_i \qquad (6.2.1.1)$$

In equation (6.2.1.1), the dependent variable is SAH, which is equal to 1 if the individual i reports "very poor" health, 2 if the individual i reports "poor" health, 3 if the individual i reports average health, 4 if the individual i reports "good" and 5 if the individual reports "very good" health. The vector E includes the individual's education level.

In each specification, we first analyze the effects of education on individual's SAH by considering years of schooling as the education control. We have discussed how we calculate the individual's average years of schooling in Chapter Three. Next, we replicate our analysis where we take the education dummies as the education controls. We have six education dummies. The first one is equal to 1 if the individual is illiterate, 0 otherwise. The second is equal to 1 if the individual is non-graduate, 0 otherwise. The third is equal to 1 if the individual finishes primary school, 0 otherwise. The fourth is equal to 1 if the individual finishes middle school, 0 otherwise. The fifth is equal to 1 if the individual is graduated from high school, 0 otherwise. Finally, the last one is equal to 1 if the individual has university or higher degree, 0 otherwise. In the analysis, the omitted category is the dummy variable which is equal to 1 if the individual is illitirate.

The vector X includes the individual's demographic factors such as gender, age, region, and marital status. In the extended model, the vector X also contains the labor market indicators for the individual as well as the household income in logarithms. We do not include labor market indicators and household income as the determinants of SAH in our original model due to the possible endogeneity problem between SAH and labor market indicators and household income (Lindeboom and Kerkhofs, 2009). Finally, m refers to the year dummy. In ordered probit analysis, the omitted category for year dummies is the year dummy for 2008, which is equal to 1 if the year is 2008 and is equal to 0 if the year is 2010 or 2012. We assume that the unobserved determinants of SAH, ε, are normally distributed. The model does not consist of intercept term. [$See\ Wooldridge$ (2002)$Page$ 505].

Ordered probit coefficients and the corresponding marginal effects are presented in tables 6.2.1.1 and 6.2.1.2. In table 6.2.1.1, the education control is the individual's years of schooling and table 6.2.1.2, the education controls are the education dummies. Both specifications support the positive and significant relationship between individual's SAH and education level. Regarding other covariates, results from both specifications reach the same conclusions. The significances of the other covariates do not change among two models and the coefficients as well as the marginal effects are nearly the same in both models. It is important to note that the magnitude of the ordered probit coefficient does not give single interpretation. It just gives the idea about the direction and significance of independent variable on health outcome. In order to examine how SAH varies with education, gender, age, region and marital status we also calculate the marginal effects of explanatory variables on each outcome. Table 6.2.1.1 reveals that additional years of schooling increases the probability of reporting very good health by 0.56 percentage points. In addition, increase in years of schooling leads to 1.45 percentage increase in probability of reporting good health. We also observe that additional years of schooling decreases the chance of reporting very poor health by 0.08 percentage points and it decreases the chance of reporting poor health by 0.67 percentage points and it leads to decrease in likelihood of reporting average health by 1.25 percentage points.

In table 6.2.1.2, for education variables, our omitted category is “illiterate”. The coefficient results clearly indicate that probability of reporting good SAH is higher for individuals with upper levels of education relative to illiterate ones. Having university or higher degree increases the chance of reporting very good health by 14.47 percentage points and it increases the probability of reporting good health by 13.07 percentage points. These percentages amount to 10.09 and 12.73 for high school graduates, 7.25 and 10.21 for middle school graduates, 4.28 and 10.23 for primary school graduates and 2.06 and 4.34 for non-graduates. Having university or higher degree decreases the probability of reporting very poor health by 0.65 percentage points and being a high school graduate decreases the probability of reporting very poor health by 0.61 percentage points. In addition, having university or higher degree decreases the chance of reporting poor health by 7.05 percentage points and having high school degree decreases the likelihood of reporting poor health by 6.24 percentage points. These percentages are higher relative to primary and middle school graduates as well as non-graduates.

Ordered probit results indicate that males are more likely to report good SAH. The marginal effects from Table 6.2.1.1 suggest that the probability of stating very good health for males is 3.48 percentage points higher than that of females. The probability of reporting good health increases by 8.55 percentage points if the individual is male. Marginal effects also show that probability of reporting very poor health decreases by 0.48 percentage points, and the chance of reporting poor health decreases by 4 percentage points if the respondent is male.

Ordered probit results also suggest that age is negatively associated with reporting good health. Table 6.2.1.1 suggests that each year of age decreases the probability of reporting very good health by 3.34 percentage points and it decreases the probability of reporting good health by 8.64 percentage points. Ordered probability results reveal that each years of age increases the probability of reporting very poor health by 0.48

percentage points and it increases the chance of reporting poor health by 4.01 percentage points. The variable "Age-squared" is also found to be positively significant. Hence, there is a significant and convex relationship between SAH and age of the individual. Thus, we can state that as people gets older, the probability of stating good health increases as people are more likely to compare their SAHs with those of their contemporaries as Zimmer et al.(2000) suggest.

In ordered probit regression, for marital status, the omitted category is "widowed/divorced". The results suggest that both married and single individuals are more likely to report good health relative to widowed/divorced individuals. Table 5.2.1.1 reveals that being married increases the probability of reporting very good and good health by 0.54 and 1.46 percentage points respectively. On the other hand being married decreases the probability of reporting very poor and poor health by 0.08 and 0.68 percentage points. The chance of reporting very good health increases by 1.27 percentage points for single individuals, whereas the likelihood of stating good health increases by 2.89 percentage points for those individuals. Being single decreases the probability of reporting very poor and poor health by 0.15 and 1.33 percentage points respectively.

In ordered probit analysis, regarding region, the variable "urban" is equal to 1 if the individual lives in urban area and it is equal to 0 if the individual lives in rural area. Surprisingly, the ordered probit results suggest that the region that the individual lives does not significantly affect his/her SAH. Finally, we observe that there is a significant increase in probability of reporting very good or good health in Turkey over time. Our results suggest that the chance of stating good or very good SAH is higher in both 2010 and 2012 compared to 2008. For both years, the rise in the chance of reporting good SAH is significant, and the rise is even higher in 2012. Accordingly, there is a significant decrease in probability of reporting very poor and poor health in 2010 and 2012 compared to 2008.

Table 6.2.1.1. Ordered Probit Results (Education Control: Years of Schooling)

Variable	Coefficients	Marginal Effects (x100) (SAH=1)	Marginal Effects (x100) (SAH=2)	Marginal Effects (x100) (SAH=3)	Marginal Effects (x100) (SAH=4)	Marginal Effects (x100) (SAH=5)
Non-Graduate	0.17*** (0.03)	-0.22*** (0.0002)	-2.00*** (0.002)	-4.20*** (0.006)	4.34*** (0.005)	2.06*** (0.003)
Primary School Graduate	0.38*** (0.02)	-0.57*** (0.0003)	-4.80*** (0.002)	-9.13*** (0.004)	10.23*** (0.004)	4.28*** (0.002)
Middle School Graduate	0.49*** (0.02)	-0.47*** (0.0002)	-4.84*** (0.001)	-12.15*** (0.005)	10.21*** (0.003)	7.25*** (0.004)
High School Graduate	0.66*** (0.02)	-0.61*** (0.0003)	-6.24*** (0.001)	-15.97*** (0.005)	12.73*** (0.002)	10.09*** (0.004)
University+	0.83*** (0.02)	-0.65*** (0.0003)	-7.05*** (0.001)	-19.84*** (0.005)	13.07*** (0.002)	14.47*** (0.005)
Male	0.30*** (0.01)	-0.45*** (0.0002)	-3.85*** (0.001)	-7.29*** (0.002)	8.24*** (0.003)	3.35*** (0.001)
Age($x10^{-1}$)	-0.32*** (0.02)	0.50*** (0.0004)	4.24*** (0.002)	7.91*** (0.005)	-9.11*** (0.006)	-3.53*** (0.002)

Table 6.2.1.1. (Continued)

Age Squared ($x10^{-3}$)	0.06*** (0.02)	-0.09*** (0.0003)	-0.83*** (0.002)	-1.54*** (0.005)	1.78*** (0.006)	0.69*** (0.002)
Urban	0.01 (0.01)	-0.02 (0.0001)	-0.21 (0.001)	-0.39 (0.002)	0.45 (0.003)	0.17 (0.001)
Married	0.04** (0.02)	-0.07** (0.0003)	-0.55** (0.002)	-1.00** (0.004)	1.19** (0.005)	0.44** (0.001)
Single	0.11*** (0.02)	-0.15*** (0.0003)	-1.38*** (0.003)	-2.78*** (0.007)	3.00*** (0.007)	1.32*** (0.003)
Year10	0.03** (0.01)	-0.04** (0.0002)	-0.41** (0.001)	-0.77** (0.003)	0.88** (0.003)	0.35** (0.001)
Year12	0.21*** (0.01)	-0.32*** (0.0002)	-2.66*** (0.001)	-4.95*** (0.003)	5.70*** (0.003)	2.23*** (0.001)
(-)Log-Likelihood	48566					
Pseudo R2	0.11					
Number of Observations	46473					

Source: Author's Computations from 2008,2010 and 2012 pooled THS data set Notes: (1) Robust standard errors are shown in parenthesis (2) *** indicates 1 per cent level of significance, ** indicates 5 per cent level of significance and * indicates 10 per cent level of significance

Table 6.2.1.2. Ordered Probit Results (Education Control: Education Dummies)

Variable	**Coefficients**	**Marginal Effects (x100) (SAH=1)**	**Marginal Effects (x100) (SAH=2)**	**Marginal Effects (x100) (SAH=3)**	**Marginal Effects (x100) (SAH=4)**	**Marginal Effects (x100) (SAH=5)**
Years of Schooling	0.05*** (0.001)	-0.08*** (0.00004)	-0.67*** (0.0001)	-1.25*** (0.0003)	1.45*** (0.0004)	0.56*** (0.0001)
Male	0.31*** (0.01)	-0.48*** (0.0002)	-4.00*** (0.001)	-7.55*** (0.002)	8.55*** (0.003)	3.48*** (0.001)
Age($x10^{-1}$)	-0.31*** (0.02)	0.48*** (0.0004)	4.01*** (0.002)	7.48*** (0.005)	-8.64*** (0.006)	-3.34*** (0.002)
Age Squared ($x10^{-3}$)	0.06** (0.02)	-0.06** (0.0003)	-0.54** (0.002)	-1.00** (0.005)	1.16** (0.006)	0.45** (0.002)
Urban	0.01 (0.01)	-0.03 (0.0001)	-0.23 (0.001)	-0.42 (0.002)	0.48 (0.003)	0.18 (0.001)
Married	0.05** (0.02)	-0.08** (0.0003)	-0.68** (0.002)	-1.24** (0.004)	1.46** (0.005)	0.54** (0.001)

Table 6.2.1.2. (Continued)

Single	0.11*** (0.03)	-0.15*** (0.0003)	-1.33*** (0.003)	-2.68*** (0.007)	2.89*** (0.007)	1.27*** (0.003)
Year10	0.03** (0.01)	-0.05** (0.0002)	-0.42** (0.001)	-0.79** (0.003)	0.91** (0.003)	0.36** (0.001)
Year12	0.21*** (0.01)	-0.03*** (0.0002)	-2.67*** (0.001)	-4.96*** (0.003)	5.72*** (0.003)	2.23*** (0.001)
(-)Log-Likelihood	49607					
Pseudo R2	0.11					
Number of Observations	46473					

Source: Author's Computations from 2008,2010 and 2012 pooled THS data set

Notes: (1) Robust standard errors are shown in parenthesis

(2) *** indicates 1 per cent level of significance, ** indicates 5 per cent level of significance and * indicates 10 per cent level of significance

6.2.2. Ordered Logit Analysis

We re-estimate equation (6.2.1.1) by assuming that the unobserved determinants of SAH, ε, are logistic distributed. Thus, we estimate an ordered logit model. The dependent and explanatory variables are exactly defined in the same way as in ordered probit regression. The sign and significance of the coefficients do not differ between the two models. Although, the magnitudes of the coefficients are somewhat higher in ordered logistic regression, the marginal effects are more or less the same in these two models. Nevertheless, we run an ordered logit regression in order to obtain the odds ratios of the covariates. We cannot obtain odds ratios with ordered probit regression. However, in the studies related to statistics, interpreting the coefficients in terms of marginal effects on the probability is not so common. Rather, the coefficients are generally inferred in terms of marginal effects on odds ratio (Cameron and Trivedi, 2005, p.470). In fact, the thing that makes the logit regression attractive to researchers is the interpretation of the coefficients in terms of log-odds ratios (Cameron and Trivedi, 2005, p.472).

In the ordered logit model we estimate a single equation over different levels of the dependent variable. Assuming that we observe the variations in levels of the dependent variable (SAH) in a cumulative sense, interpreting the coefficients in odds enable us to compare the people who belong to the groups greater than *s* versus those who belong to groups less than or equal to *s*, where *s* is the outcome of the model, individual's SAH in our study. For instance, assume that for an odds ratio y, we compare the people who report very good SAH with people who report good, average, poor and very poor SAH. If the odds ratio is greater than 1, we conclude that one unit increase in covariate x leads the odds of reporting very good SAH versus combined good, average, poor and very poor SAH categories are y times greater assuming that other variables in the model are constant[32]. The tables 6.2.2.1 and 6.2.2.2 report the ordered logit regression results.

Table 6.2.2.1. Ordered Logit Results: (Education Control: Years of Schooling)

	(1)	(2)
VARIABLES	Coefficient	Odds Ratio
Years of Schooling	0.0914***	1.096***
	(0.00244)	(0.00267)
Male	0.577***	1.781***
	(0.0198)	(0.0352)
Age ($x10^{-1}$)	-0.565***	0.568***
	(0.0409)	(0.0233)
Age Square ($x10^{-3}$)	0.0832**	1.087**
	(0.0393)	(0.0427)
Urban	0.0395*	1.040*
	(0.0213)	(0.0221)
Married	0.0887***	1.093***
	(0.0330)	(0.0361)
Single	0.223***	1.249***
	(0.0503)	(0.0629)
Year10	0.0647**	1.067**
	(0.0257)	(0.0274)
Year12	0.371***	1.449***
	(0.0226)	(0.0327)
Pseudo R2 0.11		
(-) Log-Likelihood 49469		
Observations	46,473	46,473

Source: Author's Computations from 2008,2010 and 2012 pooled THS data set

Notes: (1) Robust standard errors are shown in parenthesis

(2) *** indicates 1 per cent level of significance, ** indicates 5 per cent level of significance and * indicates 10 per cent level of significance

[32] Source: http://www.ats.ucla.edu/stat/stata/output/stata_ologit_output.htm

Table 6.2.2.2. Ordered Logit Results (Education Control: Education Dummies)

	(1)	(2)
VARIABLES	Coefficient	Odds Ratio
Non-Graduate	0.31***	1.35
	(0.04)	
Primary School Graduate	0.65***	1.92
	(0.03)	
Middle School Graduate	0.86***	2.35
	(0.04)	
High School Graduate	1.14***	3.11
	(0.04)	
University+	1.45***	4.27
	(0.04)	
Male	0.56***	1.75
	(0.02)	
Age ($x10^{-1}$)	-0.59	0.55
	(0.04)	
Age Square ($x10^{-3}$)	0.12***	1.13
	(0.04)	
Urban	0.04*	1.03
	(0.02)	
Married	0.07**	1.08
	(0.03)	
Single	0.23***	1.26
	(0.05)	
Year10	0.06**	1.06
	(0.03)	
Year12	0.37***	1.4
	(0.02)	
Pseudo R2	**0.11**	
(-)Log-Likelihood	**49443**	
Observations	46,473	

Source: Author's Computations from 2008,2010 and 2012 pooled THS data set

Notes: (1) Robust standard errors are shown in parenthesis

(2) *** indicates 1 per cent level of significance, ** indicates 5 per cent level of significance and * indicates 10 per cent level of significance

Ordered logit regression results suggest the positive association between the individual's SAH and education level. Table 6.2.2.1 reveals that for an additional year of schooling, the odds of reporting very good SAH versus the combined good, average, poor and very poor SAH categories are 1.09 times greater, given the other variables are constant in the model. Similarly, for an additional years of schooling, the odds of the combined very good, good, average and poor SAH versus very poor SAH are 1.09 times greater, given the other variables are constant.

The positive and significant relationship between individual's SAH and education level remains when we include education dummies into our model instead of years of schooling. We observe that if the individual has university or higher degree the odds of reporting very good SAH is 4.27 times greater than for high school graduates, middle school graduates, primary school graduates and non-graduates. Table 6.2.2.2 reveals that the odds of reporting very good health compared to other categories increases with education level.

Table 6.2.2.1 suggests that for males, the odds of stating very good health versus the combined good, average, poor and very poor health are 1.781 times higher than for females, assuming that other variables are held constant. Different from our previous ordered probit results, ordered logit results suggest that the region where the individual lives affect the individual's SAH at 10 per cent significance level. The results indicate that for the individuals living in urban areas, the odds of stating very good health versus the combined good, average, poor and very poor health are 1.04 times higher than for people living in rural areas, assuming that other variables are held constant.

Regarding marital status, Table 6.2.2.1 shows that for married people, the odds of stating very good health versus the combined good, average, poor and very poor health are 1.09 times greater than for single and widowed/divorced people when other variables are held constant. For single individuals, the odds of stating very good health versus the combined good, average, poor and very poor health are 1.25 times higher than for married and widowed/divorced individuals when other variables are held constant.

6.2.3. Extension of the Ordered Models

6.2.3.1. Ordered Probit Analysis

As an extension of the ordered probability models we add individual's employment status and logarithm of household income gradually into our model. For labor market indicators, employment status of the individual is classified into 3 groups: Employed, Unemployed and Inactive. We have defined three dummy variables for each employment status. The first dummy variable is equal to 1 if the individual is employed, 0 otherwise. The second is equal to 1 if the individual is unemployed, 0 otherwise and the third dummy is equal to 1 if the individual is inactive, 0 otherwise.

Tables 6.2.3.1.1 and 6.2.3.1.2 present the coefficients and marginal effects from the ordered probit analysis where the education control is the years of schooling, and the tables 6.2.3.1.3 and 6.2.3.1.4 present the coefficients and marginal effects where the education controls are the education dummies.

Table 6.2.3.1.1. Ordered Probit Results for the Extended Models (Education Control: Years of Schooling)

	(1)	(2)
VARIABLES	Ordered Probit Coefficient	Ordered Probit Coefficient
Years of Schooling	0.05***	0.04***
	(0.001)	(0.001)
Male	0.25***	0.28***
	(0.01)	(0.01)
Age (x10^{-1})	-0.32***	-0.36***
	(0.02)	(0.02)
Age Square (x10^{-3})	0.06***	0.09***
	(0.02)	(0.02)
Urban	0.04***	-0.02*
	(0.01)	(0.01)
Married	0.07***	0.04*
	(0.02)	(0.02)
Single	0.13***	0.08***
	(0.03)	(0.03)
Employed	0.15***	0.13***
	(0.01)	(0.01)
Unemployed	0.02	0.10***
	(0.03)	(0.03)
Log Household Income		0.22***
		(0.01)

Table 6.2.3.1.1. (Continued)

Year10	0.03**	0.0004
	(0.01)	(0.01)
Year12	0.21***	0.14***
	(0.01)	(0.01)
Pseudo R2	0.11	0.11
(-) Log-Likelihood	49544	48864
Observations	46,473	46,004

Source: Author's Computations from 2008,2010 and 2012 pooled THS data set

Notes: (1) Robust standard errors are shown in parenthesis

(2) *** indicates 1 per cent level of significance, ** indicates 5 per cent level of significance and * indicates 10 per cent level of significance

Table 6.2.3.1.2 Marginal Effects for the Extended Models [33]

Variable	Marginal Effects (x100) (SAH=1)	Marginal Effects (x100) (SAH=2)	Marginal Effects (x100) (SAH=3)	Marginal Effects (x100) (SAH=4)	Marginal Effects (x100) (SAH=5)
Years of Schooling	-0.05***	-0.46***	-0.87***	1.00***	0.38***
Male	-0.41***	3.55***	-6.83***	7.73***	3.06***
Age($x10^{-1}$)	0.52***	4.59***	8.73***	-10.05***	-3.79***
Age Squared ($x10^{-3}$)	-0.13***	-1.18***	-2.24***	2.58***	0.97**
Urban	0.03**	0.3**	0.57**	-0.66**	-0.25**
Married	-0.05***	-0.46***	-0.87***	1.01***	0.37***
Single	-0.11***	-1.01***	-2.04***	2.24***	0.92***

[33] The marginal effects are for the extended model where we include both the labor market indicators and the household income.

Table 6.2.3.1.2. (Continued)

Employed	-0.18***	-1.60***	-3.10***	3.51***	1.37***
Unemployed	-0.12***	-1.18***	-2.41***	2.60***	2.11***
Log Household Income	-0.31***	-2.75***	-5.21***	6.00***	2.27***
Year10	-5.99e-04	-0.005	-0.009	0.01	0.004
Year12	-0.20***	-1.73***	-3.29***	3.78***	1.44***

Source: Author's Computations from 2008,2010 and 2012 pooled THS data set

Note: *** indicates 1 per cent level of significance, ** indicates 5 per cent level of significance and * indicates 10 per cent level of significance

Table 6.2.3.1.3. Ordered Probit Results for the extended models (Education Control: Education Dummies)

	(1)	(2)
VARIABLES	Coefficent	Coefficient
Non-Graduate	0.17***	0.16***
	(0.03)	(0.03)
Primary School Graduate	0.38***	0.33***
	(0.02)	(0.02)
Middle School Graduate	0.49***	0.39***
	(0.02)	(0.03)
High School Graduate	0.65***	0.51***
	(0.02)	(0.02)
University+	0.79***	0.59***
	(0.02)	(0.03)
Male	0.23***	0.26***
	(0.01)	(0.01)
Age($x10^{-1}$)	-0.33***	-0.38***
	(0.02)	(0.02)
Age Squared ($x10^{-3}$)	0.09***	0.12***
	(0.02)	(0.02)
Urban	0.04***	-0.03**
	(0.01)	(0.01)
Married	0.06***	0.03
	(0.02)	(0.02)
Single	0.13***	0.09***
	(0.03)	(0.03)
Employed	0.16***	0.14***
	(0.01)	(0.01)

Table 6.2.3.1.3. (Continued)

Unemployed	0.02	0.11***
	(0.03)	(0.03)
Log Household Income		0.22***
		(0.01)
Year10	0.03**	-0.001
	(0.01)	(0.01)
Year12	0.21***	0.13***
	(0.01)	(0.01)
Pseudo R2	0.11	0.11
(-) Log-Likelihood	49506	48821
Observations	46,473	46,004

Source: Author's Computations from 2008,2010 and 2012 pooled THS data set

Notes: (1) Robust standard errors are shown in parenthesis

(2) *** indicates 1 per cent level of significance, ** indicates 5 per cent level of significance and * indicates 10 per cent level of significance

***Table 6.2.3.1.4. Marginal Effects for the Extended Models*[34]**

Variable	Marginal Effects (x100) (SAH=1)	Marginal Effects (x100) (SAH=2)	Marginal Effects (x100) (SAH=3)	Marginal Effects (x100) (SAH=4)	Marginal Effects (x100) (SAH=5)
Non-Graduate	-0.18***	-1.81***	-3.85***	4.02***	1.83***
Primary School Graduate	-0.46***	-4.06***	-7.88***	8.84***	3.56***
Middle School Graduate	-0.38***	-4.00***	-9.73***	8.82***	5.31***
High School Graduate	-0.48***	-5.10***	-12.59***	11.06***	7.11***
University+	-0.51***	-5.55***	-14.56***	11.77***	8.85***
Male	-0.37***	-3.29***	-6.35***	7.18***	2.83***
Age($x10^{-1}$)	0.54***	4.86***	9.24***	-10.62***	-4.02***
Age Squared ($x10^{-3}$)	-0.17***	-1.54***	-2.92***	3.36***	1.27**

[34] The marginal effects are for the extended model where we include both the labor market indicators and the household income.

Table 6.2.3.1.4 (Continued)

Urban	0.04**	0.32**	0.61**	-0.71**	-0.27**
Married	-0.04	-0.33	-0.61	0.71	0.26
Single	-0.12***	-1.09***	-2.21***	2.41***	1.00***
Employed	-0.19***	-1.76***	-3.42***	3.86***	1.52***
Unemployed	-0.13***	-1.25***	-2.59***	2.78***	1.20***
Log Household Income	-0.31***	-2.79***	-5.31***	6.11***	2.31***
Year10	0.001	0.01	0.03	-0.04	-0.01
Year12	-0.19***	-1.71***	-3.24***	3.73***	1.42***

Source: Author's Computations from 2008,2010 and 2012 pooled THS data set

Note*** indicates 1 per cent level of significance, ** indicates 5 per cent level of significance and * indicates 10 per cent level of significance.

As new control variables are added into the model, we see that the marginal effect of years of schooling on the individual's SAH decreases, but it is still significant. For instance, while additional years of schooling the increases the probability of reporting good health by 1.45 percentage points when the labor market indicators and the household income are not controlled, it decreases to 1.00 when these variables are controlled. Similarly, while more years of schooling increases the probability of reporting very good health by 0.56 percentage points in the original model, the percentage decreases to 0.38 in the extended model.

When we add indicators of employment status and logarithm of the household income gradually, we observe two main departures from our main model. First, we observe that the chance of reporting very good or good health significantly increases in urban areas when we add employment indicators. Second, when we introduce the logarithm of household income into our model we observe a significant increase in probability of stating very good or good health in rural areas.

In ordered probit regression analysis, for labor market indicator dummies, our omitted category is inactive individuals. The results reveal that employed people have more tendency to report good health relative to inactive people. However, the tendency to report good health does not significantly differ for unemployed people relative to inactive people when we do not introduce logarithm of household income into our model. Table 6.2.3.1.2 suggests that having a regular job increases the probability of reporting very good and good health by 1.37 and 3.51 percentage points respectively. We also observe that being employed decreases the probability of reporting very poor SAH by 0.18 percentage points and it decreases the chance of reporting poor SAH by 1.60 percentage points.

In the ordered probit regression analysis, we show that the logarithm of the household income positively and significantly affects the reported SAH of the individual. Table 6.2.3.1.2 suggests that each increase in the logarithm of household income level leads to 2.27 percentage points rise in probability of reporting very good health and it leads to 6 percentage points rise in the chance of reporting good health.

6.2.3.2. Ordered Logit Analysis

In this section, we replicate the previous extended models by assuming that the error terms are logistic distributed. Tables 6.2.3.2.1 and 6.2.3.2.2 report the results of the extended ordered logit models.

Table 6.2.3.2.1 Ordered Logit Results for the Extended Models (Education Control: Years of Schooling)

	(1)	(2)	(3)	(4)
VARIABLES	Coefficient	Odds Ratio	Coefficient	Odds Ratio
Years of Schooling	0.0875***	1.091***	0.0641***	1.066***
	(0.00246)	(0.00269)	(0.00275)	(0.00293)
Male	0.474***	1.606***	0.527***	1.693***
	(0.0235)	(0.0377)	(0.0237)	(0.0402)
Age ($x10^{-1}$)	-0.574***	0.563***	-0.655***	0.520***
	(0.0408)	(0.0230)	(0.0412)	(0.0214)
Age Square ($x10^{-3}$)	0.114***	1.121***	0.173***	1.189***
	(0.0393)	(0.0441)	(0.0396)	(0.0471)
Urban	0.0688***	1.071***	-0.0338	0.967
	(0.0215)	(0.0230)	(0.0221)	(0.0214)
Married	0.113***	1.119***	0.0581*	1.060*
	(0.0331)	(0.0371)	(0.0334)	(0.0354)
Single	0.247***	1.280***	0.173***	1.189***
	(0.0505)	(0.0646)	(0.0510)	(0.0606)
Employed	0.240***	1.271***	0.206***	1.229***
	(0.0250)	(0.0318)	(0.0252)	(0.0309)
Unemployed	0.0226	1.023	0.164***	1.178***
	(0.0602)	(0.0616)	(0.0614)	(0.0723)
Log Household Income			0.376***	1.457***
			(0.0190)	(0.0276)

Table 6.2.3.2.1. (Continued)

Year10	0.0629**	1.065**	0.00908	1.009
	(0.0257)	(0.0274)	(0.0259)	(0.0262)
Year12	0.370***	1.447***	0.246***	1.279***
	(0.0226)	(0.0326)	(0.0234)	(0.0300)
Pseudo R2	0.11		0.11	
(-) Log-Likelihood	49417		48742	
Observations	46,473	46,473	46,004	46,004

Source: Author's Computations from 2008,2010 and 2012 pooled THS data set

Notes: (1) Robust standard errors are shown in parenthesis

(2) *** indicates 1 per cent level of significance, ** indicates 5 per cent level of significance and * indicates 10 per cent level of significance

Table 6.2.3.2.2. Ordered Logit Results for the Extended Models (Education Control: Education Dummies)

	(1)	(2)	(3)	(4)
VARIABLES	Coefficient	Odds Ratio	Coefficient	Odds Ratio
Non-Graduate	0.32***	1.371	0.27***	1.31
	(0.04)		(0.04)	
Primary School Graduate	0.66***	1.934	0.55***	1.74
	(0.03)		(0.03)	
Middle School Graduate	0.85***	2.347	0.68***	1.97
	(0.04)		(0.04)	
High School Graduate	1.12***	3.074	0.88***	2.41
	(0.04)		(0.04)	
University+	1.39***	4.018	1.03***	2.81
	(0.04)		(0.05)	
Male	0.44***	1.556	0.49***	1.64
	(0.02)		(0.02)	
Age ($x10^{-1}$)	-0.61***	0.546	-0.69***	0.50
	(0.04)		(0.04)	
Age Square ($x10^{-3}$)	0.16***	1.170	0.22***	1.25
	(0.04)		(0.04)	
Urban	0.07***	1.071	-0.04	0.96
	(0.02)		(0.02)	
Married	0.10***	1.103	0.04	1.04
	(0.03)		(0.03)	

Table 6.2.3.2.2. (Continued)

Single	0.26***	1.293	0.18***	1.20
	(0.05)		(0.05)	
Employed	0.26***	1.296	0.23***	1.25
	(0.03)		(0.03)	
Unemployed	0.03	1.033	0.18***	1.19
	(0.06)		(0.06)	
Log Household Income			0.38***	1.47
			(0.02)	
Year10	0.06**	1.063	0.006	1.006
	(0.03)		(0.03)	
Year12	0.37***	1.445	0.24***	1.28
	(0.03)		(0.02)	
Pseudo R2	0.11		0.11	
(-)Log-Likelihood	49384		48704	
Observations	46,473		46,004	

Source: Author's Computations from 2008,2010 and 2012 pooled THS data set

Notes: (1) Robust standard errors are shown in parenthesis

(2) *** indicates 1 per cent level of significance, ** indicates 5 per cent level of significance and * indicates 10 per cent level of significance

First, we observe that the odds ratio for years of schooling decreases to 1.066 when we include labor market indicators and logarithm of household income to our model gradually. We also observe that for most of the covariates the coefficients and the odds ratios decrease but the signs and significances of the covariates remain same. The only exception is for the region dummy and marital status (married) dummy. We see that the significance of the region dummy variable disappears when we include the logarithm of the household income into our model. Moreover, the significance level of marital status (married) dummy raises to 10 percent when we add the logarithm of the household income to our model. The results also show that while being unemployed do not significantly affect individual's SAH, when we include the household income into our model, being unemployed positively and significantly affect the individual's SAH.

Regarding employment status, Table 6.2.3.2.1 shows that for employed people, the odds of stating very good health versus the combined good, average, poor and very poor health are 1.23 times greater than for unemployed and inactive people when other variables are held constant. For unemployed individuals, the odds of stating very good health versus the combined good, average, poor and very poor health are 1.18 times higher than for employed and inactive individuals when other variables are held constant.

6.3. IV Estimation Process

6.3.1. Empirical Specification

In order to see the effects of education on individual's years of schooling, we establish a two-stage model where the first stage is OLS regression, and the second stage regression is ordered probit regression. We simply call this model as "IVoprobit" model. We believe that the second stage regression should be an ordered probit regression since the dependent variable SAH is a categorical variable, hence running Two-Stage Least Squares (2SLS) regression, where one assume SAH is defined in continuous intervals, may be problematic.[35] Mullahy (1997) is one of the first studies that suggest using non-linear IV methods in which the outcome is discrete variable.

We use education reforms in the early 1960s in Turkey as the instrument for individual's education as we discuss in Chapter Five. In the empirical analysis we define the instrument as follows: The main education reform law has passed on

[35] Nevertheless, we also apply 2SLS estimation techniques by assuming that the dependent variable is measured on a continous interval scale. Both first stage and second stage of 2SLS estimation results are consistent with the estimation results from ivoprobit model. 2SLS estimation results are presented at the end of this chapter in the appendix part.

October 1960 and January 1961. Therefore, we assume that the pupils who start school or who are already enrolled in primary schooling in September 1961 or later are affected by the educational reforms. By considering the school starting age is 6 in Turkey[36], we assume that the individuals who were between 6 and 9 years old in September 1961, or who were born in or later than 1952 are affected by the education expansion. We can observe individual's age from THS data set. Hence, we calculate the birth year of the individual by subtracting the reported age from the survey year. Therefore, we are able to define a policy dummy which equals to 1 if the individual was born in 1952 or later, 0 if the individual was born before 1952. However, it is important to note that in Turkey, some parents prefer their children to start primary school at age 7, (especially the children who were born after September). Therefore, we replicate our estimations by considering the people who were born in or after 1951 are also affected from the educational reforms. Last, we redo our analysis by excluding the 1951 birth cohort. Because, some of the individuals born in that cohort may benefit from the reforms, and some of them may not.

The two stage model can be written as follows:

$$E_i = \alpha + X_i \alpha_1 + R61_i \alpha_2 + \varepsilon_i \qquad (6.3.1)$$

$$SAH_i = \delta + \hat{E}_i \beta_1 + X_i \beta_2 + u_i \qquad (6.3.2)$$

In the first-stage regression, our dependent variable is E, which shows the years of schooling completed by individual i. In the IV-estimation process we define the individual's years of schooling different from the discussion in chapter three. Because, the major education reform that we use as the instrument for the individual's years of schooling states that the compulsory schooling increases from 3 years to 5 years in the villages. Unfortunately, we cannot observe whether the individual lives in village or not when he/she attends to the primary school if he/she

[36] Source: "Milli Eğitim Temel Kanunu", http://mevzuat.meb.gov.tr/html/temkanun_1/temelkanun_1.html)

is enrolled in primary school before 1952. Therefore, we calculate a different average years of schooling by collecting the Ministry of Education statistics. From ministry of education statistics, we observe how many people are graduated from rural primary schools and from urban primary schools[37]. Then we calculate the average years of schooling for primary school graduates before 1952 in the following manner:

$$Years\ of\ Schooling = \frac{(Urban\ Graduates\ x\ 5) + (Rural\ Graduates\ x\ 3)}{Urban\ Graduates + Rural\ Graduates}$$

For the primary school graduates who finishes school in 1952 or later, years of schooling is equal to 5. Accordingly, we define the policy dummy "R61" is equal to 1 if the individual was born in or later than 1952, 0 otherwise. In the second stage regression, our dependent variable is SAH, which is equal to 1 if the individual i reports "very poor" health, 2 if the individual i reports "poor" health and 3 if the individual i reports "average" health, 4 if the individual i reports "good" health, 5 if the individual i reports "very good" health. The vector $\hat{E}$ includes the predicted values of education variable estimated from the first stage regression. In our analysis, the main parameter of interest is β_1 which shows the variations in SAH as the years of schooling increases by one unit.

In addition, the vector X includes the individual's demographic factors, gender, region and age dummies. For gender, we define a dummy variable which is equal to 1 if the individual is male, 0 if the individual is female. We define age dummies for each separate age. For instance, agedummy61 is equal to 1 if the individual is 61 years old at the time when the survey takes place, and it is equal to 0 otherwise. Including age dummies is more flexible specification than including age and trend and their higher polynomials (Aydemir and Kırdar, 2014). Since, it is probable that our instrument is correlated with cohort-level changes that affect educational outcomes, in some specifications we also control for cohort dummies in the empirical

[37] Rural areas include villages and districts and urban areas include cities and towns.

analysis. Cohort dummies are defined in 5- year intervals. By including the cohort dummies, our objective is to control for the cohort specific variations.

For region, we define a dummy variable which is equal to 1 if the individual lives in urban areas, 0 if the individual lives in rural areas. According to 2006 Family Structure Survey, approximately 96 per cent of Turkish people marry at least once in their life. Therefore, we may assume that there is no two way causality between SAH and marital status in Turkey. Hence, in some specifications, we also include marital status of the individual in vector X as an exogenous covariate. In the empirical analysis, we define three dummy variables for each marital status category (married, single, widowed/divorced). All the standard errors are clustered by birth year.

Following Roodman (2011), we apply "Conditional Mixed Process" (CMP) estimation techniques in order to estimate our two stage regression with user written STATA command "cmp", where the first stage regression is OLS and the second stage is ordered probit regression[38]. Roodman states that CMP fits with substantially larger class of simultaneous equation systems which have properties recursivity and observability. Recursitivity implies that the model should have clearly defined at least two stages with one or more equations in each stage. Observability implies that the equation include observed endogeneous variables on the right hand side. Roodman suggests that CMP is appropriate for the models in which there is simultaneity but instruments allow for the recursivity as in the two-stage least squares (2SLS) estimation. Therefore, implementing IV methodology by using "CMP" estimation techniques is appropriate for our model.

CMP is a limited-information maximum likelihood (LIML) estimator which assumes the errors are jointly distributed normally in the structural and in first-stage equations. The LIML estimator is asymptotically equivalent to 2SLS estimator (Staiger and Stock, 1997). Recent studies suggest that LIML has even more desirable finite sample properties especially if the instruments are not strong. For instance,

[38] STATA does not have a special command for ivoprobit regression.

some studies show that LIML lead to smaller biases compared to 2SLS or GMM estimators. (See Cameron and Trivedi (2010), p. 204). We simply call the regression which is implemented by CMP estimation techniques as "IVoprobit" regression. IVoprobit regression methods which are implemented by CMP techniques was previously used by Ndiaye et al (2013).

6.3.2. Empirical Results

6.3.2.1. Descriptive Statistics

The descriptive statistics indicate that in Turkey, people tend to report average or good health in general. For the whole sample, we observe that 21.64 per cent of the individuals are not affected from the education reforms in 1960s. In the pooled data set the percentage of females is 55 and the percentage of males is 45. The statistics show that approximately 71 per cent of individuals live in urban areas. Regarding marital status, we observe that 82 per cent of the individuals in our sample are married. Single and widowed/divorced individuals constitute 18 per cent of our sample. Finally, we detect that 50 per cent of our observations comes from 2012 THS Survey. The data from 2008 and 2010 constitutes the remaining fifty per cent of our sample.

Table 6.3.2.1.1. presents the summary statistics for the variables that we use to establish the causal link between SAH and education. We investigate the descriptive statistics for males and females separately. The first column presents the descriptive statistics for both treated and untreated groups for males, and the second column

presents the same statistics for females. Treated samples consist of the individuals who are affected by 1960s educational reforms, and the untreated samples consist of the individuals who are not affected from the reforms.

Table 6.3.2.1.1. Summary Statistics

Variable	Male		Female	
	Untreated	Treated	Untreated	Treated
SAH	3.22 (0.86)	3.82 (0.71)	2.84 (0.84)	3.57 (0.77)
Years of Schooling	4.82 (4.14)	8.29 (4.02)	2.31 (3.31)	6.43 (4.31)
Urban*	0.59 (0.49)	0.75 (0.43)	0.59 (0.49)	0.74 (0.44)
Rural*	0.41 (0.49)	0.25 (0.43)	0.41 (0.49)	0.26 (0.44)
Age	69.02 (7.52)	40.97 (9.66)	69.47 (7.84)	40.24 (9.59)
Married*	0.88 (0.32)	0.86 (0.35)	0.55 (0.50)	0.84 (0.37)
Single*	0.01 (0.10)	0.12 (0.33)	0.01 (0.10)	0.08 (0.27)
Widowed/Divorced*	0.11 (0.31)	0.02 (0.14)	0.44 (0.49)	0.08 (0.26)
Year 08*	0.28 (0.45)	0.24 (0.43)	0.27 (0.45)	0.25 (0.43)
Year 10*	0.27 (0.44)	0.24 (0.43)	0.27 (0.44)	0.26 (0.44)
Year 12*	0.45 (0.50)	0.51 (0.50)	0.45 (0.50)	0.49 (0.50)
Number of Observations	4507	16508	5547	19911

Source: 2008, 2010 and 2012 Turkish Health Survey

Notes : (1)*indicates a dummy variable

(2) The numbers in the paranthesis are standard deviations

From Table 6.3.2.1.1, we infer that males who belong to the untreated sample attend to school for 4.82 years on average, whereas the males who belong to the treated groups have 8.29 years of education on average. The descriptive statistics indicate that females have less average years of schooling compared to males in both treated and untreated samples. For females, the average years of schooling in the untreated sample is equal to 2.31, whereas it is equal to 6.43 in the untreated sample. Regarding the region of residence, 59 per cent of males who are not affected from the reforms live in urban areas, whereas the percentage of urban residence is 75 per cent for the males in the treated group. These percentages are very similar for females. For both males and females, we observe that the individuals in untreated sample are much older than the individuals in the untreated sample. Finally we observe that 88 per cent of males are married, 1 per cent of males are single and 11 per cent of males are widowed or divorced in the untreated sample. These percentages amount to 86 per cent, 12 per cent and 2 per cent for the treated sample respectively. In contrast, we observe that 55 per cent of females are married, 1 per cent of them are single and 44 per cent of them are widowed or divorced in the untreated sample. These numbers differ significantly from those for males, which indicates females live longer than males, hence they are more likely to be widowed. For the treated sample, we observe that 84 per cent of females are married, 8 percent of females are single and 8 per cent of them are widowed or divorced.

6.3.2.2. First Stage Results

Tables 6.3.2.2.1, 6.3.2.2.2, 6.3.2.2.3 present the first-stage regression results and Durbin endogeneity test statistics for three different specifications: In the first specification, we assume the individual who were born in or later than 1952 are affected from the reforms. In the second sample, we assume that the individuals who were born in 1951 or later are affected from the educational reforms and in the last sample we exclude 1951 birth cohort. In each specification, we make the estimations for three different samples. The first sample includes the whole sample, the individuals who were born between 1909 and 1987. We call this sample as Sample A. Although, Sample A gives idea about the relationship between the educational expansion in the early 1960s and the individual's years of schooling, the differences in the individual's reported SAH may not be purely explained by the differences by the individual's education level. For instance, older individuals tend to have worse health conditions or they may report better health statuses when they compare their health statuses with their contemporaries. Therefore, we establish two "Regression Discontinuity Design" samples in order to capture the effect of educational reforms in both the years of schooling and the reported SAH of the individuals. Sample B includes the birth cohort 1947 and 1957, that is the individuals who were born 5 years earlier than the individuals who are assumed to be affected from the educational reforms and it includes also the respondents who were born 5 years later than those individuals. Similarly, Sample C includes the individuals who were born between 1942 and 1962. Namely, it includes the individuals who were born 10 years earlier and later than the respondents who are assumed to be affected by the reforms. By restricting the birth cohorts in closer intervals, our objective is to see the exact impact of the increase in the years of schooling on individual's SAH.

Following Staiger and Stock (1997), we assume that if the F-statistics of excluded instrument is greater than 10, then the instrument is valid. As our F-statistics are greater than 10 for all the samples considered, the first stage statistics results indicate that the instrument is strongly valid. The results show that educational reforms that take place in the beginning of 1960s positively and significantly affect individuals' years of schooling. In addition, for endogeneity, Durbin test statistic indicates that SAH and years of schooling are endogenous. Therefore, it is necessary to implement the IV techniques in order to obtain consistent estimates[39].

[39] For all the estimation results in this section,Source: Authors' Computations from 2008, 2010 and 2012 Turkish Health Survey Notes: (1)*** indicates 1% level of significance, **indicates 5% level of significance, *indicates 10% level of significance.(2) Standard errors clustered by birth year are shown in parenthesis.

Table 6.3.2.2.1. First- Stage Estimation Results (The individuals who were born in and after 1952 are affected from the change in education policies)

Dependent Variable: Years of Schooling		
Sample A: Birth Cohort 1909-1987		
Policy Dummy	1.16*** (0.15)	1.04*** (0.16)
Observations	46473	46473
R-Squared	0.27	0.27
F-Statistics	38.39	34.57
Endogeneity Test-Statistics	7.04*** (p-value:0.007)	3.35* (p-value: 0.06)
Cohort Controls	No	Yes
Sample B: Birth Cohort 1947-1957		
Policy Dummy	1.15*** (0.15)	1.03*** (0.16)
Observations	7876	7876
R-Squared	0.19	0.19
F-Statistics	35.28	31.46
Endogeneity Test Statistics	5.43*** (0.01)	2.47* (0.10)
Cohort Controls	No	Yes
Sample C: Birth Cohort 1942-1962		
Policy Dummy	1.16*** (0.15)	1.03*** (0.16)
Observations	14946	14946
R-Squared	0.20	0.21
F-Statistics	36.95	33.06
Endogeneity Test Statistics	5.36** (0.02)	2.38* (0.10)
Cohort Controls	No	Yes

Table 6.3.2.2.2. First- Stage Estimation Results (The individuals who were born in and after 1951 are affected from the change in education policies)

Dependent Variable: Years of Schooling		
Sample A: Birth Cohort 1909-1987		
Policy Dummy	1.17*** (0.15)	1.01*** (0.16)
Observations	46473	46473
R-Squared	0.27	0.27
F-Statistics	37.00	28.23
Endogeneity Test-Statistics	14.67*** (p-value:0.000)	9.07*** (p-value: 0.002)
Cohort Controls	No	Yes
Sample B: Birth Cohort 1947-1957		
Policy Dummy	1.15*** (0.15)	1.06*** (0.16)
Observations	7876	7876
R-Squared	0.19	0.19
F-Statistics	30.77	38.64
Endogeneity Test Statistics	11.34*** (0.0008)	6.76*** (0.009)
Cohort Controls	No	Yes
Sample C: Birth Cohort 1942-1962		
Policy Dummy	1.15*** (0.15)	0.99*** (0.16)
Observations	14946	14946
R-Squared	0.21	0.21
F-Statistics	33.00	24.96
Endogeneity Test Statistics	11.36*** (0.0007)	6.81*** (0.009)
Cohort Controls	No	Yes

Table 6.3.2.2.3. First- Stage Estimation Results (Birth Cohort 1951 is excluded)

Dependent Variable: Years of Schooling		
Sample A: Birth Cohort 1909-1987		
Policy Dummy	1.29*** (0.18)	1.19*** (0.18)
Observations	45946	45946
R-Squared	0.27	0.27
F-Statistics	37.41	49.17
Endogeneity Test-Statistics	9.68*** (p-value:0.001)	6.34* (p-value: 0.002)
Cohort Controls	No	Yes
Sample B: Birth Cohort 1947-1957		
Policy Dummy	1.26*** (0.18)	1.17*** (0.18)
Observations	7349	7349
R-Squared	0.20	0.20
F-Statistics	29.99	38.09
Endogeneity Test Statistics	11.34*** (0.007)	4.51** (0.03)
Cohort Controls	No	Yes
Sample C: Birth Cohort 1942-1962		
Policy Dummy	1.27*** (0.18)	1.18*** (0.18)
Observations	14419	14419
R-Squared	0.21	0.21
F-Statistics	32.79	42.15
Endogeneity Test Statistics	7.25*** (0.007)	4.55*** (0.003)
Cohort Controls	No	Yes

First stage results show that the education reforms cause between 0.99 and 1.27 years increase in the individual's years of schooling. The increase in years of schooling is 1.16 years on average. We also look at the birth cohorts who were born between 1932 and 1972, but we do not report the results. For that sample, the result also indicates that the education reforms lead to 1.16 years increase in the individual's years of schooling. In terms of the LATE concept, the result indicates that people who are affected from the educational expansion in the early 1960s attend to schooling 1.16 years more than the case where they are not affected from the reforms. The first stage result for the education variable is a bit higher than the previous studies for developed countries. In the studies which are discussed in section 5.3.1 and chapter four generally find that increase in years of schooling after the reforms are between 0 and 1. Nevertheless, this is a study for developing country, Turkey. Glewwe (2002) mentions that, in general, developing countries expend more on education to increase the society's education level in order to promote growth and development. Thus, it is normal to expect that the effects of education policies in developing countries are more rapid. Hence, it is sensible to see the higher impact of education reforms on individual's years of schooling in developing countries compared to developed countries.

6.3.2.3. Second Stage Results

In the second stage regression, the dependent variable is the individual's SAH which varies from 1 to 5. Tables 6.3.2.3.1, 6.3.2.3.2 and 6.3.2.3.3 show the second stage results from CMP approach and the corresponding marginal effects for the education variable. In the second stage regression, we estimate the effect of rise in years of schooling on individual's SAH. In some specifications, we also include marital status dummies as explanatory variables. We do not report the coefficients of those dummies, but we discuss them at the end of this section[40].

Table 6.3.2.3.1. Estimates of SAH Returns to Schooling (The individuals who were born in 1952 or later are affected from the educational reforms):

Dependent Variable: SAH		
Sample A: Birth Cohort 1909-1987		
Years of Schooling	0.13*** (0.02)	0.12*** (0.04)
Male	0.13** (0.07)	0.18** (0.09)
Urban	-0.16** (0.07)	-0.12 (0.08)
Age Dummies	Yes	Yes
Cohort Controls	No	Yes
Marginal Effects for Years of Schooling Variable (x100)		
SAH=1	-0.29	-0.22
SAH=2	-1.86	-1.58
SAH=3	-3.01	-2.70
SAH=4	3.52	3.14
SAH=5	1.65	1.36
Sample B: Birth Cohort 1947-1957		
Years of Schooling	0.13*** (0.03)	0.11*** (0.03)
Male	0.12 (0.10)	0.18 (0.12)
Urban	-0.16** (0.08)	-0.12 (0.10)
Age Dummies	Yes	Yes

[40] For all the tables in this section, Source: Authors' Computations from 2008, 2010 and 2012 Turkish Health Survey

Notes: (1)*** indicates 1% level of significance, **indicates 5% level of significance, *indicates 10% level of significance. (2) Standard errors clustered by birth year are shown in parenthesis

Table 6.3.2.3.1. (Continued)

Cohort Controls	No	Yes
Marginal Effects for Years of Schooling Variable (x100)		
SAH=1	-0.67	-0.54
SAH=2	-2.48	-2.15
SAH=3	-1.99	-1.79
SAH=4	4.06	3.59
SAH=5	1.09	0.88
Sample C: Birth Cohort 1942-1962		
Years of Schooling	0.13*** (0.03)	0.11*** (0.03)
Male	0.15 (0.10)	0.21* (0.11)
Urban	-0.17** (0.08)	-0.13 (0.09)
Age Dummies	Yes	Yes
Cohort Controls	No	Yes
Marginal Effects for Years of Schooling Variable (x100)		
SAH=1	-0.61	-0.48
SAH=2	-2.46	-2.12
SAH=3	-2.18	-1.96
SAH=4	4.18	3.69
SAH=5	1.08	0.87

Table 6.3.2.3.2. Estimates of SAH Returns to Schooling (The individuals who were born in 1951 or later are affected from the educational reforms):

Dependent Variable: SAH		
Sample A: Birth Cohort 1909-1987		
Years of Schooling	0.16*** (0.03)	0.15*** (0.03)
Male	0.05 (0.09)	0.08 (0.11)
Urban	-0.23*** (0.08)	-0.21** (0.09)
Age Dummies	Yes	Yes
Cohort Controls	No	Yes
Marginal Effects for Years of Schooling Variable (x100)		
SAH=1	-0.47	-0.41
SAH=2	-2.38	-2.22
SAH=3	-3.42	-3.33
SAH=4	4.03	3.91
SAH=5	2.24	2.05
Sample B: Birth Cohort 1947-1957		
Years of Schooling	0.16*** (0.03)	0.14*** (0.04)
Male	0.03 (0.13)	0.07 (0.15)
Urban	-0.24*** (0.09)	-0.20** (0.11)
Age Dummies	Yes	Yes
Cohort Controls	No	Yes
Marginal Effects for Years of Schooling Variable (x100)		
SAH=1	-1.01	-0.83
SAH=2	-3.04	-2.77
SAH=3	-2.24	-2.14

Table 6.3.2.3.2. (Continued)

SAH=4	4.74	4.44
SAH=5	1.55	1.31
Sample C: Birth Cohort 1942-1962		
Years of Schooling	0.16*** (0.03)	0.15*** (0.04)
Male	0.06 (0.13)	0.08 (0.15)
Urban	-0.24*** (0.09)	-0.23** (0.11)
Age Dummies	Yes	Yes
Cohort Controls	No	Yes
Marginal Effects for Years of Schooling Variable (x100)		
SAH=1	-0.91	-0.80
SAH=2	-3.03	-2.87
SAH=3	-2.47	-2.41
SAH=4	4.88	4.71
SAH=5	1.53	1.37

Table 6.3.2.3.3. Estimates of SAH Returns to Schooling (Birth Cohort 1951 is excluded):

Dependent Variable: SAH		
Sample A: Birth Cohort 1909-1987		
Years of Schooling	0.14*** (0.03)	0.13*** (0.04)
Male	0.10 (0.10)	0.13 (0.11)
Urban	-0.20** (0.09)	-0.17* (0.10)
Age Dummies	Yes	Yes
Cohort Controls	No	Yes
Marginal Effects for Years of Schooling Variable (x100)		
SAH=1	-0.36	-0.30
SAH=2	-2.11	-1.91
SAH=3	-3.24	-3.08
SAH=4	3.79	3.58
SAH=5	1.94	1.71
Sample B: Birth Cohort 1947-1957		
Years of Schooling	0.14*** (0.04)	0.13*** (0.04)
Male	0.07 (0.15)	0.12 (0.17)
Urban	-0.21** (0.11)	-0.18 (0.12)
Age Dummies	Yes	Yes
Cohort Controls	No	Yes
Marginal Effects for Years of Schooling Variable (x100)		
SAH=1	-0.82	-0.70
SAH=2	-2.81	-2.59
SAH=3	-2.14	-2.20

Table 6.3.2.3.3. (Continued)

SAH=4	4.48	4.19
SAH=5	1.29	1.12
Sample C: Birth Cohort 1942-1962		
Years of Schooling	0.15*** (0.04)	0.13*** (0.04)
Male	0.10 (0.13)	0.14 (0.15)
Urban	-0.21** (0.10)	-0.19 (0.12)
Age Dummies	Yes	Yes
Cohort Controls	No	Yes
Marginal Effects for Years of Schooling Variable (x100)		
SAH=1	-0.73	-0.62
SAH=2	-2.77	-2.55
SAH=3	-2.36	-2.24
SAH=4	4.59	4.31
SAH=5	1.28	1.11

Second stage regression results indicate that increase in the years of schooling is positively and significantly related to individual’s SAH. Our results suggest that as the individual’s years of schooling increase, the individuals tend to report very good or good SAH more. Tables 6.3.2.3.1, 6.3.2.3.2 and 6.3.2.3.3 show that the coefficient of years of schooling variable from second stage CMP results is greater than the coefficient from ordered probit results that we present in Section 6.2. For instance, for the whole sample, while the marginal effects from the ordered probit results suggest that additional years of education leads to 0.56 percentage points increase in occurrence of reporting very good SAH, Table 6.3.2.3.1 shows that the occurrence of reporting very good SAH increases by 1.65 percentage points with additional years of schooling. Similarly, in the ordered probit regression analysis we observe that additional years of schooling causes 1.45 percentage points rise in probability of reporting good health, whereas it leads to 0.67 percentage points decrease in probability of reporting poor health, while for the second stage CMP estimation results from Table 6.3.2.3.1, these marginal effects are equal to 3.52 and -1.86 percentage points respectively. The apparent differences between the marginal effects from the two different estimation results provide us strong evidence that individual's years of schooling is endogenous to the SAH. It is also important to note that the standard error for years of schooling is greater in CMP estimation than in ordered probit estimation, which indicates that we lose efficiency due to IV estimation.

CMP estimation results suggest that in general, the reported SAH levels do not significantly differ between males and females, except for the whole sample where we assume that individuals who were born in or later than 1952 are affected from the education reforms. This result contradicts with the ordered probit results. Our results also suggest that SAH deteoriates with age. We also observe that people living in rural areas report better health conditions. When we include marital status dummies into the regression, our results suggest that married and single individuals report better health statuses compared to widowed or divorced individuals. This result is in line with ordered probability results.

6.4. Conclusion

In this chapter, we test the effect of years of schooling on individual's SAH. In section 6.2., we do not consider the possible endogeneity problem between the individual's years of schooling and SAH and we apply both ordered probit and ordered logit models in order to investigate the effect of higher levels of education on individual's SAH. The results indicate a positive association between health and education in Turkey as in the case of developed countries. In Section 6.2.3, we extend our original models by including labor market indicators and household income as the determinants that explain the variations in individual’s SAH. We observe that both employed and unemployed people tend to report better SAH levels than inactive individuals. We also find that household income is positively correlated with individual’s SAH.

The diagnostic tests suggest strong endogeneity between individual’s SAH and education level. Therefore, we implement IV techniques in order to obtain consistent estimates. Since the dependent variable (SAH) is a categorical variable, we establish an “IVOPROBIT” model and estimate it with Conditional Mixed Process (CMP) techniques, suggested by Roodman (2011). We use the educational reforms that take place in October 1960 and January 1961 in Turkey as the instrument for years of schooling. The first stage results indicate that our instrument is valid and it is positively correlated with individual’s years of schooling. We observe that educational reforms have more impact on males’ years of schooling than that of females. This result is valid for three different specifications and three different samples. Second stage regression results suggest that additional years of schooling positively affect the individual’s SAH. We observe that the reported SAH levels do not significantly differ between males and females. Our results also indicate people living in rural areas are more likely to report better SAH. Last, we point out that married and single individuals have better health statuses relative to widowed or divorced individuals.

In conclusion, we find a positive association between health and education in Turkey as in the case of developed countries. Therefore, we can conclude that educated people are better informed about the benefits of good health, and they are more aware of the importance of using the preventative health care services in Turkey. Concluding remarks and policy implications will be discussed in detail in Chapter Nine.

Appendix 6.A.1: 2SLS Estimation Results

In this section we present Two Stage Least Squares (2SLS) estimation results where the second stage regression is OLS. In this specification, we make a very strong assumption that SAH is defined in continuous intervals. However, it is important to note that the theory of 2SLS estimation methodology is acceptable worldwide, whereas CMP estimation techniques need to be improved more, as it is a new tool in microeconometrics. Therefore, replicating our results by implementing 2SLS estimation methods may be useful to support the validity of our results. Table 6A.1 provides the OLS and Second Stage IV Estimation Results

Table 6A1. OLS and Second Stage IV Estimation Results of SAH Returns to Schooling:

Dependent Variable: SAH	OLS (I)	2SLS (II)	2SLS (III)	2SLS (IV)	OLS (V)	2SLS (VI)
Years of Schooling	0.03*** (0.001)	0.08*** (0.03)	0.12*** (0.04)	0.10*** (0.04)	0.03*** (0.001)	0.08*** (0.04)
Male	0.21*** (0.01)	0.10 (0.06)	0.02 (0.09)	0.06 (0.08)	0.20*** (0.01)	0.10* (0.06)
Urban	0.01*** (0.006)	-0.10 (0.07)	-0.18** (0.09)	-0.14 (0.09)	0.02*** (0.0064)	-0.10 (0.07)
Married					0.05*** (0.01)	0.03** (0.01)
Single					0.07*** (0.01)	-0.04 (0.07)
Age Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Cohort Controls	Yes	Yes	Yes	Yes	Yes	Yes

Source: Authors' Computations from 2008, 2010 and 2012 Turkish Health Survey

Notes: (1) Column (1) shows the OLS results. Column (II) shows the 2SLS estimation results where we assume people who were born in or later than 1952 are affected from the reforms. Column (III) shows the 2SLS estimation results where we assume people who were born in or later than 1951 are affected from the reforms. Column (IV) shows the 2SLS estimation results where we exclude the 1951 birth cohort. Column (V) shows OLS results where we include marital status dummies and assume that people who were born in or later than 1952 are affected from the educational reforms.

Column (VI) shows 2SLS estimation results where we include marital status dummies and assume that people who were born in or later than 1952 are affected from the educational reforms.

(2)*** indicates 1% level of significance, **indicates 5% level of significance, *indicates 10% level of significance. Standard errors clustered by birth year are shown in parenthesis.

Both OLS and 2SLS estimation results suggest positive relationship between SAH and education. We observe that IV estimate of education coefficient is significant and it is higher than the OLS estimates. While OLS estimation suggests that additional years of schooling increases individual's SAH by 0.03 units, IV estimation suggests that additional education raises SAH level by 0.08 units if we assume the individuals who were born in or later than 1952 are affected from the education reforms. When we assume rather the people who were born in or later than 1951 are affected from education expansion, individual's SAH level increases by 0.12 units. Finally, we observe that increase in years of schooling leads to 0.10 units increase in the individual's SAH if we exclude the 1951 birth cohort.

Like in the CMP results, we observe that gender variable loses its significance when we take the endogeneity between the reported SAH and years of schooling into account. We observe that if we include marital status dummies into the regression, we find that males report better health statuses than females. Moreover, while OLS results suggest that people who live in urban areas have better SAH levels, IV estimation results indicate that the region that the individual lives does not significantly affect the individual's SAH level, except for the specification where we assume individuals who were born in or later than 1951 are affected from the education reforms. In that specification, we find that individuals who live in rural areas report better health statuses.

Next in columns (V) and (VI) we include marital status dummies. Education variable is still significant and IV estimate is higher than the OLS estimate. Regarding other covariates, first we observe that region variable loses its significance. Second, although OLS estimation results suggest that both married and single individuals have better SAH levels than widowed and divorced individuals, from 2SLS estimation results we observe that the SAH levels of single and widowed/divorced individuals do not significantly differ from each other.

Appendix 6.A.2. IV-Estimation Revisited

In this section, we re-estimate the effect of education on individual's SAH by including the trend controls, rather including the age dummies. The two stage regression model is written as:

$$E_i = \alpha + X_i\ \alpha_1 + R61_i\alpha_2 + trend\ controls + \varepsilon_i \qquad (6.A.2.1)$$

$$SAH_i = \delta + \hat{E}_i\beta_1 + X_i\beta_2 + trend\ controls + u_i \qquad (6.A.2.2)$$

The variables are defined in the similar way as in section 6.3.1. In this specification, the vector X includes the individual's demographic factors, gender, region, age, survey year trend controls and birth year trend controls. All the standard errors are clustered by birth year. Since, it is probable that our instrument is correlated with cohort-level changes that affect educational outcomes, we also control for birth year trends in the empirical analysis. Since, our empirical analysis covers a long time period, we also include larger power of trend controls (square and cubic powers) to our model. In this specification, we find that the educational expansion leads to 0.82 years increase in the individual's years of schooling and F-statistics suggest that our instrument is strongly valid. However, from second stage results we observe that the effect of education is negative and insignificant on individual's SAH level. Table 6.A.2 below presents the first stage and second stage regression results from this specification. For this analysis, we assume that individuals who were born in or later than 1952 are affected from the educational expansion in the early 1960s.

Table 6.A.2. 2SLS Estimation Results[41]

First-Stage Regression Results	
Dependent Variable: Years of Schooling	
Policy Dummy	0.82*** (0.09)
Male	2.04*** (0.03)
Age ($x10^{-1}$)	2.41*** (0.91)
Age Squared ($x10^{-3}$)	-1.48*** (0.34)
Urban	2.22*** (0.03)
Trend	0.04 (0.09)
Trend Squared	0.008 (0.007)
Trend Cube	0.000007 (0.000004)
Survey Year Trend Controls	Yes
F-statistics	26.05
R^2	0.16
N	46473
Second Stage Regression Results	
Dependent Variable: SAH	
Years of Schooling	-0.01 (0.02)
Male	0.32*** (0.05)
Age ($x10^{-1}$)	0.45*** (0.10)
Age Squared ($x10^{-3}$)	-0.23 (0.16)
Urban	0.13*** (0.05)
Trend	0.01 (0.01)
Trend Squared	0.0005 (0.0003)
Trend Cube	-0.000002 (0.000001)

[41] Note:*** indicates 1% level of significance, **indicates 5% level of significance, *indicates 10% level of significance. Standard errors clustered by birth year are shown in parenthesis.

CHAPTER 7

HEALTH BEHAVIORS AND EDUCATION

7.1. Introduction

In Chapter Seven we focus on the relationship between health behaviors and education for the individuals who are 25 years old or older in Turkey. The health behaviors considered in this chapter are as follows: Smoking, alcohol consumption, fruit and vegetable consumption, exercising and Body-Mass Index (BMI). In addition, this chapter examines the variation in health related behaviors due to other determinants such as demographic factors, labor market status and household income. It is important to analyze the variation in health behaviors because differences in health outcomes (such as life expectancy, mortality) are mainly related to differences in health behaviors. For instance, Mokdad et al. (2004) estimate that almost half of the adult mortality in the US is ascribed to risky health behaviors. Cutler and Lleras-Muney (2010) emphasize that health outcome differences by education need to be explained by health behavior differences by education. The differentials in health behaviors by education are studied mostly in developed countries. However, there is less evidence on this issue in developing countries. Chapter Seven investigates the health behaviors by education in Turkey which is a middle income and developing country.

This chapter is the first in the sense that investigates the relationship between various health behaviors and education for elder individuals in Turkey in one study. Previous literature on the determinants of health behaviors in Turkey is very limited and mostly focuses on one health behavior at a time. For instance, Tansel (1993) examines the tobacco consumption in Turkey. Erem et al. (2004) investigate the determinants of obesity in Trabzon, a city located in the Black Sea Region of Turkey. Hatemi et al. (2003) study the relationship between hypertension and obesity for 11 different cities located in four different regions of Turkey. Other studies related to the determinants of health behaviors in Turkey include Metintaş et al. (1998), Erbaydar et al. (2005), Kocabaş et al. (1994) and Yumuk (2005).

Among the health behaviors we analyze, we refer smoking and high levels of BMI as "Risky health behaviors". Risky health behaviors negatively affect the individual's general health status. For instance, as the frequency of risky health behaviors increase, people are more likely to report poor self-assessed health (SAH) (Brunello et al., 2015). Risky health behaviors also lead to serious diseases such as cardiovascular diseases, cancer, diabetes, etc… Some of the studies related to the harmful effects of risky health behaviors are Cawley and Ruhm (2011), Hung et al. (2004) and Stewart et al. (2009) among others. We do not consider alcohol consumption as risky health behavior in this chapter, since both the heavy drinkers and social drinkers are referred as alcohol drinkers in our study[42].

In addition, this chapter is important in the sense that it is one of the recent study that investigates the relationship between different health behaviors and education in Turkey by considering endogeneity problems between health and education. The only previous study which considers the endogeneity problem between health behaviors and education in Turkey is conducted by Cesur et al. (2014). However, in this study the authors examine the impact of education on different health behaviors for the individuals who are between 18 and 30 years old. Nevertheless, in this chapter, we focus on older individuals (25 years old or above) in order to see the

[42] We have discussed in how we define the alcohol consumption behavior in chapter three

exact effect of education on health behaviors, as the individuals complete their schooling approximately at the age of 25 in Turkey. Like in Chapter Six, we use the education reforms that took place in Turkey in the early 1960s as the instrument for the individual's years of schooling.

In general, our results suggest that education has the strongest effect on all of the health behaviors considered. Probit and OLS results suggest that university graduates tend to smoke less, consume more fruits and vegetables, and exercise more frequently than less educated individuals. In addition, they have lower BMI levels compared to the less educated ones. Alcohol consumption is an exception to these evidences. Our results indicate that highly educated people tend to consume more alcohol than less educated people. However, when we take the endogeneity between health and education into account, we find that alcohol consumption also decreases with higher levels of education.

The rest of this chapter is organized as follows. First, in section 7.2 we examine the literature on the relationship between health behaviors and education both for developed countries and Turkey. Then, in section 7.3 we present descriptive statistics for each health behavior. Next, section 7.4 presents the probit (OLS for BMI) results. Next, in section 7.5. we revisit the endogeneity problem between health and education, and use instrumental variable (IV) approach to solve the endogeneity problem between health behaviors and individual's years of schooling. Finally, section 7.6 presents the concluding remarks.

7.2. Literature Review on Health Behaviors

Examining the variation in health behaviors is important because differences in health behaviors explain the differences in important health outcomes, as Mokdad et al. (2004) suggest. In that paper the authors examine the leading causes of mortality in the USA in 2000. According to their study, the percentages of leading causes of mortality in the USA are as follows: Tobacco consumption: 18.1 per cent; poor diet and physical activity: 16.6 per cent and alcohol consumption: 3.5 per cent. The remaining causes of mortality are microbial agents, toxic agents, motor vehicle crashes, incidents involving firearms, sexual behaviors, illicit use of drugs. This constitutes the 11.8 per cent of all deaths. Hence, they conclude that 50 per cent of deaths in the USA stem from risky health behaviors.

In their study, Mokdad and his friends use the MEDLINE search of English language articles that identified epidemiological, clinical, and laboratory studies and link risky health behaviors to mortality. They implement appropriate formula to calculate assignable fractions for each disease by considering relative risk, percentage of individuals in USA who are included or not included in engaging the certain risky health behavior[43]. They conclude that smoking is the major cause of mortality. However, they stress that poor diet and physical inactivity will soon overtake smoking as the leading cause of death. Henceforth, maybe because of being the major causes of death, smoking and obesity are the most popular health behaviors that are examined in the literature.

Cutler and Lleras-Muney (2010) is one of the important studies which emphasize that health outcome differences by education need to be explained by health behavior differences by education. Cutler and Lleras-Muney study this relationship in the USA and the UK. The authors conduct their empirical analysis by using a rich combination of data sets from both the USA and the UK. The data sets they use are as follows: National Health Interview Survey (NHIS), National Longitudinal Survey of Youth (NLSY), the National Survey of Midlife Development in the United States (MIDUS), the Health and Retirement Study (HRS), the Survey on Smoking (SOS)

[43] They find the relevant data from the following sources: Published causes of death reported to the Centers for Disease Control and Prevention (CDC) for 2000, relative risks (RRs), and prevalence estimates from published literature and governmental reports .

and the National Childhood Development Study (NCDS) in the UK. In the study, they focus on the individuals who are assumed to complete their formal schooling, namely the individuals who are 25 years old or older. In the empirical analysis, for the USA, the education variable is the individual's completed years of schooling and for the UK, the education variable is a dummy variable which indicates whether the individual passed any A level examinations in the UK. The authors use Linear Probability Methods (LPM) throughout the analysis. They note that the results of linear and non- linear probability models do not significantly different from each other. Three important contributions of this study are as follows: First, they compute the average reduction in the explanatory power of the education variable when new explanatory variables are added into regression. Second, they weight smoking, heavy drinking and obesity by their impact on mortality and finally, they compute the average effect of education on different health outcomes. The health behaviors considered in their study are, smoking, diet/exercise, alcohol use, illegal drugs, automobile safety, household safety, preventative care, and care for people with chronic diseases (diabetes or hypertension).

The authors observe that the magnitude of the years of schooling variable reduces slightly as new control variables are included in the regression but its significance never changes. Their results suggest a negative relationship between smoking and the individual's years of schooling. In addition, they find a negative association between probability of being heavy alcohol drinker and education. Moreover, they conclude that increase in years of schooling results with normal ranges of BMI.

According to Cutler and Llears-Muney, income, health insurance and family background explain 30 per cent of the validity of the education gradient of health. The regression results applied to several data sets suggest that family income explains some of the education effect. For instance, the coefficients on smoking, consuming alcohol, BMI decreases by considerable amounts when family income is taken into account in the model. They find out other family and individual variables such as family size, employment status and marital status have also impact on education gradient of health. However, these effects are not as large as the effects of family income.

The authors state that another 30 per cent of education gradient of health is explained by knowledge and measures of cognitive ability. The authors argue that cognitive ability increases with education, therefore, health behavior also improves with cognitive ability. Next, they point out that 10 per cent of the education gradient is explained by social networks. Finally, the authors conclude that the proxies such as discounting, risk aversion, or the value of the future do not significantly explain the education gradient of health.

Later on, Conti and Hansman (2012) test the sensitiveness of Cutler and Lleras-Muney's results related to cognitive ability. In the study, they use National Child Development Study (NCDS) data set in which both cognitive ability measures and the measures of non-cognitive skills acquired in childhood. Cutler and Lleras-Muney (2010) have used indicators of self-efficacy (whether the respondent gets what they want out of life, how much control they have over their lives, and whether they can live their lives how they want) and measures of mental health and stress. In their study, Conti and Hansman consider alternative measures for personality and replicate the methods of Cutler and Lleras-Muney. The additional measurements that they use for personality are the Rutter Behavior Scale and syndrome scores from the British Social Adjustment Guide. The Rutter scale measures behavior of difficulties in child, which is obtained by asking the parents for the children at ages 7 and 11, and asking both parents and the teacher for children at age 16. The British Social Adjustment Guides (BSAG) describe the child's behavior in particular settings, and are obtained for children at ages 7 and 11.

The replication results show that cognitive ability and social integration explain the large portion of the education-health gradient whereas the current and future satisfaction and personality have low impact on education-health gradient. However, when the additional measures of depression are used in the analysis, the authors find that personality measures play a significant role as high as the cognitive skills. Among the personality measures, depression and anxiety have the most important impact on education-health gradient. Their findings also suggest that the personality of the individual explains his/her health-related behaviors such as smoking and alcohol consumption. Therefore, Conti and Hansman underline the importance of personality to explain the differences in health behaviors.

Conti et al. (2010) conduct another study on the education gradient of health. In this study, the authors' objective is to strengthen the link between educational attainment and health outcomes, especially by focusing on the relationship between early childhood conditions and adult outcomes by using British Cohort Study (BCS70) data set, in which all babies born in one week in April 1970. In the paper, the authors examine the relationship between health outcomes and cognitive ability as well as the personality traits.

Conti et al. use semi-parametric structural model to test the relationship between choice of schooling and poor health, smoking and obesity. The authors have both the measures of early childhood environmental factors and proxies for a vector of latent abilities for early life cognition, personality and health endowments. In the study schooling is taken as a binary variable which shows whether the individual continues on his/her education after compulsory schooling is completed. The authors have concluded that early cognitive ability has a strong and statistically significant effect for health outcomes when non-cognitive traits are not added into the model. However, if non-cognitive traits are added into the model the estimated effect of cognitive traits become smaller. Hence, the authors suggest that non-cognitive traits should not be ignored while estimating the relationship between education and health. Otherwise, the effect of cognitive abilities on the child's health will be overestimated.

Mocan and Altindag (2013) also test the effect of cognitive ability on education gradient of health by using National Longitudinal Survey of Youth (NLSY) data set for 1997. This is an important study in the sense that it uses a panel data set in order to test the impact of education on health related behaviors. They conduct similar analysis as Cutler and Lleras-Muney (2010) and the health outcomes they use are similar to those of Cutler and Lleras-Muney. Their results suggest that cognitive ability is not a significant factor that explains the education-health gradient.

A famous study that examines the relationship between different health behaviors and education is conducted by Kenkel (1991). He states that examining the variation in health behaviors is important because risky health behaviors increase the probability of mortality. In other words, he refers the health behaviors as the inputs to the health production function, where the health outcome is mortality. In that study, the author investigates the relationship between years of schooling and the consumption of cigarettes, alcohol and exercise by using Health Promotion/Disease Prevention (HPDP) supplement to the 1985 Health Interview Survey. Using this data set is advantageous since it measures the respondents' knowledge about the health consequences of the risky health behaviors. Apart from the usual explanations for the positive relationship between health and education, Kenkel believes that additional years of education benefits people in the sense that it allows them to choose healthier life-styles because education improves their knowledge of the associations between health behaviors and health outcomes.

For each health behavior, Kenkel conducts separate tobit regression analyses for empirical specification. The dependent variables he uses in different analysis are as follows: The number of cigarettes the respondent typically smokes a day, total number of alcoholic drinks in the past 2 weeks, the number of days in the past year on which the respondent had five or more drinks, and the number of minutes in the past 2 weeks the individual engaged in exercises, sports or physically active hobbies. The independent variables used in the analysis are age, family income, race, marital status, employment status, veteran status, self-reported stress levels, health knowledge and years of schooling. As indicator of health knowledge for cigarette and alcohol consumption, Kenkel uses the number of illnesses which are supposed to be related to cigarette smoking and heavy drinking by the respondent. The results point out higher probability of light drinking among higher socioeconomic groups, positive relationship between schooling and exercise and negative relationship between smoking and the years of schooling. Kenkel's results also suggest that part of the relationship between schooling and the consumption of cigarettes, alcohol, and exercise is explained by differences in health knowledge. He also stresses that the effect of years of schooling on different health behaviors remains to be significant after differences in knowledge are controlled for.

Johnston et al. (2015) test whether the higher education and health knowledge are related to each other for UK. In the analysis they use Health and Life Style Surveys (HALS) for 1984-1985 and 1991-1992 periods. In the HALS data set, individual's health knowledge is measured by the answers of questions that ask the respondents about the main causes of 10 major health problems. To overcome the problems of endogeneity between health knowledge and education, the authors use the increase in UK minimum school leaving age in 1947 and 1972 and implement 2SLS estimation procedure as well as OLS. In their study, while the OLS estimates suggest that health knowledge and education are positively correlated with each other, 2SLS regression results do not indicate a significant association between the two.

Cutler and Gleaser (2005) also support that the people who have higher education levels and who have higher incomes believe that risky health behaviors are harmful for their health and therefore the prevalence of risky health behaviors is low among those groups. They conduct the empirical analysis by using NHIS for 1990. The health related behaviors considered in this study are being current smoker, heavy drinker, obese, using a hypertension medication and for women whether having mammogram in the past 3 years. In the study, the authors focus on the individuals who are 45 years old or above. They estimate the effect of education and other socio-demographic factors on these health behaviors by establishing linear probability models (LPM) for each health behavior. The authors also test the effects of behavior specific and genetic factors on health related behaviors. In order to test the effects of the behavior specific situation on the individual's health behaviors, the authors use data from Health and Retirement Survey, and to examine the effects of genetics on individual's health behaviors they use Minnosata Twin Registry data set. By establishing a random effects model, the authors calculate the correlations between these two factors and the prevalence of different health behaviors. Cutler and Gleaser conclude that both behavior specific situations and genetics are important to explain the variations in individual's health behaviors.

Kemptner et al. (2011) conduct an important study on the validity of education gradient of health by using West-Germany microcensus data set for the years 1989, 1995, 1999, 2002, and 2003. In that study, the authors examine the effect of years of schooling on health and health related behaviors, such as smoking and obesity. They measure the current health by a dummy variable which indicates whether the individual has a serious long-term illness and whether this illness causes limitation in individual's activities. Another measure of current health in the study is the Body-Mass Index (BMI) of the individual. The authors use the BMI to create two dummy variables: The first dummy shows whether the individual is overweight or not (i.e BMI>25), and the second dummy shows whether the individual is obese or not (i.e BMI>30). The authors measure health behaviors first with the answer of the question that asks whether the individual currently smokes. If the respondent states that he/she does not smoke, there is another question that asks whether he/she has ever smoked. They estimate the relationship with both OLS and IV methods. Due to the possible endogeneity problem between education and health, the authors are in favor of the estimation results by IV methodology. The instrument for the years of education in that study is the changes in compulsory schooling laws. In the study, they estimate the IV model controlling for state-specific linear trends. Because, they claim that state-specific trends (such as a boom in the per-capita growth) may be correlated with both the change in compulsory schooling laws (such as state's resources which are available for reforms of the schooling system) and unobserved health determinants of individual's health (such as income of households). From the estimation results, the authors conclude that a rise in the years of schooling significantly reduces the probability of having long-term illness for men but they do not find a significant relationship for women. In addition, they find that increase in the years of schooling decreases the likelihood of being obese for both men and women. However, they do not find a significant link between education and smoking behavior.

Jürges et al. (2011) conduct another study for West Germany which examines the relationship between health behaviors and education. In the study the health outcomes are smoking and obesity. The authors use German Microcensus data set for 1999 and 2003. The authors also establish both OLS ans 2SLS models to test the impact of education on smoking and obesity. In the study, the instrument for education is the number of academic track schools, which is called as “Gymnasium” in Germany, in Western Germany. For German pupils, attending to gymnasium means to attend high school for three additional years, therefore the authors prefer believe that enrollment in gymnasium provides an exogenous variation in the individual’s years of schooling. Both the OLS and 2SLS estimation results indicate that higher levels of education lead to lower smoking rates for both men and women. However, the authors do not find a significant relationship between education and being overweight and obesity.

Arendt (2005) investigates the relationship between education and BMI and smoking for Denmark. He uses the Danish National Work Environment Cohort Study (WECS) data set of Danish Workers who has been interviewed in 1990 and 1995. The data consists of individuals who are above 25 in 1990. He uses both a quantal response panel model and a linear regression model in order to test the effects of education on smoking and BMI. In order to overcome the possible endogeneity problem, he uses two school reforms as instruments for years of schooling. In Denmark two school reforms were established in 1958 and 1975 respectively. They form two dummy instrumental variables to address whether the individual is affected by these two reforms. His results provide evidence that years of education and normal ranges of BMI have negative relationship. The results also indicate that additional years of education reduces the probability of being smoker and increases the likelihood of being a non-smoker.

There are some studies which do not indicate a positive relationship between health behaviors and education by using IV approach. Braakman (2011) is one of the examples of these studies. In this study, he examines the effect of education on health behaviors for January and February born individuals for 13 academic cohorts in England by using British Labor Force Survey and Health Survey data set. Braakman also implements IV methodology in order to detect the direct causality between health and education. The instrument is a dummy variable which is equal to 1 if the individual was born in February, 0 otherwise. The author argues that the instrument is valid and does not suffer from the weak instrument problem as in Angrist and Krueger (1991), who also use quarter of birth as instrument for education. He argues that quarter of birth may be invalid instrument since the families may plan the month that they want to give birth to their children, however, January and February belongs to same quarter. In addition, differences in maternal nutrition, weather conditions and sunlight exposure may explain the in mental and physical health of children. These factors are more or less equal for the children born in January or February. In fact, empirical results support his claim that, the instrument is valid since it passes the weak instrument test as the F-statistics are greater than 10. Braakman observes that the individuals who were born in February have 2-3 per cent higher probability of obtaining qualification. However, he cannot find a significant relationship between education, health outcomes and health related behaviors (smoking, drinking or eating various types of food).

The effects of other factors such as income and region together with education on health behaviors are also examined in the literature. In a famous study, Lantz et al. (1998) look for possible reasons for the higher mortality rates in lower socioeconomic groups. They investigate whether this is because of the higher prevalence of risky health behaviors among lower social economic groups. The risky health behaviors considered in that study are smoking, alcohol consumption, sedentary lifestyle, relative body weight. The authors use Americans' Changing Lives Survey data set, which allows them to calculate the impacts of education, income and health behaviors on the risk of dying within next 7.5 years. Like Mokdad et al. (2004) the authors also conduct The Cox Proportional Hazard Model in order to estimate the relative risk of mortality in terms of various background, socioeconomic and health behavior variables. The results suggest that people with lower education levels and incomes tend to have risky health behaviors more than the others. Next, they find that the people with low education levels tend to die more than the ones with higher education levels. The authors find out that the health behaviors considered are accounted for 12 per cent to 13 per cent predictive effect of income on mortality. Finally, they conclude that improving health behaviors of lower socioeconomic groups does not reduce the higher mortality rates among this group.

Kling et al. (2007) conduct an experimental analysis in order to test the effects of living in more favored places on different health behaviors of individuals. They believe that moving to a favorable area may increase the educational attainment of the youths since education opportunities are higher in favorable regions. In the experiment called as the "Opportunity Program", families offered vouchers moved to safer neighborhoods, whereas the families who are not offered vouchers continue to live in high poverty public housing projects. The families which are offered vouchers are determined by random assignment, and the authors observe the behaviors of individuals living in those families from 4 to 7 years.

Kling et al. find no significant impact of moving to more favorable area on adult economic self-sufficiency or physical health. However, they find this enhances the mental health, especially for female teenagers. Moreover, they find that after moving to more favorable place, education of female teenagers improves, prevalence of risky health behavior among female youth decreases and physical health of female youth also improves. However, their findings suggest that male youths are more likely to show risky health behaviors (such as smoking or using illicit drugs) when they move to more favorable region. The results also suggest that prevalence of obesity significantly decreases among the adults who move to favorable areas.

Among all the health behaviors considered, smoking is the most popular health behavior that is examined in the health economics literature most probably due to the fact that it leads to serious health problems. First, smoking is leading cause of lung-cancer. Chaloupka and Warner (2000) suggest that 90 per cent of lung-cancer deaths is seen due to heavy cigarette consumption. Smoking is also leading cause for chronic bronchitis and emphysema as well as heart disease and stroke. Laryngeal cancer, oral cavity, aneurysms, atherosclerotic peripheral vascular disease, neonatal death, bladder, renal, gastric, pancreatic, cervical cancers, vision and hearing problems, quiet recovering from injuries are other causes of heavy smoking (Chaloupka and Warner, 2000).

De Walque (2007) examines the effect of education on smoking by using IV methodology. The data sets they use come from nine smoking supplements from different years of the National Health Interview Survey between 1983 and 1995. They focus on the individuals who are 25 or above in their study. The author uses Vietnam War Draft as the instrument for education. Both IV and OLS estimation results suggest that higher levels of education reduce the probability of smoking. The negative effect remains when gender, race, age, birth cohort, date of interview and income are also controlled. Other studies that consider the endogeneity between smoking and education include Grimard and Parent (2007) and Sander (1995). Both of these studies use Vietnam War draft as the instrument for education.

The second important health behavior (or problem) examined in the literature is obesity. The amount of studies on obesity is higher than the amount of studies on smoking in recent decades as prevalence of smoking decreases over years, but at the same time obesity becomes an important health problem, not only in the USA, but also in other developed countries. Following the increase in the obesity rates, and the corresponding decrease in smoking rates in the USA in the past 30 years, Stewart et al. (2009) forecast the life expectancy and quality adjusted life expectancy for an 18 years old individual. The authors use National Health Interview Survey (NHIS) to observe the trends in smoking, National Health and Nutrition Examination Survey for the trends in BMI levels and finally they use 2003 Medical Expenditure Panel Survey in order to examine impacts of smoking and obesity on individual's quality of life. In this study the authors' objective is to learn whether the decline in the prevalence of smoking overwhelms the negative effects of obesity. Their findings suggest that the benefits of the decrease in smoking rates, unfortunately, cannot beat the negative impacts of obesity. Their results are valid for multiple scenarios and different projections.

Chou et al. (2002) examine the possible reasons for rapidly increasing obesity rates in the USA by using 1984-1999 Behavioral Risk Factor Surveillance System data set. They consider the adult obesity and focus on the individuals who are 25 years old and older. They attribute the increase in obesity rates to 3 major economic changes. First, as a result of the increase in the value of time, especially of women, female labor force participation rates and females' hours of work also increase. Therefore, females less spend time at home in order to prepare home-made foods. Therefore, the demand for convenience food increases. Second, since cigarette prices increase, people give up smoking, and reduction in smoking results with gaining body weight. Last, the number of fast food restaurants have increased over time. Since the cost of eating in fast food restaurants are less than the cost of eating in another restaurants which provides healthy food and since it is also less than the cost of making food at home, people prefer to eat in fast food restaurants more. This results with huge caloric intake.

Chou et al. use OLS estimation techniques in order to test these three hypotheses. In their empirical analysis, the determinants of the individual's BMI levels are per capita fast-food restaurants, per capita number of full service restaurants, the price of a meal in each type of a restaurant, the price of food consumed at home, the price of cigarettes, clean indoor airlaws, and hours of per week and hourly wage rates. These variables are all in state levels. In addition, age, gender, years of schooling completed, race and marital status are also controlled. The results indicate that males have higher BMI males, but females are more likely to be obese. Married and widowed individuals have higher BMI levels relative to single and divorced individuals Years of formal schooling completed and real household income have negative effects on BMI and the probability of being obese, which supports the preventive effects of household income and education on obesity. In contrast to their expectations, the authors find no significant relationship between smoking and probability of being obese as their results do not support a significant effects of clean indoor air laws, cigarette prices and private workplace restrictions on weight of the individuals.

Later on, Cutler et al. (2003) also examine the possible reasons for rising in the obesity rates in the USA. The authors attribute the increase in obesity rates to technological improvement. Because, they think that technological enhancement lead to both an increase in the variation and frequency of food consumption and it gives rise to switching to high-calorie/high-flavor prepared foods that are previously unavailable. With the improvement in technology, the tools that induce the prevalence of convenience food, such as vacuum packing, improved preservatives, deep freezing, artificial flavors and microwaves become more available then before, hence home-made cooking decreases as the food manufacturers are able to cook centrally and ship it to the households. This is easier than cooking at home especially for the spouses who are both working. As a result of the decrease in cost of food production consumption of convenience food increases, therefore individuals start to get more calories. By using National Health and Nutrition Examination Surveys (NHANES) data sets for the year 1994, the authors conclude that increase in the consumption of mass food lead to increase in obesity rates in both USA and most of the OECD countries, such as UK, Australia and Canada. The authors note that the obesity rates in Germany and France are relatively lower because these countries support the traditional food consumption rather than fast food consumption. The results of the study also indicate that caloric intake increases among the individuals over time but caloric expenditure does not, which also gives rise to prevalence of obesity. Last, unlike Chou et al. (2002), the authors do not support the idea that increased in labor force participation of women leads to rise in the obesity rates.

Phillopson and Posner (2003) also examine the reasons for huge growth in obesity by establishing a theoretical model which links to increase in prevalence of obesity to technological improvement. Apart from the the rise in calorie intake and reduction in exercise, the authors attribute the increasing trend in obesity rates to the fact that individuals can be addictive to eating after some time.

Next, Rashad and Grossman (2004) also investigate the possible reasons for obesity. According to them, the reasons are as follows: First, prices of groceries fall and consumers demand more groceries. As a result of this caloric intake increase. Second, microwavable meals and other foods that are easy to cook are desirable because they are quicker to prepare; they are also fattier and higher in caloric content. This is exactly what Cutler et al. (2003) suggest. Third, decline in physical activity and urban sprawl also leads to increase in obesity. Fourth, eating out at fast food restaurants and full service restaurants also raise the prevalence of obesity and they stress that this is due to the increase in female labor force participation as females prefer to eat out or prepare convenience food as they work more than before. This is consistent with Chou et al.(2002)'s ideas. Finally, the authors refer smoking as a tool to control weight, therefore high taxes on smoking lead to decrease in smoking and consequently increase in obesity rates. The authors stress that there are several consequences of obesity such as hypertension, high cholesterol, coronary heart disease, type 2 diabetes, depression, and various types of cancer. In order to prevent the negative costs of obesity, the authors suggest some policies such as food taxes, rewarding exercise i.e. subsidizing facilities and programs for exercise.

Phillopson and Posner (2008) state that obesity is not only a public health problem but also an economic problem. First, they argue that the increase in obesity rates cannot be just explained with biological factors and genes, obesity also depends on the choice of people. Second, they believe that obesity may create both social costs and private costs, and therefore the government may imply policies to reduce the obesity. Hence, the economists should advise cost reducing methods to policy makers. Different from Chou et al. (2002) and Cutler et al. (2003), Phillopson and Posner support that obesity is rising over time as a result of switching from agricultural economies to the manufacturing sector. They argue that when economies are based on agriculture, people make exercise as a part of their job. However, after the economies become more industrialized, exercise is not an element of the individual's job anymore. Exercise is a component of leisure time now, i.e. it is included in gym and jogging activities. In addition, due to technological improvement, food becomes cheaper whereas exercise becomes more expensive. The authors also note that the wide use of computers and television give rise to the childhood obesity. Like Chou et al. (2002) and Cutler et al. (2003), Phillopson and Posner also argue that reduction in prices, lead to increase in consumption of fast food. The authors also consider the relationship between obesity and cigarette and alcohol consumption. They claim that lower relative prices of alcohol and higher relative prices of smoking lead to increase in obesity rates. Different from the previous studies, Phillopson and Posner discuss that the social aspects of obesity may play an important role in decreasing the obesity rates. They believe that if prevalence of obesity is rare in society, the individuals consider obese ones as abnormal. This negative perception can help to keep obesity under control. However, if the prevalence of obesity raises, the abnormal image of obesity disappears as it becomes more common, and this may cause increase in obesity rates.

Next, Phillopson and Posner focuses on the relationship between income and obesity. They maintain that the individuals who live in countries with more improvement but sedentary technologies are more likely to be employed. This causes a rise in both income and weight of these individuals living in those countries. The authors test the relationship between income and obesity by assuming health is a normal good, which provides non-monotonic association between income and weight. Hence, the authors conclude that for poor and underweight people, increase in income level leads to more food consumption and this results with weight increase. Similarly, for rich and overweight people, rise in income might lead to weight loss since these individuals have more opportunities invest resources, such as gym, in order to reach their ideal weight. Therefore, increase in income leads to increase in weight among poorest groups and it leads to decrease in weight throughout the upper half of the income distribution.

Grabner (2009) also focuses proposes three major policies to prevent the rise in obesity rates. First, he stresses the importance of education as a factor that can reduce the prevalence of obesity. The tools of education can be nutrition labeling, advertising on the health costs of obesity, programs that teach correct diet and exercise, and general education in the form of additional years of schooling. Second, he recommends that taxation and fast-food regulation can be important tools to decrease the prevalence of obesity. Last, he promotes the campaigns against obesity. Among the tools he proposes, by assuming education is the most relevant, in his study, Grabner examines the relationship between obesity and education. He also implements instrumental variable (IV) techniques in order to test the relationship between individual's BMI and years of schooling consistently. In his analysis, the author uses the variation due to state specific compulsory schooling laws between 1914 and 1978 as instrument for education. He conducts his empirical analysis by using National Health and Nutrition Examination Survey (NHANES) conducted from 1971 to 1975 and 1976 to 1980. His results suggest a negative significant relationship between additional schooling and BMI. His results indicate that the effect is even stronger for females.

Brunello et al. (2013) look at the relationship between BMI and years of schooling for both males and females in 13 different European countries. In the empirical analysis, the authors look at the effect of education on both BMI and the propensity of being overweight and obese. They take the possible endogeneity problem between education and BMI, hence they also implement IV techniques in order to test the relationship between years of schooling and BMI. They use the change in compulsory schooling laws in several European countries during the 1960s and 1970s as the instrument for education and hence they select the countries in the following manner: First, the effect of the change in compulsory schooling laws and the effect of years of schooling on BMI should not differ significantly across countries. Second, changes in compulsory schooling laws should have positive and significant effect on individual's years of schooling in all the countries. The authors establish a sample design which they call as "Pooled Regression Discontinuity Design". In that design the authors form pre-treatment and post-treatment samples for each country. They choose a cohort, C1, which are assumed to be affected by the educational reforms. Then in the next step, they calculate the distance between the year of birth cohort C and C1. They include the individuals, who were born at most 7 years before the pivotal cohort, as the pre-treatment group and the individuals who were born at most 7 years after the pivotal cohort as the post-treatment group. They determine this width (7 years) so that the effects of other reforms can be omitted. Among the countries that they investigate, the impact of years of schooling is found to be significant for females, whereas it is insignificant for males in 9 countries[44]. The effect for females is smaller than the same effect for USA. Their results suggest that the negative effect of years of schooling on individual's BMI is even stronger for overweight females.

Webbink et al. (2010) also examines the relationship between education and the probability of being overweight by using a longitudinal data of Australian identical twins. They use within twin estimation methodology to overcome possible biases due to unobservable factors. The methodology they use make sense to prevent the measurement errors since identical twins share the same genes and socioeconomic background. The authors conclude that education has a protective effect for obesity

[44] These countries are Austria, Denmark, Germany, Greece, Italy, Portugal, Spain, Sweden and United Kingdom.

from men, however, they do not find significant relationship between education and obesity for women. They attribute this result to different peer effects for men and women.

The relationship between smoking and obesity is one of the issue that attracts many authors. Gruber and Frakes (2005) is one of the examples for these studies. They test whether smoking has negative relationship with obesity and BMI as some studies in health economics literature suggests such as Rashad and Grossman (2004). Gruber and Frakes find no evidence for decrease in smoking leads to increase in body weight, in contrast, they find negative association between cigarette taxes and body weight, this implies that reduced smoking leads to decrease in body weight. Therefore, they cannot confirm that increasing obesity rates are explained by the fall in smoking rates. Different from the previous studies, in the empirical analysis, they control the time trends in smoking and use cigarette taxes rather than endogenous cigarette tax rates. They use Behavioral Risk Factor Surveillance System data for the years between 1984 and 2002. In their study, they focus on the individuals who are between 18 and 65 years old. For cigarette tax, the authors use Monthly Cigarette Exercise Tax Data from the Tax Burden On Tobacco. In the empirical analysis, Gruber and Frakes regress BMI on real cigarette tax rate in state, age, sex, income, marital status, education unemployment rate in the state, fixed effects for state and year by applying panel data fixed effects estimation techniques. Their results suggest strong negative effect of education on individual’s BMI and the probability of being obese. Younger individuals have lower weight outcomes. Married and widowed individuals have higher levels of BMI and obesity odds relative to divorced and never married individuals.

Rashad (2006) examines the relationship between smoking and obesity by using First, Second, and Third National Health and Nutrition Examination Surveys. The basic objective of this study is to estimate the structural model of the determinants of adult obesity, in their study smoking is one of the determinants of obesity. Rashad also accounts for the possible endogeneity problem between smoking and BMI, therefore he implements IV techniques in addition to OLS estimation techniques. In the empirical analysis, BMI is a function of energy balance which is found by subtracting energy expenditure (or activity) from caloric intake, as well as education, age, state of residence, race, marital status and income. Activity-adjusted caloric intake and smoking are assumed to be endogenous to each other. The instruments he uses are price of a restaurant meal, the cigarette tax, average January and July temperatures, and clean indoor air laws. Although, OLS estimation results suggest that smoking and BMI are negatively associated with each other, the author finds that the possible negative effect of smoking on BMI decreases when the possible endogeneity problem between the two is taken into account. In addition to this, the effect of caloric intake and exercise disappears when the endogeneity problem is taken into account. Therefore, the author concludes that accumulated behaviors over time and genetic factors might have greater influence on BMI.

Both Chapter Four and Chapter Seven reveal that SAH, smoking and BMI (or obesity) are the most popular health behaviors that are investigated in the literature. However, we also see that in most of the studies other health behaviors especially, alcohol consumption is also examined. There are few studies in health economics literature that "specifically" examines the determinants of being heavy drinker. Ettner (1996) is one of the examples for these studies. In her study, Ettner focuses on the relationship between income and average daily alcohol consumption as well as SAH and other alcoholic variables by using National Survey of Families and Households data for the year 1987, the Survey of Income and Program Participation by using 1986 and 1987 panels, and the National Health Interview Survey for the year 1988. In addition to SAH and alcohol consumption, the author also investigates the relationship between income and work and functional limitations, bed days and scales of depressive symptoms. The authors implement both OLS and IV estimation techniques in order to circumvent the possible endogeneity problems between income and health outcomes considered. They use state unemployment rate, work experience, parental education and spouse characteristics as instruments for family income. For alcohol consumption, the author uses average daily alcohol drinking during the previous two weeks and the number of alcoholic behaviors shown by the respondent during the previous year as health outcome. In the empirical analysis, they also control for several demographic factors such as gender, marital status, the number of individuals living in the household, a dummy variable which indicates whether the individual lives in metropolitan area, race, age, education and the environment that the individual grows up (for instance, they define a dummy variable which is equal to 1 if the individual grows up with an alcoholic). Their results suggest that if income increases, physical and mental health improves, however alcohol consumption also improves. According to Ettner, the positive relationship between income and alcohol consumption occurs most probably due to the fact that higher prevalence of light social drinking among the individuals with high socioeconomic status.

Although it is expected that individuals with higher levels of education consume fruit and vegetables more than the others, the relationship between fruit and vegetable consumption and education is examined almost never. Rather, the relationship between mortality and morbidity and fruit and vegetables consumption is investigated in the literature. Generally, the relationship between mortality and consumption of certain fruits and vegetables are found to be significant. For instance, Gaziano et al. (1995) examine the relationship between mortality due to cardiovascular diseases and carotene containing fruits and vegetables[45] among 1299 elderly Massachusetts residents by using Massachusetts health care panel study for the year 1976. They focus on the age group who are 66 years old or older in their study. In the empirical analysis, they define the frequency of consuming carotene containing fruits and vegetables in the following manner: 1 if the individuals

[45] The foods which include high levels of caratone content are: Carrots and squash, tomatoes, salads and green leafy vegetables such as spinach and greens, dried fruits such as apricots, prunes and raisins, fresh strawberries and fresh lemon and broccoli and brussels sprouts (Source: Gaziano et al. (1995)).

consume everyday, 1/7 if the individual consumes once a week, 1/14 if the individual consumes for less than once a week, 0 if the individual never consumes fruits and vegetables which include carotene. In the empirical analysis, in addition to age and sex, the authors also control relative risk of mortality for smoking, alcohol consumption, exercising and cholesterol intake as these health behaviors can also be causes of cardiovascular diseases. Gaziano et al. conclude that carotene containing fruits and vegetables decrease the risk of cardiovascular disease mortality.

Later on, Hung et al. (2004) examine the relationship between fruits and vegetable consumption and the incidence of cardiovascular disease and cancer among nurses, physicians, dentists, pharmacists, veterinarians residing in the USA. For the study, for female health professionals, the authors use Nurses' Health Study data set for the year 1984 and for male health professionals, they use Health Professionals Follow Up Study data set for the year 1986. They calculate the multivariable adjusted relative risk of diseases by using Cox proportional hazard analysis. Their results indicate that fruit and vegetable consumption decreases the prevalence of cardiovascular diseases, but it does not significantly affect the prevalence of cancer. Their results suggest that especially green leafy vegetable intake shows the strongest inverse association with cardiovascular diseases.

There are few studies on determinants of health behaviors in Turkey. Most of these studies do not use a large data set and these studies are based on self- administrated questionnaires for small sample groups. For instance, Metintaş et al. (1998) examines how prevalence of smoking and alcohol consumption varies between the university students in Turkey. They conduct their analysis only for Eskisehir, which is located in the central region of Turkey and which has two public universities. The authors conduct self-administrated questionnaires to 1474 students attending the university in Eskisehir. They use multiple logistic regression techniques in order to analyze the determinants of smoking habits. Their results suggest that the prevalence of smoking is considerably high among the university students. The prevalence of smoking also differentiates between different faculties. Their findings suggest that the students who attend faculty of arts have the highest probability of smoking, whereas the students who attend medical school have the lowest probability of smoking. They also find that almost 66.5 per cent of the university students are regular or occasional drinkers and they state that there is a significant association between smoking and alcohol consumption as their findings indicate that the students who drink alcohol regularly or occasionally are more likely to be potential smoker.

Kocabas et al. (1994) examine the smoking patterns among medical school students in Turkey. Their results indicate that medical students tend to smoke less than the non-medical students. However, they stress that smoking is more widespread among Turkish medical students relative to the medical students who reside in USA, Europe, Australia and Asia.[46]

Erbaydar et al. (2005) also analyze the determinants of smoking in Turkey. However, they conduct their analysis only for adolescents who are between 13 and 17 years old. As Metintaş et al. (1998), the authors conduct self-administrated questionnaires for 15 cities in Turkey. They find that close friend's smoking habit is the major determinant of the adolescent's smoking behavior. In addition, mother's education and having no information on the negative aspects of smoking are the other factors that induce the prevalence of smoking among the adolescents. Moreover, their results suggest that the adolescents who are not enrolled in education are more likely to smoke relative to those who attend to some school.

Tansel (1993) examines the smoking patterns in Turkey using annual data for the years between 1960 and 1988. In this study the author investigates the demand characteristics of cigarette consumption. She finds that if people are educated about the adverse effects of smoking, the demand for cigarettes may decrease by more than the decrease in the demand that may result from an increase in cigarette prices. Her results also indicate that university graduates are better informed about the adverse effects of smoking.

[46] Russia is not included in Asia region in that study.

Erem et al. (2004) conduct one of the few studies that examine the determinants of obesity in Turkey. In that study Erem and his friends examine the prevalence of overweight and obesity in the Trabzon region by applying both descriptive and logistic regression analyses. They collect the data from different parts of the Trabzon city. The covariates they control in the empirical analysis are; age, sex, marital status, reproductive history in women, level of education, socioeconomic factors (household income and occupation), family history of selected medical conditions (diabetes, hypertension and obesity), life-style factors (smoking habits, physical activity, and alcohol consumption), and indicators for hypertension (systolic blood pressure and diastolic blood pressure levels). Their results indicate that the prevalence of obesity is seen higher among women than men in Trabzon. Regarding marital status, they find that obesity is highest among widowed individuals. Their results also indicate that obesity is positively related to being married, giving up smoking, alcohol consumption, and household income and inversely with level of education, cigarette use and physical activity. They conclude that obesity is also positively related to hypertension.

Hatemi et al. (2003) also examine the prevalence of overweight and obesity in Turkey. They use the Turkish Obesity and Hypertension Study (TOHS) data set from April 1999 to April 2000. In the data set 11 cities from four different geographic regions in Turkey are randomly selected. The survey is conducted for the individuals above 20. The distinguishing feature of TOHS survey is that the respondents' weight, height and blood pressure are not self-reported, they are measured by survey team members (a practitioner, a nurse and a medical secretary). The authors conduct Two-Way ANOVA and logistic regression analysis to analyze the determinants of obesity. In their analysis, the major determinants of obesity are age, gender and blood pressure. The authors find that obesity and overweight are seen widely in Turkey. They also stress that obesity and overweight constitute risk factors for hypertension.

Yumuk (2005) examines the prevalence of obesity in Turkey. He uses Turkish Adult Risk Factor Study (TEKHARF) which is conducted in 1990. It is the first population based health survey data set in Turkey. He makes a descriptive analysis. His results suggest that obesity is an alarming risk factor in Turkey like in the case of developed countries. The author also concludes that men tend to be overweight more than women, however, women are more likely to be obese than men.

To our knowledge, Cesur et al. (2014) is the only study that uses a large micro data set and examines the effect of education on health behaviors. They also use THS data set in that study. Their study is similar to this thesis in the sense that they also take the endogeneity problem between health outcomes and education into account. The authors take the change in compulsory schooling laws in 1997 as the instrument for the individual's education level[47]. Since the people who are affected by this reform are very young, they examine the association between health behaviors and education for the individuals who are between 18 and 30 years old. In contrast, in this thesis, we focus on the individuals who are assumed to complete their schooling, i.e. the individuals who are 25 years old or older.

In the study, they use SAH, obesity, several measures of tobacco consumption, fruit and vegetable consumption and flu vaccination. They do not find a significant linkage between the health behaviors considered and education for women. They do not find a significant relationship between smoking and education level for men for the same age group either. However, they conclude that men with higher levels of education are likely to be obese or overweight more than the less educated ones.

[47] In the study they use the instrument as a proxy whether the individual complete middle school or not.

7.3. Descriptive Statistics

7.3.1. Smoking

OECD (2010) Health Data set reports that 25.4 per cent of adult population in Turkey is regular smoker which is the second highest after Estonia (26.2 per cent) among the OECD countries. Smoking is one of the most harmful health behaviors. Regular smokers are in great risk for cardiovascular disease, chronic lung disease and several types of cancer (Stewart et al., 2009; Chalupka and Warner, 2000). In THS data set smoking does not imply tobacco consumption only. It also includes other types of tobacco products such as cigars. We define an individual as smoker if he/she reports that he/she has been a regular smoker and he/she currently smokes. Table 7.3.1 presents the descriptive statistics for smoking for the pooled 2010 and 2012 survey[48].

Table 7.3.1. Descriptive Statistics for Smoking

Variable	Smoker	Non-Smoker	Total
Male*	0.43 (0.49)	0.57 (0.49)	1.00
Female*	0.17 (0.38)	0.83 (0.38)	1.00
Age (Years)	43.32 (12.04)	49.51 (15.47)	47.74 (14.83)
Age-Squared	2022 (1169)	2690 (1638)	2499 (1549)
Urban*	0.31 (0.46)	0.69 (0.46)	1.00
Rural*	0.22 (0.42)	0.78 (0.42)	1.00
Marital Status			
Married*	0.29 (0.45)	0.71 (0.45)	1.00
Single*	0.36 (0.48)	0.64 (0.48)	1.00
Widowed/Divorced*	0.20 (0.40)	0.80 (0.40)	1.00
Education			

[48] Smoking data is available only for 2010 and 2012.

Table 7.3.1. (Continued)

Years of Schooling	7.78 (4.00)	6.21 (4.58)	6.66 (4.48)
Years of Schooling-Squared	76.57 (70.75)	59.69 (72.23)	64.51 (72.21)
Illiterate*	0.08 (0.27)	0.92 (0.27)	1.00
Non-Graduate*	0.16 (0.36)	0.84 (0.36)	1.00
Primary School*	0.29 (0.45)	0.71 (0.45)	1.00
Middle School*	0.42 (0.49)	0.58 (0.49)	1.00
High School*	0.40 (0.49)	0.60 (0.49)	1.00
University+*	0.30 (0.46)	0.70 (0.45)	1.00
Labor Market Status			
Employed*	0.41 (0.49)	0.59 (0.49)	1.00
Unemployed*	0.51 (0.50)	0.49 (0.50)	1.00
Out of Labor Force*	0.18 (0.39)	0.82 (0.39)	1.00
Log Household Income (TL)	7.05 (0.58)	6.97 (0.60)	7.00 (0.60)
Number of Observations in 2010	3469	8311	11780
Number of Observations in 2012	6446	16490	22936
Total Observations	9915	24801	34716

Source: 2010 – 2012 Turkish Health Survey

Notes : (1)*indicates a dummy variable

(2) The numbers in the paranthesis are standard deviations

Table 7.3.1 reveals that years of schooling is higher among smokers than nonsmokers. Nevertheless, the prevalence of smoking is the highest among middle school graduates. We also note that among males the percentage of smokers is larger than nonsmokers. The fraction of smokers is higher in urban areas than in rural areas. When we compare the married, single and widowed/divorced we see that the proportion of the smokers is highest among the singles. Regarding employment status, we see that percentage of smokers among unemployed is higher than those among the employed and inactive individuals. Finally, we see that household income is slightly higher among smokers.

7.3.2. Alcohol Consumption

According the OECD (2010) Health Data set, only 1.5 per cent of adult population in Turkey consumes alcohol. This amount is very low compared to other OECD countries. The low percentage of alcohol consumption in Turkey is most probably due to religious traditions which prohibit alcohol consumption. Similarly, in THS the proportion of daily alcohol drinkers is very low, less than one percent (0.5 per cent, 0.4 per cent and 0.2 per cent in 2008, 2010 and 2012 respectively). In order to capture the variation in alcohol consumption, we combine the daily and occasional alcohol drinkers and call them as "alcohol drinkers" in our analysis. In other words, we refer to an individual as alcohol drinker if the individual states that he/she currently consumes alcohol regularly or occasionally. Table 7.3.2 reports the descriptive statistics for alcohol consumption[49].

Table 7.3.2. Descriptive Statistics for Alcohol Consumption

Variable	**Alcohol Consumer**	**Non-Alcohol Consumer**	**Total**
Male*	0.20 (0.40)	0.80 (0.40)	1.00
Female*	0.04 (0.21)	0.96 (0.21)	1.00
Age (Years)	43.52 (12.00)	47.91 (15.06)	47.41 (14.81)
Age-Squared	2038 (1150)	2522 (1577)	2466 (1541)
Urban*	0.13 (0.34)	0.87 (0.34)	1.00
Rural*	0.08 (0.27)	0.92 (0.27)	1.00
		Marital Status	
Married*	0.11 (0.32)	0.89 (0.32)	1.00
Single*	0.20 (0.40)	0.80 (0.40)	1.00
Widowed/Divorced*	0.06 (0.25)	0.94 (0.25)	1.00
		Education	

[49] For all remainig tables in this section Source: 2008,2010,2012 Turkish Health Survey Notes : (1)*indicates a dummy variable (2) The numbers in the paranthesis are standard deviations

Table 7.3.2. (Continued)

Years of Schooling	9.63 (4.16)	6.14 (4.33)	6.54 (4.45)
Years of Schooling-Squared	110.11 (81.60)	56.35 (67.17)	62.52 (71.08)
Illiterate*	0.01 (0.07)	0.99 (0.07)	1.00
Non-Graduate*	0.02 (0.15)	0.98 (0.15)	1.00
Primary School*	0.08 (0.28)	0.92 (0.28)	1.00
Middle School*	0.16 (0.37)	0.84 (0.37)	1.00
High School*	0.18 (0.39)	0.82 (0.39)	1.00
University+*	0.27 (0.45)	0.73 (0.45)	1.00
Labor Market Status			
Employed*	0.20 (0.40)	0.80 (0.40)	1.00
Unemployed*	0.20 (0.40)	0.80 (0.40)	1.00
Out of Labor Force*	0.05 (0.22)	0.95 (0.22)	1.00
Log Household Income (TL)	7.19 (0.60)	6.88 (0.61)	6.92 (0.62)
Number of Observations in 2008	1338	10439	11777
Number of Observations in 2010	1495	10285	11780
Number of Observations in 2012	2500	20436	22936
Total Observations	5333	41160	46493

From Table 7.3.2., we see that average years of schooling is higher among alcohol drinkers (9.63 years) than among non-alcohol consumers (6.14 years). The occurrence of alcohol consumption increases as level of education increases and it is highest among the university or higher graduates. Percentage of alcohol drinkers is higher among males than among females. Alcohol consumers are younger than the non-alcohol consumers and the urban residents consume more than the rural residents. Regarding the marital status, we note that the percentage of alcohol drinkers is higher among singles than the married and widowed/divorced. The fraction of alcohol drinkers among employed and unemployed are equal and significantly higher than that of the inactive. Lastly, household income is significantly higher among alcohol consumers.

7.3.3. Fruit and Vegetable Consumption

Fruit and vegetables are necessary for healthy life. Sufficient daily consumption of fruit and vegetables could prevent several kinds of diseases, such as cardiovascular diseases, certain cancers and diabetes (Hung et al. 2004). Hence, it is important to investigate the variation in fruit and vegetable consumption. In our analysis, we define an individual as a regular fruit and vegetable consumer if the individual reports that he/she consumes fruits, vegetables and/or their juice at least once a week. Table 7.3.3 presents the descriptive statistics for fruit and vegetable consumption in Turkey.

Table 7.3.3. Descriptive Statistics for Fruit and Vegetables (FV) Consumption

Variable	Consume FV Regularly	Not Consume FV	Total
Male*	0.59 (0.49)	0.41 (0.49)	1.00
Female*	0.57 (0.49)	0.43 (0.49)	1.00
Age (Years)	46.72 (14.68)	48.34 (14.93)	47.40 (14.81)
Age-Squared	2398 (1516)	2559 (1571)	2466 (1541)
Urban*	0.60 (0.49)	0.40 (0.49)	1.00
Rural*	0.54 (0.50)	0.46 (0.50)	1.00
Marital Status			
Married*	0.58 (0.49)	0.42 (0.49)	1.00
Single*	0.65 (0.48)	0.35 (0.48)	1.00
Widowed/Divorced*	0.54 (0.50)	0.46 (0.50)	1.00
Education			
Years of Schooling	6.85 (4.46)	6.11 (4.39)	6.54 (4.45)
Years of Schooling-Squared	66.80 (72.81)	56.68 (68.19)	62.55 (71.08)
Illiterate*	0.50 (0.50)	0.50 (0.50)	1.00
Non-Graduate*	0.54 (0.50)	0.46 (0.50)	1.00
Primary School*	0.57 (0.50)	0.43 (0.50)	1.00

Table 7.3.3. (Continued)

Middle School*	0.60 (0.49)	0.40 (0.49)	1.00
High School*	0.62 (0.48)	0.38 (0.48)	1.00
University+*	0.64 (0.48)	0.36 (0.48)	1.00
Labor Market Status			
Employed*	0.59 (0.49)	0.41 (0.49)	1.00
Unemployed*	0.57 (0.49)	0.43 (0.49)	1.00
Out of Labor Force*	0.57 (0.49)	0.43 (0.49)	1.00
Log Household Income (TL)	6.93 (0.61)	6.89 (0.62)	6.92 (0.62)
Number of Observations in 2008	9665	2089	11754
Number of Observations in 2010	6074	5696	11770
Number of Observations in 2012	11199	11731	22930
Total Observations	26938	19516	46454

We observe that years of schooling is higher among fruit and vegetable consumers than non-consumers. The prevalence of regular fruit and vegetable consumption is the highest among individuals who have university or higher degree compared to other education groups. The fruit and vegetable consumption is higher among males. Fruit and vegetable consumers are younger and urban residents consume more fruits and vegetables than rural ones. Singles consume more fruits and vegetables than the married and widowed/divorced. Fruit and vegetable consumption is similar among the employed unemployed and inactive. Finally, household income is somewhat higher among the regular fruit and vegetable consumers.

7.3.4. Exercising

Regular physical activity is an important factor that improves the individual’s health. It could prevent many diseases such as heart disease and stroke, high blood pressure, diabetes, obesity, back pain, osteoporosis, and can improve the psychological condition of the individual (Fletcher et al., 1996). Therefore, it is important to examine the variation in physical activity of the individuals. In THS we observe the respondent's frequency of exercise, divided into three categories: High level exercise (such as aerobic exercise or working in construction sector), medium level exercise (such as riding a bicycle or house work) and low level exercise (such as walking). If the individual does not experience one of these activities at least 10 minutes in the reference week, then we assume that individual does not make regular physical activity. Table 7.3.4 presents the descriptive statistic for exercising.

Table 7.3.4. Descriptive Statistics for Exercising

Variable	Exercise Regularly	Not Exercise Regularly	Total
Male*	0.71 (0.45)	0.29 (0.45)	1.00
Female*	0.63 (0.48)	0.37 (0.48)	1.00
Age (Years)	46.21 (13.84)	50.07 (16.33)	47.50 (14.83)
Age-Squared	2326 (1405)	2774 (1757)	2475 (1546)
Urban*	0.67 (0.47)	0.33 (0.47)	1.00
Rural*	0.66 (0.47)	0.34 (0.47)	1.00
Marital Status			
Married*	0.68 (0.47)	0.32 (0.47)	1.00
Single*	0.69 (0.46)	0.31 (0.46)	1.00
Widowed/Divorced*	0.54 (0.50)	0.46 (0.50)	1.00
Education			
Years of Schooling	7.00 (4.36)	5.78 (4.48)	6.60 (4.44)
Years of Schooling-Squared	68.18 (72.66)	53.56 (67.69)	63.31 (71.37)
Illiterate*	0.48 (0.50)	0.52 (0.50)	1.00
Non-Graduate*	0.61 (0.49)	0.39 (0.49)	1.00
Primary School*	0.68 (0.46)	0.32 (0.46)	1.00
Middle School*	0.72 (0.45)	0.28 (0.45)	1.00
High School*	0.72 (0.45)	0.28 (0.45)	1.00
University+*	0.73 (0.44)	0.27 (0.44)	1.00
Labor Market Status			
Employed*	0.73 (0.44)	0.27 (0.44)	1.00
Unemployed*	0.69 (0.46)	0.31 (0.46)	1.00

Table 7.3.4. (Continued)

Out of Labor Force*	0.62 (0.49)	0.38 (0.49)	1.00
Log Household Income (TL)	6.95 (0.61)	6.87 (0.62)	6.93 (0.61)
Number of Observations in 2008	7029	4078	11107
Number of Observations in 2010	7673	3118	10791
Number of Observations in 2012	14366	7341	21707
Total Observations	29068	14537	43605

Table 7.3.4. reveals that the individuals who exercise regularly have higher years of schooling and prevalence of exercise is higher among university graduates and among males. Individuals who exercise regularly are younger. Urban residents exercise slightly more than the rural ones. Married and single individuals exercise more than the widowed/divorced. Employed people exercise more than the unemployed and the inactive. Finally, average household income is slightly higher among individuals who exercise regularly.

7.3.5. BMI

Obesity is an increasing health problem in Turkey. It is important to analyze the determinants of obesity as it is a major source of certain diseases such as cardiovascular diseases, diabetes, and joint problems (Stewart et al., 2009). OECD (2010) Health Data indicate that 21 percent of females and 13.2 percent of males in Turkey were obese in 2010. World Bank (2008) reported that the adoption of Western diets high in refined carbohydrates, saturated fats and sugars and a more sedentary lifestyle are major contributors to the increase in overweight and chronic diseases in Turkey. BMI is used as a tool for determining if an individual is overweight or obese. An individual is considered as obese if his/her BMI is greater than 30, overweight if his/her BMI is greater than 25 and underweight if his/her BMI is under 18.5 according to World Health Organization (WHO) criteria. The BMI in our study is computed from the self-reported height (in centimeters) and weight (in kilograms) in the THS. We calculate individual's BMI by dividing the self-reported weight of respondent (in kilograms) to the square of the self-reported height in meters. Table 7.3.5. presents the descriptive statistics for five different BMI groups as well as the statistics for BMI for the whole sample.

Table 7.3.5. Descriptive Statistics for BMI

Variable	Underweight (BMI<18.5)	Normal Weight (18.5<=BMI<=24.99)	Overweight (25<=BMI<30)	Obese (BMI>=30)	Total
Male*	0.01 (0.10)	0.37 (0.48)	0.45 (0.50)	0.17 (0.38)	1.00
Female*	0.02 (0.15)	0.38 (0.48)	0.34 (0.47)	0.26 (0.44)	1.00
Age (Years)	42.57 (16.79)	43.77 (14.96)	47.47 (13.94)	50.51 (13.08)	46.65 (14.44)
Age-Squared	2093 (1759)	2139 (1534)	2447 (1443)	2722 (1385)	2385 (1489)
Urban*	0.02 (0.13)	0.37 (0.48)	0.39 (0.49)	0.22 (0.41)	1.00
Rural*	0.02 (0.14)	0.37 (0.48)	0.39 (0.48)	0.22 (0.41)	1.00
Marital Status					
Married*	0.01 (0.12)	0.35 (0.48)	0.41 (0.49)	0.23 (0.42)	1.00
Single*	0.05 (0.23)	0.62 (0.49)	0.25 (0.44)	0.08 (0.26)	1.00
Widowed/Divorced*	0.02 (0.14)	0.35 (0.48)	0.35 (0.48)	0.28 (0.45)	1.00
Education					
Years of Schooling	7.73 (4.82)	7.56 (4.54)	6.99 (4.30)	5.81 (4.02)	6.96 (4.39)
Years of Schooling-Squared	82.87 (80.44)	77.77 (76.78)	67.39 (71.75)	49.98 (61.99)	67.78 (72.65)
Illiterate*	0.02 (0.14)	0.34 (0.47)	0.34 (0.48)	0.30 (0.46)	1.00
Non-Graduate*	0.02 (0.12)	0.31 (0.46)	0.37 (0.48)	0.30 (0.46)	1.00
Primary School*	0.01 (0.11)	0.34 (0.47)	0.40 (0.49	0.25 (0.43)	1.00
Middle School*	0.02 (0.14)	0.40 (0.49)	0.40 (0.49)	0.18 (0.38)	1.00
High School*	0.02 (0.14)	0.44 (0.50)	0.39 (0.49)	0.15 (0.36)	1.00
University+*	0.02 (0.16)	0.47 (0.50)	0.38 (0.49)	0.13 (0.33)	1.00
Labor Market Status					
Employed*	0.02 (0.13)	0.41 (0.49)	0.41 (0.49)	0.16 (0.37)	1.00
Unemployed*	0.03 (0.17)	0.53 (0.50)	0.35 (0.48)	0.09 (0.28)	1.00
Out of Labor Force*	0.02 (0.14)	0.34 (0.47)	0.38 (0.48)	0.26 (0.44)	1.00

Table 7.3.5. (Continued)

Log Household Income (TL)	6.88 (0.64)	6.95 (0.63)	6.98 (0.59)	6.95 (0.59)	6.96 (0.61)
Number of Observations in 2008	209	3998	3764	1943	9914
Number of Observations in 2010	201	3844	3928	2317	10290
Number of Observations in 2012	322	7570	8331	4647	20870
Total Observations	732	15412	16023	8907	41074

Accordingly, average years of schooling is lower among obese individuals. The prevalence of obesity is higher among illiterate and non-graduate individuals. Females are more obese than males and males are more overweight than females. Obese individuals are slightly older. There is no difference in the weight ranges of urban and rural residents. The occurrence of being overweight is higher among married and prevalence of obesity is higher among widowed/divorced individuals. Prevalence of being obese is higher among inactive individuals. Lastly, we observe that household income is slightly higher among overweight individuals.

7.4. Empirical Results

In this section we present the effects of individual characteristics on each of the health behavior separately. In this section we assume that health behaviors and education are exogenous to each other. The definitions of health outcomes are told in detail in chapter 3. Smoking, alcohol consumption, fruit and vegetable consumption, and exercising are discrete binary variables, whereas BMI is a continuous dependent variable. Hence assuming that error terms are normally distributed with zero mean and constant variance, we establish a probit model for the health behaviors except BMI, and for BMI we implement OLS estimation techniques.

We run five different probit regressions (OLS for BMI) in the following manner: In the first regression, explanatory variables included are; years of schooling, square of years of schooling and gender dummy. In the second regression, we add age, square of age and a dummy that indicates whether the individual lives in urban or rural areas. Next, in the third regression we add marital status dummies. In the fourth regression we add dummies which shows labor market status of the individual. Lastly, in the fifth regression we include the individual's household income (in logarithms). Here, our objective is to see whether the magnitude and significance of years of schooling variable changes when other controls are added into the regression. We observe that the magnitude of the years of schooling variable reduces slightly as new control variables are included in the regression but its significance never changes, which is consistent with the findings of Cutler and Lleras-Muney (2010).

Next, we replicate the five regressions for each health behavior by adding five dummy variables which indicate the education levels of the individuals in place of years of education. By doing this, first, we aim to see if there would be a change in the interpretation of our findings. We observe that the neither the explanatory power nor the sign of other control variables do not change. Second, we aim to see how the health behaviors vary among different education groups. Indeed, unlike the previous studies, THS data set enable us to see the variation of health behaviors among different education levels.

We can write the model more specifically as follows:

$$H_i = \beta_0 + \beta_1 E_i + X_i{}'\beta_2 + m\beta_3 + \varepsilon \qquad (7.4.1)$$

In equation 7.4.1 H_i refers to health behavior of the individual i. E_i shows the education variable. X_i is the rich set of vector covariates that are gradually added into the regression and m refers to the year dummies.

Year dummies for 2010 and 2012 are included in all of the regressions. For alcohol consumption, the coefficient of the 2010 year dummy is estimated as a positive indicating that compared to 2008 there has been an increase in alcohol consumption in 2010. On the other hand, the coefficient estimate of the 2012 year dummy is negative which implies that in 2012 compared to 2008 alcohol consumption decreased in Turkey. We also find that fruit and vegetable consumption decreases in 2010 and 2012 compared to 2008. The coefficient estimate of the year dummy for 2010 and 2012 are positive for exercise and individual's BMI level, which indicate that in 2010 and 2012 compared to 2008 there has been an increase in the prevalence of exercise and BMI levels. For smoking, only year dummy for 2012 is included in the regressions because smoking data are not available for 2008. We find that the coefficient estimate of the 2012 year dummy is negative. This suggests that in 2012 compared to 2010, there has been a raise in smoking in Turkey. In the rest of this section we examine the variations in each health behavior separately.

7.4.1. Smoking

Table 7.4.1.1. reports the marginal effects (in percentages) from probit estimation results. We observe probability of smoking increases by 3.84 percentage points when schooling increases by a year. The positive association between years of schooling and smoking in our study contradicts with previous studies such as Cutler and Lleras-Muney (2010), Kenkel (1991) and Lantz et al. (2001). They all find a negative and significant relationship between years of schooling and smoking in the USA and UK which are developed countries. Turkey is considered as a developing country. For this reason, the relationship between education and smoking may differ in Turkey from that in the developed countries. Indeed the coefficient estimate of the years of schooling squared indicate that that smoking and years of schooling and smoking has an inverted U-shaped relationship. This implies that the probability of smoking decreases among the highly educated. Table 7.4.1.2. reports the marginal effects (in percentages) (see also Figure 7.4.1.) of probit estimation results where education dummies are used as education control variables instead of years of

schooling. These results support our previous findings. We again find a positive relationship between smoking and education level which declines if the individual has college or higher degree. For instance, the probability of smoking increases by 14.63 percentage points for high school graduates and by 5.42 percentage points for college graduates relative to an illiterate individual. Hence, we conclude that individuals who have college degree are better informed about adverse health effects of smoking than the lower educated people[50].

Table 7.4.1.1. Marginal Effects from Probit Estimation Results for Smoking (%)with Years of Schooling

Variable	(1)	(2)	(3)	(4)	(5)
Years of Schooling	6.05*** (0.20)	3.72*** (0.21)	3.82*** (0.21)	3.87*** (0.21)	3.84*** (0.21)
Years of Schooling Square	-0.30*** (0.01)	-0.21*** (0.01)	-0.21*** (0.01)	-0.22*** (0.01)	-0.22*** (0.01)
Male	20.60*** (0.43)	22.90*** (0.43)	23.92*** (0.43)	20.83*** (0.53)	21.00*** (0.53)
Age($x10^{-1}$)		8.77*** (1.18)	9.53*** (1.21)	8.60*** (1.19)	8.56*** (1.21)
Age Square($x10^{-3}$)		-13.96*** (1.18)	-15.32*** (1.21)	-13.55*** (1.20)	-13.59*** (1.21)
Urban		5.42*** (0.52)	5.13*** (0.53)	5.59*** (0.53)	5.61*** (0.55)
			Marital Status		
Married			-10.96*** (0.86)	-10.07*** (0.86)	-10.28*** (0.86)
Single			-11.57*** (1.19)	-11.52*** (1.19)	-11.78*** (1.21)
			Labor Force Status		
Employed				5.43*** (0.58)	5.31*** (0.59)
Unemployed				11.05*** (1.31)	10.97*** (1.34)
Log Household Income					0.36 (0.47)

[50] For all the calculations in this section Notes: (1)*** indicates 1% level of significance, **indicates 5% level of significance, *indicates 10% level of significance.(2) Robust standard errors are shown in paranthesis. (3) Marginal effects are computed at the means of the variables

Table 7.4.1.1. (Continued)

Dummy12	-2.38*** (0.48)	-2.31*** (0.47)	-2.36*** (0.47)	-2.28*** (0.47)	-2.35*** (0.48)
Pseudo R2	0.09	0.12	0.12	0.13	0.13
(-) Log-Likelihood	18802	18240	18163	18102	17916
N	34716	34716	34716	34716	34350

Source:Authors' computations using 2010-2012 Turkish Health Survey

Table 7.4.1.2. Marginal Effects from Probit Estimation Results for Smoking (%) with Education Levels

Variable	(1)	(2)	(3)	(4)	(5)
			Education		
Non-Graduate	8.50*** (1.34)	6.27*** (1.33)	6.75*** (1.32)	6.84*** (1.31)	6.66*** (1.32)
Primary School	20.16*** (0.91)	10.98*** (0.94)	11.71*** (0.95)	11.67*** (0.94)	11.50*** (0.95)
Middle School	29.45*** (1.09)	17.20*** (1.14)	17.81*** (1.14)	17.62*** (1.14)	17.27*** (1.16)
High School	28.08*** (1.01)	14.74*** (1.08)	15.34*** (1.08)	14.99*** (1.08)	14.63*** (1.12)
University+	19.47*** (1.07)	6.29*** (1.14)	7.11*** (1.14)	5.89*** (1.15)	5.42*** (1.22)
Male	20.61*** (0.43)	22.90*** (0.43)	23.90*** (0.43)	20.79*** (0.53)	20.96*** (0.53)
Age($x10^{-1}$)		9.41*** (1.18)	10.05*** (1.20)	9.15*** (1.19)	9.09*** (1.20)
Age Square($x10^{-3}$)		-14.59*** (1.18)	-15.83*** (1.21)	-14.08*** (1.20)	-14.10*** (1.21)
Urban		5.30*** (0.53)	5.04*** (0.52)	5.49*** (0.53)	5.51*** (0.54)
			Marital Status		
Married			-10.76*** (0.86)	-9.85*** (0.86)	-10.07*** (0.86)
Single			-11.49*** (1.19)	-11.43*** (1.19)	-11.70*** (1.20)
			Labor Force Status		
Employed				5.48*** (0.58)	5.36*** (0.58)
Unemployed				11.07*** (1.31)	10.99*** (1.34)

Table 7.4.1.2. (Continued)

Log Household Income					0.39 (0.46)
Dummy12	-2.42*** (0.48)	-2.34*** (0.47)	-2.38*** (0.47)	-2.31*** (0.47)	-2.37*** (0.48)
Pseudo R2	0.10	0.12	0.13	0.13	0.13
(-) Log-Likelihood	18791	18227	18152	18091	17905
N	34716	34716	34716	34716	34350

Source:Authors' computations using 2010-2012 Turkish Health Survey

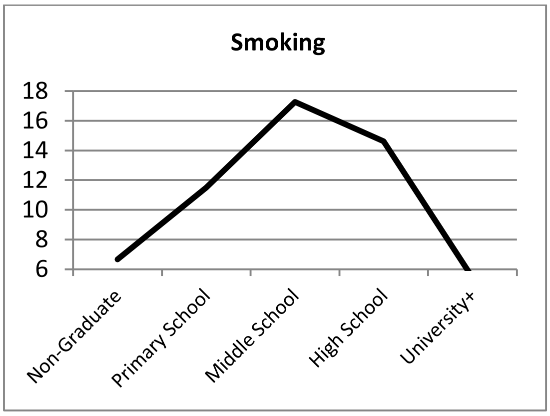


Figure 7.4.1.Marginal Effects for Smoking (x100) by Education Level

Source: Author's Computations from THS data set (2010 and 2012)

The results for the other covariates are similar in Tables 7.4.1.1 and 7.4.1.2. We comment on them briefly. We find that men are more likely to smoke than women. Smoking and age have concave relationship. Moreover, as it is expected, urban residents tend to smoke more than rural residents. Next, we see that married and single individuals are less likely to smoke than widowed / divorced. Further, we note that the probability of an unemployed smoking (10.99) is twice as large as the probability of an employed smoking (5.36) relative to inactive. This result can be attributed to the stressful work life for employed individuals, and to being anxious while looking for a job for the unemployed. Finally, we find that household income does not significantly affect individual's smoking behavior. Cutler and Lleras-Muney (2010) also include labor market status along with other main covariates. They suggest that the inclusion of labor market status variables reduces the education coefficient by 10 percentage points. In our case inclusion of all other covariates reduces the coefficient of years of schooling by almost half (Table 7.4.1.1, model 1 and model 5).

7.4.2. Alcohol Consumption

Table 7.4.2.1 presents the marginal effects (in percentages) from probit estimation results. Our results indicate that there is a positive relationship between education level and alcohol consumption. We find that the probability of alcohol consumption increases by 2.36 percentage points when schooling increases by one year. We replicate our model by adding education dummies in place of years of schooling. The marginal effects (in percentages) (see also Figure 7.4.2) from this regression are reported in Table 7.4.2.2. Table 7.4.2.2 shows that the probability of alcohol consumption increases with education level. It increases by 5.64 per cent for non-graduates, 12.59 percentage points for primary school graduates 16.91 percentage points for middle school graduates by 17.95 percentage points for high school graduates and 20.91 percentage points for the individuals with university or higher degree, compared to illiterates. This result can be attributed to two facts: First, highly educated people participate in social activities more than the low educated due to their larger social networks and they tend to consume more alcohol during the social activities. Second, as Kenkel (1991) suggests, more educated people may know that some drink is good for health, hence they drink more than the others[51].

[51] For all the remaining estimation results' tables and figures in this section, Source:Authors' computations using 2008-2010-2012 Turkish Health Survey

Table 7.4.2.1. Marginal Effects from Probit Estimation Results for Alcohol Consumption (%)with Years of Schooling

Variable	(1)	(2)	(3)	(4)	(5)
Years of Schooling	2.93*** (0.14)	2.44*** (0.15)	2.47*** (0.15)	2.57*** (0.15)	2.36*** (0.15)
Years of Schooling Square	-0.08*** (0.008)	-0.06*** (0.008)	-0.06*** (0.008)	-0.07*** (0.008)	-0.07*** (0.008)
Male	12.69*** (0.29)	12.93*** (0.30)	13.24*** (0.30)	11.11*** (0.35)	11.60*** (0.35)
Age($x10^{-1}$)		7.58*** (0.73)	9.09*** (0.75)	8.45*** (0.74)	7.64*** (0.75)
Age Square($x10^{-3}$)		-8.54*** (0.74)	-10.04*** (0.76)	-8.66*** (0.75)	-8.07*** (0.76)
Urban		1.11*** (0.33)	1.07*** (0.33)	1.41*** (0.33)	0.40 (0.34)
Marital Status					
Married			-4.63*** (0.56)	-4.20*** (0.56)	-4.75*** (0.57)
Single			-1.70*** (0.72)	-1.40** (0.72)	-2.10*** (0.73)
Labor Force Status					
Employed				4.37*** (0.36)	3.96*** (0.37)
Unemployed				4.74*** (0.73)	5.92*** (0.75)
Log Household Income					3.91*** (0.28)
Dummy10	1.22*** (0.38)	1.30*** (0.38)	1.26*** (0.38)	1.23*** (0.38)	0.64* (0.38)
Dummy12	-1.42*** (0.33)	-1.37*** (0.33)	-1.44*** (0.33)	-1.50*** (0.33)	-2.76*** (0.35)
Pseudo R2	0.15	0.16	0.16	0.17	0.17
(-) Log-Likelihood	13984	13883	13836	13761	13537
N	46493	46493	46493	46493	46024

Table 7.4.2.2. Marginal Effects from Probit Estimation Results for Alcohol Consumption (%) with Education Levels

Variable	(1)	(2)	(3)	(4)	(5)
			Education		
Non-Graduate	5.85*** (1.37)	5.58*** (1.36)	5.91*** (1.35)	5.98*** (1.35)	5.64*** (1.35)
Primary School	15.57*** (1.05)	13.31*** (1.06)	13.78*** (1.06)	13.82*** (1.06)	12.59*** (1.06)
Middle School	21.23*** (1.11)	18.62*** (1.13)	18.92*** (1.13)	18.91*** (1.13)	16.91*** (1.14)
High School	23.25*** (1.08)	20.46*** (1.11)	20.72*** (1.11)	20.62*** (1.11)	17.95*** (1.12)
University+	28.21*** (1.08)	25.43*** (1.11)	25.60*** (1.11)	24.83*** (1.11)	20.91** (1.14)
Male	12.59*** (0.29)	12.84*** (0.30)	13.16*** (0.30)	11.06*** (0.35)	11.56*** (0.35)
Age($x10^{-1}$)		7.31*** (0.73)	8.80*** (0.76)	8.18*** (0.75)	7.38*** (0.75)
Age Square($x10^{-3}$)		-8.24*** (0.75)	-9.74*** (0.77)	-8.38*** (0.76)	-7.82*** (0.76)
Urban		1.15*** (0.33)	1.12*** (0.33)	1.44*** (0.33)	0.44 (0.34)
			Marital Status		
Married			-4.78*** (0.57)	-4.35*** (0.57)	-4.90*** (0.57)
Single			-1.80*** (0.73)	-1.51*** (0.72)	-2.21*** (0.73)
			Labor Force Status		
Employed				4.30*** (0.36)	3.89*** (0.37)
Unemployed				4.67*** (0.73)	5.85*** (0.75)
Log Household Income					3.90*** (0.28)
Dummy10	1.22*** (0.38)	1.30*** (0.38)	1.25*** (0.38)	1.22*** (0.38)	0.63* (0.38)
Dummy12	-1.44*** (0.33)	-1.39*** (0.33)	-1.46*** (0.33)	-1.52*** (0.33)	-2.77*** (0.35)
Pseudo R2	0.16	0.16	0.17	0.17	0.18
(-) Log-Likelihood	13961	13867	13818	13746	13523
N	46493	46493	46493	46493	46024

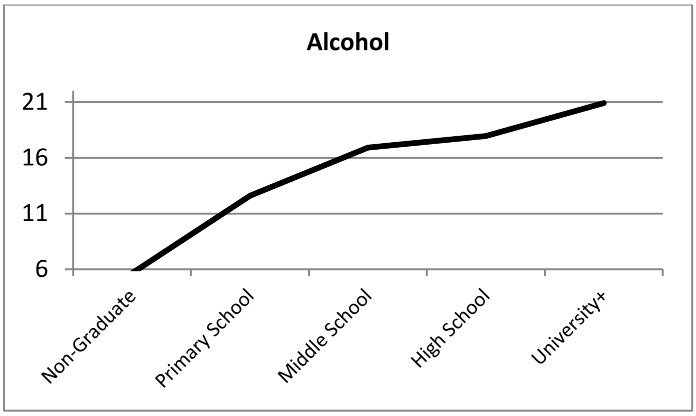


Figure 7.4.2. Marginal Effects for Alcohol Consumption (x100) by Education Level

We next consider the rest of the covariates. We see that males tend to consume more alcohol than females. Like in the case of smoking, the relationship between alcohol consumption and age is an inverted U-shaped. Alcohol consumption increases with age and roughly at age 48 alcohol consumption reaches a maximum after which it starts to decrease. The urban dwellers consume significantly more alcohol than the rural ones albeit it looses its significance when we control for income. The probability of alcohol consumption of married and single people are significantly less than that of widowed/divorced. Moreover, being in labor force also positively and significantly affects the probability of alcohol consumption. The probability of alcohol consumption increases by approximately 4 percentage points and 6 percentage points for the employed and unemployed respectively compared to an inactive person. This finding may again be attributed to larger social networks for the employed and the anxiety/stress for the unemployed. Finally, we find that an increase in log of household income leads to 3.90 percentage points increase in probability of alcohol consumption.

In short, our results are consistent with Kenkel (1991) as well as Ettner (1996) who suggest higher probability of light drinking among higher socioeconomic groups. However, Cutler and Lleras-Muney (2010) find a negative association between probability of being heavy alcohol drinker and education It is important to note that the number of heavy drinkers in our data set is very small for a separate analysis. Thus; our results mostly explain the variations in light alcohol consumption by education and other determinants like Kenkel and Ettner.

7.4.3. Fruit and Vegetable Consumption

Table 7.4.3.1 presents the marginal effects (in percentages) from probit estimation results We find that fruit and vegetable consumption is positively and significantly associated with education level. We observe that the probability of fruit and vegetable consumption increases by 1.42 percentage points when years of schooling increase by a year. We re-estimate the regression by dropping years of schooling and including education level dummies in place. Marginal effects (in percentages) (see also Figure 7.4.3) from these probit estimation results are reported in Table 7.4.3.2. We again find a positive relationship between fruit and vegetable consumption and education level. the probability of alcohol consumption increases with education level. It increases by 3.58 percentage points for non-graduates, 5.95 percentage points for primary school graduates 8.12 percentage points for middle school graduates by 8.56 percentage points for high school graduates and 8.81 percentage points for the individuals with university or higher degree compared to illiterates. We can conclude that people with higher levels of education are better informed about the benefits of fruit and vegetable consumption.

Table 7.4.3.1. Marginal Effects from Probit Estimation Results for Fruit and Vegetable Consumption (%)with Years of Schooling

Variable	(1)	(2)	(3)	(4)	(5)
Years of Schooling	1.87*** (0.16)	1.76*** (0.17)	1.77*** (0.17)	1.74*** (0.17)	1.42*** (0.17)
Years of Schooling Square	-0.05*** (0.01)	-0.05*** (0.01)	-0.05*** (0.01)	-0.05*** (0.01)	-0.05*** (0.01)
Male	-0.16 (0.45)	0.11 (0.46)	-0.09 (0.47)	1.25** (0.54)	1.94*** (0.55)
Age($x10^{-1}$)		-3.02*** (1.01)	-2.19** (1.04)	-2.04** (1.04)	-3.34*** (1.05)
Age Square($x10^{-3}$)		3.02*** (0.97)	2.44** (1.00)	1.99** (1.01)	3.02*** (1.01)
Urban		4.64*** (0.50)	4.71*** (0.50)	4.47*** (0.50)	3.01*** (0.51)
			Marital Status		
Married			1.01 (0.78)	0.63 (0.79)	-0.21 (0.79)
Single			4.89*** (1.16)	5.03*** (1.16)	4.09*** (1.17)
			Labor Force Status		
Employed				-2.42*** (0.57)	-3.01*** (0.58)
Unemployed				-6.33*** (1.36)	-4.52*** (1.38)
Log Household Income					5.85*** (0.44)
Dummy10	-32.25*** (0.59)	-32.30*** (0.59)	-32.30*** (0.59)	-32.28*** (0.59)	-33.13*** (0.59)
Dummy12	-35.23*** (0.51)	-35.34*** (0.51)	-35.35*** (0.51)	-35.38*** (0.50)	-37.29*** (0.52)
Pseudo R2	0.07	0.07	0.07	0.08	0.08
(-) Log-Likelihood	29298	29251	29240	29225	28854
N	46454	46454	46454	46454	45990

Table 7.4.3.2. Marginal Effects from Probit Estimation Results for Fruit and Vegetable Consumption (%) with Education Levels

Variable	(1)	(2)	(3)	(4)	(5)
			Education		
Non-Graduate	4.53*** (1.05)	4.27*** (1.05)	4.22*** (1.05)	4.11*** (1.05)	3.58*** (1.06)
Primary School	8.12*** (0.68)	7.75*** (0.73)	7.78*** (0.74)	7.73*** (0.73)	5.95*** (0.75)
Middle School	12.21*** (0.95)	11.06*** (1.02)	10.96*** (1.02)	10.97*** (1.02)	8.12*** (1.04)
High School	14.02*** (0.83)	12.56*** (0.92)	12.35*** (0.92)	12.45*** (0.92)	8.56*** (0.97)
University+	16.27*** (0.88)	14.58*** (0.97)	14.12*** (0.98)	14.63*** (0.99)	8.81*** (1.08)
Male	-0.18 (0.45)	0.08 (0.46)	-0.11 (0.47)	1.23** (0.54)	1.92*** (0.55)
Age($x10^{-1}$)		-3.07*** (1.02)	-2.24*** (1.05)	-2.11** (1.05)	-3.37*** (1.06)
Age Square($x10^{-3}$)		3.07*** (0.98)	2.49*** (1.01)	2.04** (1.01)	3.05*** (1.02)
Urban		4.65*** (0.49)	4.72*** (0.50)	4.48*** (0.50)	3.02*** (0.52)
			Marital Status		
Married			0.98 (0.78)	0.58 (0.79)	-0.24 (0.79)
Single			4.86*** (1.16)	5.00*** (1.16)	4.06*** (1.17)
			Labor Force Status		
Employed				-2.44*** (0.57)	-3.02*** (0.58)
Unemployed				-6.33*** (1.36)	-4.51*** (1.38)
Log Household Income					5.85*** (0.44)
Dummy10	-32.26*** (0.59)	-32.32*** (0.59)	-32.32*** (0.59)	-32.30*** (0.59)	-33.15*** (0.59)
Dummy12	-35.24*** (0.51)	-35.35*** (0.51)	-35.36*** (0.51)	-35.39*** (0.51)	-37.30*** (0.53)
Pseudo R2	0.07	0.07	0.07	0.08	0.08
(-) Log-Likelihood	29297	29250	29239	29224	28853
N	46454	46454	46454	46454	45990

Source:Authors' computations using 2008-2010-2012 Turkish Health Survey

Notes:

(1)*** indicates 1% level of significance, **indicates 5% level of significance, *indicates 10% level of significance.

(2) Robust standard errors are shown in paranthesis.

(3) Marginal effects are computed at the means of the variables

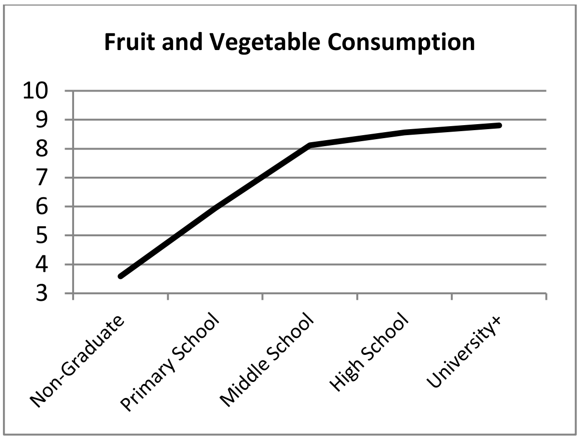


Figure 7.4.3. Marginal Effects for Fruit and Vegetable Consumption (x100) by Education Level

For the rest of the covariates, we observe that males are more likely to consume fruits and vegetables than females. The probability of fruit and vegetable consumption and age has a U-shape relationship indicating an initial decline up to age 56 and an increase afterwards. Urban residents tend to consume more fruit and vegetables than rural ones. Singles tend to consume more fruits and vegetables than the widowed/divorced while the marginal effect for the married is not significantly different from the latter group. The probability of consuming fruit and vegetables for the employed and unemployed individuals is significantly smaller than that of the inactive. Finally, as household income increases the probability of consuming fruits and vegetables also increases.

7.4.4. Exercise

Table 7.4.4.1. shows the marginal effects (in percentages) from probit estimation results for exercise. We find that regular exercise is positively and significantly associated with education level. The probability of exercise increases by 2.65

percentage points when years of schooling increase by a year. We re-estimate the regression by dropping years of schooling and including education level dummies in place. Marginal effects (in percentages) (see also Figure 7.4.4) from these probit estimation results are reported in Table 7.4.4.2. We again find a positive relationship between exercising and education level. The probability of exercising increases with education level. It increases by 8.80 percentage points for non-graduates, 11.57 percentage points for primary school graduates 14.94 percentage points for middle school graduates, by 15.10 percentage points for high school graduates and 15.57 percentage points for the individuals with university or higher degree compared to illiterates. Our results are consistent with previous literature findings. For instance, the studies for developed countries such as Kenkel (1991) and Lantz et al. (2001) in the USA also find a positive relationship between schooling and exercise.

Table 7.4.4.1. Marginal Effects from Probit Estimation Results for Exercise (%)with Years of Schooling

Variable	(1)	(2)	(3)	(4)	(5)
Years of Schooling	3.82*** (0.16)	2.64*** (0.17)	2.60*** (0.17)	2.66*** (0.17)	2.65*** (0.18)
Years of Schooling Square	-0.17*** (0.01)	-0.11*** (0.01)	-0.10*** (0.01)	-0.11*** (0.01)	-0.11*** (0.01)
Male	4.67*** (0.46)	5.43*** (0.46)	5.27*** (0.47)	3.46*** (0.55)	3.53*** (0.55)
Age($x10^{-1}$)		16.07*** (1.00)	15.22*** (1.04)	14.97*** (1.04)	14.83*** (1.05)
Age Square($x10^{-3}$)		-18.01*** (0.96)	-17.15*** (0.99)	-16.49*** (1.00)	-16.37*** (1.00)
Urban		-4.44*** (0.51)	-4.44*** (0.51)	-3.94*** (0.51)	-3.82*** (0.53)
			Marital Status		
Married			1.94** (0.78)	2.29*** (0.78)	2.10*** (0.78)
Single			-0.29 (1.14)	0.12 (1.15)	-0.09 (1.15)
			Labor Force Status		
Employed				4.06*** (0.58)	4.11*** (0.58)
Unemployed				0.37 (1.37)	0.18 (1.39)
Log Household Income					0.21 (0.45)
Dummy10	7.78*** (0.62)	8.24*** (0.62)	8.26*** (0.62)	8.23*** (0.62)	8.22*** (0.62)
Dummy12	2.13*** (0.53)	2.64*** (0.53)	2.66*** (0.53)	2.64*** (0.53)	2.76*** (0.55)

Table 7.4.4.1. (Continued)

Pseudo R2	0.02	0.03	0.04	0.04	0.04
(-) Log-Likelihood	27092	26783	26777	26752	26486
N	43605	43605	43605	43605	43206

Table 7.4.4.2. Marginal Effects from Probit Estimation Results for Exercise (%) with Education Levels

Variable	(1)	(2)	(3)	(4)	(5)
			Education		
Non-Graduate	10.21*** (1.06)	8.89*** (1.06)	8.78*** (1.06)	8.96*** (1.06)	8.80*** (1.06)
Primary School	16.80*** (0.68)	11.77*** (0.73)	11.56*** (0.74)	11.64*** (0.74)	11.57*** (0.75)
Middle School	20.26*** (0.96)	15.29*** (1.03)	15.19*** (1.03)	15.11*** (1.03)	14.94*** (1.05)
High School	20.56*** (0.83)	15.59*** (0.92)	15.55*** (0.93)	15.26*** (0.93)	15.10*** (0.97)
University+	21.49*** (0.88)	16.82*** (0.97)	16.90*** (0.98)	15.82*** (0.99)	15.57*** (1.08)
Male	4.58*** (0.46)	5.35*** (0.46)	5.21*** (0.47)	3.40** (0.55)	3.47*** (0.55)
Age($x10^{-1}$)		15.88*** (1.01)	15.07*** (1.04)	14.83*** (1.04)	14.69*** (1.05)
Age Square($x10^{-3}$)		-17.84*** (0.97)	-17.03*** (1.00)	-16.38*** (1.00)	-16.26*** (1.01)
Urban		-4.41*** (0.51)	-4.38*** (0.51)	-3.92*** (0.51)	-3.81*** (0.53)
			Marital Status		
Married			1.81** (0.78)	2.16*** (0.78)	1.97** (0.78)
Single			-0.43 (1.14)	-0.01 (1.15)	-0.21 (1.15)
			Labor Force Status		
Employed				4.06*** (0.58)	4.11*** (0.58)
Unemployed				0.38 (1.37)	0.18 (1.39)
Log Household Income					0.20 (0.45)
Dummy10	7.74*** (0.62)	8.20*** (0.62)	8.22*** (0.62)	8.19*** (0.62)	8.17*** (0.62)
Dummy12	2.13*** (0.53)	2.64*** (0.53)	2.66*** (0.53)	2.64*** (0.53)	2.76*** (0.55)
Pseudo R2	0.02	0.04	0.04	0.04	0.04
(-) Log-Likelihood	27076	26773	26767	26741	26476
N	43605	43605	43605	43605	43206

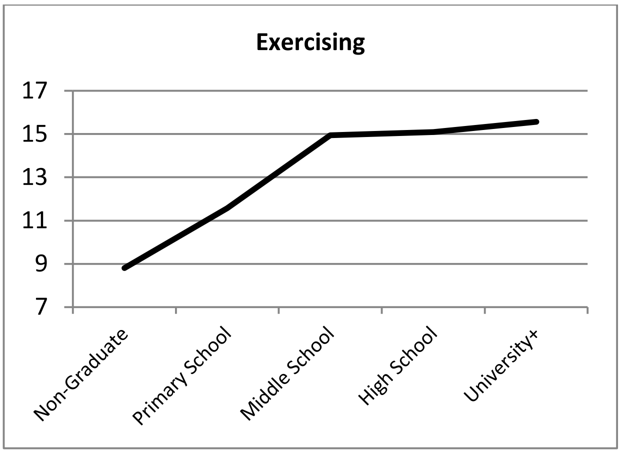


Figure 7.4.4. Marginal Effects for Exercising (x100) by Education Level

We now consider the rest of the covariates. Males tend to exercise more than the females. The probability of exercise and age has a concave relationship: Individuals are more likely to exercise as they get older, but roughly at age 46, probability of exercise starts to decrease. Rural residents tend to exercise more than the urban ones. The probability of exercise is significantly higher for the married people than for the widowed/divorced, whereas exercise behavior of single individuals do not significantly differ from that of widowed/divorced. Employed individuals tend to exercise significantly more than inactive while the exercise behavior of the unemployed is not significantly different from that of the inactive. The probability of exercise increases by 4.11 percentage points if the individual is employed. These results are consistent with our expectations, since employed people are physically more active than the unemployed or the inactive. Finally, our results suggest that household income does not significantly affect the exercise behavior of the individual.

7.4.5. Body Mass Index (BMI)

Table 7.4.5.1 presents the OLS estimation results where the dependent variable is the individual's BMI. We find that an increase in years of schooling results with normal ranges of BMI. This result is similar to the results of the previous literature such as Kemptner et al. (2011), Brunello et al. (2011), Webbink et al (2010), Cutler and Lleras-Muney (2010) and Lantz et al. (2001). We find that one year increase in years of schooling leads to 0.11 unit decrease in individual's BMI level. Table 7.4.5.2 reports OLS estimation results where we drop years of schooling and add dummy variables for education levels instead (see also Figure 7.4.5). Our results suggest that BMI levels of illiterate and non-graduate people do not significantly differ from each other when we add logarithms of household income into our analysis. For other education groups, we observe that as education level increases, the individual's BMI level decreases. The BMI level decreases even more (by 1.72 units) if the individual has university or higher degree. This result suggests that more educated people are better informed about the risks of overweight or obesity.

Table 7.4.5.1. OLS Estimation Results for BMI with Years of Schooling

Variable	(1)	(2)	(3)	(4)	(5)
Years of Schooling	-0.13*** (0.02)	-0.07*** (0.02)	-0.08*** (0.02)	-0.09*** (0.02)	-0.11*** (0.02)
Years of Schooling Square	-0.003** (0.001)	-0.003*** (0.001)	-0.001*** (0.001)	-0.001*** (0.001)	-0.001*** (0.001)
Male	-0.22*** (0.04)	-0.41*** (0.04)	-0.36*** (0.04)	-0.13*** (0.05)	-0.08*** (0.05)
Age(x10^{-1})		5.11*** (0.10)	4.73*** (0.10)	4.76*** (0.10)	4.64*** (0.10)
Age Square(x10^{-3})		-4.50*** (0.10)	-4.20*** (0.10)	-4.28*** (0.10)	-4.19*** (0.10)
Urban		0.46*** (0.05)	0.44*** (0.05)	0.40*** (0.05)	0.30*** (0.05)
Marital Status					
Married			-0.12 (0.09)	-0.19** (0.09)	-0.25*** (0.09)
Single			-1.55*** (0.12)	-1.52*** (0.12)	-1.60*** (0.12)
Labor Force Status					
Employed				-0.41*** (0.05)	-0.44*** (0.05)
Unemployed				-1.12*** (0.11)	-0.96*** (0.11)
Log Household Income					0.43*** (0.04)
Dummy10	0.41*** (0.06)	0.31*** (0.06)	0.32*** (0.06)	0.32*** (0.06)	0.27*** (0.06)
Dummy12	0.57*** (0.05)	0.44*** (0.05)	0.44*** (0.05)	0.44*** (0.05)	0.29*** (0.05)
R2	0.03	0.10	0.10	0.10	0.11
Adjusted R2	0.03	0.10	0.10	0.10	0.11
N	41074	41074	41074	41074	40699

Table 7.4.5.2. OLS Estimation Results for BMI with Education Levels

Variable	(1)	(2)	(3)	(4)	(5)
Education					
Non-Graduate	0.20 (0.14)	0.24* (0.13)	0.25* (0.13)	0.23* (0.13)	0.21 (0.13)
Primary School	-0.33*** (0.09)	-0.22** (0.09)	-0.25*** (0.09)	-0.26*** (0.09)	-0.38*** (0.09)
Middle School	-1.19*** (0.11)	-0.69*** (0.11)	-0.67*** (0.11)	-0.67*** (0.11)	-0.87*** (0.11)
High School	-1.76*** (0.10)	-1.15*** (0.10)	-1.09*** (0.10)	-1.08*** (0.10)	-1.35*** (0.11)
University+	-2.12*** (0.10)	-1.55*** (0.10)	-1.40*** (0.10)	-1.32*** (0.10)	-1.72*** (0.11)
Male	-0.24*** (0.04)	-0.42*** (0.04)	-0.37*** (0.04)	-0.14*** (0.05)	-0.09 (0.05)
Age($x10^{-1}$)		5.07*** (0.10)	4.71*** (0.10)	4.73*** (0.10)	4.61*** (0.11)
Age Square($x10^{-3}$)		-4.46*** (0.10)	-4.18*** (0.10)	-4.26*** (0.10)	-4.16*** (0.10)
Urban		0.46*** (0.05)	0.45*** (0.05)	0.41*** (0.05)	0.30*** (0.05)
Marital Status					
Married			-0.14 (0.09)	-0.21** (0.09)	-0.27** (0.09)
Single			-1.56*** (0.12)	-1.54*** (0.12)	-1.62*** (0.12)
Labor Force Status					
Employed				-0.41*** (0.05)	-0.45*** (0.05)
Unemployed				-1.12*** (0.11)	-0.96*** (0.11)
Log Household Income					0.43*** (0.04)
Dummy10	0.42*** (0.06)	0.31*** (0.06)	0.32*** (0.06)	0.32*** (0.06)	0.26*** (0.06)
Dummy12	0.58*** (0.05)	0.44*** (0.05)	0.44*** (0.05)	0.44*** (0.05)	0.29*** (0.05)
R2	0.03	0.10	0.10	0.10	0.11
Adjusted R2	0.03	0.10	0.10	0.10	0.10
N	41074	41074	41074	41074	40699

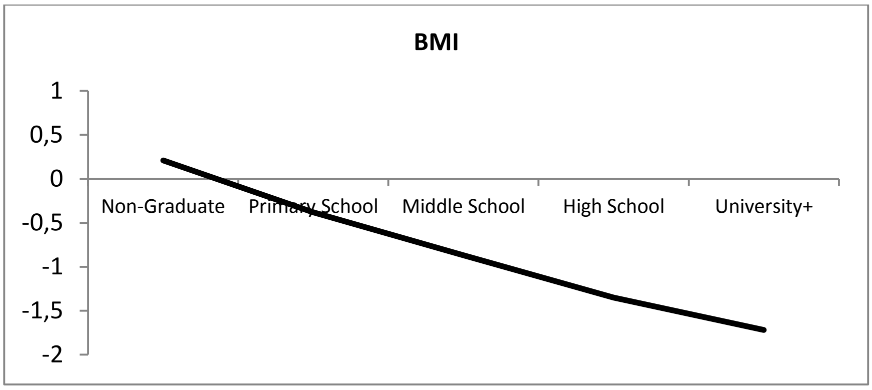


Figure 7.4.5. Coefficients for BMI by Education Level

For the rest of the covariates, we first observe that females have significantly higher levels of BMI than males. However, when we add the logarithm of household income into our analysis, we see that BMI levels of males and females do not significantly differ from each other. BMI level increases with age at a decreasing rate. Urban residents have higher BMI levels than rural ones. Our results indicate that the BMI levels of married and single individuals are significantly lower from that of the widowed/divorced people. Next, we find that both employed and unemployed individuals have lower BMI levels than the inactive. This may be due to the more sedentary life-styles of inactive people. Finally, we find that as household income increases so does BMI.

7.5. Instrumental Variable (IV) Approach

In section 7.5 we revisit the endogeneity problem between health and education. This time, the health outcomes are the different health behaviors. We assume that the individual's health behaviors may also be endogenous to the individual's years of schooling because of the reasons discussed in chapter five. Therefore, usual OLS and probit estimates may give inconsistent results. In this section, we also implement instrumental variable (IV) approach in order to obstruct the problems due to endogeneity. Most of the studies in the literature that we discuss in section 7.2 use IV methodology to test the relationship between health behaviors and education. In section 7.5., we again use the evidence in educational expansion in Turkey in the early 1960s as the instrument for the individual's years of schooling.

In the IV estimation setting we only include purely exogenous covariates as control variables. Namely, we only include age dummies, gender and region controls in order to see the impact of years of schooling on different health behaviors. There are several reasons for it. First, individual's employment status and household income may also be endogenous to health behaviors. In addition, these variables may be directly related to the individual's years of schooling. It is important to note that including age dummies is more flexible specification than including polynomials of trends and age for controlling the birth year dummies. In this specification, at first, we do not include the marital status, then in the second step, marital status dummies are also added into the regression. Because, we discuss in chapter five that approximately 97 per cent of the individuals are married at least once in their life in Turkey. Therefore, marital status and health behaviors may not create a two-way causality. In this specification, we assume that individuals who were born in or after 1952 are affected from the education reforms in the early 1960s assuming that school starting age is 6 in Turkey (The individuals who are 9 years old in September 1961 are assumed to be affected from those reforms). Individual's years of schooling is calculated in the same way as in Chapter Six if the individual was born before 1952 and if the individual is graduated from primary school. All the results are clustered by birth year.

In this empirical analysis, we establish an IVprobit model and for BMI we establish an two stage least squares (2SLS) IV model in order to find the causal link between the individual's years of schooling and health behaviors. The model in this section can be defined as follows:

$$E_i = \delta + X_i \alpha_1 + R61_i \alpha_2 + \varepsilon_i \qquad (7.5.1)$$

$$H_i = \alpha + \hat{E}_i \beta_1 + X_i \beta_2 + u_i \qquad (7.5.2)$$

In the above equations, the vector X includes individual's gender, age dummies and the region where he/she lives. Age dummies are defined in the following manner: For instance, for a person who is 30 years old in the time of the survey agedummy30 is equal to 1, and it is equal to 0 otherwise. The age dummies are defined in the similar manner for other age groups as well. In equation (7.5.1), R61 is the instrument for individual's years of schooling, which is equal to 1 if the individual was born in or later than 1952. In equation (7.5.1) the dependent variable E shows the individual's completed years of schooling. Next, in equation (7.5.2) H refers to different health behavior in different specifications. The construction of H is the same as in equation (7.4.1). In some specifications we also add cohort controls which are defined in five year intervals. We do not report results with cohort controls in the tables below. They are nearly the same as the results without cohort controls. Last, in some specifications we add marital status dummies. We do not report the marital status dummies coefficient, we discuss on them at the end of this section. Last, we present the results for three different samples (Sample A, Sample B and Sample C). The construction of these samples are discussed in Chapter Six.

Table 7.5.1. First Stage Results

Dependent Variable: Years of Schooling					
Sample A: Birth Cohort 1909-1987					
Health Behavior	Smoking	Alcohol Consumption	Fruit and Vegetable Consumption	Exercising	BMI
Policy Dummy	1.14*** (0.17)	1.16*** (0.19)	1.15*** (0.15)	1.13*** (0.22)	1.07*** (0.17)
Observations	34716	46246	46454	43599	41074
R-Squared	0.27	0.27	0.27	0.27	0.23
F-Statistics	41.07	37.44	38.27	25.90	23.19
Endogeneity Test Statistic	442.62*** (p-value:0.007)	28.66*** (p-value:0.00)	184.02*** (p-value:0.00)	2.20 (p-value:0.13)	5.87*** (p-value:0.01)
Age Dummies	Yes	Yes	Yes	Yes	Yes
Sample B: Birth Cohort 1947-1957					
Policy Dummy	1.13*** (0.17)	1.15*** (0.19)	1.14*** (0.19)	1.12*** (0.22)	1.07*** (0.17)
Observations	5763	7880	7875	7372	6789
R-Squared	0.19	0.19	0.19	0.19	0.16
F-Statistics	45.32	34.54	35.25	24.44	21.88
Endogeneity Test Statistic	183.53*** (p-value: 0.00)	27.15*** (p-value: 0.00)	39.86*** (p-value: 0.00)	3.38* (p-value:0.06)	4.68** (p-value: 0.03)
Age Dummies	Yes	Yes	Yes	Yes	Yes
Sample C: Birth Cohort 1942-1962					
Policy Dummy	1.13*** (0.17)	1.15*** (0.19)	1.14*** (0.18)	1.12*** (0.22)	1.07*** (0.17)
Observations	10959	14953		14050	12907
R-Squared	0.21	0.21	0.21	0.20	0.18
F-Statistics	44.97	36.14	36.91	25.41	22.93

Table 7.5.1. (Continued)

Endogeneity Test Statistic	322.51*** (p- value: 0.00)	31.09*** (p- value: 0.00)	42.02*** (p- value: 0.00)	3.79** (p-value: 0.05)	4.58** (p- value: 0.03)
Age Dummies	Yes	Yes	Yes	Yes	Yes

Notes: (1) *** indicates 1 per cent level of significance, ** indicates 5 per cent level of significance and * indicates 10 per cent level of significance.

(2)Standard Error which are clustered by birth year are shown in parenthesis.

(3) For discrete dependent variables, we present Wald Exogeneity test statistic, for BMI we present Durbin Endogeneity Test Statistic.

First stage results suggest that the instrument is valid (F-statistics are greater than 10), and the education reforms positively affect the individual's years of schooling. The first statistics results suggest that the educational expansion in the early 1960s lead to approximately 1.13 years increase in individual's years of schooling. In addition, as the p-values of endogeneity test statistics are smaller than 0.05, we can reject the null hypothesis that individual's years of schooling are exogenous to the health outcomes but one exception. We cannot find evidence that years of schooling and the probability of exercising are endogenous to each other for the whole sample, which includes all the individuals who are above 25 years old. Nevertheless, we also present the IV-Probit estimation results for that sample as well.

Table 7.5.2. Second Stage Results

Dependent Variable:Health Behaviors	Smoking	Alcohol Consumption	Fruit and Vegetable Consumption	Exercising	BMI
Sample A: Birth Cohort 1909-1987					
Years of Schooling	-0.12*** (0.004)	-0.13*** (0.004)	-0.62*** (0.09)	-0.009 (0.02)	0.32* (0.17)
Male	0.97*** (0.02)	0.95*** (0.03)	1.33*** (0.20)	0.24*** (0.04)	-1.23*** (0.37)
Urban	0.43*** (0.02)	0.46*** (0.04)	1.52*** (0.22)	-0.04 (0.05)	-0.50 (0.41)
Age Dummies	Yes	Yes	Yes	Yes	Yes
Cohort Controls	No	No	No	No	No
Sample B: Birth Cohort 1947-1957					
Years of Schooling	-0.12*** (0.006)	-0.13*** (0.03)	-0.24*** (0.01)	-0.009 (0.02)	0.30** (0.15)
Male	1.12*** (0.04)	1.19*** (0.04)	0.67*** (0.03)	0.33*** (0.05)	-3.11** (0.40)
Urban	0.48*** (0.04)	0.50*** (0.04)	0.57*** (0.01)	0.02 (0.06)	-0.23 (0.44)
Age Dummies	Yes	Yes	Yes	Yes	Yes
Cohort Controls	No	No	No	No	No
Sample C: Birth Cohort 1942-1962					
Years of Schooling	-0.12*** (0.005)	-0.13*** (0.02)	-0.24*** (0.01)	-0.009 (0.02)	0.31** (0.16)
Male	1.05*** (0.03)	1.13*** (0.04)	0.64*** (0.03)	0.32*** (0.05)	-2.81*** (0.39)

Table 7.5.2. (Continued)

Urban	0.44*** (0.03)	0.51*** (0.04)	0.59*** (0.02)	0.01 (0.05)	-0.32 (0.42)
Age Dummies	Yes	Yes	Yes	Yes	Yes

Notes: (1) *** indicates 1 per cent level of significance, ** indicates 5 per cent level of significance and * indicates 10 per cent level of significance.

(2)Standard Error which are clustered by birth year are shown in parenthesis.

For all the health behaviors, the sign of years of schooling coefficient differs from probit (OLS for BMI) results. First, we observe the prevalence of smoking decreases as the individual's years of schooling increases. It is consistent with the findings for developed countries, and in fact it is a satisfying result since this implies individual's probability of smoking decreases with the education level. However, this finding contradicts with the probit results[52]. In the LATE context, we can conclude that the individuals who are affected from the educational reforms are likely to smoke less since they have more years of schooling. Similar case also occurs for alcohol consumption, in the IV estimation context, we see that alcohol consumption reduces by education level. The same contradicting result also occurs for BMI. We observe that individual's BMI level rises with individual's education level, which contradicts with OLS results. Surprisingly, we see that individuals who have higher levels of schooling eat less fruits and vegetables and they exercise less as well.

The results also suggest that males tend to smoke more, consume more alcohol, eat more fruits and vegetables and exercise more. We also observe that females have higher BMI levels in Turkey, which is consistent with the OLS results. Last, we observe that individuals in urban areas smoke more, consume more alcohol and eat

[52] It is important to note that in the previous sections, we also control for other variables such as household income and employment status.

more fruits and vegetables. However, the region where the individual lives does not significantly affect the individual's probability of exercising and his/her BMI level.

Regarding marital status, in the IV estimation setting, we find that married and single individuals consume less alcohol and they have less BMI levels compared to widowed/divorced individuals. The results also suggest that married people smoke less than widowed/divorced individuals, whereas the smoking behavior of single and widowed/divorced individuals do not significantly differ from each other.

7.6. Conclusion

This chapter investigates the determinants of health behaviors in Turkey in particular with respect to education. This chapter is the first study that analyzes the variations in health behaviors in Turkey for adult men and women in Turkey. The health behaviors considered are smoking, alcohol consumption, fruit and vegetable consumption, exercise and individual's BMI. We considered education as well as demographic factors, such as gender, age, the region where the individual lives (urban/rural), employment status of the individual and the household income.

In conclusion, education is found to be an important factor that could reduce the probability of risky health behaviors in Turkey. Probit results suggest that the probability of smoking increases with education. However, the relationship between tendency to smoke and additional years of schooling is concave, which implies that as education level rises, the probability of smoking decreases. Thus the results indicate that smoking is positively associated with education at all levels with a decreasing effect with the level of education. In fact IV estimation results also suggest that probability of smoking decreases as education level increases. In the LATE context, we can state that the individuals who are affected from the educational reforms tend to smoke less since they have higher levels of education. Therefore, although smoking is a serious public health problem in Turkey at all levels of education, individuals with higher levels of education are more aware of the dangers of smoking. In addition, we observe that higher educated individuals clearly eat more fruits and vegetables compared to less educated and illiterate We also find that higher educated individuals exercise more. However, IV estimation results destroy those findings and suggest that the prevalence of these behaviors decreases by education. In the LATE context, this result may be normal since IV estimands focus only on the individuals who are affected by those reforms. Next, OLS results suggest that higher educated individuals have BMI levels in the normal range compared to less educated. However, when we take the endogeneity problem into account, we observe that higher levels of education lead to higher BMI levels. We do not observe a negative relationship between alcohol consumption and education, as in the case of developed countries. Our results suggest that higher educated individuals tend to consume more alcohol in Turkey than the less educated. However, for the individuals who are affected by the education reforms and have higher years of schooling are found to consume less alcohol. As a conclusion, we can say that higher education may be a factor that heightens sensitivity towards adverse effects of risky behaviors. The contradicting results between Probit (OLS) and IV estimates can be explained by first the differences in control variables, second the volatility in health behaviors across different cohorts. In general, we can conclude that policy makers should pay more attention to increasing education levels.

Further, it is worthy to note that household income significantly increases alcohol consumption, fruit and vegetable consumption and BMI while income does not affect the probability of exercise. In addition, our results suggest that males tend to consume more tobacco and alcohol than females. They are also more likely to consume fruits and vegetables. Finally, in Turkey we observe that BMI of females is higher than that of males. The policy implications will be discussed in detail in Chapter Nine.

APPENDIX 7.A.1. IV Estimation Revisited

In this section, we re-estimate the effect of education on individual's health behaviors by including the trend controls, rather including the age dummies. The two stage regression model is written as:

$$E_i = \delta + X_i\, \alpha_1 + R61_i \alpha_2 + trend\ controls + \varepsilon_i \qquad (7.A.1.1)$$

$$H_i = \alpha + \hat{E}_i \beta_1 + X_i \beta_2 + trend\ controls + u_i \qquad (7.A.1.2)$$

The descriptions of dependent and explanatory variables are explained in section 7.5. Since our analysis covers a long period of time we control for birth year trends and their higher powers. Table 7.A.1 below presents the first stage results and the second stage results from this empirical specification.

Table 7.A.1. IV-Probit (2SLS for BMI) Estimation Results

Dependent Variable: Years of Schooling	**Smoking**	**Alcohol Consumption**	**Fruit and Vegetable Consumption**	**Exercise**	**BMI**
Sample: Individuals who were born in or later than 1952 are affected from educational reforms					
Policy Dummy	0.81*** (0.17)	0.82*** (0.16)	0.82*** (0.16)	0.84*** (0.17)	0.84*** (0.10)
Male	2.04*** (0.07)	2.04*** (0.06)	2.04*** (0.06)	2.01*** (0.06)	1.83*** (0.03)
Age (x10^{-1})	1.70*** (0.71)	2.41*** (1.05)	2.37*** (1.06)	2.67*** (1.12)	1.26 (0.99)
Age Squared (x10^{-3})	-1.20* (0.67)	-1.48*** (0.48)	-1.49*** (0.48)	-1.48*** (0.47)	-1.47*** (0.39)
Urban	2.24*** (0.04)	2.22*** (0.03)	2.22*** (0.03)	2.20*** (0.03)	2.17*** (0.04)
Trend	0.05 (0.07)	0.04 (0.10)	0.04 (0.10)	0.06 (0.10)	-0.03 (0.10)
Trend Square	-0.00006 (0.001)	-0.0008 (0.001)	0.0008 (0.001)	0.0008 (0.001)	-0.00008 (0.0008)
Trend Cube	0.00001 (0.000007)	0.000007 (0.000006)	0.000007 (0.000006)	0.000006 (0.000006)	0.00001 (0.000005)
Observations	34716	46493	46454	43605	41074

Table 7. A.1 (Continued)

R-squared	0.27	0.27	0.27	0.27	0.24
F-statistics	19.28	25.91	25.96	24.63	21.33
Second Stage Results					
Years of Schooling	-0.02 (0.04)	0.11* (0.06)	0.08** (0.04)	-0.03 (0.03)	0.26* (0.15)
Male	0.87*** (0.08)	0.74*** (0.16)	-0.12 (0.09)	0.29*** (0.07)	-1.12*** (0.35)
Age ($x10^{-1}$)	-0.09 (0.24)	2.93*** (0.44)	-9.49*** (0.53)	4.81*** (0.34)	5.50*** (0.57)
Age Squared ($x10^{-3}$)	-0.38 (0.24)	-0.90*** (0.25)	-0.84*** (0.18)	-0.80*** (0.13)	-3.22*** (0.86)
Urban	0.25*** (0.10)	0.001 (0.15)	-0.01 (0.10)	0.007 (0.07)	-0.38 (0.39)
Trend	-0.14*** (0.03)	0.15*** (0.04)	-1.09*** (0.06)	0.40*** (0.03)	0.11 (0.11)
Trend Square	0.002*** (0.0003)	0.0005 (0.0005)	-0.0001 (0.0003)	-0.0001 (0.0002)	0.001 (0.002)
Trend Cube	-0.00001*** (0.000001)	-0.000001 (0.000003)	0.000007*** (0.000001)	0.000002** (0.000001)	-0.00002**** (0.0000078)

Note:*** indicates 1% level of significance, **indicates 5% level of significance, *indicates 10% level of significance. Standard errors clustered by birth year are shown in parenthesis.

Table 7.A.1 shows that educational expansion in the early 1960s has positive and significant impact on the individual’s years of schooling (it leads to an increase in the years of schooling approximately by 0.82 units). Second stage results suggest that additional years of schooling does not have significant impact on probability of smoking and exercising. However, it leads to an increase in the prevalence of alcohol consumption and fruit and vegetable consumption. Second stage regression results also suggest that an increase in the years of schooling leads to an increase in the individual’s BMI level, which contradicts with the case of developed countries.

CHAPTER 8

QUANTILE REGRESSION ANALYSIS: THE RELATIONSHIP BETWEEN BMI AND EDUCATION IS REVISITED

8.1. Introduction

We examine the relationship between individual's body-mass index (BMI) level and education in Chapter Seven. Chapter Seven reveals that when we assume years of schooling is exogenous to individual's BMI level, there is a negative association between years of schooling and BMI. This is consistent with our expectations: Individuals with higher levels of education are more aware of the dangers of having high BMI levels relative to the less-educated. However, when we take the possible endogeneity between the individual's BMI and years of schooling, we have conflicting results: For instance, in Chapter Seven, we see that when we restrict the whole sample to middle aged and older individuals, the relationship between individual's BMI level and years of schooling becomes positive if the endogeneity problem between BMI and years of schooling are taken into account. This result is valid for the whole sample as well in the IV estimation setting. This finding contradicts with our expectations.

There can be several reasons for the contradicting results. For instance, the educational reforms may affect the older individuals more, hence there is positive relationship with years of schooling and BMI levels, as the individuals tend to have higher BMI levels when they become old. From econometric perspective, it is important to note that both Ordinary Least Squares (OLS) and two-Stage least squares (2SLS) estimation processes give the effect of covariates upon the mean.

This may overlook the fact that covariates are not homogeneous across the quantiles of the distribution of BMI. Therefore, in Chapter Eight, in order to see more robust relationship between education and BMI, we reinvestigate the determinants of BMI by using quantile regression (QR) techniques following Koenker and Basset (1978). By implementing the quantile regression methodology, we aim to see how education as well as other factors affect the individual's BMI level at different quantiles of the distribution of BMI.

Quantile regression is a popular methodology in health economics literature. For instance, Costa-Font et al. (2009) examine the BMI gaps among females in Spain and Italy by using quantile regression techniques. Garcia Villar and Quintana-Domeque (2009) investigate the association between household income (along with other factors) and BMI in nine European countries[53] by using quantile regression methods. In a different study Atella et al. (2008) examine the impact of BMI on wages across the different quantiles of the wage distribution for nine[54] countries using European Community Household Panel data for period 1998-2001. Section 8.2. provides brief literature review on applied studies that use quantile regression approach.

Next, in section 8.3, we present quantile regression results for the individuals who are 25 years old or older where we assume that years of schooling is exogenous to the individual's BMI level. Our objective in this section is to show how education affects individuals' BMI levels at different quantiles of BMI. In addition, we control for exogenous demographic factors, labor market indicators and household income and see the effects of those variables on different quantiles as well. Quantile regression results also suggest individual's BMI decreases with additional years of schooling. Moreover, the results indicate that females tend to be more obese than males and prevalence of obesity increases over time in Turkey.

[53] The countries they analze are: Austria, Belgium, Denmark, Finland, Greece, Ireland, Italy, Portugal and Spain

[54] The countries they examine are the same in Villar and Quintina-Domeque (2009).

In section 8.4, we revisit the endogeneity problem between BMI and education and estimate the effects of individual’s years of schooling together with other covariates by implementing Instrumental Variable Quantile Regression (IV-QR) methodology. IV-QR methodology is previously used by Brunello et al. (2009)[55] and Atella et al. (2008). It is a very new approach in microeconometrics, which is first suggested by Chernozhukov and Hansen (2006). Implementing user-written STATA command “ivqreg” is really hard, and most of the time we face with the problem “Convergence Not Achieved”. In order to solve this problem, I decide to enlarge the sample size. Therefore, I include the individuals who are between 18 and 24 to my analysis. The details of the new model will be discussed in Section 8.4. IV-QR models also suggest the negative association between individual's years of schooling and BMI level, and we observe that the highest impact of years of schooling is seen on the highest quantile distribution.

This chapter is new in the sense that it is the first that induces Quantile Regression approach to study conducted with a large micro data set in Turkey. It is interesting to see the impacts of different factors on different quantiles of BMI. In Section 8.5. we present the concluding remarks regarding QR and IV-QR methodology.

8.2. Literature Survey

In this section we examine the limited empirical studies on BMI that is done by implementing QR and IV-QR techniques. Costa-Font et al. (2009) is one of the important examples. The authors examine the association between individual’s BMI level and education using QR analysis. Costa-Font and his friends conduct cross-country comparison analysis between Italy and Spain, namely, two countries which have similar GDP and socioeconomic conditions. The authors implement the analysis by using National Survey on Daily Life for Italy and Spanish National Health Survey

[55] In fact Brunello et al. (2009) use Quantile Treatment Effect (QTE) Approach rather than (IV-QR) methodology. QTE Approach is proposed by Chernozhukov and Hansen (2005).

for Spain. In order to detect the cross country differences, the authors use Blinder-Oaxaca Decomposition. Their findings suggest that the defensive impact of education on BMI is highest upon highest quantiles of BMI. In addition, they also find that younger females in Spain are more likely to obese than their contemporaries in Italy. The authors attribute this result to cross country-differences.

Brunello et al. (2009) look at the relationship between BMI and years of schooling for both males and females in 13 different European countries. We investigate the published version of this study (Brunello et al. 2013) in detail in chapter six. In this working paper version, in the last part of their study they implement IV-Quantile Treatment Effects (IV-QTE) methodology in order to overcome the endogeneity problem between BMI and education level. Similar to their previous findings, IV-QTE results also suggest a negative relationship between individual's BMI and years of schooling.

Villar and Quintana-Domeque (2009) examine the association between individual's household income and BMI by using quantile regression methods for nine European Union countries[56]. In that study, they use the data from the European Community Household Panel (ECHP). The authors stress that OLS can give misleading results due to non-linear or non-monotonic relationship between BMI and household income. According to the authors, a non-linear relationship between BMI and household income may exist if the association is valid only for high BMI levels. They also point out that endogeneity problem can arise between them because of unobservable determinants of both BMI and household income, such as the individual's discount rates. In the study, they decompose household income to two components. The authors call these components as "own labor earnings" and "other household income". In the study, Villar and Quintana-Domeque also look at the effects of completed levels of education on BMI. They consider education as exogenous in the empirical analysis. Their results do not suggest a significant

[56] The countries they study are Austria, Belgium, Denmark, Finland, Greece, Ireland, Italy, Portugal and Spain.

relationship between household income and BMI for men, in contrast, they find a negative relationship between the two for women. The authors state that the difference for men and women mainly comes from the own-labor earnings for women.

There are some studies which use BMI as independent variable to explain the variations in different quantiles of wages. This kind of studies is relevant for this chapter in the sense that they are good guides for implementing quantile regression techniques and they are recent famous studies in labor economics literature. Atella et al. (2008) is one of the examples for these studies. In this study, the authors examine the association between obesity and wages using both QR and IV-QR techniques. For the study, they use the European Community Household Panel (ECHP) over the period 1998-2001 for 9 countries[57]. They focus on individuals who are between 25 and 64 years old. The authors stress that they apply quantile regression methodology so that they can detect the diverse effects of obesity at different points of the wage distribution. They also state that their objective is to see at what level of wages obesity represents a problem. Last, the authors point out that quantile regression analysis gives more relevant results since the usual OLS regression overlooks the effects of the heterogeneity across different countries and different wage quantiles.

In the study, Atella and his friends find that being obese negatively and significantly affect the wages of women across different wage quantiles. For men, they find that being obese results with lower wages at lower quantiles of wage distribution. However, obesity causes rise the men’s wages at higher wage quantiles. The authors stress that in the lowest points of the wage distribution workers perform manual activities that require effort and greater muscle mass. Similarly, at the highest points of the wage distribution intellectual activity is needed and obesity may not be a problem.

---

[57] They examine wage and obesity relationship for Denmark, Belgium, Ireland, Italy, Greece, Spain, Portugal, Austria and Finland.

In order to overcome the endogeneity problem between BMI and wages, Atella and his friends also apply quantile regression techniques together with IV approach. The authors use biological BMI as an instrument for the individual’s BMI. Biological BMI is calculated as the average BMI level of household members. However, the authors mention that they face with weak instrument problem during that analysis, since they find that F-statistics of excluded instruments are low.

Johar and Katayama (2012) also examine the effect of BMI on wages by using quantile regression techniques using National Longitudinal Survey of Youth data set for USA in which the individuals were interviewed first in 1979. They also find the incidence of wage penalty for females, in the sense that the negative effects of higher levels of BMI are seen on females in each quantiles of wages. The effect is even higher upon the highest distribution of wages. Like Atella et al. (2008), the authors also implement IV-QR techniques in order to overcome the possible endogeneity problem between wages and BMI. They use same-sex sibling’s body mass measure as instrument for BMI. Different from Atella et al., the authors also control for job-type in order to investigate the relationship between BMI and wages, as they think that the role of BMI may be more important for social jobs, where social skills are needed, such as physicians and teachers.

8.3. QR Methodology and Empirical Results

Brunello et al. (2009) state that just concentrating on the conditional mean effect may overlook the fact that covariates are not homogeneous across the quantiles of the distribution of BMI. Therefore, by applying quantile regression methodology, we are able to examine how these covariates affect the BMI at different quantiles of the distribution of BMI. Hence, with quantile regression methodology, we are able to see a more extensive view of the effects of the determinants on BMI. Apart from providing more general results, quantile regression has some other advantages over OLS. For instance, quantile regression is more robust to outliers than OLS. In addition, unlike OLS, we do not need to make strong stochastic assumptions in quantile regressions (See Cameron and Trivedi (2005), p. 85). Hence, quantile regression analysis provides stronger implications about the effects of demographic factors, labor market status, income and education on BMI. In order to have more efficient and consistent estimates we used bootsrapped techniques[58] in order to compute standard errors suggested by Koenker and Hallock (2001).

Based on the previous literature, we establish our model as follows:

$$Q_{BMI|X,E}(\tau|X,E)= \beta_0+ \beta_1(\tau)E_i+ \beta_2(\tau)X_i \qquad (8.3.1)$$

In equation (8.3.1), τ is the τ^{th} quantile, Q_{BMI} shows the body-mass index of the individual in the τ^{th} quantile. $\beta_1(\tau)$ shows the impact of education level on individual's BMI level, whereas $\beta_2(\tau)$ shows the effect of other covariates such as gender,age region, marital status, labor market status, household income and year dummies on the individual's BMI for the τ^{th} quantile.

Table 8.3.1 presents both OLS estimation and quantile regression estimation results. The quantile regression is estimated for the 5th,10th,25th, 50th,75th, 90th, and 95th quantiles of BMI distribution.

[58] We replicated the bootstrapped quantile regressions 400 times and get the robust results

Table 8.3.1. OLS and QR Estimation Results (Dependent Variable: BMI)

Variable	OLS	Q5	Q10	Q25	Q50	Q75	Q90	Q95
Education								
Years of Schooling	-0.10*** (0.02)	-0.03 (0.03)	-0.003 (0.03)	-0.05** (0.03)	-0.13*** (0.03)	-0.14*** (0.03)	-0.22*** (0.04)	-0.19*** (0.06)
Years of Schooling Square	-0.002 (0.001)	-0.002 (0.001)	-0.004** (0.001)	-0.002 (0.002)	0.0004 (0.001)	-0.0004 (0.001)	0.003 (0.002)	0.0004 (0.003)
Demographic Factors								
Male	-0.10* (0.05)	0.79*** (0.07)	0.83*** (0.07)	0.76*** (0.06)	0.38*** (0.06)	-0.34*** (0.09)	-1.27*** (0.12)	-1.71*** (0.19)
Age($x10^{-1}$)	4.54*** (0.10)	2.93*** (0.15)	3.21*** (0.15)	3.88*** (0.11)	4.42*** (0.11)	4.88*** (0.15)	5.34*** (0.21)	5.41*** (0.32)
Age Square ($x10^{-3}$)	-4.11*** (0.10)	-2.76*** (0.15)	-2.94*** (0.16)	-3.56*** (0.11)	-4.01*** (0.11)	-4.39*** (0.15)	-4.81*** (0.20)	-4.84*** (0.32)
Urban	0.30*** (0.05)	0.25*** (0.08)	0.30*** (0.07)	0.28*** (0.06)	0.26*** (0.06)	0.27*** (0.07)	0.16 (0.10)	0.25 (0.16)

Table 8.3.1. (Continued)

	Marital Status							
Married	-0.30*** (0.10)	0.06 (0.13)	0.12 (0.11)	-0.13 (0.10)	-0.19* (0.11)	-0.59*** (0.15)	-0.59*** (0.23)	-0.98** (0.28)
Single	-1.37*** (0.12)	-0.86*** (0.18)	-0.94*** (0.15)	-1.24*** (0.13)	-1.34*** (0.14)	-1.73*** (0.19)	-1.50*** (0.29)	-1.56*** (0.38)
	Labor Force Status							
Employed	-0.44*** (0.06)	-0.20*** (0.08)	-0.25*** (0.08)	-0.44*** (0.06)	-0.43*** (0.07)	-0.54*** (0.10)	-0.55*** (0.14)	-0.54*** (0.19)
Unemployed	-0.92*** (0.12)	-0.26 (0.21)	-0.31** (0.14)	-0.69*** (0.17)	-0.86*** (0.12)	-1.29*** (0.20)	-1.34*** (0.27)	-1.23*** (0.38)
Log Household Income	0.42*** (0.05)	0.50*** (0.07)	0.50*** (0.06)	0.49*** (0.05)	0.47*** (0.06)	0.35*** (0.07)	0.25** (0.11)	0.20* (0.12)
Dummy10	0.26*** (0.06)	0.10 (0.10)	0.13* (0.08)	0.16** (0.07)	0.17*** (0.08)	0.32*** (0.10)	0.37*** (0.12)	0.43*** (0.16)
Dummy12	0.29*** (0.06)	0.14 (0.09)	0.18*** (0.07)	0.24*** (0.06)	0.24*** (0.07)	0.28*** (0.09)	0.34*** (0.12)	0.48*** (0.16)
(Pseudo) R^2	0.11	0.05	0.06	0.06	0.06	0.06	0.08	0.08
N	40699	40699	40699	40699	40699	40699	40699	40699

Source:Authors' computations using 2008, 2010 and 2012 Turkish Health Survey

*** indicates 1% level of significance, **indicates 5% level of significance, *indicates 1% level of significanc

Table 8.3.1 suggests that the effect of individual's years of schooling is insignificant in 5th and 10th quantile. But, starting from the 25th quantile we see the negative impact of the individual's years of schooling on the individual's BMI. The effect rises systematically from 25th to 90th quantile. It decreases slightly in 95th quantile (from 0.22 to 0.19).

The gender effect on BMI is mixed. Quantile regression results suggest that at lower quantiles of BMI, males are more likely to have higher BMI levels. However, at the highest quantiles of BMI (75th, 90th and 95th quantile) of BMI, females have higher BMI levels. The gender effect increases in absolute values across the quantiles. This result indicates that prevalence of obesity and overweight is higher among females in Turkey.

The relationship between age and BMI level is concave as it was the case in the OLS results. The impact of age on BMI level increases monotonically from the lower to the upper quantiles. This implies that the occurrence of overweight and obesity raises with age.

The people who live in urban areas have significantly higher BMI levels than people who live in rural areas. The marginal effect of region increases up to 90th quantile. In the top two quantiles we observe that the effect of the region variable on individual's BMI becomes insignificant. This shows us that the prevalence of obesity do not differ between urban areas and the rural areas. The life style is more sedentary in urban areas and unhealthy food consumption is seen in urban areas more often. These two can lead to increase in obesity rates in urban areas. However, we should note that the consumption of foods with high calories such as white bread, biscuits, butter, sugar, honey and molasses is higher in rural areas[59]. Therefore, it is not surprising that the prevalence of obesity does not significantly differ between two regions.

[59] Source:Turkish Nutrition and Health Survey 2010.

The single people have significantly lower BMI levels relative to widowed/divorced individuals. Table 8.3.1 reveals that the effect of being single increases systematically in absolute value with the level of BMI, rising from 0.86 in 5^{th} quantile to 1.56 in 95^{th} quantile. This result can be explained by the age effect, since widowed/divorced people are more likely to be older they tend to have higher levels of BMI. Being married has no significant effect on BMI level in lower quantiles. Nevertheless, starting from the 50^{th} quantile married people have significantly lower BMI levels than widowed/divorced individuals. The effect of being married on BMI level raises monotonically from 50^{th} to 95^{th} quantile. However, the impact of being married is not as strong as that of being single on individual's BMI levels.

Regarding with labor force status, from Table 8.3.1 we see that being unemployed do not significantly affect the BMI level in the 5^{th} quantile. In the 10^{th} quantile, our results suggest that being unemployed leads to 31 per cent decrease in BMI level. The impact of being unemployed raises monotonically from 10^{th} quantile and it reaches its highest value (1.34 in absolute value) in the 90^{th} quantile. The impact of being employed increases systematically beginning from the 5^{th} quantile and the negative effect of being employed on BMI level is highest in the 90^{th} quantile (54 per cent). In short, our findings imply that both employed and unemployed people have significantly lower BMI levels than inactive people.

Table 8.3.1 reports that the logarithm of household income has positively significant effect on BMI levels in all quantiles, In general, our results suggest that people with higher income levels have higher BMI levels. This is an expected result since people can spend on various and healthy food more if their income is high.

Finally, we find both year dummies are positively significant across all quantiles. The impact of both 2010 dummy and the 2012 dummy increase monotonically from the lower to upper quantiles. Further, there has been a larger increase in the BMI in all of the quantiles in 2012 than in 2010, except the 90^{th} quantile. This result indicates that in Turkey, the prevalence of obesity have been increasing over time.

8.4. Quantile Regression Analysis with IV Methodology

In this section, we readdress the endogeneity problem between BMI and education. We implement quantile regression techniques together with IV methodology in order to see the effects of education across different quantiles of BMI. The model we use in this chapter is called as instrumental variable quantile regression (IV-QR).

IV-QR analysis is very new issue in microeconometrics literature. Angrist and Pischke (2009) argue that IV-QR is not flexible and convenient as 2SLS methodology. During the estimation process, I find out that STATA user command “ivqreg” is really difficult to implement and generally it gives the error “convergence not achieved”. In order to solve this problem, I decide to enlarge the sample a bit more and include the individuals whose ages vary between 18 and 24.

By including this sample of individuals to our model, we include another instrument for education variable to our analysis: The change in compulsory schooling laws in Turkey in 1997. In 1997, compulsory schooling increases from 5 to 8 years in Turkey. The law passed in 1997, and the secular government was very quick to implement the law in order to avoid the spread of the religious education, therefore the students who are in 4th grade in 1997 are subject to the change in compulsory schooling law. Hence, following Aydemir and Kırdar (2013), by considering school starting age as 6, we assume the individuals who are 10 years old in 1997, or in other words who were born in 1987 attend the compulsory schooling for 8 years. Hence, for the analysis, we define a new instrumental dummy, R97, which is equal to 1 if the individual was born in or later than 1987, 0 otherwise. The validity of this instrument is shown by Aydemir and Kırdar. We have one endogenous variable, the years of education and two excluded instruments, R61 and R97, which implies we make an overidentification.

The establishment of the IV-QR model is similar to Two Stage Least Squares (2SLS) model. We can write the general IV-QR model as follows:

$$Q_{BMI|X,E}(\tau|X,E)= \beta_0+ \beta_1(\tau)E_i+ \beta_2(\tau)X_i \qquad (8.4.1)$$

The Covariate Vectors X and E are defined in the same manner as Equation (8.3.1). This time, the vector E is considered as endogenous, and the vector X consists of education reforms in the early 1960s (R61) and the change in compulsory schooling laws in 1997 (R97), as the instrument for years of schooling. The definition of R61 is explained in Chapter six.

The estimation of (8.4.1) consists of two steps: In the first step, we regress $Q_{BMI|X,E}(\tau|X, E)$ on X and the instruments R61 and R97. In the second step, an estimate of the coefficient $\beta_1(\tau)$ is chosen so that $\beta_1(\tau)$ minimizes the absolute value of the coefficients of R61 and R97 in the first stage regression. Chernozhukov and Hansen (2005) suggest to use a grid search procedure to choose that level of $\beta_1(\tau)$. STATA 11.2 automatically runs the 200 grid search procedure and choose the level of $\beta_1(\tau)$ accordingly. Chernozhukov and Hansen (2005, 2006) call $\beta_1(\tau)$ as Quantile Treatment Effect (QTE).

In the IV estimation design we include cohort dummies defined in 5- year intervals in order to break the linearity between age, birth year and survey year. By including the cohort dummies into regression, our objective is to control for both age and birth cohort trends[60].

Following Staiger and Stock (1997), we test the validity of the overidentified instruments by implementing both Basmann and Sargan chi-square test, and we find that overidentification is valid at 5 per cent significance level. By enlarging the

[60] Unfortunately, The IV-QR estimates where we include age and trend controls seperately do not converge. In addition, IV-QR estimates are already robust results. We cannot manage to cluster the standard errors by birth year with IV-QR design.

sample and including another relevant instrument, We manage to make the grid-search process in IV-QR model and manage to have efficient and consistent estimates for both years of schooling and other covariates. The results of both 2SLS estimation and IV-QR estimation for 95th quantile (the most relevant quantile for the analysis) are presented in Table 8.4.1.

Table 8.4.1. 2SLS and IV-QR Results (Dependent Variable: BMI)

Variable	2SLS	Q95
Years of Schooling	-0.84*** (0.04)	-3.88* (2.28)
Male	0.81*** (0.06)	4.31* (2.14)
Urban	1.03*** (0.07)	5.13** (2.44)
Married	-0.10 (0.10)	-0.26 (0.62)
Single	-0.57*** (0.18)	2.93 (2.91)
Employed	0.39*** (0.08)	2.15 (2.65)
Unemployed	0.23* (0.13)	4.91 (3.11)
Log Household Income	2.42*** (0.13)	12.16** (6.44)
Dummy10	0.19*** (0.06)	-0.41 (0.81)
Dummy12	0.11* (0.06)	-1.79 (1.70)
Cohort Controls	Yes	Yes
N	46978	46978

Source:Authors' computations using 2008, 2010 and 2012 Turkish Health Survey

*** indicates 1% level of significance, **indicates 5% level of significance, *indicates 1% level of significance

The education variable is found to be negatively significant at 95th quantile. The results imply that one unit increase in education variable leads to 3.88 units decrease in the individual's BMI level. Therefore, we again conclude that individuals with higher levels of education are more careful about the dietary and exercise requirements in order to prevent obesity.

The education variable is not significant at lower quantile distributions of BMI. This may be because of the fact that we include individuals, who do not complete their education yet, in our sample. At the 95th quantile, we still observe that individuals with higher income levels tend to have higher BMI levels. We also see that males have higher BMI levels at the top quantile, which contradicts with our previous QR results. The IV-QR results for the 95th quantile also suggest that individuals who live in urban areas are more likely to be obese.

Most of the covariates loose its significance in IV-QR analysis. This may occur due to the design of our analysis. Implementing the IV-QR methodology and finding the accurate model is really hard. IV-QR analysis requires more attention from the researchers.

8.5. Conclusion

In Chapter Eight, first, we reinvestigate the determinants of BMI by using QR techniques in order to see how education as well as other factors affect the individual's BMI level at different quantiles of the distribution of BMI. QR results suggest individual's BMI decreases with additional years of schooling. The negative effect of education increases monotonically from 25th to 90th quantile. The negative effect of years of schooling is still very high in the top BMI quantile. Therefore, we can conclude that people with higher levels of education are more aware of the dangers of obesity. Quantile regression results also suggest that females tend to be more obese than males and prevalance of obesity increases over time in Turkey. Hence, QR results also indicate that obesity is becoming a serious health problem in Turkey over time like in developed and developing countries. To prevent the severe health consequences of obesity, we have some policy suggestions. The suggestions are discussed in chapter eight. Regarding age, QR results suggest that old individuals are more likely to be obese or overweight relative to younger individuals. In addition, we find that the occurrence of obesity do not differ between urban areas and the rural areas. Our findings also imply that both employed and unemployed people have significantly lower BMI levels than the individuals who are out of labor force. Finally QR results suggest that people with higher household income have higher BMI levels.

Next, in section 8.4, we take the endogeneity problem into account and implement IV-QR. In that part, we include the individuals who are between 18-24 years old and include another instrument which controls for whether the individual is affected from the compulsory schooling law change in 1997. By enlarging the sample and include another relevant instrument to our analysis, we manage to make the grid search process in a more convenient manner, and suppose to get relevant results for the analysis. However, we observe that most of the coefficients we use lose their significance in IV-QR model and we find the education variable significant at the highest quantile. Neverheless, IV-QR is newly developed process and it requires more attention for further research.

CHAPTER 9

CONCLUSION

9.1. Summary of the Main Findings and Conclusions

The objective of this thesis is to examine the relationship between the two components of human capital, health and education for a middle income, developing country, Turkey. We first analyze the factors that affect the individual's health and discuss why these factors may be valid determinants to explain the variations in health outcomes. Second, we examine the effect of education on individual's Self-Assessed Health Status (SAH) empirically first by considering education as exogenous to the individual's health and next, we assume that education is endogenous to the individual's SAH. We examine the effect of other covariates on the individual's SAH as well. Third, we examine the effect of education on various health behaviors, such as smoking, alcohol consumption, fruits and vegetable consumption, exercising and BMI. We also consider the effects of the other covariates as well when we investigate the association between health behaviors and education. Last, we specifically focus on the rising health problem both in the world and in Turkey, namely, obesity, and we examine the effects of various health behavior determining factors across the quantiles of the distribution of individual's BMI level.

This thesis contributes to the literature in several ways. First, it is the first study which examines the association between health and education for adult men and women in Turkey by considering the several components of health. Second, the empirical analyses in this thesis are implemented by using a new and rich micro data set, the Turkish Health Survey (THS) data set. It consists of a large set of observations for adult men and women in Turkey. Third, several microeconometric techniques are used in this thesis. Among these techniques the Conditional Mixed Process (CMP) and IV-Quantile Regression (IV-QR) are used for the first time for a health related research conducted on Turkey. Fourth, this thesis uses the instrumental variable (IV) methodology which is a popular strategy in order to overcome the problems that result from possible endogeneity between health and education. Unlike the previous studies on Turkey, we introduce a new instrument to the literature, namely, the education reforms that took place in the early 1960s and after the 1960 Constitution has been implemented. This thesis is the first study that uses the 1960 educational reforms in Turkey as an instrument for education. This is important for the analysis, since we are interested in the causal association between health and education for adults. Owing to these contributions, we believe that this thesis shed light on further studies regarding both health and education in Turkey.

Following the famous studies on human capital theory (Becker, 1965) and Schultz (1961), Grossman (1972) introduces a new theory into the literature: “The Health Demand Theory”. Many empirical models today are still based on Grossman’s Health Demand Theory in the health economics literature. Grossman is one of the first economists who explains the variations in health by education. In Chapter Two, we briefly explain the implications of the Grossman’s model. We also discuss recent models that extend Grossman’s model. Finally, based on these models we try to explain why we use the specific covariates in order to examine the variations in health outcomes in our empirical analysis in the following chapters.

Next, in Chapter Three we introduce the THS data set. The THS is beneficial is useful in order to examine the health related issues in Turkey. It contains a large set of observations from nationwide randomly chosen households. It includes several health related questions related to health outcomes that are used in this thesis. In addition there are questions about the psychology of the respondents, how often a respondent makes use of preventative health care services, or whether the respondents had an accident recently, etc…. We use THS data set for the years 2008, 2010 and 2012 and we pool these three data sets for the empirical analysis. We argue that the most important shortcoming of THS data is the fact that the answers to all the questions are self-reported, including the answers for questions of objective health criteria.

Unfortunately, we cannot observe some other important health outcomes such as mortality and life-expectancy of an individual in the survey. Therefore, one of the health outcomes we use in this thesis is the individual's Self-Assessed Health (SAH) status. It is proved to be a reliable health outcome and a good predictor of mortality and morbidity several times in the literature (Idler and Benyamini, 1997). We also discuss the advantages of using SAH as a health outcome in Chapter Four. Chapter Four also provides a frequency distribution of variations in SAH in Turkey in our data set. We examine the variations in SAH in Turkey by education, gender, age, marital status, region, household income and employment status. Frequency distribution analysis indicates that education and household income level is positively correlated with individual's SAH. Males tend to report better SAH levels than females. In addition, we observe that young individuals are more likely to report good SAH. Our results also suggest that individuals living in urban areas report good SAH more often than people living in rural areas. We also find that the percentage of reporting poor SAH is the highest among widowed/divorced females in Turkey. Lastly, regarding labor force status, we find that inactive individuals report poor health more often than the employed and unemployed individuals.

In Chapter Five, we describe the econometric techniques that we use in chapters six and seven. We introduce the IV methodology and discuss the validity of the instrument. We use the educational expansion in the early 1960s in Turkey as the instrument for years of schooling. This is a unique instrumental variable that is being used for the first time in any study on Turkey. We establish validity of this instrument for years of schooling with extensive investigations of the several educational outcomes in the 1950s and 1960s. Chapter Five includes an elaborate discussion of these results. We also investigate health reforms in the 1960s in Chapter Five. We conclude that health reforms are not as effective as the education reforms, therefore, they do not coincide with each other.

In Chapter Six, we investigate the variations in SAH empirically. First, we assume that education is exogenous to individual's SAH and implement ordered probit and ordered logit regressions. We find positive association between health and education as in the case of developed countries. Then, we extend the models by including labor market indicators and household income as determinants of the variations in individual's SAH. We observe that both the employed and unemployed people tend to report better SAH than inactive individuals. We also find that household income is positively correlated with individual's SAH. These results are in line with results of the descriptive analysis. Then, we move to IV analysis so that we can overcome the problems related to endogeneity between health and education. Since the dependent variable (SAH) is a categorical variable, we establish an "IVOprobit" (Instrumental Variable Ordered Probit) model and estimate it with Conditional Mixed Process (CMP) techniques as suggested by Roodman (2011). CMP is used in an empirical analysis for the first time in Turkey. The first stage results indicates that the instrument is strongly valid and education reforms significantly increase the individual's years of schooling. Our second stage regression results indicate that an additional years of schooling positively affect individual's SAH.

In Chapter Seven, we investigate the determinants of health behaviors in Turkey paying particular attention to education. The health behaviors considered in this chapter are smoking, alcohol consumption, fruit and vegetable consumption, exercise and individual's BMI. In general, our results indicate that education is an important factor that reduces the probability of risky health behaviors in Turkey. When we consider years of schooling as exogenous to the individual's smoking behavior, we find an inverted U-shaped relationship between smoking and education. This implies that the probability of smoking increases up to certain education degree (middle school) but it decreases with higher levels of education. When we re-estimate the effects of education on the probability of smoking by assuming that years of schooling is endogenous, we find a preventative effect of education on smoking, that is higher education levels reduce smoking.

In the probit model estimations, we do not observe a negative relationship between alcohol consumption and education in Turkey which is observed in the case of developed countries. Our results suggest that higher educated individuals consume higher levels of alcohol. We believe that this result is mainly due to the design of the sample, as the sample refers both heavy and the occasional drinkers as alcohol consumers. Nevertheless, when we take years of schooling as endogenous to alcohol consumption, we observe that probability of alcohol consumption decreases by education as expected and as we also observe in the case of developed countries.

In the probit model estimation results, we observe that individuals with more years of schooling clearly eat more fruits and vegetables compared to less educated and illiterate individuals. We also see that probability of exercising rises as the individual's education level rises. Next, OLS results suggest that the higher educated individuals have BMI levels in the normal range compared to the less educated. However, when we take the endogeneity between BMI and years of schooling into account, we observe that BMI level increases by education contrary to the OLS results and our expectations.

Chapter Eight is devoted to a reexamination of the determinants of BMI by using QR techniques in order to see how education as well as other factors affect the individual's BMI level at different points on the conditional distribution of the BMI. Similar to the previous chapters, we first test the effect of education on individual's BMI by assuming that schooling is exogenous to BMI. We find that individual's BMI decreases with additional years of schooling. In addition, we see that the effect of education increases monotonically across different quantiles in absolute value.

Therefore, we can conclude that higher levels of education reduces BMI and have preventative effects on being overweight or obese.

In addition, we observe an increasing trend in obesity in Turkey in recent years as in the case of developed countries. Therefore, obesity is a rising health problem in Turkey as well. Another interesting finding is that the prevalence of obesity is higher among adult females than among adult males in Turkey. Our results suggest individual’s BMI level deteriorates with age. Our findings also imply that both the employed and the unemployed people have significantly lower BMI levels than the inactive individuals. Lastly, we find that household income is positively related to the individual’s BMI level. In fact, it is an expected result since individuals with higher levels of household income are more capable of reaching various kinds of nutritious food.

Next, in Chapter Eight, we further take the endogeneity problem into account and conduct the QR analysis together with the IV methodology called “Instrumental Variable Quantile Regression” (IV-QR). IV-QR is a complicated and newly developed approach in microeconometrics. Unfortunately, we cannot have convergent estimates with the IV-QR methodology in our sample of individuals over 25 years of age. Therefore, we enlarge our sample a bit and we include the individuals who are between 18-24 years old. In addition, for the IV process, we include another instrument which controls for whether the individual is affected from the recent compulsory schooling laws in Turkey enacted in 1997. By enlarging the sample and including another relevant instrument into our analysis, we manage to make the grid search process reach convergence and produce relevant results for our analysis. In these results of the IV-QR estimation most of the coefficients lose their significance. However, we still find that the education variable is significant at the upper quantiles of BMI. Therefore, IV-QR results also indicate that higher educated individuals pay more attention to the dangers of having high BMI levels.

In conclusion, our results definitely suggest that the individual's education level is positively correlated with his/her SAH. As SAH is a good predictor of the individual's mortality or morbidity, we can state that educated people report better SAH status. They may be more aware of the preventative health care services, the dangers of risky health behaviors, the benefits of healthy nutrition and exercising. We can also attribute this result to the more rational behaviors of educated individuals relative to the less educated. They are also rational in the sense that they are able to think of the environmental conditions, the health status of the other individuals, and respond to the health related questions in a more sensible manner. Regarding the variations in health behaviors some of our results are conflicting depending on whether education is assumed exogenous or endogenous to the health behaviors. In general, we can say that higher education may be a factor that heightens sensitivity towards adverse effects of risky behaviors except that of alcohol. Regarding the two other health behaviors, fruits and vegetable consumption and exercising we also observe the protective effect of education. Thus, policy makers should pay more attention to increasing education levels in Turkey.

9.2. Policy Implications

There are only a few recent studies on health expenditures and health reforms in Turkey. Two recent studies review the health reforms (Kol, 2015) and improvement in health expenditures (Daştan and Çetinkaya, 2015) in Turkey but do not suggest solid policy implications. In the light of the results we find in this thesis, we try to suggest some policy implications in order to increase the general health status of the society. First, we will briefly review the recent studies on health reforms and health expenditures in Turkey. Next, stressing the importance of individual's education level on health, we will propose some policies and discuss how expenditures related to both health and education can be spent in a more effective way.

Kol (2015) examines the health reforms in Turkey in the light of “The Rights on Health”. According to World Health Organization (WTO), health is the right of everyone and the governments are responsible for people’s health. Kol briefly describes the health related laws in both the 1960 and the 1982 Constitutions and discuss some health reforms that took place in the 1990s. In a recent work, Daştan and Çetinkaya (2015) provide some statistics for the health system, health expenditures and health outcomes (such as infant mortality rates and life expectancy rates) in Turkey and compare them to those of the USA and OECD countries. The authors show that health expenditures per capita was 69 dollars in Turkey in 1980. Turkey had the lowest health expenditures among all the OECD countries and the USA. Although, the health expenditures per capita increase to 984 dollars in Turkey in 2012, it is still very low compared to the OECD average, which is 3500 dollars. However, it is important to note that the OECD per capita health expenditures on average is ten times higher than those of Turkey’s in 1980. However, in 2012 the average is 3 times as high as that of Turkey’s. This is in fact very important implication: Expenditures on health increased in Turkey during the last three decades, but they are still low compared to other OECD countries.

The thesis has a very important conclusion: Education has strong and significant impact on the individuals’ health in Turkey. To our knowledge, there is no previous study that suggests education related policies in order to improve the general health status of the society. The health behaviors used as health outcomes in this thesis are very relevant for future policies. Our results provide evidence that the causality goes from education to health in Turkey. Therefore, we suggest that first the expenditures on education should be increased in Turkey, as it leads to improvement in the individual’s health. Next, optimal amount of public health spending should be determined by policy makers.

Obesity is an increasing health problem in Turkey. Another conclusion of this thesis is that having high levels of education has protective effect on the individual’s BMI level. Therefore, first, health lectures in high schools should be implemented examining the consequences of obesity deeply. In schools, children should be informed about healthy nutrition processes. The consumption of the foods with high calorie such as fast-food is widespread among children. Both the schools and families should be concerned about it. Government can increase the public health expenditures by providing seminars to adults on the dangers of being obese. In some specifications, we see that individuals are likely to be overweight or obese also in the rural areas. Especially, physicians and teachers in rural areas should inform the public about the risks of consuming fat, sugar, salt and flour in large amounts. Recent medical literature warns against consumption of such foods. We discuss in Chapter Seven that, people living in rural areas tend to gain weight due to consuming such foods in large amounts. Providing seminars and campaigns on the risks of obesity in both urban and rural areas are necessary. The literature on developed countries argues that education related policies on obesity are not very successful. It will take a long time for such policies to succeed in a developing country, but at least children may grow up more informed if the health education in schools is more effective.

Apart from education, several precautions to prevent the widespread obesity can be considered. First, there are lots of fast-food restaurants with different brands, especially in urban areas. Quotas can be put on the number and operation of fast-food restaurants. For instance, every brand may have just one restaurant in certain regions. Fast-food restaurants are cheap especially for students. Government can tax fast-food consumption at a higher rate. Convenience foods which are sold especially in super markets are also easy to make and cheap. However, they also provide high calories to the body. The tax on these foods should also be increased.

Our results also suggest that education has positive impact on fruits and vegetables consumption and exercising. These health behaviors are also important for the individuals so that they can have normal ranges of BMI. Fruits and vegetable consumption should also be promoted in schools in the early years of education. Government can increase its spending on the advertisements related to the importance of these two health behaviors. Especially in high inflationary periods or due to seasonal factors various fruits and vegetables are very expensive. It is better to put price ceiling at least in certain foods so that the individuals can benefit from consuming them. The sports centers are generally operated by private sector in Turkey. Therefore, everyone cannot have chance to exercise regularly. Therefore, the sports facilities in schools should be increased, and they should be free or very cheap for the young. For the old, the government can establish new social sports centers so that the individuals can cheaply exercise.

In Turkey, in fact the government tries to reduce the prevalence of smoking and alcohol consumption. For instance, the taxes on cigarettes and alcoholic drinks are high. There are regular advertisements on televisions regarding the dangers of smoking. Smoking is prohibited in certain open areas since the beginning of the 2010s, and the prohibition will increase even more[61]. In addition, since 2013 individuals are not allowed to purchase alcohol between 10:00 p.m and 6:00 a.m. In our opinion, rather than prohibiting alcohol or smoking, the government should inform the individuals about the risks of heavy cigarette and alcohol consumption. The prevalence of smoking is high among teenagers in Turkey. The health lectures should be held in schools and the students should be informed about the dangers of such risky behaviors in a confident way.

Lastly, we propose that in addition to the improvements in education system to improve the health conditions of people in Turkey, more attention should be given to Non-Governmental Organizations (NGOs) and the media while applying policies related to health. The validity of the education gradient of health in Turkey will be useful in discussing how education and health related policies can be improved and, which precautionary policies can be applied in order to increase the general health status of the society. The policy implications require further research.

[61] Source: Milliyet NewsPaper 25.07.2014. The news is entitled as “Sigara Yasağında Yeni Yol Haritası”

9.3. Further Research

This thesis has an important contribution to the health economics literature that it in general proves the validity of the education gradient of health for the adult individuals in Turkey. In order to solve the problems that result from the endogeneity between health behaviors and education, we use an IV approach. In this procedure we use the educational reforms implemented after the 1960 Constitution. This is a good instrument for the analysis, however Turkish education statistics suggest that compulsory schooling law change in 1997 is a more relevant education reform as an instrumental variable because it forces lots of students to pursue at least middle school degree. Therefore, collecting the THS data for the coming years in which the individuals who are affected from 1997 compulsory schooling law change become older, and testing the validity of the effects of education on those individuals' health behaviors will strengthen our findings.

The THS data set has some limitations. First, we cannot observe the individual's household characteristics, except the individual's household income. For instance, we cannot observe the size of the household, the parents' education, the parents' job etc...In fact, family background variables are important determinants of the individuals' health and health behaviors. In addition, we cannot observe the individual's own wage rate. We cannot observe the region that the individual lives in. Therefore, in a more extensive view we cannot control for region specific variables. Having knowledge on these variables would also help us to establish new instruments in order to control for the endogeneity between the household income, labor market indicators and health. For instance, for earnings and GDP per capita by regions may be good instruments. Similarly, unemployment level by regions may be a relevant instrument. Lastly, the most important shortcoming of the THS data set is the fact that it does not come from a longitudinal survey, hence we cannot observe the individual's health status and health behaviors over time. To study with a longitudinal data set which includes the variables discussed above will give more relevant results, and will be a better contribution not only to the health economics but also to the labor economics literature in Turkey.

In any case, THS is an informative data set to use in the sense that it includes several questions regarding adult individual's health. In this thesis, we just use the SAH and health behaviors as individuals' health outcomes. In fact the responses related to objective health status, preventative health care services, psychology of the individuals and frequency of traffic accidents are also interesting to examine and they can also be used as health outcomes to examine the association between health and education. The relationship between several health outcomes can also be tested by using proper Econometric techniques.

Finally, such studies are very limited on developing countries. Similar data sets from several developing countries can be combined. It will be interesting to examine the validity of education gradient of health across different countries. In such study, we can see whether the cultural norms, climate, traditions, religion or race have impact on individuals' health as well. We can test the power of the impact of education on health by controlling those factors. This kind of study allows us to see whether the impact of education on health weakens or strengthens when we control for the country specific factors.

REFERENCES

Adams, S.J. (2002) "Educational Attainment and Health: Evidence from A Sample of Older Adults." *Education Economics* 10 (1): 97-109.

Akyüz,Y. (1999) *Türk Eğitim Tarihi (Başlangıçtan 1999'a)*. (The History of Turkish Education from the Beginning till 1999). Alfa Yayınevi. İstanbul.

Angrist, J.D. and A.B. Krueger (1991) "Does Compulsory School Attendance Affect Schooling and Earnings?." *Quarterly Journal of Economics* 106 (4): 979-1014.

Angrist, J.D., G.W. Imbens, and D. B. Rubin (1996) "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444-455.

Arendt, J. N. (2005) "Does Education Cause Better Health? A Panel Data Analysis Using School Reforms for Identification." *Economics of Education Review* 24(2): 149-160.

Atella, V., N. Pace, and D. Vuri (2008) "Are Employers Discriminating with respect to Weight?: European Evidence Using Quantile Regression." *Economics and Human Biology* 6 (3): 305-329.

Austad, S. N. (2006) "Why Women Live Longer Than Men: Sex Differences in Longevity." *Gender Medicine* 3 (2) : 79-92.

Aydemir, A., and M. Kırdar (2013) "Estimates of the Return to Schooling in a Developing Country: Evidence from a Major Policy Reform in Turkey." *MPRA Working Paper* No.51938. Munich.

Baker, M., M. Stabile and C. Deri (2004) “Why do Self-Reported, Objective, Measures of Health Measure?” *The Journal of Human Resources* 39 (4):1067-1093.

Becker, G. S. (1965) "A Theory of the Allocation of Time." *The Economic Journal* 75 (299): 493-517.

Becker, G.S. (1993) "Nobel Lecture: The Economic Way of Looking at Behavior." *Journal of Political Economy* 101(3): 385-409.

Benyamini, Y., E. A. Leventhal and H. Leventhal (1999) “Self-Assessments of Health: What do People Know That Predicts Their Mortality?” *Research on Aging* 21(3): 477-500.

Berger, M.C., and J.P. Leigh (1989) "Schooling, Self-selection, and Health." *Journal of Human Resources* 24 (3): 433-455.

Bobak, M., H.Pikhart, C. Hertzman, R. Rose and M. Marmot (1998) “Socioeconomic Factors, Perceived Control and Self-Reported Health in Russia. A Cross-Sectional Survey.” *Social Science in Medicine* 47 (2): 269-279.

Bolin, K. "Health Production." *Oxford Handbook of Health Economics.* Ed. S. Glied and P.C. Smith. New York: Oxford University Press, 2011. 95-123.

Bound, J. (1991) “Self- Reported versus Objective Measures of Health in Retirement Models.” *The Journal of Human Resources,* 26 (1): 106-138.

Bound, J., D. A. Jaeger, and R. M. Baker (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 (430): 443-450.

Böckerman, P., and P. Ilmakunnas (2009) "Unemployment and Self-Assessed Health: Evidence from Panel Data." *Health economics* 18 (2): 161-179.

Braakmann, N. (2011) "The Causal Relationship between Education, Health and Health Related Behaviour: Evidence from a Natural Experiment in England." *Journal of Health Economics* 30 (4): 753-763.

Brunello,G., D. Fabbri and M. Fort (2009) “Years of Schooling, Human Capital and the Body Mass Index of European Females”, *IZA Discussion Paper* No.4667. Bonn.

Brunello, G., D. Fabbri, and M. Fort (2013) "The Causal Effect of Education on Body Mass: Evidence from Europe," *Journal of Labor Economics* 31(1): 195-223.

Brunello, G., M. Fort, N. Schneeweis and R. Winter-Ebmer (2015). "The Causal Effect of Education on Health: What is the Role of Health Behaviors?." *Health Economics. (Published online DOI: 10.1002/hec.3141)*

Bulut, A. “Türkiye’de Sağlık Reformunun Tarihçesi (The History of Health Reform in Turkey)” *Avrupa’da ve Türkiye’de Sağlık Politikaları (Health Policies in Europe and Turkey)*. Ed. Ç. Keyder, N. Üstündağ, T. Ağartan and Ç. Yoltar. İstanbul: İletişim Yayınları, 2007. 111-124.

Burstörm, B. and P. Fredlund (2001) “Self Rated Health: Is it As Good Predictor of Subsequent Mortality among Adults in Lower As Well As in Higher Social Classes?” *Journal of Epidemiology and Community Health* 55 (11):836-840.

Butler, J.S., R.V. Burkhauser, J. M. Mitchell and T. P. Pincus (1987) “Measurement Error in Self-Reported Health Variables” *The Review of Economics and Statistics* 69 (4):644-650.

Cameron, C.A., and P.K. Trivedi. (2005) *Microeconometrics: Methods and Applications*. Cambridge University Press. New York.

Cameron, C.A. and P.K. Trivedi. (2010) *Microeconometrics Using Stata*. Stata Press Publication. Lakeway Drive College Station. Texas.

Card, D. “Using Geographic Variation in College Proximity to Estimate the Return to Schooling.” *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*. Ed.L. N. Christofides, E. K. Grant and R. Swidinsky. Toronto: University of Toronto Press, 1995. 201-222.

Case, A. and C. Paxson (2005) “Sex Differences in Morbidity and Mortality” *Demography* 42 (2): 189-214.

Cawley, J., and C. J. Ruhm. "The Economics of Risky Health Behaviors" *Handbook of Health Economics Vol. 2*. Ed. M.V. Pauly, T.C. McGuire, P.P. Barros. Amsterdam: Elsevier. 2012. 97-153.

Cesur, R., B. Dursun, and N. Mocan (2014) *“The Impact of Education on Health and Health Behavior in a Middle-Income, Low-Education Country”*. National Bureau of Economic Research (NBER) Working Paper No. 20764. Cambridge, MA.

Chaloupka, F. J., and K. E. Warner "The Economics of Smoking." *Handbook of Health Economics Vol.1*. Ed. A.J. Culyer and J.P.Newhouse. Amsterdam: Elsevier. 2000. 1539-1627.

Chernozhukov, V., and C. Hansen (2005) "An IV Model of Quantile Treatment Effects." *Econometrica* 73 (1) : 245-261.

Chernozhukov, V., and C. Hansen (2006) "Instrumental Quantile Regression Inference for Structural and Treatment Effect Models." *Journal of Econometrics* 132 (2): 491-525.

Chou, S. Y., M. Grossman, and H. Saffer (2004) "An Economic Analysis of Adult Obesity: Results from the Behavioral Risk Factor Surveillance System." *Journal of Health Economics* 23 (3): 565-587.

Conti, G., J. Heckman and S. Urzua (2010) “The Education-Health Gradient.” *American Economic Review* 100 (2): 234-238.

Conti, G., and C. Hansman (2013). "Personality and the Education–Health Gradient: A Note on “Understanding Differences in Health Behaviors by Education”." *Journal of Health Economics* 32 (2): 480-485.

Costa-Font, J., D. Fabbri, and J. Gil (2009) "Decomposing Body Mass Index Gaps between Mediterranean Countries: A Counterfactual Quantile Regression Analysis." *Economics and Human Biology* 7 (3): 351-365.

Crossley, T. F. and S. Kennedy (2001) “The Reliability of Self-Assessed Health Status” *Journal of Health Economics* 21 (4): 643-658.

Cutler, D.M., and A. Lleras-Muney. (2008) “Education and Health: Evaluating Theories and Evidence”, In: James, S.H., Schoeni, R.F., Kaplan, G.A., Pollack, H. (Eds), Making Americans Healthier: Social and Economic Policy as Health Policy. Russel Sage Foundation, New York.

Cutler, D. M. and A. Lleras-Muney (2010) “Understanding Differences in Health Behaviors by Education”, *Journal of Health Economics* 29(1):1-28.

Cutler, D. M., A. Lleras-Muney, and T. Vogl. “Socioeconomic Status and Health: Dimensions and Mechanisms”. *Oxford Handbook of Health Economics. Ed. S. Glied and P.C. Smith. New York: Oxford University Press,* 2011. 124-163.

Cutler, D. M., E.L. Glaeser, and J.M. Shapiro (2003) "Why Have Americans Become More Obese?," *The Journal of Economic Perspectives* 17 (3): 93-118.

Cutler, D. M., and E. L. Glaeser (2005) "What Explains Differences in Smoking, Drinking, and Other Health-Related Behaviors?." *The American Economic Review* 95(2): 238-242.

Daştan, İ and V. Çetinkaya (2015) “OECD Ülkeleri ve Türkiye'nin Sağlık Sistemleri, Sağlık Harcamaları ve Sağlık Göstergeleri Karşılaştırması” (Comparing Health Systems, Health Expenditures and Health Indicators in OECD Countries and Turkey) *Sosyal Güvenlik Dergisi – SGD* 5 (1): 104-134.

De Walque, D. (2007) "Does Education Affect Smoking Behaviors?: Evidence Using the Vietnam Draft as an Instrument for College Education." *Journal of Health Economics* 26 (5): 877-895.

Eide, E.R. and M.H. Showalter (2011) “Estimating the Relation between Health and Education: What Do We Know and What Do We Need to Know?” *Economics of Education Review* 30 (5):778-791.

Erbaydar, T., et al. (2005) "Influence of Social Environment in Smoking among Adolescents in Turkey," *The European Journal of Public Health* 15(4): 404-410.

Erdoğan, İ. (2003) *Yeni bir Binyıla Doğru Türk Eğitim Sistemi Sorunlar ve Çözümler* (Towards the New Millennium the Turkish Educational System: Problems and Solutions). Sistem Yayıncılık. İstanbul.

Erdoğan, S., D.C. Yıldırım, and O. Tosuner (2012) "The Effect of Gender Inequality in Education on Health: Evidence from Turkey." *Educational Sciences: Theory and Practice* 12 (3): 1860-1866.

Erem, C., et al. (2004) "Prevalence of Obesity and Associated Risk Factors in a Turkish Population (Trabzon City, Turkey)," *Obesity Research* 12 (7): 1117-1127.

Ettner, S. L. (1996) "New Evidence on the Relationship between Income and Health," *Journal of Health Economics* 15(1): 67-85.

Farmer, M.M. and K.F. Ferraro (1997) “Distress and Perceived Health: Mechanisms of Health Decline,” *Journal of Health and Social Behavior* 39 (3): 298-311.

Ferraro, K.F., M.M. Farmer and J.A. Wybraniec (1997) “Health Trajectories: Long-Term Dynamics Among Black and White Adults” *Journal of Health and Social Behavior* 38(1): 38-54.

Fletcher, G.F., et al. (1996) "Statement on Exercise: Benefits and Recommendations for Physical Activity Programs for All Americans A Statement for Health

Professionals by the Committee on Exercise and Cardiac Rehabilitation of the Council on Clinical Cardiology, American Heart Association," *Circulation* 94(4): 857-862.

Fletcher, J.M., and D.E. Frisvold (2009) "Higher Education and Health Investments: Does More Schooling Affect Preventive Health Care Use?." *Journal of Human Capital* 3 (2): 144-176.

Fonseca, R. and Y. Zheng. (2011) “The Effect of Education on Health Cross-Country Evidence,” *RAND Labor and Population Working Pap*er WR-864. Santa Monica.

Fuchs, V. R. (2004) "Reflections on the Socio-Economic Correlates of Health" *Journal of Health Economics* 23(4): 653-661.

Gaziano, J. M., et al. (1995) "A Prospective Study of Consumption of Carotenoids in Fruits and Vegetables and Decreased Cardiovascular Mortality in the Elderly." *Annals of Epidemiology* 5 (4): 255-260.

Glewwe, P. (2002) "Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes." *Journal of Economic Lliterature* 40 (2): 436-482.

Grabner, M. (2009) "The Causal Effect of Education on Obesity: Evidence from Compulsory Schooling Laws." (Published Online) *Available at SSRN 1505075*.

Grimard, F., and D. Parent (2007) "Education and Smoking: Were Vietnam War Draft Avoiders Also More Likely to Avoid Smoking?." *Journal of Health Economics* 26 (5): 896-926.

Grossman, M. (1972) “On the Concept of Health Capital and The Demand for Health,” *Journal of Political Economy* 80 (2): 223-255.

Grossman, M., "The Human Capital Model." *Handbook of Health Economics Vol.1. Ed. A.J. Culyer and J.P.Newhouse. Amsterdam: Elsevier,* 2000. 348-408.

Grossman, M., (2004) "The Demand for Health, 30 Years Later: A Very Personal Retrospective and Prospective Reflection." *Journal of Health Economics* 23 (4): 629-636.

Grossman, M. (2008) “The Relationship between Health and Schooling”, *Eastern Economic Journal* 34 (3): 281-292.

Gruber, J., and M. Frakes (2006) "Does Falling Smoking Lead to Rising Obesity?," *Journal of Health Economics* 25 (2): 183-197.

Haan, P. and M. Myck (2009) “Dynamics of Health and Labor Market Risks”, *Journal of Health Economics* 28 (6): 1116-1125.

Harmon, C., and I. Walker (1995) "Estimates of the Economic Return to Schooling for the United Kingdom." *The American Economic Review* 85(5) : 1278-1286.

Hatemi, H., et al. (2003) "Prevalence of Overweight and Obesity in Turkey," *Metabolic Syndrome and Related Disorders* 1(4): 285-290.

Hung, H.C, et al. (2004) "Fruit and Vegetable Intake and Risk of Major Chronic Disease." *Journal of the National Cancer Institute* 96 (21): 1577-1584.

Ichino, A., and R. Winter-Ebmer (2004) "The Long-Run Educational Cost of World War II." *Journal of Labor Economics* 22 (1): 57-87.

Idler, E.L. and Y.Benyamini (1997) “Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies”, *Journal of Health and Social Behavior* 38(1): 21-37.

Johar, M., and H. Katayama (2012) "Quantile Regression Analysis of Body Mass and Wages." *Health Economics* 21 (5): 597-611.

Johnston, D.W., C. Propper and M. A. Shields (2009) “Comparing Subjective and Objective Measures of Health: Evidence from Hypertension for the Income/Health Gradient.”*Journal of Health Economics* 28 (3): 540-552.

Johnston, D.W., G. Lordan, M.A. Shields and A. Suziedelyte (2015) “Education and Health Knowledge: Evidence from UK Compulsory Schooling Reform.” *Social Science and Medicine* 127: 92-100.

Jürges, H., S. Reinhold and M. Salm (2011) “Does Schooling Affect Health Behavior? Evidence from Educational Expansion in Western Germany” *Economics of Education Review* 30: 862-872.

Kane, T. J., and C. E. Rouse (1995) "Labor-Market Returns to Two-and Four-Year College." *The American Economic Review* 85 (3): 600-614.

Karakuş, Abdullah. “Sigara Yasağında Yeni Yol Haritası” (New Road Map for Smoking Ban) *Milliyet.* 25th July 2014. Accessed from http://www.milliyet.com.tr/sigara-yasaginda-yeni-yol-haritasi-gundem-1916821/. Accessed on 11th May 2015.

Kemptner,D., H. Jürges and S. Reinhold (2011) “Changes in Compulsory Schooling and the Causal Effect of Education on Health: Evidence from Germany”, *Journal of Health Economics* 30 (2): 340-354.

Kenkel, D. S. (1991) "Health Behavior, Health Knowledge, and Schooling," *Journal of Political Economy* 99(2): 287-305.

Kırdar, M.G., M. Dayıoğlu-Tayfur, and İ. Koç (2014) “Does Longer Compulsory Education Equalize Schooling by Gender and Rural/Urban Residence?” *IZA Discussion Paper* No.7939. Bonn.

Kırdar, M.G., M. Dayıoğlu-Tayfur, and İ. Koç (2012) "The Effect of Compulsory Schooling Laws on Teenage Marriage and Births in Turkey," *MPRA Working Paper* No.38735. Munich.

Kling, J.R., J.B. Liebman, and L.F. Katz (2007) "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83-119.

Kocabaş, A., et al. (1994) "Smoking Patterns at Different Medical Schools in Turkey", *Tobacco Control* 3 (3): 228-235.

Koenker, R., and G. Bassett (1978) "Regression Quantiles" *Econometrica: Journal of the Econometric Society* 46 (1): 33-50.

Koenker, R., and K. Hallock (2001) "Quantile Regression: An Introduction," *Journal of Economic Perspectives* 15 (4): 43-56.

Kol, E. (2015) "Türkiye'de Sağlık Reformlarının Sağlık Hakkı Açısından Değerlendirilmesi." (Evaluation of Healthcare Reforms in Turkey in Terms of Health Right) *Sosyal Güvenlik Dergisi - SGD* 5 (1): 135-164.

Koray, M. (2000). *Sosyal Politika (Social Policy)*. Ezgi Kitabevi. Bursa.

Lantz, P. M., et al. (1998) "Socioeconomic Factors, Health Behaviors, and Mortality." *JAMA: The Journal of the American Medical Association* 279 (21): 1703-1708.

Lantz, P. M., et al. (2001) "Socioeconomic Disparities in Health Change in a Longitudinal Study of US Adults: The Role of Health-Risk Behaviors," *Social Science & Medicine* 53(1): 29-40.

Liang, J., et al. (2005) "How Does Self-Assessed Health Change with Age? A study of Older Adults in Japan." *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 60 (4): 224-232.

Lindeboom, M., and M. Kerkhofs (2009) "Health and Work of The Elderly: Subjective Health Measures, Reporting Errors and Endogeneity in The Relationship between Health and Work." *Journal of Applied Econometrics* 24 (6): 1024-1046.

Lleras-Muney, A., (2005) “The Relationship between Education and Adult Mortality in the U.S.” *Review of Economic Studies* 72 (1): 189-221.

Mackenbach, Johan P., et al. (2005) "The Shape of the Relationship between Income and Self-Assessed Health: An International Study." *International Journal of Epidemiology* 34 (2): 286-293.

Mankiw, N.G., D. Romer and D. N.Weil (1992) “A Contribution to the Empirics of Economic Growth” *The Quarterly Journal of Economics* 107(2): 407-437.

Metintaş, S., et al. (1998) "Smoking Patterns of University Students in Eskişehir, Turkey," *Public Health* 112 (4): 261-264.

Miilunpalo, S., I. Vuori, P. Oja, M. Pasanen, and H. Urponen (1997) “Self-Rated Health Status as a Health Measure: The Predictive Value of Self-Reported Health Status on the Use of Physician Services and on Mortality in the Working-Age Population” *Journal of Clinical Epidemiology* 50(5): 517-528.

Mocan, N., and D. T. Altindag (2014) "Education, Cognition, Health Knowledge, and Health Behavior." *The European Journal of Health Economics* 15 (3): 265-279.

Mokdad, A. H., J. S. Marks, D. F. Stroup, and J. L. Gerberding (2004) “Actual Causes of Death in the United States 2000,” *Journal of American Medical Association* 291 (10): 1238-1245.

Mullahy, J. (1997) "Instrumental-Variable Estimation of Count Data Models: Applications to Models of Tobacco Smoking Behavior," *Review of Economics and Statistics* 79 (4): 586-593.

Ndiaye, A.S., I.T. Diop and S.F.A. Dedehouanou. (2013) “Effect of Governance on Public Service Delivery in the Water Sector in Senegal,” *Global Development Network Working Paper Series* No.84. New Delhi.

OECD Health Data Set from: http:// stats.oecd.org. (Accessed on October 2012).

Oreopoulos, P. (2006) "Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter." *The American Economic Review* 96 (1): 152-175.

Philipson, T. J., and R. A. Posner (2003) "The Long-Run Growth in Obesity as a Function of Technological Change." *Perspectives in Biology and Medicine* 46 (3): 87-107.

Philipson, T. J., and R. A. Posner (2008) "Is the Obesity Epidemic a Public Health Problem? A Review of Zoltan J. Acs and Alan Lyles's" Obesity, Business and Public Policy." *Journal of Economic Literature* 46 (4): 974-982.

Pischke, J.S., and T. Von Wachter (2008) "Zero Returns to Compulsory Schooling in Germany: Evidence and Interpretation." *The Review of Economics and Statistics* 90 (3): 592-598.

Pischke, J.S., and J. D. Angrist (2009) *Mostly Harmless Econometrics*. Princeton University Press. Princeton.

Quevedo, C. H., A. M. Jones and N. Rice (2005) "Reporting Bias and Heterogeneity in Self-Assessed Health. Evidence from The British Household Panel Survey" *Health, Econometrics and Data Group (HEDG) Working Paper 05/04* ISSN 1751-1976.

Rashad, I. and M. Grossman (2004) "The Economics of Obesity." *Public Interest* 156: 104-113.

Rashad, I. (2006) "Structural Estimation of Caloric Intake, Exercise, Smoking, and Obesity." *The Quarterly Review of Economics and Finance* 46 (2): 268-283.

Reijneveld, S.A.(2002) "Neighbourhood Socioeconomic Context and Self Reported Health and Smoking: a Secondary Analysis of Data on Seven Cities" *Journal of Epidemiology and Community Health* 56 (12): 935-942.

Roodman, D. (2011) "Fitting Fully Observed Recursive Mixed-Process Models with CMP." *The Stata Journal* 11(2): 159-206.

Sander, W. (1995) "Schooling and Smoking." *Economics of Education Review* 14 (1): 23-33.

Schultz, T.W. (1961) "Investment in Human Capital." *The American Economic Review* 51(1): 1-17.

Silles, M. A. (2009) "The Causal Effect of Education on Health: Evidence from the United Kingdom," *Economics of Education Review* 28 (1):122-128.

Staiger, D.O., and J.H.Stock (1997) "Instrumental Variables Regression with Weak Instruments", *Econometrica* 65 (3): 557-586.

Stata Annotated Output Ordered Logistic Regression. UCLA: Statistical Consulting Group from : http://www.ats.ucla.edu/stat/stata/output/stata_ologit_output.htm (accessed August 28, 2014).

Stewart, S. T., D. M. Cutler, and A. B. Rosen (2009) "Forecasting the Effects of Obesity and Smoking on US Life Expectancy." *New England Journal of Medicine* 361 (23): 2252-2260.

Şen, A. (2013) “Osmanlı’dan Günümüze Eğitimde Modernleşme Çabaları.” (The Modernization Attempts in Education Since the Ottomans till Today) *EKEV Akademi Dergisi* 17(57): 477-492.

Tansel, A. (1993) “Cigarette Demand, Health Scares and Education in Turkey”, *Applied Economics* 25(4): 521-529.

Tansel, A. “Health System of Turkey.” *World Health Systems, Challenges and Perspectives*, Ed. B.J. Fried and L. M. Gaydos. Chicago, Illinois: Health Administration Press, 2012. 381-407.

Turkish Nutrition and Health Survey Report, 2010 from: http://www.sagem.gov.tr/TBSA_Beslenme_Yayini.pdf. (Accessed on November 2014).

Türkiye Cumhuriyeti Milli Eğitim Bakanlığı (Ministry of Education of Turkish Republic).“Milli Eğitim Temel Kanunu” from http://mevzuat.meb.gov.tr/html/temkanun_1/temelkanun_1.html) (accessed July 10, 2014)

Vaillant N. and F.C. Wolff (2012) “On the Reliability of Self-Reported Health: Evidence from Albanian Data” *Journal of Epidemiology and Global Health* 2 (2): 83-98.

Van Doorslaer, E. and A. M. Jones (2003) “Inequalities in Self-Reported Health: Validation of A New Approach to Measurement” *Journal of Health Economics* 22(1): 61-87.

Villar, J. G., and C. Quintana-Domeque (2009) "Income and Body Mass Index in Europe." *Economics and Human Biology* 7 (1): 73-83.

Webbink, D., N. G. Martin and P. M. Visscher (2010) “Does Education Reduce the Probability of Being Overweight?”, *Journal of Health Economics* 29 (1): 29-38.

Weil, D.N. (2007) “Accounting for the Effect of Health on Economic Growth” *The Quarterly Journal of Economics* 122(3):1265-1306.

Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press. Cambridge.

Worldbank Data Set from: http:// data.worldbank.org/indicator. (Accessed on October 2012).

Yıldırım, H.H. (2013). *Türkiye Sağlık Sistemi: Sağlıkta Dönüşüm Programı Değerlendirme Raporu (Turkish Health System: An Evaluation Report on Health Transformation Program).* Sağlık-Sen Yayınları. Ankara.

Yıldırım, S. (1994). *Sağlık Hizmetlerinde Harcama ve Maliyet Analizi (Expenditure and Cost Analysis on Health Care Services).* Uzmanlık Tezi. DPT.

Yen, S. T.,W. D. Shaw and Y. Yuan (2010) "Cigarette Smoking and Self-Reported Health in China" *China Economic Review* 21 (4): 532-543.

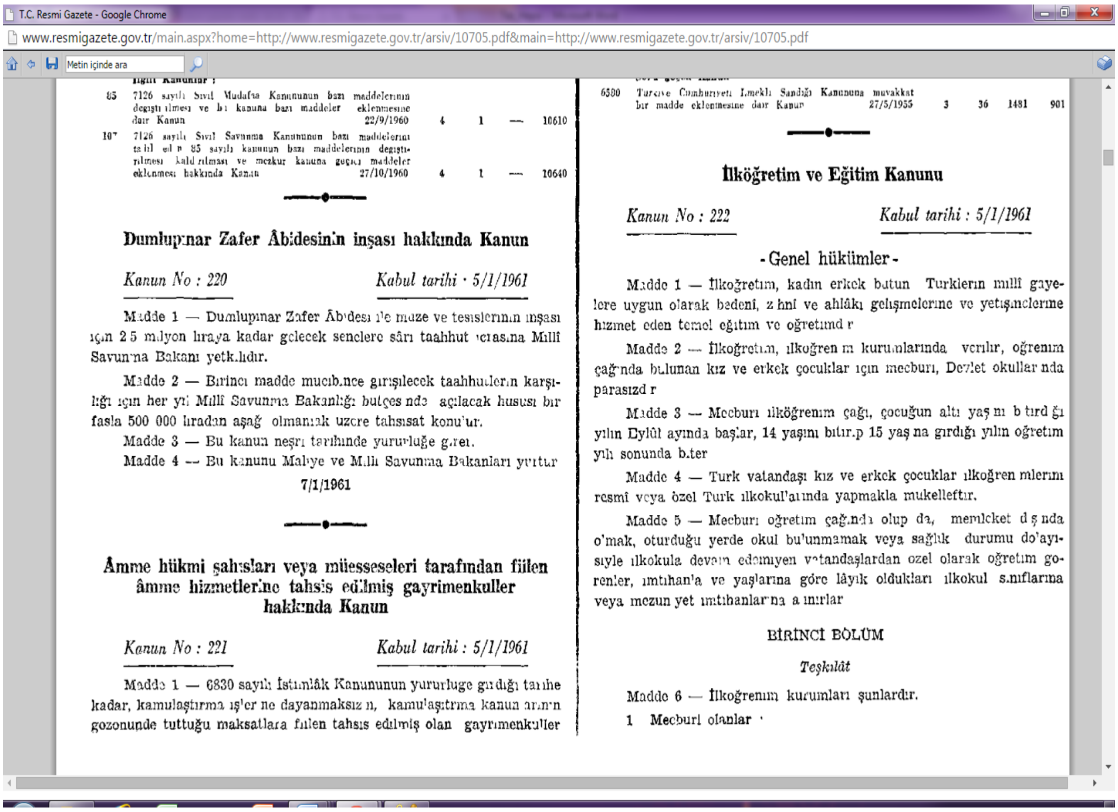
Yumuk, V. D. (2005) "Prevalence of Obesity in Turkey." *Obesity Reviews* 6 (1): 9-10.

Zimmer, Z., J. Natividad, H.S. Lin and N. Chayovan (2000) "A Cross-National Examination of the Determinants of Self-Assessed Health" *Journal of Health and Social Behavior* 41 (4): 465-481.

APPENDICES

A. ORIGINAL LAWS

In this part of the thesis we show the sources of original education laws that we discuss in chapter five. The sources are the official newspapers dated 12.01.1961 and 15.10.1961.



T.C. Resmi Gazete - Google Chrome

www.resmigazete.gov.tr/main.aspx?home=http://www.resmigazete.gov.tr/arsiv/10705.pdf&main=http://www.resmigazete.gov.tr/arsiv/10705.pdf

Metin içinde ara

85	7126 sayılı Sivil Müdafaa Kanununun bazı maddelerinin değiştirilmesi ve bu kanuna bazı maddeler eklenmesine dair Kanun 22/9/1960	4	1	—	10610
107	7126 sayılı Sivil Savunma Kanununun bazı maddelerini tadil eden 85 sayılı kanunun bazı maddelerinin değiştirilmesi, kaldırılması ve mezkûr kanuna geçici maddeler eklenmesi hakkında Kanun 27/10/1960	4	1	—	10640

Dumlupınar Zafer Âbidesinin inşası hakkında Kanun

Kanun No : 220 *Kabul tarihi : 5/1/1961*

Madde 1 — Dumlupınar Zafer Âbidesi ile müze ve tesislerinin inşası için 2,5 milyon liraya kadar gelecek senelere sâri taahhüt icrasına Millî Savunma Bakanı yetkilidir.

Madde 2 — Birinci madde mucibince girişilecek taahhütlerin karşılığı için her yıl Millî Savunma Bakanlığı bütçesinde açılacak hususi bir fasla 500 000 liradan aşağı olmamak üzere tahsisat konulur.

Madde 3 — Bu kanun neşri tarihinde yürürlüğe girer.

Madde 4 — Bu kanunu Maliye ve Millî Savunma Bakanları yürütür.

7/1/1961

Âmme hükmi şahısları veya müesseseleri tarafından fiilen âmme hizmetlerine tahsis edilmiş gayrimenkuller hakkında Kanun

Kanun No : 221 *Kabul tarihi : 5/1/1961*

Madde 1 — 6830 sayılı İstimlâk Kanununun yürürlüğe girdiği tarihe kadar, kamulaştırma işlerine dayanmaksızın, kamulaştırma kanunlarının gözönünde tuttuğu maksatlara fiilen tahsis edilmiş olan gayrimenkuller

6580	Türkiye Cumhuriyeti Emekli Sandığı Kanununa muvakkat bir madde eklenmesine dair Kanun 27/5/1955	3	36	1481	901

İlköğretim ve Eğitim Kanunu

Kanun No : 222 *Kabul tarihi : 5/1/1961*

- Genel hükümler -

Madde 1 — İlköğretim, kadın erkek bütün Türklerin millî gayelere uygun olarak bedenî, zihnî ve ahlâki gelişmelerine ve yetişmelerine hizmet eden temel eğitim ve öğretimdir.

Madde 2 — İlköğretim, ilköğrenim kurumlarında verilir, öğrenim çağında bulunan kız ve erkek çocuklar için mecburi, Devlet okullarında parasızdır.

Madde 3 — Mecburi ilköğrenim çağı, çocuğun altı yaşını bitirdiği yılın Eylûl ayında başlar, 14 yaşını bitirip 15 yaşına girdiği yılın öğretim yılı sonunda biter.

Madde 4 — Türk vatandaşı kız ve erkek çocuklar ilköğrenimlerini resmî veya özel Türk ilkokullarında yapmakla mükelleftir.

Madde 5 — Mecburi öğretim çağında olup da, memleket dışında olmak, oturduğu yerde okul bulunmamak veya sağlık durumu dolayısıyle ilkokula devam edemiyen vatandaşlardan özel olarak öğretim görenler, imtihanla ve yaşlarına göre lâyık oldukları ilkokul sınıflarına veya mezuniyet imtihanlarına alınırlar.

BİRİNCİ BÖLÜM

Teşkilât

Madde 6 — İlköğrenim kurumları şunlardır:

1. Mecburi olanlar:

b) Tamamlayıcı sınıflar ve kurslar.

Madde 7 — İlkokul; 1 inci maddede belirtilen amacı gerçekleştirmek için kurulmuş bir Milli Eğitim ve Öğretim Kurumudur.

Madde 8 — İlkokulun öğretim süresi en az beş yıldır.

Madde 9 — Çeşitli sebeplerle henüz bir ilkokul açılmamış olan birbirlerine yakın birkaç köyün bulunduğu yerlerde veya evleri ve ev grupları dağınık olan köylerde gündüzlü, yatılı, pansiyonlu bölge okulları ve gezici okullar açılabilir ve gezici öğretmenler görevlendirilir.

Bu okullar 6 ncı maddede gösterilen yetiştirici sınıfları ve kursları da ihtiva edebilir.

Şehir ve kasabalarda da, ihtiyaca göre yatılı ve pansiyonlu okullar kurulabilir.

Madde 10 — Bir köy halkı, iş ve üretim hayatının gereği olarak veya olağanüstü sebeplerle yayla, otlak ve bağ gibi yerlere taşındığında köy okulu da, köy halkiyle birlikte göçecek şekilde düzenlenmiş olmalı ve gidilen yerde hemen günlük çalışmasını ve görevini devam ettirmelidir.

Madde 11 — Yetiştirici ve tamamlayıcı sınıflar ve kurslar, mecburi ilköğrenim çağında bulundukları halde, öğrenimlerini yaşıtlariyle birlikte zamanında yapmamış olan çocuklara kısa yoldan ilköğrenim vermek ve ayrıca yetişmelerine lüzum görülen çocukları ilkokula hazırlamak veya ilkokulu bitirmiş olup da henüz mecburi öğrenim çağında bulunan ve üst dereceli öğrenim kurumlarına gidemiyecek olanların genel bilgilerini artırmak ve kendilerine iş ve üretim hayatında faydalı olacak bilgi ve maharetleri kazandırmak amaciyle gerçek ve tüzel kişilerle, belediyeler, özel idareler ve Devlet tarafından açılabilir.

Madde 12 — Mecburi ilköğrenim çağında bulundukları halde, zihnen, bedenen, ruhan ve sosyal bakımdan özürlü olan çocukların özel eğitim ve öğretim görmeleri sağlanır.

Madde 13 — Okul öncesi kurumlarında mecburi öğrenim çağına gelmemiş olan çocuklar eğitilir.

İsteğe bağlı tamamlayıcı sınıflarda ve kurslarda, ilköğrenim çağı dışına çıkmış olup da üst dereceli öğrenim kurumlarına gidememiş olan yurtdaşlardan genel bilgilerini artırmak ve kendilerinin daha iyi bir iş ve üretim unsuru olarak yetiştirilmeleri amaciyle öğretim yapılır.

Bu kurumlar gerçek ve tüzel kişilerle belediyeler, özel idareler ve Devlet tarafından açılabilir.

Okul müdür ve müdür yardımcıları ile ilköğretim müdürleri yönetmeliğindeki esaslara göre ayrıca müstakil sınıf veya ders okutmakla da yükümlüdürler.

Madde 20 — İlkokul müdürlerine müdür yardımcılarına ve tek öğretmenli okulları yönetenlere bu kanunun 88 inci maddesine göre makam tazminatı verilir.

Madde 21 — İlkokullara müdür olabilmek için en az iki yıl başarı ile öğretmenlik veya bir yıl müdür yardımcılığı yapmış olmak şarttır.

Madde 22 — İlköğretim müdürlüklerine :

a) En az üç yıl başarı ile ilkokul müdürlüğü veya müdür yardımcılığı yapmış olanlardan,

b) En az beş yıl başarı ile ilkokul öğretmenliği yapmış olanlardan, Milli Eğitim Müdürlüğünce gösterilecek üç aday arasından,

c) İlköğretim müfettişlerinden, Milli Eğitim Bakanlığınca uygun görülenler atanır.

Madde 23 — İlköğretim müfettişleri :

a) Asil ilkokul öğretmenliği yaptıktan sonra, eğitim enstitülerini veya yabancı memleketlerdeki bu enstitülere denk okulları bitirenlerden;

b) İlkokullarda, en az 8 yıl öğretmenlik, veyahut en az iki yılı ilköğretim veya ilkokul müdürlüklerinde, ilkokul müdür yardımcılığında geçmiş olmak şartiyle en az 6 yıl öğretmenlik yapmış olanlardan Milli Eğitim Bakanlığınca kabul edilecek esaslara göre bu görevlerdeki başarıları ve müfettişlik için gerekli nitelik ve yeterlikleri bulunduğu saptanmış bulunan ve açılacak kurslarda başarı gösteren ilkokul öğretmenlerinden Bakanlıkça atanırlar.

İlk öğretim müfettişliğinde başarı sağlayamadıkları usulü veçhiyle tesbit edilenlerle hastalık ve başka sebeplerle bu görevi yapamıyacak durumda olanlar öğrenim derecelerine uygun okul öğretmenliklerine almakta oldukları aylıklarla atanırlar.

İlköğretim müfettişlikleri, yönetmeliğine göre, bölgelerindeki özel ve resmi ilköğretim kurumları ile bütün dersane ve kursların ve yetişkinler eğitimi kurumlarının eğitim, öğretim ve yönetim işlerini ve çocuk kitaplıklarını teftiş ve murakabe etmek; bu kurumlarda çalışan öğretmenleri özellikle stajyer vekil ve askerlik müddetlerini okullarda yapmakta olan öğretmenleri yetiştirmek, öğrenci devamını takip ve devamsızlığı önleyici tedbirler almak, okul idareleri ve öğretim kurullarınca devam-

1076 sayılı Yedek Subay ve Yedek Askeri Memurlar Kanununun 3 üncü maddesinin değiştirilmesine ve bu kanuna geçici maddeler ilâvesine dair Kanun

Kanun No : 97 *Kabul tarihi 11/10/1960*

Madde 1 — 1076 sayılı Yedek Subay ve Yedek Askeri Memurlar Kanununun 3923 sayılı kanunla tadil edilen, 3993, 4107, 4175, 4497, 4690, 5221, 5674 ve 6137 sayılı kanunlarla fıkralar eklenen 3 üncü maddesi aşağıda yazılı olduğu şekilde değiştirilmiştir :

Madde 3 — Her sene yapılan son yoklama neticesinde sağlam veya sakat asker edilenlerden yüksek okul ve fakültelerden mezun oldukları anlaşılanlarla bu derecede olduğu Milli Eğitim Bakanlığınca tasdik edilen yurt dışındaki yüksek okul ve fakültelerden mezun olanlardan 8 inci maddede yazılı şartları haiz olanlar, yedek subay yetiştirilmek üzere kayıt ve kabul merkezlerine yollanırlar.

Kayıt ve kabul merkezlerine gelenler Silâhlı Kuvvetlerin ihtiyacı gözönünde tutularak test ve mülâkatlarla tesbit edilen kabiliyetlerine, meslek ve ihtisaslarına göre sınıflandırılarak Yedek Subay Okul ve talimgâhlarına gönderilirler.

Yedek subay yetiştirileceklerin hizmet süreleri 2 yıl olup, bunun 6 ayı yedek subay adayı olarak Yedek Subay Okul veya talimgâhlarında, 18 ayı da subay olarak sınıfları birlik veya müesseselerinde eğitim ve hizmette geçer. Okul veya talimgâhlarda geçen 6 aylık tahsil süresi, gerektiğinde Milli Savunma Bakanlığınca lüzumu kadar uzatılabilir. Bu uzatmadan dolayı 2 senelik hizmet süresi değişmez.

8 inci maddede yazılı şartları haiz fakülte ve yüksek okul mezunları yaş hadlerini doldurmuş olsalar bile yukarıdaki fıkrada belirtilen askerlik hizmetlerini yapmadıkça çağdan çıkarılamaz.

Okul veya talimgâhlardaki yetiştirilme sırasında Askeri Ceza Kanununun 154 üncü maddesine göre hükmen çıkarılanlar ve diğer gayriahlâki vasıfları ve disiplinsizlikleri sebebiyle yedek subaylığa yaramıyacakları okul veya talimgâhlar eğitim ve disiplin kurullarında tesbit edilenler dönem sonu beklenmeksizin er olarak; okul veya talimgâhlardaki öğrenim ve eğitim dönemi sonunda başarı gösteremiyenler çavuş olarak geri kalan muvazzaflık hizmetlerini tamamlamak üzere sınıfları kıta ve müesseselerine yollanırlar. Bu gibilerin, Askerlik Kanunundaki kısa hizmet şartlarını taşısalar dahi (Sakatlar hariç), muvazzaflık hizmetleri tamamlattırılır. Bunların Yedek Subay Okul veya talimgâhlarında geçen müddetleri askerlik hizmetlerinden sayılır.

Yedek Subay Okul veya talimgâhlarındaki öğrenim ve eğitim dönemi sonunda yapılan imtihanlarda başarı gösterenler asteğmen nasbolunurlar ve sınıfları kıta ve müesseselerine atanırlar. Buralarda asteğmen olarak 6 ay hizmetten sonra ehliyetleri sicilleriyle sabit olanlar, teğmenliğe yükseltilir ve hizmetleri sonunda teğmen olarak terhis edilirler. Asteğmen olarak 6 aylık hizmetleri sonunda teğmenliğe liyakatleri görülmiyenler, 6 ay daha asteğmen olarak hizmet görürler. Bu sure sonunda liyakatleri sicilleriyle tesbit edilenler teğmenliğe yükseltilirler. Bu süre sonunda da liyakati görülmiyenler terhislerine kadar asteğmen olarak kalırlar ve asteğmen olarak terhis edilirler.

Ordunun çeşitli hizmetlerinde, sivil hayattaki mesleki ihtisasları dahilinde çalıştırılmalarında zaruret görülenler ile teknik eleman ve tercüman olarak vazifelendirilecek olanların yukarıdaki fıkrada yazılı kıta ve sınıfı görevi kaydına tabi olmaksızın tâyin edildikleri yerlerden alacakları sicilleri üzerine terfileri yapılır.

Yedek subay nasbedilenlerin hizmet süreleri olağanüstü hallerde Bakanlar Kurulu kararıyle lüzumu kadar uzatılabilir. Bu uzatılan süreler, yaş haddinin sonundan iki kat düşülür.

Yedek asteğmen ve teğmen olarak kıta ve müesseselerdeki hizmetleri sırasında hastalananlar, muvazzaf subay emsalleri gibi tedavi ve sıhhi izin süresine tabi tutulurlar. Sıhhi izin süresini tamamlıyanlar terhis edilirler. Ancak sıhhi izin süresini tamamlamadan (Faal kıta görevi yapabilir) raporu alanların hizmet süreleri 6 ayı doldurmadıkça terhisleri yapılmaz.

Yedek subaylardan sağlık özürlerinden ötürü orduda istihdama elverişli olmadıklarına ait (Sürekli çürük) raporu alanlar haklarında hastalık ve ârızalarının durum ve derecelerine ve görev alıp alamıyacaklarına nazaran terhis ve mâlûliyet işlemleri özel kanunlarına göre yapılır.

Geçici madde 1 — Bu kanunun mer'iyete girdiği tarihte, yedek subay olanlarla Yedek Subay okullarında talebe olanlar eski kanun hükümlerine tabidirler.

Geçici madde 2 — Bu kanunun yürürlüğe girdiği tarihte, yürürlükten kaldırılan hükümler gereğince yedek subay olmak hakkını veren üniversite ve yüksek okullardan gayri okullardan mezun olup, 1076 sayılı kanunun 8 inci maddesindeki şartları haiz olanlar Milli Eğitim Seferberliğinde kullanılmak üzere öğretmen olarak Milli Eğitim Bakanlığı emrine verilirler ve askerlik şubelerinden sevkedildikleri andan itibaren askerlik hizmetleri başlamış sayılır. Bunlar iki senelik askerlik mükellefiyetleri zarfında Milli Eğitim Bakanlığınca gösterilen yerlerde öğretmenlik yapmağa mecburdurlar.

Bunlardan halen Milli Eğitim Bakanlığı kadrolarında öğretmen olanlar müktesep maaşlarını almağa devam ederler. Bunun haricinde olanlar fiilen öğretmenliğe başladıkları tarihten itibaren baremin 13 üncü derece maaşını alırlar. Bunlar askerlik şubelerinden sevklerinden itibaren iki senelik askerlik mükellefiyetleri esnasında firar ve izin tecavüzü bakımından Askeri Ceza Kanununa tabidirler. Bunların, okulların yaz tatili devresinde ne şekilde askeri eğitime tabi tutulacakları Milli Sa-

B. TURKISH SUMMARY

Beşeri sermaye bir ülkenin üretiminin, refahının ve gelirinin gelişmesini hızlandırıcı bir faktördür. Schultz (1961) beşeri sermayeye yatırımın önemini işaret eden ilk çalışmalarından birini yapmıştır. Bu çalışmada Schultz eğitime ve sağlığa yapılan doğrudan harcamaları, daha iyi koşullarda yaşayabilmek için yapılan iç göçleri, öğrencilerin daha fazla okula giderek kazanmaktan vazgeçtiği gelirleri ve işçilerin mesleki eğitimlere katılarak kazanmaktan vazgeçtiği maaşlarını beşeri sermayenin bileşenleri olarak saymıştır. Schultz'a göre beşeri sermayeye yapılan bu yatırımlar kişi başına düşen reel geliri arttıran temel faktörlerdir. Becker (1993) beşeri sermayenin kötü alışkanlıkları da içerdiğini öne sürmüştür. Örneğin, Becker'a göre aşırı alkol tüketmek de beşeri sermayenin bir bileşenidir. Becker beşeri sermayenin bu tip bileşenlerinin pazar ve pazardışı sektörlerde üretim üzerinde negatif bir etkiye sahip olduklarını söylemiştir.

Sağlık ve eğitim beşeri sermayenin en önemli iki bileşenidir. Sağlık ve eğitime yapılan yatırımların bir ekonomin büyümesi üzerinde pozitif etkilerinin olduğu daha önce yapılan çalışmalarda defalarca gösterilmiştir. Örneğin, Mankiw ve diğerleri (1992) fiziki semaye ile birlikte beşeri sermayeyi (bu çalışmada sadece eğitim) içeren genişletilmiş Solow modelinin ülkeler arası gelir farklılığını beşeri sermayeyi içine katmadan oluşturulan Solow modelinden daha iyi açıkladığını göstermişlerdir. Daha sonra Weil (2007) bu genişletilmiş Solow Modeline sağlığı da eklemiştir. Weil bu çalışmasında temel olarak sağlığın büyüme üzerindeki etkisini incelemiştir. Weil, bir ülkede yaşayan insanların daha sağlıklı olması durumundan daha üretken işçi oldukları savını öne sürmüş ve bunun bir ülkenin büyüme oranını artırdığını savunmuştur. Weil ülkeler arası gelir farklılıklarını belirlemede sağlığın önemini vurgulamıştır.

Bu tezin amacı beşeri sermayenin iki önemli bileşeni olan sağlık ve eğitim arasındaki ilişkiyi gelişmekte olan bir ülke, Türkiye, için incelemektir. Bu tezde öncelikli olarak kişinin sağlık durumunu etkileyen sosyo ekonomik faktörler incelenmiş ve bu faktörlerin kişinin sağlığı üzerinde ne gibi bir şekilde etkili olabilecekleri tartışılmıştır. Bu faktörler başlıca eğitim, cinsiyet, yaş, yaşanılan bölge (kentsel/kırsal), medeni durum, istihdam durumu ve hane halkı geliri olarak sıralanabilir. Daha sonra sağlık ve eğitim arasındaki amprik ilişkiye iki şekilde bakılmıştır. İlk olarak kişinin eğitim seviyesinin sağlığa göre dışsal bir faktör olduğu varsayılmış ve sıralı olasılık modelleri ile iki bileşen arasındaki ilişki test edilmiştir. İlişkiyi test ederken, eğtimle birlikte diğer sosyo-ekonomik faktörler de kontrol edilmiştir. Daha sonra sağlıkla eğitim arasında var olan içsellik problemi ele alınmış ve araç değişken metodları kullanılarak iki bileşen arasındaki ilişki tekrar test edilmiştir. Bu amprik modellerde kişinin kendi rapor ettiği sağlık durumu (Çok iyi/İyi/Orta/Kötü/Çok Kötü) sağlık çıktısı olarak kullanılmıştır. Tezin bir sonraki amprik kısmında ise kişilerin sağlık ile ilgili olan davranışları sağlık çıktıları olarak kullanılmıştır. Bu davranışlar sigara tüketimi, alkol tüketimi, meyve ve sebze tüketimi ile kişinin düzenli egzersiz faaliyetleridir. Bunların dışında kişinin veri setinde mevcut olan boy ve kilosu kullanılarak Vücüt Kitle İndeksi (VKİ) hesaplanmış ve eğitimle aralarındaki ilişki test edilmiştir. Bu bölümde de kişinin eğitim seviyesi önmce dışsal sonra içsel olarak ele alınmıştır. Son olarak gelişmiş ve gelişmekte olan ülkelerde çağın en büyük hastalığı olan obezite üzerinde özellikle durulmuş ve kişinin sosyo-ekonomik faktörlerinin kişinin VKİ si üzerindeki etkisi kesirli regresyon modelleri kullanılarak incelemiştir. Burdaki temel amaç eğitim başta olmak üzere kişinin sosyoekonomik faktörlerinin kişinin VKİ'sini değişik yüzdelerde nasıl etkilediğidir.

Bu tezin literatüre değişik açılardan katkısı vardır: Birinci olarak, bu tez Türkiye'de eğitimin sağlık ile ilişkisini, sağlığı olası etkileyen diğer faktörler de ele alınarak, 25 yaş ve üstü insanlar için inceleyen ilk çalışmadır. İkinci olarak, bu tezdeki amprik bulgular Türkiye Sağlık Araştırması (TSA) veri seti kullanılarak elde edilmiştir. TSA Türkiye'de çok yeni kullanılan bir veri setidir. Bildiğimiz kadarıyla bu tez ve Cesur ve diğerlerinin (2014) çalışmaları dışında başka bugüne kadar başka bir çalışmada kullanılmamıştır. TSA oldukça fazla olan gözlem sayısı ve kişi sağlığı ile ilgili değişik sorularıyla ileriki sağlık çalışmaları için de kullanılabilecek zengin bir mikro veri seti özelliğini taşımaktadır. Üçüncü olarak bu tezin amprik kısmında çok değişik mikroekonometrik modeller kullanılmıştır. Bunlardan koşullu karmaşık işlemler ve kantil araç değişken teknikleri bildiğimiz kadarıyla Türkiye için yapılan bir uygulamalı çalışmada ilk defa kullanılmıştır. Dördüncü olarak, sağlık ve eğitim arasındaki ilişkiyi çözmek için ilgili litratür takip edilerek araç değişken teknikleri kullanılmış ve litratüre eğitim için yeni bir araç değişken kazandırılmıştır. Bu araç değişken 1960'lı yılların başında Türkiye'de yapılan eğitim reformlarıdır. O yıllarda yapılan üç önemli reform şu şekilde sıralanabilir: İlk olarak köylerde daha önce üç yıl olan zorunlu eğitim süresi beş yıla çıkarılmıştır. Böylece kişiler daha fazla sene okula gitmişlerdir. 1961'de çıkarılan bir başka yasa ile ortaokul mezunu kişiler belirli bir kursu geçtikleri takdirde ilkokul öğretmeni olma hakkını, aynı şekilde lise mezunları da ortaokul öğretmeni olma hakkını kazanmışlardır. Bu şekilde Türkiye'de genel olarak öğretmen sayısı artmıştır. Son olarak da,Kasım 1960'da çıkarılan bir yasayla yedek subaylara da askerliklerini öğretmen olarak yapma hakkı verilmiş, ve askerlikleri bitse dahi bu kişilerin istedikleri takdirde öğretmenliğe devam etme hakkı tanınmıştır. Bu kanun da Türkiye'de öğretmen sayısını arttırmada önemli bir faktör olmuştur.

1961 yılında çıkarılan başka bir yasa sağlıkta sosyalleşme reformları ile ilgilidir. Bu yasanın temel amacı birincil sağlık hizmetleri olan sağlık ocakları ve sağlık evlerinin öncelikli kullanımını yaygınlaştırarak, birincil ve ikincil sağlık hizmetlerinin (büyük ölçekli hastahaneler) entegre bir şekilde çalışmasını sağlamaktır. Ancak kaynak yetersizliği, politik sorunlar, kişilerin öncelikli olarak ikincil sağlık hizmetini seçmeleri gibi sebeplerden dolayı sağlıkta sosyalleşme o yıllarda ve sonraki 15 yılda bir türlü etkin bir şekilde gerçekleştirilememiştir. Sağlıkta etkin reformların başlaması ancak 1990'lı yılların başında olabilmiştir. Dolayısıyla 1960 'lı yıllarda sağlık ve eğitim reformları ortak bir noktada kesişememiştir. Yani, takip eden dönemlerde kişinin sağlık durumunda bir iyileşme var ise bunu eğitim reformları kadar sağlık reformlarına da bağlamak çok doğru bir sav olamaz.

Türkiye’de daha önce yapılan gelir ve eğitim üzerine yapılan çalışmalar eğitimin araç değişkeni olarak 1997’de kabul edilen zorunlu eğitimin 5 yıldan 8 yıla çıkma politika değişikliğini kullanmaktadır. Bu tezde 1960 eğitim reformlarının araç değişken olarak kullanılması 25 ve üstü kişilerin sağlıkları ve sağlık davranışlarının değişiminin daha doğru bir şekilde incelenmesi açısından önemlidir, 1997’deki reformlardan etkilenen çoğu kişinin henüz çok genç olmasından dolayı sağlık durumlarında eğitimlerine bağlı olarak bir iyileşme ve kötüleşme olmasının mümkün olmadığı varsayılmıştır.

Türkiye’de eğitim ve sağlık arasındaki ilişki daha önce nerdeyse hiç incelenmemiştir. Bu alanda yapılan çalışmalar oldukça sınırlıdır. Erdoğan ve diğerleri (2012) zaman serileri veri seti kullanarak eğitimde cinsiyet eşitsizliği ve sağlık arasındaki ilişkiyi incelemiştir. Erdoğan ve arkadaşları bu çalışmalarında doğum anındaki yaşam beklentisi, beş yaş altı ölüm oranı ve de yeni doğan bebek ölüm oranlarını sağlık çıktısı olarak kullanmışlardır. Bir başka çalışmada Tansel (1993) Türkiye’de sigara tüketimindeki talebi etkileyen faktörleri incelemiştir. Tansel çalışmasının sonucunda kişilerin sigaranın sağlığa zararlı olması hakkında yeterince bilgilendirilmeleri durumunda sigaraya olan talebin düştüğünü söylemiştir. Tansel bu çalışmasında ayrıca üniversite mezunlarının diğer eğitim gruplarındaki kişilere göre daha az sigara içtiği sonucuna varmıştır. Çok yakın bir zamanda yapılan diğer bir çalışmada Cesur ve diğerleri (2014) Türkiye’de kişilerin sağlık davranışları ve eğitim seviyesi arasındaki ilişkiyi incelemiştir. Cesur ve arkadaşları çalışmalarını 18 ve 30 yaş arası kişiler için yapmışlardır. Çalışmalarının sonucunda sigara tüketimi ve eğitim seviyesi arasında anlamlı bir ilişki bulamamışlar ve fakat eğitim yüksek seviyesi yüksek insanların obez ve daha fazla kilolu olmaya eğilimli olduklarını göstermişlerdir.

Beşeri sermaye üzerine en önemli çalışmalar ilk olarak Becker (1965) ve Schultz (1961) tarafından yapılmıştır. Bu calışmaların ışığında Grossman (1972) litratüre “Sağlık Talebi” teorisini kazandırmıştır. Grossman’in bu teorisi sağlık ekonomisinde halen en önemli teori olma özelliğini taşımaktadır. Bu teorinin tez açısından en önemli özelliği sağlık ve eğitim arasındaki pozitif ilişkiyi ilk gösteren teori olmasıdır.

Bu model hakkında Grossman'e yöneltilen en önemli kritiklerden biri eğitim ve sağlık arasındaki içsellik probleminin modelde ele alınmamis olmasıdır. Tezin ikinci bölümünde önce Grossman'in ünlü modeli ana hatlarıyla incelenmiş, daha sonra günümüze daha yakın zamanlarda Grossman'in modeli temel alınarak oluşturulan diğer modeller üzerinde durulmuştur. Tezin ilgili bölümünde son olarak bu modellerin sonuçları faz alınarak tezde kullanılacak ve kişinin sağlığını ve sağlık davranışlarını etkileyebileceği düşünülen sosyoekonomik faktörler belirlenmiş ve bu faktörlerin sağlığı ne gibi yönlerde etkileyebileceği tartışılmıştır.

Tezin üçüncü bölümünde ise TSA detaylı bir şekilde incelenmiştir. TSA ileride sağlık ile ilgili çalışmalarda kullanılabilecek önemli bir veri setidir. TSA rassal olarak seçilmiş hanelerde yaşayan kişilerle ilgili cinsiyet, yaş, eğitim durumu, medeni durum, istihdam durumu, hanehalkı geliri gibi bilgiler haricinde sağlıkla ilgili zengin bir soru havuzuna sahiptir. Bu tezde kullanılan sorular haricinde veri setinde bulunan sağlık bilgileri kısaca şöyle sıralanabilir: Kişinin psikolojik durumu, ne kadar sıklıkla koruyucu sağlık hizmetleri aldığı (aşılanma, tansiyon ölçümü, kolestrol ölçümü, kan şekeri ölçümü, mamografi, simir testi, prostat muayenesi vb...), yakın zamanda trafik kazası geçirip geçirmediği vb...TSA'nın en büyük handikapı tüm soruların cevaplarının kişinin kendi rapor ettiği cevaplar olmasıdır. Dolayısıyla tüm soruların subjektif bir bakış açısıyla cevaplandırılması çok olası bir ihtimaldir.

TSA'nın bir diğer handikapı ise ölüm, yaşam beklentisi gibi önemli sağlık çıktılarının veri setinde bulunmamasıdır. Bu yüzden tezde ana sağlık çıktımız kişinin kendi rapor ettiği sağlık durumu olarak belirlenmiştir. Kişinin kendini iyi veya kötü hissetmesinin kişinin hasta olup olmaması veya ölümüyle (Panel çalışmalarda görülebliniyor) doğrudan pozitif bir ilişkiye sahip olduğu litratürde daha önce defalarca gösterilmiştir (Idler ve Benyamini, 1997). Kişinin rapor ettiği sağlık durumunun avantajları ve dezavantajları tezin dördüncü bölümünde detaylı bir şekilde tartışılmıştır. Bu bölümde ayrıca kişinin kendi hiseettiği sağlık durumundaki varyasyonlar eğitim, cinsiyet, yaş, yaşanılan bölge, medeni durum, istihdam durmu ve hanehalkı gelirine göre frekans analizleri yapılmıştır. Frekans analizlerinin sonuçları kısaca şöyle özetlenebilir: Türkiye'de eğitim seviyesi ve sağlık durumu arasında pozitif bir ilişki vardır. En yüksek eğitim seviyesi olan kişilerin (Üniversite mezunuları ve yüksek lisans, doktora mezunları) diğer eğitim gruplarına göre daha fazla "Çok İyi" sağlık durumu rapor ettikleri görülmüştür. Buna karşılık "Çok Kötü" sağlık durumu rapor etme frekansının en yüksek okuma ve yazma bilmeyen kişilerde görüldüğü saptanmıştır. Hanehalkı gelirinin de kişinin sağlık durumunu pozitif yönde etkilediği önemli bulgularımız arasındadır. Cinsiyete göre yapılan frekans analizinde erkeklerin kadınlardan daha fazla "Çok İyi" veya "İyi" sağlık durumları rapor ettikleri gözlenmiştir. Ayrıca, şehirlerde yaşayan insanların kırsal kesimlerde yaşayan insanlardan sağlık açısından daha iyi oldukları bulduğumuz bir diğer sonuçtur. Frekans analizi evli veya bekar kişilerin eşini kaybetmiş veya eşinden boşanmış kişilere göre daha iyi sağlık durumunda olduklarını göstermektedir. Son olarak, istihdam durumuna göre yapılan frekans analizi hiç bir yerde çalışmayan veya iş aramayan (aktif olmayan) kişilerin istihdamda olan veya işsiz olan (aktif bir şekilde iş arayan) kişilerden daha fazla "Kötü" veya "Çok Kötü" sağlık durumları rapor ettiklerini göstermektedir.

Tezin beşinci bölümünde bu çalışmada kullanılan belirli mikroekonometrik tekniklerin metodolojisi ve teorisi incelenmiştir. Bu bölümde ayrıca 1960'lı yıllarda yapılan eğitim ve sağlık reformları detaylı bir şekilde ele alınmıştır.

Tezin altıncı bölümünde kişinin sağlık durumunu belirleyici etkenler amprik olarak incelenmiştir. Bu bölümde ilk olarak sağlık ve eğitim arasındaki içsellik problemi ele alınmamış ve sıralı olasılık modelleri kullanılarak sağlık ve eğitim arasındaki ilişki diğer sosyoekonomik faktörler de kontrol edilerek incelenmiştir. Bu bölümde hem sıralı probit hem de sıralı logit modelleri kullanılmıştır. Her iki model de benzer sonuçlar vermiştir. Ancak sıralı logit modellerinde olasılık oranlarını elde edebildiğimiz için istatistiki açıdan daha anlamlı yorumlar yapmamız sıralı logit modellerini kullanarak sağlanmıştır. Her iki modelde de öncelikle tamamiyle sağlığa dışsal oldukları düşünülen bağımsız değişkenler kontrol edilmiştir: Bunlar, eğitim, yaş, cinsiyet, bölge ve medeni durumdur. Medeni durum bazı modellerde sağlıkla içsel olarak algılansa da, bu tezde Türkiye için tamamiyle dışsal olarak kabul edilmiştir. Çünkü, Türkiye İstatistik Enstitüsü (TÜİK) verilerine göre Türkiye'de insanların yüzde 97 si hayatlarının bir döneminde en az bir kere evlenmektedir. Dolayısıyla, Türkiye için medeni durum ve sağlık arasında iki yönlü bir ilişkinin olamayacağı var sayılmıştır.

Tamamiyle dışsal faktörleri kontrol ederek oluşturduğumuz modellerde sağlık ve eğitim arasındaki pozitif ilişki teyit edilmiştir. Daha sonra sağlıkla içsel bir ilişkide olabilecek istihdam durumu ve hanehalkı geliri (logaritmik olarak) modellere dahil edilmiştir. Sağlık ve eğitim arasındaki pozitif ilişki bu genişletilmiş modellerin sonucuna göre de değişmemiştir. Genişletilmiş modeller ayrıca, istihdamda olan veya aktif olarak iş arayan kişilerin istihdam piyasasında aktif olmayan kişilere göre daha iyi sağlık durumunda oldukları ve hanehalkı gelirinin kişinin sağlık durumuyla pozitif yönde ilişkili olduğunu göstermiştir. Bu bulguların hepsi dördüncü bölümde yaptığımız frekans analizleri ile tutarlı sonuçlardır.

Tezin altıncı bölümünün ikinci kısmıda sağlık ve eğitim arasındaki olası içsellik problemi göz ardı edilmemiş ve daha tutarlı sonuçlar almak için araç değişken teknikleri kullanılmıştır. Araç değişken tekniklerinde sapmalı bir sonuç doğurmaması için kişinin istihdam durumu ve hanehalkı geliri modele hiç katılmamıştır. Araç değişken modellerinde kullandığımız araç değişken 1960'lı yılların başında Türkiye'de yapılan eğitim reformlarıdır. Kanunların hepsi 5 Ocak 1961'de tamamen geçtiği için yeni kanunlardan yararlanan kişilerin Eylül 1961'de okula başladıkları veya bu tarihte ilkokul üçüncü sınıfta olmalarıdır (İlk paragrafta bahsettiğimiz gibi kanunların birinci maddesi köy ilkokullarının üç yıldan beş yıla çıkarılmasını öngörüyordu). Yani, Türkiye'de okula başlama yaşının 6 olduğunu varsayarak, veri setimizde 1952 ve daha sonra doğan kişilerin bu reformlardan etkilendiği düşünülmüştür. Ayrıca, özellikle Eylül ayından sonra doğan çocukların okula bir sene geç başladığı (7 yaşında) gerçeği de göz ardı edilmeyerek bazı model tanımlamalarında bu kişilerin de reformlardan etkilendiği ele alınmıştır. Birinci aşama modelimiz en küçük kareler (EKK) modelidir. Burda bağımlı değişkenimiz kişinin okula devam ettiği yıl sayısıdır. Bu sayı okuma-yazma bilmeyenler için sıfır, okuma-yazma bilen veya herhangi bir kurumdan mezun olmayanlar için iki, ortaokul mezunları için sekiz, lise mezunları için onbir ve üniversite veya daha yüksek eğitim kurumlarından mezun olanlar için onbeşe eşittir. İlkokul mezunları için ise yıl sayısı şu şekilde belirlenmiştir: Öncelikle Milli Eğitim Bakanlığı istatistiklerinden 1952 yılı öncesi köy ve şehir ilkokullarından mezun olan kişilerin sayısı alınmıştır. Daha sonra köy okullarından mezun olan kişilerin sayısı üç ve şehir ilkokullarından mezun olan kişilerin sayısı beş ile çarpılmış, bulunan rakamlar toplanmış ve köy ve şehirlerdeki toplam mezun sayısına bölünmüştür. Her yıl için çıkan ortalama 1952 senesinden önce doğan kişilerin, eğer ilkokul mezunu iseler, eğitim yılı olarak yazılmıştır. 1952 ve daha sonraki yıllarda doğan kişiler içinse bu sayı beş olarak alınmıştır. 1952'den önce doğan kişiler için bu şekilde bir ağırlıklı ortalama almamızın sebebi TSA veri setinde kişinin okula gittiği yıllarda köyde mi şehirde mi yaşadıklarını görememizden kaynaklanmaktadır. Birinci aşamada kullanılan araç değişken ise kukla değişken olarak tanımlanmıştır. Kukla değişkenin değeri eğer kişi 1952 yılında veya daha sonra doğduysa 1'e, bu seneden önce doğduysa 0'a eşittir.

Bu aşamanın sonucunda yapılan içsellik testleri sağlık ve eğitim içsel bir ilişki içinde olduğunu tespit etmiştir. Bu durumda araç değişken metodlarının uygulanmasının gerekliliği kaçınılmazdır. Gene birinci aşama sonucunda yapılan F-istatistikleri ise litratürde önerilen eşik değerinden (10) büyük çıkmıştır. Bu sonuç kullandığımız araç değişkenin geçerli bir araç değişken olduğuna işaret etmektedir. Birinci aşama sonucunda eğitim reformlarının kişinin okula gitme katsayısını 0.99-1.27 birim aralığında ((yapılan değişik model spesifikasyon sonuçlarına göre) arttırdığı sonucuna ulaşılmıştır. Modelin ikinci aşamasında ise bağımlı değişken kişinin kendi rapor ettiği sağlık durumudur. Bu değişken ordinal bir değişken olduğu için aralıklarının sürekli olduğunu varsayarak EKK tahminlemesi uygulanması tutarsız sonuçlar verebilir. Bu yüzden bu aşamada Türkiye için yapılan bir çalışmada ilk defa "Sıralı Araç Değişken " modeli kullanılmıştır. Bu modeli oluşturmada da dünyada ilk defa Roodman (2011) in önerdiği "Koşullu Karmaşık İşlemler" tekniklerinden yararlanılarak yapılmıştır. Bu modellemenin sonucunda kişinin eğitim seviyesinin arttıkça daha iyi bir sağlık durumu belirttiği sonucu süregelmiştir. Yapılan bu reformların sonucunda erkekler ve kadınların eğitim seviyesinin sağlık üzerindeki etkisinin farklılığının anlamlı olmadığı saptanmıştır.

Tezin yedinci bölümünde ise bu sefer sağlık çıktıları olarak sağlık davranışları ele alınmıştır. Bu davranışlar sigara tüketimi, alkol tüketimi, meyve ve sebze tüketimi, düzenli egzersiz yapma ve kişinin VKİ sidir. VKİ haricindeki tüm değişkenler kukla değişken olarak tanımlanmıştır. Örneğin eğer kişi sigara içicisi olduğunu rapor ettiyse bu kişi için sigara değişkeni 1, eğer değilse 0 olarak tanımlanmıştır. Tezin bu bölümünde de sağlık davranışları ve eğitim arasındaki ilişki basit probit (VKİ için EKK) yöntemleri ile test edilmiştir. Bu regresyonlarda sağlık ve eğitim arasındaki içsellik problemi ele alınmıştır. Bu modellemelerde öncelikle her sağlık davranışı için ayrı ayrı eğitimin ilgili davranış üzerindeki etkisi incelenmiştir. Daha sonra aşamasal olarak cinsiyet, yaş, bölge, medeni durum, istihdam durumu ve hanehalkı geliri modellere eklenmiştir. Eklenen her bağımlı değişkenden sonra eğitimin her sağlık davranışı üzerindeki etkisinin düştüğü fakat hala yüksek bir şekilde anlamlı olduğu görülmüştür. Bulunan sonuçlar kısaca şöyle özetlenebilir: Sigara tüketimi olasılığının eğitim seviyesi arttıkça yükseldiği, fakat bu ilişkinin konkav olduğu gözlenmiştir. Yani, belirli bir eğitim seviyesine kadar (ortaokul) sigara tüketimi olasılığı artarken, belli bir eğitim seviyesinden sonra düştüğü görülmüştür. Üniversite veya daha yüksek derecesi olan kişilerde sigara içme olasılığının iyice düştüğü gözlenmiştir. Bu durum yüksek eğitim seviyeli kişilerin sigaranın zararları konusunda daha farkında olduklarını kanıtlamaktadır. Bir başka sonuç ise alkol tüketiminin eğitim seviyesinin yükselmesiyle beraber görülme olasılığının artmasıdır. Bu durumun bizim veri setimiz için beklenen bir durum olduğu söylenebilir. Zira, veri setimizde sürekli içici olduğunu rapor eden kişilerin oranı yüzde 0.5 olarak hesaplanmıştır. Dolayısıyla, alkol tüketimi ve alkol arasındaki ilişkiyi incelerken yapılan analizlerde daha kapsamlı sonuçlar alabilmek adına ara sıra alkol içtiğini söyleyen kişiler ve sürekli içici olduğunu söyleyen kişilerin hepsi "alkol tüketen" kişiler olarak birleştirilmiştir. Daha yüksek eğitim seviyesi olan kişilerin sosyal ortamlara daha çok girdikleri varsayıldığında, bu kişilerin daha çok alkol tükettiği sonucuna ulaşmamız bizce sürpriz değildir.

İncelediğimiz sağlık davranışlarından, insan sağlığına yararlı olabilecek davranışlardan sebze ve meyve tüketimi ile düzenli egzersiz yapma olasılıklarının eğitim seviyesindeki artış ile birlikte yükseldiği görülmüştür. Yine bu sonuca bağlı olarak, eğitim seviyesi yüksek kişilerin sağlığa yararlı davranışların yapmalarının faydalarının kendileri için daha iyi sonuçlar doğuracağının farkında oldukları tespit edilmiştir. Son olarak, kişinin VKİ'sinin eğitim seviyesindeki arışla birlikte daha normal aralıklara dödüğü görülmüştür. Başka bir deyişle günümüzün en yaygın hastalığı olan obezitenin eğitim seviyesi yüksek olan kişilerde görülme olasılığı eğitim seviyesi düşük olan kişilerde görülme olasılığından daha azdır. Benzer şekilde, fazla kilolu olmak da eğitim seviyesi düşük olan kişilerde daha yaygındır.

Tezin yedinci bölümünün ikinci kısmında sağlık davranışları ve eğitim arasındaki içsellik problemi ele alınarak tüm tahminler tekrar yapılmıştır. Birinci ve ikinci aşamada kullanılan bağımsız değişkenler beşinci bölümde kullanılan değişkenler ile aynıdır. Bu bölümde bağımlı değişkenin sigara içme, alkol tüketme, meyve ve sebze tüketimi ve düzenli egzersiz olduğu modeller için ikinci aşama regresyonu araç değişkenli probit regresyonu, bağımlı değişkeni VKİ olan modeller içinse ikinci aşama regresyonu tekrar EKK regresyonudur. Yani, kişinin VKİ'si ve eğitim seviyesindeki ilişkiyi ölçmek için İki aşamalı en küçük kareler (2EKK) yöntemleri kullanılmıştır. Ancak içsellik problemi ele alındığında, eğitim ve sağlık davranışları arasındaki ilişkinin değiştiği gözlenmiştir. Bu durum araç değişken tahmin sonuçlarının yerel ortalama işlem etkisi (Local Average Treatment Effect) ile açıklanabilir. Bu etkiye göre, araç değişkenin verdiği tahmin sonucu sadece araç değişkenden (yapılan eğitim reformlarından) etkilenen kişilerin eğitim seviyesindeki değişmenin bu kişilerin sağlık davranışları konusundaki etkisini göstermekte ve bu kişilerin eğer bu reformlardan yararlanmasalardı davranışlarının ne yönde olacağı hakkında fikir vermektedir. Yaptığımız araç değişken regresyon analizleri sonucunda, sigara ve alkol tüketiminin eğitim seviyesi ile birlikte azaldığı görülmüştür. Bununla birlikte meyve ve sebze tüketimi de eğitim seviyesindeki artışla birlikte azalmaktadır. Egzersiz yapma ve eğitim arasındaki içsellik ilişkisi, bazı spesifikasyonlarda doğrulanmamakla beraber, gene de araç değişken model sonuçlarına bakılmış ve burda da negatif fakat anlamlı olmayan bir ilişki bulunmuştur. Bu çelişkili sonuçların başka bir nedeni de EKK ve Probit modellerinde istihdam ve hanehalkı gelirini kontrol etmemiz, bu tanımlamalar da ise kontrol etmememiz olabilir. Sigara ve alkol tüketiminin eğitim seviyesiyle birlikte azaldığını görmek yüz güldürücü bir sonuçken, meyve ve sebze tüketiminin eğitim seviyesiyle birlikte azaldığını görmek düşündürücüdür. Son olarak kişinin VKİ sinin de eğitim seviyesi ile birlikte arttığı görülmüştür. Ancak bu çelişkili sonuç yedinci bölümde uygulanan kantil regresyon modelleriyle çözülmüştür.

Tezin sekizinci bölümünde kişinin VKİ'si ni belirleyen faktörler kantil regresyon modelleri ile incelenmiştir. Bu modellerin özelliği VKİ'nin farklı yüzdelik dilimlerinde her regresyonu tekrar yapmasıdır ve VKİ'yi belirleyen faktörleri VKİ'nin farklı yüzdelik dilimlerinde incelemesidir. Çalışmada VKİ'nin 5'lik, 10'luk, 25'lik, 50'lik, 75'lik, 90'lık ve 95'lik yüzdelik dilimlerinde analizler yapılmıştır. Bu analizlerde eğitimin VKİ üzerinde negatif bir etkisi olduğu ve bu etkinin yüzdelik dilimleri büyüdükçe arttığı gözlenmiştir. Bu sonuç, VKİ'nin en yüksek dilimlerinde eğitimin etkisinin çok yüksek olduğunu, dolayısıyla fazla kilolu veya obez olmanın tehlikelerinin yüksek eğitim seviyesine sahip olan kişiler tarafından daha anlaşılabilir olduğunu kanıtlamaktadır.

Kantil regresyon modellerinin başka bir değişik sonucu ise Türkiye'de kadınların erkeklerden obez veya fazla kilolu olmaya daha fazla eğilimli olduğudur. Yapılan analiz sonuçları cinsiyetin etkisinin yüzde 50 lik dilime kadar pozitifli anlamlı, Yüzde 75 lik dilimden sonra ise negatifli anlamlı olduğunu ortaya çıkarmıştır (Cinsiyet bir kukla değişkendir. Eğer kişi erkek ise bu değişkenin değeri 1'e, eğer kişi kadın ise değişkenin değeri sıfıra eşittir). Ayrıca, kantil regresyon sonuçları kentte yaşayan kişilerin kırsal kesimlerde yaşayan kişilere göre daha yüksek VKİ'ye sahip olduğunu göstermektedir. Ancak bu ilişkinin en yüksek yüzdelik dilimlerinde (90 ve 95) kaybolduğu ve ilişkinin anlamsıza dönüştüğü gözlenmiştir. Şehirlerde yaşayanlar daha çok hazır yemek ve fast-food tarzı ürünler tükettikleri için VKİ'lerinin daha yüksek çıkması beklenen bir sonuçtur. Obezite düzeyindeki VKİ dilimlerinde kırsal ve kentte yaşayan insanların arasında anlamlı bir fark olmaması ise kırsalda (özellikle köylerede) yaşayan kişilerin çok fazla beyaz ekmek, reçel ve pekmez gibi glisemik indeksi yüksek olan besinler tüketmesiyle açıklanabilir. Bir başka deyişle, tükettikleri besinlerin özelliklerinden dolayı kırsal kesimlerde yaşayan kişiler de obez olmaya yatkındırlar. Kişilerin VKİ'sin kişinin yaşlanmısıyla beraber artması ve bu ilişkinin konkav olması kantil regresyon analizlerinin verdiği bir diğer sonuçtur. Bunların dışında evli veya bekar olan kişilerin eşinden boşanmış veya eşi ölmüş kişilere göre daha normal seviyelerde VKİ'ye sahip olmaları, ayrıca istihdamda olan veya aktif olarak iş arayan kişilerin aktif olarak iş aramayan kişilere göre daha düşük VKİ seviyelerine sahip olmaları kantil regresyon analizlerinin önerdiği bir başka sonuçtur. Bu sonuçlar EKK tahmin yöntemleriyle elde edilen sonuçlara eş değerdir. Son olarak da tüm yüzdelik dilimlerde hane halkı gelir düzeyinin kişinin VKİ'sini anlamlı olarak arttırdığı, anlamlılık derecesinin en yüksek yüzdelik dilimlerde düştüğü ancak geçerliliğini hala koruduğu gözlenmiştir.

Tezin sekizinci bölümün ikinci kısmında kişinin eğitim seviyesi ve VKİ'si arasındaki içsellik problemi tekrar ele alınmış ve araç değişkenli kantil regresyon modelleri kullanılarak bu ilişki tekrar incelenmiştir. Araç değişkenli kantil regresyon modelleri bildiğimiz kadarıyla bu alanda yapılan bir çalışmada Türkiye için ilk defa kullanılmıştır. Araç değişkenli kantil regresyon modelleri uygulamalı ekonometri litratüründe yeni yeni kullanılmaya başlanmıştır ve henüz gelişmeye açıktır. Modellerde araç değişkenleri mutlak değerde minimize eden katsayılar regresyonlar 200 kere replike edilerek hesaplanmıştır. Kurduğumuz araç değişkenli kantil regresyon modellerinde ne yazıkki yakınsama gerçekleşememiştir. Bu sorunu çözmek için tezin bu bölümünde kullandığımız örneklemin büyüklüğü 18-24 yaş arası kişiler de eklenerek arttırılmıştır. Bu kişileri örneklemimize ekledikten sonra önceki kullandığımız araç değişkene ek olarak yeni bir araç değişken kullanma ihtiyacı doğmuştur. Çünkü, bu yaş grubundaki kişiler 1997 yılında yapılan yeni eğitim reformundan etkilenmişlerdir. 1997'de çıkarılan bir yasayla Türkiye'de daha önce beş yıl olan zorunlu eğitim süresi sekiz yıla çıkarılmıştır, ve yasa çıktığı yıl uygulamaya konulmuştur. Dolayısıyla o sene içinde ilkokul dördüncü sınıfta olan kişiler (1987 senesinde veya daha sonra doğanlar) sekiz yıllık zorunlu eğitime tabi tutulmuşlardır. Bu bilgi çerçevesinde çalışmanın bu bölümünde yeni bir araç değişken tanımlanmıştır. Bu araç değişken yine bir kukla değişkendir. Değişkenin değeri eğer kişi 1987 ve sonrası doğumluysa 1'e, daha önce doğmuşsa 0'a eşittir. Yapılan ilk aşama analizlerinde iki araç değişkenin fazladan idantifikasyon testlerini geçtiği, her iki araç değişkenin kişinin okula gitme senesini pozitif ve anlamlı olarak etkilediği ve F-istatistiklerinin yeterince yüksek olduğu gözlenmiştir. Genişleyen örneklemimizle kantil regresyon analizlerini tekrar yaptığımız zaman yakınsamanın VKİ'nin en yüksek yüzdelik diliminde gerçekleştiği görülmüştür. Bu dilimde kantil regresyon analizlerinin sonucuyla tutarlı bir şekilde kişinin okula gitme yılının kişinin VKİ'si ile negatif bir ilişki içerisinde olduğu sonucuna ulaşılmıştır.

Tez genel olarak eğitimin insan sağlığı üzerindeki olumlu etkilerini tespit etmiştir. Bu sonucun ışığında tezin son kısmında uygulanabilecek politikalar tartışılmıştır. Türkiye üzerine sağlık reformları ve sağlık harcamaları üzerine yapılan çalışmalar oldukça yenidir ve yapılan çalışmalar sınırlıdır. Yakın zamanda yapılan çalışmalrdan Kol (2015) Türkiye'deki sağlık reformlarını analiz etmiş; bir başka çalışmada ise Daştan ve Çetinkaya (2015) Türkiye'deki sağlık harcamalarının artışına dikkat çekmiş ancak somut bir politika önerememiştir. Eğitimin sağlık üzerindeki önemli etkisini gördükten sonra bu yönde uygun olan politika önermelerini yapmak tezin daha sonraki çalışmalara ışık tutması bakımından önemlidir. Ayrıca, bu tezde önerilen politikaların, sağlık reformlarının ve sağlık harcamalarının daha verimli bir şekilde kullanılmasına katkı sağlayacağı beklenmektedir. Kullanılan sağlık çıktısı (kişinin kendi rapor ettiği sağlık durumu) ve sigara, alkol tüketimi gibi sağlık davranışları uygulanacak politikalar için bizce anlamlı değişkenlerdir.

Dolayısıyla öncelikle, eğitim alanında reformların yapılmasının (kesintisiz eğitim, 12 yıl gibi) toplum sağlığını olumlu yönde etkileyeceği düşünülmektedir. Bulduğumuz sonuçlar kişilerin ne kadar çok okula giderse sağlıklarının daha iyi bir hal alacağına ve sağlık açısınndan riskli olarak kabul edilen davranışların görülme olasılığının azalacağına işaret etmektedir.

Sosyal politika alanında yapılan harcamaların öncelikle eğitim üzerine yapılmasının gerekliliği savunduğumuz bir başka savdır. Çünkü, eğitim seviyesi sağlık ile birlikte diğer sosyo-ekonomik değişkenleri (kişinin geliri gibi) de olumlu yönde etkilemektedir. Eğitime yapılan harcamları belirledikten sonra sağlık harcamalarının optimum seviyesini belirlemek politika yapıcılarının dikkat etmesi gereken bir konudur. Çünkü Daştan ve Çetinkaya'ya göre son yıllarda yapılan sağlık harcamaları artmış olmakla birlikte, bir çoğu gereksizdir. Harcamalarda optimum seviyeyi bulmak sağlık reformlarının da daha verimli uygulanması bakımından önemlidir.

Bu tezin önemli bir kısmını kişinin VKİ'sinin eğitim ile ilişkisini test eden analizler oluşturmuştur. VKİ obezitenin tespit edilmesi açısından önemli bir araçtır. Dünya Sağlık Örgütü (DSÖ)'nün tanımına göre eğer kişinin VKİ si 25'den büyükse kişi fazla kilolu, eğer VKİ si 30'dan büyük ise de kişi obez olarak adlandırılmaktadır. Obezite hem dünyada hem de Türkiye'de yaygın bir sağlık sorunu halini almıştır. Anket yıllarını kontrol ederek yaptığımız analizlerde kişilerin VKİ'lerinin yillar geçtikçe yükseldiği görülmüştür. Bu durum obezitenin de yıllar itibariyle artmasına bir işaret olarak gösterilebilinir. Obezitenin önüne geçebilmek için de öncelikle eğitime verilen önemin arttırılması gerekmektedir. Bunun için öncelikle liselerde sağlık bilgisi derslerine daha ağırlık verilmelidir, ve bu derslerde obezitenin ne kadar önemli bir sağlık sorunu olduğu, ne tip sağlık sorunlarına yol açabileceği özellikle vurgulanmalıdır. Okullarda ayrıca sağlıklı beslenmenin üzerinde önemle durulmalıdır. Özellikle okul öncesi çağlarda ailenin de sağlık eğitimi konusundaki önemi göz ardı edilmemelidir. Çocukların yüksek kalorili besinlerden uzak durmaları hakkında uyarmak hem aile hem de okuldaki öğretmenlerin başlıca görevlerinden biri olmalıdır.

Bir diğer önemli politika da büyükerin de obezitenin tehlikeleri konusunda bilgilendirilmeleridir. Bunun için yapılacaklardan biri, devlet obezite konusunda uyarıcı seminerler yapmaya önem vermesi olabilir. Tezin sonucunda kırsal kesimlerde de obezitenin yaygınlaştığı görülmüştür. Özellikle kırsal kesimlerde çalışan doktorlar ve öğretmenler çok fazla un, yağ, tuz ve şeker tüketiminin kişi sağlığı üzerindeki olumsuz etkileri hakkında devlet destekli seminerler vermelidir. Aynı şekilde obezite karşıtı kampanyalar da devlet ve sivil toplum örgütleri tarafından desteklenmelidir. Gelişmiş olan ülkeler için yapılan çalışmalar bu tip uygulamaların obeziteyi engelleme konusunda etkili olamadığını söylemektedir. Burdan gelişmekte olan bir ülke için bu uygulamalarının sonuçlarının çok daha geç meyve vereceği çıkarımı yapılabilinir. Ancak, en azından okul öncesi ve ilkokul çağlarında çocukları bu konuda bilgilendirmek ve çocuk yaşta önlem alınmasını sağlamak gereklidir.

Obezitenin önüne geçmek için eğitim kadar başka önlemlerin alınması da önemlidir. Örneğin, özellikle büyük şehirlerde neredeyse aynı mahallelerde bir fast-food zincirinden iki tane bulunmaktadır. Bu restoranların açılabilmesi için bir takım kotaların konulması gereklidir. Çünkü, fast-food yemekleri hem ucuz hem de doyurucudur, bu da özellikle öğrencileri daha çok fast-food restoranlarından yemeye itmektedir. Aynı şekilde marketlerden alınan ve evde mikrodalgaya konularak kolayca pişen yemekler özellikle çalışan çiftler tarafından oldukça tüketilen besin maddeleridir. Halbuki bu besinlerde de çok sayıda katkı maddesi ve çok fazla oranda yağ bulunmaktadır. Marketlerde satılan bu tip yemekler üzerindeki vergileri artırmak obezitenin önüne geçmek için başka bir önlem olarak sayılabilinir.

EKK yöntemleriyle yaptığımız analizler eğitimin düzenli egzersiz ve meyve-sebze tüketimi ile pozitif etkisini göstermiştir. Bu iki sağlık ile ilgili davranış kişinin

VKİ'sini normal seviyelerde tutabilmesi bakımından önemlidir. Okullarda bu konularla ilgili eğitime de önem verilmelidir. Özellikle enflasyonun yüksek olduğu dönemlerde Türkiye'de meyve ve sebze fiyatları çok yükselmektedir. Dolayısıyla, bu tip ürünlere tavan fiyat uygulaması getirilmesi her kesimden insanın meyve ve sebze tüketebilmesi açısından önemlidir. Türkiye'de spor merkezleri genelde özel sektörün elindedir. Dolayısıyla gelir seviyesi belirli bir aralığın üzerindeki kişiler daha fazla düzenli egzersiz yapabilmektedir. Okullarda spor olanakları genişletilmeli ve parasız olmalıdır. Aynı şekilde parasız egzersiz alanları devlet ya da ilgili belediyeler tarafından genişletilmelidir.

Şu andaki hükümet sigara ve alkolün kullanımının kısıtlanması konusunda çeşitli önlemler almaktadır. Örneğin, sigara ve alkollü içecekler üzerindeki vergiler yüksektir. Ayrıca, "Kamu Spotu" başlığı altında sigara ve alkolün zararlarını gösteren kısa televizyon reklamları sık sık yapılmaktadır. 2010 yılının başında belli kriterleri sağlayan açık alanlarda sigara içilmesi yasaklanmıştır ve bu yasağın daha da genişletilmesi beklenmektedir. 2013 yılından itibaren de gece 10:00 ve sabah 06:00 saatleri arasında alkol satışı yasaklanmıştır. Ancak bu yasakların sigara ve alkol tüketiminin azalması konusunda çok da etkili olmadığı görülmüştür. Bize göre, devletin bir takım yasaklar koymasından ziyade sigara ve alkol tüketiminin sağlığa zararları konusunda özellikle de okullarda daha çok bilgilendirme çalışmaları yapması daha etkili sonuçlar doğuracaktır. Aynı zamanda ilköğretim seviyesindeki çocuklar sağlık bilgisi derslerinde bu maddelerin zararları konusunda daha çok bilgilendirilmelidir.

Yukarıda bahsettiğimiz olası politikaların uygulanmasında devlet kadar sivil toplum örgütleri ve medya da önem teşkil etmektedir. Bu kuruluşların yapacağı kampanyalar, tanıtımlar ve aktiviteler kişilerin daha çok bilgilenmesi açısından önemlidir. Başka ne tür politikaların yapılabileceği tartışmaya açıktır ve daha fazla araştırma gerektirmektedir.

Tezin en son kısmında ise TSA veri seti kullanılarak başka hangi çalışmaların yapılabileceği tartışılmıştır. Öncelikle, şunu vurgulamak gerekir ki konu gelişmekte olan bir ülke için çok yeni bir alan olduğundan gelişmeye ve yeni sonuçlara açıktır. Tüm test sonuçlarında verdiği gibi sağlık ve eğitim arasında içsellik problemi vardır ve bu problemin çözülmesi için araç değişken teknikleri önemli bir metod olmaya devam edecektir. Bu çalşmada 25 ve yaş üstü kişiler için eğitim ve sağlık arasındaki ilişkiye bakıldığından 1960'ların başında kabul edilen eğitim reformları kişinin eğitim yılı yerine araç değişken olarak alınmıştır. Ancak, 1997'de yapılan eğitim reformlarının kişilerin okula gitme yılını arttırma konusunda çok daha etkili olduğu çeşitli çalışmalarda ve Milli Eğitim Bakanlığı istatistiklerinde gösterilmiştir. Ancak, bu reformlara tabi olan kişilerin çoğu henüz çok gençtir. Dolayısıyla, kötü veya çok kötü sağlık durumları rapor etmeleri beklenilmemektedir. Daha sonraki yıllarda hazırlanacak olan TSA veri setleri ile 1997 reformundan yararlanan kişilerin sağlık durumları ve eğitimleri arasındaki ilişkiyi incelemek ilginç bir çalışma olacaktır.

TSA'nın güzel bir veri seti olmasına rağmen bir takım handikapları vardır. Örneğin hanehalkı geliri dışında hane ile ilgili hiçbir değişken gözlenememektedir (Hanehalkının büyüklüğü, anne ve babanın eğitim durumu, anne ve babanın meslekleri vb...). Halbuki ailenin eğitim seviyesi, geçmişi gibi faktörler kişinin hem sağlık hem de sağlık ile ilgili davranışlarını açıklayan önemli belirliyicilerdir. Aynı zamanda kişinin kendi geliri de TSA veri setinde gözlenememektedir. Ayrıca, kişinin yaşadığı bölge sadece kent ve kır düzeyinde verilmektedir. Bu sebepten, bölgesel faktörler tam olarak incelenememektedir. Bu bilgilerin olması kişinin sağlığı ve istihdam durumu ile gelir düzeyiyle içsellik problemini çözmemizde araç değişken bulmaya yardımcı olabilir. Örneğin, gelir için bölgenin kişi başına düşen gayri safi yurt içi hasılası iyi bir araç değişken olabilir. Benzer şekilde, işsizlik için kişinin yaşadığı bölgenin işsizlik oranı iyi bir araç değişken olabilir. Son olarak da TSA ne yazık ki bir panel veri seti değildir. Bu yüzdem kişilerin sağlık durumlarının yıllar itibariyle ne yönde değiştiğini takip etmemiz imkansızdır. TSA panel şeklinde düzenlenip geçerli değişkenler de eklendiği zaman daha verimli çalışmaların yapılabileceği bir veri setine dönüşecektir.

Şunu belirtmekte fayda var: TSA tüm handikaplarına rağmen kişinin sağlığı ile ilgili zengin bir veri setidir. Bu tezde sadece kişinin sağlık durumu ve sağlık davranışları sağlık çıktısı olarak ele alınsa da hastalıkların, kişinin ne kadar koruyucu sağlık hizmetlerinden yararlandığının ve kişinin psikolojik durumunun eğitim ve diğer sosyo ekonomik faktörler ile ilişkisini incelemek başka bir ilginç çalışma olacaktır.

Son olarak, daha önce de belirtildiği gibi bu yönlü çalışmalar gelişmekte olan ülkeler için neredeyse hiç yapılmamıştır. Diğer gelişmiş ülkeler için TSA benzeri veri setlerini araştırmak ve incelemek ve bu ülkeler arasında karşılaştırma yapmak ülkelerin kendilerine özel özelliklerinin de toplum sağlığı üzerindeki etkisini incelemek gelişmekte olan ülkeler için yapılacak güzel bir çalışma olacaktır. Çünkü, kültür, gelenek, görenek, iklim, din ve ırk gibi faktörlerin de toplum sağlığı üzerinde etkilerinin olduğu düşünülmektedir.

C. CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Karaoğlan, Hanife Deniz
Nationality: Turkish (TC)
Date and Place of Birth: 1 January 1985 , Karşıyaka
Marital Status: Single
Phone: +90 312 210 3069
Fax: +90 312 210 22 97
email: kadeniz @ metu.edu.tr

EDUCATION

Degree	**Institution**	**Year of Graduation**
M.A	Bilkent University, Economics	2009
B.S.	Bilkent University, Economics	2007
High School	Izmir Çiğli Milli Piyango Anatolian High School	2003

WORK EXPERIENCE

Year	**Place**	**Enrollment**
December 2011-September 2015	METU, Department of Economics	Research Assistant
October 2011-December 2011	Hacettepe University, Department of Economics	Research Assistant
April 2010-September 2011	Yaşar University, Department of Economics	Research Assistant
December 2009-March 2010	T.C. Vakıflar Bankası	Researcher
September 2007-September2009	Bilkent University	Teaching Assistant

FOREIGN LANGUAGES

Turkish (Native), English (Excellent), German (Beginner), Italian (Beginner)

PUBLICATIONS

Gümüş, Gülüzar K., A.Tülay Yücel, **Deniz Karaoglan** and Şaban Çelik. (2011). "The Impact of Domestic and Foreign Macroeconomic News on Stock Market Volatility: İstanbul Stock Exchange". *Boğaziçi Journal Review of Social, Economic and Administrative Studies* 25(1):123-137

Dalgıç Başak, Burcu Fazlıoğlu and **Deniz Karaoğlan** (2015). "Entry to Foreign Markets and Productivity: Evidence from a Matched Sample of Turkish Manufacturing Firms*" Journal of International Trade and Economic Development: An International and Comparative Review* 24(5): 638-659. (SSCI)

Karaoglan, Deniz and Cagla Okten. (2015). "Labor force participation of married women in Turkey: A study of the added worker effect and the discouraged worker effect". *Emerging Markets Finance and Trade*. 51(1): 274-290. (SSCI)

AWARDS

Scientific Activity Support Scholarship, TÜBİTAK, 2014 (for 35th MEEA Conference)

Scientific Activity Support Scholarship, TÜBİTAK, 2011 (for 25th ESPE Conference)

Graduate Student Scholarship, TÜBİTAK, 2007-2009

M.A Fellowship, Bilkent University 2007-2009

B.A. High Honour Fellowship, Bilkent University 2004-2007

D. TEZ FOTOKOPİSİ İZİN FORMU

ENSTİTÜ

Fen Bilimleri Enstitüsü [ ]

Sosyal Bilimler Enstitüsü [X]

Uygulamalı Matematik Enstitüsü [ ]

Enformatik Enstitüsü [ ]

Deniz Bilimleri Enstitüsü [ ]

YAZARIN

Soyadı : KARAOĞLAN
Adı : HANİFE DENİZ
Bölümü : İKTİSAT

TEZİN ADI (İngilizce) : Essays on the Education Gradient of Health in Turkey

TEZİN TÜRÜ : Yüksek Lisans [ ] Doktora [X]

1. Tezimin tamamından kaynak gösterilmek şartıyla fotokopi alınabilir. [X]
2. Tezimin içindekiler sayfası, özet, indeks sayfalarından ve/veya bir bölümünden kaynak gösterilmek şartıyla fotokopi alınabilir. [ ]
3. Tezimden bir (1) yıl süreyle fotokopi alınamaz. [ ]

TEZİN KÜTÜPHANEYE TESLİM TARİHİ: