EMPIRICAL ANALYSIS OF HEALTH AND EDUCATIONAL ATTAINMENT IN TURKEY

Dilek KILIC

A thesis submitted to the University of Sheffield for the Degree of Doctor of Philosophy
in the Department of Economics

Date Submitted: June 2012

ABSTRACT

The overall aim of this thesis is to examine and analyse three separate yet related issues in health and education that have been attracting increasing attention on the policy agenda in Turkey, especially over the recent decade. Public policy attention has particularly been devoted to issues such as financial protection against health care costs, decreasing the smoking prevalence rate and closing the gender gap in educational attainment. One of the main reasons behind the interest of policy makers in these areas may be attributed to need of the country to enhance human capital as well as to its aim to fulfill the requirements for EU membership. This provides the motivation for the empirical analyses presented in this thesis which investigate these issues in depth and, where relevant, provide policy implications.

The first empirical study presented in Chapter 2 investigates the prevalence of 'catastrophic' out-of-pocket health care expenditure in Turkey and identifies the factors which are associated with its risk. The results indicate that poverty is an important barrier to seeking health care and that poverty is inversely associated with the risk of incurring catastrophic health expenditure. Chapter 3 examines the determinants of adult smoking propensity and intensity from a gender perspective. The findings of this chapter are twofold. First, the factors associated with smoking participation differ from the factors associated with the level of cigarette consumption. Second, the determinants of smoking behaviour are found to be different across males and females. The third empirical study is presented in Chapter 4 and aims to examine the educational attainment of girls and boys at the primary and secondary education levels. The findings indicate that the determinants of educational attainment and the potential factors that are associated with gender inequality differ by the level of education.

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to both of my supervisors, Sarah Brown and Arne Risa Hole for their precious time, expert guidance and unwavering support. This thesis would not have been completed without their professionalism and encouragement. I am also grateful to Jennifer Roberts for her valuable comments and suggestions. I would like to thank Steven McIntosh for his help and interest during this process as well as his valuable suggestions and comments. My thanks also go to the Department of Economics of the University of Sheffield for the warm welcome and their help anytime I needed it.

I would like to thank my family for their unconditional love and encouraging me in this tough process. I would also like to thank Selcen Turanli, who provided her unique friendship at every stage of this PhD process, for her unconditional support. I thank Selcen for many hours of patient listening and for her help and useful comments on this thesis. I would like express my thanks to my friends, Onur Yeni and Derya Guler Aydin for their warm support they gave over this time. My special thanks go to Sureyya and Metin Mehmet for their friendship and hospitality.

Most of all, I would like to thank Seyfi Kilic, whose love and belief in me made this all possible, for his encouragement, sacrifices and understanding.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

2SLS: Two-stage Least Squares

AHRQ: Agency for Healthcare Research and Quality

Bag-Kur: Social Insurance Agency of Merchants, Artisans and the Self-employed

BK: Bag-Kur

BRFS: Behavioural Risk Factor Survey

CCTP: Conditional Cash Transfer Program

CDC: Centre for Disease Control and Prevention

CPI: Consumer Price Index

CSFII: Continuing Survey of Food Intakes by Individuals

DHS: Demographic and Health Survey

EC: European Commission

EU: The European Union

FCTC: Framework Convention on Tobacco Control

GATS: Global Adult Tobacco Survey

GCE: General Certificate of Education

GCSE: General Certificate of Secondary Education

GDP: Gross Domestic Product

GERF: Government Employees Retirement Fund

GNP: Gross National Product

GYTS: Global Youth Tobacco Survey

HBS: Household Budget Survey

HDI: Human Development Index

HICES: Household Income and Consumption Expenditure Survey

HTP: Health Transformation Programme

IIA: Independence of Irrelevant Alternatives

ILO: International Labour Organisation

IV: Instrumental Variable

LR: Likelihood Ratio

MoH: Ministry of Health

NB: Negative Binomial

NHIS: National Health Interview Survey

NSFH: National Survey of Families and Households

NTCP:	National Tobacco Control Programme
OECD:	The Organisation of Economic Co-operation and Development
OLS:	Ordinary Least Squares
OSYM:	Student Selection and Placement Centre
PCA:	Principal Component Analysis
PRILIF:	Programme of Research into Low Income Families
PSU:	Primary Sampling Units
SHS:	Second-hand Smoke
SPO:	State Planning Office
SSI:	Social Security Institute
SSK:	Social Insurance Organisation
TFSS:	Turkish Family Structure Survey
TurkStat:	Turkish Statistical Institute
UHI:	Universal Health Insurance
UK:	The United Kingdom
UNDP:	United Nations Development Programme
US:	The United States of America
USDA:	United States Department of Agriculture
VCSE:	Vocational Certificate of Secondary Education
WHO:	World Health Organisation
YOK:	Higher Education Council of Turkey
ZI:	Zero-inflated
ZINB:	Zero-inflated Negative Binomial
ZIP:	Zero-inflated Poisson

CHAPTER 1: INTRODUCTION

1.1 Aims and Motivation of the Thesis

Health and education are two areas of economic research that have increasingly been on the political agenda in recent years in both developing and developed countries. As is well documented in the existing literature, health and education, which are arguably the most important elements of human capital, not only play an important role in social development but also contribute to economic growth through a more qualified labour force (Schultz, 1961; Becker, 1962). In this respect, the large differences in economic growth rates across countries may reflect the differences in health and education. To be specific, growth in economic output that is not explained by increases in inputs of physical capital can be attributed to improvements in the population's health and education (Schultz, 1961). One might argue that better health improves human capital and enhances workers' labour market outcomes. Similarly, the education level of society is also an indicator of the development of human capital since a labour force that has a high level of education is likely to be more productive.

In this context, human capital is particularly important for Turkey, which is a developing country that aims to stabilise the economy and increase economic growth and development. Additionally, Turkey started the accession negotiations for full membership of the EU in October, 2005. In the ninth development plan that covers the 2007-2013 period, the State Planning Office (SPO) stated the future vision of Turkey as follows: "Turkey, a country of information society, growing in stability, sharing more equitably, globally competitive and fully completed her coherence with the European Union (p. 11)." (SPO, 2006). This aim can be achieved since Turkey is among the twenty largest economies in the world and has a growth rate of over 8 percent annually (OECD, 2010). However, it can be argued that none of these aims can be fully achieved or sustained without enhancing the country's human capital. In this respect, in the reports on the progress made by Turkey in preparing for EU membership, it is stated that Turkey should invest more in upgrading the country's human capital to reach the level of the EU member countries (EC, 2008; EC, 2010). Turkey's Human Development Index (HDI) value in 2010 was 0.679, which puts Turkey well behind all EU member states (this value was 0.890 for the Netherlands, 0.885 for Germany, 0.863 for Spain and 0.849 for the UK) as well as below the OECD average, which was 0.879

in 2010. Furthermore, Turkey ranks the lowest among the five EU candidate countries (the HDI value was 0.767 for Croatia, 0.701 for the Former Yugoslav Republic of Macedonia, 0.869 for Iceland and 0.769 for Montenegro) (Human Development Report, 2010). These figures suggest that government policies should focus on health and education to enhance human capital in Turkey.

In recent years, attention has been devoted to three specific policy areas in health and education in Turkey, which provide the motivation for the research presented in this thesis. Firstly, for many years, the Turkish health care system has been problematic which is mostly due to its complex structure. Since a large number of institutions are included in the decision-making and implementation processes and each social security scheme provides different benefit packages both in terms of depth and breadth of financial protection, the health care system arguably fails to effectively meet the health needs of the country (Tatar and Kanavos, 2006). Although the Health Transformation Programme, which has been in place since 2003, has resulted in some improvements in this complex structure by unifying all insurance schemes under one scheme, a large proportion of the population still do not have adequate financial protection. However, financial protection against high levels of out-of-pocket health care expenditure is a strong determinant of seeking health care and those without any form of financial protection or those who have less financial resources to pay for health care costs are less likely to seek health care or they tend to delay meeting medical care needs. This may result in more severe illnesses and higher health care costs in the future, and ultimately increased mortality. Therefore, it is important to focus on delivering health care services in an effective and efficient way to obtain a more healthy population and higher health status indicators.

The area of financial protection is also emphasised in the process of Turkey's membership to the EU in the European Commission (EC) reports and it is stated that responding to the health needs of the population on the basis of accessibility is one of the most important problems that the country must address and efforts should be directed towards protecting vulnerable groups against health care costs (EC, 2008). Similarly, in the 2008 OECD report, evaluating the performance of the Turkish health care system in terms of its fairness in financing and identifying the risk factors related to incurring high out-of-pocket health care expenditure are highlighted as important research areas (OECD, 2008). In this context, the aim of the first empirical study

(Chapter 2) is to investigate the prevalence of 'catastrophic' out-of-pocket health care expenditure in Turkey and to identify the factors which are associated with its risk. However, it may be the case that many poor households may delay dealing with their medical needs as they cannot afford out-of-pocket health expenditure and, thus, they may be seen as not incurring catastrophic health expenditure. Therefore, another aim of this chapter is to investigate the determinants of catastrophic out-of-pocket health care expenditure in the context of attempting to adjust for the medical care seeking behaviour of households.

Secondly, Turkey has arguably attached increasing importance to anti-smoking policies in recent years in order to decrease the high smoking prevalence rate and to fulfil the EU membership requirements. The Anti-Tobacco Law, which was the first legal regulation controlling the consumption of tobacco, was enacted in 1996. Furthermore, a comprehensive tobacco regulation law, which includes a set of measures such as a ban on smoking in all indoor public areas and a ban on tobacco advertising, was adopted in 2008. A key reason behind the interest of policy makers in this area is that Turkey has an alarmingly high rate of smoking prevalence, nearly one-third (30%) of adults aged 15 and over (47.7% among males and 14% among females) were current smokers in 2008 (Global Adult Tobacco Survey (GATS), 2008). Furthermore, smoking attributed diseases are the highest risk factors for deaths among males and the second most common cause of all deaths in Turkey (Yurekli *et al.*, 2010). The diseases caused by smoking constitute a considerable burden on Turkey's health care system. The deaths and diseases caused by tobacco smoking also mean a loss in productive labour and this leads to a decrease in the human capital accumulation of the country. Additionally, there has been growing attention paid by policy makers to curb the increase in female smoking in Turkey since the smoking prevalence rate of females has increased sharply in recent years, by nearly 40% between 1997 and 2009 (Yurekli *et al.*, 2010). Therefore, special attention should be devoted to investigating the reasons behind this increase and public policies should be designed to prevent further increases in smoking prevalence among females. In this respect, the second empirical study (Chapter 3) in this thesis examines the potential factors associated with both smoking participation and the level of cigarette consumption in Turkey for males and females separately.

Finally, in addition to policy reforms in health, another important policy area in Turkey relates to the educational attainment of children. It is well known that similar to health,

education also plays an important role in both economic growth and development by improving the productive potential of workers. In addition, improvements in the education of girls have non-market benefits, such as better child nutrition, increased bargaining power in the household, fertility reduction and lower levels of infant mortality (Strauss and Thomas, 1995). Therefore, gender inequality in educational attainment has been accepted as one of the major barriers to social and economic development for developing countries by international development agencies (United Nations Development Program (UNDP), 1995). In this respect, in the EC report (2010), it is stated that there has been limited progress in increasing the educational attainment of children and that gender inequality in educational attainment remains one of the most important problems in Turkey.

The extension of compulsory primary education from five to eight years in 1997 has increased the total enrolment rate for primary education (from 84.7% in the 1997/1998 academic year to 96.5% in the 2008/2009 academic year). However, boys have slightly higher enrolment rates than girls (90.2% for boys and 78.9% for girls in the 1997/1998 academic year and 97% for boys and 96% for girls in the 2008/2009 academic year) (National Education Statistics, 2009). This trend indicates that the education reforms have served to close the gender gap significantly in the primary education enrolment rates. There figures, however, reflect the current enrolment status of individuals at primary school rather than the rate at which individuals complete primary education. Perhaps most importantly, the net enrolment rate for secondary education is lower than the primary school enrolment rate and there is greater gender inequality in secondary school enrolment rates as compared to primary education. The net enrolment rates for secondary education were 60.3% for boys and 56.3% for girls in the 2008/2009 academic year (National Education Statistics, 2009). Although it appears that there is not a considerable gender difference in the enrolment rates of primary and secondary education, these figures do not reflect the number of individuals who have completed primary and secondary education. Dropping out of school is a common phenomenon in Turkey especially among girls and there are significant gender differences particularly in the secondary school educational attainment in the eastern and south-eastern parts of the country (EC, 2010; Goksen *et al.*, 2006). Moreover, the female labour force participation rate (29.8% in 2011) and the access of females to education are the lowest in Turkey among the EU states and the OECD countries (EC, 2010; TurkStat, 2011). In

this context, identifying the factors underlying household decisions regarding the education of children may be informative for policy-makers. Thus, the aim of the third empirical study (Chapter 4) is to investigate the determinants of the educational attainment of girls and boys in Turkey at the primary and secondary education levels.

1.3 Structure and Content of the Thesis

Chapters 2, 3 and 4 present the empirical studies of the three separate yet related issues relating to the aims and motivations set out above. Each of the empirical studies has the same structure and constitutes a self-contained chapter while Chapter 5 provides an overall conclusion of the thesis. The contents of Chapters 2, 3 and 4 are summarised below.

1.3.1 Chapter 2

The empirical analysis in Chapter 2 investigates the prevalence of 'catastrophic' out-of-pocket health expenditure and the factors associated with its risk at the household level. Out-of-pocket health expenditure is defined as 'catastrophic' if the ratio of the household health expenditure to total household expenditure or non-food expenditure exceeds a pre-defined threshold level. Since there is no agreed threshold value for defining catastrophic health expenditure, a number of threshold values are used in order to explore the robustness of the results. The analysis in this chapter uses the Household Budget Surveys (HBS) for Turkey from 2002 to 2008, which are a time-series of repeated cross-sections. A standard probit model is estimated for the pooled data set as a first step in order to explore the association between poverty and experiencing catastrophic health expenditure as well as other risk factors. However, if the household expects health care costs to be unaffordable and, thus, chooses not to seek health care, there could be a sample selection problem. In this respect, the Sartori selection model (Sartori, 2003) and the Heckman selection model (Heckman, 1979) are also estimated in order to control for the potential selection problem and to further explore the robustness of the results.

The results indicate that poor households are less likely to seek health care relative to non-poor households and that there is a negative association between poverty and incurring catastrophic health expenditure even after adjusting for the potential selection bias. The results further indicate that households with a disabled or ill member and households with more preschool or elderly members are more likely to seek health care

and are more likely to experience catastrophic health expenditure. Higher levels of education, living in an urban area and insurance coverage are all found to be protective factors against the risk of catastrophic health expenditure.

1.3.2 Chapter 3

Chapter 3 examines gender differences in cigarette consumption in Turkey using the Global Adult Tobacco Survey (GATS) for the year 2008. This survey is the most comprehensive individual level survey focusing on tobacco consumption for adults in Turkey. This chapter also aims to investigate the potential distinction among the factors affecting the decision to smoke (i.e., smoking participation) and how much to smoke (i.e., the level of consumption). With this aim, count data models, namely poisson, negative binomial (NB) and zero-inflated models, are estimated. The results of the Vuong and Likelihood-Ratio (LR) tests indicate the zero-inflated negative binomial (ZINB) model as the preferred specification. In addition, the two-part/hurdle count data models are estimated to explore the robustness of the findings.

The main finding of this chapter is that the factors associated with smoking participation differ from the factors associated with the level of consumption for both genders. This implies that policy-makers should consider both the propensity and intensity dimensions of smoking behaviour, which are associated with different factors, while designing anti-smoking policies. The findings also highlight the potential effectiveness of specific anti-smoking policies for males and females since the determinants of smoking behaviour are found to be generally different between males and females. Using a cigarette tax policy may be more effective in decreasing smoking prevalence among females since females are found to be more responsive to changes in cigarette prices as compared to males. However, public awareness programs on the health risks of tobacco consumption and tobacco advertising bans may also be used as policy instruments to decrease the cigarette consumption of males.

1.3.3 Chapter 4

Chapter 4 investigates the effects of parental, household and community characteristics on the educational attainment of girls and boys in Turkey at the primary and secondary education levels, separately. The analysis in this chapter uses the Household Budget Survey (HBS) for the year 2003 which is the most comprehensive household survey for Turkey. The most important reason for using the 2003 HBS is that it differs from the

other HBSs by providing information on the residence of the household at the regional and province level which enables the exploration of the association between community characteristics and educational attainment. In order to analyse gender differences in educational attainment, ordered probit models are estimated as a first step. However, children from the same household are likely to share common unobserved family characteristics that cannot be measured by the data and this may bias the estimation results (Lillard and King, 1984). Therefore, random effects ordered probit models are also estimated to allow for this.

The findings indicate that the determinants of educational attainment and the potential factors that are related to gender inequality differ by the level of education. Higher levels of parental education and income are found to be the main determinants of both primary and secondary school educational attainment for both genders. Moreover, the number of very young siblings (aged 0-5) in the household and community characteristics are the main factors associated with the schooling of girls, particularly at the secondary school level, whilst having a self-employed father or having a father working in agriculture are among the key determinants of the primary school educational attainment of boys. Another main finding relates to the econometric methodology employed in this chapter since some differences are identified between the results of the ordered probit and the random effects ordered probit models, which highlights the importance of accounting for unobserved family characteristics in the empirical analysis.

CHAPTER 2: THE DETERMINANTS OF 'CATASTROPHIC' OUT-OF-POCKET HEALTH CARE EXPENDITURE IN TURKEY

2.1 INTRODUCTION

In the 2000 World Health Organization (WHO) report, it is stated that financial protection against 'catastrophic' health expenditure should be one of the primary goals in designing health sector reform strategies (WHO, 2000). Out-of-pocket health expenditure is defined as 'catastrophic' if it is more than a 'reasonable proportion of income' or it causes a household to fall into poverty. The premise is that households should not spend more than a specific percentage of their income on health care to maintain other basic needs (Wagstaff and Doorslaer, 2003).

The impact of health care financing systems on the welfare of households, particularly poor households, is regarded as an important issue by health policy makers in developing health systems and insurance mechanisms (Xu *et al.*, 2003). However, out-of-pocket health care expenditure plays a dominant role in health care financing systems in a number of countries such as Korea, Mexico, Switzerland, Hungary and Poland. The share of out-of-pocket health expenditure in total health expenditure was 35% in Korea, 49.3% in Mexico, 30.8% in Switzerland, 23.9% in Hungary and 22.4% in Poland in 2008 (OECD, 2010).

Turkey is an interesting country for investigating issues concerning financial protection due to several reasons. First, the Turkish health care system has been restructured with the Health Transformation Programme (HTP) and health reforms since 2003 which is mostly due to its complex structure. There are a large number of social security schemes, which provide different benefit packages both in terms of depth and breadth of financial protection, and this arguably results in inequitable distribution of health care services (Tatar and Kanavos, 2006). The Universal Health Insurance (UHI) system, which unifies all public insurance schemes under the one scheme, was implemented in 2008. However, implementation of unifying all insurance schemes has not been completed and a large proportion of the population still do not have adequate financial protection. In this respect, one of the main aims of the HTP is to ease the burden on households by decreasing the proportion of out-of-pocket health expenditure in total health expenditure. This share has decreased from 29% in 1999 to 22% in 2007, but it is

still quite high relative to the EU countries (TurkStat, 2010). This share was 7.4% in France, 13% in Germany, 5.7% in Netherlands, 11.1% in the UK and 15.1% in Austria in 2008 (OECD, 2010). Second, the Turkish health care system is in alignment with that of the EU countries (OECD, 2008) and the area of financial protection in terms of responding to the health needs of the population on the basis of financial accessibility is among the most important issues in the EU membership negotiations (EC, 2008). Third, despite the recognition that financial protection against high levels of out-of-pocket health expenditure is an important issue for Turkey, there are only a few studies that have focused on identifying the risk factors associated with catastrophic health expenditure. However, as stated by Xu *et al.*, 2003, it is important for policy makers to know which characteristics make households more vulnerable to experiencing catastrophic health expenditure when developing the health policy.

In this context, the aim of this chapter is to investigate the prevalence of catastrophic health expenditure in Turkey and to identify the factors which are associated with its risk at the household level using the Household Budget Surveys (HBS) from 2002 to 2008. In the existing literature, there is no agreed threshold value for defining catastrophic health expenditure. Therefore, in order to explore the robustness of the results, using a number of threshold levels is regarded as appropriate for the definition of catastrophic health expenditure rather than specifying a single threshold level. Catastrophic health expenditure is defined at five threshold levels, 2.5%, 5%, 10%, 15% and 20% for total expenditure, and five threshold levels, 10%, 15%, 20%, 30% and 40% for non-food expenditure in this study and the choice of these threshold levels is based on the earlier studies in this area which are discussed in detail in the literature review presented in Section 2.3 below.

The impact of out-of-pocket health expenditure, on the other hand, cannot be fully captured by analysing catastrophic health expenditure. Many poor households may choose not to seek health care at all since they feel that health care costs are unaffordable for them and this may lead to a sample selection problem (Xu *et al.*, 2003). In this case, the households that choose not to seek health care are excluded from the analysis. Therefore, poor households may be seen as less likely to experience catastrophic health expenditure as compared to non-poor households if the potential selection problem is not considered in the analysis. However, investigating this sample selection problem may produce different results for Turkey since Turkey's health care

system has interesting aspects, which are discussed in detail in Section 2.2. One of them is that it is very common among rich households in Turkey to use private facilities and public health services are often regarded as being unsatisfactory. After the health reforms, access to private facilities was improved for patients from all social insurance institutions via increased contracts with private facilities to deliver outpatient and inpatient health services (OECD, 2008). However, it is still the case that using private facilities requires an extra charge imposed by the private provider. This improvement has arguably benefited particularly the non-poor segment of the population, who can afford the extra charge imposed by the private provider. Since this extra cost was reduced by the recent health reforms, non-poor households' interest in private health care may create demand inducement. Therefore, it may also be the case that non-poor households in Turkey are more likely to incur catastrophic health expenditure than poor households even after accounting for the medical care seeking behaviour of households. In that regard, in the 2008 OECD report on the performance of the Turkish health system and its determinants, it is stated that the Turkish health system can seem to be equitable and progressive but this could also reflect the fact that rich households spend a large amount of their budget on private health care services due to public services being regarded as poor quality while the poor households delay their medical needs as they cannot afford them. This problem is highlighted as an important research area for the evaluation of the Turkish health care system (OECD, 2008).

Although the potential selection problem may bias the estimation results, most of the existing studies ignore households that do not seek treatment and this measurement problem is accepted as a limitation (Russell, 2004; Xu *et al.* 2003). In this respect, the most important contribution of this study to the existing literature is that, it is the first attempt to investigate the determinants of catastrophic health expenditure in Turkey in the context of attempting to adjust for the medical care seeking behaviour of households. Furthermore, this study extends the existing literature by using the Sartori selection model to investigate the implications of this potential selection bias (Sartori, 2003). The most important feature of the Sartori selection model, which is discussed in detail in Section 2.4.2 below, is based on the assumption of identical error terms in both the selection and outcome equations which relaxes the exclusion restriction in the Heckman selection model. However, the Heckman selection model is also analysed in the chapter to explore the robustness of the results from the Sartori selection model.

The rest of the chapter is structured as follows; Section 2.2 provides background information on the Turkish health system and information related to out-of-pocket health expenditure in Turkey. Section 2.3 evaluates the background literature on defining and measuring catastrophic health expenditure as well as its determinants. Section 2.4 describes the data and the variables used in the analysis as well as the key descriptive statistics and the methodology employed in this chapter. Section 2.5 discusses the results and summarises the key findings and Section 2.6 discusses the main findings and policy implications, the shortcomings of the analysis and potential directions for future research.

2.2 THE TURKISH HEALTH CARE SYSTEM

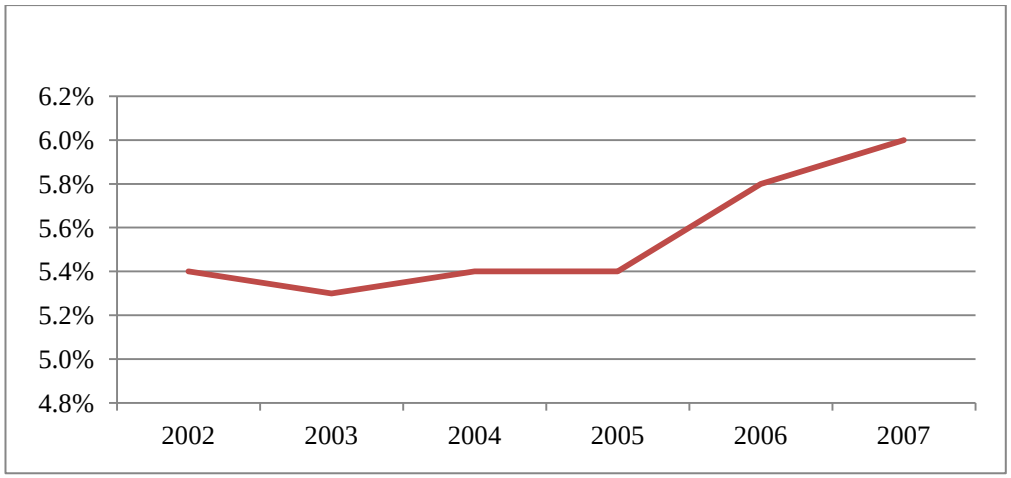
Turkey is an upper-middle income country in Europe and the Middle East, with a population of 73.7 million and per capita gross national product (GNP) of $14,580 in 2010 (TurkStat, 2010; World Bank, 2010). The Turkish economy is one of the 20 largest economies of the world. Turkey is also a candidate for membership of the EU. However, the population's overall health status and the structure and quality of the health care system are poor relative to the country's general level of development. Life expectancy at birth stands at 74 years for women and 70 years for men, with an overall average of 72 years for the whole population. This is well below the EU average life expectancy at birth of 82 years for women, 75 years for men and 79 for the total population. The latest estimates of infant mortality per 1000 live births is 23.9, which is much higher than the average for Europe of 5.6 per 1000 live births (World Factbook, 2010).

On the other hand, inequality in health status is an important problem in Turkey. There is a significant difference in health status indicators between western and eastern Turkey and between rural and urban areas. It was reported that the infant mortality rate was 22 per 1,000 live births in the west compared with 41 per 1,000 live births in the eastern part of the country in the 2003 Demographic and Health Survey (DHS, 2003; Tatar and Kanavos, 2006). Reducing such inequalities among geographical regions and between rural and urban areas has been one of the major aims of every five-year plan since 1963 (Savas *et al*, 2002). Moreover, Turkey has a sizeable poor population; the rate of poverty is 17.11% in 2008 which is calculated with a measure of absolute poverty (TurkStat, 2009). In this respect, poor people are also at greater risk of having poor health due to the lack of any health insurance coverage for any health expenditure they incur. The wide gaps in health status between the poor and non-poor and between regional areas (and urban and rural areas) are among the major health care challenges facing Turkey's health care system.

Since the late 1980s, economic liberalisation has led to a great extension of the private health care sector in Turkey. In addition, there is a positive trend, which can be seen in Figure 2.1 below, in the proportion of total health expenditure in gross domestic product (GDP). Total health expenditure accounted for 6% of GDP in 2007. This proportion was 4.8% in 1999, 4.9% in 2000 and 5.2% in 2001. While per capita health expenditure was

78.7 Turkish Liras (186 USD) in 1999, it had increased to 724.6 Turkish Liras (553 USD) in real terms in 2007 (TurkStat, 2010).

Figure 2.1: Proportion of Total Health Expenditure to Gross Domestic Product



Source: TurkStat, 2010.

The WHO European Region Committee accepted a new health policy framework in 1998. Turkey has also accepted the regulations of this health policy and designed health reforms to align the Turkish health care system with that of the EU and the OECD countries (OECD, 2008; Sulku and Bernard, 2009). With this aim, the 'Health Transformation Programme' (HTP) has been implemented since 2003.

The HTP was designed as a ten-year reform package covering the period 2003-2013 to address long-standing problems such as low insurance coverage rates, urban/rural and regional inequalities in health status, fragmentation in the financing and delivery of health care, inequitable distribution of health services and health care personnel, low quality of care and high levels of out-of-pocket health care expenditure in the Turkish health care system (OECD, 2008). Furthermore, the HTP of the Ministry of Health (MoH) includes the implementation of a Universal Health Insurance (UHI) system (also called as the General Health Insurance System), which unifies all insurance schemes under the Social Security Institute (SSI).

The main institutional and organisational changes designed under the UHI are that enrolment in UHI is compulsory and that contribution rates are proportional to ability to pay. If individuals cannot pay the premiums, the premiums are paid from the national budget on the basis of a means-tested system. All beneficiaries are entitled to the same benefit package. The SSI is the single purchaser in the health sector which sets up

contracts with public and private facilities to deliver health care (OECD, 2008). In April 2008, the legislation was passed and implementation of UHI began.

Prior to the HTP reforms, Turkey's health care system was centralised but fragmented and highly complex. The MoH was the main primary and secondary health care provider through its own primary health care facilities and hospitals and the only provider of preventive health services. Besides the MoH, the Social Insurance Organisation (SSK) and universities were the other main public providers. Each medical school has its own university hospital and they were the major providers of tertiary care but their share in the overall delivery system was small. The SSK, which was founded in 1945, provided health care services through its own 120 hospitals and other health facilities (Savas *et al*, 2002; Sulku and Bernard, 2009).

In the liberalisation period, the late 1980s, the government provided considerable incentives for private hospital investment. At the end of the 1990s, there were 100 new hospitals focused in the largest cities. In 2007, there were 365 private hospitals which accounted for 28.6% of all hospitals. However, the private sector constituted 7.2% of total hospital beds (TurkStat, 2007). Most of the doctors working for public agencies were also allowed to work in their private clinics after office hours (Tatar *et al*, 2007). Working for both public and private facilities is common since public facilities are seen as unsatisfactory and patients think that they can obtain a better service from private facilities. Before the health reforms, patients visiting private facilities were paying for services out-of-pocket even if they have a health insurance coverage (Savas *et al*, 2002).

Before the HTP reforms, the health care financing system in Turkey was also complicated and fragmented by the high number of institutions involved in providing and financing health care (Savas *et al*, 2002). Turkey had three main social security schemes, each to provide different benefit packages both in terms of depth and breadth of health insurance coverage: the SSK, the Government Employees Retirement Fund (GERF) and the Social Insurance Agency of Merchants, Artisans and the Self-employed (Bag-Kur).

The SSK insured private sector employees, blue-collar public sector employees and their dependants and provided health services through its own facilities. SSK beneficiaries were allowed to use MoH and university facilities with referral. Its members were entitled to benefits covering all inpatient and outpatient facilities. Health

services provided by the SSK were financed almost entirely through employees and employers' contributions. Additional sources of funding included fees paid by non-members of the SSK and co-payments for part of the outpatient drug costs[1] (Savas *et al*, 2002; Sulku and Bernard, 2009).

Bag-Kur (BK) was the insurance scheme for self-employed individuals. In contrast to the SSK, Bag-Kur did not operate its own health facilities. Beneficiaries of BK were entitled to benefits which cover all outpatient, inpatient treatment and pharmaceuticals[2]. Bag-Kur provided health services through health organisations which were contracted such as the MoH, SSK facilities, university and private hospitals and non-governmental organisations and pharmacies (World Bank, 2003).

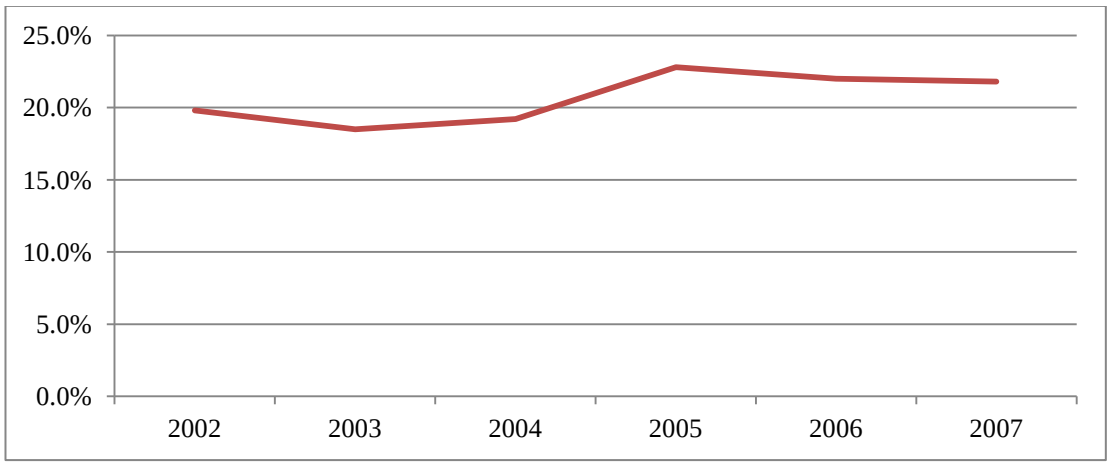
The GERF covered retired civil servants and their health services were funded by the contributions from members while the health expenditure of active civil servants and their dependents was financed by the national budget. The GERF covered all health expenditure of outpatient and inpatient care including medical and non-medical services. Like other insurance schemes, there was a requirement to pay a co-payment of 10% for the cost of drugs. GERF members were allowed to use all types of health facilities such as public facilities, universities and the private facilities. Active civil servants were also entitled to use public facilities and the private facilities with referral. In addition to those schemes, the Green Card scheme, established in 1992, provided free access to health services for poor people earning less than a minimum level of income which is defined by law and was directly financed by the national budget[3]. Green Card members were allowed to benefit from the outpatient and inpatient care at MoH hospitals and university hospitals with referral. This scheme paid for inpatient drugs but did not pay for outpatient drugs (Savas *et al*, 2002; Sulku and Bernard, 2009).

Apart from the general government budget and social security contributions, out-of-pocket expenditure, which refers to direct payments to private facilities, insurance premiums paid for voluntary health insurance and co-payments for drugs and health services, is also one of the main sources of health care financing. High levels of out-of-

[1] 20% for active members and their dependents and 10% for pensioned members.
[2] BK members were required to pay a co-payment of 20% for active members and their dependants and 10% for pensioned members. BK did not pay for general consultations or the purchase of prostheses .
[3] The national budget is financed by tax revenue and allocated primarily to the MoH, university hospitals, the Ministry of Defence, other public agencies and the health care expenditure of active civil servants (Savas *et al*., 2002).

pocket health expenditure, however, violate the vertical equity principle, which requires that payment should be related to ability to pay (WHO, 2000).

Figure 2.2: Proportion of Out-of-Pocket Health Expenditure to Total Health Expenditure



Source: TurkStat, 2010.

The recent share of out-of-pocket health expenditure in total health expenditure (Figure 2.2) suggests that out-of-pocket health expenditure accounts for 21.8% of total expenditure on health care in 2007, rising from 19.8% in 2002. It can be, however, argued that out-of-pocket health expenditure may be higher than this figure indicates. There are several possible reasons for this.

First, there was a great extension in the private health care sector at the end of the 1980s. During this period, not only did the number of private enterprises increase, but also the service amount of each enterprise increased rapidly. According to Savas *et al.* (2002), it is commonly accepted that for tax reasons, most of the private enterprises report lower income levels than their actual income. Thus, some private hospitals are likely to declare lower revenue in the reports. In this context, it is possible to argue that private health care expenditure in Turkey is likely to be higher than the official statistics suggest. Second, it is common in Turkey to make informal payments such as giving presents to doctors or hospital personnel in order to receive better care (Adaman, 2003; Erus and Aktakke, 2012). Tatar *et al.* (2007) found that about 70% of total out-of-pocket health expenditure is formal while the rest is informal in Turkey using a questionnaire adopted from a wider international study. Their results also indicated that drug expenditure constitutes half of formal expenditure and medical expenditure on physicians is the most significant component of the informal payments. Third, only two-

thirds of the population have health insurance coverage in Turkey. Uninsured people are usually not working in the formal economy and, thus, they and their dependants are not covered by any health insurance scheme. This has resulted in high levels of out-of-pocket health expenditure and informal payments to health facilities (Tatar *et al.*, 2007).

In the context of the Health Transformation Programme, the period 2003-2008 has witnessed several reforms to harmonise health benefits across the different public health insurance schemes which may have important effects on out-of-pocket health expenditure in different ways. One of the most important health reforms is that SSK and BK enrolees gained access to outpatient and inpatient services in contracted private hospitals. With these changes, the benefits of SSK and BK members were improved to the level of the GERF members[4] (OECD, 2008). People who prefer private facilities over public facilities, because of, for example, overcrowding in public hospitals, long waiting times to see a doctor and a perceived low quality of health care, would have benefited from this change in private facilities. In order to encourage private sector to contract with the SSI, private hospitals are allowed to implement 'extra billing' (OECD, 2008). Therefore, it is still the case that using private facilities requires an extra charge imposed by the private provider, though the charge was reduced by the recent health reforms[5] (Erus and Aktakke, 2012).

With regard to other reforms, Green Card members gained access to outpatient care and pharmaceuticals in 2005[6]. With this change, they have become entitled to the same benefits as SSK, BK and GERF beneficiaries. In 2007, it was accepted that all individuals would have access to free primary care even if they are not entitled to any health insurance scheme. In order to improve productivity among public sector health personnel, a performance-based payment system was initiated in MoH facilities in 2004 (OECD, 2008). In 2008, implementation of the UHI began. The UHI aims to extend GERF benefits to all insurance schemes. However, the completion of unifying all insurance schemes will take some time (Sulku and Bernard, 2009).

[4] GERF members and active civil servants had access to private hospitals after 2003 (OECD, 2008).
[5] This extra charge was limited to 30% of the payment by the government to the provider (Erus and Aktakke, 2012).
[6] In the same year, it was accepted that co-payments are required from the Green Card holders for pharmaceuticals (OECD, 2008).

It can be argued that one of the main objectives of these reforms is to enhance the financial protection of all members of the various health insurance schemes. However, even under full insurance coverage, patients are required to pay a co-payment for drug expenditure. Furthermore, informal payments to doctors and hospital personnel are also still an issue in Turkey's health care system. Improvement in access to private hospitals has increased both the number of and the service volume of health care facilities in the country. However, this improvement has satisfied patients who prefer private facilities over public facilities and who are able to pay for private health services. Thus, this improvement has a potential to aggravate existing inequalities in access to health care among people with different income levels.

2.3 LITERATURE REVIEW

2.3.1 Defining Catastrophic Health Expenditure

Fairness in financing health care is the main issue behind the concerns about catastrophic health expenditure in the existing literature. It is stated that people should contribute to the health care cost according to their ability to pay (vertical equity), while utilising necessary medical care (van Doorslaer, Wagstaff, van der Burg *et al.*, 1999).

A number of recent policy reports have emphasised the importance of the financial protection of households from high levels of out-of-pocket health care expenditure. In its 2000 World Health Report, the WHO defined fair financing in health systems as "the risk each household faces due to the costs of the health system is distributed according to ability to pay rather than to the risk of illness: a fairly financed system ensures financial protection for everyone." (p. 35). The report also emphasised that a health system is considered as unfair if the households or individuals are forced into poverty as a result of purchasing needed care or are forced to do without care because of the high costs.

In another recent policy document, the 2002 report of the International Labour Organisation (ILO), it is stated that a health system should guarantee that no one is impoverished by paying for health care. An active role should be given to the insurance system to protect households from catastrophic health expenditure (Baeza *et al.*, 2002). Financial protection is defined as follows: "financial protection means that no family or household should contribute any more than a reasonable proportion of their income to finance a system of social protection in health and/or specific health services." (p. 7).

There is no specific definition of catastrophic health expenditure in the existing literature. Russel (2004) provided a comprehensive definition of catastrophic health expenditure as follows: "the term catastrophic implies that such expenditure levels are likely to force household members to cut their consumption of other minimum needs, trigger productive asset sales or high levels of debt and lead to impoverishment." (p. 147).

The most common approach is to set the threshold in terms of out-of-pocket health expenditure as a percentage of income (Berki, 1986, Wyszewianski, 1986). However, it is not clear what threshold levels of income can be considered 'catastrophic'. As

Wyszewianski (1986) notes, even high out-of-pocket health expenditures are not always catastrophic in terms of imposing a severe financial burden on the household. However, a small amount of expenditure on health care can be financially devastating for poor households. Therefore, it is common to use and present results for a range of threshold values in the existing literature (see, for example, Wagstaff and Doorslaer, 2003 for Vietnam).

In addition, there is no consensus in terms of defining denominators. Some studies use total household income as a denominator whereas some studies consider 'disposable' household income ('ability to pay' or 'capacity to pay' in the WHO's terminology) which means effective income remaining after basic needs, usually food, have been met. To calculate household subsistence expenditure, the poverty line is based on the proportion of food expenditure in total expenditure (Xu *et al.*, 2003). The motivation behind using the latter approach is that poor households spend more on basic necessities, particularly food, whereas well-off households have more capacity for other spending including health care expenditure. Food expenditure is considered as an approximation to the household's relative income (WHO, 2000).

In this context, it can be argued that the greater vulnerability of the poor can be emphasised by using non-food expenditure as a denominator. The reason behind this choice is that the poor spend a larger proportion of their income on food than the relatively well-off households, leaving a relatively small proportion of income to meet other expenditure. In this case, a small amount spent on health care can cause poor households to exceed a given non-food expenditure threshold and increase their vulnerability to catastrophic health expenditure. However, it should be stated that the decision about the choice of denominator is somewhat subjective and arguably political (Hatt, 2006). As an example of this issue, Xu *et al.* (2003), in their study of 59 countries, used 40% of a household's total expenditure minus the average food expenditure of households in the 45^{th} and 55^{th} percentiles, adjusted for household size in accordance with a consumption equivalence scale[7].

In the existing literature, the total expenditure of the household is generally used as a proxy for income in analyses of household survey data for developing countries since

[7] The WHO's initial suggested thereshold was 50% of non-food expenditure, which was found to be quite high in the sense of reflecting a small proportion of households with catastrophic health expenditure and was revised downward to 40% (WHO, 2000; Kawabata, 2002).

expenditure tends to be more accurately reported, is easier to measure and is measured with less error relative to current income measures. Furthermore, it has been argued that total household expenditure is a better proxy for household income since savings allow smoothing of expenditure over time whereas income may be subject to transitory fluctuations (Deaton, 1997; Tansel, 2002).

On the other hand, when using total expenditure or total income in the denominator, lower thresholds levels have been more commonly used in the literature. Wagstaff and van Doorslaer (2003) presented results using a range of thresholds and both total expenditure and total expenditure minus food expenditure as denominators, from 2.5% to 15% of total expenditure and from 10% to 40% of non-food expenditure for Vietnam. Berki (1986) used a 5% threshold level of total household income for the analysis of catastrophic health expenditure in the US. He stated that the reason for choosing this threshold is that households were allowed to deduct health expenditures exceeding 5% of their income with the US tax code in 1986. Waters *et al.* (2004) again used total household income for their analysis of financial protection in the US but chose a 40% threshold level. Ranson (2002) and Pradhan and Prescott (2002) used 10% of actual self-reported household income for Indonesia and 10% of total household expenditure for Gujarat (India), respectively.

The numerator has most extensively been defined as total health expenditure including co-payments, consultation fees, purchase of medicine, hospital bills and other types of out-of-pocket expenditure on health and generally excludes insurance premiums (Ranson, 2002; Wagstaff and van Doorslaer, 2003; Hatt, 2006). The reason for excluding insurance premiums and taxation is that this type of health expenditure is arguably not made at the time the household received the service and, moreover, can be anticipated in advance. Therefore, it can be argued that the numerator includes unplanned out-of-pocket health expenditure[8]. However, there are some studies which included insurance premiums and social insurance contributions in the numerator (Knaul, 2000; Murray *et al.*, 2000).

[8] Any reimbursement from a health insurance scheme is also deducted from the out-of-pocket health expenditure of households.

2.3.2 The Distribution of Catastrophic Health Expenditure

There is no coherent pattern of the socioeconomic distribution of health expenditure within developing countries. Makinen *et al*. (2000) published a review of household survey data from eight developing countries and countries in transition. They found that there was no clear pattern among countries concerning health expenditure as a proportion of income by income quintiles. In Burkina Faso, Paraguay and Thailand, regressive trends have been found (i.e. the wealthier quintiles spend a lower percentage of their total consumption on health care than poorer quintiles) whereas in Guatemala and South Africa, progressive trends have been identified. Moreover, wealthier households were found to be more likely to seek health care when they need it than poorer households. In other words, an upward trend by quintile in health care use was observed in all countries included in the analysis with the exception of Kyrgyzstan. It might be the case that poor households may fail to seek health care when they need it since they believe that they cannot afford it. As Pradhan and Prescott (2002) state, in this case, these poor families who need financial protection are excluded from the measurement. Therefore, it is important to investigate the differences in the health seeking behaviour of households.

There are also some limitations of Makinen *et al*. (2000)'s study; firstly, there are important differences in the household surveys across countries. For example, the sample sizes are different and some of the surveys are not nationally representative; secondly, the information about health seeking behaviour or spending on health care obtained from the surveys is arguably not comparable since the surveys were conducted independently and the questions in those surveys were not asked in the same way.

In a similar vein, Xu *et al.* (2003) used household survey data from 59 countries to investigate the levels and determinants of catastrophic health expenditure. The percentage of households that incurred catastrophic health expenditure was calculated using the threshold of 40% of non-food expenditure. The findings indicated a range of different patterns of catastrophic health expenditure across countries. In countries with advanced social protection systems such as Canada, Czech Republic, Denmark, the UK, Germany and France, the proportion of households incurring catastrophic health expenditure is less than 0.1%. However, catastrophic health expenditure is found to be common in some countries in transition, middle-income countries, in certain Latin

American countries and several low-income countries with over 10% in Vietnam and Brazil and over 5% in Azerbaijan and Colombia.[9] They also investigated the factors that are likely to be associated with the risk of catastrophic health expenditure and found that there is a strong positive relationship between the proportion of households with catastrophic health expenditure and the share of out-of-pocket payments in total national health expenditure. Moreover, it was found that the proportion of households experiencing catastrophic health expenditure is positively related with the share of total health expenditure in GDP and the percentage of households living below the estimated poverty line. Generally, lower income groups are more likely to incur catastrophic health expenditure as compared to higher income groups. However, the interesting finding of this study is that the highest rate of catastrophic health expenditure was not observed in the lowest income group. This finding can be attributed to the health care seeking behaviour of households. They may choose not to seek heath care due to its high cost and, hence, report zero or low health expenditure.

In the Xu *et al.* (2003) study, the threshold levels and poverty definitions were changed to check the robustness of the results and the results stayed the same across all threshold levels and poverty definitions. However, there were many important limitations to their analysis. First, it is clear that there are country-specific factors that affect the health system's organisation and financing system but these factors were not included in the analysis. Second, the study provided no information on household-level correlates of catastrophic health expenditure. Third, non-medical and indirect costs of care seeking such as transport, food and accommodation costs or lost working time and earnings were not included so the effect of out-of-pocket health expenditure was potentially underestimated. Most household surveys, however, do not include any information regarding non-medical or indirect costs of care seeking and they only include direct costs[10]. Fourth, the household surveys of countries used in the study were not representative of the world population. Therefore, generalisations made from these results can be misleading. Fifth, since information on household expenditure was obtained by different methods across countries such as recall of expenditure in household interviews or recording of expenditure in diaries, it can be argued that there

[9] Turkey was not included in the analysis. This could be because the Household Budget Surveys have been conducted since 2002 and, hence, they were not available until recently.

[10] Since the Turkish Household Budget Surveys do not include such information, this study presented in this chapter cannot take into account non-medical and indirect costs of care seeking.

was potentially a high level of measurement error in each of the predictor and outcome variables. Hence, comparisons of the countries in terms of the proportion of households with catastrophic health expenditure are potentially problematic (Hatt, 2006). Finally, as they stated in the study, many poor households may delay their medical needs due to high health care costs so the effect of out-of-pocket expenditure is not completely captured by only examining the levels of catastrophic health expenditure.

Wagstaff and van Doorslaer (2003) developed two alternative approaches for the definition and measurement of catastrophic health expenditure based on a binary indicator. The first approach can be defined as out-of-pocket health expenditure that is more than a pre-specified threshold level of household income or expenditure (catastrophic health expenditure). The second approach measures whether a household is forced into poverty by out-of-pocket health expenditure or the poverty status of the household has been exacerbated as a result of this expenditure (impoverishment). They developed indices for the measurement of intensity and incidence of catastrophic health expenditure and for the measurement of poverty-impact incidence (i.e. crossing the poverty line) and intensity (i.e. how far out-of-pocket health expenditure pushes the household below the poverty line). These methods were used to investigate the incidence and socioeconomic distribution of catastrophic health expenditure in Vietnam using out-of-pocket health expenditure data for 1993 and 1998. Their findings indicated that the proportion of households spending more than 5% of their income on health care was 38% in 1993 and this rate had decreased to 33% in 1998 and became less concentrated among the poor households. Out-of-pocket health expenditure increased the normalised poverty gap (i.e. poverty gaps are divided through by the poverty line) by 1.4 percentage points in 1993 and by 0.8 percentage points in 1998. In both years, nearly three quarters of the addition to the poverty gap was from previously poor people becoming even poorer and only one quarter resulted from previously non-poor people being pushed below the poverty line by out-of-pocket health expenditure. In other words, they emphasised that most of the 'poverty impact' resulting from catastrophic health expenditure was due to the poor becoming even poorer rather than the non-poor becoming poor.

Wagstaff and van Doorslaer's (2003) study is arguably one of the pioneer studies of the socioeconomic distribution of catastrophic health expenditure. However, it has been criticised for assuming a fixed household income (Hatt, 2006). The definition used by

Wagstaff and Doorslaer (2003) is insensitive to how these expenditures are financed. This approach ignores financing coping strategies such as spending savings, selling assets or borrowing money from friends or relatives. It is possible that households might protect consumption of other goods at least in the short term (Flores *et al.*, 2008; Sauerborn *et al.*, 1996).

2.3.3 The Determinants of Catastrophic Health Expenditure

Although some studies have focused on the factors that are likely to be associated with the risk of catastrophic health expenditure, only a limited number of empirical analyses have been conducted. Therefore, little is known about which segments of the population are most at risk of experiencing catastrophic health expenditure. In general, catastrophic health expenditure is associated with poverty or low income, unemployment, low levels of insurance coverage and having disabled, chronically ill or aging household members. Wyszewianski (1986), for example, found that ageing, unemployment and poverty were the most important risk factors in the US for incurring catastrophic health expenditure. Similarly, in an earlier analysis of the US health system, Berki (1986) stated that poverty and not having health insurance coverage were among the risk factors associated with catastrophic expenditure on health care.

O'Donnell *et al*. (2005) investigated sources of variation in the incidence of catastrophic expenditure on health care across six Asian countries using household surveys. They used a 10% threshold level of total household expenditure following Pradhan and Prescott (2002) and Ranson (2002). They compared the estimation results of a standard probit model to one which considered the endogeneity of total household expenditure. They argued that total expenditure could be endogenous since households generally use a range of strategies, such as borrowing, using savings or selling assets, to meet health care costs and do not necessarily choose to cut other types of consumption within a fixed, single period income constraint. Different indicators of access to savings and credit, such as land holdings or land size, were used as instruments for total expenditure. It was assumed that such access affected the household's total expenditure but was not correlated with the out-of-pocket budget share. The reason for choosing such measures of access as instruments was that, for a given initial income, households with easy access to credit have more capacity to extend the household budget to cover unexpected health expenditure. The authors found that the probability of incurring catastrophic

health expenditure increases with total household expenditure ('richer' households incurred such expenditure more than poor households) when endogeneity was not controlled for. However, when endogeneity was allowed for by including instruments, the positive coefficient of total expenditure changed from being statistically significant to being statistically insignificant which implied that endogeneity led to a strong upward bias in the estimation results indicating the impact of the household income on its out-of-pocket expenditure share. They also found that having a highly educated household head, insurance coverage and living in an urban area were all inversely associated with the probability of incurring catastrophic health expenditure.

The study by O'Donnell *et al.* (2005) is important in terms of introducing a new methodology addressing a widely discussed dimension of catastrophic health expenditure. However, the methodology is arguably somewhat controversial because the instruments used in the analysis are potentially problematic in terms of their liquidity. They are large assets to sell in order to cover health care costs. Furthermore, inter-temporal adjustment to health care costs has a potential to extend its burden over time and may lead to greater debt in the future. In order to analyse the role of financial coping strategies, it seems that using longitudinal data is an interesting avenue to pursue.

Although as mentioned above, an important conceptual challenge should be again highlighted at this point. As Kawabata *et al.* (2002) emphasise, it is possible and common that the highest proportion of catastrophic health expenditure is not always experienced by the lowest income group. The reason is that 'catastrophic health expenditure' can only be experienced if the household seeks health care and expenditure occurs (Hatt, 2006). Poor households usually experience a delay in meeting their health care needs due to their economic situation. In this case, catastrophic health expenditure indicators are subject to potential selection bias. Medical care seeking behaviour is not accounted for in most of the analyses and this measurement problem is accepted as a limitation and the analysis is conducted with the available household survey data. However, this problem has a critical importance in terms of financial accessibility to health care (McIntyre *et al,* 2006). Thus, ignoring this dimension of catastrophic health expenditure arguably does not provide a complete picture of the issue of financial protection.

Pradhan and Prescott (2002) used a simulation model to construct a distribution of needed health expenditure using household survey data for Indonesia which provides household health care expenditure as well as health care utilisation information. Catastrophic health expenditure was defined as out-of-pocket health expenditure exceeding 10% of the household's total expenditure. The distribution of catastrophic health expenditure by expenditure quintiles indicated progressive trends, which implies that richer households are more likely to spend 10% of their income on health care as compared to poor households in Indonesia. It was claimed that it is not possible to investigate directly whether the poor households suffer disproportionally from catastrophic health expenditure from the household survey data. Therefore, they used observed health service utilisation and the expenditure pattern of the middle-income group and randomly applied this group's pattern to the rest of sample. Then, the obtained and actual expenditure of the household were compared to each other to obtain the stochastic distribution of 'needed' health care. The reason for choosing the middle income quintile was justified as it is a starting point to obtain information for underutilisation by the poor households and overutilisation by the rich households (Hatt, 2006). The age and sex composition of households was used as a proxy for health status which identifies 'needed' health expenditure.

Pradhan and Prescott (2002) also highlighted the relationship between equity and financial protection. Besides financial protection, ensuring equity in access to health care is also one of the most important objectives of health policy. They noted that equity requires a subsidy for low-cost primary care that would generate benefit for a large proportion of the population, whereas financial protection focuses on subsidies for the smaller proportion of the population who are more likely to incur high health care costs and impoverishment. The results of their simulation analysis indicated that subsidising inpatient care would result in the greatest decrease in the proportion of households with catastrophic health expenditure while subsidising outpatient care would provide benefits particularly for the very poor segment of the population. They concluded that if the aim is to provide financial protection for poor households, the free inpatient regime is not the preferred regime for Indonesia. One limitation of their study is that they only allowed the prices of health care to change and treated all the utilisation rates as fixed and, in this case, increasing utilisation rates arising from subsidising outpatient care could not be identified (Hatt, 2006). On the other hand, Pradhan and Prescott (2002)

attempted to shed light on the differences in the utilisation of health care between the poor and the rich households which can be regarded as an important contribution to the literature.

The relationship between health insurance and its effect on out-of-pocket health expenditure is also a widely analysed issue in the existing literature (Sepehri *et al.*, 2006). It is expected that insurance coverage provides financial protection from catastrophic health expenditure. However, it is also possible for health insurance to create demand inducement, in the case of, for example, small benefit packages or inadequate insurance coverage, and this demand increase may result in high levels of out-of-pocket health expenditure (Wagstaff and Lindelow, 2008). In some countries, insurance packages cover only some of the total costs, requiring households to pay a co-payment. If the total cost is quite high relative to the budget, even a 20% co-payment can be classified as catastrophic health expenditure. As Kawabata *et al.* (2002) state, under insurance coverage, catastrophic health expenditure may not simply go away if the benefit package does not cover most of the health expenditure.

In addition, the study by Foster (1994) is considered as one of the leading studies in the existing literature, which emphasises studying the consequences of illness in terms of productivity. Foster (1994) noted that in the health-productivity literature the effect of illness is more important than the effect of nutrition because, in contrast to the case of nutrition over which households arguably have relatively direct control, illness has an unexpected nature, which households may have little control over. In this context, there is also a growing literature on indirect costs, which focus on productive time losses for the ill individual and for other household members. However, most of household surveys do not include indirect costs and, thus, studies generally include only direct costs due to data unavailability and methodological challenges (McIntyre *et al.*, 2006). In this respect, most of the studies that include both direct and indirect costs emphasised that indirect costs are generally more than direct costs (see, for example, McIntyre *et al.*, 2006 for a review of studies carried out in low and middle income countries; Sauerborn *et al.* 1996 for Burkina Faso; Koopmenschap and Rutten, 1994 for eight different countries and Gertler and Gruber, 2002 for Indonesia).

On the other hand, in addition to hospitalisation costs, fees and medicine costs, direct costs include transportation costs, costs of nutrition (e.g. special food for a sick member

of the household) and accommodation costs. According to Attanayake *et al.* (2000), transport costs for ill household members may also impose a considerable burden on household's budget and may constitute 20% of total direct costs (McIntyre *et al.*, 2006). Analysis focusing on the economic consequences of out-of-pocket health expenditure including other components of direct costs is, however, limited due to data availability.

In the case of Turkey, there are only a few studies focusing on the factors which are associated with catastrophic health expenditure risk. However, there have been important policy changes in the health area since 2003 which may have important effects on health expenditure in Turkey. Yardim *et al*., (2009) investigated the level of catastrophic health expenditure and identified the factors associated with catastrophic health expenditure risks in Turkey. The HBS for 2006 and the methods introduced by Xu *et al.,* (2003) were used. The results of the logistic regression analysis indicated that the health insurance coverage of the household head and living in an urban area were closely related to the catastrophic status of households. There are, however, two main limitations of this study; firstly, they estimated the model only for 2006 and used only one threshold level (40% of non-food expenditure). Hence, it can be argued that the time dimension and sensitivity checks of the results were ignored. Therefore, policy implications drawn from these results could be potentially misleading. Secondly, they overlooked the problems arising from the distribution of catastrophic expenditure across income quintiles and the selection problem in terms of the difference in the treatment seeking position of the poor and the rich households were not taken into account.

In another study focusing on Turkey, Sulku and Bernard (2009) examined the role of the health insurance system in terms of providing adequate financial protection against high out-of-pocket health expenditure in the population aged less than 65 years using Turkey's 2002/2003 National Household Health Expenditure Survey. They found that 19% of the non-elderly population were living in households whose health expenditure exceeds 10% of their income. For poor households, 23% of the non-elderly population were living in households whose expenditure on health care is more than 20% of their income.

Kisa *et al.* (2009), on the other hand, investigated the delayed use of health care services among the urban poor in Turkey. They conducted a field study among the 92 poorest households in the Etimesgut region in Ankara in order to collect information about

health service delays among the poor as well as the factors associated with those delays. Household members were asked whether they had experienced difficulty in paying for health care services and the results indicated that about 63.3% of poor households did not seek health care due to inability to pay and 17.4% of poor households reported that they had experienced extreme financial difficulty when attempting to pay for health care. They concluded that overall; two out of three poor households had delayed or not sought health care because they thought they could not afford it. The results of their study suggest that the medical care seeking behaviour of poor households is an important problem in Turkey. Thus, difficulties in access to health care among the poor should be one of the primary goals of health policy in Turkey. In this context, this chapter extends the analysis of the determinants of catastrophic health expenditure in Turkey by controlling for potential selection bias in terms of treatment seeking behaviour.

2.4 DATA AND METHODOLOGY

2.4.1 Data

The Turkish Statistical Institute (TurkStat) has been conducting the Household Budget Survey (HBS) annually since 2002. Prior to this, the Household Income and Consumption Expenditure Survey (HICES) was conducted to provide information on the distribution of income and consumption with respect to the population, region and rural and urban areas in 1987. The latest HICES was conducted in 1994 but a different method from previous surveys was followed. In order to eliminate the influences of unusual conditions such as natural disasters or economic crises from the survey, it was decided to construct a systematic survey structure in 2002 (TurkStat, 2006). The HBSs, which are nationally representative of Turkey, are a time-series of repeated cross-sections rather than panel data that follow the same households over time.

In this study, the HBSs from 2002 to 2008 have been used. The questionnaires have been standardised across all the years analysed with the changes in international standards. The questionnaire used in the survey includes the following sections: household composition; socio-economic status of the household; general consumption patterns of the household; stocks of durables and related expected expenditures; household consumption expenditure; commodities consumed from the household's own production; in-kind income received by the household; income in terms of gifts and aid received by the household; goods and services purchased by the household to be given as a gift, aid or transfer to other households or institutions; expenditure other than consumption (savings, debt or loan payments); employment and income status of the household; information related to agricultural holdings and the balance of income and expenditure. Basically, the survey includes three main groups of variables: variables relating to the socio-economic status of the households (heating system, type of dwelling and ownership, facilities in the house, owned durables and transportation vehicles); expenditure on consumption and variables related to individuals (age, gender and education); variables related to employment (occupation, profession and position in the job), income from main economic activity and from subsidiary economic activities.

In the health category, the HBS provides information on out-of-pocket health care expenditure as well as its components. It covers expenditure on medicine and pharmacy products, treatment equipment, dentistry services, laboratories and X-ray services,

nursing care services and hospitalisation. In addition, the survey gives information about the health insurance status of each individual and the presence of disabled and ill members in the household. It does not unfortunately provide information on the use of health facilities/providers.

In order to minimise seasonal influences on consumption, households were changed on a monthly basis by one with similar characteristics (TurkStat, 2006). Each HBS includes a large number of households with about 800 households that changed each month of the survey year (except 2003 in which about 2200 households were surveyed per month) [11].

Table 2.1: Sample Sizes of the Household Budget Surveys

Years	Urban	Rural	Total
2002	8091	1464	9555
2003	18278	7486	25764
2004	5985	2559	8544
2005	5985	2574	8559
2006	5930	2628	8558
2007	5893	2655	8548
2008	5958	2591	8549
Total	**56120**	**21957**	**78077**

Table 2.1 presents the number of households included by each survey year. A total of 78077 household-level observations (56120 for urban and 21957 for rural) are included in the data set pooled over seven years.

2.4.2 Estimation Methods

In order to identify the factors which are associated with the risk of catastrophic health expenditure, a standard probit model is estimated for the pooled data as a first step. Probit regression is a commonly used framework where the outcome of interest is a binary variable. The dependent variable is a discrete variable which takes the value of 1 if the household's total out-of-pocket health expenditure exceeds the threshold level (i.e., if the household experienced catastrophic health expenditure) and 0 otherwise.

[11] The sample size of the 2003 HBS was increased in order to cover the whole of Turkey including all regions and 26 provinces to construct the new harmonised consumer price index in the context of the EU's harmonisation studies (TurkStat, 2006).

The probit model takes the following form:

$$y^* = x'\beta + e \tag{2.1}$$

where y^* denotes a latent or unobserved variable, β is a set of parameters, x is a vector of explanatory variables and the error term is normally distributed with zero mean and unit variance, $e \sim N(0,1)$. We observe $y = 1$ if $y^* > 0$ and $y = 0$ otherwise. In this formulation, $x'\beta$ is called the index function (Greene, 2003).

Equation (2.2), below, gives the likelihood for both cases $y = 0$ and $y = 1$. The cumulative standard normal distribution, $\Phi(.)$, constrains the probability to lie between 0 and 1.

$$Pr(y = 1|x) = \Phi(x'\beta) \tag{2.2}$$

The household-level covariates include a dummy indicator for survey year (base year= 2002); household size including its squared term to capture a potential non-linear relationship with the outcome variable; a dummy indicator for urban residence (base= rural); a dummy indicator for the household head not having health insurance (base= the presence of health insurance coverage); a dummy indicator for the presence of a disabled or ill member in the household; a dummy indicator for a male household head; dummy indicators for the highest level of education of the household head (base= primary education or less); a dummy indicator for a poor household; the number of preschool children in the household (under aged 5); the number of school children in the household (aged 6-14), the number of elderly household members (aged 65 and over) and dummy indicators for the employment status of the household head (base= not employed)[12]. The set of parameters β denotes the effect of changes in x on the probability of incurring catastrophic health expenditure. Estimation of the probit model is based on the maximum likelihood method.

As mentioned above, if the household expects health costs to be unaffordable and, thus, chooses not to seek health care, there could be a sample selection problem, which may bias the estimation results. The sample selection bias is defined as a systematic error where the dependent variable is observed only for a non-random sample. In this case, the households who choose not to seek treatment are not equally represented with the

[12] Detailed variable definitions are given in Section 2.4.3 below.

households who seek treatment in the estimation results. The implications of taking into account the medical care seeking behaviour of households should be explored in order to accurately assess the risk factors associated with catastrophic health expenditure. Unfortunately, the HBS does not include information on the use of health services. Therefore, a binary variable for the selection equation is based on having or not having positive health expenditure since all members of all the insurance schemes are required to pay a co-payment for drug expenditure which provides information on their medical visits (Erus and Aktakke, 2012). It should be noted that the sample selection problem controlled for in this chapter may be seen as a specific part of a wider problem since individuals with the same health condition may choose to spend different levels of health expenditure. The health status of the individuals, however, cannot be included in the analysis due to unavailability of a measure of the health status of individuals in the data set.

Since the probit model in equation (2.1) above is based on a binary dependent variable, the health care seeking behaviour of households (i.e. having positive health expenditure) cannot be accounted for in this model[13]. In order to control for the potential sample selection problem, the Sartori selection model is explored in the second specification (Sartori, 2003). In this context, the selection equation predicts the probability that the household sought health care and determines whether the household would be observed in the outcome equation. The outcome equation predicts the probability that the household incurred catastrophic health expenditure, adjusted for the selection process.

Studies generally use the Heckman selection approach (Heckman, 1979) to deal with potential selection bias (Van de Ven and Van Praag, 1981). In the Heckman selection model, the equation that determines the sample selection is as follows:

[13] An alternative estimation method is a multinomial logit model which is used when there are multiple categories which cannot be ordered. The potential categories are: households with catastrophic health expenditure; households with positive but not catastrophic health expenditure; and households with zero health expenditure. The multinomial logit model relies on the assumption of the Independence of Irrelevant Alternatives (IIA) which states that odds for each specific pair of outcomes are not affected by the other available outcomes. It is important to test this assumption prior to the estimation of this model. The suest-based Hausman test, which is a modification of the Hausman and McFadden (Greene, 2003) test, is used to test this assumption because the Hausman test and small-Hsiao test provide conflicting information in determining violations of the assumption of the IIA (Long and Freese, 2006). However, this assumption is rejected for every definition of catastrophic health expenditure. It implies that the categories are not distinct and weighted independently. The nested logit model which relaxes this assumption has also been explored and the results of the nested logit model are found to be consistent with the other models used in this chapter.

$$z_i^* = w'\gamma_i + u_i \quad (2.3)$$

and the equation of primary interest (the outcome equation) is as follows:

$$y_i = x_i'\beta + \varepsilon_i \quad (2.4)$$

The sample selection rule is that y_i is observed only when z_i^* is greater than zero. It is assumed that ε_i and u_i have a bivariate normal distribution with zero means and correlation ρ. The likelihood that the household seeks health care and, hence, is observed in the outcome equation is estimated with a probit model:

$$Prob\ (y_i = 1) = prob(w'\gamma_i + u_i > 0) \quad (2.5)$$

$$= prob(u_i > -w'\gamma_i) \quad (2.6)$$

$$= 1 - \Phi(-w'\gamma_i) \quad (2.7)$$

$$= \Phi(w'\gamma_i) \quad (2.8)$$

Similarly, the binary outcome is only observed when $z_i^* = 1$. The probability of the outcome equation can be estimated via a partial bivariate probit model (since not all possible combinations of outcomes are observed, it is a partial bivariate probit):

$$Prob\ (y_i = 1) = prob(x_i'\beta + \varepsilon_i > 0) \cap prob(w'\gamma_i + u_i > 0) \quad (2.9)$$

$$= \Phi_2(x_i'\beta, w'\gamma_i, \rho) \quad (2.10)$$

The log likelihood function includes three combinations: (z_i=1, y_i=1), (z_i=1, y_i=0), (z_i=0), (Greene, 2003; Hatt, 2006).

To identify the Heckman selection model, there should be at least one extra explanatory variable that affects the selection equation but not the outcome equation. This means that the selection equation should have at least one explanatory variable which is excluded from the outcome equation and the choice of this variable should ideally be based on theory. Although Heckman selection models can be estimated with identical sets of explanatory variables, the results are based upon distributional assumptions about the error terms rather than upon variation in the explanatory variables (Maddala, 1999; Sartori, 2003). However, if the economic theory which determines sample selection is very similar to that determining the outcome of interest, an exclusion restriction for the Heckman selection model may not exist.

When theory suggests identical explanatory variables for both the selection and the outcome equations, Sartori (2003) argues that authors usually try to identify an arguably theoretically 'unjustifiable' exclusion restriction for the selection equation (which generally results in specification error if the variable does not have an influence in this equation) to reinforce the model requirements.

To overcome this problem, Sartori (2003) develops a new estimator where identical explanatory variables are used in the selection and outcome equations and the key identification is based on the assumption of identical error terms in both equations.

$$U_{1i} = x_i\gamma' + v_{1i} \tag{2.11}$$

$$U_{2i} = x_i\beta' + v_{2i} \tag{2.12}$$

Equation (2.11) is the selection equation and equation (2.12) is the outcome equation under the case of non-random selection. In these equations, U represents an unobserved continuous dependent variable. The explanatory variables, x, are the same in both equations, but the coefficients, γ and β are usually different. Each equation contains a normally distributed mean zero error term, v_1 and v_2. Rather than observing the U terms, two dichotomous variables, Z_{1i} and Z_{2i}, are observed, which are shown in equations (2.13) and (2.14) below (Sartori, 2003).

$$Z_{1i} = 0 \; if \; U_{1i} < 0, \; Z_{1i} = 1 \; if \; U_{1i} \geq 0 \tag{2.13}$$

$$Z_{2i} = 0 \; if \; U_{2i} < 0, \; Z_{2i} = 1 \; if \; U_{2i} \geq 0 \tag{2.14}$$

Z_{1i} represents whether or not the observation is selected and Z_{2i} represents the observed outcome equation. The main assumption is that the error terms in equations (2.13) and (2.14) are identical. Sartori (2003) describes this situation as follows: "In the situation that the observed dependent variables are dichotomous, the underlying dependent variables are on the same scale (standardised) and both error terms are normally distributed. Thus, it is initially plausible that two equations have similar error terms". (p.117)

Sartori, then, derives a maximum-likelihood estimator for the impact of the explanatory variables on the outcome variable of interest allowing for the selection process. Sartori describes three random variables, Y_{ij}, as follows:

$$Y_{i0} = 1 \; if \; Z_1 = 0 \; and \; 0 \; otherwise; \tag{2.15}$$

$$Y_{i1} = 1 \; if \; Z_1 = 1 \; and \; Z_2 = 0 \; and \; 0 \; otherwise; \tag{2.16}$$

$$Y_{i2} = 1 \; if \; Z_1 = 1 \; and \; Z_2 = 1 \; and \; 0 \; otherwise; \tag{2.17}$$

Y_{i0} takes the value of 1 if the observation is not selected, Y_{i1} takes the value of 1 if the observation is selected but the value of the outcome variable is 0 and Y_{i2} takes the value of 1 if the observation is selected and the value of the outcome variable is 1. The second step is to specify the data generating process for the probability that $Y_{ij} = 1$ in each case. Sartori (2003) defines these probabilities as follows[14]:

$$Pr(Y_{i0} = 1) = \Phi(-x_i\gamma') \tag{2.18}$$

$$Pr(Y_{i1} = 1) = \Phi(-x_i\beta') - \Phi(-x_i\gamma') \;\; if \;\; (\gamma' - \beta')x_i = 0 \; and \; 0 \; otherwise \tag{2.19}$$

$$Pr(Y_{i2} = 1) = \; \Phi(x_i\beta') \;\; if \;\; (\gamma' - \beta')x_i > 0 \; and \; \Phi(x_i\gamma') \;\; if \;\; (\gamma' - \beta')x_i \leq 0 \tag{2.20}$$

Sartori (2003) states that the assumption of identical error terms can be expected to hold if the processes behind the selection and outcome of interest are very similar or involve similar goals; if the selection and outcome have the same causes and if these two processes are close to each other in time frame and/or space[15].

In the current application, it seems reasonable to assume that the processes behind seeking health care and incurring catastrophic health expenditure are very similar and that they have the similar determinants. The primary determinant for seeking health care and incurring catastrophic health expenditure is poor health. In addition, as explained above, the numerator only takes into account unexpected out-of-pocket health expenditures, (i.e., it excludes insurance premiums which are arguably anticipated in advance). Therefore, seeking health care and experiencing catastrophic health expenditure arguably occur at the same time and place. It can, therefore, be argued that this case largely satisfies the conditions for the assumption of identical error terms.

In order to obtain the Sartori selection estimation results, the dependent variable is coded as follows; it takes the value of 0 if the observation does not select in (i.e.

[14] For further details see Sartori (2003), p. 118-120.

[15] When the same factors have opposite effects on the selection and outcome processes, their errors can be assumed to be in the opposite direction to each other. For such cases, the Sartori selection model also allows for opposite error terms. However, in this chapter, all of the Sartori estimations are obtained by assuming identical error terms.

households having zero health expenditure); it takes the value of 1 if the observation selects in and has a value of 0 in the outcome equation (i.e. households not experiencing catastrophic health expenditure but having positive health expenditure) and it takes the value of 2 if the observation selects in and has a value of 1 in the outcome equation (i.e. households incurring catastrophic health expenditure and having positive health expenditure). It should be highlighted that in the Sartori selection estimation, the same sample (all households) used in the standard probit model is analysed rather than the sample including only those households who need health care because the questionnaires of the HBS do not include any questions related to the health status of the individuals.

In Section 2.5, in addition to the results of the probit and the Sartori selection models, the results of the Heckman selection model without exclusion restrictions (i.e. the same explanatory variables are used in both the selection and outcome equations) are presented and discussed in order to compare and explore the robustness of the results from the Sartori selection model. In addition, since the basic assumption of the Sartori selection model is based on identical error terms in both the outcome and the selection equations, it can be argued that the Sartori selection model is a special case of the Heckman selection model. Therefore, an alternative robustness check is explored by defining a constraint for the correlation coefficient between the selection and outcome error terms (ρ) for the Heckman selection model. The results of the 'constrained' Heckman selection models are also discussed in Section 2.5.

2.4.3 Variable Construction and Definitions

In this study, the household is the unit of analysis. It is assumed that when an individual faces a serious illness, the household jointly makes a decision about seeking or not seeking health care. Thus, the bargaining positions of the household members are not taken into consideration here. In other words, households are treated as a single individual, endowed with a single set of preferences (Russell, 2004).

The decision about seeking treatment is crucial in the context of catastrophic health expenditure. If the household does not seek health care, it may result in, for example, disability or death. On the other hand, if the household decides to seek health care, the household will face decisions relating to the type of health care provider such as public or private and this provider choice will influence the catastrophic health expenditure

status of the household (McIntyre *et al.*, 2006; Hatt, 2006). However, as stated above, the HBS does not provide information on the health care utilisation of the household. Therefore, the effect of the choice of health care provider on catastrophic health expenditure cannot be included in the empirical analysis. In addition, even under full public health insurance coverage, individuals must pay a co-payment for drug expenditure in Turkey. Therefore, it can be argued that having positive health expenditure provides information on medical visits (Erus and Aktakke, 2012). In this context, the health care seeking behaviour of the household is based on having or not having positive health expenditure.

The variables used in the analysis are defined as follows:

Out-of-pocket health expenditure (monthly): This category includes all types of out-of-pocket health expenditure (co-payments, drug expenditure, expenditure on other pharmacy products, other medical products, treatment equipment, nursing care and hospitalisation expenditure, laboratory tests and X-ray services, dentistry services and other health services not related to hospital) reported at the household level. It excludes health insurance expenditure, which is available in the surveys, because it can be argued that it has an anticipated nature and is not a type of health expenditure made at the time the household received the service.

Total household expenditure (monthly):[16] This variable is used as a proxy for household income. There are twelve main expenditure categories in the surveys: food and non-alcoholic beverages; alcoholic beverages and tobacco products; clothing; housing and rental payment; housing facilities; health; transportation; communication; leisure activities; education; hotels and restaurants and other goods and services. All types of expenditure are aggregated to obtain total household expenditure[17].

Total non-food expenditure (monthly): Food expenditure includes all types of food but excludes expenditure on alcoholic beverages and tobacco products and foods consumed outside the home (restaurant, hotels and bars) to avoid the 'luxury' items and reflect the

[16] Although this variable is labelled total expenditure because of the catastrophic health expenditure terminology, it also includes non-monetary consumption such as commodities consumed from own production and gifts and aid received. As Deaton (1997) states, total consumption measures welfare more accurately than total expenditure since rural families are more likely to use more home grown food than urban families and focusing on only monetary expenditure can make these families appear poorer (Hatt, 2006).

[17] All Turkish Lira (TL) values are adjusted for price inflation using the 2002 general CPI levels for monthly expenditure.

main category of the household budget more accurately (Hatt, 2006). Total food expenditure is subtracted from total expenditure in order to obtain total non-food expenditure.

Catastrophic health expenditure (the dependent variable): Ten dummy variables (i.e. ten alternative dependent variables) indicating the presence of catastrophic health expenditure in the household are created based on two different denominator approaches. The dummy variable equals 1 if the household's total out-of-pocket health expenditure exceeds the threshold value and 0 if it does not exceed the threshold value. A range of the most commonly used threshold values in the literature for each denominator are chosen in this study. Five threshold levels, 2.5%, 5%, 10%, 15% and 20% are used for the denominator of total household expenditure and five threshold levels, 10%, 15%, 20%, 30% and 40% are used for the denominator of total non-food expenditure.

Health Insurance: The HBS gives information on the health insurance status of each member of the household. Since the observation unit is the household, the insurance status of the household head is used to control for the effect of health insurance coverage on catastrophic health expenditure. A dummy variable equals 1 if the household head has no health insurance coverage and equals 0 if the household head has any health insurance coverage.

Health Status: It is apparent that the illness type (its severity and duration) is likely to have an important effect on the magnitude of out-of-pocket health expenditure. Unfortunately, the HBS does not provide information on the health status of each household member and illness types. The only information related to health status is the presence of a disabled household member or a household member who cannot work due to illness[18] (sickness disabled). A dummy variable is constructed which takes the value of 1 if an ill or/and disabled person lives in the household.

Health care seeking: As stated above, the HBS does not give information on the health care seeking behaviour of the households. Since all insurance schemes' members must

[18] This variable is obtained from the survey question related to the reasons for not looking for a job. The possible responses are: found a job but waiting to start, student, housewife, retired, old (aged 60+), disabled, ill, personal and family reasons, seasonal worker and other. However, it should be stated that it could also be case that the individual can be both sick and for example student. Therefore, it is a limitation that using this variable can lead to underestimate the effect of the presence of an ill household member on catastrophic health expenditure.

pay a co-payment for drug expenditure, a dummy variable is created as follows: if the household has positive (non-zero) health expenditure, this variable takes the value of 1 and if the household has zero health expenditure, it takes the value of 0.

Poverty: The relative poverty definition of the OECD is used to denote poor households. According to this definition, the poverty line is set at 60% of equivalised median total expenditure. In order to obtain equivalised median consumption, the total expenditure of each household is divided by the OECD equivalence household size which is an aggregate indicator of household size. This equivalence approach assigns a value of 1 to the first adult, of 0.5 to each additional adult aged 14 and above and of 0.3 to each child under the age 14. Then, the median value of equivalised expenditure is calculated. If the household`s equivalised total expenditure does not exceed 60% of equivalised median total expenditure, it is labelled a poor household (Guio, 2004).

Education: The categories for this variable are: primary education or less (including elementary education), secondary education (including any secondary level education and vocational schools) and higher education (including any post-secondary education). The highest level of the education status of the household head is used in the analysis. These categories are converted to dummy variables and 'primary or less education' is used as the omitted category.

Employment Status: This variable has three different categories: not working (including the unemployed, students, the retired, unpaid family workers, housewives, ill/disabled members), employed (not including the self-employed) and self-employed[19]. Again, the employment status of the household head is taken into consideration. These categories are included as dummy variables and the 'not working' category is used as the omitted category.

Urban/Rural residence: This dummy variable equals 1 if the household resides in an urban area and 0 if the household resides in a rural area.

Household size: The number of all individuals living in the same household is included in the analysis.

Gender: Gender of the household head is a dummy variable which equals 1 if the household head is male and 0 if the household head is female.

[19] The self-employment dummy variable is included to control for the impact of the unsalaried work position of the self-employed on catastrophic health expenditure.

Age groups: To proxy for the household-level risk of illness, the number of preschool children (aged 0-5), the number of school children (aged 6-14)[20] and the number of elderly household members (aged 65 and over) are included.

Survey years: Year dummy variables are included.

2.4.4 Descriptive Statistics

Tables presenting descriptive statistics are presented in the Appendix to this chapter. Table A2.1 presents descriptive statistics for the continuous variables and percentage distributions for the categorical variables by survey year. This table provides information on the change in the characteristics across the survey years. All values in Turkish Lira (TL) were converted to the British Pound (£) using 2002 as a base year.

Table A2.1 indicates that there are significant differences in the mean values of total household expenditure (the proxy for income) but there is not a consistently increasing trend across the years. Total expenditure decreases from 412.7 in 2002 to 328.4 in 2003 and decreases from 491.7 in 2005 to 396.9 in 2006. The mean value of monthly total (out-of-pocket) health expenditure ranges from 6.8 to 10.7, which initially appears quite small[21]. However, the definition of catastrophic health expenditure implies that risk factors can influence either the numerator or denominator of the indicator or both, and in this case, out-of-pocket health expenditure can become catastrophic when the numerator (the magnitude of out-of-pocket health expenditure) is large or the denominator (the household's total income) is small (Hatt, 2006). Average household size is around 4 over the 7-year period. The mean number of preschool children is about 0.4 and the mean number of school children is between 0.7 and 0.8 across the survey years. Similarly, the average number of elderly members living in the household does not change significantly over the period.

Table A2.1 also presents the proportion of households with catastrophic health expenditure across the survey years. There are no significant differences in the proportion of households incurring catastrophic health expenditure across the years except in 2008. The proportion of households with catastrophic health expenditure

[20] There is a difference in terms of the measurement of the age of the individuals between the surveys before and after 2006. The *exact* age of the individuals is shown in the 2002, 2003, 2004 and 2005 HBSs. However, since 2006, the age of individuals has been measured in five-year bands. Therefore, the "6-14 age group" is constructed as a "school children" group.

[21] The maximum value ranges from 1017.2 to 2922.6 across the years.

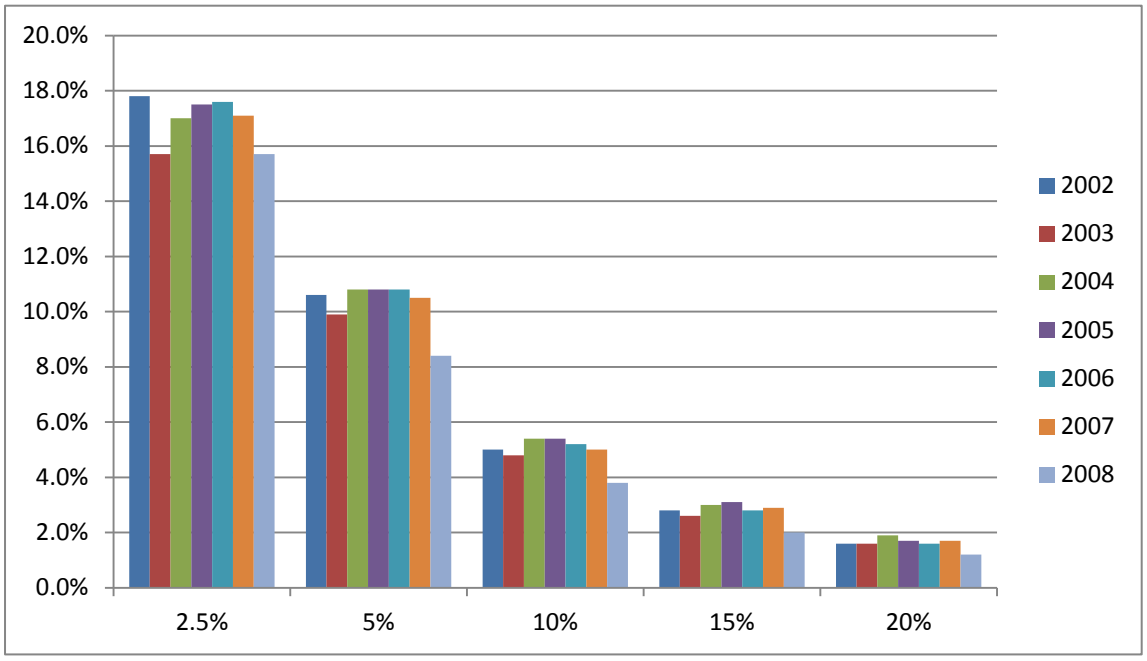
decreases in 2008, the year in which implementation of the Universal Health Insurance system was started. From 2002 to 2008, between 16% and 18% of households spent more than 2.5% of their income on health care and between 1% and 2% of households reported health expenditure exceeding 20% of their total expenditure. Similarly, between 6% and 8% of households reported health expenditure exceeding 10% of their non-food expenditure and between 0.5% and 1% of households reported health expenditure exceeding the 40% threshold level. In general, the proportion of households with catastrophic health expenditure is lower with the threshold levels of the non-food expenditure denominator than the threshold levels of the total expenditure denominator. The similarity in the proportions of catastrophic health expenditure across the survey years can be seen from Figures 2.3 and 2.4 below.

With respect to other characteristics, in 2002, 20% of the sample is labelled as 'poor' and this rate has increased to 22.6% in 2004. After 2004, there is a consistently decreasing trend in the poverty rates. If the year 2008 is excluded, it is possible to say that there are no considerable differences in the education status of the household head over time.

The pattern of the employment status of the household head stays the same over time and nearly two-thirds of the households live in an urban area rather than a rural area over the survey period (except in the year 2002). Approximately 10% of household heads are female and this pattern remains the same for most of the years.

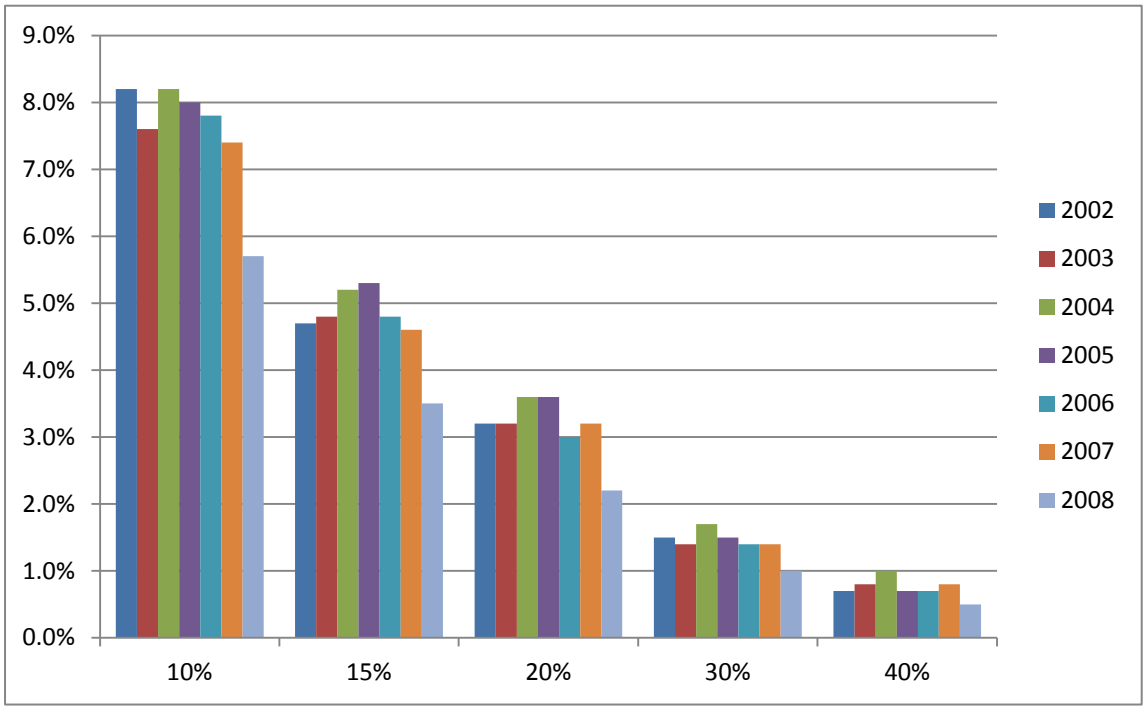
Because of a change in the questionnaire (the presence of a disabled household member was not asked in the 2002 and 2003 HBSs), the rates of the presence of a disabled or ill household member are smaller in 2002 and 2003 (4.1% and 3.7%, respectively) than the other survey years. The number of household heads with health insurance coverage has generally increased over the survey period. About 91% of household heads have health insurance in 2008.

Figure 2.3: Percentage of Households with Catastrophic Health Expenditure, by Survey Year and Threshold (Denominator: Total Expenditure)



Source: Household Budget Surveys (2002-2008), TurkStat.

Figure 2.4: Percentage of Households with Catastrophic Health Expenditure, by Survey Year and Threshold (Denominator: Non-food Expenditure)



Source: Household Budget Surveys (2002-2008), TurkStat.

Table A2.2 presents the results of bivariate analysis which evaluates statistical relationships between the household characteristics and the rates of catastrophic health expenditure for the pooled data set. Chi-squared tests are used to identify statistically significant associations. In order to obtain the results of the chi-squared tests, the weighted categories and incidence of catastrophic health expenditure were tabulated for each threshold level[22]. This table provides information about how the relationship between the key characteristics and the proportion of households with catastrophic health expenditure changes when the definition of catastrophic health expenditure (i.e., threshold level and denominator) changes.

Rates of catastrophic health expenditure by pooled consumption quintiles indicate that there is a significant increasing trend in the risk of incurring catastrophic health expenditure from the poorest to the richest quintiles. Richer households are significantly more likely to spend a higher percentage of their income on health care than poorer households (the joint chi^2 p-value is between 0.000 and 0.074). However, when non-food expenditure is used as a denominator, the risk of incurring catastrophic health expenditure is increased for the households in the poorest quintile relative to using total expenditure as a denominator. It is important to emphasise that the progressive trend (i.e., the risk of catastrophic health expenditure increasing with income level) is apparent for most definitions of catastrophic health expenditure.

The association between poverty status and catastrophic health expenditure indicates the same progressive pattern because 'non-poor' households are more likely to exceed all threshold levels than 'poor' households. This relationship is statistically significant for seven definitions of catastrophic health expenditure. Although the system appears progressive, this pattern could reflect potential selection bias. If the households, which choose not to seek health care because they expect the health costs to be unaffordable, are not considered, this may lead to biased and misleading results. Therefore, it is important to consider the relationship between health care seeking behaviour and incurring catastrophic health expenditure. Table 2.4 below presents the cross-tabulation results for poor households and households with positive health care expenditure for the pooled data.

[22] Stata version 10 was used for the analysis. A sample weight variable which was provided by Turkstat for each HBS is used in order to correct for sampling bias.

Table 2.2: The Cross-tabulation Results for Poverty and Positive Health Care Expenditure

Poor households with positive health expenditure	36.4%
Poor households with zero health expenditure	63.6%
Non poor households with positive health expenditure	49.8%
Non poor households with zero health expenditure	50.2%
If there is at least one disabled or ill member in the household	
Poor households with positive health expenditure	45.9%
Poor households with zero health expenditure	54.1%
Non poor households with positive health expenditure	64.8%
Non poor households with zero health expenditure	35.2%

It can be seen that 49.8% of non-poor households seek health care (i.e. have positive health expenditure) while 36.4% of poor households have positive health expenditure. When the proxy for household health status is taken into consideration, the relative health seeking position of poor households becomes more apparent. 65% of non-poor households, in which at least one household member is disabled or ill, have positive health expenditure whereas 46% of poor households seek health care when there is at least one disabled or ill member in the household.

Returning to Table A2.2, the relationship between catastrophic health expenditure and the highest level of education of the household head indicates that the households in which the head of household has 'primary or less' education are at greatest risk of experiencing catastrophic health expenditure except for the 2.5% threshold level for the total expenditure denominator. As the education level increases, the probability of incurring catastrophic health expenditure decreases for nearly every definition of catastrophic health expenditure and these relationships are all statistically significant at the 1% level of significance.

The association between the employment status of the household head and rates of catastrophic health expenditure is interesting since the households in which the household head is self-employed are at greatest risk of incurring catastrophic health expenditure for every definition except for the highest threshold level for the total expenditure denominator. However, households in which the household head is employed are at less risk of incurring catastrophic health expenditure compared to households in which the household head is not employed.

Household size is not linearly related to the risk of incurring catastrophic health expenditure. For six definitions of catastrophic health expenditure, the households in

which two members are living are more likely to experience catastrophic health expenditure than households with 3 or more members and individuals living alone. For the other definitions of catastrophic health expenditure, individuals living alone have the highest risk of experiencing catastrophic health expenditure.

The gender of the household head indicates that female-headed households are more likely to incur catastrophic health expenditure than male-headed households at the 10%, 15% and 20% threshold levels of total expenditure and all threshold levels of non-food expenditure except the lowest threshold level but these associations are not statistically significant at most of the threshold levels.

The proxy for the health status of the household (the presence of a disabled or ill household member in the household) is strongly and statistically significantly associated with exceeding all threshold levels for both denominators. The absence of health insurance of the household head is associated with higher rates of catastrophic health expenditure and this pattern is statistically significant and consistent across all definitions.

The association between certain age groups, the number of preschool children and the number of elderly household members, and rates of catastrophic health expenditure is as expected whereas this is not the case for the group of school children. Households with more members in the school age children group are less likely to incur catastrophic health expenditure. However, households with more members in the preschool and elderly age groups are more likely to experience catastrophic health expenditure which accords with expectations since higher health care costs are generally associated with these groups.

The urban/rural region variable indicates that rural households are statistically significantly associated with experiencing catastrophic health expenditure compared to urban households. Urban households generally have better access to health care facilities compared to rural households and, therefore, this pattern is consistent with expectations.

Table 2.3 below presents the percentages of types of out-of-pocket health expenditure by survey year. It is apparent that the category of ‘medicines and other medical products’ constitutes the largest part of health expenditure. This is consistent with the

recent health reforms in Turkey. Individuals are required to pay a co-payment for drug expenditure which was decreased after the implementation of the HTP but is still burdensome for households. Although 'medical services' is the second largest category, it includes three different components namely physician services, nursing care services and health expenditure not related to hospital. It can also be said that hospitalisation costs and the cost of dentistry services are also among the important sources of out-of-pocket health expenditure in Turkey.

Table 2.3: Types of Out-of-pocket Health Expenditure by Survey Year

	2002	2003	2004	2005	2006	2007	2008
Medicines and other medical products	37.8%	37.8%	34.2%	32.4%	30.4%	31.4%	31.5%
Treatment equipment	4.9%	5.2%	4.5%	5.1%	7.4%	5.7%	6.8%
Medical services*	27.1%	27.2%	29.1%	30.6%	30.9%	27.6%	27.9%
Dentistry services	10.4%	9.5%	13.8%	12.3%	11.9%	14.8%	12.1%
Laboratories and X-ray Services	7.3%	7.4%	7.1%	6.7%	8.9%	6.9%	8%
Hospital services	12.5%	12.9%	11.3%	12.9%	10.5%	13.6%	13.7%

*Medical services include physician services, nursing care services and health expenditure not related to hospital.

In the context of analysing the determinants of catastrophic out-of-pocket health care expenditure in Turkey, the next section discusses the results from the econometric analysis.

2.5 RESULTS

In this section, the results of estimating the probit and Sartori selection models are presented and discussed. In addition, the results of the Heckman probit model (without an exclusion restriction) are compared to the results of the Sartori selection model in the context of the potential sample selection bias. Tables presenting the results for all definitions of catastrophic health expenditure are presented in the Appendix.

As a sensitivity analysis, ten different definitions of catastrophic health expenditure (five threshold levels for total expenditure and five threshold levels for non-food expenditure) are used and the results for both the probit and the Sartori selection models are generally robust across all definitions. There are only some slight differences in the statistical significance levels of the education and employment status of the household head and some of the year controls. Furthermore, many of the statistically significant risk factors in the Sartori selection model are found to be similar to those in the probit model. The most pronounced and arguably important finding relates to the association between the poverty status of the household and the likelihood of incurring catastrophic health expenditure after accounting for the household's health care seeking position which will be discussed in detail below[23].

The results of the probit model indicate that poor households are less likely to experience catastrophic health expenditure compared to non-poor households. Table A2.3 indicates that, considering all of the explanatory variables at their mean values, poor households have a 6.6% lower probability of experiencing catastrophic health expenditure compared to non-poor households at the 2.5% threshold level of total expenditure. This probability decreases in magnitude at higher threshold levels for both denominators. Tables A2.7 and A2.12, for example, indicate that poor households have 0.8% and 0.3% lower probability of incurring catastrophic health expenditure at the

[23] In order to control for the socio-economic status of the households, poverty status is included in the analysis since one of the aims is to analyse the influence of being poor in the context of the risk of catastrophic health expenditure. An alternative option is to use total expenditure (a proxy for income) as an independent variable in the analysis. However, the denominator of the dependent variable and the poverty indicator are constructed using total expenditure so including total expenditure as an explanatory variable may lead to biased estimation results due to the high correlation between poverty and income. When both were included as explanatory variables, the estimated coefficient of the poverty indicator became statistically insignificant in all models. On the other hand, both the dependent variable and the poverty indicator are functions of total expenditure. However, as discussed in Section 2.3, this approach is very common in the existing literature (see, for example, Xu *et al.*, 2003; Berki, 1986; Wyszewianski, 1986) to explore the association between the socio-economic status of the household and the probability of experiencing catastrophic health expenditure.

highest threshold level of total expenditure and non-food expenditure, respectively. Health insurance coverage is inversely associated with the probability of catastrophic health expenditure and households with a disabled or ill member are more likely to incur catastrophic health expenditure. These effects are highly statistically significant in the case of both catastrophic health expenditure definitions and at any threshold. Household heads with higher levels of education are associated with a lower risk of catastrophic health expenditure at seven definitions of catastrophic health expenditure compared to households with a head with primary or less education. The employment status of the household head is also a significant determinant of the probability of experiencing catastrophic health expenditure. Having a self-employed head of household is positively associated with the probability of catastrophic health expenditure for eight threshold levels. The results of the probit model further indicate that there is a statistically significant and negative association between the likelihood of catastrophic health expenditure and household size. At this point, it should be noted that the quadratic term of household size is also controlled for in the analysis and the marginal effect of household size represents one unified marginal effect for household size and its squared term. One of the more consistently statistically significant variables is the region of residence of the household. Living in an urban area is highly statistically significantly associated with a lower risk of experiencing catastrophic health expenditure at all threshold levels. Similarly, having more preschool children and more elderly household members are both positively associated with the likelihood of catastrophic health expenditure.

The results of the Sartori selection model are generally similar to the results of the probit model. However, the magnitude of the marginal effects becomes larger when the health care seeking behaviour of the households is controlled for. The selection equation regression results indicate that poor households are much less likely to seek health care than non-poor households for all threshold levels and that poverty status has the largest marginal effect (ranging from -0.204 to -0.206 across all threshold levels[24]) compared to the other explanatory variables. This finding is consistent with the idea that poor households may not seek health care due to high health care costs[25]. However, when the Sartori selection model is used to adjust for the health care seeking position of

[24] The marginal effects are larger in the case of higher thresholds for both denominators.

[25] The reason for not seeking treatment may also reflect households that report zero health expenditure having better health.

households, the results surprisingly suggest that poor households are less likely to experience catastrophic health expenditure compared to non-poor households[26]. This finding can arguably be attributed to using more private facilities among non-poor households. It is commonly acknowledged that patients can receive a better quality service in private health facilities due to the lack of confidence in the public health care facilities in Turkey (Savas *et al.*, 2002). As an example of this issue, Yildirim *et al.*, (2011) used maximum likelihood logistic analysis to investigate the determinants of out-of-pocket health expenditure in Turkey using survey data obtained from the major hospitals in Ankara. They used a binary dependent variable based on whether the patients had any out-of-pocket health expenditure while accessing the health care facilities and their particular interest was placed upon the relationship between the attitudes of patients towards the health care system and the decision of making out-of-pocket health expenditure. Their findings indicated distrust in the public health care facilities since they found that patients with higher levels of income chose private health care rather than the public health care services. They also found that there is a positive association between the level of income and the likelihood of incurring out-of-pocket health expenditure. Moreover, as stated in Section 2.2 above, before the health reforms, patients using private health care were paying for services out-of-pocket, even if they have health insurance. After the health reforms, access to private facilities was improved with contracted private hospitals but in order to encourage private sector to contract with the Social Security Institute, private hospitals are allowed to implement 'extra charge' (OECD, 2008). Therefore, it can be argued that this improvement in access to private health care particularly benefited the non-poor segment of the population who can pay for the extra charge imposed by the private provider. This may create demand inducement among non-poor households who prefer private health care rather than public health care and may lead to a higher probability of incurring catastrophic health expenditure. However, this hypothesis cannot be explored in the empirical analysis since the HBS does not provide information on the type of health care provider such as public or private.

The results of the Sartori selection model confirm the results of the probit model in terms of the protective effect of health insurance coverage against catastrophic health

[26] The magnitude of this effect is larger (more than twice as large at all definitions) in the Sartori selection model compared to the probit model.

expenditure. Furthermore, as expected, the results of the selection equation indicate that households without any health insurance coverage are less likely to seek health care as compared to households with health insurance coverage. This finding confirms the important role of insurance coverage in terms of providing financial protection[27]. However, 19.7% of household heads do not have any health insurance in this sample and only 15.6% of poor household heads have health insurance. This highlights the vulnerability of poor households to the risk of catastrophic health expenditure.

The results of estimating the selection equation suggest that the presence of a disabled or ill individual in the household is positively associated with the probability of seeking health care. This positive marginal effect is the second largest effect in magnitude following the effect of the poverty status of the household in the selection equation. Similarly, households with a disabled or ill member are more likely to experience catastrophic health expenditure. This association is statistically significant for all definitions of catastrophic health expenditure. Furthermore, the household health status proxy has the largest marginal effect in the outcome equation. This finding indicates that the presence of a disabled or ill individual in the household appears to be the most important risk factor for catastrophic health expenditure for the period considered in this study[28]. According to this relationship, for example in Table A2.5, considering all of the explanatory variables at their mean values, the presence of a disabled or ill member in the household will cause 7.4% increase the probability of incurring catastrophic health expenditure.

The estimation results of the Sartori selection model support the protective effect of education on the probability of catastrophic health expenditure. Household heads with the secondary or higher level of education as the highest level of education are less likely to incur catastrophic health expenditure compared to household heads with primary education or less. Furthermore, the inverse association between higher levels of education and the risk of catastrophic health expenditure has become more pronounced

[27] It should be noted that private health insurance is also likely to be subject to selection bias because it is generally purchased by relatively well-off households or the households with the highest risk of illness or the most risk averse households (Hatt, 2006). However, private insurance is not common in Turkey and only 1.1% of household heads have private health insurance in this sample.

[28] As highlighted above, the measure of out-of-pocket health expenditure used in this study covers all health expenditure. Having a disabled or ill member in the household may lead to both expected and unexpected health expenditure.

when health care seeking is adjusted for with the Sartori selection model as compared to the probit model. Similarly, the employment status of the household head is among the important predictors of catastrophic health expenditure. Households with a self-employed head are associated with a higher probability of incurring catastrophic health expenditure whereas households with an employed head are inversely associated with the risk of experiencing catastrophic health expenditure. The inverse relationship between households with an employed head and the risk of experiencing catastrophic health expenditure is statistically significant at all threshold levels of non-food expenditure with the exception of the highest threshold level. The selection equation results, on the other hand, indicate that households whose heads are self-employed or employed are more likely to seek health care compared to households whose heads are not employed. These findings imply that the unsalaried work position of the self-employment is not a barrier to seeking health care but is associated with a higher probability of experiencing catastrophic health expenditure.

With regard to household size, the results of the Sartori selection model reinforce the results of the probit model in terms of the negative association between the likelihood of catastrophic health expenditure and household size. In general, larger households are more likely to be concentrated in the lower socioeconomic quintiles and have more dependent individuals and, thus, they arguably have less resources for health care (Hatt, 2006). In this context, it is expected that catastrophic health expenditure risk increases with household size. However, it may be the case that large households may pool their income which may decrease the risk of experiencing catastrophic health expenditure. Having more income earners in the household, for example, may also lead to such an association. The results from the selection equation indicate that larger households have a higher probability of seeking health care compared to smaller households. This finding highlights the possibility that a risk factor resulting in a higher probability of seeking health care does not always mean a higher probability of experiencing catastrophic health expenditure.

The age composition of the household is also important in terms of the risk of experiencing catastrophic health expenditure. The number of members of the household belonging to more risky groups in terms of health status, those aged under 5 or above 65 are both positively associated with a high risk of experiencing catastrophic health expenditure across all definitions. The results from the selection equation indicate that

an increase in the number of preschool children and the number of elderly members in the household are both associated with the household's health care seeking propensity. These findings are not surprising since elderly household members generally need more frequent and expensive health care and, in addition, they tend to have significantly reduced income (or they may be living as a dependent of other family members) (Hatt, 2006). Similarly, having more preschool children may lead to an increased demand for health care since they may experience early age illnesses as well as needing preventive health care services. The number of school children, on the other hand, is found to be positively related to the probability of incurring catastrophic health expenditure despite being associated with a lower probability of seeking health care. However, the positive association between the number of children and the probability of catastrophic health expenditure is only statistically significant at the 15%, 20%, 30% and 40% threshold levels of non-food expenditure.

The location of the household is also a statistically significant determinant of catastrophic health expenditure when the health care seeking behaviour of households is controlled for with the Sartori selection model. Urban residence is found to be inversely associated with the risk of incurring catastrophic health expenditure and this association is statistically significant in every model and at all threshold levels. Urban residence is also associated with a higher probability of seeking health care as compared to rural residence which may be due to, for example, better access to health care facilities.

Finally, the estimation results of the Sartori selection model indicate that the gender of the household head is not a statistically significant determinant of catastrophic health expenditure for all definitions of catastrophic health expenditure and the year controls show that the rate of catastrophic health expenditure falls slightly in the later years until 2008 and there is a sharp decline in 2008.

In order to compare the findings from the Sartori selection model with those from the Heckman selection model, the estimation results of the Heckman selection model without an exclusion restriction[29] (i.e. the same explanatory variables are included in both the selection and outcome equations) are presented in Tables A2.13 and A2.14 for the middle threshold levels for each denominator. As stated in Section 2.4.2 above, the

[29] The choice of the middle threshold levels reflects the fact that these levels are the most commonly used threshold levels among all threshold levels in the existing literature (see, for example, Pradhan and Prescott, 2002; O'Donnell *et al.*, 2005; Wagstaff and van Doorslaer, 2003; Ranson, 2002).

Heckman selection model can be estimated with identical sets of explanatory variables and, in that case, the results are based upon distributional assumptions about the error terms rather than upon variation in the explanatory variables (Maddala, 1999; Sartori, 2003). Such an approach seems appropriate given that the main aim of analysing the Heckman selection model is to explore the robustness of the results from the Sartori selection model.

In general, the risk factors of catastrophic health expenditure identified by the Heckman selection model are similar to those identified in the probit and the Sartori selection models. In particular, the results of the Heckman selection model confirm the findings from the Sartori selection model related to the poverty status of the households. According to the findings, poor households are less likely to seek health care and being poor is inversely associated with the risk of catastrophic health expenditure. In addition, rho (ρ) values indicate higher positive correlation between selection and outcome error terms which can arguably be seen as an evidence for the basic assumption of the Sartori selection model.

The Sartori selection model, where the basic assumption is identical error terms in both the selection and outcome equations, is a special case of the Heckman selection model. Therefore, another robustness check is explored by imposing constraints on the correlation coefficient between the selection and outcome error terms (ρ) for different levels from 0 (i.e., no correlation of the error terms in the two equations) to 0.9. The results of the 'constrained' Heckman selection model are found to be consistent with the other models and the results where the correlation coefficient is set to 0.5 for the middle threshold levels for each denominator are presented in Tables A2.15 and A2.16 in the Appendix.

2.5.1 Summary of the Key Findings

One of the aims of this chapter was to clarify the association between the poverty status of the household and the likelihood of incurring catastrophic health expenditure while controlling for the household's health care seeking position. The results from the Sartori selection models highlight the vulnerability of poor households in terms of their health care seeking position since poor households are found to be less likely to seek health care compared to non-poor households. However, it should be acknowledged that reporting zero health expenditure may also reflect having a better health condition rather

than indicating the vulnerability of poor households in terms of their treatment seeking behaviour. Poor households are also found to be less likely to experience catastrophic health expenditure compared to non-poor households after accounting for the potential selection bias. The reason for the higher probability of incurring catastrophic health expenditure for non-poor households may arguably be attributed to their preferences towards private facilities, which reflects another existing inequality between poor and non-poor households in Turkey.

With regard to the risk factors of catastrophic health expenditure, the results of the Sartori selection model also indicate that the presence of a disabled or ill member in the household is the most important risk factor for catastrophic health expenditure. Similarly, households with more elderly members or preschool children are at increased risk of experiencing catastrophic health expenditure. Higher levels of education, living in an urban area and insurance coverage are all found to be protective factors against the risk of catastrophic health expenditure.

The results of the Heckman selection and 'constrained' Heckman selection models also reinforce the findings from the Sartori selection model related to the determinants of catastrophic health expenditure. In this context, it can be argued that the attempt to account for the potential selection bias does not dramatically influence the observed relationships between the risk factors and the probability of experiencing catastrophic health expenditure but adjusting for health care seeking choices does increase the effects of all of the risk factors in terms of magnitude. Furthermore, attempting to account for the health care seeking position of the households has also provided information on the relationships between the potential factors and the probability of seeking health care. In this respect, the marginal effects of the selection equation generally remain unchanged for each different outcome variable and the coefficient signs of the explanatory variables are as expected. Moreover, the levels of statistical significance are very high for most of the explanatory variables.

2.6 CONCLUSION

This chapter has explored the important research question of identifying the risk factors associated with experiencing catastrophic health expenditure at the household level in Turkey and has provided a number of interesting insights on this issue. Out-of-pocket health expenditure is defined as 'catastrophic' if the ratio of the household health expenditure to total household expenditure or non-food expenditure is more than a pre-specified threshold value. In this analysis, a number of threshold levels, ranging from 2.5% to 20% for the denominator of total household expenditure and from 10% to 40% for the denominator of total non-food expenditure, are used to provide a comprehensive picture of catastrophic health expenditure and its determinants and to explore the sensitivity of the results. The rates of catastrophic health expenditure are quite similar across the years with the exception of 2008. There is a significant decrease in the rate of households incurring catastrophic health expenditure in 2008, which is the year that the implementation of the UHI system was started. Nearly 5% of households spend at least 10% of their budget on health care which is equivalent to more than 3.5 million individuals. However, this high rate of catastrophic health expenditure arguably underestimates the full 'cost burden' of health care which refers to both direct (e.g. medical costs and non medical costs such as transportation, accommodation and special food) and indirect costs (e.g. lost working time) related to health care seeking because non medical and indirect costs unfortunately cannot be included in this analysis due to data unavailability.

To the author's knowledge, the empirical analysis presented in this chapter is the first attempt to investigate the determinants of catastrophic health expenditure, particularly the position of poor households, in the context of adjusting for the health care seeking behaviour of households in Turkey. The results suggest that poor households are much less likely to seek health care relative to non-poor households. This finding is consistent with the situation where poor households choose not to seek health care when they face an illness possibly due to lack of access or inability to pay for health care costs. On the other hand, the results of the Sartori selection model surprisingly indicate that poor households are less likely to incur catastrophic health expenditure after adjusting for the potential selection bias. This finding could be explained by the poor and non-poor households' choices of type of health care provider, such as public or private, which cannot be included in the analysis due to data limitations. As stated in detail in Section

2.2, public health services are often criticised as being unsatisfactory and patients think that they can obtain a better service from private facilities. Before the health reforms, patients using private facilities were paying for services out-of-pocket regardless of their membership of any social insurance institutions (Savas *et al*., 2002). However, after health reforms, patients from all social insurance institutions gained access to private facilities in contracted private hospitals but in order to stimulate private sector interest in contracting with the SSI, private hospitals were allowed to implement 'extra billing' (OECD, 2008). This improvement in access to private health care may lead to demand inducement for non-poor households for the time period considered in the analysis and to increase their out-of-pocket health expenditure due to extra charges imposed by the private provider and finally result in catastrophic health expenditure.

The results also indicate that households with a disabled or ill member and households with more preschool or elderly members are more likely to seek health care and are more likely to experience catastrophic health expenditure. This finding is consistent with the hypothesis that these groups are at most need of protection against catastrophic health expenditure risk since they tend to have more health problems and require more expensive and repeated treatments. In addition, temporary health shocks tend to be more serious for these groups (Hjortsberg, 2003; Hatt, 2006). Higher levels of education, employment and insurance coverage are all found to be protective factors against the risk of catastrophic health expenditure. Heads of household who are self-employed appear to be less able to provide financial protection for their families but households with self-employed and employed heads are more likely to seek health care than unemployed heads. Urban households and households with insurance coverage have a higher probability of seeking health care and a lower probability of incurring catastrophic health expenditure relative to rural households and households whose head has no health insurance. Interestingly, the results suggest that, as household size increases, the probability of seeking health care increases but the probability of incurring catastrophic health expenditure decreases. This finding can be explained by the potential presence of more income earners in the household and a higher propensity for income pooling. In the context of these findings, it is possible to state that a risk factor which is positively associated with the probability of seeking health care is not necessarily associated with an increase in the risk of catastrophic health expenditure. The findings are also relatively robust to changes in the definition of catastrophic health

expenditure and the results of the three estimation methods used in the analysis, namely the probit model, the Sartori selection model and the Heckman selection model, are consistent in terms of the risk factors associated with catastrophic health expenditure.

The empirical analysis presented in this chapter sheds light on the household characteristics associated with catastrophic health expenditure and this may help policy-makers to design appropriate policies such as targeting vulnerable households, which have specific risk factors, with for example exemptions from health care costs or the provision of subsidised health care services. One of the aims of the Health Transformation Programme in Turkey, which has been implemented since 2003, is to organise and deliver health care services on the basis of financial accessibility. The findings presented in this chapter related to the socio-economic distribution of catastrophic health expenditure suggest that non-poor households spend more on health care, potentially leading to better health relative to poor households, and that poor households are much less likely to seek health care. The results also indicate that health insurance coverage potentially provides important protection against the probability of catastrophic health expenditure. Given that the analysis uses data for the period of 2002-2008, it can be argued that the steps towards unifying health insurance schemes have started to be effective on financial protection. Therefore, it appears that health reforms should continue to cover those households that are more vulnerable to catastrophic health expenditure risk and do not seek health care since they think that the cost of health care is unaffordable and that universal health insurance should include the groups, which currently face a lack of formal coverage such as informal-sector workers and other non-contributory groups. Furthermore, insurance coverage mechanisms could be expanded because an effective and sustainable insurance system can avert the risk of catastrophic health expenditure by pooling the resources of a great number of people (Hatt, 2006). In order to create a fair insurance premium, the individual's expected costs arising from a health problem rather than the risk of illness should be taken into account (WHO, 2000). In the context of the UHI, the means-tested system, which provides financial protection for individuals, who cannot afford their insurance premiums, is already in place but a carefully regulated private health insurance system may also be useful in decreasing the burden of catastrophic health costs. In addition, the low probability of seeking health care for poor households implies substantial barriers to health care access. Therefore, special attention should be devoted to overcome these

cost barriers particularly for poor households. A lack of adequate and timely health care may increase the severity of illness, increase the risk of death and finally result in increased costs to society.

Another potential policy relates to abating the co-payments for drug expenditure particularly for households with elderly members or disabled/ill members because expenditure on drug is the largest proportion of out-of-pocket health expenditure in Turkey. This can be seen as an important step towards a more comprehensive insurance coverage.

Throughout this chapter, some empirical and conceptual limitations are stated and acknowledged. It is important to summarise these limitations and to mention a few additional limitations in the context of evaluating the findings. In order to fully estimate the economic effect of illness costs, it is necessary to take into consideration both indirect and direct costs. Evaluating the indirect costs of illness cannot be fully undertaken without analysis of the methods used by households to cope with them (Russell, 2004). Gertler and Gruber (2002) note that ideally longitudinal data would be used to examine how household expenditure on non-medical goods and services changes following a health shock. If health costs are financed by cutting back on current consumption, the opportunity cost (reducing consumption of other goods and services in order to pay for health care) may be incurred in the short term but if health care costs are financed by credit or savings, the opportunity cost may be incurred in the long term. It is difficult to discern between the short and long term effects with cross-section data (O'Donnell *et al.*, 2008). However, the Household Budget Surveys do not follow the same households through time and, thus, how the households manage and cope with health costs over time cannot be included due to the lack of longitudinal data.

It is also important to note that the results of the Sartori selection model should be treated with caution. There are two particular limitations of this approach. Firstly, as mentioned in Section 2.4.2, the sample used in the Sartori selection model is not restricted to the individuals who 'needed' health care. The aim is to analyse the risk of catastrophic health expenditure taking account of the households who need health care but cannot afford it. However, due to data limitations it is difficult to define an indicator that reflects the need for health care. If the presence of a disabled or sick member in the household is selected as an indicator for needing health care (only 8.1% of total

households), this may lead to the exclusion of some households in which there is a member with another type of illness. A more precise measure such as a health status index or measure of illness types for every individual in the household would be needed to control for the nature of the selection bias. Unfortunately, such information is not available in the data set. Secondly, the 'having or not having positive health expenditure' dichotomy modelled in the selection equation may have led to the loss of some information such as the type of provider or the number of household members seeking health care (Hatt, 2006). However, data capturing such dimensions of health care seeking behaviour are not available.

An additional limitation related to all the models concerns the aggregation process. The household is the unit of analysis and it is assumed that if one of the members of the household gets sick, this will affect the entire household. Therefore, the findings are evaluated as a summary of individual risk factors at the household level and some information about the interaction between individual risk and household treatment seeking behaviours may be lost in such a process.

Notwithstanding these limitations, analysis of the expenditure on health care in excess of a substantial fraction of the household's budget is informative in the context of the economic consequences of illness. For future research, examining the effects of catastrophic health expenditure over time would be an important contribution to the literature but this depends on the availability of the relevant longitudinal data which are not currently available for Turkey. Decomposing total health expenditure into categories and investigating which of the components is mostly related to the risk of experiencing catastrophic health expenditure would also be potentially useful for policy makers in identifying which segments of the population are more likely to incur catastrophic health expenditure for each component.

APPENDIX TO CHAPTER 2

Table A2.1: Descriptive Statistics for the Continues Variables and Percentage Distributions for the Categorical Variables across Survey Years

	2002	2003	2004	2005	2006	2007	2008
Continuous Variables							
Total expenditure*	412.7	328.4	400.7	491.7	396.9	419.6	451.9
(St. Dev.)	(449.8)	(293.9)	(341.7)	(407.8)	(317.3)	(320.7)	(338.2)
(Min)	11.7	8.7	26.2	11.1	11.9	16.1	6.4
(Max)	18993.0	7328.9	7473.2	6126.7	4535.9	7852.0	4782.8
Total health exp.*	8.6	6.8	8.8	10.7	8.5	9.1	8.1
(St. Dev.)	(32.3)	(32.9)	(34.5)	(45.4)	(37.4)	(42)	(43.4)
(Min)	0	0	0	0	0	0	0
(Max)	1017.2	2174.5	1022.3	2197.5	1671	1290.7	2922.6
Household size	4.25	4.17	4.14	4.14	4.08	4.04	3.89
(St. Dev.)	(2.01)	(2.04)	(2.02)	(2.04)	(1.89)	(1.99)	(1.83)
(Min)	1	1	1	1	1	1	1
(Max)	20	23	19	22	23	22	23
Preschool children	0.47	0.43	0.42	0.41	0.41	0.39	0.37
(St. Dev.)	(0.74)	(0.72)	(0.72)	(0.70)	(0.69)	(0.69)	(0.66)
(Min)	0	0	0	0	0	0	0
(Max)	6	6	6	6	7	8	5
Children	0.81	0.78	0.78	0.80	0.76	0.74	0.69
(St. Dev.)	(1.08)	(1.05)	(1.05)	(1.07)	(1.03)	(1.02)	(0.97)
(Min)	0	0	0	0	0	0	0
(Max)	10	9	7	11	9	8	10
Elderly	0.22	0.25	0.24	0.25	0.24	0.24	0.26
(St. Dev.)	(0.52)	(0.56)	(0.54)	(0.55)	(0.54)	(0.55)	(0.56)
(Min)	0	0	0	0	0	0	0
(Max)	4	4	3	3	3	3	3
Categorical Variables (%)							
Households with out-of-pocket health expenditure exceeding % of total expenditure							
2.5%	17.8	15.7	17.0	17.5	17.6	17.1	15.7
5%	10.6	9.9	10.8	10.8	10.8	10.5	8.4
10%	5.0	4.8	5.4	5.4	5.2	5.0	3.8
15%	2.8	2.6	3.0	3.1	2.8	2.9	2.0
20%	1.6	1.6	1.9	1.7	1.6	1.7	1.2
Households with out-of-pocket health expenditure exceeding % of non-food expenditure							
10%	8.2	7.6	8.2	8.0	7.8	7.4	5.7
15%	4.7	4.8	5.2	5.3	4.8	4.6	3.5
20%	3.2	3.2	3.6	3.6	3.0	3.2	2.2
30%	1.5	1.4	1.7	1.5	1.4	1.4	1.0
40%	0.7	0.8	1.0	0.7	0.7	0.8	0.5
Education of household head (reference: primary or less)							
Primary (or less)	74.8	74.0	72.8	74.6	73.8	73.3	70.0
Secondary	16.2	16.6	17.7	16.3	16.5	17.2	18.2
Higher	9.0	9.4	9.5	9.1	9.7	9.5	11.8
Employment status of household head (reference: not employed)							
Not employed	29.8	30.1	30.6	30.3	28.6	31.2	32.0
Employed	45.6	40.0	41.1	41.6	44.6	42.8	42.8
Self-employed	24.6	29.9	28.3	28.1	26.8	26.0	25.2

Table A2.1 continued: Descriptive Statistics for the Continuous Variables and Percentage Distributions for the Categorical Variables across Survey Years

Gender of household head							
Male	89.9	90.5	89.4	89.7	89.9	89.3	88.6
Female	10.1	9.5	10.6	10.3	10.1	10.7	11.4
Health Status (Presence of disabled or ill member in the household)							
Yes	4.1	3.7	11.1	10.8	11.5	11.9	13.0
No	95.9	96.3	88.9	89.2	88.5	88.1	87.0
Health insurance status of household head							
No	23.2	26.1	21.9	18.9	13.9	11.5	8.9
Yes	76.8	73.9	78.1	81.1	86.1	88.5	91.1
Location of residence							
Urban	84.7	70.9	70.0	69.9	69.3	68.9	69.7
Rural	15.3	29.1	30.0	30.1	30.7	31.1	30.3
Poverty							
Poor	20.0	20.5	22.6	21.7	21.5	20.7	20.8
Not Poor	80.0	79.5	77.4	78.3	78.5	79.3	79.2

**Notes: ***Turkish Lira (TL) values are converted to British Pound (£) values using 2002 as a base year.

Table A2.2: Household Characteristics in relation to the Rates of Catastrophic Health Expenditure (Full Dataset: 2002-2008)

	Denominator: Total expenditure						*Denominator: Non-food expenditure*				
Threshold	2.5%	5%	10%	15%	20%		10%	15%	20%	30%	40%
Pooled expenditure quintile (from poorest to richest) (%)											
quintile 1	12.8	7.7	4.1	2.5	1.5		6.8	4.6	3.4	1.6	0.7
quintile 2	15	9.2	4.6	2.3	1.2		7.8	5	3.2	1.2	0.6
quintile 3	16.3	10.2	5	2.5	1.3		7.8	4.8	3	1.2	0.5
quintile 4	18	10.9	4.7	2.6	1.6		7.5	4.3	2.8	1.3	0.7
quintile 5	21.2	12.8	5.9	3.5	2.4		8	4.9	3.4	1.9	1.2
Chi2 p-value	*.000*	*.000*	*.000*	*0.000*	*.000*		*.001*	*.074*	*.018*	*.000*	*.000*
Poverty status (%)											
Poor	13.1	7.8	3.8	2	1.2		7.1	4.5	3.2	1.4	0.6
Non poor	17.6	10.8	5.2	2.9	1.7		7.7	4.7	3.2	1.4	0.8
Chi2 p-value	*.000*	*.000*	*.000*	*.000*	*.000*		*.004*	*.234*	*.983*	*.724*	*.060*
Education of household head (%)											
Primary (less)	16.8	10.4	5.2	2.9	1.8		8.2	5.3	3.6	1.6	0.8
Secondary	15.7	9.3	4.2	2.1	1.2		6.3	3.4	2.1	0.9	0.4
Higher	17.1	9.6	3.7	1.9	1.2		5.4	2.7	1.7	0.8	0.4
Chi2 p-value	*.005*	*.000*	*.000*	*.000*	*.000*		*.000*	*.000*	*.000*	*.000*	*.000*
Employment status of household head (%)											
Not employed	16.9	10.4	5.3	3.1	2		7.9	5.1	3.5	1.7	0.9
Employed	15.8	9.3	4	2.1	1.2		6.3	3.5	2.2	0.9	0.5
Self-employed	17.8	11.3	5.8	3.1	1.9		9.2	6.1	4.1	1.9	1
Chi2 p-value	*.000*	*.000*	*.000*	*.000*	*.000*		*.000*	*.000*	*.000*	*.000*	*.000*
Household size (%)											
1	16.5	10.8	6.5	4.7	3.3		8.1	5.6	4.5	2.8	1.3
2	19	12.3	6.9	4.2	2.7		9.8	6.6	4.8	2.5	1.3
3	16.8	10.4	4.7	2.5	1.5		7.2	2.3	2.7	1.2	0.7
4 and more	16	9.5	4.3	2.3	1.3		7.1	4.3	2.8	1.2	0.6
Chi2 p-value	*.000*	*.000*	*.000*	*.000*	*0.000*		*.000*	*.000*	*.000*	*.000*	*.000*
Gender of household head (%)											
Male	16.7	10.2	4.9	2.6	1.6		7.6	4.7	3.1	1.4	0.7
Female	16.5	10.1	5.2	3.2	2		7.5	4.9	3.5	1.8	0.9
Chi2 p-value	*.732*	*.713*	*.168*	*.003*	*0.001*		*.802*	*.418*	*.053*	*.005*	*.192*
Health Status (Presence of disabled or ill member in the household) (%)											
Yes	25.3	16.3	9.2	5.4	3.4		13.5	9.2	6.5	3.3	1.8
No	15.9	9.6	4.5	2.5	1.4		7.1	4.3	2.9	1.3	0.7
Chi2 p-value	*.000*	*.000*	*.000*	*.000*	*.000*		*.000*	*.000*	*.000*	*.000*	*.000*
Health insurance (%)											
No	18	11.8	6.2	3.6	2.7		9.9	6.5	4.6	2.1	1.3
Yes	16.3	9.8	4.6	2.5	1.4		7	4.3	2.8	1.2	0.6
Chi2 p-value	*.000*	*.000*	*.000*	*.000*	*.000*		*.000*	*.000*	*.000*	*.000*	*.000*
Number of preschool children (%)											
0	15.8	9.7	4.8	2.8	1.7		7.1	4.5	3.1	1.5	0.8
1	18.2	10.8	5	2.5	1.4		7.9	4.7	2.9	1.2	0.6
2 and more	19.2	12.2	5.8	2.9	1.5		10.2	6.3	4	1.7	0.9
Chi2 p-value	*.000*	*.000*	*.001*	*.061*	*.082*		*.000*	*.000*	*.000*	*.004*	*.025*

Table A2.2 continued: Household Characteristics in relation to the Rates of Catastrophic Health Expenditure (Full Dataset: 2002-2008)

	Denominator: Total expenditure						*Denominator: Non-food expenditure*				
Threshold	2.5%	5%	10%	15%	20%		10%	15%	20%	30%	40%
Number of children (%)											
0	17.6	11	5.5	3.1	1.9		8.1	9.2	3.5	1.7	0.9
1	16.1	9.6	4.3	2.2	1.2		6.8	3.9	2.5	1	0.5
2 and more	14.9	8.9	4.1	2.2	1.3		7.2	4.4	3	1.3	0.7
Chi2 p-value	*.000*	*.000*	*.000*	*.000*	*.000*		*.000*	*.000*	*.000*	*.000*	*.000*
Number of elderly members (%)											
0	15.6	9.4	4.4	2.3	1.3		6.8	4.1	2.7	1.1	0.6
1	20.2	12.7	6.4	3.8	2.4		10.3	6.3	4.4	2.1	1.1
2 and more	22.8	15	8.7	5.2	3.6		12.7	9	6.6	3.8	2.3
Chi2 p-value	*.000*	*.000*	*.000*	*.000*	*.000*		*.000*	*.000*	*.000*	*.000*	*.000*
Location of residence (%)											
Urban	16.4	9.8	4.5	2.4	1.4		6.7	3.9	2.5	1.1	0.6
Rural	17.3	11.1	6	3.4	2.2		9.7	6.6	4.7	2.3	1.2
Chi2 p-value	*.001*	*.000*	*.000*	*.000*	*.000*		*.000*	*.000*	*.000*	*.000*	*.000*

Table A2.3: Estimation Results of the Probit Model and the Sartori Selection Model for Catastrophic Health Expenditure (out-of-pocket health expenditure > 2.5% total expenditure)

	Probit Model		**Sartori Selection Model**			
			Selection Equation		*Outcome Equation*	
	Probability of catastrophic health expenditure		Probability of seeking health care		Probability of catastrophic expenditure	
Variable	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.
Poor	-0.066***	0.003	-0.204***	0.012	-0.153***	0.015
No health insurance	0.027***	0.003	-0.016**	0.012	0.058***	0.014
Has disabled or ill person	0.096***	0.005	0.136***	0.017	0.176***	0.018
Secondary education	-0.008**	0.003	-0.024***	0.012	-0.018**	0.015
Higher education	0.007	0.004	-0.001	0.016	0.014	0.019
Employed	0.004	0.003	0.015**	0.012	0.010	0.015
Self-employed	0.019***	0.004	0.029***	0.013	0.040***	0.016
Male	0.0008	0.004	0.010	0.016	0.004	0.019
Household size	-0.004***	0.001	0.019***	0.001	-0.010***	0.003
Urban	-0.007**	0.003	0.021***	0.011	-0.013**	0.013
Preschool (under age 5)	0.029***	0.002	0.034***	0.007	0.064***	0.009
Children (age 6 to 14)	-0.002	0.001	-0.012***	0.005	-0.004	0.007
Elderly (age 65+)	0.034***	0.002	0.059***	0.008	0.072***	0.010
2003	-0.023***	0.004	-0.063***	0.014	-0.026***	0.014
2004	-0.013***	0.005	-0.007	0.017	-0.002	0.019
2005	-0.008	0.005	0.050***	0.018	-0.009	0.019
2006	-0.007	0.005	0.072***	0.018	-0.011	0.019
2007	-0.011**	0.005	0.055***	0.018	-0.009	0.019
2008	-0.026***	0.005	0.091***	0.019	-0.060***	0.019
Log likelihood	-34592.927		-76119.436			
LR chi2 / Wald chi2	1171.19 (20)		2652.84 (19)			
Prob>chi2	0.0000		0.0000			
N	78067		78067			

Notes: 1) ***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the highest level of education status of the household head is 'primary or less' and the omitted category for the employment status of the household head is 'not employed'. 2002 is the reference year for survey years.

3) One marginal effect for household size and its squared term is calculated manually.

4) The marginal effects are calculated at the mean values of the explanatory variables.

Table A2.4: Estimation Results for the Probit Model and the Sartori Selection Model for Catastrophic Health Expenditure (out-of-pocket health expenditure > 5% total expenditure)

	Probit Model		**Sartori Selection Model**			
			Selection Equation		*Outcome Equation*	
	Probability of catastrophic health expenditure		Probability of seeking health care		Probability of catastrophic expenditure	
Variable	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.
Poor	-0.046***	0.002	-0.205***	0.012	-0.109***	0.017
No health insurance	0.025***	0.003	-0.011**	0.012	0.049***	0.016
Has disabled or ill person	0.068***	0.005	0.131***	0.017	0.123***	0.021
Secondary education	-0.007***	0.002	-0.026***	0.012	-0.016**	0.017
Higher education	-0.003	0.003	-0.004	0.016	-0.006	0.022
Employed	0.001	0.003	0.015**	0.012	0.003	0.017
Self-employed	0.015***	0.003	0.029***	0.013	0.031***	0.018
Male	0.003	0.003	0.010	0.016	0.011	0.022
Household size	-0.004***	0.0007	0.019***	0.001	-0.013***	0.003
Urban	-0.008***	0.002	0.026***	0.011	-0.018***	0.015
Preschool (under age 5)	0.020***	0.001	0.034***	0.007	0.045***	0.010
Children (age 6 to 14)	-0.001	0.001	-0.012***	0.005	-0.002	0.008
Elderly (age 65+)	0.023***	0.001	0.059***	0.008	0.048***	0.011
2003	-0.010***	0.003	-0.062***	0.015	-0.023***	0.020
2004	-0.002	0.004	-0.006	0.018	-0.005	0.025
2005	-0.003	0.004	0.050***	0.018	-0.006	0.025
2006	-0.002	0.004	0.072***	0.018	-0.005	0.025
2007	-0.004	0.004	0.055***	0.018	-0.009	0.025
2008	-0.025***	0.003	0.091***	0.019	-0.059***	0.026
Log likelihood	-25203.236		-71338.604			
LR chi2 / Wald chi2	998.62 (20)		2751.18 (20)			
Prob>chi2	0.0000		0.0000			
N	78067		78067			

Notes: 1) ***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the highest level of education status of the household head is 'primary or less' and the omitted category for the employment status of the household head is 'not employed'. 2002 is the reference year for survey years.

3) One marginal effect for household size and its squared term is calculated manually.

4) The marginal effects are calculated at the mean values of the explanatory variables.

Table A2.5: Estimation Results for the Probit Model and the Sartori Selection Model for Catastrophic Health Expenditure (out-of-pocket health expenditure > 10% total expenditure)

	Probit Model		**Sartori Selection Model**			
			Selection Equation		*Outcome Equation*	
	Probability of catastrophic health expenditure		Probability of seeking health care		Probability of catastrophic expenditure	
Variable	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.
Poor	-0.023***	0.001	-0.205***	0.012	-0.055***	0.022
No health insurance	0.017***	0.002	-0.012**	0.012	0.032***	0.019
Has disabled or ill person	0.045***	0.003	0.132***	0.017	0.074***	0.025
Secondary education	-0.005**	0.002	-0.026***	0.012	-0.011**	0.022
Higher education	-0.009***	0.002	-0.004	0.016	-0.021***	0.029
Employed	-0.003	0.002	0.014**	0.012	-0.006	0.022
Self-employed	0.007***	0.002	0.028***	0.013	0.016***	0.022
Male	0.002	0.002	0.011	0.016	0.007	0.028
Household size	-0.004***	0.0005	0.019***	0.001	-0.013***	0.003
Urban	-0.010***	0.001	0.027***	0.011	-0.021***	0.018
Preschool (under age 5)	0.011***	0.001	0.034***	0.007	0.025***	0.013
Children (age 6 to 14)	0.0002	0.001	-0.012***	0.005	0.0008	0.010
Elderly (age 65+)	0.012***	0.001	0.059***	0.008	0.027***	0.013
2003	-0.004**	0.002	-0.062***	0.015	-0.011**	0.025
2004	-0.0006	0.003	-0.007	0.018	-0.0006	0.031
2005	-0.001	0.002	0.050***	0.018	-0.002	0.031
2006	-0.001	0.002	0.073***	0.018	-0.003	0.031
2007	-0.002	0.002	0.055***	0.018	-0.006	0.032
2008	-0.014***	0.002	0.091***	0.019	-0.034***	0.033
Log likelihood	-14874.493			-64446.001		
LR chi2 / Wald chi2	848.15 (20)			2759.00 (20)		
Prob>chi2	0.0000			0.0000		
N	78067			78067		

Notes: 1) ***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the highest level of education status of the household head is 'primary or less' and the omitted category for the employment status of the household head is 'not employed'. 2002 is the reference year for survey years.

3) One marginal effect for household size and its squared term is calculated manually.

4) The marginal effects are calculated at the mean values of the explanatory variables.

Table A2.6: Estimation Results for the Probit Model and the Sartori Selection Model for Catastrophic Health Expenditure (out-of-pocket health expenditure > 15% total expenditure)

	Probit Model		**Sartori Selection Model**			
			Selection Equation		*Outcome Equation*	
	Probability of catastrophic health expenditure		Probability of seeking health care		Probability of catastrophic expenditure	
Variable	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.
Poor	-0.014***	0.001	-0.206***	0.012	-0.034***	0.028
No health insurance	0.012***	0.001	-0.012**	0.012	0.022***	0.024
Has disabled or ill person	0.027***	0.002	0.131***	0.017	0.042***	0.029
Secondary education	-0.004***	0.001	-0.026***	0.012	-0.010***	0.028
Higher education	-0.006***	0.001	-0.004	0.016	-0.014***	0.037
Employed	-0.001	0.001	0.014**	0.012	-0.003	0.027
Self-employed	0.002*	0.001	0.028***	0.013	0.006**	0.027
Male	0.0001	0.001	0.011	0.016	0.001	0.033
Household size	-0.003***	0.0004	0.019***	0.001	-0.010***	0.003
Urban	-0.007***	0.001	0.027***	0.011	-0.014***	0.022
Preschool (under age 5)	0.005***	0.0009	0.034***	0.007	0.011***	0.016
Children (age 6 to 14)	0.0004	0.0007	-0.012***	0.005	0.0009	0.013
Elderly (age 65+)	0.007***	0.0009	0.059***	0.008	0.015***	0.016
2003	-0.003**	0.001	-0.062***	0.015	-0.009**	0.031
2004	-0.001	0.002	-0.007	0.018	-0.001	0.038
2005	-0.0002	0.002	0.051***	0.018	-0.0006	0.038
2006	-0.002	0.002	0.073***	0.018	-0.006	0.039
2007	-0.001	0.002	0.055***	0.018	-0.003	0.039
2008	-0.009***	0.001	0.091***	0.019	-0.023***	0.042
Log likelihood	-9389.7149		-60313.299			
LR chi2 / Wald chi2	635.26 (20)		2757.03 (20)			
Prob>chi2	0.0000		0.0000			
N	78067		78067			

Notes: 1) ***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the highest level of education status of the household head is 'primary or less' and the omitted category for the employment status of the household head is 'not employed'. 2002 is the reference year for survey years.

3) One marginal effect for household size and its squared term is calculated manually.

4) The marginal effects are calculated at the mean values of the explanatory variables.

Table A2.7: Estimation Results for the Probit Model and the Sartori Selection Model for Catastrophic Health Expenditure (out-of-pocket health expenditure > 20% total expenditure)

	Probit Model		**Sartori Selection Model**			
			Selection Equation		*Outcome Equation*	
	Probability of catastrophic health expenditure		Probability of seeking health care		Probability of catastrophic expenditure	
Variable	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.
Poor	-0.008***	0.0008	-0.206***	0.012	-0.020***	0.034
No health insurance	0.008***	0.001	-0.012**	0.012	0.015***	0.029
Has disabled or ill person	0.017***	0.002	0.131***	0.017	0.026***	0.035
Secondary education	-0.002**	0.001	-0.026***	0.012	-0.005**	0.034
Higher education	-0.002*	0.001	-0.004	0.016	-0.005*	0.045
Employed	-0.001	0.001	0.014**	0.012	-0.002	0.033
Self-employed	0.001	0.001	0.028***	0.013	0.003	0.033
Male	0.00006	0.001	0.011	0.016	0.001	0.040
Household size	-0.002***	0.0002	0.019***	0.001	-0.006***	0.004
Urban	-0.005***	0.001	0.026***	0.011	-0.010***	0.027
Preschool (under age 5)	0.002***	0.0007	0.034***	0.007	0.006***	0.020
Children (age 6 to 14)	0.0007	0.0005	-0.012***	0.005	0.001	0.016
Elderly (age 65+)	0.005***	0.0006	0.059***	0.008	0.011***	0.019
2003	-0.001	0.001	-0.062***	0.015	-0.005*	0.038
2004	0.0007	0.001	-0.007	0.018	0.001	0.046
2005	-0.0009	0.001	0.051***	0.018	-0.002	0.047
2006	-0.001	0.001	0.073***	0.018	-0.003	0.048
2007	-0.0008	0.001	0.055***	0.019	-0.001	0.047
2008	-0.005***	0.001	0.091***	0.019	-0.013***	0.051
Log likelihood	-6187.8886		-57759.204			
LR chi2 / Wald chi2	497.78 (20)		2756.73 (20)			
Prob>chi2	0.0000		0.0000			
N	78067		78067			

Notes: 1) ***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the highest level of education status of the household head is 'primary or less' and the omitted category for the employment status of the household head is 'not employed'. 2002 is the reference year for survey years.

3) One marginal effect for household size and its squared term is calculated manually.

4) The marginal effects are calculated at the mean values of the explanatory variables.

Table A2.8: Estimation Results for the Probit Model and the Sartori Selection Model for Catastrophic Health Expenditure (out-of-pocket health expenditure > 10% non-food expenditure)

	Probit Model		**Sartori Selection Model**			
			Selection Equation		*Outcome Equation*	
	Probability of catastrophic health expenditure		Probability of seeking health care		Probability of catastrophic expenditure	
Variable	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.
Poor	-0.027***	0.002	-0.205***	0.012	-0.063***	0.018
No health insurance	0.023***	0.002	-0.011**	0.012	0.046***	0.017
Has disabled or ill person	0.060***	0.004	0.131***	0.017	0.104***	0.022
Secondary education	-0.008***	0.002	-0.026***	0.012	-0.019***	0.019
Higher education	-0.015***	0.003	-0.004	0.016	-0.036***	0.026
Employed	-0.003	0.002	0.015**	0.012	-0.009*	0.019
Self-employed	0.011***	0.002	0.028***	0.013	0.021***	0.019
Male	0.004	0.003	0.011	0.016	0.012*	0.024
Household size	-0.004***	0.0007	0.018***	0.001	-0.014***	0.002
Urban	-0.019***	0.002	0.027***	0.011	-0.040***	0.016
Preschool (under age 5)	0.017***	0.001	0.034***	0.007	0.040***	0.011
Children (age 6 to 14)	0.0009	0.001	-0.012***	0.005	0.003	0.008
Elderly (age 65+)	0.020***	0.001	0.059***	0.008	0.044***	0.012
2003	-0.009***	0.002	-0.062***	0.015	-0.021***	0.022
2004	-0.005	0.003	-0.006	0.018	-0.011	0.027
2005	-0.007**	0.003	0.050***	0.018	-0.017**	0.027
2006	-0.009***	0.003	0.072***	0.018	-0.019**	0.027
2007	-0.011***	0.003	0.055***	0.018	-0.025***	0.028
2008	-0.026***	0.003	0.090***	0.019	-0.066***	0.029
Log likelihood	-20399.037		-68198.819			
LR chi2 / Wald chi2	1144.29 (20)		2751.22 (20)			
Prob>chi2	0.0000		0.0000			
N	78067		78067			

Notes: 1) ***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the highest level of education status of the household head is 'primary or less' and the omitted category for the employment status of the household head is 'not employed'. 2002 is the reference year for survey years.

3) One marginal effect for household size and its squared term is calculated manually.

4) The marginal effects are calculated at the mean values of the explanatory variables.

Table A2.9: Estimation Results for the Probit Model and the Sartori Selection Model for Catastrophic Health Expenditure (out-of-pocket health expenditure > 15% non-food expenditure)

	Probit Model		**Sartori Selection Model**			
			Selection Equation		*Outcome Equation*	
	Probability of catastrophic health expenditure		Probability of seeking health care		Probability of catastrophic expenditure	
Variable	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.
Poor	-0.016***	0.001	-0.205***	0.012	-0.039***	0.021
No health insurance	0.016***	0.002	-0.011**	0.012	0.031***	0.019
Has disabled or ill person	0.043***	0.003	0.132***	0.017	0.071***	0.025
Secondary education	-0.009***	0.001	-0.026***	0.012	-0.022***	0.023
Higher education	-0.015***	0.002	-0.004	0.016	-0.037***	0.032
Employed	-0.004**	0.002	0.014**	0.012	-0.010**	0.022
Self-employed	0.007***	0.002	0.028***	0.013	0.014***	0.022
Male	0.002	0.002	0.011	0.016	0.007	0.028
Household size	-0.003***	0.0005	0.018***	0.001	-0.012***	0.003
Urban	-0.017***	0.001	0.027***	0.011	-0.035***	0.018
Preschool (under age 5)	0.011***	0.001	0.034***	0.007	0.026***	0.012
Children (age 6 to 14)	0.001	0.0009	-0.012***	0.005	0.004**	0.010
Elderly (age 65+)	0.012***	0.001	0.059***	0.008	0.028***	0.013
2003	-0.003	0.002	-0.062***	0.015	-0.007	0.026
2004	-0.0007	0.002	-0.006	0.018	-0.0005	0.032
2005	-0.0004	0.002	0.050***	0.018	-0.0006	0.032
2006	-0.003	0.002	0.073***	0.018	-0.008	0.032
2007	-0.005***	0.002	0.056***	0.018	-0.012**	0.033
2008	-0.015***	0.002	0.091***	0.019	-0.038***	0.035
Log likelihood	-14309.013		-63978.577			
LR chi2 / Wald chi2	1023.91 (20)		2756.41 (20)			
Prob>chi2	0.0000		0.0000			
N	78067		78067			

Notes: 1) ***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the highest level of education status of the household head is 'primary or less' and the omitted category for the employment status of the household head is 'not employed'. 2002 is the reference year for survey years.

3) One marginal effect for household size and its squared term is calculated manually.

4) The marginal effects are calculated at the mean values of the explanatory variables.

Table A2.10: Estimation Results for the Probit Model and the Sartori Selection Model for Catastrophic Health Expenditure (out-of-pocket health expenditure > 20% non-food expenditure)

	Probit Model		**Sartori Selection Model**			
			Selection Equation		*Outcome Equation*	
	Probability of catastrophic health expenditure		Probability of seeking health care		Probability of catastrophic expenditure	
Variable	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.
Poor	-0.010***	0.001	-0.205***	0.012	-0.025***	0.024
No health insurance	0.012***	0.001	-0.012**	0.012	0.023***	0.022
Has disabled or ill person	0.030***	0.003	0.132***	0.017	0.048***	0.028
Secondary education	-0.008***	0.001	-0.026***	0.012	-0.020***	0.028
Higher education	-0.011***	0.001	-0.004	0.016	-0.027***	0.038
No self-employed	-0.003**	0.001	0.015**	0.012	-0.008**	0.026
Self-employed	0.003**	0.001	0.028***	0.013	0.007**	0.025
Male	0.0009	0.002	0.011	0.016	0.003	0.032
Household size	-0.003***	0.0004	0.019***	0.001	-0.011***	0.002
Urban	-0.014***	0.001	0.027***	0.011	-0.028***	0.021
Preschool (under age 5)	0.006***	0.0009	0.034***	0.007	0.015***	0.015
Children (age 6 to 14)	0.001*	0.0007	-0.012**	0.005	0.003**	0.012
Elderly (age 65+)	0.009***	0.0009	0.059***	0.008	0.020***	0.015
2003	-0.003**	0.001	-0.062***	0.015	-0.008**	0.029
2004	-0.0007	0.002	-0.007	0.018	-0.0005	0.036
2005	-0.001	0.002	0.050***	0.018	-0.002	0.036
2006	-0.005***	0.002	0.072***	0.018	-0.012**	0.038
2007	-0.003	0.002	0.055***	0.018	-0.007	0.037
2008	-0.012***	0.001	0.091***	0.019	-0.030***	0.040
Log likelihood	-10525.374		-61156.886			
LR chi2 / Wald chi2	856.12 (20)		2757.81 (20)			
Prob>chi2	0.0000		0.0000			
N	78067		78067			

Notes: 1) ***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the highest level of education status of the household head is 'primary or less' and the omitted category for the employment status of the household head is 'not employed'. 2002 is the reference year for survey years.

3) One marginal effect for household size and its squared term is calculated manually.

4) The marginal effects are calculated at the mean values of the explanatory variables.

Table A2.11: Estimation Results for the Probit Model and the Sartori Selection Model for Catastrophic Health Expenditure (out-of-pocket health expenditure > 30% non-food expenditure)

	Probit Model		**Sartori Selection Model**			
			Selection Equation		*Outcome Equation*	
	Probability of catastrophic health expenditure		Probability of seeking health care		Probability of catastrophic expenditure	
Variable	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.
Poor	-0.004***	0.0008	-0.206***	0.012	-0.011***	0.033
No health insurance	0.006***	0.001	-0.012**	0.012	0.011***	0.029
Has disabled or ill person	0.015***	0.002	0.132***	0.017	0.023***	0.036
Secondary education	-0.003***	0.0009	-0.026***	0.012	-0.007***	0.038
Higher education	-0.003***	0.001	-0.004	0.016	-0.009***	0.052
Employed	-0.002*	0.001	0.014**	0.012	-0.004**	0.035
Self-employed	0.002*	0.001	0.028***	0.013	0.004*	0.034
Male	0.00007	0.001	0.011	0.016	0.0005	0.041
Household size	-0.001***	0.0002	0.019***	0.001	-0.005***	0.005
Urban	-0.007***	0.001	0.027***	0.011	-0.014***	0.028
Preschool (under age 5)	0.002***	0.0006	0.034***	0.007	0.006***	0.020
Children (age 6 to 14)	0.001**	0.0005	-0.012***	0.005	0.002**	0.016
Elderly (age 65+)	0.005***	0.0006	0.059***	0.008	0.011***	0.019
2003	-0.002**	0.001	-0.062***	0.015	-0.005**	0.039
2004	-0.0001	0.001	-0.006	0.018	0.00002	0.047
2005	-0.001	0.001	0.051***	0.018	-0.004	0.049
2006	-0.003**	0.001	0.073***	0.018	-0.007**	0.050
2007	-0.002**	0.001	0.055***	0.019	-0.005*	0.050
2008	-0.005***	0.001	0.091***	0.019	-0.014***	0.054
Log likelihood	-5560.5012		-57261.106			
LR chi2 / Wald chi2	567.63 (20)		2756.24 (20)			
Prob>chi2	0.0000		0.0000			
N	78067		78067			

Notes: 1) ***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the highest level of education status of the household head is 'primary or less' and the omitted category for the employment status of the household head is 'not employed'. 2002 is the reference year for survey years.

3) One marginal effect for household size and its squared term is calculated manually.

4) The marginal effects are calculated at the mean values of the explanatory variables.

Table A2.12: Estimation Results for the Probit Model and the Sartori Selection Model for Catastrophic Health Expenditure (out-of-pocket health expenditure > 40% non-food expenditure)

	Probit Model		**Sartori Selection Model**			
			Selection Equation		*Outcome Equation*	
	Probability of catastrophic health expenditure		Probability of seeking health care		Probability of catastrophic expenditure	
Variable	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.
Poor	-0.003***	0.0005	-0.206***	0.012	-0.008***	0.043
No health insurance	0.004***	0.0008	-0.012**	0.012	0.008***	0.037
Has disabled or ill person	0.008***	0.001	0.131***	0.017	0.012***	0.045
Secondary education	-0.001**	0.0006	-0.026***	0.012	-0.003**	0.050
Higher education	-0.001	0.0008	-0.004	0.016	-0.003	0.065
Employed	-0.0006	0.0007	0.014**	0.012	-0.001	0.045
Self-employed	0.001	0.0007	0.028***	0.013	0.002	0.043
Male	0.0003	0.0008	0.011	0.016	0.001	0.054
Household size	-0.001***	0.0001	0.019***	0.001	-0.003***	0.005
Urban	-0.003***	0.0007	0.026***	0.011	-0.007***	0.035
Preschool (under age 5)	0.001***	0.0004	0.034***	0.007	0.003***	0.026
Children (age 6 to 14)	0.0006*	0.0003	-0.012***	0.005	0.001*	0.021
Elderly (age 65+)	0.003***	0.0004	0.059***	0.008	0.007***	0.024
2003	-0.0005	0.0008	-0.062***	0.015	-0.001	0.051
2004	-0.0003	0.001	-0.006	0.018	0.0009	0.061
2005	-0.001	0.0009	0.051***	0.018	-0.002	0.065
2006	-0.001	0.0009	0.073***	0.018	-0.003	0.066
2007	-0.0006	0.0009	0.056***	0.019	-0.001	0.064
2008	-0.003***	0.0007	0.091***	0.019	-0.008***	0.073
Log likelihood	-3245.307		-55361.265			
LR chi2 / Wald chi2	390.86 (20)		2755.47 (20)			
Prob>chi2	0.0000		0.0000			
N	78067		78067			

Notes: 1) ***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the highest level of education status of the household head is 'primary or less' and the omitted category for the employment status of the household head is 'not employed'. 2002 is the reference year for survey years.

3) One marginal effect for household size and its squared term is calculated manually.

4) The marginal effects are calculated at the mean values of the explanatory variables.

Table A2.13: Estimation Results for the Heckman Probit Model for Catastrophic Health Expenditure (out-of-pocket health expenditure > 10% total expenditure)

	Selection equation Probability of seeking Health care		*Outcome equation* Probability of catastrophic expenditure	
Variable	Marg. Eff.	St. Err.	Marg.Eff.	St. Err.
Poor	-0.170***	0.004	-0.016***	0.004
No health insurance	-0.010**	0.004	0.039***	0.004
Has disabled or ill person	0.111***	0.006	0.060***	0.006
Secondary education	-0.022***	0.005	-0.006	0.004
Higher education	-0.003	0.006	-0.019***	0.005
Employed	0.012**	0.005	-0.008**	0.004
Self-employed	0.024***	0.005	0.011**	0.004
Male	0.009	0.006	0.005	0.005
Household size	0.019***	0.001	-0.005***	0.001
Urban	0.022***	0.004	-0.028***	0.004
Preschool (under age 5)	0.029***	0.003	0.019***	0.002
Children (age 6 to 14)	-0.010***	0.002	0.002	0.002
Elderly (age 65+)	0.050***	0.003	0.016***	0.002
2003	-0.052***	0.006	-0.00009	0.005
2004	-0.005	0.007	0.0005	0.006
2005	0.043***	0.007	-0.010*	0.005
2006	0.062***	0.007	-0.015***	0.005
2007	0.047***	0.007	-0.015***	0.005
2008	0.078***	0.007	-0.041***	0.004
Log likelihood	-64445.8			
Rho	0.75			
Wald chi2 (20)	769.41			
Prob>chi2	0.0000			
N	78067			

Notes: 1) ***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the highest level of education status of the household head is 'primary or less' and the omitted category for the employment status of the household head is 'not employed'. 2002 is the reference year for survey years.

3) One marginal effect for household size and its squared term is calculated manually.

4) The marginal effects are calculated at the mean values of the explanatory variables.

Table A2.14: Estimation Results for the Heckman Probit Model for Catastrophic Health Expenditure (out-of-pocket health expenditure > 20% non-food expenditure)

	Selection equation Probability of seeking Health care		*Outcome equation* Probability of catastrophic expenditure	
Variable	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.
Poor	-0.169***	0.004	-0.010***	0.003
No health insurance	-0.010**	0.004	0.028***	0.003
Has disabled or ill person	0.111***	0.006	0.044***	0.005
Secondary education	-0.021***	0.005	-0.015***	0.003
Higher education	-0.003	0.006	-0.023***	0.003
Employed	0.012**	0.005	-0.009***	0.003
Self-employed	0.023***	0.005	0.005	0.003
Male	0.009	0.006	0.002	0.004
Household size	0.019***	0.001	-0.006***	0.0001
Urban	0.022***	0.004	-0.035***	0.003
Preschool (under age 5)	0.029***	0.003	0.012***	0.002
Children (age 6 to 14)	-0.010***	0.002	0.005***	0.001
Elderly (age 65+)	0.050***	0.003	0.013***	0.002
2003	-0.052***	0.006	-0.003	0.004
2004	-0.005	0.007	0.0005	0.005
2005	0.043***	0.007	-0.009*	0.004
2006	0.062***	0.007	-0.018***	0.004
2007	0.047***	0.007	-0.013***	0.004
2008	0.078***	0.007	-0.032***	0.003
Log likelihood	-61237.18			
Rho	0.74			
Wald chi2 (19)	697.94			
Prob>chi2	0.0000			
N	78067			

Notes: 1) ***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the highest level of education status of the household head is 'primary or less' and the omitted category for the employment status of the household head is 'not employed'. 2002 is the reference year for survey years.

3) One marginal effect for household size and its squared term is calculated manually.

4) The marginal effects are calculated at the mean values of the explanatory variables.

Table A2.15: Estimation Results for the Heckman Probit Model for Catastrophic Health Expenditure (out-of-pocket health expenditure > 10% total expenditure) (rho=0.5)

	Selection equation Probability of seeking Health care		*Outcome equation* Probability of catastrophic expenditure	
Variable	Marg. Eff.	St. Err.	Marg.Eff.	St. Err.
Poor	-0.170***	0.004	-0.015***	0.004
No health insurance	-0.010**	0.004	0.040***	0.004
Has disabled or ill person	0.111***	0.006	0.061***	0.006
Secondary education	-0.022***	0.005	-0.006	0.004
Higher education	-0.003	0.006	-0.019***	0.005
Employed	0.012**	0.005	-0.008**	0.004
Self-employed	0.023***	0.005	0.012**	0.004
Male	0.009	0.006	0.005	0.005
Household size	0.019***	0.001	-0.007***	0.0007
Urban	0.022***	0.004	-0.028***	0.004
Preschool (under age 5)	0.029***	0.003	0.019***	0.002
Children (age 6 to 14)	-0.010***	0.002	0.002	0.002
Elderly (age 65+)	0.050***	0.003	0.016***	0.002
2003	-0.052***	0.006	0.0001	0.005
2004	-0.005	0.007	0.0005	0.006
2005	0.043***	0.007	-0.011*	0.005
2006	0.062***	0.007	-0.015***	0.005
2007	0.047***	0.007	-0.015***	0.005
2008	0.078***	0.007	-0.042***	0.004
Log likelihood	-64445.9			
Wald chi2 (20)	760.82			
Prob>chi2	0.0000			
N	78067			

Notes: 1) ***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the highest level of education status of the household head is 'primary or less' and the omitted category for the employment status of the household head is 'not employed'. 2002 is the reference year for survey years.

3) One marginal effect for household size and its squared term is calculated manually.

4) The marginal effects are calculated at the mean values of the explanatory variables.

Table A2.16: Estimation Results for the Heckman Probit Model for Catastrophic Health Expenditure (out-of-pocket health expenditure > 20% non-food expenditure (rho=0.5)

	Selection equation Probability of seeking Health care		*Outcome equation* Probability of catastrophic expenditure	
Variable	Marg.Eff.	St. Err.	Marg.Eff.	St. Err.
Poor	-0.170***	0.004	-0.011***	0.003
No health insurance	-0.010**	0.004	0.029***	0.003
Has disabled or ill person	0.111***	0.006	0.043***	0.005
Secondary education	-0.022***	0.005	-0.015***	0.003
Higher education	-0.003	0.006	-0.023***	0.003
Employed	0.012**	0.005	-0.009***	0.003
Self-employed	0.023***	0.005	0.005	0.003
Male	0.009	0.006	0.002	0.004
Household size	0.019***	0.001	-0.005***	0.0006
Urban	0.022***	0.004	-0.035***	0.003
Preschool (under age 5)	0.029***	0.003	0.012***	0.002
Children (age 6 to 14)	-0.010***	0.002	0.005***	0.001
Elderly (age 65+)	0.050***	0.003	0.013***	0.002
2003	-0.052***	0.006	-0.001	0.004
2004	-0.005	0.007	0.0002	0.005
2005	0.043***	0.007	-0.008*	0.004
2006	0.062***	0.007	-0.018***	0.004
2007	0.047***	0.007	-0.013***	0.004
2008	0.078***	0.007	-0.033***	0.003
Log likelihood	-61156.32			
Wald chi2 (20)	842.54			
Prob>chi2	0.0000			
N	78067			

Notes: 1) ***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the highest level of education status of the household head is 'primary or less' and the omitted category for the employment status of the household head is 'not employed'. 2002 is the reference year for survey years.

3) One marginal effect for household size and its squared term is calculated manually.

4) The marginal effects are calculated at the mean values of the explanatory variables.

CHAPTER 3: GENDER DIFFERENCES IN CIGARETTE CONSUMPTION IN TURKEY

3.1 INTRODUCTION

Smoking is accepted as a principal cause of numerous diseases such as cancer, cardiovascular and respiratory diseases (WHO, 2009). Tobacco also causes more than 5 million deaths each year and, if the current trends continue, it is estimated that the annual number of people killed by tobacco use will increase to 10 million every year by 2030 and 7 million of these deaths will occur in developing countries (Jha and Chaloupka, 1999). In this context, Turkey has an alarmingly high rate of smoking prevalence, nearly one-third (30%) of adults aged 15 and over (47.7% among males and 14% among females) were current smokers in 2008 (Global Adult Tobacco Survey (GATS), 2008). The smoking prevalence rate of males in Turkey is higher than in any Western European Country and is one of the highest rates in Central Asia. Furthermore, smoking attributed diseases are the highest risk factors for deaths among males and the second most common cause of all deaths in Turkey (Yurekli *et al.*, 2010). On the other hand, smoking attributed diseases constitute a considerable burden on Turkey's health care system. The Ministry of Health states that 23% of total patient days and 52% of total hospital deaths were caused by diseases attributable to smoking in 2000. Tobacco use caused 54,699 deaths in Turkey in 2003, which is approximately 13% of total deaths, and 596,684 years of life lost (Ministry of Health, 2009). If the current pattern continues, it is estimated that the figure will increase to more than 127,000 deaths in 2050. Effective anti-tobacco policies have been predicted to lead to a decrease in smoking prevalence rate to 10% by 2050 and to save nearly 47,000 lives every year (Yurekli *et al.*, 2010).

In this respect, policy-makers in Turkey focus on identifying the most effective public policy to reduce smoking prevalence and the associated healthcare costs. Consequently, the tobacco control regulations have been restructured over the past few decades in Turkey. The anti-tobacco law, which was the official milestone for the tobacco control policies of Turkey, was enacted in 1996. In 2008, Turkey adopted comprehensive tobacco legislation including a set of measures such as a ban on smoking in all indoor public places and a ban on tobacco advertising and promotion. In addition, Turkey started the accession negotiations for full membership of the EU in 2005. In order to

fulfil the EU requirements and become a member state, Turkey has to revise many of its laws including anti-smoking laws (Bilir *et al.*, 2009; Yurekli *et al.*, 2010). Thus, one of the aims of the 1996 and 2008 legislation is arguably to meet the criteria for EU membership since the EU has attached increasing importance to anti-smoking policy in recent years. Furthermore, public policy aims to decrease the smoking prevalence rate and, in this way, to close the gap in tobacco use between the EU and Turkey. In this context, the debate exposes the crucial importance of information on the potential factors, which affect the demand for cigarettes in Turkey, to design effective tobacco control policies. However, studies on the determinants of cigarette smoking in Turkey, which is a tobacco producing country with a high prevalence of cigarette smoking, are surprisingly limited.

The aim of this chapter is to examine the socio-economic and demographic factors associated with both smoking participation and the level of cigarette consumption in Turkey, the understanding of which are crucial to the formulation and implementation of public policies aimed at intervention. Furthermore, the existence of potential gender differences in cigarette consumption patterns has been empirically proved to be relevant in the existing literature. This may indicate the need for specific anti-smoking policies for males and females (Yen, 2005; Aristei and Pieroni, 2009). In the most developed countries, the share of smokers among females has recently approached the respective share among males, mainly because of a sharply decreasing rate among males (Bauer *et al*., 2007). In Turkey, in addition to a decrease in the male smoking prevalence rate, the smoking prevalence rate of females has increased dramatically in recent years, by nearly 40% between 1997 and 2009 from 10.13% to 14% (Yurekli *et al.*, 2010). From a policy perspective, it is important to focus on gender differences in cigarette consumption in Turkey to design anti-smoking policies in a more efficient way by addressing specific target groups. In order to decide whether male and female sub-samples should be pooled together or should be treated separately in the empirical analysis, the hypothesis of equal consumption parameters for males and females is tested using a likelihood ratio (LR) test (Yen, 2005). The results of the LR test suggest that the hypothesis of equal parameters for males and females is rejected for all models, which means that males and females should be treated separately in the empirical analysis. However, no known study has addressed such differences in Turkey, which may be due to the limitation of

the data sets used. Thus, it can be argued that this study is the first attempt to investigate cigarette consumption in Turkey from a gender perspective.

In order to investigate the factors associated with the cigarette demand for males and females separately, the Global Adult Tobacco Survey (GATS) for the year 2008, which is the most comprehensive individual level survey focusing on tobacco consumption for adults in Turkey, is used in this chapter. As recommended by the WHO European Strategy for Tobacco Control, the GATS, which is a rich source of important variables related to tobacco consumption and is a nationally representative data with 9030 completed interviews, can be used to develop a National Tobacco Control Program (NTCP) (GATS Turkey Report, 2010). However, no known study has used the GATS to investigate the determinants of adult cigarette demand in Turkey since it has become available only recently[30]. Perhaps most importantly, this study extends the existing literature by using individual level rather than household level or aggregate time-series data, which differs from the existing studies carried out for Turkey.

In general, one important characteristic of cigarette consumption data is that the dependent variables include a high number of zero observations. Furthermore, cigarettes are treated as special consumer goods as they differ from other goods in which zero observations are considered as resulting from a corner solution. In cigarette consumption, zero observations may belong to either non-smokers who pay no regard to cigarette consumption in their decision process (i.e., abstention) or potential smokers who become consumers if the price is lower or income is higher (i.e., a corner solution) (Blundell, 1988; Harris and Zhao, 2007). Therefore, the methodological challenge is to distinguish the nature of the zero observations. Since the dependent variable, the number of cigarettes smoked per day, consists of nonnegative integer values, zero observations are accommodated in this study by using the zero-inflated negative binomial (ZINB) model as well as other count data models to explore the robustness of the findings. The results of the LR test and Vuong (1989) test indicate the ZINB model as the preferred specification which describes the data best for both gender groups. The most important feature of the ZINB model is that, which is discussed in detail in Section 3.4.2 below, it adds extra weight to the likelihood of observing a zero by a mixing specification by assuming the zeros can arise from the two different sources (Mullahy,

[30] The GATS has become available from May 2010 (Turkish Statistical Institute, 2010).

1986). Although the results of the LR and Vuong (1989) tests suggest the ZINB model as the preferred specification, the results of the NB model are also discussed in this chapter since it provides a baseline model for modelling count data and can be used as a basis for comparison. The two-part/hurdle count data model, where the NB specification is used in the second part, is also estimated to explore the robustness of the results.

The rest of the chapter is structured as follows; Section 3.2 provides information on various aspects of smoking prevalence in Turkey. Section 3.3 reviews the background literature focusing on the econometric methodologies of the studies. Section 3.4 describes the data, the variables used in the analysis and the descriptive statistics as well as the estimation methods employed in this chapter. Section 3.5 discusses the results and summarises the key findings and Section 3.6 discusses the main findings and policy implications, the limitations of the analysis and potential directions for future research.

3.2 SMOKING PREVALENCE AND TOBACCO CONTROL POLICIES IN TURKEY

This section aims to provide an overview of various aspects of smoking prevalence such as across occupational groups, age and gender in Turkey as well as smoking control policies and their implications for Turkey. It should be noted that tobacco has been more than a consumption product for Turkey. Tobacco has been an important revenue item for more than a century (Bilir *et al*., 2009). Turkey is the fifth largest tobacco producer in the world and is one of the significant cigarette exporters due to its location and low cost production (Onder, 2002). Furthermore, Turkey ranks first in oriental tobacco production in the world. According to the US Department of Agriculture (USDA), Turkey was the leader in terms of tobacco exporting among oriental tobacco exporting countries between 2003 and 2006 (USDA, 2007). Therefore, it can be argued that tobacco production is another dimension of the economics of tobacco in Turkey, which has made significant contributions to the Turkish economy in, for example, employment, exports and tax income. However, tobacco production and its market share in Turkey will not be further discussed in this section since the main aim of this chapter is to investigate the determinants of cigarette demand in Turkey and hence this chapter focuses on the demand dimension of cigarette consumption[31].

3.2.1 Smoking Prevalence among Adults

According to the USDA, total cigarette consumption decreased by 4% between 1990 and 1999 in the world, whereas it increased by 52% in Turkey which is the third largest increase after Pakistan and Bulgaria (Onder, 2002). Turkey is still among the top 10 tobacco consuming countries in the world and consumes 2% of the total world tobacco production and 14% of the WHO European Region (Bilir *et al.*, 2009; GATS Turkey Report, 2010). Therefore, it is possible to argue that Turkey continues to be one of the largest cigarette consuming countries.

One of the rare studies covering the whole country, based on a representative sample of adults aged 15 and over in 1988, indicated that the smoking prevalence rate was 44% for all adults, 62.8% for males and 24% for females in Turkey (Bilir, 1997). Table 3.1 below presents the prevalence rate of tobacco use among adults aged 18 and over from

[31] Although, there is no noticeable relationship between recent policy changes, which are discussed in Section 3.2.5 below, and tobacco production, it should be stated that tobacco production and its market share may be linked to the anti-smoking policies which were adopted in Turkey.

two different surveys of a representative sample of adults. The Health Services Utilisation Survey implemented in 1993 revealed that the smoking prevalence rates were 57.8% and 13.5% for males and females, respectively. The National Household Survey (2003) indicated that 33.8% of adults aged 18 years and over were daily smokers and that the male smoking rate (52.9%) was significantly higher than the female smoking rate (19.5%) (Bilir *et al*., 2009).

Table 3.1: Smoking Prevalence among Adults (aged 18+ years)

Year	Male	Female	Total
1993*	57.8%	13.5%	33.6%
2003**	52.9%	19.5%	33.8%

Sources: *1993 Health Services Utilisation Survey (Toros and Oztek, 1993), **2003 National Burden of Disease and Cost-Effectiveness Study (Unuvar *et al.*, 2006), Bilir *et al*., 2009.

The GATS, which is the most recent and comprehensive tobacco survey for Turkey, indicates that nearly one-third (29.9%) of adults aged 15 and over were current smokers in 2008. The smoking rate for males (47.7%) is remarkably higher than the female smoking rate (14%), which corresponds to approximately 12 million males and 4 million females.

3.2.2 Smoking Prevalence among Selected Occupational Groups

Although smoking is a primary concern for the whole population, some segments of the population have special roles and responsibilities related to smoking such as physicians, teachers, politicians and sportsmen and women. Based on this fact, most of the tobacco surveys were conducted for various interest groups in Turkey. In 1998 and 1999, two country based surveys including 12,500 individuals from various occupational groups revealed that smoking prevalence was relatively high in occupational groups that are considered as 'role models' in society. Nearly 48% of teachers, 43% of physicians and 35% of sportsmen and women reported current smoking (Bilir *et al.*, 2009).

3.2.3 Youth Smoking Prevalence

Adolescents may arguably be seen a major target group for several public policies since adolescence is an important and common period for starting to smoke. The Global Youth Tobacco Survey (GYTS) was administrated to 15,957 students aged 13-15 years in 2003 and the results showed that almost 3 in 10 (26.3%) schoolchildren, 22.3% of girls and 33.1% of boys reported that they had smoked cigarettes at some time in their

lives. 30.7% of individuals who stated that they had smoked at some point reported that they started smoking before the age of 10 years, significantly this was the case for more boys (34.9%) than girls (23.7%). Overall, 6.9% of students stated that they currently smoked cigarettes and the proportion of students who were current smokers was greater for boys (9.1%) than girls (3.5%) (Erguder *et al.*, 2008). This high prevalence of youth tobacco use was confirmed again by the GYTS in 2009, which was administrated to 5,045 students mainly aged 13-15 years. The results of this survey revealed that the prevalence of tobacco use among youths was increasing. Overall, 8.4% of schoolchildren reported that they were current smokers (10.2% of boys and 5.3% of girls) (GYTS, 2009). These figures imply that the prevalence of youth smoking has increased especially among girls.

Furthermore, in recent years, using water pipes to smoke tobacco has been a rising form of tobacco use in Turkey and has been marketed particularly to young people and women. Although the prevalence rate of water pipe use is not known certainly due to the unavailability of relevant data, observations imply its increasing use especially among young adults (Bilir *et al.*, 2009).

3.2.4 Gender Differences in Smoking Prevalence

The anti-smoking policies have generally had different effects on males and females. In general, males have always smoked at much higher rates than females all over the world, but the differences between males and females have become much smaller since the 1960s (Chaloupka, 1990). Although some significant progress has been achieved in reducing smoking participation among males and females, the rate of decrease is generally smaller for females. Female cigarette smoking is decreasing in most developed countries such as Canada, Australia, the UK and the US but in several southern, central and eastern European countries, the smoking prevalence rates of females have not shown any change or there is an increasing trend in female tobacco consumption rate. Therefore, there is growing attention paid by policy makers to curbing the increase in female smoking and it is emphasised that health policies, which are designed to prevent increases in smoking prevalence among females especially in low and middle-income countries, will have a greater impact on global health than any other single intervention (Jha and Chaloupka, 1999).

In fact, tobacco consumption behaviour is generally different between males and females potentially due to different smoking risk judgements and attitudes (Chaloupka, 1990; Yen, 2005). Females face additional health risks from smoking compared to males because of the well-established link between smoking and a variety of fetal diseases. Smoking during pregnancy increases the risk of complications and is well known to have a negative effect on pregnancy outcomes such as lower birth-weight and a higher rate of still births. Furthermore, children of women who smoke during pregnancy have greater infant mortality rates compared to children of women who are non-smokers (Chaloupka, 1990). Additionally, it can be argued that females have some specific reasons which are different from males for smoking such as coping with both responsibilities of caring for children and growing careers as well as weight control and female physical health issues.

Gender differences in smoking prevalence are also the case for Turkey. In nearly every age group, there is a considerable gender difference in smoking in Turkey. Furthermore, the smoking prevalence rate of females has increased sharply in recent years, by nearly 40% between 1997 and 2009 from 10.13% to 14% (Yurekli *et al.*, 2010). Turkey is a country experiencing a process that changes old traditional culture to the popular culture that encompasses the whole world through mass media. This may lead to a significant change in the traditional female role in Turkey. In this respect, the increasing prevalence of female smokers may be explained by an increasing economic independence among Turkish women and clever tobacco marketing campaigns targeting them (Erten and Aslan, 2008). Some of the determinants of this independence include the level of education, location of residence, improvements in the bargaining position of women in the household with an increase in their earned income. Table 3.2 below presents cigarette smoking profiles of women who have been married at some point in their lives in Turkey in 2003[32].

As it can be seen from the table, cigarette consumption among women who live in urban areas is higher than for those who live in rural areas. Similarly, the cigarette consumption of women living in the western part of the country is higher than for those living in the eastern part of the country. In contrast to the general consensus regarding the negative effect of higher education on smoking behaviour, smoking is more

[32] The reason for presenting cigarette consumption profiles of women only for those who have been married at some point in their lives is data availability.

common among women who have highly educated mothers. This pattern may indicate the effect of income through higher education. Interestingly, 15% of pregnant women reported that they currently smoked and 14.6% of whom smoked more than 10 cigarettes per day. These figures reveal interesting aspects of female smoking in Turkey and indicate the need for the examination of the reasons behind this situation in depth.

Table 3.2: Cigarette Consumption Profiles of Women in Turkey (2003)

Characteristics	**Percentage who smoke cigarettes**	**Percentage distribution of women who smoke cigarettes according to number of cigarettes smoked per day**				
		0	1-2	3-5	6-9	10+
Age						
15-19	16.5	1.5	22.5	43.8	13.3	19.3
20-34	30.1	3.2	21.5	26.9	14.0	34.2
35-49	25.7	2.1	17.4	18.0	10.6	51.8
Location						
Urban	32.8	2.1	17.8	22.0	13.1	44.8
Rural	14.9	6.0	30.1	29.9	9.3	24.7
Region						
West	32.3	1.7	17.1	22.0	12.5	46.7
South	26.7	6.7	17.2	22.5	14.8	38.6
Central	25.8	2.1	23.5	24.1	12.1	37.9
North	21.5	3.8	28.6	25.7	9.6	32.4
East	22.1	2.9	21.8	26.3	12.3	36.7
Mother's education						
No education	18.4	2.1	25.4	22.9	8.0	41.7
First level primary	24.5	3.1	22.0	22.9	12.8	39.2
Second level primary	40.1	2.4	9.8	27.1	16.5	43.5
High school	43.9	2.5	16.7	22.6	12.8	45.4
Maternity Status						
Pregnant	15.0	3.7	26.6	41.1	13.9	14.6
Breastfeeding	19.6	4.5	26.1	29.4	17.3	22.7
Not pregnant or breastfeeding	29.7	2.5	18.9	22.1	12.0	44.4
Total	27.6	2.7	19.7	23.3	12.5	41.7

Source: Turkish Statistical Institute (2003).
Notes: 1) Table shows the percentage of women who have been married and who smoke regularly or occasionally where '0' cigarettes smoked per day indicates occasional smoking.

3.2.5 Tobacco Control Policies in Turkey

Tobacco control policies in Turkey were almost non-existent until 1996 and the strong anti-smoking policies have been put into place fairly recently. Prior to the 1980s, the only tobacco control policy was a ban on pro-cigarette marketing on television, radio and billboards. In 1988, there was an anti-smoking campaign, initiated by the Ministry of Health, with required posters in public areas that included information on the adverse effects of smoking on health (Yurekli *et al.*, 2010).

The Anti-Tobacco Law, (Law No. 4207), which was the first legal regulation controlling consumption of tobacco products, was enacted in 1996. This law included mainly a ban on cigarette smoking in some public places, specifically in education, health and cultural service locations, enclosed sports facilities, on public transportation[33] and in waiting rooms. The ban on pro-cigarette marketing was extended to other types of advertising including newspapers and magazines and the use of tobacco brand names on non-tobacco products was prohibited. Selling tobacco products to people under the age of 18 was prohibited. A warning label 'Legal Warning: Harmful to Health' was required on all imported and domestically produced cigarette packages. In addition, public education against smoking via television was mandatory. Although Law 4207 was designed to be comprehensive, there were important problems with its implementation which limited its effectiveness in decreasing smoking prevalence mostly due to the non-existence of an agency which is responsible for the enforcement and collection of penalties (Bilir *et al.*, 2009; Yurekli *et al.*, 2010).

The WHO Framework Convention on Tobacco Control (FCTC) was adopted by the World Health Assembly in 2003 due to the need for an international action to reduce smoking prevalence. The aim of the WHO FCTC is stated as to protect present and future generations from the adverse health, social, economic and environmental impacts of tobacco use and exposure to tobacco smoke (WHO, 2003). In this context, Turkey signed the WHO FCTC in April 2004, which was adopted in November 2004 and put into place in Turkey in February 2005. This international convention requires the government to adopt measures to decrease both the supply and the demand for tobacco products (Yurekli *et al.*, 2010). The National Tobacco Control Plan (NTCP) has been prepared for 2008-2012, which includes awareness-raising of tobacco-related health

[33] This was arguably the major success of the law since 'non-smoking' and 'smoke-free' norms among the Turkish population are internalised by this law (GATS Turkey Report, 2010).

hazards, smoking cessation, prevention of second-hand exposure to tobacco, media and pro-cigarette marketing, accessibility to young people, monitoring and evaluation of tobacco use and pricing and taxation (Bilir *et al.*, 2009). Prohibition of smoking in certain places is accepted as a very effective anti-tobacco regulation since it makes easier for smokers to quit smoking and reduces smoking initiation. Therefore, the new legislation, Law No. 4227, was adopted in January 2008 in two phases. The first phase started in May 2008 with an extended ban on smoking in all public buildings, public transportation, taxis, health care facilities, sport facilities and the inside and outside of all schools. In July 2009, in the second phase, the ban was further extended to the entertainment sector, restaurants, bars and Turkish coffee houses. These recent extensions in the anti-tobacco law support the fact that tobacco consumption is one of the major problems and public policy areas in Turkey.

3.3 LITERATURE REVIEW

Over recent decades, the empirical literature has focused on the economic and non-economic determinants of cigarette demand. This section provides a review of the existing literature which relates to the determinants of cigarette demand on the basis of the econometric methods used in the analysis. Furthermore, this section provides information on studies on cigarette demand for Turkey.

3.3.1 Time Series Studies

One branch of the existing literature has used aggregate time-series data to examine the effects of price and taxes on cigarette demand while controlling for income, tobacco control regulations and socio-economic and demographic factors (see, for example, Keeler *et al.*, 2001 for the US; Yurekli and Zhang, 2000 for the US). Most of these studies suggest a strong inverse association between cigarette prices and cigarette demand. However, most of the previous studies investigate the impact of cigarette prices on average adult cigarette consumption since they use aggregate level data to estimate cigarette demand equations.

Although aggregate level data tend to be straightforward to obtain compared to individual level survey data, as Chaloupka and Warner (2000) point out, there are some limitations of studies using aggregate level time-series data. First, the high correlations among many of the explanatory variables and prices are an issue for most of the cigarette demand studies. In this case, simultaneity bias may occur because cigarette prices, sales and consumption are simultaneously determined and, thus, it becomes hard to interpret causal directions. In addition, including highly correlated variables in the analysis may lead to multicollinearity and unstable estimation results. Excluding potentially important variables, on the other hand, may result in biased estimation results[34]. Second, studies using aggregate time-series data often evaluate the effects of cigarette prices, income and other factors on average or per capita measures of cigarette consumption. Therefore, these studies are generally not able to examine price sensitivity in different segments of the population, such as smokers of different ages, genders or race/ethnicity. A third problem with time-series analysis is that it cannot examine the

[34] Some studies, however, have addressed these limitations (see, for example, Keeler *et al.*, 1996 for the US; Sung *et al.*, 1994 for 11 western states) by using econometric methods, such as recursive models, natural experiments and simultaneous equations models (Chaloupka and Warner, 2000).

impact of prices on smoking propensity (i.e. being a smoker or not) or on intensity (such as daily cigarette consumption) separately.

Another aspect of the existing empirical literature using aggregate time-series data is the evaluation of the nature of consumer behaviour in the context of the rational addiction model proposed by Becker and Murphy (1988). According to the theory of rational addiction, there are inter-temporal linkages in the consumption of some goods and these linkages should be taken into account when estimating its demand (Dorsett, 1999). The partial adjusted (myopic) model of addiction assumes that current cigarette consumption is affected by current price and past consumption (Baltagi and Levin, 1986). However, the rational addiction paradigm argues that current cigarette consumption is associated not only with the current price and lagged consumption, but also with future cigarette consumption (Becker *et al.*, 1994; Baltagi and Griffin, 2001). It should be noted that the rational addiction model is also subject to some econometric problems such as serial correlation in the error terms and endogeneity due to the existence of both current and lagged consumption in the regression model.

3.3.2 Household and Individual Level Studies

Econometric studies that have used household or individual level data to investigate the determinants of cigarette demand have increased in recent years. The main reason behind this increase is that the use of microeconomic data in the empirical analysis has allowed the modelling of cigarette demand to account for zero observations and also to control for heterogeneous individual behaviour by using a range of socio-demographic variables (Aristei and Pieroni, 2008). Furthermore, the use of individual level data helps to avoid some of the problems inherent in using aggregate data[35] and enables the exploration of some issues that typically cannot be addressed with aggregate data (Chaloupka and Warner, 2000). For example, a majority of studies using individual level data consider the effects of various factors on smoking participation and the level of consumption separately (Jones, 1989a; Blaylock and Blisard, 1992a, Atkinson *et al.*, 1984). In the existing literature there are also some individual and household level studies that focused only on smoking participation and did not take into account the level of cigarette consumption. For example, Marsh and McKay (1994) investigated the

[35] For example, potential simultaneity biases are less likely in individual level analysis since an individual cannot smoke enough to affect the market price. It can also be argued that individual level income data and the socio-economic determinants of cigarette demand are less likely to be correlated with price and policy regulations compared to aggregate data (Chaloupka and Warner, 2000).

relationship between smoking and a range of household characteristics using logistic regression analysis for low income families in the UK. They found that income does not have an important effect on the probability of smoking. However, Fry and Pashardes (1988) also examined the relationship between smoking and household characteristics using logistic regression analysis for the UK and they found that smoking prevalence increases with income but at a decreasing rate. This result can be attributed to not restricting the analysis to low-income households[36].

It should be acknowledged that individual level analysis also has some limitations. First, in the existing literature it is stated that self-reported cigarette consumption may be inaccurate and survey-based self-reported cigarette consumption may significantly understate actual consumption (Warner, 1978). In this regard, Hatziandreu *et al*., (1989) investigated the hypothesis that self-reported cigarette consumption is not a reliable measure to examine tobacco consumption pattern. With this aim, they compared the total self-reported cigarette consumption using data from the National Health Interview Surveys and the National Household Surveys on Drug Abuse with the adjusted consumption data from cigarette excise taxes using data from the US Department of Agriculture for the period of 1974-1985. They found no statistical difference between them from year to year, which indicates that there is no underreporting of cigarette smoking in these surveys. In light of this finding, they concluded that cross-sectional surveys of self-reported cigarette consumption are reliable data to evaluate cigarette smoking behaviour. The second limitation is that the price variable used in the individual level analysis is generally an aggregate level weighted average price. However, this price measure may not be representative of the real cigarette price in the regions where smokers live and, in addition, cannot capture the different type and quality of cigarettes. Third, as stated above, the majority of the existing individual level studies mainly focused on the effect of cigarette prices on cigarette demand. However, tobacco control policies other than price and tax may also have significant impacts on cigarette consumption. Studies excluding appropriate policy variables in the model may produce biased estimation results.

[36] It should be noted that Fry and Pashardes (1988) derived the smoking indicator by examining household tobacco expenditure and it only gives information on whether or not there is a smoker in the household (Dorsett, 1999). This derivation process may suffer from the infrequency of expenditure and recall and response bias problems which are common in expenditure surveys (Jones, 1995).

Notwithstanding these limitations, the most important point in terms of using individual-level data in the analysis is that it allows the investigation of both smoking propensity and intensity. A distinguishing feature of the distribution of individual level tobacco consumption data is that it is generally skewed to the right and includes a high number of zero observations. Several econometric estimation methods have been used in previous studies using micro level data to deal with these distributional characteristics. These are the two-part/hurdle count data models (Lewit and Coate, 1982 for the US; Wasserman, 1991 for the US; Hu *et al.*, 1995 for California), the double-hurdle models (Jones, 1989a and 1989b for the UK; Blaylock and Blisard, 1992a and 1992b for the US; Yen and Jones, 1996 for the UK) and the zero-inflated models (Mullahy, 1997 for the US; Sheu *et al.*, 2004 for California; Bauer *et al*., 2007 for Germany).

3.3.2.1 The Double-Hurdle Model

Most of the studies have highlighted the shortcoming of the standard Tobit model (Tobin, 1958) in cross-sectional analysis of cigarette consumption since the Tobit model assumes that zero consumption represents a corner solution of a standard consumer demand problem and ignores zero consumption arising from abstention (Aristei and Pieroni, 2008) (see, for example, Blundell and Meghir, 1987 for the UK; Yen and Jones, 1996 for the UK; Garcia and Labeaga, 1996 for Spain). In addition, as stated above, accounting for cigarette consumption data features, such as excess zeros, is important in terms of obtaining unbiased estimation results.

Some studies have used the double-hurdle model, which was originally proposed by Cragg (1971), to deal with excess zeros in the data. The main characteristic of the double-hurdle model, which is based on bivariate normality of unobserved errors, is that the participation decision and the level of consumption are generated by separate individual choices and the factors affecting these two decisions are allowed to differ (Jones, 2000). In other words, the main assumption of the double-hurdle model is that a strictly positive level of consumption can be observed if an individual passes two separate hurdles (Blaylock and Blisard, 1992a). This model is generally used to analyse household or individual tobacco expenditure patterns (Aristei and Pieroni, 2008)

In this context, Jones (1989a) employed the double-hurdle technique to model starting smoking, quitting smoking and cigarette consumption as separate decisions using data

from the UK General Household Survey. This approach provided information on how socio-economic characteristics affect these different decisions and allows 'non-starters' to be separated from 'quitters'. He stated that it is possible to exploit certain characteristics, which are directly associated with the qualitative distinction between 'smokers' and 'non-smokers' and which are independent of the quantity approach. Furthermore, he emphasised that the empirical results suggest the importance of modelling starting and quitting smoking as separate decisions. He found that education has a negative influence on the probability of being a smoker and its influence on starting is more than on quitting. The presence of other smokers in the household is also found to be positively associated with all of the participation equations.

Blaylock and Blisard (1992a) extended the study of Jones (1989a) by including an extensive set of demographic explanatory variables. They also used a double-hurdle model to examine both participation and consumption of tobacco for low-income women in the US using the USDA Continuing Survey of Food Intakes by Individuals (CSFII) for 1985-1986. Their results indicated that the number of children, region, education, ethnicity and the presence of an adult male are the most important factors affecting the smoking participation decision whereas age, race, region, ethnicity and health status are the most significant factors influencing the level of consumption.

Although the double-hurdle model is commonly used in the existing literature, there is an important limitation of this model. The standard double-hurdle model is based on the assumption of bivariate normal distribution for the error terms (Blundell and Meghir, 1987; Jones 1989a). However, the maximum likelihood estimation results will be inconsistent if the normality assumption for the error terms does not hold (Arabmazard and Schmidt, 1982). While the normal distribution may be appropriate for applications based on cigarette expenditure data, this limitation becomes particularly important if the model is applied to a dependent variable that has a highly skewed distribution, which is often the case for survey data on cigarette consumption (Jones, 2000). Jones and Yen (2000), therefore, generalised the double-hurdle model with dependence between smoking participation and the level of consumption by applying a Box-Cox transformation on the dependent variable. This transformation relaxes the normality assumption on the conditional distribution of the dependent variable and is used by Yen and Jones (1996), Jones (1995) and Garcia and Labeaga (1996).

3.3.2.2 The Zero-Inflated Model and the Two-Part/Hurdle Count Data Model

Count data models, where the dependent variable consists of nonnegative integer values (i.e. the number of cigarettes smoked in a specific time period), are being increasingly used in applied econometrics over recent decades. In particular, the zero-inflated negative binomial (ZINB) model, which is designed to deal with two common issues that occur with the application of the Poisson model to count data, namely overdispersion and excess zeros (Greene, 1994), is widely used in the existing literature. Overdispersion means that the variance of the count-dependent variable usually exceeds its mean and excess zeros can be defined as the existence of more zeros in the data than the Poisson model predicts (Cameron and Trivedi, 1998).

Sheu *et al*. (2004), for example, have attempted to investigate the price sensitivity of smokers in California using the Behavioural Risk Factor Survey (BRFS). They employed the ZINB model and found that the increase in cigarette prices during the period 1996-1999 has no effect on smoking participation. However, it was found to be effective in reducing the number of cigarettes smoked. In terms of socio-economic factors, their results indicated that white, black, male, single, less educated, unemployed, younger individuals, individuals with less income and with poorer health are more likely to smoke whereas the volume of cigarette consumption is higher for those who are white, male, divorced, older and with poorer health. Similarly, Bauer *et al.* (2007) investigated smoking behaviour by estimating count data models for males and females separately using the German Socio-Economic Panel for 1998, 2002 and 2004 and the ZINB model was chosen as the best specification for the data based on the LR and Vuong tests. They found some important differences in the cigarette smoking behaviour of males and females particularly in marital status and employment status.

The two-part/hurdle count data models, which are based on the assumption that the zero and the non-zero values occur from two separate processes, have also been extensively used in the existing literature. Wasserman *et al.* (1991), for example, estimated a generalised linear model and a two-part/hurdle model to investigate adult cigarette demand in the US using data from the National Health Interview Survey (NHIS) over the period 1970-1985 and youth cigarette demand with data from the Second National Health and Nutrition Examination Survey. They found that among adults, price is inversely associated with the decision of initiation and that the price elasticity of cigarette demand is similar for teenagers and adults. They also stated that certain socio-

economic characteristics such as higher levels of education, being single and being white are positively associated with cigarette consumption. In addition, tobacco control policies, which ban smoking in public places, have an important effect on both adult and teenage cigarette demand.

3.3.3 Panel Data Studies

The most important shortcoming of the studies based on cross-section data is that information on a given sampling unit is only observable in one specific time period. The absence of longitudinal data does not allow for taking into account addictive behaviours such as cigarette consumption while controlling for socio-demographic characteristics. Panel data on cigarette smoking is particularly useful since it allows for the effect of previous cigarette consumption behaviour on current consumption behaviour (Dorsett, 1999). However, there are only a few studies which can take advantage of panel data in analysing event dependence, in the context of smoking, mostly due to the unavailability of panel data. For example, Dorsett (1999) investigated the demand for tobacco among lone mothers over the period 1991-1996 in the UK using panel data from the Programme of Research into Low Income Families (PRILIF). He found that smoking in the previous time period is more likely to affect smoking for older individuals compared to younger individuals. In addition, higher levels of education were negatively associated with smoking, which may reflect a positive relationship between education and the knowledge about the health risks associated with smoking.

3.3.4 Studies on Different Aspects of Cigarette Demand

As stated above, one of the important advantages of using individual level data is that individual level analyses can examine the factors associated with the probability of being smoker and the level of cigarette consumption in different population segments. In this context, some cross-sectional studies have focused on different aspects of cigarette demand such as education and gender (Farrell and Fuchs, 1982; Yen, 2005; Bauer *et al.*, 2007). Farrell and Fuchs (1982) investigated the association between education and smoking and found evidence that the level of education is not statistically significant in explaining smoking behaviour. They concluded that a causal relationship between education and health is not always the case since a negative education and smoking correlation could be the result of an unobserved third variable, which has an influence both on education and smoking. Fuchs (1982), on the other hand, argued that

both education and smoking are associated with individual differences in time discounting rates, which implies that individuals with low rates of time preference are more likely to invest in both education and health. In this respect, Kenkel (1991), who used a direct measure of health knowledge, found that most of the effects of schooling on smoking behaviour remain after accounting for differences in health knowledge.

Yen (2005) investigated gender differences in cigarette consumption using data from the CSFII for the period 1994-1996 in the US and estimating double-hurdle models. He criticised other studies focusing on gender differences on the basis of only including a gender dummy variable in their analyses or for only focusing on differences in the price and income elasticities of tobacco consumption between males and females. The results of the LR test indicated that the assumption of equal consumption patterns is rejected and the price elasticities of cigarette demand are found to differ between males and females. Chaloupka (1990) is also one of the few studies which treated males and females separately in the analysis by estimating cigarette demand equations which were derived from the rational addictive model. One of the interesting findings was that males are sensitive to changes in the cigarette prices whilst females are virtually unresponsive to changes in the price of the cigarettes in the US.

3.3.5 Studies on the Cigarette Demand in Turkey

Studies on the cigarette demand in Turkey, which is a tobacco producing country with a high prevalence of cigarette smoking, are surprisingly limited. Tansel (1993) used annual time-series data on cigarette consumption per adult over 15 years old for the period from 1960 to 1988 to estimate a cigarette demand model for Turkey. A series of double-log models that include income, price and an indicator for the years in which health warning labels were required on cigarette packages were estimated. In addition, an indicator for the years when anti-smoking media campaigns were in place, the enrolment rates of secondary and higher education and lagged cigarette consumption were included in the analysis. In all specifications, she found a negative effect of price on cigarette demand while, consistent with the assumption of addictive behaviour theory, a positive effect of lagged cigarette consumption on current consumption. She also stated that health warning labels have a negative but small effect on cigarette smoking and income has a strong positive effect on cigarette demand in Turkey. It should be noted that the findings in this study were based on aggregate time-series data,

which can lead to biased results due to problems of aggregation, as discussed above. In addition, using aggregate data cannot provide information on the effect of individual-specific characteristics on cigarette consumption.

Onder (2002) is another study on cigarette demand in Turkey which includes both aggregate time-series and cross-sectional analyses. In the first part of the study, she used aggregate level data from 1960 to 2000 to estimate price and income elasticities for cigarette consumption using Generalised Least Squares. It was found that income, which was the only statistically significant factor, has a positive relationship with cigarette demand which implies that cigarettes are normal goods. The findings also indicated that the regulation variable, which represents banning smoking in public places and the advertising of cigarettes, has no significant effect on cigarette demand. In the second part, she examined the demand for cigarettes at the household level using the 1994 HICES and estimated a two-step model to analyse the relationship between poverty and cigarette consumption. She found that price and income both have a negative relationship with smoking participation for most income quintiles. Other findings suggested that males are more likely to smoke whilst higher levels of education and being employed are both associated with a lower probability of smoking. Households are found to be more responsive to price when deciding how much to smoke than when deciding whether or not to smoke. However, it can be argued that the estimation results are potentially biased since she used a censored Tobit model, which ignores zero consumption arising from abstention and assumes that all zero observations represent a corner solution of a standard consumer demand problem.

Bilgic *et al.* (2009) analysed household cigarette demand in Turkey using the 2003 Household Budget Survey (HBS) for households with and without teenagers[37]. They estimated a ZINB model for cigarette demand. The results indicated that there are some potentially relevant differences between the factors that affect the participation decision and the number of cigarette packs smoked per week. Both Bilgic *et al.* (2009) and Onder (2002) used the household as the decision unit in their analyses due to data limitations. Thus, the cigarette consumption for each household member is aggregated at the household level. However, some information related to the interaction between individual-level characteristics and cigarette consumption may be lost in such a process.

[37] The reason for treating each group separately is based on the parents' or family 'stigma' effect, which implies that households with teenagers tend to underreport cigarette consumption.

Moreover, using the household as the unit of analysis may generate biased estimation results since tobacco consumption is mainly the result of an individual decision.

In the case of Turkey, it is apparent that there is very little existing analysis of the determinants of cigarette smoking and, to my knowledge, there is no study focusing on cigarette demand at the individual level. However, in recent years, the existing literature emphasises the use of microeconomic data, particularly individual level data, in the empirical analysis of cigarette consumption since understanding the profile of a cigarette smoker is needed to develop effective policies for the segments of population at most risk of suffering from smoking related health problems. Furthermore, so far, no study appears to have formally analysed gender differences in cigarette consumption from a microeconometric approach for Turkey. In this context, another focus of this study is to fill this gap in the existing literature by exploring gender differences in cigarette consumption patterns.

3.4 DATA AND METHODOLOGY

3.4.1 Data

The Global Adult Tobacco Survey (GATS) for 2008 forms the basis of the empirical analysis presented in this chapter. The GATS enables countries to collect data on the consumption of tobacco products and the key anti-tobacco regulations using a standardised sample design procedure and a standard questionnaire for adults aged 15 years and over (GATS Turkey Report, 2010).

The GATS, which was carried out within the context of a project with the WHO and the US Centre for Disease Control and Prevention (CDC), was implemented with the aim of assessing the outcomes of the introduction of tobacco regulations and examining and evaluating the frequency of tobacco use in Turkey for the first time in November 2008. The survey was administrated by the Turkish Statistical Institute (TurkStat) as a nationally representative household survey of individuals aged 15 years and over (GATS Turkey Report, 2010). The most important feature of this survey for Turkey is that it is a unique individual-level data source in terms of providing data on the use of tobacco and tobacco products by adults as well as information on exposure to tobacco smoke, information on individuals' attitudes and perceptions on media and health warnings.

Basically, the questionnaire used in the survey includes eight main sections; background characteristics (age, gender, education and employment status), tobacco smoking (tobacco use status of the individuals (i.e., daily smoker, occasional smoker, non-smoker, former/past smoker, consumption volume of different products such as cigarettes, pipes and cigar), smokeless tobacco, cessation (advice to quit smoking by health care providers, methods used to try to stop smoking), second-hand smoke (SHS) (smoking allowed in the home, exposure to SHS at home and indoor smoking policy at the workplace), economics (type of tobacco products and quantity bought, cost of the tobacco product), media (exposure to pro-cigarette marketing on television, radio, billboards, newspapers, magazines, cinema, internet and public transportation, exposure to tobacco promotion activities, reaction to health warning labels on cigarette packages), and knowledge, attitudes and perceptions (knowledge about the health effects of both smoking and smokeless tobacco) (GATS, 2008).

The sample design included all settlements in Turkey but excluded villages with populations less than 200. A sample consisted of 11,200 households was used for the GATS based on the results of the 2006 Turkey Time Use Survey[38]. A three-staged, stratified systematic cluster sample was employed as a sampling technique of the survey. In the first stage a total of 400 clusters (primary sampling units (PSU)), 200 from urban and 200 from rural areas, were selected. In the second stage, 28 households were selected systematically within each selected PSU. In the final stage, one eligible individual (aged 15 years and over) was selected randomly within each selected household by using a household roster that includes all eligible members of the household by gender. The GATS was administrated to the selected 11,200 households throughout the country and the overall household response rates were 88% overall, 89.3% for urban and 86.7% for rural. The household roster was completed in a total of 9,322 households. The individual response rates were 97% overall, 97.7% for urban and 96.3% for rural (GATS Turkey Report, 2010). In total, 9,030 individual interviews were completed, 4,584 urban and 4,446 rural and 4269 males and 4761 females.

3.4.2 Estimation Methods

In order to investigate gender differences in cigarette smoking, count data models are estimated since the dependent variable, the number of cigarettes smoked per day, consists of nonnegative integer values. Count data models are appropriate when the sample is concentrated on a few small discrete values. In contrast to the classical regression model, the dependent variable is discrete with a distribution that places the probability mass only at nonnegative integer values. On the other hand, similar to other limited or discrete dependent variable models, count data regression models are nonlinear with many properties and special features related to discreteness and nonlinearity (Cameron and Trivedi, 1998; 2005).

Although it is generally accepted as restrictive, the Poisson regression model is the starting point for count data models. In a basic Poisson regression model, the number of events that have occurred, y_i, (i.e. the number of cigarettes smoked per day in this study) corresponding to individual i follows a Poisson distribution with a conditional mean or rate parameter λ depending on the characteristics, $\boldsymbol{x}_i$, of the individual:

[38] The sample size was determined according to the requirements of the GATS Sample Design Manual which requires at least 8,000 completed respondent questionnaires, with 2,000 each for males and females living in urban areas and males and females living in rural areas (GATS Turkey Report, 2010).

$$\lambda_i = E(y_i|\boldsymbol{x}_i) = e^{x_i\beta} \tag{3.1}$$

The probability density function of y given $\boldsymbol{x}$ is;

$$Pr[y_i|\boldsymbol{x}_i] = \frac{e^{-\lambda_i}\ \lambda_i{}^{y_i}}{y_i!} \qquad y = 0,1,2,\ldots \tag{3.2}$$

The first two features of the Poisson distribution are;

$$E[y] = \lambda \ \ and \ \ V[y] = \lambda \tag{3.3}$$

As it can be seen from (3.3), the basic feature of the Poisson distribution is the equidispersion property which means that the count-dependent variable has the same mean and variance (Cameron and Trivedi, 1998; Winkelmann, 2008). The Poisson regression model has been usually criticised because of this assumption. The problem arises from the fact that the distribution is parameterised according to a single scalar parameter (λ) so that all moments of *y* depend on λ. In contrast, the normal distribution has separate parameters for location (λ) and scale (σ^2) (Cameron and Trivedi, 2005).

In this respect, one of the shortcomings of the Poisson model is a feature of the data called overdispersion which means that the variance usually exceeds the mean[39]. In the case of overdispersion, the estimation results based on Poisson regression will be inefficient. Furthermore, another shortcoming of the Poisson model is that its density predicts the probability of a zero count to be significantly less than it is actually often observed in the sample. This is called the excess zeros problem which can be defined as the existence of more zeros in the data than the Poisson model predicts (Cameron and Trivedi, 1998; Jones, 2000).

In general, unobserved heterogeneity and/or excess zeros in the data cause overdispersion (Cameron and Trivedi, 1998; Winkelmann, 2008). Unobserved individual heterogeneity can be relevant to this study since the sample includes individuals with a wide variation of smoking status such as former smokers, new smokers, regular smokers, social smokers, heavy smokers and non-smokers. Although, a range of socio-demographic characteristics are used in this study, which are explained in detail in Section 3.4.3 below, they might not be able to capture these different heterogeneous cigarette consumption behaviours (Sheu *et al.*, 2004). Therefore, the

[39] The existence and magnitude of overdispersion can be obtained by comparing the sample mean and variance of the count dependent variable (Cameron and Trivedi, 1998).

remaining unobserved cigarette smoking behaviour has to be taken into account in the model. This issue provides the motivation for many of the methodological developments in the existing literature.

One of these approaches is to use the Negative Binomial (NB) model which is extended to deal with unobserved heterogeneity and, thus, overdispersion in the data. In the NB model, unobserved heterogeneity is accounted for by adding an error term, ε, to the conditional mean of the Poisson distribution. In equation (3.4) below, *exp* (ε_i) is generally assumed to have a gamma distribution with mean 1 and variance α so that the conditional mean of y_i can be shown to be Negative Binomial with a conditional mean of λ_i and the conditional variance of $\lambda_i(1 + \alpha\lambda_i)$.

$$\widetilde{\lambda_i} = E(y_i|\boldsymbol{x}_i) = e^{(x_i\beta + \varepsilon_i)} \tag{3.4}$$

The NB distribution with a nonzero dispersion parameter, α, fits highly positively skewed data and is given by the following equation:

$$f(y_i) = \frac{\Gamma(y_i + \alpha)}{y_i!\,\Gamma(\alpha)} \left(\frac{\alpha}{\lambda_i + \alpha}\right)^{\alpha} \left(\frac{\lambda_i}{\lambda_i + \alpha}\right)^{y_i} \qquad y_i \geq 0\,; \quad \lambda_i, \alpha > 0 \tag{3.5}$$

The variance exceeds the mean since $\alpha > 0$ and $\lambda_i > 0$. From equation (3.6) below, it can be seen that if α approaches zero, *y* becomes a Poisson distribution and when it becomes larger, the distribution will be more dispersed (Cameron and Trivedi, 2005; Sheu *et al.*, 2004). Thus, it can be argued that the dispersion parameter controls the deviation from the Poisson distribution.

$$\frac{Var(y_i)}{E(y_i)} = 1 + \alpha E(y_i) \tag{3.6}$$

The solution for the overdispersion issue provided by the NB model, however, is not adequate if the overdispersion is not only resulted from heterogeneity but also from excess zeros in the data. The phenomenon of excess zeros is an issue for this study since most individuals do not report any cigarette consumption in a given day (74.3% overall, 57.4% for males and 89.5% for females). The issue of excess zeros can be addressed by applying zero-inflated (ZI) models which are designed as a solution to this problem by modelling the number of predicted zero observations explicitly and also by allowing for overdispersion (Lambert 1992; Greene, 1994; Long, 1997).

Although overdispersion can account for a large number of zero observations in the sample, it may be the case that there is something special about the zero observations and these excess zero counts may not be related to increased dispersion throughout the distribution (Jones, 2000).

The basic assumption of standard consumer choice theory is that every decision maker has a potential to consume all goods. However, this may not be true for some goods like tobacco. 'True-zero' observations, which are also called 'automatic zeros', indicate the individuals who are non-smokers irrespective of income and cigarette prices (i.e., abstention). These zero counts can also include 'false-zero' observations which imply being potential smokers who are unable to afford cigarettes at the current income and prices (i.e., a corner solution). Thus, these two types of zeros are originated by different consumer behaviour systems (Blundell, 1988; Fry and Pashardes, 1994; Harris and Zhao, 2007; Bilgic *et al.*, 2009).

The micro level implications of the difference between non-smokers and potential smokers are widely emphasised in the existing literature. If zero counts representing potential smokers are excluded from the estimation, then a potential sample selection problem may occur and parameter estimates will be biased[40]. Therefore, it is important to distinguish the nature of the high number of zero observations. As stated above, the zero-inflated count data models which allow for overdispersion can also deal with the high number of zero observations in the data (Gurmu and Trivedi, 1996; Winkelmann, 2008).

Zero-inflated (ZI) count data models, such as the zero-inflated Poisson (ZIP) or the zero-inflated negative binomial (ZINB) models, add extra weight to the probability of observing a zero by a mixing specification by assuming the zeros can arise from two different sources, as it is a mixture of a point mass at zero and a Poisson or NB distribution (Mullahy, 1986). Conceptually, this property can be seen as a splitting mechanism, which divides individuals into non-smokers with probability π_i and potential smokers with probability $(1 - \pi_i)$ (Jones, 2000). The former follows a binary distribution (such as a logit or probit model) with the probability 'spike' at zero whilst the latter follows the basic count distribution. In other words, a 'false-zero' count

[40] In particular, when the sample is divided between smokers and non-smokers on the basis of actual expenditure on tobacco, it is more likely to misclassify potential smokers as the non-smokers (Fry and Pashardes, 1994).

integer, or being a potential smoker, is generated under the basic count distribution and is added to the probability that generates a 'true-zero' state that accounts for the total probability of being non-smoker. In this case, the total probability of observing a zero count, or being a non-smoker, consists of two probabilities. The first one is the probability from a binary model that takes into account the non-smokers irrespective of income or cigarette prices, and the other one is the probability that the current observed non-smokers include potential smokers (Bilgic *et al.*, 2009). If P_0 is the probability of a zero, then the general form of the zero-inflated model is as follows:

$$Pr(y_i) = \begin{cases} P_0 + (1 - P_0)\, f(y_i) & y_i = 0 \\ (1 - P_0)\ f(y_i) & y_i > 0 \end{cases} \tag{3.7}$$

The ZIP distribution (Lambert, 1992) allows one to model the probability of a zero with probability π_i and the probability that the data arise from a Poisson distribution with probability $(1 - \pi_i)$. The ZIP model can be written as follows:

$$Pr(y_i) = \begin{cases} \pi_i + (1 - \pi_i)e^{-\lambda_i} & y_i = 0 \\ (1 - \pi_i)\frac{e^{-\lambda_i}\lambda_i{}^{y_i}}{y_i!} & y_i > 0 \end{cases} \tag{3.8}$$

Greene (1994) has extended the ZIP model by applying a NB distribution as a second process to deal with overdispersion. The ZINB model (Greene, 1994) is given by:

$$Pr(y_i) = \begin{cases} \pi_i + (1 - \pi_i)\left(\frac{1}{1+\lambda_i\alpha}\right)^{\alpha^{-1}} & y_i = 0 \\ \left[\frac{\Gamma(y_i+\alpha^{-1})}{y_i!\Gamma(\alpha^{-1})}\left(\frac{1}{1+\alpha\lambda_i}\right)^{\alpha^{-1}}\left(\frac{\alpha\lambda_i}{1+\alpha\lambda_i}\right)^{y_i}\right] & y_i > 0 \end{cases} \tag{3.9}$$

where Γ and α are the gamma function and the dispersion parameter, respectively. Since the Poisson model and ZIP models are not nested, similar to the NB and ZINB models, the Vuong non-nested test (Vuong, 1989) is used to see which model has a better fit for the sample (Sheu *et al*., 2004).

$$V = \frac{\sqrt{N}\bar{m}}{s_m} \tag{3.10}$$

where $m_i = ln\,[\widehat{P_1}(y_i|x_i)/\,\widehat{P_2}(y_i|x_i)]$ and $\widehat{P_1}(.)$ and $\widehat{P_2}(.)$ are the predicted probability densities of the two competing models. $\bar{m}$ is the mean and s_m is the standard deviation for the sample of m_i. Under the null hypothesis that the two models are indistinguishable, V asymptotically follows a standard normal distribution, so V $\geq$ 1.96

distinctly favours the ZI count model, whilst $V < -1.96$ distinctly favours the standard count model; otherwise, neither model is preferred (Greene, 1994).

The 'best' model among the ZINB, ZIP, NB and Poisson models can be chosen based on the following steps as proposed by Greene (1994). If the Vuong test indicates that the ZINB model is rejected in favour of the NB model, the splitting mechanism is not an issue. In this case, the NB model has a better fit for the sample. If the Vuong test indicates that the NB model is rejected in favour of the ZINB model, then the splitting mechanism is a source of overdispersion (Sheu *et al*., 2004). The source of unobserved heterogeneity can be tested by the LR test applied to α by comparing the ZINB model with the ZIP model.

In addition to the ZI models, the two-part/hurdle count data models are alternative approaches for this type of data. In the case of cigarette consumption, the participation decision is usually estimated assuming a discrete choice model (probit or logit) and determines whether or not the dependent variable is a zero whilst the second part, how much to consume, is a count model truncated at zero that only focuses on individuals who have cigarette consumption greater than zero (Jones, 2000). If P_0 is the probability of a zero, the mixture distribution can be written as follows:

$$Pr(y_i) = \begin{cases} P_0 & y_i = 0 \\ (1 - P_0)\ f(y_i) & y_i > 0 \end{cases} \qquad (3.11)$$

where $f(y_i)$ is a probability density function defined for $y_i > 0$.

The choice of the model, namely the ZI model or the two-part/hurdle count data model, depends on the question being answered since there is no consensus over which model is better. Sheu *et al.* (2004) argued that it is reasonable to assume that the decisions as to whether to smoke and how much to smoke are conceptually integrated rather than being separate. They also argued that the distinction between non-smokers and smokers, which is based on zero cigarette consumption or not, is oversimplified in the two-part models. In this respect, they concluded that the ZI models are more appropriate on *a priori* grounds. Grootendorst (1995) compared the two-part/hurdle count data model with the ZI model empirically where he analysed the effect of co-payments on the utilisation of prescription drugs by the elderly, using data from the 1990 Ontario Health Survey. The results indicated that the two-part/hurdle count data model outperforms

other specifications. Zorn (1996), however, using data on Congressional responses to Supreme Court decisions from 1979 to 1988, concluded that both ZI models and two-part/hurdle count data models lead to similar results in estimating and interpreting them in practice since each of these models is a special case of a more general data generating process. In the study presented in this chapter, in addition to the ZINB model, the two-part/hurdle count data model, where the NB specification is used in the second part, is estimated in order to explore the robustness of the results.

3.4.3 Variable Construction and Definitions

The individual is the unit of analysis for the study presented in this chapter. As stated above in Section 3.4.1, the GATS includes detailed socio-economic, demographic and behavioural information on the individuals. The choice of independent variables used in this chapter is based on the previous empirical literature (see, for example, Yen, 2005; Blaylock and Blisard, 1992a; Jones, 1989a and 1989b, Sheu *et al.*, 2004). The variables used in the analysis are defined as follows:

The number of cigarettes smoked per day (the dependent variable): The measure of cigarette consumption used in the analysis is the number of cigarettes smoked per day[41]. In the survey, each individual was asked "what is your tobacco usage status at present? Every day, less than every day or none" and if the response was every day, the individual was asked "How many cigarettes do you smoke per day?" If the respondent uses cigarettes but less than one per day, these observations are treated as zero consumption[42]. This volume based dependent variable is common in tobacco surveys and, unlike a tobacco expenditure variable, does not control for differences in the price or quality of cigarettes smoked (Yen and Jones, 1996). However, it can be argued that it is less likely suffer from the problems which are likely to arise in expenditure surveys such as the infrequency of expenditure, recall and response bias (Jones, 1995).

Age groups: In the existing literature, to control for how age-related health problems influence the cigarette consumption of individuals and to explore the existence of a lifecycle pattern in smoking behaviour, the age of individuals is included in the analysis (Aristei and Pieroni, 2008). Since age is provided in age groups instead of the actual age of the individual in the survey, age groups of 15-24, 25-44, 45-64 and 65 and over are

[41] This variable includes not only manufactured cigarettes but also hand-rolled cigarettes.
[42] The proportion of these observations is 3.5% for all individuals, 3.7% for the male sample and 3.3% for the female sample.

included in the analysis. These categories are included as dummy variables and the 'aged 15-24' category is the omitted category.

Employment Status: The employment status of the individual is included in the analysis. This variable consists of three different categories: not working (including the unemployed, students, the retired, housewives and individuals who are looking for a job), employed (including public and private employees but excluding the self-employed) and the self-employed[43]. These categories are included as dummy variables and the 'not working' category is used as the omitted category.

Education Status: Education has often been used in previous empirical studies (see, for example, Yen, 2005; Jones, 1989a; Blaylock and Blisard, 1992a) to control for the possibility that individuals with higher levels of education may be more aware of the adverse health effects of smoking[44]. The highest educational attainment of the individual converted into years is used in the analysis.

Cigarette Price: Cigarette price is a key determinant of cigarette consumption and is important for policy intervention. It allows for inferences to be made about the level of cigarette consumption and the burden of tobacco taxes. In this chapter, individual specific cigarette price, which has both advantages and limitations as compared to an aggregate level weighted average price, is used. As an advantage, this price measure can capture the different type and quality of cigarettes. However, it is self-reported and thus, depends on the accuracy of respondents in reporting the price of cigarettes that they use. In this context, the price of a packet of cigarettes is available in the survey and is obtained from the following survey question: "How much do you usually pay per packet of cigarettes you buy?" However, if the individual is not a smoker, the cigarette price is missing. This creates an important problem in terms of dealing with the missing observations which may bias the estimation results. To solve this problem, the average cigarette price is used for the missing observations in the analysis. To calculate the average cigarette price, both the reported and the estimated prices, which are obtained by regressing the natural logarithm of prices on individual characteristics, are used,

[43] This set of variables is constructed using the following survey question: 'Which of the following best describes your employment status within last 12 months?' The possible responses are: public and private employees, self-employed, student, housewife, retired, looking for a job and other.

[44] It should be noted that an interesting aspect of smoking behaviour in Turkey is that smoking is more prevalent among individuals, particularly females, with higher levels of education in contrast to most high-income countries (Bilir *et al.*, 2009).

which is an approach often employed in the existing literature (see, for example, Ross and Chaloupka, 2003; Yen and Lin, 2006; Bilgic, *et al.*, 2009).

Asset Index Quintile: The income of the individual is another key determinant of cigarette consumption. Unfortunately, the GATS does not ask individuals about their income. However, Yurekli and Perucic (2010) stated that asset (wealth) indices, which are generally based upon household ownership of consumer goods (such as televisions and washing machines) and characteristics of the dwelling (for example, type of roof and floor) can be used to proxy household socioeconomic status. Principal component analysis (PCA) is one of the standard techniques used to construct weights and generate a single linear index. In the GATS, the individual is asked whether the household owns, possesses or has access to any of the following items: electricity, water, fixed telephone, cell phone, television, radio, refrigerator, car, motorcycle and washing machine. To construct an asset index, all available items are included and, following Filmer and Pritchett (2001), PCA is conducted by using weightings on the first component (i.e. scoring factors)[45]. Each scoring factor is divided by the standard deviation of the sample size for the asset to obtain the final asset weight. The asset indices are then used to generate the break points that define five asset quintiles from the poorest to the richest and the poorest quintile is used as the omitted category.

Presence of other smokers in the household[46]: Social interaction is another important indicator that has an influence on consumer behaviour. An individual's daily behaviour is influenced by the environment in which they live and by the customs, laws and conventions of their culture. Social interaction has an important role in the process of learning about the availability or quality of new products and this is particularly the case for addictive goods (Jones, 1995). The standard approach used in the existing literature is followed, which regards social interaction as a form of 'externality' following Becker (1974), by including a dummy variable indicating the presence of other smokers in the household. In this analysis, a dummy variable is constructed which takes the value of 1

[45] To calculate the principal components, the *pca* code in Stata, which computes the correlation or covariance matrix of the variables, is used.

[46] This variable is arguably not ideal since it does not capture interaction outside the household and allows a potential simultaneity problem as members of the same household may all affect each other (Jones, 1995). However, in general, social interaction proxied by the presence of other smokers in the household was found to be one of the most important variables influencing cigarette consumption in the existing studies (see, for example, Blaylock and Blisard (1992a), Jones, 1989a, 1995, Bilgic *et al*., 2009).

if other smokers exist in the household using the 'passive tobacco use' part of the survey.

Media effect: In the GATS, there is a media section which includes some questions related to encounters with the media and advertisements within the last 30 days which differs from most of the other data sets used in the existing literature. Since the media has the potential to attract and hold the attention of all individuals, it is interesting to explore the effect of the media on smoking behaviour in order to formulate effective media campaigns. To examine the effect of the media on cigarette consumption, a control for whether the individual has seen any pro-cigarette marketing in the mass media is used in the analysis. The dummy variable equals 1 if the individual has noticed any images, announcements, signboards or advertisements encouraging smoking in shops, television, radio, billboards, posters, newspapers/magazines, cinemas, internet or mass transportation vehicles or bus stops or has noticed any sports or sports events which are associated with cigarette brands or firms during the last 30 days.

Knowledge, attitudes and perceptions: To investigate the effect of knowledge about the health effects of smoking on cigarette consumption, a dummy variable is used in the analysis which equals 1 if the individual states that he/she knows that tobacco use causes serious diseases[47]. Similar to the media variable, this variable is also different from commonly used explanatory variables in the existing literature.

Urban/Rural residence: This dummy variable equals 1 if the individual resides in an urban area and 0 if the individual resides in a rural area.

3.4.4 Descriptive Statistics

Descriptive statistics of all the dependent and independent variables used in the analysis are reported in the Appendix in Table A3.1. This table provides information on descriptive statistics for the continuous variables and percentage distributions for the categorical variables for the pooled data set and for males and females separately. The samples include 9030 individuals overall, 4269 males and 4761 females.

Table A3.1 indicates that the average number of cigarettes smoked per day is 4.5 for the sample including both smokers and non-smokers and is substantially different between

[47] This variable is constructed using the following survey question: 'To your knowledge and in your opinion, does tobacco use cause serious diseases?'

males and females, i.e., 8.1 and 1.2, respectively. Similarly, there is a considerable difference between males and females in the average number of cigarettes smoked per day (19.1 for males and 11.7 for females) for the sample including only smokers. The mean value of years of education is nearly two years greater for males than females. However, the mean value of years of education is higher in the female sample than in the male sample when only smokers are included. In addition, in the sample most of the males are employed (or self-employed) whilst the vast majority of females are not working.

Table A3.1 also shows that the sample is nearly equally divided between urban and rural areas and the proportions of individuals in each age group are similar across the pooled sample and the gender samples due to the sampling design. The variables related to the individuals' attitudes and perceptions of the media and health effects of cigarette smoking indicate that 11.7% of adults noticed pro-cigarette marketing in the past 30 days although it was prohibited by law in 1996. The proportion of individuals having seen pro-cigarette marketing is greater for males (15.6%) than females (8.1%). Over 95% of adults reported that they believe that smoking has adverse health effects and there is no significant difference in the level of knowledge of the effects of smoking on health between males and females. In addition, 37% of individuals (37% for males and 38% for females) reported that there is at least one other smoker in the household.

The prevalence of cigarette smoking is one of the most important indicators of the effectiveness of a country's anti-tobacco policies. Table 3.3 below presents the prevalence of smoking among current smokers, former smokers and those who have never smoked in the GATS sample. Current smokers are presented in four sub-groups: daily smokers, occasional smokers, occasional but formerly daily smokers and occasional but never daily smokers. Similarly, former smokers are divided into two sub-groups: former daily smokers and former occasional smokers. The results from the data indicate that nearly one-third (29.9%) of adults aged 15 and over were current smokers in 2008, representing 16 million adults which supports the case that Turkey is a country characterised by a high smoking prevalence rate. The prevalence of smoking for males (47.7%) is remarkably higher than the female smoking rate (14%) which corresponds to approximately 12 million males and 4 million females. Similarly, almost half of males (43.9%) and 10.7% of females were current daily smokers. Perhaps most importantly, more than 90% of males (43.9% out of 47.7%) and 76.4% of females (10.7% out of

14%), who were currently smoking, reported that they are daily smokers. However, the proportion of occasional smokers for males and females are similar (3.5% for males and 3.3% for females). Therefore, it can be said that daily smoking is more common as a form of cigarette smoking than occasional or social smoking in Turkey. In addition, one in four males (26%) and three quarters of females (76.1%) stated that they have never smoked cigarettes. The remaining 17.7% (26.3% of males and 9.9% of females) were daily or occasional former smokers at the time of the survey.

Table 3.3: Percentage of Adults by Smoking Status and Gender

Smoking Status	Overall	Male	Female
Current smoker	**29.9%**	**47.7%**	**14.0%**
Daily smoker	26.4%	43.9%	10.7%
Occasional smoker (formerly daily)	1.7%	2.1%	1.4%
Occasional smoker (never daily)	1.8%	1.7%	1.9%
Occasional smoker (total)	3.5%	3.8%	3.3%
Former smoker	**17.7%**	**26.3%**	**9.9%**
Former daily smoker	12.0%	20.8%	4.1%
Former occasional smoker	5.7%	5.5%	5.7%
Never smoker	**52.4%**	**26.0%**	**76.1%**

Source: Turkey Global Adult Tobacco Survey (GATS), 2008.
Note: The current smoker category includes both daily and occasional (less than daily) smokers.

Table A3.2 provides information on how the rate of smoking participation (i.e., the smoking prevalence rate) and intensity of cigarette consumption (i.e., the average number of cigarettes smoked per day) change across the different socio-economic characteristics of the individuals for the pooled sample and for males and females separately. Table A3.2 indicates that until the age of 45, there is an increasing trend in smoking participation, and then it follows a decreasing trend. Daily smoking is not a common form of smoking among elderly females (1.2%) whereas almost one in five (18.7%) elderly males is a daily smoker. The youngest group, aged 15-24 group, smoke fewer cigarettes per day as compared to the other age groups for both males and females. In addition, the youngest age group has the highest occasional smoking rate. Therefore, two important features of smoking patterns for the youngest group are observed: they smoke fewer cigarettes and most of them smoke irregularly.

With regard to education, Table A3.2 indicates that the prevalence of smoking is common among higher educated individuals whilst individuals with higher levels of

education smoke fewer cigarettes compared to individuals who have primary or secondary education. For females, both the propensity and intensity of smoking increase by the level of education, however there is no clear pattern among males. The data also indicate that the highest rate of tobacco use, among both the pooled sample and males, is found among self-employed individuals (47% and 50% for the pooled sample and for males respectively). Self-employed individuals also have the highest rate of average cigarette consumption per day across all samples. However, there is a different picture for the female sample. Employed females have the highest smoking prevalence rate (22.5%) whilst females who are not working have the lowest smoking prevalence rate and the lowest cigarette consumption per day.

The urban/rural residence variable indicates that there is no considerable difference in smoking prevalence between urban and rural areas (29.1% and 23.7% for urban and rural areas, respectively). Individuals living in rural areas consume more cigarettes, on average, (18.8) compared to individuals living in urban areas (16.5). The urban-rural difference is most apparent among females (15.8% for urban versus 5.4% for rural) whereas cigarette consumption behaviour is similar within urban/rural areas for males and both smoking prevalence and the level of cigarette consumption are higher among males in rural areas. With regard to socio-economic status proxied by the asset index quintiles, Table A3.2 indicates that there is no clear pattern for all samples. However, it is possible to say that the level of cigarette consumption is highest among the poorest quintiles for all samples whereas smoking participation is common among relatively rich quintiles especially for the female sample.

3.5 RESULTS

The count data models, described in Section 3.4.2 above, are estimated to investigate gender differences in cigarette consumption since the dependent variable, the number of cigarettes smoked per day, consists of nonnegative integer values. The different count data models are tested against each other using the Likelihood Ratio (LR) and Vuong tests (Vuong, 1989) for non-nested models. The Vuong test of the Zero-inflated Negative Binomial (ZINB) and Negative Binomial (NB) models for this study shows that V=36.65, V=34.53 and V=13.92 (see Tables A3.5, A3.6 and A3.7 in the Appendix) for all individuals, males and females, respectively. These values indicate that the ZINB model is the preferred model and, thus, the splitting mechanism is confirmed as one source of overdispersion which implies that abstention and corner solutions (i.e., true-zero and false-zero states) for excess zeros are in effect. Furthermore, the ZINB model is more appropriate than the Zero-inflated Poisson (ZIP) model as indicated by the LR tests in Tables A3.5, A3.6 and A3.7. The LR test (LR= 6524.13 for all individuals, LR= 5134.54 for males and LR= 1238.84 for females) rejects the null hypothesis of no overdispersion which indicates that the ZINB model is the best specification for this study ($p < 0.001$)[48].

In this section, the results of estimating both the NB and ZINB models for all individuals and for males and females separately are presented and discussed. As a sensitivity analysis, the two part/hurdle count data models are also estimated and the results of this model are also discussed in this section. All of the tables presenting the results for the NB, the ZINB models and the two part/hurdle count data models are presented in the Appendix.

An important focus of this chapter is on gender differences since cigarette consumption behaviours are likely to be different for males and females due to different physical conditions and life styles (Chaloupka, 1990). Yen (2005) emphasises that using a gender dummy variable in the empirical analysis implies that gender differences in cigarette consumption can be fully captured by this variable whereas all other coefficients are the same for males and females. Therefore, he argues that male and

[48] In fact, the ZINB model is plausible for this study because it accounts for both the excess zeros and the unobserved heterogeneity, which are the main sources of overdispersion. In this case, unobserved factors such as individual health concerns, individual implicit preferences about cigarette consumption or cigarette-pack specific factors may be identified as possible unobserved sources of heterogeneity (Bilgic *et al*., 2009).

female samples should be treated separately in modelling cigarette demand to explore the factors affecting cigarette consumption for males and females, which is crucial to the formulation and implementation of public policies. As suggested by Yen (2005), the role of gender and appropriateness of pooling the samples are tested for all models using the LR tests to decide whether the male and female samples should be pooled or included separately in modelling cigarette consumption. The results from the LR tests are presented in Table 3.4 below. The unrestricted regression is estimated for the male and female samples which excludes the gender dummy variable and the restricted regression is estimated for all individuals using a gender dummy variable[49].

Table 3.4: Likelihood-Ratio Tests for Gender Differences

	Log-likelihood value (number of parameters)					
Model	Pooled	Males	Females	LR	*Df*	p-value
NB	-14171.67(16)	-10556.58(15)	-3265.87(15)	698.45	16	<0.001
ZINB	-11794.60(16)	-8832.30(15)	-2822.80(15)	279.00	31	<0.001
Two-part	-11794.46(16)	-8832.28(15)	-2822.55(15)	279.24	31	<0.001

The results of the LR tests indicate that the hypothesis of equal parameters for males and females is rejected for every model (p-value < 0.001). This means that it is not appropriate to pool the samples for all models considered in this chapter and the models should be estimated separately for males and females. In this respect, as discussed in detail below, the results obtained for the pooled sample differ substantially from the results for both of the samples split by gender and the results obtained for males are different in several respects from those obtained for females.

3.5.1 The Results of the Negative Binomial Analysis

The results of the NB model are discussed as a first step since it provides a baseline model for modelling count data[50]. The results from the NB models are presented in Tables A3.3 and A3.4 in the Appendix. The results for all individuals indicate that males are more likely to smoke more cigarettes in a given day compared to females. For the age groups, the level of cigarette consumption is positively associated with

[49] The LR statistic is as follows: $LR=2(\text{Log } L_m + \text{Log } L_f - \text{Log } L_p)$ where $\text{Log } L_m$, $\text{Log } L_f$ and $\text{Log } L_p$ represent the maximum log-likelihoods for males, females and the pooled sample with the corresponding numbers of parameters k_m, k_f and k_p. The LR statistic is chi-squared distributed with degrees of freedom equal to $k_m + k_f - k_p$ (Yen, 2005).

[50] Although the Poisson regression model is generally accepted as a benchmark model for count data (Cameron and Trivedi, 1998), the results of the LR test indicate that overdispersion is an important issue for this study.

individuals between 25 and 64 years of age and is negatively associated with the oldest age group, aged 65 and over, compared to the reference group defined as individuals between 15 and 24 years of age, for both genders. However, the positive relationship between the level of cigarette consumption and individuals between 45 and 64 years of age is not statistically significant for females.

The most important difference in the level of cigarette consumption between males and females is apparent in the education variable. For females, there is a positive association between education and smoking intensity similar to the results for all individuals whereas education does not have a statistically significant effect on the level of cigarette consumption for males. In the existing literature, it is stated that education may improve the cognitive skills of an individual regarding the adverse health effects of cigarette consumption and, thus, it plays an important role in decreasing cigarette consumption (Blaylock and Blisard, 1992a and 1992b; Jones 1989a and 1989b; Yen, 2005)[51]. The results, however, are not in line with the existing literature that argues that higher human capital endowments make individuals more aware of the health risks associated with smoking. The results of the NB model, on the other hand, are similar across genders in terms of the employment status of the individuals. Employed and self-employed individuals are found to be associated with a higher probability of smoking more cigarettes in a given day compared to the reference group defined as individuals who are not working.

Social interactions are accepted as important factors behind many behavioural and economic outcomes based on the premise that the utility that an individual obtains from a certain activity depends on the behaviours of the other individuals in the individual's reference group. The net utility obtained by consuming a given good will then increase with the consumption of the same good by other individuals (Aristei and Pieroni, 2009). In the existing literature, it is assumed that the smoking behaviour of other individuals has a direct impact on an individual's decision to smoke (Jones, 1995). In line with previous studies[52] (see, for example, Jones, 1989a, 1995; Yen and Jones, 1996), the presence of other smokers in the household is found to be positively related to the level

[51] The positive effect of education may also reflect the influence of income, which can be interpreted as females with higher levels of education are more likely to participate in labour market and are more likely to have economic power to buy cigarette and this may increase their cigarette consumption.

[52] It should be noted that the presence of other smokers in the household may be endogenous. Therefore, the results for the proxy of social interaction should be treated with caution. It should also be noted that the estimation results in this study relate to correlations between the dependent and explanatory variables rather than causal relationships.

of cigarette consumption for both genders. Furthermore, the marginal effect of this variable is the largest in magnitude for both genders and for the pooled sample.

In terms of the relationship between cigarette prices and the level of cigarette consumption, the results of the NB model give the price elasticity of cigarette demand since the results of the NB model give semi-elasticities and the logarithm of cigarette price is used in the analysis. The results for the response to price changes are found to be substantially different across the male and female sub-samples. The price elasticity of cigarette demand is -0.49 and -1.96 for males and females, respectively. Since the price elasticity of cigarette demand for females is found to be higher than for males, it can be argued that females may be more sensitive to changes in cigarette prices arising from increased taxation of cigarettes. It should be kept in mind that an individual specific price variable, which has both some advantages and limitations as compared to aggregate price measures as stated in Section 3.4.3, is used in the analysis. The price elasticities of cigarette demand found in recent previous studies using both time-series and cross-sectional data fall within the range of -0.14 to -1.23, but most fall in the narrower range between -0.3 and -0.5 (Chaloupka and Warner, 2000). The price elasticity of cigarette demand reported in this study for all individuals, -1.25, approaches the upper bound of the wide range of the estimated price elasticities of cigarette demand found in previous studies. The results of the NB model also indicate that there is no clear pattern or relationship between socioeconomic status as measured by the asset index quintiles and the level of cigarette consumption for both genders. However, the results of the pooled sample indicate that individuals in wealthier quintiles are more likely to smoke more cigarettes compared to the poorest quintile.

It is also important for policy makers to know whether pro-cigarette marketing affects cigarette consumption in order to construct effective policies to decrease smoking prevalence. The Law 4207 has banned direct and indirect mass media advertising and promotion of cigarettes and other tobacco products in Turkey since 1996. Furthermore, this law was revised to ban sponsorship of all sports and cultural events by tobacco companies in 2008 (GATS Turkey Report, 2010). This advertising ban in Turkey constitutes an important component of the country's anti-tobacco strategy. Although tobacco advertising has been prohibited for more than 10 years, the descriptive statistics, presented in Table A3.1, demonstrate that one in eight of the adults noticed some advertising promotion and sponsorship. However, the results of the NB models

indicate that pro-cigarette marketing has no effect on the level of cigarette consumption for the pooled sample and for both genders.

Knowledge of the health risks associated with cigarette smoking is also included in the analysis. In the existing literature, it is widely acknowledged that educating and informing individuals, especially young individuals, about the adverse health effects of smoking play an important role in decreasing smoking prevalence. The results of the NB models indicate that perception or knowledge of the health risks of smoking is inversely associated with the level of cigarette consumption for males whilst this relationship is found to be statistically insignificant for females.

Finally, with regard to region of residence, the results indicate that the region of residence does not have a statistically significant effect on the level of cigarette consumption for males. However, females living in urban areas are found to be more likely to smoke more cigarettes in a given day compared to females living in rural areas.

3.5.2 The Results of the Zero-Inflated Negative Binomial Analysis

As stated above, the ZINB model is found to be the preferred specification for this study based on the results of the LR and Vuong (1989) tests. Although the NB model captures unobserved heterogeneity, the excess zeros are also an important source of overdispersion in this study. The results of the ZINB model[53], which are presented in Tables A3.5 to A3.7, are discussed and compared with those from the NB model in this section. In addition, for the purposes of comparison and interpretation of the results, the marginal effects of each variable are provided in the last column of the tables. The marginal effects indicate the average predicted number of cigarettes smoked per day.

The results of the ZINB model for all individuals indicate that the gender variable (i.e., being male) is positively associated with being a smoker and males are more likely to smoke more cigarettes in a given day compared to females. This finding supports the existing literature using cross-section data (Blaylock and Blisard, 1992a; Yen, 2005; Bauer *et al.*, 2007). The results further indicate that smoking participation and the level

[53] The ZINB model includes a logit part (i.e., the decision not to smoke) and a NB part which identifies the factors that influence the number of cigarettes smoked per day. The logit part of the ZINB model predicts the outcome of zero observations and, thus, indicates the likelihood of being in the non-smoking group (Sheu *et al.*,2004; Bilgic *et al.*, 2009). In order to compare the results from the logit part to those from the NB part, the coefficient signs of the logit part are interpreted as reverse signs.

of cigarette consumption reveal different patterns among different age groups, which is in line with the findings from the NB model and the descriptive statistics. In the existing literature, it is hypothesised that when an individual grows older, he or she is likely to become more aware of the health risks associated with cigarette smoking and more concerned about health (Yen, 2005). In particular, being aged 65 and over, when there is an increasing risk of chronic diseases, some of which are related to tobacco use, could lead to individuals giving up smoking. This hypothesis is partially supported by the findings in this study. Older individuals of the age of 65 and over have a lower probability of smoking for both males and females. However, women between the ages of 25 and 64 are found to be more likely to participate in cigarette smoking compared to individuals between 15 and 24 years of age whereas older men aged between 45 and 64 have a lower probability of smoking compared to younger individuals. The negative association between age and the probability of smoking for males arguably reflects the fact that most smokers start smoking at a younger age and this finding is consistent with the findings from existing cigarette demand studies (Jones, 1995; Sheu *et al.*, 2004; Yen, 2005). With regard to the amount of cigarettes consumed among smokers, those between the ages of 25 and 64 smoke more cigarettes than the reference group for both males and females. In this context, two main characteristics of the smoking patterns of individuals at different age groups are identified. First, younger adults are more likely to smoke fewer cigarettes than older adults and this finding is consistent across genders. Second, older male adults in particular have a lower probability of smoking compared to younger male adults.

In accordance with the results from the NB model, the most important difference in smoking prevalence and intensity across genders is evident in the education variable. For males, there is an inverse relationship between education and smoking prevalence and intensity, whereas the opposite relationship, surprisingly, exists for females. However it is expected that, as the level of education increases, participation and the level of cigarette consumption are likely to decrease (Yen, 2005). In this respect, the positive relationship between education and smoking behaviour may be explained by the increasing economic independence of females with higher levels of education. Escardibul (2005) stated, in her study for Spain, that the higher participation of females with higher levels of education in the labour market and the kind of jobs in which they are employed arguably imply greater interaction with males who have a higher

probability of smoking. Therefore, they are arguably induced to follow patterns of behaviour that are traditionally adopted by males such as smoking. Furthermore, participating in the labour market is potentially associated with higher levels of stress due to, for example, increasing responsibilities of caring for children and developing careers, which could encourage smoking. A similar positive association between the level of education and smoking behaviour has also been observed in some former Eastern Bloc countries (Jha and Chaloupka, 1999).

This striking difference between males and females in terms of the association between education and smoking behaviour highlights the importance of dividing the data into two separate groups. The results of the ZINB model for all individuals, where males and females are not treated separately, indicate that a positive relationship exists between the level of education and smoking behaviour. This finding could lead to misleading policy provision since education may be an effective instrument to decrease smoking prevalence especially among males in Turkey.

The findings regarding employment status indicate that employment has different influences on smoking propensity and intensity and the influence also differs in some respects across genders. Employed and self-employed individuals are found to be associated with a higher probability of being a smoker compared to the reference category defined as individuals who are not working. This finding is consistent between genders and the pooled sample. For males, employed individuals are more likely to participate in smoking but they are more likely to smoke fewer cigarettes in a given day than the reference group. It is plausible that an individual who is not working may smoke more cigarettes due to less structured daily activities than an employed individual who has to allocate time to work, travel to and from the work and perform other household related responsibilities on a daily basis (Bilgic *et al.*, 2009). For females, however, employment has no effect on the level of cigarette consumption and self-employment is positively associated with the level of cigarette consumption in a given day by a factor of 3.52 (=exp (1.26)). This means that compared to the average predicted number of events of 1.26, this represents a 35% increase in the number of cigarettes smoked per day for self-employed females. This finding may be explained by the increased economic power of working females and their social environment. Employed females are arguably more exposed to the opinions and habits of males in a working environment and they are less subject to the conventional constraints on female

behaviour. In addition, they have more access to financial resources with which to buy cigarettes and acquire more power in the bargaining position in the household (Waldron, 1988).

With respect to cigarette prices, the results of the ZINB models indicate that an increase in cigarette prices is statistically significantly associated with a decrease in the probability of being a smoker for both males and females in accordance with consumer theory. In the NB part of the model, the price coefficient is not statistically significant for males indicating that cigarette price does not affect the level of cigarette consumption. Similarly, the results of the pooled sample indicate that cigarette price has no effect on the level of cigarette consumption. However, the results indicate a negative relationship between cigarette prices and the level of cigarette consumption for females. This finding suggests that using taxes as a policy instrument may be particularly effective in reducing cigarette consumption among females in Turkey. Similar to the results of the NB model, the results of the ZINB models indicate that there is not a clear relationship between socioeconomic status as measured by asset index quintiles and both smoking propensity and the level of cigarette consumption in contrast to the expectation of increased smoking participation in response to the ascending socioeconomic status of individuals. It should be noted, however, that the asset index rankings only consider whether the household (or individual) owns, possesses or has access to different items rather than the current real income of the individuals, due to a lack of data, which may lead to this unexpected result. The effects of certain variables, however, such as the age, sex, employment and education status of the individual support the existence of an effect of the socioeconomic and cultural environment on smoking behaviour.

The results of the ZINB model confirm the results of the NB model in terms of the positive relationship between the presence of other smokers in the household and the level of cigarette consumption. The logit part of the ZINB models also indicates a positive relationship between the presence of other smokers in the household and smoking propensity for both males and females. Furthermore, the marginal effect of this variable is largest in magnitude for males and the pooled sample and this variable has the second largest marginal effect following that of cigarette price for the female sample. Due to the high prevalence of smoking, exposure to second-hand smoking at home is very common in Turkey. For this reason, this finding is not surprising for

Turkey. In addition, as Bilgic *et al.*, (2009) states, family members' relationships are very strong and close in Eastern societies, which is different from some Western societies. For example, offering and sharing cigarettes with one another between family members such as siblings, mothers and daughters are often encountered in Turkey.

Perhaps the most interesting variables used in this study are those related to pro-cigarette marketing in the media and perception or knowledge of the health risks of smoking behaviour. There is a vigorous debate over the association between pro-cigarette marketing in the media and smoking behaviour. On the one hand, it is argued by public health advocates that such advertising increases the consumption volume. On the other hand, the tobacco industry argues that pro-cigarette marketing does not recruit new smokers, but rather it encourages existing smokers to stay with, or switch to, a particular brand (Jha and Chaloupka, 1999). The existing literature focusing on the association between pro-cigarette marketing and cigarette consumption state either that pro-cigarette marketing has no effect on cigarette consumption or that there is a very small positive relationship between them (Jha and Chaloupka, 1999; Chaloupka and Warner, 2000). In this context, the results of the ZINB models indicate that pro-cigarette marketing has no effect on smoking participation for both males and females. This finding reinforces the hypothesis that pro-cigarette marketing has no effect on the decision to start smoking. However, a positive association is found between noticing pro-cigarette marketing in the media and the level of cigarette consumption for males, in contrast to the results of the NB model. This finding arguably implies that males are more likely to notice cigarette advertising and are more affected by pro-cigarette marketing than females. Similarly, perception or knowledge of the health risks of smoking is found to be inversely associated with the level of cigarette consumption for males while this relationship is found to be statistically insignificant for females. These differences between males and females highlight again the importance of dividing the data into two separate groups.

With respect to region of residence, the results of the ZINB models indicate that females living in urban areas are more likely to participate in cigarette smoking compared to females living in rural areas. This association is statistically insignificant for males. In accordance with the results from the NB model, the urban/rural residence dummy variable has no effect on the level of the cigarette consumption of males. Similarly, the urban/rural residence dummy variable is not a statistically significant determinant of the

level of cigarette consumption of females in the ZINB model. This finding differs from the results of the NB model, which indicate a positive relationship between urban residents and the level of cigarette consumption of females.

The findings related to region of residence are consistent with the hypothesis that the area where people live affects their smoking behaviour and indicates the importance of differences in lifestyles between urban and rural areas. This difference is particularly pronounced for females. Several factors may lead to the positive association between urban residence and smoking participation such as the social and physical environment in which people live. With regard to the social environment, urban areas may have a more tolerant environment toward smoking especially for females which encourages starting smoking (Eiser *et al.*, 1989). The physical environment may affect smoking habit by making life more stressful in urban areas, which then makes females turn to smoking as a way to overcome stress. Furthermore, other factors such as a large number of tobacco selling points and more illegal sales to younger individuals may contribute to higher smoking participation in urban areas (Idris *et al*., 2007).

In order to explore the robustness of the results, the two-part/hurdle count data model is also estimated for all samples. The estimation results for the two-part/hurdle count data model where a NB model is estimated for the second part are presented in Tables A3.8 to A3.10 in the Appendix for the pooled sample, males and females respectively. The findings suggest that the results of the two-part/hurdle count data models do not differ qualitatively from the results of the ZINB models for all samples.

3.5.3 Summary of the Key Findings

Growing attention in the existing cigarette smoking literature is devoted to the potential differences in the determinants of the decision to smoke (i.e. smoking participation) and the determinants of how much to smoke (i.e. the level of cigarette consumption). The differences between the factors affecting the two-observed choices provide an opportunity to improve the effectiveness of public policy that aims to decrease smoking prevalence (Bilgic *et al*., 2009). The ZINB model allows for examining both the determinants of smoking participation and the level of cigarette consumption. In this respect, the main finding of this chapter is that the factors associated with smoking participation are different from those that are related to the level of consumption among smokers.

Overall, in terms of the socio-demographic variables, those who are more likely to smoke are males, adults aged between 25 and 64, those with higher levels of education, employed and self-employed individuals, individuals who are living with other smokers and urban residents. Among potential smokers, the level of cigarette consumption is positively associated with being male, being aged 25-64, self-employment, living with other smokers, noticing pro-cigarette marketing in the mass media and is negatively associated with knowledge of the adverse health effects of smoking. Furthermore, when the sample is divided into males and females, as suggested by the LR test, the determinants of smoking propensity and intensity for males and females differ substantially from those of the overall sample. The most important difference is observed in the relationship between education and smoking behaviour. The findings suggest a negative relationship between education and smoking participation and intensity for males whilst the opposite relationship exists for females. Furthermore, males aged between 45 and 64 are associated with decreased smoking participation whereas the opposite pattern is observed for females. Similarly, employment is found to be negatively associated with the level of cigarette consumption for males whilst self-employed females are found to be more likely to smoke more cigarettes compared to females who are not working. Additionally, living in an urban area is positively associated with smoking participation for females while it has no effect on the level of cigarette consumption for both genders.

In terms of the other variables, cigarette price is negatively associated with smoking participation for both genders and cigarette price is negatively associated with the level of cigarette consumption for females. Knowledge about the adverse effects of smoking has no effect on both smoking participation and the level of cigarette consumption for females but it has a negative relationship with the level of cigarette consumption for males. Similarly, pro-cigarette marketing is positively associated with the level of cigarette consumption for males. These findings indicate that using tax policy may be more effective in decreasing smoking prevalence among females whereas anti-smoking policies based on increasing public awareness about the adverse health effects of smoking and tobacco advertising bans may also be used as policy instruments to decrease the cigarette consumption of males.

3.6 CONCLUSION

There has been long-standing attention to the empirical analysis of tobacco consumption since the deleterious effects of tobacco use are clearly an important public health issue. In order to reduce the rate of smoking prevalence among adults with effective anti-smoking policy and appropriate tobacco smoking prevention programs, policy makers need to know the potential factors which are associated with cigarette consumption. In this respect, this chapter examines the potential factors associated with adult smoking participation and the level of cigarette consumption at the individual level for males and females in 2008 by employing a large nationally representative data set for Turkey.

Turkey is an important country for investigating such issues since Turkey is characterised by a high smoking prevalence rate. In 2008, nearly one-third (29.9%) of adults aged 15 and over were current smokers in Turkey, which corresponds to 16 million adults. Furthermore, smoking is the leading cause of death among males and smoking prevalence is becoming increasingly higher among females. The male and female smoking rates are 47.7% and 14%, respectively which corresponds to approximately 12 million males and 4 million females (GATS, 2008). In addition, the tobacco environment has changed significantly in recent years as a result of new anti-tobacco policies adopted by the government. Turkey has achieved some progress in reducing tobacco consumption: the smoking prevalence rate among adults decreased from 44% in 1988 to 29.9% in 2008, yet it still has one of the highest male smoking rates in the WHO European Region (Bilir 1997; GATS Turkey Report, 2010). Thus, tobacco use is still among the major problems and public policy areas in Turkey.

In this chapter, count data models were estimated to investigate both the propensity and intensity of cigarette consumption in Turkey and the results of the Vuong and LR tests indicated that the ZINB model is the preferred specification for this study. The results indicated that the factors associated with smoking participation differ from the factors associated with the level of cigarette consumption. Following Sheu *et al*., (2004), the justification for using the ZINB model rather than the two part/hurdle count data model is that the decisions as to whether or not to smoke and the level of consumption are assumed to be related rather than being separate and that both decisions are made by the same individual. However, the two-part/hurdle count data model is also estimated to explore the robustness of the results and the results of both models are found to be

consistent, which suggests that each of these models is a special case of a more general data generating process (Zorn, 1996).

An important focus of this chapter is to investigate cigarette consumption in Turkey from a gender perspective. With this aim, the hypothesis of equal parameters between males and females was tested using the LR test and the results indicated that the models should be estimated for males and females separately. However, for the purposes of comparison the results of all models which are estimated for all individuals were also presented and discussed. The results indicated strong empirical evidence of gender differences in tobacco use in terms of parameter estimates and the marginal effects of the independent variables. In this context, gender specific estimations are found to be more informative than those obtained from the pooled sample which confirms the results of the LR test suggesting that the male and female samples should be treated separately.

The gender specific estimations highlight arguably the most interesting finding related to the association between education and smoking behaviour. Higher levels of education are found to be positively associated with both smoking participation and the level of consumption for females and the pooled sample but the opposite association is found for males. These findings indicate that public education through, for example, mass media, about the health dangers of smoking could be an effective policy instrument to decrease smoking prevalence and the level of cigarette consumption especially among males. The results also indicate that knowledge about the adverse effects of smoking is only statistically significantly negatively associated with the level of cigarette consumption for males. In addition, living in an urban area is found to be positively associated with smoking participation for females. This finding indicates that it may be appropriate for the forms of anti-smoking policies to differ between urban and rural areas, which may affect the cigarette consumption decisions of females in particular.

Another important difference between the results of the ZINB models for males and females is apparent in the age categories. Being aged between 45 and 64 is negatively associated with smoking participation for males compared to individuals between 15 and 24 years of age whereas the opposite relationship exists for females. This finding arguably implies that males are more likely to participate in smoking at an early age in Turkey while females participate in smoking at later ages which is generally true for

females living in low and middle income countries and is explained by gender inequality (Waldron, 1988). On the one hand, the existence of gender inequality is associated with a delayed increase in the rate of female smoking participation which may be due to social disapproval of female smoking. On the other hand, the decrease in gender inequality due to emancipation may lead to an adaptation of male tobacco consumption behaviour by females (Flandorfer *et al*., 2010). In light of the findings related to the relationship between age and smoking behaviour, it can be argued that education campaigns about the health dangers of tobacco use should be disseminated to schools to prevent initiation of smoking, particularly for males, to have a long-term impact. Furthermore, the results indicate that younger adults are more likely to smoke fewer cigarettes compared to older adults for both genders. In this respect, as Sheu *et al*., (2004) argues, increases in cigarette prices alone may not be an effective way to reduce the level of cigarette consumption among youths since they do not smoke much and have not been addicted to smoking for a long time.

The use of tax policy, on the other hand, has been proven to be effective in reducing smoking prevalence in Turkey (Onder, 2002; Bilgic *et al*., 2009). In this context, the structure and the level of cigarette taxes in Turkey has changed significantly in the recent decade. The Special Consumption Tax on tobacco products increased by 20% in January 2010 and the total tax increased to 78% (GATS Turkey Report, 2010). This increase is a little above the level recommended by WHO, which is 75% (WHO, 2009). The number of cigarettes sold during the first four months of 2010 was 25% lower than the number of cigarettes sold during the same period in 2009 (GATS Turkey Report, 2010). The results of this study also indicate that cigarette price is negatively associated with smoking participation for both genders. Furthermore, females are found to be more sensitive to price increases compared to males for the year considered in the analysis. It is possible to argue that increases in the cigarette price could be an effective policy to reduce smoking prevalence especially among females. However, one important problem related to cigarette taxes is to determine the appropriate level of taxation since the distributional consequences of higher taxes should be considered. Under the current cigarette tax system in Turkey, the tax on high-quality brands (nearly 50%) is lower than the tax on lower-quality brands (nearly 70%) (Bilgic *et al.*, 2009). This tax difference has a distributional impact because poorer individuals generally tend to purchase low-priced cigarettes contributing more to total tax revenues compared to

richer individuals. Therefore, the government should consider this dimension of the tax policy while designing the tax rates. Additionally, there is no regulation related to nicotine and tar levels in cigarettes in Turkey which makes Turkish cigarettes more addictive than those in developed countries. Addiction, in turn, arguably may lead to less sensitivity to changes in the cigarette prices (Bilgic *et al.*, 2009). The anti-tobacco policies could be more effective if stricter regulation of nicotine levels is applied to Turkish cigarettes.

The ban on pro-cigarette marketing in Turkey has been successfully imposed on direct advertising in the print and broadcast media. However, additional measures are required to eliminate indirect and hidden advertising since the descriptive statistics demonstrate that exposure to tobacco advertising is still noticed by individuals despite the fact that these forms of advertising were illegal at the time of the survey. Furthermore, the results indicate a positive association between pro-cigarette marketing and the level of cigarette consumption for males which highlights the need for better enforcement. It is particularly important to note that the presence of other smokers in the household is found to be positively associated with both smoking participation and the level of cigarette consumption across all samples for the year considered in the analysis. This finding, which is consistent across all samples, reinforces the finding by Jones (1995) that social interaction is an important determinant of cigarette consumption. In view of this finding, a possible public policy recommendation may be to continue to create smoke-free environments by prohibiting smoking in all indoor public places, places of employment, public transportation and places of entertainment. This policy is important not only for preventing exposure to second-hand smoking but also for reducing the effect of social interaction on smoking behaviour.

There are nearly 16 million smokers in Turkey and there are strong addictive dimensions of smoking which require more information on the costs and achievements of alternative anti-tobacco policies. It is, therefore, important to recognise that designing effective and economically efficient policies to decrease smoking prevalence is inherently difficult on the basis of the findings of this chapter. However, this study extends the existing literature by being the first attempt to investigate gender differences in cigarette consumption in Turkey at the individual level using a large nationally representative data set. At this point, it is important to acknowledge that there are some limitations of this study. First, the survey based on self-reported tobacco use which may

underestimate actual consumption particularly among young adults (Warner, 1978). This measure depends on the accuracy of individuals in reporting their cigarette consumption. In this respect, the limitation of using self-reported cigarette consumption measures is that they are not based on objective data that would not be affected by changing attitudes towards smoking. Second, due to the cross-sectional nature of the survey, current cigarette consumption is assumed to be affected only by current price. Therefore, it is assumed that any change in past or future prices will not influence current cigarette consumption. Consequently, this framework does not account for the inter-temporal dependence of tobacco consumption which is an important feature of addictive behaviour (Dorsett, 1999). Third, some of the variables that are commonly used in the existing literature such as health status, the marital status of individuals and alcohol consumption, which arguably occurs simultaneously with cigarette consumption, cannot be included in the analysis since information on these variables is unfortunately unavailable in the GATS. However, as Bilgic *et al.*, 2009 state, whether cigarette and alcohol consumption occur simultaneously depends on the cultural environment and cigarette consumption does not always imply alcohol consumption in a Muslim country. Finally, the asset index quintiles allow the exploration of the relationship between the socio-economic status of the individual and cigarette consumption behaviour. However, as Yen and Jones (1996) emphasise, the use of proxy variables for income may result in measurement error that may be correlated with other variables such as education. Hence, the results related to the socio-economic variables should be treated with caution.

In this context, future studies might investigate the inter-temporal dependence of cigarette consumption at the individual level over time but this depends on the availability of the relevant longitudinal data, which are not currently available for Turkey. Furthermore, identification of the determinants of youth smoking would also be an important contribution to the existing literature since the high smoking prevalence rate among youths implies that Turkey will face additional health and economic consequences from cigarette smoking in the future. In addition, the distinction between daily and occasional smokers is important since considering only the dichotomy between smokers and non-smokers may not be sufficient for developing effective anti-smoking policies (Sheu *et al*., 2004). The effect of nicotine addiction and the decision to quit smoking may also be potentially useful areas for future research.

APPENDIX TO CHAPTER 3

Table A3.1: Descriptive Statistics for the Continuous Variables and Percentage Distributions for the Categorical Variables

Continuous Variables			
Variables	**Overall** (n=9030)	**Male** (n=4269)	**Female** (n=4761)
Number of cigarettes smoked per day	4.5 (17.5)**	8.1 (19.1)	1.2 (11.7)
(St. Dev.)	9.5 (11.1)	11.9 (11.1)	4.5 (8.6)
(Min)	0 (1)	0 (1)	0 (10)
(Max)	100 (100)	100 (100)	70 (70)
Years of education	6 (7.1)	7 (7.1)	5.1 (7.4)
(St. Dev.)	4.1 (3.7)	3.8 (3.5)	4.1 (4.0)
(Min)	0 (0)	0 (0)	0 (0)
(Max)	21 (21)	21 (21)	17 (17)
Cigarette Pack Price*	1.5 (1.5)	1.5 (1.5)	1.5 (1.4)
(St. Dev.)	0.2 (0.4)	0.3 (0.4)	0.1 (0.3)
(Min)	0.4 (0.4)	0.4 (0.4)	0.4 (0.4)
(Max)	3.7 (3.7)	3.7 (3.7)	2.7 (2.7)
Categorical Variables (%)			
Employment status of the individual (reference: not working)			
Not working	62%	32.6%	88.3%
Employed	22.7%	37.7%	9.2%
Self-employed	15.3%	29.7%	2.5%
Presence of other smokers in the household			
Yes	37.3%	36.9%	37.8%
No	64.7%	63.1%	62.2%
Urban/Rural Residence			
Urban	50.8%	50.8%	50.7%
Rural	49.2%	49.2%	49.3%
Age groups (reference: aged 15-24)			
Aged 15-24	13.6%	12.3%	14.8%
Aged 24-44	44%	43.7%	44.3%
Aged 45-64	30.1%	31.9%	28.5%
Aged 65 and over	12.3%	12.1%	12.4%
Has the individual noticed pro-cigarette marketing in the mass media during the last 30 days?			
Yes	11.7%	15.6%	8.1%
No	88.3%	84.4%	91.9%
Knowledge about the health effects of smoking			
Yes	97%	97.5%	96.7%
No	3%	2.5%	3.3%

***Turkish Lira (TL) values are converted to British Pound (£) values using the 2008 (November) exchange rate which is obtained by Central Bank of Republic of Turkey.

**Numbers in brackets present the descriptive statistics for smokers.

Table A3.2: Prevalence Rates of Smoking and the Average Number of Cigarettes Smoked per day by Selected Socio-Demographic Characteristics

Characteristics	Smoking Status			Average number of cigarettes smoked per day*
	Daily	Occasional	Current non-smoker	
Overall	26.4%	3.5%	70.1%	17.5
Age (years)				
15-24	21.0%	3.6%	75.4%	14.9
25-44	33.3%	4.7%	62.0%	17.0
45-64	25.8%	2.3%	71.9%	19.4
65 and over	9.3%	1.8%	88.9%	18.1
Education**				
Primary or less	23.0%	3.1%	73.9%	18.2
Secondary	29.7%	4.4%	65.9%	16.7
Higher	34.6%	4.3%	61.1%	16.6
Employment				
Not working	14.9%	3.2%	81.9%	15.6
Employed	43.9%	4.5%	51.6%	17.6
Self-employed	47.0%	3.4%	49.6%	20.0
Residence				
Urban	29.1%	4.1%	66.8%	16.5
Rural	23.7%	2.8%	73.5%	18.8
Asset index				
Quintile 1	20.5%	3.1%	76.4%	19.2
Quintile 2	28.6%	4.5%	66.9%	17.2
Quintile 3	27.6%	3.1%	69.3%	17.2
Quintile 4	29.1%	3.4%	67.5%	15.9
Quintile 5	26.7%	3.3%	70.0%	16.4
Males	43.9%	3.8%	52.3%	19.1
Age (years)				
15-24	35.4%	4.5%	60.1%	16.9
25-44	54.4%	4.6%	41.0%	18.8
45-64	42.5%	2.4%	55.1%	20.5
65 and over	18.7%	3.1%	78.2%	18.5
Education				
Primary or less	45.0%	3.2%	51.8%	20.0
Secondary	41.1%	5.7%	53.2%	17.8
Higher	43.5%	3.7%	52.8%	17.9
Employment				
Not working	31.8%	3.4%	64.8%	19.4
Employed	49.7%	4.4%	45.9%	18.2
Self-employed	50.0%	3.2%	46.8%	20.1
Residence				
Urban	43.9%	3.8%	52.3%	18.3
Rural	44.0%	3.7%	52.3%	19.9

Table A3.2 continued: Prevalence Rates of Smoking and the Average Number of Cigarettes Smoked per day by Selected Socio-Demographic Characteristics

Asset index				
Quintile 1	41.3%	4.4%	54.3%	20.4
Quintile 2	47.9%	4.4%	47.7%	18.3
Quintile 3	48.0%	3.4%	48.6%	18.8
Quintile 4	46.1%	3.3%	50.6%	17.4
Quintile 5	39.0%	3.4%	57.6%	18.0
Females	10.7%	3.3%	86.0%	11.7
Age (years)				
15-24	10.4%	2.8%	86.8%	9.7
25-44	14.6%	4.8%	80.6%	11.1
45-64	8.9%	2.3%	88.8%	14.2
65 and over	1.2%	0.7%	98.1%	13.8
Education				
Primary or less	7.7%	3.0%	89.3%	10.9
Secondary	14.1%	2.5%	83.4%	12.3
Higher	21.6%	5.1%	73.3%	12.6
Employment				
Not working	9.3%	3.0%	87.7%	11.3
Employed	22.5%	5.0%	72.5%	12.7
Self-employed	14.5%	5.1%	80.4%	14.9
Residence				
Urban	15.8%	4.4%	79.8%	12.0
Rural	5.4%	2.1%	92.5%	10.9
Asset index				
Quintile 1	4.9%	2.2%	92.9%	11.0
Quintile 2	11.2%	4.7%	84.1%	13.0
Quintile 3	10.0%	2.7%	87.3%	10.2
Quintile 4	13.3%	3.6%	83.1%	11.0
Quintile 5	14.2%	3.3%	82.5%	11.7

Source: Turkey Global Adult Tobacco Survey (GATS), 2008.

*****Average number of cigarettes smoked per day is reported by daily smokers and average values are calculated by only including individuals who have positive cigarette consumption.

**The Education category indicates the highest level of education status of the individual which consists of three categories: primary and less (including elementary school), secondary education (including any secondary level of education and vocational schools) and higher education (including any post-secondary education).

Table A3.3: Estimation Results for the Negative Binomial Model

(*Dependent Variable: The number of cigarettes smoked per day*)

Variables	All individuals		Males		Females	
	Coef.	St.Err.	Coef.	St.Err.	Coef.	St.Err.
Male	2.208***	0.079	-	-	-	-
Age 25-44	0.519***	0.100	0.592***	0.116	0.428**	0.211
Age 45-64	0.381***	0.106	0.445***	0.121	0.090	0.231
Age 65 and over	-0.546***	0.141	-0.522***	0.156	-1.382***	0.349
Years of education	0.062***	0.009	-0.012	0.011	0.114***	0.022
Employed	0.401***	0.088	0.308***	0.096	0.590**	0.235
Self-employed	0.554***	0.099	0.299***	0.095	0.886**	0.440
Ln(price)	-1.256***	0.134	-0.495***	0.142	-1.963***	0.348
Quintile 2	0.162	0.106	0.132	0.120	-0.038	0.251
Quintile 3	0.157	0.109	0.292**	0.121	-0.453*	0.262
Quintile 4	0.263**	0.116	0.239*	0.129	-0.081	0.276
Quintile 5	0.388***	0.109	0.149	0.120	0.371	0.252
Other smokers	1.943***	0.066	1.503***	0.073	2.519***	0.150
Media	0.126	0.143	0.139	0.151	0.209	0.353
Smoking attitude	-0.279	0.187	-0.417*	0.223	0.136	0.460
Urban	0.409***	0.071	-0.087	0.077	1.120***	0.166
Log-likelihood	-14171.67		-10556.58		-3265.87	
Pseudo R Squared	0.0583		0.0239		0.0626	
LR chi2	1754.71 (16)		516.36 (15)		436.38 (15)	
N	9030		4269		4761	

Notes: 1)***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the age groups is ‘age 15-24’, the omitted category for the employment status of the individual is ‘not employed’ and the omitted category for the asset index quintiles is ‘quintile 1’ (poorest quintile).

Table A3.4: Marginal Effects for the Negative Binomial Model

(*Dependent Variable: The number of cigarettes smoked per day*)

Variables	All individuals		Males		Females	
	Marg.Eff.	St.Err.	Marg.Eff.	St.Err.	Marg.Eff.	St.Err.
Male	9.291***	0.995	-	-	-	-
Age 25-44	3.305***	0.836	5.141***	1.360	0.647*	0.403
Age 45-64	2.678***	0.910	4.242***	1.448	0.146	0.390
Age 65 and over	-2.720***	0.556	-3.540***	0.825	-1.192***	0.212
Years of education	0.396***	0.070	-0.103	0.095	0.180***	0.048
Employed	2.660***	0.734	2.720***	0.995	1.088*	0.602
Self-employed	4.099***	0.983	2.704***	1.006	2.174	1.657
Ln(price)	-7.990***	1.069	-4.219***	1.259	-3.075***	0.739
Quintile 2	1.087	0.775	1.177	1.140	-0.059	0.382
Quintile 3	1.056	0.796	2.740**	1.329	-0.604**	0.298
Quintile 4	1.838**	0.928	2.218	1.353	-0.124	0.406
Quintile 5	2.708***	0.937	1.323	1.159	0.608	0.498
Other smokers	11.699***	0.128	12.717***	1.306	3.189***	0.651
Media	0.849	1.032	1.262	1.475	0.360	0.679
Smoking attitude	-2.037*	1.193	-4.353**	1.887	0.200	0.722
Urban	2.526***	0.572	-0.748	0.629	1.393***	0.392

Notes: 1)***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the age groups is ‘age 15-24’, the omitted category for the employment status of the individual is ‘not employed’ and the omitted category for the asset index quintiles is ‘quintile 1’ (poorest quintile).

Table A3.5: Estimation Results for the Zero-Inflated Negative Binomial Model for All Individuals

Variables	Logit *Decision not to smoke*[1]		Negative Binomial *Number of cigarettes smoked per day*		Marginal Effects	
	Estimate	St. Err.	Estimate	St. Err.	Estimate	St. Err.
Male	-2.090***	0.082	0.497***	0.035	5.896***	0.313
Age 25-44	-0.920***	0.100	0.124***	0.041	2.350***	0.327
Age 45-64	-0.622***	0.107	0.237***	0.045	2.442***	0.378
Age 65 and over	0.325**	0.159	0.109	0.072	-0.137	0.443
Years of education	-0.030***	0.009	0.002	0.003	0.067***	0.023
Employed	-0.537***	0.085	0.027	0.033	1.178***	0.249
Self-employed	-0.720***	0.093	0.062*	0.035	1.760***	0.289
Ln(price)	1.895***	0.190	-0.079	0.048	-3.913***	0.411
Quintile 2	-0.147	0.106	-0.013	0.042	0.217	0.280
Quintile 3	-0.318***	0.108	-0.048	0.042	0.374	0.288
Quintile 4	-0.314***	0.113	-0.081*	0.045	0.213	0.295
Quintile 5	-0.185*	0.105	-0.035	0.043	0.183	0.279
Other smokers	-2.455***	0.069	0.260***	0.027	6.736***	0.254
Media	0.068	0.133	0.103**	0.052	0.348	0.372
Smoking attitude	0.097	0.191	-0.200***	0.072	-1.214**	0.482
Urban	-0.281***	0.067	-0.013	0.027	0.465***	0.181
Ln α	-1.279	z:-34.36***				
α (dispersion)	0.278					
Log-likelihood	-11794.6					
Likelihood-ratio test of α=0: chibar2(01) = 6524.13***						
Vuong test of ZINB versus standard NB: z = 36.65***						
(cigarette >0)	25.7%					
LR chi2(16)	378.78					
N	9030					

Notes: 1) This model predicts the outcome of zero observations and thus reported signs for the estimates in the logit part of the model relate to the probability of choosing not to smoke.

2)***p<0.01, **p<0.05, *p<0.1

3) Omitted category for the age groups is 'age 15-24', the omitted category for the employment status of the individual is 'not employed' and the omitted category for the asset index quintiles is 'quintile 1' (poorest quintile).

Table A3.6: Estimation Results for the Zero-Inflated Negative Binomial Model for Males

Variables	Logit *Decision not to smoke*[1]		Negative Binomial *Number of cigarettes smoked per day*		Marginal Effects	
	Estimate	St. Err.	Estimate	St. Err.	Estimate	St. Err.
Age 25-44	-1.237***	0.133	0.123***	0.046	4.770***	0.661
Age 45-64	0.850***	0.135	0.182***	0.049	4.166***	0.703
Age 65 and over	0.253	0.182	0.025	0.074	-0.541	0.784
Years of education	0.024**	0.011	-0.006*	0.024	-0.123***	0.047
Employed	-0.458***	0.105	-0.024*	0.036	1.135**	0.446
Self-employed	-0.463***	0.106	0.004	0.036	1.402***	0.461
Ln(price)	0.973***	0.218	-0.030	0.051	-3.080***	0.755
Quintile 2	-0.154	0.130	-0.045	0.043	0.077	0.514
Quintile 3	-0.405***	0.132	-0.031	0.044	0.918*	0.548
Quintile 4	-0.325***	0.138	-0.085*	0.047	0.224	0.548
Quintile 5	-0.109	0.128	-0.034	0.045	0.034	0.519
Other smokers	-2.527***	0.082	0.246***	0.027	11.683***	0.456
Media	-0.050	0.161	0.112**	0.054	1.128	0.728
Smoking attitude	0.201	0.241	-0.205***	0.075	-2.569***	0.926
Urban	0.057	0.083	-0.038	0.028	-0.481	0.327
Ln α	-1.393	z:-33.69***				
α (dispersion)	0.248					
Log-likelihood	-8832.303					
Likelihood-ratio test of α=0: chibar2(01) = 5134.54***						
Vuong test of ZINB versus standard NB: z = 34.53***						
(cigarette >0)	42.6%					
LR chi2(15)	131.54					
N	4269					

Notes: 1) This model predicts the outcome of zero observations and thus reported signs for the estimates in the logit part of the model relate to the probability of choosing not to smoke.

2)***p<0.01, **p<0.05, *p<0.1

3) Omitted category for the age groups is 'age 15-24', the omitted category for the employment status of the individual is 'not employed' and the omitted category for the asset index quintiles is 'quintile 1' (poorest quintile).

Table A3.7: Estimation Results for the Zero-Inflated Negative Binomial Model for Females

Variables	Logit *Decision not to smoke*[1]		Negative Binomial *Number of cigarettes smoked per day*		Marginal Effects	
	Estimate	St. Err.	Estimate	St. Err.	Estimate	St. Err.
Age 25-44	-0.590***	0.157	0.143	0.095	0.657***	0.212
Age 45-64	-0.379**	0.182	0.407***	0.108	0.965***	0.309
Age 65 and over	0.774*	0.427	0.396	0.280	-0.165	0.447
Years of education	-0.111***	0.015	0.027***	0.009	0.123***	0.017
Employed	-0.408**	0.160	0.115	0.084	0.527**	0.215
Self-employed	-0.709**	0.319	0.287*	0.173	1.268**	0.611
Ln(price)	4.367***	0.433	-0.344***	0.132	-3.923***	0.377
Quintile 2	-0.101	0.211	0.059	0.129	0.159	0.251
Quintile 3	-0.189	0.216	-0.186	0.136	-0.076	0.222
Quintile 4	-0.238	0.222	-0.098	0.134	0.068	0.246
Quintile 5	-0.382*	0.207	-0.035	0.128	0.270	0.253
Other smokers	-2.269***	0.130	0.321***	0.086	2.221***	0.224
Media	0.246	0.255	0.034	0.150	-0.150	0.247
Smoking attitude	-0.109	0.381	-0.118	0.222	-0.057	0.406
Urban	-0.872***	0.127	0.098	0.079	0.779***	0.166
Ln α	-0.937	z:-10.80***				
α (dispersion)	0.391					
Log-likelihood	-2822.803					
Likelihood-ratio test of α=0: chibar2(01) = 1238.84***						
Vuong test of ZINB versus standard NB: z = 13.92***						
(cigarette >0)	10.5%					
LR chi2(15)	57.38					
N	4761					

Notes: 1) This model predicts the outcome of zero observations and thus reported signs for the estimates in the logit part of the model relate to the probability of choosing not to smoke.

2)***p<0.01, **p<0.05, *p<0.1

3) Omitted category for the age groups is 'age 15-24', the omitted category for the employment status of the individual is 'not employed' and the omitted category for the asset index quintiles is 'quintile 1' (poorest quintile).

Table A3.8: Estimation Results for the Two-Part/Hurdle Count Data Model for All Individuals

Variables	Logit *Decision to smoke*		Negative Binomial *Number of cigarettes smoked per day*	
	Estimate	St. Err.	Estimate	St. Err.
Male	2.094***	0.082	0.497***	0.035
Age 25-44	0.921***	0.100	0.125***	0.041
Age 45-64	0.625***	0.107	0.238***	0.045
Age 65 and over	-0.323**	0.158	0.110	0.072
Years of education	0.030***	0.008	0.002	0.003
Employed	0.536***	0.085	0.027	0.033
Self-employed	0.720***	0.093	0.062*	0.035
Ln(price)	-1.895***	0.190	-0.078*	0.048
Quintile 2	0.146	0.106	-0.012	0.042
Quintile 3	0.317***	0.107	-0.048	0.042
Quintile 4	0.312***	0.113	-0.081*	0.045
Quintile 5	0.184*	0.105	-0.035	0.043
Other smokers	2.455***	0.069	0.260***	0.027
Media	-0.066	0.133	0.103**	0.052
Smoking attitude	-0.099	0.191	-0.200***	0.072
Urban	0.280***	0.066	-0.014	0.027
Log-likelihood	-11794.463			
Wald chi2(16)	1848.3			
N	9030			

Notes: 1)***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the age groups is ‘age 15-24’, the omitted category for the employment status of the individual is ‘not employed’ and the omitted category for the asset index quintiles is ‘quintile 1’ (poorest quintile).

Table A3.9: Estimation Results for the Two-Part/Hurdle Count Data Model for Males

Variables	Logit *Decision to smoke*		Negative Binomial *Number of cigarettes smoked per day*	
	Estimate	St. Err.	Estimate	St. Err.
Age 25-44	1.237***	0.132	0.123***	0.046
Age 45-64	-0.850***	0.135	0.182***	0.049
Age 65 and over	-0.253	0.182	0.025	0.074
Years of education	-0.024**	0.011	-0.006	0.004
Employed	0.457***	0.105	-0.024*	0.036
Self-employed	0.463***	0.106	0.004	0.036
Ln(price)	-0.972***	0.218	-0.030	0.051
Quintile 2	0.153	0.130	-0.044	0.043
Quintile 3	0.404***	0.132	-0.031	0.044
Quintile 4	0.324**	0.138	-0.085*	0.047
Quintile 5	0.108	0.128	-0.034	0.045
Other smokers	2.525***	0.082	0.246***	0.027
Media	0.051	0.161	0.112**	0.054
Smoking attitude	-0.203	0.241	-0.205***	0.075
Urban	-0.057	0.083	-0.038	0.028
Log-likelihood	-8832.2882			
Wald chi2(15)	1038.60			
N	4269			

Notes: 1)***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the age groups is 'age 15-24', the omitted category for the employment status of the individual is 'not employed' and the omitted category for the asset index quintiles is 'quintile 1' (poorest quintile).

Table A3.10: Estimation Results for the Two-Part/Hurdle Count Data Model for Females

Variables	Logit *Decision to smoke*		Negative Binomial *Number of cigarettes smoked per day*	
	Estimate	St. Err.	Estimate	St. Err.
Age 25-44	0.595***	0.156	0.144	0.095
Age 45-64	0.394**	0.181	0.409***	0.108
Age 65 and over	-0.758*	0.426	0.402	0.280
Years of education	0.112***	0.015	0.027***	0.009
Employed	0.409***	0.159	0.115	0.084
Self-employed	0.716**	0.318	0.287*	0.173
Ln(price)	-4.371***	0.428	-0.343***	0.133
Quintile 2	0.103	0.211	0.061	0.129
Quintile 3	0.183	0.216	-0.184	0.137
Quintile 4	0.234	0.221	-0.097	0.135
Quintile 5	0.379*	0.206	-0.033	0.128
Other smokers	2.275***	0.129	0.322***	0.086
Media	-0.244	0.254	0.034	0.150
Smoking attitude	0.104	0.380	-0.120	0.223
Urban	0.874***	0.126	0.097	0.079
Log-likelihood	-2822.5549			
Wald chi2(15)	549.63			
N	4761			

Notes: 1)***p<0.01, **p<0.05, *p<0.1

2) Omitted category for the age groups is 'age 15-24', the omitted category for the employment status of the individual is 'not employed' and the omitted category for the asset index quintiles is 'quintile 1' (poorest quintile).

CHAPTER 4: GENDER DIFFERENCES IN EDUCATIONAL ATTAINMENT AT THE PRIMARY AND SECONDARY EDUCATION LEVELS IN TURKEY

4.1 INTRODUCTION

The important role of education in both economic growth and development by enhancing the technical and allocative efficiencies of economic actors has been emphasised in the existing literature (Schultz, 1963; Wolfe and Behrman, 1984; King and Lillard, 1983). In addition to economic development, education also has an important role for social development. There is growing attention paid by policy makers to improvements in the education of girls, in particular, which has additional non-market benefits on family welfare through improved child nutrition, fertility reduction and lower levels of infant mortality (Strauss and Thomas, 1995). Therefore, in general, increasing education levels has been one of the major goals of development programs and, in particular, gender gap in educational attainment has been accepted as a main impediment to economic and social development in developing countries by international development agencies (United Nations Development Program (UNDP), 1995).

In many developing countries including Turkey, the gender of the child is potentially important since girls are less likely to be literate or to attend school as compared to boys (European Commission (EC), 2010). Although, at the primary education level, there is a high enrolment rate for boys and girls, boys have consistently higher enrolment rates than girls in Turkey. In the 1997/98 academic year, when the primary schools (five years of education) were combined with the middle schools (three years of education) and eight years of compulsory education were put into place, the total enrolment rate for primary education was 84.7% overall, 90.2% for boys and 78.9% for girls. This reveals that a significant number of school-aged children, mostly girls, were not enrolled in primary school despite the compulsory education law. However, the primary school enrolment rates increased to 96.5% overall, 97% for boys and 96% for girls in the 2008/2009 academic year (National Education Statistics, 2009). This indicates that the education reforms have served to close the gender gap significantly in the primary education enrolment rates. These figures, however, do not reflect dropping out from primary education since these rates relate to current enrolment at primary school rather

than the number of individuals who completed primary education. According to the findings of a qualitative study on drop-outs in Turkey's basic education, dropping out of school is a common phenomenon in Turkey (nearly 13% of children dropped out of primary education when they were at the 5th grade in the academic year 2000-2001) and the dropout rate for girls is about 10% higher than that of boys (Goksen *et al*., 2006). The issue of early school leaving particularly in rural areas is also emphasised in the European Commission Regular Report on Turkey (2010) as an important problem in the education system in Turkey.

In the case of secondary education, there is a higher gender gap as compared to primary education, with net enrolment rates of 60.3% for boys and 56.3% for girls in the 2008/2009 academic year (National Education Statistics, 2009). In the EC report (2010) it is also stated that the gender inequality in the secondary and higher levels of education persists and ensuring gender equality particularly at the secondary education is an important challenge for Turkey in order to achieve the goal of EU accession. Furthermore, there is a significant difference in the primary and secondary school enrolment rates between the eastern and western regions of Turkey. Gender inequality is also more apparent in rural compared to urban areas and the female illiteracy rate is highest in the Eastern and Southeastern regions of the country. The female illiteracy rate is 25.1% in the Southeastern Anatolia region while this rate is 7.2% in the Marmara region, which is the most developed region of the country (TurkStat, 2010). In this respect, enhancing educational attainment, especially in the socio-economically disadvantaged parts of the country, is a crucial instrument in reducing inequalities between the regions.

In light of the importance of investing in education, an understanding of the factors underlying household decisions regarding the education of children, and particularly decisions about the education of girls, may be useful for policy makers in adopting strategies to increase the education levels of society and to close the gender gap in educational attainment (Glick and Sahn, 2000; Holmes, 2003). In this respect, the aim of this chapter is to investigate the determinants of the educational attainment of girls and boys at the primary and secondary education levels in Turkey using the Household Budget Survey (HBS) for 2003, which differs from the other HBSs by providing information on the residence of the household at the regional and province level, focusing on the roles of parental, household and community characteristics. It should be

noted that the results of this chapter provide information on the associations between educational attainment and potential factors at the primary and secondary education levels in 2003 when there is a greater gender gap in schooling as compared to the years after 2003 since the gender gap has consistently been closing after the compulsory education law implemented in 1997. In this respect, the results of the analysis are also important in terms of providing information on the determinants of primary and secondary school educational attainment after the policy change in 1997, which may be useful to evaluate the changes in educational attainment after the compulsory education law. Although there is a vast literature on educational attainment, only a limited number of studies have focused on the determinants of educational attainment in Turkey. However, Turkey is an interesting country to study for two main reasons. First, there have been many improvements in the education system and educational policies in the country, which are discussed in Section 4.2 below, due to its political desire to increase educational attainment and establish gender equality in education. Second, Turkey is a developing country and it may be useful to investigate changes in education over time in the context of a developing country to design future policies both in developing and underdeveloped countries.

Prior to investigating gender differences in educational attainment, one should first consider the potential factors that may cause this inequality. The gender based differences in educational attainment can be explained by several factors, some of which are specific to Turkey. First, as Tansel (2005) states, females in Turkey face discrimination in the private sector in terms of access to employment or in earnings where the returns to education are higher than in the public sector. Furthermore, the labour force participation rate of males is much higher than the rate for females in Turkey (70.8% for males and 27.6% for females in 2010) (Turkey's Statistical Yearbook, 2010). This may lead to lower investment in girls' education, as investing in the education of boys seems to be more rational due to higher employment opportunities for boys. Second, especially in rural areas, the opportunity cost of girls' education may be greater than that for boys since females are often required to perform household tasks and to care for younger siblings (Glick and Sahn, 2000). Third, parents may predict that the expected return to education of girls may be small relative to the expected return to education of boys, as girls join their spouse's family by marriage while boys are more likely to provide help for their parents in old age (Holmes, 2003). Finally, there may be

considerable returns to girls' education in terms of non-market benefits, such as nutrition and health, but parents may not be aware of these benefits or may value them less than monetary benefits (King and Hill, 1993; Glick and Sahn, 2000).

In this context, numerous econometric studies in the existing literature have examined the educational attainment of girls and boys separately to identify potential factors that cause the significantly lower level of attainment for girls (Deolalikar, 1993; Binder, 1998; Tansel, 2002). It should be noted that there are important econometric issues which should be considered in educational attainment analyses to obtain unbiased and consistent estimation results. One issue relates to the right-censored observations of enrolled students. For children who were enrolled in school at the time of the survey, final grade attainment is unknown and treating their education level as identical to those who have completed their education at that level may potentially bias the estimation results of the educational attainment model (King and Lillard, 1983; Tansel, 2002). The second issue relates to the intrafamily correlation of educational attainment among children from the same household. Education outcomes among siblings from the same households can be correlated because of common unobserved family and household characteristics (King and Lillard, 1983; Lillard and King, 1984). The third issue relates to treating the dependent variable, final educational attainment, as a continuous variable and using the OLS method (see, for example, Behrman and Wolfe, 1987; Handa, 1996; Case and Deaton, 1996). In fact, the schooling choice should be treated as a discrete variable but the OLS method does not take into account the discreteness of the data (King and Lillard, 1983; Holmes, 2003).

The empirical methodology used in this chapter, which is discussed in Section 4.4.2 below, effectively deals with these three issues. The ordered probit model, which treats final grade attainment as a discrete variable, is estimated on different samples restricted to individuals above graduation ages from different levels of education to deal with censoring in the data. Furthermore, the ordered probit model is extended to allow a random household specific component in the error term to circumvent the problem that arises because siblings coming from the same household within the sample might share the same unobserved family characteristics which causes the error terms to be correlated for these individuals (King and Lillard, 1983; Hisarciklilar *et al*., 2010). Although most of the recent studies have considered censoring in the data, there are only a few studies which have accounted for common family and household characteristics (see, for

example, Lillard and Willis, 1994; Glick and Sahn, 2000). Therefore, this chapter extends the existing literature by considering both censoring and common family characteristics, which are key econometric issues in modelling educational attainment, for Turkey. Furthermore, to my knowledge, this study is the first attempt to use the HBS for the year 2003 in an educational attainment analysis. The 2003 HBS differs from the other HBSs by providing detailed information on the residence of the household and, thus, enables the exploration of the association between community characteristics and educational attainment.

The rest of the chapter is structured as follows; Section 4.2 provides information on the structure and recent developments in the education system in Turkey. Section 4.3 reviews the background literature focusing on the econometric methodologies of the studies and discusses the contributions and shortcomings of the studies on gender differences in educational attainment in Turkey. Section 4.4 describes the data and the variables used in the analysis as well as the key descriptive statistics and the estimation methods employed in this chapter. Section 4.5 discusses the results and summarises the key findings and Section 4.6 discusses the main findings and policy implications, the shortcomings of the analysis and potential directions for future research.

4.2 EDUCATION AND THE EDUCATION SYSTEM IN TURKEY

4.2.1 Structure of the Education System

The education system in Turkey was formed by the Basic Law of National Education (Law No. 1739) which was put into place in 1973. This law arranges formation of the education system in two main parts: formal and non-formal education. The formal education system is defined as the regular education which covers individuals in a certain age group and has four categories: pre-school education, primary education, secondary education and higher education. Non-formal education includes education, training, guidance and applied activities outside the formal education system and aims to teach individuals how to read and write for those who have never been within the formal education system and to provide or improve the knowledge and skills for those who have left their studies at any stage or who are currently at a particular stage of their education (Ministry of Education, 2010).

4.2.1.1 Pre-school Education

Pre-school education includes the education of children in the age group of 3 to 5 who are not eligible for primary education. This level of education in Turkey is non-compulsory. The aim of pre-school education is to develop the physical, mental and emotional capacity of children and to ensure that they acquire good habits while they are preparing for primary education. However, the government does not have the necessary capacity to meet the demand for pre-school education (Ministry of Education, 2010; Eurydice, 2010). The enrolment rate of children between 4 and 5 years of age for pre-school education has increased to 39% in the 2009-2010 academic year from 33% in the 2008-2009 academic year (EC, 2010).

4.2.1.2 Primary Education

Compulsory primary education constitutes the first grade of formal education and covers the education of children between 6 and 14 years of age[54]. Prior to 1997, the formal education system in Turkey consisted of primary, middle, high school and tertiary levels of education and it had been compulsory to attend primary school and have at least five years of education since 1980. In August 1997, as part of an

[54] The compulsory education period starts in September of the year that the individual reaches age 6 and ends by the end of the academic year when the individual reaches age 14. However, in the case of not graduating from school, even if the individual reaches the end of the compulsory primary education period, they are offered a maximum of four further years of education (Eurydice, 2010).

educational modernisation program, a law (Law No. 4306) was put into place which expanded the minimum years of basic education to eight years covering the middle schools, which used to take three years to complete.

There is an e-school based system which allows for monitoring children who are at the age of the compulsory primary education, but are not enrolled or have a problem of attendance. If parents do not send their children to school, the public administration authorities are entitled to fine them[55]. In the context of supporting attendance of pupils of compulsory education age at school, the Primary Education Law also stated that children who are at the compulsory primary education age cannot be employed in any business (Eurydice, 2010). However, Table 4.1 below, which presents the net enrolment rates of primary education after 1997 when the compulsory education law was put into place, indicates that there are still children who are at the compulsory primary education age and are not enrolled in primary education.

Table 4.1: Net Enrolment Rates of Primary Education from the 1997/1998 to 2008/2009 academic years

Academic Year	Males	Females	Total
1997/98	90.25%	78.97%	84.74%
1998/99	94.48%	83.79%	89.26%
1999/00	98.41%	88.45%	93.54%
2000/01	99.58%	90.79%	95.28%
2001/02	96.20%	88.45%	92.40%
2002/03	94.49%	87.34%	90.98%
2003/04	93.41%	86.89%	90.21%
2004/05	92.58%	86.63%	89.66%
2005/06	92.29%	87.16%	89.77%
2006/07	92.25%	87.93%	90.13%
2007/08	98.53%	96.14%	97.37%
2008/09	96.99%	95.97%	96.49%

Source: National Education Statistics, 2009.

Table 4.1 indicates that the overall enrolment rate of primary education increased from 84.74% in 1997 to 96.49% in 2009. With regard to gender differences, there is an increase in both boys' and girls' school enrolment rates although the increase in the girls' enrolment rate is higher than the increase in the boys' enrolment rate. Therefore, it

[55] This fine is 100 Turkish Liras (nearly 35 British Pounds) per day the child did not attend the school (Eurydice, 2010). However, this fine is not being adequately enforced. There are some problems in the school record keeping systems and updating the data. In addition, some schools do not have necessary data processing mechanisms and, thus, it is difficult to maintain accurate information in schools (Goksen *et al*., 2006).

can be argued that the compulsory education law has led to an increase in the enrolment rate of primary education. Since the compulsory education law has provided an increase in the enrolment rate of primary education, the Ministry of Education is planning to extend the period of compulsory education from 8 to 12 years to increase the enrolment rate of secondary education.

4.2.1.3 Secondary Education

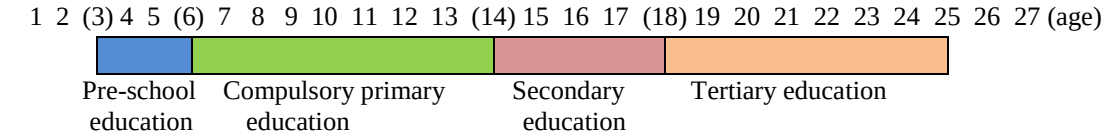
The second grade of the national formal education system involves secondary education. According to the Basic Law of National Education (Law No. 1739), every student who graduated from primary education is also entitled to attend secondary school and benefit from the opportunities of secondary education. The aim of secondary education is stated as to provide students with the general background to prepare them for university education (Eurydice, 2010).

Secondary education includes different kinds of high schools which require a minimum of a four-year education. The types of high schools are as follows: General High Schools, Anatolian High Schools, Science High Schools, Vocational High Schools, Theology High Schools and Technical High Schools. Admission to secondary education institutions is determined in two ways. General high schools, vocational high schools, theology high schools and high schools with multiple curriculums admit students by direct application. Admission to the anatolian, science and private high schools is based on a centralised and nationwide admission exam. In these high schools, in which the language of education is English or another foreign language, there is a one-year preparation class to learn the language and, during the entire high school period, all lessons are taught in that foreign language. Prior to 2005, high school took three years to complete (four years in the case of technical, anatolian and science high schools). Beginning from the 9th grade as of the 2005-2006 academic year, secondary education was expanded from three years to four years (Eurydice, 2010). The age category for secondary education covers ages 14-18[56]. In order to enrol for the first grade of

[56] As compared to the UK education system, primary level education in Turkey corresponds to combined junior and infant school or a first stage covering infant schools and a second stage covering junior schools, which cover six years of education in total from age 5 to age 11. The secondary level education in Turkey corresponds to the General Certificate of Secondary Education (GCSE) or Vocational Certificate of Secondary Education (VCSE) or a combination of both in the UK, which covers education from the age of 11 to the minimum school leaving age of 16. Furthermore, the secondary level education in Turkey also corresponds to the General Certificate of Education Advanced Level (GCE A Level) or the General Certificate of Education Advanced Subsidiary Examinations (GCE AS Examinations), which covers schooling from the age of 16 to 18 in the UK (Euro Education, 2011).

secondary education, an individual must be under the age of 19. However, the high schools might admit individuals who are one year older with the permission of the governor (Eurydice, 2010). Figure 4.1 below presents the organisation of the education system in terms of the education levels and age categories in Turkey.

Figure 4.1: Organisation of the Education System in Turkey, 2009/2010



Source: Eurydice, 2010.

Table 4.2 below presents net enrolment rates for secondary education. This table suggests that overall net enrolment has increased from 37.87% in 1997 to 60.63% in 2009. Similar to primary education, girls' secondary school enrolment rates lag behind boys' enrolment rates in every year. Moreover, the gender gap is higher in secondary education as compared to primary education. In the 2008/2009 academic year, the secondary education enrolment rate has increased to 56.3% for girls and 60.63% for boys.

Table 4.2: Net Enrolment Rates of Secondary Education from the 1997/1998 to 2008/2009 academic years

Academic Year	Males	Females	Total
1997/98	41.39%	34.16%	37.87%
1998/99	42.34%	35.22%	38.87%
1999/00	44.05%	36.52%	40.38%
2000/01	48.49%	39.18%	43.95%
2001/02	53.01%	42.97%	48.11%
2002/03	55.72%	45.16%	50.57%
2003/04	58.01%	48.50%	53.37%
2004/05	59.05%	50.51%	54.87%
2005/06	61.13%	51.95%	56.63%
2006/07	60.71%	52.16%	56.51%
2007/08	61.17%	55.81%	58.56%
2008/09	60.63%	56.30%	58.52%

Source: National Education Statistics, 2009.

4.2.1.4 Tertiary Education

Tertiary education constitutes the third grade of the education system in Turkey and can be split into two groups as the public and foundation (non-profit) tertiary education institutions. Both public and private universities are under the control of the Higher Education Council of Turkey (YOK). There is also an opportunity for distance learning under the structure of the Open University, where students can follow the lectures via media, such as television and radio.

The levels of higher education are as follows: associate degree level (2 years, vocational tertiary education schools), undergraduate level (4 years[57], faculties and tertiary education schools), master's level (2 years) and doctoral level (3-4 years). In Turkey, public universities charge students tuition fees for each semester and the amount of the fees differ on the basis of the type of the program, status, semester and features. Tuition fees are also different in public and private universities. These fees are subject to a legal framework. Private universities determine the tuition fees in the context of the Higher Education Law whereas public universities do not determine the tuition fees since these fees are determined centrally by the Board of Ministers. There are scholarships/grants provided by the government for those who cannot afford university fees. There are also student loans for living costs and tuition fees and financial support for accommodation and subsistence (Eurydice, 2010).

On the other hand, the demand for higher education cannot be completely met by the existing tertiary education system. Therefore, admission to undergraduate education is centralised and based on a very competitive nationwide examination administrated and controlled by the Student Selection and Placement Centre (OSYM). For example, in 2003, 1,451,811 candidates took the university entrance examination and only 311,498 (21.5%) applicants were placed on a university program (Tansel and Bircan, 2006).

Table 4.3 below presents the net enrolment rates for tertiary education. It indicates that there was a significant increase in the enrolment rates in the 2008/2009 academic year. The overall net enrolment rate was 10.25% in 1997 while it was 20.11% in 2009 which is the highest rate since 1997. Furthermore, it can be said that there is a gender gap in tertiary education since the enrolment rate for males is higher than the enrolment rate

[57] However, the faculty of medicine requires education for 6 years, the faculty of dentistry and veterinary requires education for 5 years and colleges training teachers for upper secondary education as affiliated to the faculty of education require education for 5 years (Eurydice, 2010).

for females in every year which is similar to the primary and secondary education levels. The net enrolment rate for tertiary education has reached 29.4% for males whereas this number has reached 25.92% for females in the 2008/2009 academic year.

Table 4.3: Net Enrolment Rates of Tertiary Education from the 1997/1998 to 2008/2009 academic years

Academic Year	Males	Females	Total
1997/98	11.28%	9.17%	10.25%
1998/99	11.81%	9.67%	10.76%
1999/00	12.68%	10.52%	11.62%
2000/01	13.12%	11.38%	12.27%
2001/02	13.75%	12.17%	12.98%
2002/03	15.73%	13.53%	14.65%
2003/04	16.62%	13.93%	15.31%
2004/05	18.03%	15.10%	16.60%
2005/06	20.22%	17.41%	18.85%
2006/07	21.56%	18.66%	20.14%
2007/08	22.37%	19.69%	21.06%
2008/09	29.40%	25.92%	27.69%

Source: National Education Statistics, 2009.

4.2.2 Education Expenditure and the Methods of Financing Education

In Turkey, public sector institutions are the main providers for education services and these institutions are financed mainly by the national budget. Primary and secondary schooling are provided by the government free of charge in the public schools but private primary and secondary schools are also available. All primary and secondary education institutions, public and private, are controlled by the Ministry of Education (Tansel, 2002). Public sector pre-school, primary and secondary education institutions are funded by the public budget and have limited autonomy with regard to financing because of the centralised structure of the Turkish education system. With regard to the public higher institutions, there are three different sources of their financing: the national budget funds, tuition fees and self-generated revenues. However, the private (foundation) universities have complete financing autonomy but they also obtain partial public funding (Eurydice, 2010).

Table 4.4 below shows public education expenditure and it indicates that the share of public education expenditure in total public expenditure has not shown a significant increase since 2005. However, the share of public education expenditure in GDP has increased considerably from 3.21% in 2008 to 5.1% in 2009.

Table 4.4: Public Education Expenditure

Indicator	2005	2006	2007	2008	2009
Public education expenditure/total public expenditure (%)	11.72%	12.47%	12.60%	13.43%	13.37%
Public education expenditure/GDP (%)	2.88%	2.93%	3.05%	3.21%	5.10%

Source: Eurydice, 2010.

4.2.3 Education of Girls and Boys

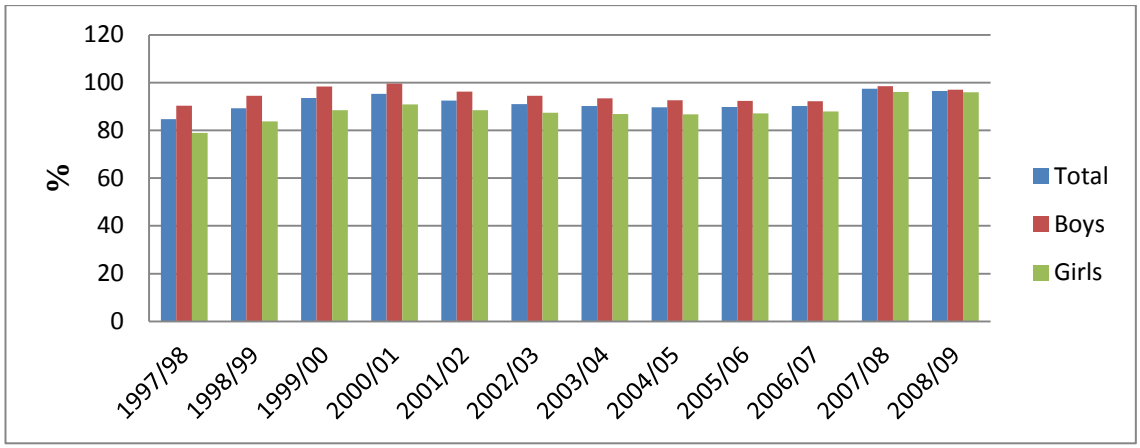
In the EU progress reports for Turkey, it is widely emphasised that the education system in Turkey is in alignment with the system in the EU but there are still significant differences between Turkey and the EU in terms of access to education and the low enrolment rates of girls. It is also stated that improving the provision of education to children from poor families, in particular girls, constitutes the most important problem in the process of the EU membership of Turkey (EC, 2000; EC, 2010).

In this context, there are two important projects in Turkey which aim to increase the educational attainment of girls. The 'Girls to Schools Now' Campaign, which was implemented with the cooperation of the Ministry of Education and UNICEF and has been applied in 81 provinces in 2008, aims to ensure 100% attendance of girls at the ages of the primary education period (ages of 6-14) and to bring in the pupils who left school or have an attendance problem. This project has been supported by a range of institutions such as the Ministry of Internal Affairs, the Ministry of Agriculture and Rural Affairs, the Ministry of Health and the Institution of Social Services and Child Protection. With this campaign, a total of 239,112 girls excluded from the education system have enrolled in primary education from 2003 to 2008 (Eurydice, 2010; Goksen *et al*., 2006).

Another project entitled 'Increasing the Schooling Rate of Girls' differs from the 'Girls to Schools Now' project since it aims to increase the educational attainment of girls at the secondary education level. The main purpose of this project is to reduce the rate of school dropouts for girls, increase the professional quality and skills of the labour force and to increase the awareness of the public in terms of the importance of investing in education. The implementation period of this project is between 2010-2012 (Eurydice, 2010).

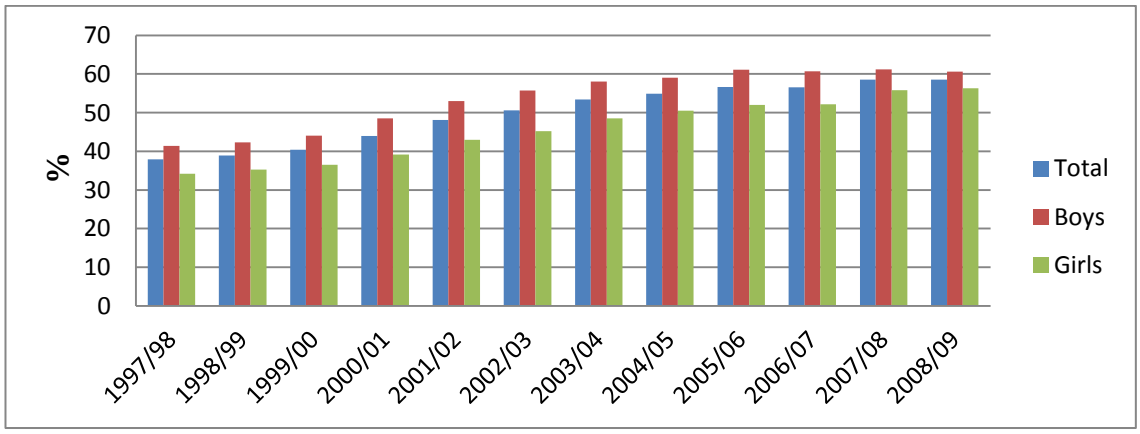
In order to ascertain the extent of the gender gap in educational attainment, it is important to consider both primary and secondary level education. Figure 4.2 below shows the figures of Table 4.1, which presents the net enrolment rates for primary (compulsory) level education after the compulsory education law in 1997. This figure suggests that the extension of compulsory education has had a considerable impact on the total enrolment rate and on the gender inequality in educational attainment. Although high levels of enrolment rates have been achieved at the primary education level for girls and boys, only nearly 60% of the primary school graduates attain secondary education (see Figure 4.3). Enrolment rates at the high school (secondary education) level are low and there are significant gender differences at this level as compared to primary level education. However, it can be argued that the compulsory education law has had a considerable knock-on impact on the secondary school enrolment rate since there is an increasing trend in enrolment rates after 1997.

Figure 4.2: Primary Level Education Net Enrolment Ratios (1997-2009)



Source: National Education Statistics, 2009.

Figure 4.3: Secondary Level Education Net Enrolment Ratios (1997-2009)



Source: National Education Statistics, 2009.

4.3 LITERATURE REVIEW

There have been numerous and diverse econometric studies on educational attainment over recent decades in terms of their research questions, measures of education, and econometric methodologies. In particular, gender-based differentials in educational attainment have received growing attention in the existing literature. This section provides a literature review of this area focusing mainly on the studies for developing countries and the econometric methodologies employed in the studies and the methods used in these studies to overcome various econometric problems. Furthermore, this section provides information on studies on educational attainment in Turkey.

4.3.1 Estimation Issues

4.3.1.1 Measures of Education

An important part of examining the educational outcomes of individuals is to choose a measure of education for the study. There are a range of measures of the education of individuals in the existing literature. The most frequently used measures of education include completed years of schooling, the highest level of completed education, enrolment in a specific level of education, the age-relevant educational attainment and drop outs from school. The measure of education is generally selected according to the system of education in a country, the aims of the study and the available information.

The highest level of completed education is argued to be the most suitable measure of education in the existing literature since it has the potential to reflect the cumulative process of the education of an individual (Maitra, 2003; Hisarciklilar, 2002). In that regard, Sawada and Lokshin (2001), in their study for Pakistan, emphasised that educational attainment levels are stock rather than flow variables and, thus, the current schooling outcome depends both on past and current decisions about the education of individuals. Another commonly used measure of education is ‘current enrolment at school’ (see, for example, Pal, 2004 for India, Smits and Hosgor, 2006 for Turkey). Enrolment models are criticised, however, as they only reflect current decision-making whereas the completed level of education reflects the decisions made over an extended period. On the other hand, some studies, particularly for developing or underdeveloped countries, have used two or more measures of education such as the highest grade attained, enrolment in a specific grade and leaving school in their analyses (see, for example, Chernichovsky, 1985 for rural Botswana; Glick and Sahn, 2000 for Guinea).

In contrast to developed countries, most of the population in developing countries has low levels of education and a small proportion of the population has secondary or higher educational qualifications. Therefore, the studies attempting to explore the determinants of educational outcomes in a developing country context usually use a low cut-off value of education such as current enrolment in a given grade (generally below the secondary school level) for a specific age group or they restrict the analysis to completion of primary education. For example, Keng (2004) used 'ever enrolled' or 'whether dropping out of school before completing grade four' as measures of education for Cambodia. Similarly, Rose and Al-Samarrai (2001) used 'whether enrolled' or 'whether complete primary school' for Ethiopia.

Another estimation issue related to the measure of education should be stated here. Since all measures of education relate to the age of the unit of analysis, determining age restrictions is an important part of investigating the educational attainment of individuals. The age restrictions related to the unit of analysis are determined according to the aims of the study and the structure of the education system in the country. For example, Maitra (2003) included the complete educational age range (6 to 24 years) of individuals in her analysis of Bangladesh. However, Tansel (2002) included individuals aged from 14 to 20 in her analysis for Turkey covering the primary, middle and high school completion ages of the individuals.

4.3.1.2 Dependent Variables and Estimation Methods

After selecting a measure of education and the dependent variable, the next step is to identify the estimation method which largely depends on the dependent variable. When the dependent variable takes the form of years of schooling, school leaving age, highest level of educational attainment or maximum education obtained, OLS is the most common method used to model educational attainment in the existing literature (see, for example, Behrman and Wolfe, 1987; Wolfe and Behrman, 1984 and 1986 for Nicaragua; Chernichovsky, 1985 for rural Botswana; Jamison and Lockheed, 1987 for Nepal; Parish and Willis, 1993 for Taiwan). However, the studies, which use the OLS estimation method, have an important limitation since this approach does not take into account the discreteness of the data. Furthermore, there are generally a high number of zero observations for those who have no educational qualifications in the sample and similar probability spikes exist at primary and secondary education levels, where continuation to the next grade might be delayed because of fees or entrance

examinations. The OLS estimation method is arguably not appropriate due to these issues (Holmes, 2003).

In this respect, the ordered probit model, proposed by King and Lillard (1983 and 1987) and Lillard and King (1984) for modelling educational attainment, is more appropriate than the OLS model when the dependent variable is the highest level of education (or highest grade attained). The probit model is used when the dependent variable is current enrolment at school. Glick and Sahn (2000) examined gender differences in educational attainment using different measures of education (i.e., final educational attainment, current enrolment and withdrawal from school) for Guinea, West Africa. In the first model, they used final grade attainment as the dependent variable and estimated an ordered probit model whereas in the second and third models, the determinants of school enrolment status at the time of the survey and the determinants of leaving school were estimated by using probit models. The main findings of this study include that maternal education and household income are the main determinants of girls' educational attainment and current enrolment but these variables have no effect on boys' education. They also found that the presence of siblings under age 5 is negatively associated with girls' educational attainment and current enrolment and is positively related to the probability of leaving school for girls. In accordance with Glick and Sahn (2000), Kabubo-Mariara and Mwabu (2007) investigated the determinants of school enrolment using a probit model and the highest level of educational attainment for Kenya using an ordered probit model. They found that parental education, child and household characteristics and the quality and cost of schooling are the main factors behind the demand for education in Kenya.

Although there are other studies using alternative estimation methods such as ordered logit models (see, for example, Dreze and Kingdon, 2001 for India) and IV estimation methods (see, for example, Dayioglu *et al.*, 2009 for Turkey), when the dependent variable is the highest level of educational attainment, it can be argued that a relatively small, but growing, number of studies have used an ordered probit model to investigate educational attainment for developing countries (see, for example, Tansel, 2002 for Turkey; Holmes, 2003 for Pakistan, Maitra, 2003 for Bangladesh). The main reason for using the ordered probit model is that it allows for analysing different levels of educational attainment and, hence, comparing the impacts of factors for each educational outcome.

4.3.2 Econometric Issues

It is important to consider some econometric issues while investigating the highest level of educational attainment to obtain unbiased estimation results. These include the censoring problem of the final attainment of enrolled children, intrafamily correlation among siblings and selection bias related to children currently residing in the household.

4.3.2.1 The Censoring Problem

As stated above, educational attainment has generally been investigated using current school enrolment and/or the highest level of education attained in the existing literature. The current school enrolment analysis is generally examined by using a probit model and is somewhat standard. However, estimation of the highest level of educational attainment is more complicated and may lead to some econometric problems. Firstly, children who are enrolled at school at the time of the survey constitute an important problem since, for those children, it is unknown whether they will complete this level of education or drop out of school. Consequently, their highest educational attainment is unknown. Such observations are right censored, which occurs when the survey (used in the analysis) ends before the event (completed education) has occurred, and may lead to biased estimation results[58] (King and Lillard, 1983 and 1987; Behrman and Knowles, 1999; Maitra, 2003). Several approaches have been used to deal with this problem in the existing literature.

The first method is to exclude currently enrolled children (i.e., the censored observations) and to estimate the model only for individuals who have completed their education (Lazear, 1977; DeTray, 1978). Excluding such censored observations, however, may result in sample selection bias since this method may lead to a selection of older individuals and individuals with lower levels of education (Lillard and King, 1984). An alternative to this approach is to restrict the samples to include only older individuals who have completed education, as opposed to excluding currently enrolled children from the sample (Tansel, 1997; Holmes, 2003). Kalmijn (1994), for example, investigated the relationship between mothers' socio-economic status and children's schooling for the US using the National Survey of Families and Households (NSFH) 1987/1988. The author estimated logistic regression models only for individuals who were aged 24 or older at the time of the survey to overcome the censoring problem. This

[58] For a further discussion of the censoring problem, see Glick and Sahn (2000).

method, however, may lead to a significantly reduced sample size. This method may also not be useful for developing countries experiencing constant change in their education systems since an analysis of more recent periods may be more relevant for policy. In addition, the impact of family background may become harder to ascertain since older individuals are less likely to live with their parents and household surveys generally do not provide information on the childhood environment of adults (Holmes, 2003).

Another way to deal with the censoring problem used in the existing literature is to incorporate age and its squared term as regressors in the model (Behrman and Wolfe, 1987). Including the age of the individual as a covariate may account for differences in educational attainment between young and enrolled children and older individuals who have completed education but it is not an effective way to overcome the censoring problem, as it cannot make a distinction between completers and enrolees[59] (Lillard and King, 1984). This method is further complicated especially in underdeveloped countries because of late entry to school and repeating years, which are frequently encountered in these countries, and reduce the power of age as a predictor of educational attainment (Holmes, 2003).

Another method to overcome the censoring problem is to construct an 'age and sex specific education index' which is computed as the ratio of the current educational attainment of child *j* in family *i*, of age *m*, and sex *k* ($S_{ji,mk}$) to the mean observed educational attainment of children of age *m* and sex *k* in the sample (S^*_{mk})[60]. This index has been used as a dependent variable by Rosenweig (1978), Birdsall (1982) and Wolfe and Behrman (1986). However, for enrolled children, this index indicates only whether the child lags behind his/her cohort by, for example, grade repetition, late start to school, temporary leave or whether the child is ahead rather than accounting for the censoring problem (Lillard and King 1984).

There are also other studies which used different methods to address the censoring problem in the existing literature. Chernichovsky (1985) used the OLS approach to

[59] Consider two individuals of the same age, one of them is enrolled and the other one has already dropped out of school. In this case, including age as a covariate is not an effective way to deal with the censoring problem since both children are considered as identical in the estimation (Lillard and King, 1984; Holmes, 2003).

[60] In other words, an 'age and sex specific education index' indicates the ratio of child schooling to the mean schooling of children in the relevant age-sex group.

examine the socio-economic and demographic correlates of school enrolment and educational attainment in rural Botswana using the Rural Income Distribution Survey 1974/1975. The author addressed the censoring problem by examining separately the factors associated with the demand of education for enrolled children and non-enrolled children in the age group of 6-18 years old which includes both primary and secondary school children. Barros and Lam (1996), on the other hand, estimated OLS and 2SLS models to examine the determinants of educational attainment at the household level for 14-year-olds only in urban regions of Sao Paulo, Brazil using the 1982 Brazilian Annual Household Survey. They used this age group since the children in this age group should have completed the compulsory education in Brazil. However as Holmes (2003) argued, it is not clear whether using only one 'schooling cohort' is a good predictor of the completed level of education.

In order to circumvent the censoring problem, another alternative method is proposed by King and Lillard (1983 and 1987) and later used by Glewwe and Jacoby (1992), Behrman *et al*. (1997) and Holmes (2003), which is an ordered probit model that takes into account right censoring explicitly. In this censored ordered probit model, it is assumed that individuals will at least complete their last grade and this assumption is incorporated into the likelihood function (Lillard and Willis, 1994; Glick and Sahn, 2000).

4.3.2.2 Intrafamily Correlations in Educational Attainment

Another econometric issue in the educational attainment literature relates to common family characteristics for children from the same household since household surveys generally collect multiple observations per family. Children belonging to the same household are likely to share unobserved (by the researcher) characteristics, which affect their performance in school and their demand for education in a similar way. In this case, the residual terms are unlikely to be independent and they will be correlated through a common household-level component for children from the same household. Therefore, it can be argued that the highest level of educational attainment of siblings may be correlated as they have common or highly correlated values of regressors (Glick and Sahn, 2000; Lillard and King, 1984). Failure to account for this problem may lead to substantially underestimated standard errors. One way to deal with this problem is to select a single child from a household. For example, Parish and Willis (1993), in their analysis for Taiwan, used a single child from a household because they claimed that

including more than one child from a household may result in over-representation of large families. However, the authors did not state any selection criteria in the case of the existence of multiple eligible children in the household. Moreover, using this method may severely reduce the sample size.

An alternative method is to allow for intrafamily correlation explicitly in the estimation model. In this context, an ordered probit model, which allows for such correlations in the model through random effects, is proposed by King and Lillard (1983 and 1987) and Lillard and King (1984). In other words, the assumption is that the error terms in the index functions for educational attainment in the ordered probit model composed of a common household heterogeneity component and an idiosyncratic individual error (Glick and Sahn, 2000). The random effects ordered probit model, which is an extension of the ordered probit model by allowing for a household random specific component in the error term, has been used in only a few educational attainment studies particularly for underdeveloped or developing countries (see, for example, Lillard and King, 1984 for the Philippines; Glick and Sahn, 2000 for Guinea).

4.3.2.3 The Potential Selection Problem

In general, educational attainment studies use the child as the unit of analysis to explore mainly the association between the parental characteristics and child schooling. The importance of conducting child-specific research is emphasised by Holmes (2003) as follows: "Using children as the unit of observation permits the use of information about current parental, household and community characteristic, and, thus, the environment in which the schooling decisions are made (p.252)". Furthermore, child-specific research is particularly important for developing countries and more relevant to policy since many developing countries have been attempting to restructure their education system and children are affected mostly by these changes.

The studies using children as the unit of observation often, however, include only children who live with their parents because most surveys do not provide information on children who do not reside in their parents' house. Therefore, another econometric issue is the potential selection problem which arises from the fact that children leave the household of their parents after a certain age and those who are observed still live in the household (i.e., home-resident children) may be an unrepresentative and non-random sample (Tansel, 2002). It should be acknowledged that there may be a close association

between leaving home and educational attainment. For example, if the least capable or least supported children in the household leave home at an early age and are less likely to obtain higher level of education, then the correlation between the error terms in any leaving home and educational attainment equations may result in sample selection bias when educational attainment is estimated only for children living with their parents (Holmes, 2003). There is only limited discussion of this potential selection problem in the existing literature. While some studies include information on all living children and have estimated the models for all children and, thus, do not face this problem (see, for example, Glewwe and Jacoby, 1992 for Ghana; Tansel, 1997 for Cote d'Ivoire and Ghana), a majority of the educational attainment studies do not state the probability of this form of bias probably due to the unavailability of information on children not living with their parents (see, for example, Birdsall, 1980 for Columbia; Handa, 1996 for Jamaica). However, Holmes (2003) investigated the determinants of educational attainment in Pakistan by specifically examining two potential sources of bias, the censoring bias and sample selection bias. She estimated censored ordered probit models for all children and for home-resident children only using the 1991 Pakistan Integrated Household Survey. The results of this study indicated that the sample used for the estimation of educational attainment can alter the estimation results and, in particular, samples including only home-resident children lead to a bias in the estimation results of the demand for education. Although this study is one of the few studies, which considers sample selection bias in the analysis, it also has an important limitation since the author did not consider the bias resulting from the intra-family correlations for children from the same household.

4.3.3 Studies on Gender Differences in Educational Attainment in Turkey

The determinants of the demand for education, particularly the gender gap in educational attainment, have been widely examined in the existing literature. This area is particularly important for Turkey since in the process of becoming a member of the EU, increasing the enrolment rates at both the primary and secondary education levels and providing gender equality in access to education are among the most important issues in the negotiations (EC; 2008; EC, 2010). Although understanding the factors causing the low level of educational attainment and the large gender inequality particularly at the secondary education level is useful for policy makers in Turkey, there are only a few studies exploring educational attainment for Turkey.

Tansel (2002) estimated an ordered probit model with separate regressions for boys and girls in order to investigate the main determinants of the gender differences using the Household Income and Consumption Expenditure Survey (HICES) of 1994. In this study, individual and household factors, such as the age of the child, household income, the employment status of parents and parental education, and location characteristics were considered as the main determinants of the demand for education. One of the differences of this study from the other studies for Turkey is that the determinants of educational attainment are examined at different education levels, namely primary, middle and high school levels, in order to provide policy implications for different education levels. Furthermore, this study differs from other studies by emphasising the importance of location characteristics and using different location characteristics, such as the distance to a metropolitan centre and Istanbul and dummy variables for undeveloped streets, developed streets and squatter settlements. In this context, the most important finding of this study is that there is a positive relationship between distance to a regional metropolitan centre and lower educational attainment for both boys and girls at the middle and high school levels while distance to Istanbul is related to lower educational attainment for both girls and boys at the primary education level. This study is arguably the seminal study on educational attainment in Turkey but it has an important limitation. Although the censoring problem was overcome by estimating the model on different samples, which include individuals above graduation ages, the intrafamily correlation of educational attainment among children from the same household, which is discussed in detail in Section 4.3.2.2 above, was not considered in the analysis which may have biased the estimation results.

There are also two more important studies for Turkey which differ from the other studies for Turkey by focusing primarily on gender role attitudes and cultural factors. Smits and Hosgor (2006) investigated the effect of family background on enrolment of children at primary and secondary school using the 1998 Turkish Demographic and Health Survey (DHS). They estimated a multivariate logistic regression model for boys and girls separately. Since the focus of their study is on socio-cultural characteristics, they included variables such as gender role attitudes[61], dependency, which is a dummy variable indicating whether or not the mother was married under the age of 17, and the

[61] The gender role attitude dummy variables were based on whether or not the mother agrees with the questions which elevate the position of the father in the family.

Turkish language proficiency of the mother, which is a dummy variable indicating whether or not the mother is able to speak Turkish and has another language as her mother tongue[62]. The results of the study indicated that household income, the employment status of the father, parental education, the number of siblings and whether or not the mother is able to speak Turkish are the main factors behind school enrolment. They also found that the traditional gender role attitudes of the mother are negatively associated with the secondary school enrolment of girls whereas the economic position of the household is the major determinant for participation of boys. It should be stated that their study has an important limitation in that they estimated enrolment models rather than the completed education level and, thus, their results give information on schooling for one point in time.

Similarly, Rankin and Aytac (2006) investigated the determinants of the gender gap in schooling in Turkey by focusing on cultural attitudes and practices[63] as well as individual, household and macro-structural variables[64]. They estimated multivariate logistic models for girls and boys separately using the 1988 Turkish Family Structure Survey (TFSS). The results of their study indicated that the effects of cultural variables on educational attainment and gender gap are mainly apparent at the post-primary education level whereas cultural attitudes have no relationship with primary education.

Another study by Dayioglu *et al*. (2009) examined the role of the number of siblings, birth order and sibling sex composition in school enrolment in urban Turkey using the 1998 DHS. This study is different from most of the other studies for Turkey in terms of their estimation method. They used an instrumental variable (IV) estimation method in order to overcome the problem of endogeneity of the number of siblings due to the parents' fertility and schooling decisions. Furthermore, the importance of this study for the existing literature on Turkey is that it is one of the very few studies which focuses on sibling composition. In this context, they found that the number of siblings has no effect on school enrolment. Their results also indicated that there is a parabolic pattern for the birth order of children where middle-born children fare worst for nearly all

[62] The Kurdish language is common in Turkey especially in the south-eastern and eastern parts of the country.

[63] For the cultural attitudes, variables such as the father's attitude toward female participation in the labour market and the family's preference toward having son or daughter were used.

[64] As macro-structural variables, they used variables such as region controls, urban/rural residence as well as the proportion of employed individuals in agriculture and the ratio of males to females for all employed individuals at the province level for the year of 1985.

income groups and this parabolic pattern is at its strongest for the poorest households. Similarly, sibling sex composition was found to be more important for poorer households, which indicates that limited financial resources play a crucial role in sibling composition effects.

Hisarciklilar *et al*. (2010) investigated changes in the effects of family background and socio-economic characteristics on the educational attainment of boys and girls using the 1988 and 2006 Household Labour Force Surveys for Turkey. They considered both the censoring problem and the unobserved family effects using a censored ordered probit model incorporating family random effects into the model. However, they estimated only one model for all education outcomes including both primary and secondary education and, hence, the results of their study do not provide information on the determinants of educational attainment separately for primary and secondary education levels. They also used the sample including individuals aged between 15 and 24 and using such a high upper age cut-off may have increased the possibility of incurring sample selection bias. They found that maternal education, household size and financial limitations of the family all have important effects on the educational attainment of girls whereas household size, location characteristics, the occupational status of the father are the main determinants of educational attainment for both genders.

In the case of Turkey, it can be argued that every study examined a different aspect of gender differences in educational attainment by focusing on different potential determinants of educational attainment and using different estimation methods. To my knowledge, there is no study estimating separate samples for primary and secondary education levels while considering both the censoring problem and unobserved common family characteristics for children coming from the same household for Turkey. Furthermore, there is no study for Turkey using the HBS for 2003 which includes a range of individual, household and location characteristics and differs from the other HBSs by providing information on the residence of the household at the regional and province level. In this context, this chapter makes an important contribution to the existing literature by using the most comprehensive HBS and by considering both the censoring problem and unobserved family random effects in the estimation method while estimating the models for primary and secondary education levels separately.

4.4 DATA AND METHODOLOGY

4.4.1 Data

The Household Budget Survey (HBS) for 2003, which is the most comprehensive household survey administered by the Turkish Statistical Institute, is used in this chapter. The survey was administered to 25,764 households, 107,614 individuals, 52,552 of whom are males and 55,062 are females, around the country. This survey is particularly suited to the analysis because it contains detailed information on both parents and their children which constitutes a crucial part in examining the effects of household characteristics on educational attainment as discussed in Section 4.3 above and Section 4.4.2.1 below[65]. Another reason for using this data set is that the 2003 HBS differs from the other HBSs for Turkey by providing information on the residence of the household at the regional and province level which enables us to control for community characteristics[66]. The surveyed households were located in 12 regions and 26 provinces of Turkey[67]. This survey is nationally representative of all households with 7,486 households from rural areas and 18,278 households from urban areas.

Since one of the aims of this chapter is to examine the effects of parental characteristics on the educational attainment of their children, information on both the child and parents is needed. Therefore, a new data set where each child in the household is matched with their parents was created. Furthermore, the new sample was limited to include only the children of the household head. In fact, most of the children in the data set were children of the household head and the spouse if present[68]. However, there were also other children, such as grandchildren and other relatives' children, in the household but some crucial variables, such as parental characteristics, were not available for these children. In addition, it may also be case that the household head may consider their children, grandchildren and the children of relatives differently in terms of the investments in their human capital (Tansel, 2002). Therefore, the new

[65] The content of the questionnaire is discussed in detail in Chapter 2, Section 2.4.1.

[66] As discussed in Section 4.4.2.3 below, the importance of community characteristics is widely emphasised in the existing literature (see, for example, Wolfe and Behrman, 1984 for Nicaragua; Glick and Sahn, 2000 for Guiana, Tansel, 2002 for Turkey).

[67] The survey was administrated to 2,842 households in Istanbul, 1,705 households in Western Marmara, 3,906 households in Aegean, 2,008 households in Eastern Marmara, 2,648 households in Western Anatolia, 3,328 households in Mediterranean, 1,688 households in Central Anatolia, 2,501 households in Western Black Sea, 1,169 in Eastern Black Sea, 604 households in Northeast Anatolia, 1054 households in Middle East Anatolia and 2,311 households in Southeast Anatolia.

[68] The percentage of the children of the household head is 87.6% while other children in the household constitute 12.4% of the sample.

sample is restricted to the children of the household head to provide a more homogenous sample in terms of the incentives for investing in their own children's human capital.

In this chapter, the aim is to estimate the determinants of the educational attainment of children at the primary and secondary education levels in Turkey. As stated in Section 4.2 above, in Turkey, children start primary education at the age of 6 and finish their compulsory education when they are 14 years old and finish secondary education (high school) when they are 18 years old in normal circumstances. However, it is common in Turkey to repeat a grade or to start school a year late. Therefore, for the primary (compulsory) level of education, children[69] in the 14-17 age group were analysed and, for the secondary level of education, children in the 18-20 age group were included in order to obtain a representative sample of children who should have finished primary education and secondary education, respectively[70]. The choice of the upper age cut-off for the secondary education level was based on the fact that older children have a greater propensity to leave the household which implies that children who live with their parents may not be a representative sample (Dayioglu *et al*., 2009). Furthermore, older children are more likely to make their own decisions about their educational attainment implying that the decision process may be different than the one investigated in this chapter. Thus, the final sample for the primary education level includes 4,229 observations for girls, 4,109 observations for boys and, for the secondary education level, 2,610 observations for girls and 2,268 observations for boys.

4.4.2 Estimation Methods

The potential parental, household and community factors affecting the primary and secondary school educational attainment of children are analysed by estimating ordered probit and random effects ordered probit models. As stated in Section 4.3 above, in the existing literature, a common way of analysing educational attainment decisions is to apply the OLS model, which is based on the assumption of a continuous distribution for the dependent variable. However, applying a discrete choice model is more convenient since it treats educational attainment as a discrete and non-normally distributed variable.

[69] The individuals could be described as young adults. However, the use of the term 'child' in this study refers to being the parent's child because one of the aims of this chapter is to explore the effects of parental characteristics on their children's educational attainment.

[70] These age restrictions are generally the same as those used in the existing educational attainment studies for Turkey (see, for example, Tansel, 2002; Rankin and Aytac, 2006).

Since the dependent variable is categorical in nature, an ordered probit model, where different levels of education can be analysed and, hence, the effects of factors for each educational transition can be compared, is estimated for the primary and secondary levels of educational attainment, respectively[71]. In addition, the ordered probit model is extended to allow a random household specific component in the error term to deal with the intra-family correlations for children coming from the same household. Furthermore, since the aim of this chapter is to explore gender differences in educational attainment, the models are estimated for girls and boys separately in order to allow their regression coefficients and threshold values to differ.

There are two main reasons for estimating the models for the primary and secondary levels of educational attainment separately. The first one is to overcome the censoring problem, which is discussed in Section 4.3.2.1 above. As Holmes (2003) suggests, one way to deal with this censoring problem is to restrict the samples to include only children above the graduation age. It should be acknowledged, however, that this method is criticised in the existing literature as it leads to losing observations for younger children (Holmes, 2003; Lillard and King, 1984). However, in this chapter, the sample is divided into two groups focusing on primary and secondary school educational attainment according to their ages: 14-17 year-old children for the primary education sample and 18-20 year-old children for the secondary education sample, as discussed in detail in Section 4.4.1 above. Therefore, it can be argued that estimation on two different samples enables us to include children in a wide age range and to minimise the limitations from omitting younger observations. Secondly, estimation on two different samples (i.e., one for primary level education and one for secondary level education) enables us to explore the hypothesis that the potential factors associated with gender inequality in educational attainment differ by the level of education.

In this respect, the dependent variable for the primary education sample takes the value of 0 if the child is illiterate, 1 for children[72] who are literate but who did not graduate

[71] University graduation is not included in the analysis because, as stated in Section 4.2.1.4 above, admission to a university is based on a very competitive centralised exam in Turkey and, thus, it cannot be completely seen as reflecting preferences towards education (Hisarciklilar *et al*., 2010). Furthermore, individuals generally go to another province to study at a university in Turkey and, thus, they are not observed in the sample since the sample provides information only for resident household members. In this case, including university graduation in the model would potentially lead to biased estimation results.

[72] In Turkey, there are also open primary education institutions which target illiterate individuals or individuals with low levels of literacy who did not graduate from primary school and who exceed the age limit designated for the compulsory primary education (i.e., age 14) (Eurydice, 2010).

from any school assuming that literacy is acquired in school and 2 for children who graduated from eight-years of compulsory primary school (i.e., the new compulsory education system) [73]. The dependent variable for the primary education sample does not include the category of five-years of primary education since children aged 17 and below in the data set are subject to the new compulsory education law.

For the secondary education sample, on the other hand, the dependent variable takes the value of 0 for illiterate children, 1 for literate children but who did not graduate from any school, 2 for children who graduated from five-years of primary education, 3 for children who graduated from middle school (or eight-years of primary education) and 4 for children who graduated from high school (i.e., secondary education). In the data set, children aged 19 and over are subject to the old compulsory education law. However, age 18 is a borderline age where some of the children are subject to the new compulsory education law (8 years) whereas some are subject to the old compulsory education law (5 years). Therefore, the dependent variable for the secondary education sample has two different categories for primary school educational attainment (i.e., one for five-years of primary education and one for eight-years of primary education or middle school).

One important point for the methodology used in this chapter for modelling educational attainment is that the models focus on the completed education level of children (i.e., the highest level of educational attainment), which differs from analysing enrolment. As stated above, enrolment models estimate the demand for education at one point in time and, thus, they do not take into account grade repetition or drop-outs[74] while models for the completed education level allow for analysing the effects of the cumulative decision process for the education of children made over an extended period.

4.2.2.1 Ordered Probit and Random Effects Ordered Probit Models

An ordered probit model is a commonly used framework when the outcome of interest is categorical in nature. In this study, the ordered probit model, where different levels of education can be analysed, is estimated as a first step. However, as discussed in Section

[73] The possible survey responses for the question of the highest level of education completed is as follows: illiterate; literate but not graduated from any school; primary school (5 years); primary school (8 years); middle school; vocational school at the middle school level; high school; vocational school at the high school level; university (2-years associate degree); university (4 years-undergraduate); master's and doctorate.

[74] Parents may enrol both their sons and daughters at school, but it may be the case that they support their sons more and give priority to the education of their sons by allocating more resources to them while girls, for example, are required to perform household tasks (Hisarciklilar, 2002).

4.3.2.2 above, one important econometric issue arising in the estimation of the model is that there is potentially more than one child from the same family in the data set. The same unobserved family characteristics such as parental genetic information, parent's supervision ability and the adequacy of child care in the household, which cannot be measured by the data, might be shared by children belonging to the same household and this problem may lead to incorrect standard errors (Lillard and King, 1984; Lillard and Willis, 1994; Hisarciklilar, 2002). The ordered probit models estimated in this chapter allow for unobserved family characteristics by including a random error component which takes the same value for children coming from the same household[75]. In this respect, the random effects ordered probit model is estimated in the second specification.

The ordered probit model takes the following form:

$$y_i = \beta' x_i + u_i \tag{4.1}$$

where y_i is the propensity of schooling for the i^{th} individual, β is a $k \times 1$ parameter vector, x_i is a $k \times 1$ vector of individual characteristics and u_i is the stochastic disturbance term.

In the model, we observe y_i such that:

$$S_i = s \quad if \quad \mu_s \leq y_i \leq \mu_{s+1} \quad for \quad s = 0,1,2,3,4 \tag{4.2}$$

where *s* takes the following form for the primary education sample:

$$s = \begin{cases} 0 \text{ if the individual is illiterate} \\ 1 \text{ if the individual is literate but has no educational qualification} \\ 2 \text{ if the individual has completed (eight years) primary school} \end{cases}$$

and where s takes the following form for the secondary education sample:

$$s = \begin{cases} 0 \text{ if the individual is illiterate} \\ 1 \text{ if the individual is literate but has no educational qualification} \\ 2 \text{ if the individual has completed (five years) primary school} \\ 3 \text{ if the individual has completed (eight years) primary school or middle school} \\ 4 \text{ if the individual has completed secondary school} \end{cases}$$

where the μ's denote the threshold values where $\mu_0 < \mu_1 < \cdots < \mu_5$, $\mu_0 = -\infty$ and $\mu_5 = +\infty$. The conditional probability of observing the s^{th} category is as follows:

[75] Although unobserved characteristics at the household level are accounted for in the analysis, unobserved influences at the individual level, such as ability, motivation and genetics, mental and physical health, cannot be controlled for in the analysis. The existing literature, however, ignores this dimension of the analysis or acknowledges it as a limitation of the analysis.

$$Pr(S_i = s \mid x_i) = Pr(\mu_s \leq \beta' x_i + u_i \leq \mu_{s+1}) \quad (4.3)$$

If a standard normal distribution for the stochastic disturbance term ($u_i \sim N(0,1)$) is assumed, the conditional probabilities are as follows:

$$Pr(S_i = s \mid x_i) = \Phi(\mu_{s+1} - \beta' x_i) - \Phi(\mu_s - \beta' x_i) \quad (4.4)$$

where Φ denotes the cumulative standard normal distribution with $\Phi(-\infty) = 0$ and $\Phi(+\infty) = 1$ (Maddala, 1999; Greene, 2003; Hisarciklilar, 2002).[76]

In the random effects ordered probit framework (Frechette, 2001), the propensity for the educational attainment of the i^{th} individual in the k^{th} household can be written as follows:

$$y_{ki} = \beta' x_{ki} + \theta_k + u_{ki} \quad (4.5)$$

where θ denotes unobserved family characteristics and is assumed to be shared by all children in the household. It is also assumed to be normally distributed with a variance of $\sigma^2_{\theta_k}$ ($\theta_i \sim N\left(0, \sigma^2_{\theta_k}\right)$. Hence, the correlation for the disturbance terms for different family members is as follows:

$$\rho = \sigma^2_\theta / (\sigma^2_\theta + \sigma^2_u) = \sigma^2_\theta / (\sigma^2_\theta + 1) \quad (4.6)$$

The conditional probabilities for the schooling outcomes of the individuals are as follows (Hisarciklilar, 2002).

$$Pr(s_{ki}) = \Phi(\mu_{s+1} - \beta' x_{ki} - \theta_k) - \Phi(\mu_s - \beta' x_{ki} - \theta_k) \quad (4.7)$$

It is possible to argue that this study is one of the few educational attainment studies, in particular, for Turkey, which allows for common unobserved family characteristics for children belonging to the same household. However, it should be stated that the sample used in this chapter includes only children who reside with their parents due to the unavailability of information on all living children whether they reside with their parents or not. As stated in Section 4.3.2.3 above, such information is rare among surveys and only a few studies have exploited it to investigate the selection bias related

[76] The marginal effects for both the ordered probit models and random effects ordered probit models are also calculated. For the marginal effects of a continuous variable, the first derivatives of the probability function with respect to the independent variable are calculated at the sample means of the variables whereas for the marginal effects of a dummy variable, the probabilities in the cases where the dummy variable takes the value of 1 and 0 are compared at the sample means of the variables (see, Greene 2003, p.738-39, for further details of the calculation).

to the exclusion of children not living with their parents (see, for example, Holmes, 2003). However, for this chapter it can be argued that the aim of using a low upper age cut-off, which is 20 for the secondary education sample, may serve to minimise this kind of selection bias[77].

4.4.2 Variable Construction and Definitions

In the context of modelling educational attainment, it should be considered that the educational attainment of children is largely a family decision which is constrained by the family budget and affected by factors that have an impact on the costs and benefits of investing in the education of children (King and Lillard, 1983). The household's decision-making can be represented by a simple theoretical model of investment in the human capital of children. Such a model, which is developed by Alderman and King (1988), is summarised in the Appendix as an example of discussion of the differences in investment in the education of girls and boys. According to this model, the differences in the expected returns to boys' and girls' education, in terms of enhanced future earnings, direct costs (for example, school tuition fees) and indirect costs (i.e., foregone earnings associated with time spent at school or the opportunity cost of children's time), may lead to differences in the educational attainment of boys and girls in the family. In this respect, gender based differences in educational attainment can be explained by parental and household characteristics as well as by community characteristics.

There are three main groups of factors which are commonly used in the existing literature (Gertler and Glewwe, 1990; Strauss and Thomas, 1995). These groups are: parental characteristics, such as parental education and the employment status of parents; household characteristics, such as household income and household composition; and location characteristics such as urban/rural residence and regional controls.

4.4.2.1 Parental Characteristics

Parental Education: The highest level of the educational attainment of parents is included to account for the genetic ability of children, parents' taste for education and their provision of a supportive learning environment. Furthermore, parental education may serve as a measure of household income or the parent's market earning potential that could be invested in the education of their children. In this context, parental

[77] This upper age bound is also used in Tansel (2002)'s study for Turkey.

education may influence the educational attainment of children in different ways. First, parental education may lead to positive attitudes towards the accumulation of human capital and, thus, it results in higher investment in their children's education (Al-Samarrai and Reilly, 2000). Second, parents with higher education levels can provide an environment convenient to better learning at home through, for example, helping children with their school work which may reduce the cost of education and may increase the human capital received per year by the child (Gertler and Glewwe, 1990; Handa, 1996; Kabubo-Mariara and Mwabu, 2007). Third, higher education for mothers, in particular, may result in increased bargaining power in the household and better educated mothers may decide to allocate more resources toward their children's education than their husbands would (Thomas, 1994; Holmes, 2003). The mother and father's education are likely to have different associations with different levels of schooling and each gender. Therefore, the categories for the education variables are included separately for mothers and fathers in the analysis. The categories for those variables (defined for the mother and father separately) are: parent not having any educational qualification; primary school (five-year primary education); middle school (or eight-year primary education); secondary education (including any secondary level education and vocational schools); and university education (including any post-secondary education). These categories are converted to dummy variables and parents without any educational qualification are used as the omitted category.

Labour Force Participation of the Parents: This variable has three different categories for the father's labour force participation: not working (including the unemployed, students, the retired, unpaid family workers, ill/disabled members), participation in the labour market (not including the self-employed) and self-employment. The 'not working' category is used as the omitted category. The educational attainment of children, particularly sons, may be negatively correlated with having self-employed fathers due to early entry into work. Another variable included in the regressions is a dummy variable for the mother's participation in the labour market[78]. Although having a mother participated in the labour market may lead to higher household resources being available for the children's education, it may also affect the education of older children, particularly of daughters, if help is needed to do housework.

[78] Although information on whether the mother is self-employed or not is available in the survey, the self-employment category is not used in the analysis because of an insufficient number of observations for the self-employment category.

Agricultural Controls: Since it is common to use child labour in the household in Turkey (Rankin and Aytac, 2006) and the children of farming families are generally expected to work in the fields, two dummy variables representing whether the mother and the father work in agricultural activities are included. These variables only include the mother and the father who reported working in agricultural activities in the survey. However, individuals are generally working as unpaid family workers in rural areas in Turkey and, therefore, they may report themselves as not working in agriculture. This may lead to underestimation of the effect of the agricultural controls on educational attainment.

4.4.2.2 Household Characteristics

Total Household Expenditure (monthly): Financial limitations and credit constraints are considered as important factors influencing educational investment decisions. The premise is that if families are credit constrained then current income may affect a family's capacity to invest in their children's education. The effect may be larger for girls since poorer families may consider boys as potential breadwinners and, thus, may give priority to a boy's education (Hisarciklilar, 2002). Total household expenditure is used as a measure of household permanent income. As Tansel (2002) argued, the primary reason for choosing total expenditure is that it is easier to measure than total household income and it is measured with less error. Furthermore, income may be subject to transitory fluctuations while saving allows for the smoothing of expenditure over time.

Household Wealth: In addition to total household expenditure, household wealth is proxied by two other variables: a dummy variable which equals 1 if the household owns its own dwelling house and a variable indicating the number of rooms in the house.

Extended Family and the Number of Adults in the Household: The opportunity cost of children's education, particularly girls' education, depends on family size and composition and relates to a women's traditional role in the family. In traditional families, girls are generally expected to perform household tasks such as helping with housework or helping to take care of younger siblings. However, the existence of an adult (or grandparent) who provides child care assistance or helps with other household tasks, as is common in Turkey, is expected to decrease the opportunity cost of the education of girls and may lead to higher educational attainment. On the other hand, the

existence of elderly family members in the household, which is particularly common in extended families, who need to be taken care of, is expected to increase the opportunity cost of schooling (Hisarciklilar, 2002). Two variables are included in the models to capture the opportunity cost of schooling. The first one is a dummy variable indicating whether the family is an extended family, which consists of at least three generations living together. The second one is the number of adults, aged 15 and over, in the household.

The Number of Younger Siblings (aged 0-5) in the Household: Another way to account for the opportunity cost of children's schooling is to include the number of very young siblings (aged 0-5) in the household. Most children contribute to household resources by freeing up the adult family member's time such as helping with housework or taking care of younger siblings rather than through direct participation in the labour market and these responsibilities are likely to be imposed more heavily on daughters than on sons (Tunali, 1996; Rankin and Aytac, 2006). In this context, an increase in the number of very young children in the household may increase the demand for the labour supply of girls in childcare in the household and this will affect their educational attainment negatively relative to that of boys (Glick and Sahn, 2000).

Being the Eldest Resident Child in the Household: A dummy variable indicating being the eldest resident child in the household is included in the analysis in order to control for the fact that being the eldest child in the household potentially has the advantage of benefiting most from family income and the time resources due to the constraints of transferring resources across time periods (King and Lillard, 1987). On the other hand, the opportunity cost of child schooling may become higher with being the eldest resident child in the household because older children are generally expected to perform a range of time-consuming tasks in the home that may impede the educational attainment of children (King and Hill, 1993).

4.4.2.3 Community Characteristics

Urban/Rural Residence and Region Dummy Variables[79]: Since the residence where the children live is important for their educational attainment, a dummy variable which equals 1 if the individual resides in an urban area and 0 if the individual resides in a

[79] Urban areas are defined as settlements having a population above 20,000 and rural areas are defined as settlements having a population equal to or below 20,000 (TurkStat, 2006).

rural area is included in the analysis. Children living in a rural area may face higher opportunity costs of educational attainment because of engaging in agricultural activities as well as higher prices for transport. In addition, rural areas can be seen as more traditional and conservative in regard to tastes as compared to urban areas (Wolfe and Behrman, 1984). Furthermore, a set of region dummy variables, Marmara, Aegean, Mediterranean, Central Anatolia, Black Sea, East Anatolia and Southeast Anatolia, is also included since there is a great difference between the regions in terms of the structure of the society, traditions and values, and the levels of industrialisation and economic development. The dummy variable for the Marmara region is used as the omitted category because it is the wealthiest region and is the centre of much of the economic activity in the country. In recent decades, a significant sectoral and regional shift of the labour force out of traditional agriculture has occurred towards the west of Turkey (Tansel, 2002). In terms of the supply side of education, schools in the west generally provide higher quality education while there are an insufficient number of schools in the east due to its geographic situation and the lower investment of the government in this region (Hisarciklilar, 2002).

Distance to a Metropolitan Centre: Recent decades have also witnessed a considerable population movement (i.e., migration) in Turkey. The high level of industrialisation in the west, mechanisation of agriculture, recent terrorist activities and wide sectoral and regional differences in productivity in the east are among the most important reasons for the population movement (Tansel, 2002). Since distance is an important factor for migration, distance of a province from the regional metropolitan centre is included in the analysis in order to capture the association between migration possibilities and educational attainment following Tansel (2002). To be specific, distance to a metropolitan centre is the distance, measured in kilometres, of each of the 26 provinces in the sample from a regional metropolitan centre in each of the seven regions. A province in a region is determined as the metropolitan centre of that region if it received the most internal migration according to the Turkish Statistical Institute (2000) provincial migration data and these data are merged with the individual level data used in this chapter.[80]

[80] The provinces that were determined as metropolitan provinces of the regions are as follows: Istanbul in the Marmara region, Ankara is the Central Anatolia region, Izmir in the Aegean region, Mersin in the Mediterranean region, Samsun in the Black Sea region, Malatya in the Eastern region and Gaziantep in the Southeast region. The distance data are drawn from the General Directorate of Highways (2010).

4.4.2 Descriptive Statistics

Descriptive statistics for the variables used in the empirical analysis are reported in the Appendix in Table A4.1. This table provides information on descriptive statistics for the continuous variables and percentage distributions for the categorical variables for the samples used for the primary level and secondary level education analysis and for girls and boys separately.

Table A4.1 indicates that there is no significant difference in the descriptive statistics of the continuous variables between the primary and secondary educational attainment samples for both genders. However, the percentage distribution of the education status of the parents is slightly different between the samples for both genders. Most fathers and mothers had completed primary school (around 50%) and mothers have consistently less schooling as compared to fathers. For example, the proportion of mothers who had completed secondary school is approximately 5% whereas this proportion is nearly 12% for fathers. Similarly, 32% (or over for some samples) of mothers have no educational qualification while 9% of fathers have no educational qualification. With respect to the employment status of the parents, for the primary education sample, 21% of fathers are not working and most of the fathers are employed whilst, for the secondary education sample, 29% of fathers are not working and most of the fathers are self-employed. Furthermore, the vast majority of mothers are not working for all samples. The proportion of fathers engaged in agriculture is higher than the proportion of mothers engaged in agriculture for all samples and these proportions are generally higher for the sample of girls than for the sample for boys.

Table A4.1 also shows that most of the children's parents own their own dwelling in all samples (nearly 77% and 81% for the primary and secondary level education samples, respectively). Furthermore, 16.6% of girls and 15.9% of boys live in extended families for the primary education sample while these percentage rates increase to 18% for girls and 22.1% for boys in the secondary education sample. With regard to being the eldest resident child in the household, Table A4.1 indicates that the proportion being the eldest child is nearly 7 percentage points higher in the secondary education sample since this sample includes older children (aged 14-20) as compared to the primary education sample (aged 14-17). The proportions of children in rural and urban areas indicate that 68.3% of girls and 69.8% of boys live in urban areas for the primary education sample

whereas 65.8% of girls and 69.2% of boys live in urban areas for the secondary education sample. The percentage distributions of the regions show that the highest proportions of children live in the Marmara region (nearly 21%) whilst the lowest proportions of children live in the Eastern region for both samples (nearly 9%).

Table 4.5 below presents the percentage distribution of the highest level of educational attainment of children for the samples used for the primary and secondary level educational attainment (i.e., the dependent variables)[81]. These proportions are calculated by gender and by urban/rural residence. This table indicates that the educational attainment rates of girls are consistently lower than the rate of boys in (8-years) primary education for both samples. Similarly, the secondary school educational attainment rate of girls is lower than the rate of boys. With regard to having no educational qualification, the split between having no educational qualification by literacy status is important since 3.9% of girls are illiterate in the primary education sample while this rate is 4.3% in the secondary education sample. In the rural areas, these rates increase to 5.2% and 5.8% for the primary and secondary education samples, respectively.

Table 4.5: The Percentage Distribution of the Highest Level of Educational Attainment of Children by Gender and Urban/Rural Residence

Panel A: Primary Education (Aged 14-17)						
	Total		**Urban**		**Rural**	
Final Grade Attainment	Girls	Boys	Girls	Boys	Girls	Boys
No Qualification (illiterate)	3.9%	1.5%	3.3%	1.2%	5.2%	2.2%
No Qualification (literate)	24.3%	23.1%	22%	20.9%	29.4%	28.2%
Primary School (8 years)	71.8%	75.4%	74.7%	77.9%	65.4%	69.6%
Panel B: Secondary Education (Aged 18-20)						
	Total		**Urban**		**Rural**	
Final Grade Attainment	Girls	Boys	Girls	Boys	Girls	Boys
No Qualification (illiterate)	4.3%	1.4%	3.4%	1.4%	5.8%	1.3%
No Qualification (literate)	4.5%	2.5%	3.3%	1.8%	6.9%	4.1%
Primary School (5 years)	28.1%	19.9%	20.1%	13.2%	43.4%	34.9%
Primary School (8 years)*	16.9%	23.6%	15.2%	23.2%	20.4%	24.5%
Secondary School	46.2%	52.6%	58%	60.4%	23.5%	35.2%

Source: Household Budget Survey, 2003.

*This category also includes children who graduated from middle schools which used to take three years to complete before the compulsory education law. The compulsory education law was put into place in 1997 and expanded the minimum years of compulsory education to eight years covering the middle school.

[81] As stated in Section 4.4.1 above, for the primary education model only children in the 14-17 age group and, for the secondary education model, only children in the 18-20 age group are included in order to obtain samples of children who should have finished primary and secondary education, respectively.

The differences in the highest level of educational attainment between girls and boys are most apparent in the rural areas and particularly in the secondary school graduation rates. 23.5% of girls in the rural areas graduated from secondary school while this rate is 35.2% for boys in the rural areas. Similarly, particularly for girls, there is a considerable difference in the rates of highest educational attainment between the urban and rural areas. For example, 58% of girls in urban areas are secondary school graduates while this rate is 23.5% for girls in rural areas. It can be argued that girls are less likely to complete primary school as compared to boys and when they complete primary school, they are less likely to go beyond primary school particularly in the rural areas.

Table 4.6 below presents the percentage distribution of the highest level of educational attainment of parents by urban/rural residence. This table indicates that mothers are less likely to have attended primary or higher levels of education as compared to fathers and the gender gap becomes more apparent at the secondary and higher education levels. For example, 14.9% of fathers had completed secondary education but this rate is only 6.5% for mothers for the primary education sample. Similarly, for the secondary education sample, 12.6% of fathers had completed secondary education while this rate decreases to 4.9% for mothers. With regard to the tertiary education level, the rate of university graduation of fathers is 4.1 percentage points higher than the rate of mothers for the primary education sample while this difference is 4.5 percentage points for the secondary education sample. Moreover, the proportions of mothers having no educational qualification are 33.7% and 38% for the primary and secondary education samples, respectively. Table 4.6 also indicates a considerable gender difference at all education levels between the urban and rural areas. For example, in the rural areas, the proportions having no educational qualification are 47.1% for mothers and 19.6% for fathers in the secondary education sample. In the urban areas, however, these proportions are 33.7% and 9.7% for mothers and fathers, respectively.

When Table 4.6 is compared with Table 4.5, it can be said that the highest educational attainment rates peak around the primary education (5-years) level for parents whereas the educational attainment rates of children are generally highest at the (8-years) primary education level, especially for boys. This pattern may be attributed to the compulsory education law that made 8-years of primary education compulsory for everyone and was put into place in 1997.

Table 4.6: The Percentage Distribution of the Highest Level of Educational Attainment of Parents of Children by Urban/Rural Residence

Panel A: Primary Education (Aged 14-17)						
	Total		**Urban**		**Rural**	
Final Grade Attainment	Father	Mother	Father	Mother	Father	Mother
No Qualification	11.2%	33.7%	8.5%	29.8%	17%	42.5%
Primary School (5 years)	55.8%	52.6%	51%	52.8%	66.3%	52.3%
Middle School	11.4%	4.8%	13.1%	5.8%	7.6%	2.6%
Secondary School	14.9%	6.5%	18.5%	8.4%	7%	2.1%
University	6.4%	2.3%	8.5%	3.1%	2%	0.4%
Master/Doctorate	0.3%	0.1%	0.4%	0.1%	0.1%	0.1%
Panel B: Secondary Education (Aged 18-20)						
	Total		**Urban**		**Rural**	
Final Grade Attainment	Father	Mother	Father	Mother	Father	Mother
No Qualification	13%	38%	9.7%	33.7%	19.6%	47.1%
Primary School (5 years)	57.6%	50.7%	53.2%	51.4%	66.4%	49.2%
Middle School	10.1%	4.4%	11.9%	5.6%	6.5%	2.1%
Secondary School	12.6%	4.9%	16%	6.5%	5.7%	1.2%
University	6.4%	1.9%	8.8%	2.7%	1.7%	0.3%
Master/Doctorate	0.3%	0.1%	0.4%	0.1%	0.1%	0.1%

Source: Household Budget Survey, 2003.

Table 4.7 below presents the percentage distribution of the highest level of educational attainment of children by region. Perhaps the most important figures in this table are that the worst educational attainment rates are observed in the Eastern and Southeastern regions, which can be explained by the fact that these regions are considerably poorer than the rest of the country. For example, the proportion of illiterate girls is 14.7% in the Southeastern region whilst this rate is 0.7% in the Marmara region for the primary education sample. Moreover, the proportion of girls having no educational qualification (illiterate and literate) is higher than the proportion of primary school graduates in the Southeastern region. There is also a significant difference in the secondary education level between girls and boys in the Eastern and Southeastern regions. In the Southeastern region, 17.2% of girls and 32.3% of boys graduated from secondary school and in the Eastern region, these proportions are 33.9% and 59.8% for girls and boys, respectively.

Table 4.7 also shows that splitting the having no educational qualification category into illiterate and literate individuals, which is discussed in Section 4.4.2 above, is important since 14.7% of girls and 4.8% of boys in the Southeastern region and 7.1% of girls and

1.5% of boys in the Eastern region were illiterate at the time of the survey for the primary education sample. Similarly, for the secondary education sample, the proportions of illiterate girls are 7.2% and 19.1% for the Southeastern and Eastern regions, respectively.

Table 4.7: The Percentage Distribution of the Highest Level of Educational Attainment of Children by Gender and Region

Panel A: Primary Education (Aged 14-17)				
Region	**Gender**	**No Qualification (illiterate)**	**No Qualification (literate)**	**Primary School (8 years)**
Marmara	Girls	0.7%	18%	81.3%
	Boys	0.7%	17.8%	81.5%
Aegean	Girls	0.6%	21%	78.4%
	Boys	1.2%	18.5%	80.3%
Mediterranean	Girls	1.2%	21.9%	76.9%
	Boys	1.1%	21.1%	77.8%
Central Anatolia	Girls	0.9%	21.5%	77.6%
	Boys	0.5%	19.8%	79.7%
Black Sea	Girls	1.5%	21.6%	76.9%
	Boys	0.9%	20.4%	78.7%
Eastern	Girls	7.1%	23.9%	69%
	Boys	1.5%	18.5%	80%
Southeastern	Girls	14.7%	40.8%	44.5%
	Boys	4.8%	43.6%	51.6%

Panel B: Secondary Education (Aged 18-20)					
Region	**Gender**	**No Qual.**	**Primary School (5 years)**	**Primary School (8 years)**	**Secondary School**
Marmara	Girls	2.5%	21.7%	18.4%	57.4%
	Boys	1.3%	16.6%	24%	58.1%
Aegean	Girls	4.2%	30.6%	21.1%	44.1%
	Boys	1%	16.8%	29.4%	52.8%
Mediterranean	Girls	5%	20.9%	16.5%	57.6%
	Boys	3.9%	16.8%	27%	52.3%
Central Anatolia	Girls	2.2%	21%	15.6%	61.2%
	Boys	1.7%	16.5%	25.2%	56.6%
Black Sea	Girls	1.3%	33.2%	21.3%	44.2%
	Boys	2.2%	19.8%	21.2%	56.8%
Eastern*	Girls	12.3%	37.3%	16.5%	33.9%
	Boys	2.9%	12.7%	24.6%	59.8%
Southeastern*	Girls	35.1%	37.7%	10%	17.2%
	Boys	14.7%	38.1%	14.9%	32.3%

Source: Household Budget Survey, 2003.

*For all regions, the category indicating no qualification includes both literate and illiterate individuals since the rates of literate and illiterate individuals are close to each other for all regions except the Eastern and Southeastern regions. For the Eastern region, the proportion of illiterate individuals is 7.2% for girls while this proportion is 0.5% for boys. Similarly, for the Southeastern region, the proportions of illiterate individuals are 19.1% and 5.9% for girls and boys, respectively.

Table 4.8 below presents the percentage distribution of the highest level of educational attainment of children by whether their parents' work in agriculture. This table indicates that girls have higher rates of having no educational qualification and lower rates of having higher educational attainment as compared to boys if their parents are working in agriculture for both samples. Furthermore, this table shows that the proportions of children who had completed (8-years) primary education or secondary education (i.e., higher education levels) are lower if their fathers are working in agriculture as compared to mothers working in agriculture for both genders.

Table 4.8: The Percentage Distribution of the Highest Level of Educational Attainment of Children by the Parents Working in Agriculture

Panel A: Primary Education (Aged 14-17)				
	Father Engaged in Agriculture		**Mother Engaged in Agriculture**	
Final Grade Attainment	Girls	Boys	Girls	Boys
No Qualification (illiterate)	6.2%	3.1%	7.4%	1.9%
No Qualification (literate)	33.3%	32%	26.8%	26.9%
Primary School (8 years)	60.5%	64.9%	65.8%	71.2%
Panel B: Secondary Education (Aged 18-20)				
	Father Engaged in Agriculture		**Mother Engaged in Agriculture**	
Final Grade Attainment	Girls	Boys	Girls	Boys
No Qualification (illiterate)	7.2%	1.4%	5.4%	1.5%
No Qualification (literate)	8.7%	5.7%	10.7%	1.5%
Primary School (5 years)	50.9%	39.8%	35.7%	35.8%
Primary School (8 years)	18.3%	25.2%	20.5%	25.4%
Secondary School	14.9%	27.9%	27.7%	35.8%

Source: Household Budget Survey, 2003.

Table 4.9 below presents the percentage distribution of the highest level of educational attainment of children by the employment status of parents. This table indicates that there is no significant difference in the educational attainment at the (8-years) primary education and secondary education levels between girls and boys if their fathers are employed. However, the rate of secondary school educational attainment of girls is lower than the rate of boys if their fathers are self-employed or if their mothers participate in the labour market.

Table 4.9: The Percentage Distribution of the Highest Level of Educational Attainment of Children by the Employment Status of Parents

Panel A: Primary Education (Aged 14-17)						
	Father employed		**Father self-employed**		**Mother working**	
Final Grade Attainment	Girls	Boys	Girls	Boys	Girls	Boys
No Qualification (illiterate)	2.9%	1%	4.8%	2.4%	2.8%	1.4%
No Qualification (literate)	23%	22.3%	27.6%	26.1%	23.5%	22%
Primary School (8 years)	74.1%	76.7%	67.6%	71.5%	73.7%	76.6%
Panel B: Secondary Education (Aged 18-20)						
	Father employed		**Father self-employed**		**Mother working**	
Final Grade Attainment	Girls	Boys	Girls	Boys	Girls	Boys
No Qualification (illiterate)	2.1%	1.3%	5.4%	1.3%	2.1%	1.4%
No Qualification (literate)	3%	1.8%	6.4%	4%	4.9%	1.4%
Primary School (5 years)	19.5%	11.7%	38.1%	26.8%	27.9%	18.5%
Primary School (8 years)	17.3%	23.5%	17.4%	24.4%	21.9%	26.7%
Secondary School	58.1%	61.7%	32.7%	43.5%	43.2%	52%

Source: Household Budget Survey, 2003

4.5 RESULTS

As discussed in Section 4.4 above, the ordered probit model, which treats the highest level of educational attainment as a discrete and non-normally distributed variable, is estimated as a first step, for girls and boys separately, since the dependent variable is categorical in nature (King and Lillard, 1983). Furthermore, the random effects ordered probit model is also estimated in the second specification to allow for the unobserved common family characteristics for children from the same household. These models are estimated for primary and secondary levels of educational attainment separately to account for the right-censored observations for currently enrolled children. In this respect, the samples are restricted to only include children above the 'graduation' ages (Tansel, 2002). The results of estimating the ordered probit and random effects ordered probit models are presented in the Appendix and, in this section, these results are discussed and the key findings are summarised.

4.5.1 Primary Level Education

4.5.1.1 The Results of the Ordered Probit Analysis

Table A4.2 presents the coefficients for the ordered probit models for girls and boys. Although the parameter estimates of an ordered probit model indicate the direction of the relationships and the statistical significance levels for the explanatory variables in the model, marginal effects allow for observing the magnitude of the effects of a change in the explanatory variables on each education transition. Tables A4.3 and A4.4 present the marginal effects, which are calculated at the sample means of the explanatory variables, for the ordered probit model and these results can be used to compare the magnitude of the influence of the variables for girls and boys.

The results of the ordered probit models indicate that parental education is an important determinant for the primary school educational attainment of both girls and boys as expected. Since much of the human capital accumulation of the child occurs in the home especially in the pre-school period of the life-cycle, parental education plays a significant role in this process (Mincer, 1984). To be specific, parental education might influence the educational attainment of children directly via their preferences towards education, acting as role models for their children, determining the quality and quantity of time spent with the child and providing a better environment for education and indirectly via household resources and the bargaining power of the mother in the

household. Therefore, even when the effect of family income is controlled for, a positive relationship is expected between parental education and child schooling. In this context, the results indicate that parental education is associated with higher primary school educational attainment for both genders except for the mother's education at the (8-years) primary level. The marginal effects, however, indicate that the father's education is a more important determinant of the primary school educational attainment of children than the mother's education for both genders. Table A4.3, for example, indicates that if a mother has a university degree, it increases the probability of primary school educational attainment for girls by 10.1 percentage point, whereas the impact of father's university education on the primary school educational attainment of girls is 14.1 percentage points. This finding is in line with the existing literature (Gertler and Glewwe, 1990; Al-Samarrai and Peasgood (1998)[82]. This finding is not surprising for Turkey as a developing and male-dominated country since in these countries the father is generally the main decision-maker in the family. The marginal effects also indicate that parental education is more important for the primary school educational attainment of daughters than of sons. This finding suggests that there is less intergenerational social mobility for girls than for boys since a strong association between parental education and child schooling is explained as implying less intergenerational social mobility (Tansel, 2002). There may be several reasons for this. First, it is possible that parents with relatively low levels of education are more likely to be in locations where there are some barriers to girls' education. Second, parents with lower levels of education may not be more aware of the value of girls' education and this affects their attitudes towards the education of girls (Tansel, 2002).

In terms of the employment status of parents, the results indicate that children, who have self-employed fathers, are less likely to attend primary school which supports the argument of a higher opportunity cost of attending school for those who can work with their parents. Moreover, the marginal effect for having a self-employed father is larger in the boys sample than in the girls sample which suggests that there may be a high probability of boys beginning to contribute to their family's income by working with their fathers and dropping out of school early to begin to work. It should also be stated that the relationship between having an employed father and primary school educational

[82] It should be noted that in contrast to the studies for developing countries, the studies for developed countries generally have found that mother's education is more important than the father's education (Schultz, 1993; Birdsall, 1985).

attainment is similar to the effect of having a self-employed father. The mother's participation in the labour market, on the other hand, has no effect on primary school educational attainment for both genders, which is in line with the existing literature for Turkey (see, for example, Rankin and Aytac, 2006; Hisarciklilar *et al.*, 2009).

Similar to the effect of the mother's employment status, having a mother engaged in agriculture has no effect on the primary school educational attainment of children. However, having a father engaged in agriculture is negatively associated with the primary school educational attainment of boys. This finding is not surprising for Turkey since the demands on children's time, particularly for teenage boys, are especially high in families engaged in agriculture due to the extensive use of child labour in this sector. In this respect, in the EU report (2008) for Turkey, it is stated that efforts to curb the use of child labour in agriculture should be improved since there is no comprehensive legal protection for children working in agriculture.

In the existing literature, one of the most consistent findings is the positive relationship between household income and child schooling (Wolfe and Behrman, 1984; Birdsall, 1985; Behrman and Rosenweig, 1994; Parish and Willis, 1993; Behrman and Knowles, 1999; Tansel, 2002). The results of this chapter also confirm this positive association between household income and primary school educational attainment for both genders. Furthermore, the marginal effects indicate that, as household income increases, girls are more likely to attend primary school as compared to boys. This could be entirely due to a taste effect or could be explained as better off families being able to hire help for childcare and for housework (Glick and Sahn, 2000), which potentially decrease the opportunity cost of time spent at school for girls, and, thus, girls' educational attainment may be enhanced in the case of higher household resources.

With regard to another important household factor, the results of the ordered probit models indicate that there is a positive association between the number of adults in the household and the primary school educational attainment of children. This finding may reflect the fact that children living in households with more adults, who can help with the housework or can help the father in the work place, are more likely to attend primary school since their opportunity cost of school time is lower in this case. On the other hand, living in an extended family is found to be negatively associated with primary school educational attainment for both boys and girls. As stated in Section

4.4.2, it may be the case that the presence of elderly household members who need to be taken care of in the extended families may increase the opportunity cost of schooling especially for girls. However, the marginal effects surprisingly indicate that the effect of living in an extended family is larger for boys than for girls. This finding may also be attributed to the fact that in the extended families the available resources have to be divided among more family members and this may decrease the available resources for the educational attainment of children (Smits and Hosgor, 2006).

In addition to family size, the age structure of siblings is another important factor in educational attainment studies (Glick and Sahn, 2000) since child labour is a pervasive problem especially in developing countries and children are expected to perform some low-skilled household tasks such as taking care of younger siblings or helping with housework. In Turkey, girls often stay at home and are responsible for these tasks, particularly taking care of their siblings, more often than boys (Rankin and Aytac, 2006). In this respect, the number of very young siblings (aged 0-5) is found to be negatively associated with the primary school educational attainment of both girls and boys but the marginal effect of this variable is larger in the girls sample than in the boys sample. This finding supports the fact that taking care of siblings is likely to be imposed more heavily on girls than on boys. This finding is important in terms of household time allocation decisions relating to girls' schooling opportunities and implies that subsidised child care, which may decrease the need for girls' domestic labour, may increase the probability of girls' primary school educational attainment (Glick and Sahn, 2000). Moreover, an alternative explanation for this negative association may be based on the child quantity-quality model, which is proposed by Becker (1960) and later developed by Becker and Lewis (1973), as this model predicts a negative relationship between the number of children in the household and the average education level of the children.

On the other hand, being the eldest resident child is found to be positively related to primary school educational attainment which is in line with the existing literature (King and Lillard, 1987). Being the eldest child in the household has some advantages such as spending more time with parents than later-born children and benefiting more from the family's financial resources (King and Lillard, 1987). The marginal effects of this variable indicate that boys are more likely to have the advantage of benefiting most from family income and time resources as compared to girls since the marginal effect is larger in the boys sample.

For community characteristics, the results of the ordered probit models indicate that living in an urban area has no effect on primary school educational attainment. However, distance to a metropolitan centre is negatively associated with primary school educational attainment which indicates that, as the distance to a metropolitan centre increases, children are less likely to attend primary school. Since distance is a crucial factor for migration, this finding implies that easier migration possibilities or closeness to a metropolitan centre may enhance access to primary school (Tansel, 2002). The results also indicate that girls living in the Eastern and Southeastern regions and boys living in the Southeastern region are less likely to attend primary education as compared to children living in the Marmara region. This finding is as expected given that the eastern part is different from the rest of the country due to the problems of underdevelopment, low levels of industrialisation, ethnic suppression and political conflict in this region (Rankin and Aytac, 2006). It should also be noted that the marginal effects of the variables indicating distance to a metropolitan centre and the Southeastern region are larger in the girls sample than in the boys sample which implies that community characteristics are more important for the primary school educational attainment of girls as compared to boys. In addition, the marginal effects of the variable indicating the Southeastern region are the largest marginal effects in both samples, which shows the considerable effect of the location where children live on their primary school educational attainment.

4.5.1.2 The Results of the Random Effects Ordered Probit Analysis

As stated in Section 4.4.2, the ordered probit model is extended to include a random error term which takes the same value for each individual coming from the same household (Frechette, 2001) in order to deal with the problem of unobserved family characteristics. Tables A4.5 to A4.7 present the coefficients and the marginal effects of the random effects ordered probit models for the primary school educational attainment of girls and boys.

In general, compared to the results of the ordered probit models, the signs and the statistical significance levels of the variables are same in the random effects ordered probit models for both genders. The marginal effects, however, are greater in the random effects ordered probit model particularly for girls, which supports the potential existence of bias in the estimation results when unobserved family characteristics are

not allowed for. The increase in the marginal effects is most apparent for parental education for both genders, which highlights its central importance for the primary school educational attainment of children. In particular, the marginal effects of the variables representing the mother's education increased in the results of the random effects ordered probit model for girls. Similarly, in their studies for Turkey, Tansel (2002) and Smits and Hosgor (2006) also found that mother's education is an important determinant of educational attainment of girls. This finding may be explained as mothers with higher education are more aware of the benefits of female education and they may have more power in the household bargaining process and, thus, may affect decisions relating to the education of their children.

Another difference is that the marginal effects relating to the household income variable have increased in the random effects ordered probit model as compared to the ordered probit models and this increase is greater in the girls' sample. Furthermore, the variable indicating owning their own dwelling, which is used as a proxy for household wealth, has become statistically significant for girls. These findings are consistent with the existing literature. For example, Glick and Sahn (2000) for Guinea and Kabubo-Mariara and Mwabu (2007) for Kenya found that household resources are more important for girls than boys. Even when education is provided free of charge by public schools, educational attainment requires out-of-pocket expenditure such as for learning materials, contributions to the school, school uniforms and travel expenses. For some reasons, these expenditures may differ between girls and boys. For example, parents may be more reluctant to send their daughters to school without proper school uniforms and this increases the cost of educational attainment of girls (Hill and King, 1995).

Finally, the magnitude of ρ, which represents the proportion of total variance in the dependent variable explained by the random error component, indicates that 43% of the variation in the primary school educational attainment of girls, and 40% of the variation in the primary school educational attainment of boys, coming from the same household is explained by common unobserved family characteristics. This finding has provided empirical evidence for that random effects ordered probit model rather than ordered probit model should be preferred due to the potential existence of bias in the estimation results when unobserved family characteristics are not accounted for in the analysis.

4.5.2 Secondary Level Education

4.5.2.1 The Results of the Ordered Probit Analysis

Table A4.8 presents the coefficients and Tables A4.9 and A4.10 present the marginal effects of the ordered probit model for secondary school educational attainment for girls and boys. The results for secondary school educational attainment are different from those obtained for primary level education in terms of the statistical significance levels of some variables and the magnitude of marginal effects. This indicates that educational attainment and the potential factors associated with gender inequality differ by the level of education. The results indicate that the father's education is the main determinant of the secondary school educational attainment of girls whereas both the father's and the mother's education are the main determinants of the secondary school educational attainment of boys since the marginal effects of these variables are the largest in magnitude. These findings highlight the crucial importance of parental education for the post-compulsory educational attainment of children in Turkey. As stated in Section 4.2.1.3 above, there is an e-school based system to monitor children who are not enrolled at primary school as well as legal sanctions for parents who do not send their children to compulsory primary school in Turkey. Furthermore, there are financial incentives to encourage primary school attendance such as scholarships for children from poor households. However, the secondary school attainment of children largely depends on family decisions and, thus, parental education plays a more critical role in the secondary school educational attainment process.

Similar to the primary school educational attainment model, the results for secondary school educational attainment indicate that the father's education is more important than the mother's education for both genders and the marginal effects from the father's education are significantly greater for girls than boys. In the existing literature, the father's education is regarded mainly as a measure of his income and, hence, affects the education of children through income. The education of the mother, on the other hand, has a direct role in developing the skills of their children and reflects the input of parents' time into children's education (King and Lillard, 1983). In this context, it can be argued that the father's education is more important than the mother's direct role in the schooling process, such as providing motivation for education, efficient household expenditure and production, parental role modelling (Tansel, 1997), for girls to go beyond compulsory primary education.

The results of the ordered probit model also indicate that there is a negative association between having a father engaged in agriculture and the secondary school educational attainment of children. This variable, however, was not statistically significant in the results for the primary school educational attainment of girls whereas it was one of the main determinants of the primary school educational attainment of boys. This finding supports the fact that girls who have families engaged in agriculture generally help their families in agricultural activities as unpaid family workers in Turkey (Rankin and Aytac, 2006). Furthermore, the marginal effects of this variable are greater for girls than boys, which indicate that having a father engaged in agriculture is more important for girls than boys when it comes to attending secondary school. Similarly, in his study for Turkey, Kirdar (2009) also found that the father's employment in agriculture matters only after completing compulsory primary level education.

Similar to the primary school educational attainment model, the results from the secondary school educational attainment model also support the positive association between household income and child schooling. Moreover, the marginal effects of this variable are greater in the girls sample, which is similar to the results of the primary school educational attainment model. The budget constraint shapes the allocation of resources between the family members and each child, which may make it less likely, especially for girls, to obtain education, particularly secondary level education, since there is more financial support available at the primary level than at the secondary level education in Turkey (Duman, 2010). Additionally, in contrast to primary education, the results for secondary school educational attainment confirm the importance of household wealth since the variables indicating owning their own dwelling and the number of rooms in the house are found to be statistically significant and positively associated with the secondary school educational attainment of boys.

With regard to the other household factors, the results for secondary school educational attainment are different from the results for primary school educational attainment in two respects. First, the sign of the variable indicating the number of adults in the household is negative in the results for the secondary education model, as opposed to having a positive sign in the primary education model. These findings indicate that, at the primary education level, living with more adults in the household decreases the opportunity cost of children's schooling with more adults to help with housework or the family business. However, this is not the case for secondary education due to the

negative sign of the variable indicating the number of adults in the household. This finding may reflect the fact that as family size increases, the level of household resources allocated for expenditure on the post-compulsory educational attainment of children decreases since family resources have to be shared among more family members. Second, living in an extended family and being the eldest resident child in the household have no effects on secondary school educational attainment for both genders as opposed to the results for the primary education model. However, in accordance with primary school educational attainment, the results indicate that there is a negative relationship between the number of very young siblings (aged 0-5) and the secondary school educational attainment of children. This finding implies that having younger siblings at home is especially burdensome for girls who may be expected to care for younger siblings and this may lead to an increase in the opportunity cost of their education.

The community characteristics also have a large influence on the education opportunities for girls and boys. In this regard, the results indicate that living in an urban area is positively associated with the secondary school educational attainment of children, which was statistically insignificant in the results for the primary school educational attainment model. This finding can be explained by the greater accessibility and probably higher quality of schools, lower travel expenses, and greater labour market opportunities for individuals, particularly for females, with high levels of education in the urban areas (King and Lillard, 1987; Al-Samarrai and Reilly, 2000). The number of secondary schools in rural areas is much lower than the number of primary schools in Turkey and, thus, children have to go far from home to attend secondary school[83]. However, parents may prefer not to send their children, especially girls, to distant schools and this constitutes a barrier to the educational attainment of girls (Hill and King, 1995). Moreover, the cost of education can be higher in rural areas due to travel costs and this may also lower the educational attainment of children.

In a broader sense, when the education level increases from the primary to the secondary level, it can be said that community characteristics have become more important for the education of girls as compared to the education of boys since the

[83] According to the latest statistics of the Ministry of Education, the number of primary schools in rural areas is 20717 whereas the number of secondary schools in rural areas is 822 in Turkey (National Education Statistics, 2009).

variable indicating the distance to a metropolitan area is found to be negatively associated with the secondary school educational attainment of girls whereas it has no effect on the secondary school educational attainment of boys. Furthermore, all of the regional dummy variables, except for the Eastern region, are statistically significant for girls while only the Black Sea, Eastern and Southeastern regions are statistically significant for boys. The marginal effects relating to the regional dummy variables are greater in size in the girls sample as compared to the boys sample. However, some of the regional dummy variables, such as the Mediterranean region, the Central Anatolia region and Black Sea region in the girls sample and the Black Sea region and the Eastern region in the boys sample, have unexpected positive signs which imply that the children who are living in these regions are more likely to attend secondary school as compared to the Marmara region which is the most developed region of Turkey. Other studies for Turkey (see, for example, Tunali, 1996; Tansel, 2002; Rankin and Aytac, 2006) have also found similar results and have explained these results as follows: the more developed regions have more job opportunities for young individuals both in the formal and informal sectors, which may result in dropping out of school earlier and not going beyond compulsory primary education. This may be especially true for children who are living in the shanty towns of the big cities and who have poor families (Tansel, 2002; Rankin and Aytac, 2006).

4.5.2.2 The Results of the Random Effects Ordered Probit Analysis

Table A4.11 presents the coefficients of estimating the random effects ordered probit model for the secondary school educational attainment of both girls and boys while Tables A4.12 and A4.13 present the marginal effects of the random effects ordered probit model for the secondary school educational attainment of girls and boys, separately. Compared to the results for the ordered probit model, the results of the random effects ordered probit model are different in terms of larger marginal effects particularly for girls. Similar to the primary school educational attainment model, this increase is most apparent in the marginal effects representing parental education. Having a father who graduated from university has the largest marginal effect in magnitude for both genders. Furthermore, the marginal effects for both the mother's education and the father's education have mostly increased in the girls sample. Similarly, the marginal effects relating to having a father working in agriculture have increased in size in both the samples but this increase is greater for girls.

The results of the random effects ordered probit model also support the findings from the ordered probit model in terms of the community characteristics. The marginal effects indicating urban/rural residence, distance to a metropolitan centre and the regional dummy variables are greater in the results of the random effects ordered probit models for both genders but the increases in these marginal effects are larger for girls which supports the hypothesis that community characteristics are more important for girls than for boys when it comes to going beyond compulsory primary education.

In addition, the variable indicating being the eldest resident child in the household has become statistically significant for girls in the results of the random effects ordered probit model as compared to the results of the ordered probit model. This finding indicates that there is a negative relationship between being the eldest resident child in the household and secondary school educational attainment as opposed to the results of the primary school educational attainment model which indicate a positive association between being the eldest resident child in the household and primary school educational attainment. This finding supports the existing literature which argues that the eldest daughters but not the eldest sons are less likely to go beyond compulsory primary education than their younger siblings since the eldest daughter generally faces housework and child care tasks (Greenhalgh, 1985; Rankin and Aytac, 2006).

Finally, the results of the random effects ordered probit model indicate that ρ is highly statistically significant and its magnitude shows that 70% of the total variance in the secondary school educational attainment of girls and 58% of the total variance in the secondary school educational attainment of boys, who belong to the same household, are explained by unobserved family and household characteristics. Similar to the primary school educational attainment model, this finding has provided further evidence for the importance of accounting for unobserved family characteristics for children coming from the same household in the analysis.

4.5.3 Summary of the Key Findings

To summarise the key findings of the analysis of the determinants of the primary and secondary levels of educational attainment, it can be said that the results for secondary school educational attainment are different from those obtained for primary school educational attainment for girls and boys in terms of the statistical significance levels of some variables and the magnitude of marginal effects. This indicates that educational

attainment and the potential factors related to gender inequality differ by the level of education. In this respect, having parents with higher levels of education, household income, the number of adults in the household and being the eldest resident child are all positively associated with primary school educational attainment whereas having a self-employed father, living in an extended family, the number of very young (aged 0-5) siblings in the household, the distance to a metropolitan centre and living in the Southeastern region are all negatively associated with primary school educational attainment for both genders. In addition, having a father engaged in agriculture is negatively associated with boys' primary school educational attainment whereas the marginal effects from parental education, household income, the number of very young siblings and the Southeastern region are greater in magnitude for girls.

The results further indicate that community variables are more important for secondary school educational attainment than primary school educational attainment for both genders. Having a father engaged in agriculture is negatively associated with secondary school attainment for both genders. Household income, owning their own dwelling, the number of rooms in the house and living in an urban area are all positively associated with the secondary school educational attainment of boys whilst household income, living in an urban area and closeness to a metropolitan centre are among the main determinants of the girls' secondary school educational attainment. In addition, the number of adults and the number of very young siblings in the household are negatively associated with secondary school educational attainment for both girls and boys and the marginal effects of these variables are greater for girls than for boys.

Compared to the results of the ordered probit models for both the primary and secondary school educational attainment models, the signs and the statistical significance levels of the variables are generally in line with the results of the random effects ordered probit models. However, the marginal effects are larger in the random effects ordered probit models as compared to the ordered probit models particularly for girls. This increase in the size of the marginal effects is most apparent for parental education for both genders. These findings support the potential existence of bias in the estimation results when unobserved family characteristics are not accounted for in the analysis.

4.6 CONCLUSION

The importance of education as a fundamental factor of increased labour productivity is widely emphasised in the existing literature on economic growth and development (Schultz, 1961; Becker, 1962). Therefore, it is important to promote educational attainment and human capital accumulation for long-run economic growth. This issue is particularly important for Turkey since in the European Commission (EC) Reports it is stated that despite impressive progress in recent years, there is still a low level of educational attainment and a large gender gap particularly at the secondary and higher education levels in Turkey as compared to the EU member states and the OECD countries. It is also stated that the female labour force participation rate and women's access to education are the lowest in Turkey among the EU member states and the OECD countries (EC, 2008; EC, 2010). Therefore, Turkey should continue its reform efforts in education not only to improve economic growth and development but also to obtain membership to the EU.

In this context, the aim of this study was to investigate the determinants of the educational attainment of girls and boys at the primary and secondary education levels in Turkey using the HBS for 2003 in order to shed new light on the factors behind the educational attainment process and the gender inequality in schooling. With this aim, the ordered probit model and random effects ordered probit model are estimated for separate primary and secondary school educational attainment models because the factors that are related to compulsory primary school educational attainment may differ from those related to secondary school educational attainment. The findings support the hypothesis that the educational attainment and the potential factors that are associated with gender gap in schooling differ by the level of education.

The results of this study indicate that parental education is one of the main determinants of both primary and secondary school educational attainment for both genders for the year considered in the analysis. This finding indicates the importance of the intergenerational aspect of educational attainment and endorses the need for further incentives to increase the education level of the current generation. The findings also indicate that maternal education is an important determinant of both primary and secondary school educational attainment. This finding can be seen as a further endorsement for public investment in the education of girls since the intergenerational

effects of such investments will result in greater future reductions in the gender inequality in educational attainment (Glick and Sahn, 2000).

One of the differences between the results for the primary and secondary school educational attainment models is that community characteristics are one of the main determinants of secondary school educational attainment whereas urban/rural residence is found to be statistically insignificant in the primary school educational attainment model. In addition, the marginal effects representing the community variables are greater in the girls sample which implies that the gender gap at the secondary school level may decrease with an increased level of urbanisation. In that regard, the government can potentially increase rates of secondary school educational attainment through urbanisation with improvements in the availability of schools (King and Lillard, 1983). School-building programs particularly in the least developed regions such as the eastern part of the country may decrease the cost of schooling via cheaper transportation and this may lead to higher educational attainment. Moreover, distance to the regional metropolitan centres is found to be associated with lower primary school educational attainment of children and the secondary school educational attainment of girls. This finding implies that government policies may have important effects on educational attainment by providing primary and secondary schools at critical distances from the metropolitan centres.

Consistent with the existing literature, the results of this chapter confirm the positive association between household income and child schooling both at the primary and secondary education levels, particularly for girls. In this respect, the Conditional Cash Transfer Program (CCTP) has been implemented by several developing countries including Turkey, which is based on providing money to poor families 'conditional' on certain behaviour usually related to investments in the human capital of their children such as sending children to school or taking children to health centres (Rawlings and Rubio, 2005). Turkey has been implementing the CCTP in the context of the Social Solidarity Fund (SSF) since 2001. This program targets the poorest six percent of the population and one of its aims is to improve the educational attainment of children. In this context, cash aid is provided to families on the condition that they send their children to school. In 2005, the CCTP also decided to give an extra incentive to parents who send their daughters to school in Turkey (Rawlings and Rubio, 2005). Although this program has contributed to increasing educational attainment and closing the

gender gap in Turkey, it tends to exclude the poorest and least literate families, those who need it most, since they do not have the capacity to fill the application forms or meet the pre-requisite of birth registration (McLoughney *et al*., 2007). This may result in an unequal distribution of these payments since most of aid may be given to children who arguably would have attended school anyway and may not reach the most vulnerable children and, thus, the educational attainment status of children with low propensities of school enrolment may not be changed considerably (Dayioglu *et al.*, 2009). Therefore, this program could be enhanced by considering specific targeting of the most vulnerable children to ensure that higher amounts of aid are given to these children. Moreover, the results of this chapter suggest that, in addition to income, there are other factors which are related to the low educational attainment of children.

One of these factors is having parents engaged in agriculture. The results indicate that having a father working in agriculture is negatively associated with the primary school educational attainment of boys and the secondary school educational attainment of both girls and boys. Since using child labour in agriculture is common in Turkey, mobile schools were introduced to reach children working in agriculture. However, there has been limited progress in preventing child labour in the fields (EC, 2008). Therefore, there is a need for an effective monitoring system for children working in agriculture. Moreover, there are some shortcomings of the existing labour law in Turkey as it prohibits use of child labour in the industrial sector but it does not provide protection for children working in agriculture (EC, 2008). In an indirect way, another policy recommendation may be to encourage improvements in agricultural technology or to reshape the secondary school curricula in order to make them more relevant to farmers (King and Lillard, 1983). Such changes may decrease the opportunity cost of school time or increase the expected benefits of education and encourage parents to enrol their children at school.

Another important finding of this chapter relates to the importance of household structure. The number of very young siblings (aged 0-5) is found to be negatively associated with both primary and secondary school educational attainment for both genders but this association is stronger for girls as compared to boys. This finding indicates the burden of household responsibilities, which are generally imposed more heavily on girls. Policies such as subsidised childcare, which may decrease the opportunity cost of the school time of girls and reduce the dependence of households on

the domestic labour of girls, have the potential to increase the educational attainment of girls and, thus, to close the gender inequality in educational attainment (Glick and Sahn, 2000).

Most importantly, it is possible to argue that using only the compulsory education law as a policy measure may not be sufficient in increasing the educational attainment of children and to close the gender gap since the economic structure of the country and the perceptions of society regarding education, particularly girls' education, also affect educational attainment. Limited employment opportunities for women can be seen as a barrier to the education of girls since this discourages parents from investing in their daughters' education (Hill and King, 1995). Therefore, policies should aim to increase the employment opportunities for women and to tackle gender discrimination in the labour market in Turkey. Furthermore, efforts should be directed towards changing the perceptions of women's traditional roles in society and to removing the barriers to female empowerment. Media campaigns could be designed to disseminate the value of investment in education and to inform the society about the economic and social gains from higher female education (Hill and King, 1995). From a broader perspective, macroeconomic stability is also important for Turkey to provide sustained economic growth since there is a two-way relationship between human capital accumulation and economic growth. On the one hand, human capital increases labour productivity and contributes to economic growth. On the other hand, economic growth with increased industrialisation and urbanisation may lead to changes in the occupational structure and to an increase in the demand for skilled jobs and the expected benefits of education.

Another finding of this study relates to the econometric methodology employed in this chapter. The differences between the results of the ordered probit and random effects ordered probit models indicate the importance of considering econometric issues which may bias estimation results. In this respect, this study makes an important contribution to the existing literature since it differs from the other studies for Turkey by estimating the determinants of educational attainment at the primary and secondary education levels separately while considering both the censoring problem and unobserved common family characteristics. However, there are also some limitations of this study. First, the influences of parents' preferences for favouring males for non-economic reasons or higher expected benefits from male education cannot be explored directly in the empirical analysis because of the unavailability of the relevant data. Second, supply

side variables such as the quality and quantity of schools in the relevant regions or provinces, which are other potential determinants of the demand for education and may affect access to schooling or the benefits of education, cannot be included in the analysis since information on these variables is unfortunately unavailable. However, school (or teacher) quality may affect the costs of schooling at a given level through its impact on grade repetition (Glick and Sahn, 2000). Similarly, academic performance, which indicates pupils' aptitudes to go on to the next level of education, may be a factor behind gender differences in educational attainment. Parents may consider the academic performance of their sons and daughters when allocating household resources and domestic responsibilities among them. Unfortunately, this variable is also unavailable in the data set but as Glick and Sahn (2002) argue the household level variables used in the analysis may implicitly capture these processes. Finally, the empirical analysis cannot explicitly consider the potential selection problem which arises from the fact that this study only takes into account children who live with their parents due to the unavailability of the relevant data. Individuals generally leave their parents' home for university education, which starts at age 18 in normal circumstances in Turkey. However, as in other developing countries, late entry to school or grade repetition are very common in Turkey. In this regard, as stated in Section 4.4.1 and 4.4.2, the choice of the upper age cut-off, which is 20 for the secondary education sample, arguably minimises this kind of selection bias.

In this context, future studies may investigate the implications of this potential selection problem for Turkey but this depends on data availability. Furthermore, using supply side variables such as the availability of schools in the relevant regions and provinces or school quality as well as the academic performance of children in the empirical analysis would also potentially be important contributions to the existing literature particularly for Turkey since there are no studies focusing on the supply side of the educational attainment models mostly because of the unavailability of relevant data. Moreover, in addition to the parental household and location characteristics, identification of the relationship between the specific characteristics of siblings, such as birth intervals and the educational attainment of siblings is also a potential future research area but this also depends on the availability of data.

APPENDIX TO CHAPTER 4

A4.1 A Theoretical Model of Parental Investment in the Human Capital of their Children

One of the theoretical models behind the demand for education is the human capital model, which regards education as both a consumption and an investment good to maximise lifetime wealth, developed by Schultz (1960, 1961 and 1963), Becker (1962) and Mincer (1958 and 1974). In this lifetime optimisation framework, each individual compares the expected return to education and the costs of education. Investment in education will continue up to the level at which the marginal rate of return to education equals its marginal cost. In this model, the optimal level of investment in education increases with the benefits of additional education and decreases with the cost of additional education (King and Lillard, 1983 and 1987; Behrman and Knowles, 1999; Tansel, 2002). The basic characteristic of this model is that the decision to invest in education depends both on public policies and market forces, which determine the benefits and costs of education, and on the preferences of family, income constraints and the opportunity cost of children's time (King and Lillard, 1987).

Alderman and King (1998) extended the basic model by allowing for gender differentials in the investment in the education of the children. The basic assumption of the model is that there are two periods where parents work in the first period and retire in the second period. Consumption in the first period is less than their current income and parents invest in their children's human capital in the first period and consumption in the second period is affected by the financial returns from their children (i.e., the children's human capital). Since the focus is on gender differences and, for simplicity, another assumption is that the family has two children, one of each sex. Hence, the lifetime utility function of parents is given by:

$$U = F(C_1) + G(C_2, W_m, W_f) \tag{4.8}$$

where C_1 and C_2 denote consumption in the first and second period, respectively. W_m is the future wealth of the male child and W_f is the future wealth of the female child. The returns to children's human capital and children's remittance rates are allowed to differ by gender in order to introduce market incentives. Thus, the utility function can be written as:

$$U = F(C_1) + G\left(\beta m H_m + \tau f H_f, m H_m, f H_f\right) \quad (4.9)$$

where β and τ are the rates of transfers per unit of wealth from the male and female child, respectively. H_m and H_f denote human capital where m and f are the returns to investment in human capital for male and female child, respectively. The family's budget constraint is as follows:

$$P_m H_m + P_f H_f + C_1 = Y \quad (4.10)$$

where Y is parental income and P_m and P_f are the prices of human capital for boys and girls, respectively. Parents allocate their income between current consumption and investment in their children's human capital. Parents determine the level of investment in the human capital of their son and daughter by maximising their utility (equation 4.9) subject to their budget constraint (equation 4.10). If the direct costs of education are assumed to be same for boys and girls, the expression below can be written as:

$$\frac{\partial G}{\partial C_2}\beta m + \frac{\partial G}{\partial W_m} m = \frac{\partial G}{\partial C_2}\tau f + \frac{\partial G}{\partial W_f} f \quad (4.11)$$

Equation (4.11) above shows that parents invest in the human capital of their son and daughter up to the level where the marginal benefit of the son's human capital equals the marginal benefit of the daughter's human capital. If equal remittance rates are assumed for sons and daughters and the market return to the son's human capital is greater than the return to the daughter's human capital ($m > f$), equation (4.11) is satisfied at a point at which $H_m > H_f$ (i.e., the investment in the son's human capital is higher than the investment in the daughter's human capital), since the marginal utility functions are decreasing in the level of human capital (H). Similarly, if the remittance rate of the son is higher than the daughter's ($\beta > \tau$), or if parents give more importance to their son's wealth than to their daughter's wealth ($\partial G/\partial W_m > \partial G/\partial W_f$), the marginal benefit of the son's human capital will be greater than the marginal benefit of the daughter's human capital at the same value of H (Hisarciklilar, 2002). In this case, investment in the son's human capital will be higher than that of the daughter's[84].

[84] If parents do not favor one gender over the other explicitly, the marginal utility of each child's wealth will be equal. Thus, $\partial G/\partial W_m = \partial G/\partial W_f$ and $\partial^2 G/\partial W_m \partial W_m = \partial^2 G/\partial W_f \partial W_f$ when $W_m = W_f$.

Table A4.1: Descriptive Statistics for the Continuous Variables and Percentage Distributions for the Categorical Variables

Continuous Variables				
Variables	**Primary Education (Age 14-17)**		**Secondary Education (Age 18-20)**	
	Girls	**Boys**	**Girls**	**Boys**
Household expenditure (Log)	20.2	20.2	20.2	20.3
(St. Dev.)	(0.6)	(0.6)	(0.6)	(0.6)
(Min)	17.9	18.3	18.1	18.5
(Max)	23.5	22.8	22.8	23.1
Number of rooms	3.5	3.5	3.5	3.5
(St. Dev.)	(0.8)	(0.7)	(0.7)	(0.8)
(Min)	1.0	1.0	1.0	1.0
(Max)	8.0	8.0	8.0	8.0
Number of adults	4.0	4.0	4.6	4.5
(St. Dev.)	(1.5)	(1.4)	(1.4)	(1.4)
(Min)	1.0	1.0	2.0	2.0
(Max)	14.0	14.0	14.0	14.0
Number of siblings (aged 0-5)	0.3	0.3	0.1	0.3
(St. Dev.)	(0.6)	(0.7)	(0.4)	(0.7)
(Min)	0	0	0	0
(Max)	4.0	6.0	3.0	5.0
Distance to metro centre	223.7	238.1	236.4	220.5
(St. Dev.)	(243.4)	(253.9)	(249.7)	(243.6)
(Min)	0	0	0	0
(Max)	883	883	883	883
Categorical Variables				
Father's education (reference: no educational qualification)				
No educational qualification	11.0%	9.7%	11.8%	12.1%
Primary school (5 years) graduate	52.3%	51.4%	53.2%	52.6%
Primary school (8 years) graduate	10.4%	10.8%	9.5%	9.0%
Secondary school graduate	13.3%	14.5%	11.5%	11.6%
University graduate	6.1%	6.4%	6.1%	6.3%
Mother's education (reference: no educational qualification)				
No educational qualification	34.3%	32.7%	37.3%	38.1%
Primary school (5 years) graduate	51.6%	52.9%	50.8%	49.5%
Primary school (8 years) graduate	4.6%	4.9%	4.5%	4.2%
Secondary school graduate	6.4%	6.4%	4.6%	4.9%
University graduate	2.3%	2.5%	1.8%	2.3%
Employment status of the father (reference: not working)				
Not working	21.0%	20.9%	29.1%	29.5%
Employed	43.7%	44.2%	34.7%	34.0%
Self-employed	35.3%	34.9%	36.2%	36.5%
Mother's participation in the labour market				
Yes	11.9%	12.3%	11.0%	9.7%
No	88.1%	87.7%	89.0%	90.3%

Table A4.1 continued: Descriptive Statistics for the Continuous Variables and Percentage Distributions for the Categorical Variables

Variables	Primary Education (Age 14-17)		Secondary Education (Age 18-20)	
	Girls	**Boys**	**Girls**	**Boys**
Mother engaged in agriculture				
Yes	3.5%	3.9%	4.3%	2.9%
No	96.5%	96.1%	95.7%	97.1%
Father engaged in agriculture				
Yes	17.6%	16.6%	20.7%	18.5%
No	82.4%	83.4%	79.3%	81.5%
Own dwelling				
Yes	77.4%	76.9%	82.0%	80.9%
No	22.6%	23.1%	18.0%	19.1%
Extended family				
Yes	16.6%	15.9%	18.7%	22.1%
No	83.4%	84.1%	81.3%	77.9%
Eldest resident son/daughter				
Yes	43.9%	43.0%	50.3%	51.5%
No	56.1%	57.0%	49.7%	48.5%
Urban/rural residence				
Urban	68.3%	69.8%	65.8%	69.2%
Rural	31.7%	30.2%	34.2%	30.8%
Regions				
Marmara	20.8%	21.0%	21.2%	21.2%
Aegean	11.5%	12.4%	10.9%	13.4%
Mediterranean	12.2%	11.6%	12.3%	11.3%
Central Anatolia	16.1%	15.6%	16.0%	17.9%
Black Sea	13.0%	14.1%	14.9%	12.2%
Eastern	9.0%	9.5%	9.1%	9.0%
Southeastern	17.4%	15.8%	15.6%	15.0%

Table A4.2: Estimation Results for the Ordered Probit Model (*Primary Education*)

Variables	GIRLS		BOYS	
	Estimate	St.Err.	Estimate	St.Err.
Parental Characteristics				
Father primary (5 years) school graduate	0.296***	0.060	0.224***	0.067
Father primary (8 years) school graduate	0.475***	0.090	0.289***	0.096
Father secondary school graduate	0.403***	0.091	0.254***	0.094
Father university graduate	0.513***	0.132	0.447***	0.140
Mother primary (5 years) school graduate	0.226***	0.054	0.106*	0.058
Mother primary (8 years) school graduate	0.179	0.118	0.193	0.122
Mother secondary school graduate	0.200*	0.116	0.240**	0.123
Mother university graduate	0.351*	0.202	0.360*	0.210
Father employed	-0.222***	0.061	-0.333***	0.068
Father self-employed	-0.235***	0.072	-0.323***	0.078
Mother's participation in the labour market	-0.018	0.083	-0.069	0.087
Mother engaged in agriculture	-0.063	0.138	-0.026	0.143
Father engaged in agriculture	-0.109	0.082	-0.157*	0.087
Household Characteristics				
(Ln) Household expenditure	0.136***	0.043	0.141***	0.045
Own dwelling	0.085	0.053	0.066*	0.055
Number of rooms	0.021	0.029	-0.030	0.031
Number of adults	0.174***	0.020	0.279***	0.023
Extended family	-0.339***	0.063	-0.447***	0.069
Number of young (aged 0-5) siblings	-0.165***	0.033	-0.112***	0.033
Eldest resident son/daughter	0.336***	0.052	0.442***	0.056
Community Characteristics				
Urban	-0.009	0.060	0.009	0.062
Metro	-0.0003***	0.0001	-0.0002***	0.0001
Aegean	-0.039	0.082	-0.020	0.082
Mediterranean	0.102	0.088	0.003	0.092
Central Anatolia	-0.089	0.074	0.028	0.077
Black Sea	0.067	0.086	0.085	0.087
Eastern	-0.212**	0.091	0.093	0.097
Southeastern	-0.786***	0.077	-0.709***	0.081
_cut1	1.576		1.431	
_cut2	2.980		3.093	
Log-likelihood	-2649.30		-2273.62	
Pseudo R Squared	0.1150		0.0998	
LR chi2 (28)	688.48		504.24	
N	4229		4109	

Notes: 1)***p<0.01, **p<0.05, *p<0.1.

2) Omitted category for the education status of parents is 'no education', the omitted category for the employment status of the father is 'not employed' and the omitted category for regions is the Marmara region.

Table A4.3: Marginal Effects for the Ordered Probit Model for Girls *(Primary Education)*

Variables	Illiterate	Literate	Primary -8 years-
Parental Characteristics			
Father primary (5 years) school graduate	-0.015***	-0.081***	0.096***
Father primary (8 years) school graduate	-0.016***	-0.118***	0.134***
Father secondary school graduate	-0.015***	-0.102***	0.118***
Father university graduate	-0.016***	-0.124***	0.141***
Mother primary (5 years) school graduate	-0.011***	-0.062***	0.073***
Mother primary (8 years) school graduate	-0.007*	-0.047	0.055
Mother secondary school graduate	-0.008**	-0.053*	0.061*
Mother university graduate	-0.012**	-0.088**	0.101**
Father employed	0.011***	0.061***	-0.073***
Father self-employed	0.012***	0.065***	-0.078***
Mother's participation	0.0009	0.005	-0.005
Mother engaged in agriculture	0.003	0.017	-0.021
Father engaged in agriculture	0.005	0.030	-0.036
Household Characteristics			
(Ln) Household expenditure	-0.006***	-0.037***	0.044***
Own dwelling	-0.004	-0.023	0.028
Number of rooms	-0.001	-0.005	0.007
Number of adults	-0.008***	-0.048***	0.056***
Extended family	0.021***	0.096***	-0.117***
Number of young (aged 0-5) siblings	0.008***	0.045***	-0.054***
Eldest resident son/daughter	-0.016***	-0.091***	0.108***
Community Characteristics			
Urban	0.0004	0.002	0.003
Metro	0.0001***	0.0001***	-0.0001***
Aegean	0.002	0.010	-0.012
Mediterranean	-0.004	-0.027	0.032
Central Anatolia	0.004	0.024	-0.029
Black Sea	-0.003	-0.018	0.021
Eastern	0.012**	0.060**	-0.073**
Southeastern	0.067***	0.218***	-0.285***

Notes: 1)***p<0.01, **p<0.05, *p<0.1.
2) The marginal effects are calculated at the mean values of the explanatory variables.

Table A4.4: Marginal Effects for the Ordered Probit Model for Boys *(Primary Education)*

Variables	Illiterate	Literate	Primary -8 years-
Parental Characteristics			
Father primary (5 years) school graduate	-0.004***	-0.062***	0.067***
Father primary (8 years) school graduate	-0.004***	-0.074***	0.078***
Father secondary school graduate	-0.004***	-0.066***	0.070***
Father university graduate	-0.006***	-0.107***	0.113***
Mother primary (5 years) school graduate	-0.002*	-0.029*	0.031*
Mother primary (8 years) school graduate	-0.003*	-0.050*	0.053*
Mother secondary school graduate	-0.003**	-0.062**	0.066**
Mother university graduate	-0.005***	-0.088**	0.093**
Father employed	0.007***	0.093***	-0.100***
Father self-employed	0.007***	0.092***	-0.099***
Mother's participation	0.001	0.019	-0.021
Mother engaged in agriculture	0.0005	0.007	-0.008
Father engaged in agriculture	0.003	0.045*	-0.048*
Household Characteristics			
(Ln) Household expenditure	-0.003***	-0.039***	0.042***
Own dwelling	-0.001	-0.018	0.020
Number of rooms	0.0006	0.008	-0.009
Number of adults	-0.005***	-0.077***	0.083***
Extended family	0.014***	0.133***	-0.147***
Number of young (aged 0-5) siblings	0.002***	0.031***	-0.033***
Eldest resident son/daughter	-0.009***	-0.119***	0.128***
Community Characteristics			
Urban	-0.0001	-0.002	0.002
Metro	0.00006**	0.00008***	-0.00008***
Aegean	0.0004	0.005	-0.006
Mediterranean	-0.00008	-0.001	0.001
Central Anatolia	-0.0005	-0.007	0.008
Black Sea	-0.001	-0.023	0.025
Eastern	-0.001	-0.025	0.027
Southeastern	0.028***	0.215***	-0.243***

Notes: 1)***p<0.01, **p<0.05, *p<0.1.
2) The marginal effects are calculated at the mean values of the explanatory variables.

Table A4.5: Estimation Results for the Random Effects Ordered Probit Model
(*Primary Education*)

Variables	GIRLS		BOYS	
	Estimate	St.Err.	Estimate	St.Err.
Parental Characteristics				
Father primary (5 years) school graduate	0.355***	0.084	0.253***	0.090
Father primary (8 years) school graduate	0.601***	0.127	0.346***	0.128
Father secondary school graduate	0.493***	0.126	0.299**	0.124
Father university graduate	0.633***	0.181	0.547***	0.185
Mother primary (5 years) school graduate	0.301***	0.075	0.136*	0.077
Mother primary (8 years) school graduate	0.259*	0.160	0.263*	0.160
Mother secondary school graduate	0.286*	0.157	0.302*	0.161
Mother university graduate	0.457*	0.271	0.456*	0.275
Father employed	-0.291***	0.085	-0.416***	0.092
Father self-employed	-0.306***	0.099	-0.388***	0.104
Mother's participation in the labour market	0.004	0.113	-0.079	0.114
Mother engaged in agriculture	-0.094	0.191	-0.058	0.187
Father engaged in agriculture	-0.111	0.113	-0.205*	0.116
Household Characteristics				
(Ln) Household expenditure	0.177***	0.060	0.174***	0.060
Own dwelling	0.119*	0.072	0.088	0.073
Number of rooms	0.033	0.039	-0.035	0.041
Number of adults	0.244***	0.031	0.369***	0.039
Extended family	-0.492***	0.092	-0.594***	0.099
Number of young (aged 0-5) siblings	-0.225***	0.047	-0.152***	0.045
Eldest resident son/daughter	0.480***	0.074	0.604***	0.084
Community Characteristics				
Urban	-0.005	0.082	0.018	0.082
Metro	-0.0004***	0.0001	-0.0004***	0.0001
Aegean	-0.058	0.110	-0.034	0.108
Mediterranean	0.128	0.120	0.010	0.121
Central Anatolia	-0.111	0.100	0.039	0.101
Black Sea	0.085	0.116	0.108	0.114
Eastern	-0.272**	0.126	0.124	0.129
Southeastern	-1.063***	0.119	-0.914***	0.121
_cut1	2.089*	1.169	1.727	1.169
_cut2	3.949***	1.182	3.877***	1.188
Rho	0.429***	0.059	0.398***	0.075
Log-likelihood	-2628.62		-2262.19	
LR chi2 (28)	646.43		487.48	
N	4229		4109	

Notes: 1)***p<0.01, **p<0.05, *p<0.1.

2) Omitted category for the education status of parents is 'no education', the omitted category for the employment status of the father is 'not employed' and the omitted category for regions is the Marmara region.

Table A4.6: Marginal Effects for the Random Effects Ordered Probit Model for Girls (*Primary Education*)

Variables	Illiterate	Literate	Primary -8 years-
Parental Characteristics			
Father primary (5 years) school graduate	-0.003**	-0.095***	0.099***
Father primary (8 years) school graduate	-0.006**	-0.162***	0.168***
Father secondary school graduate	-0.005**	-0.133***	0.138***
Father university graduate	-0.006**	-0.171***	0.177***
Mother primary (5 years) school graduate	-0.003**	-0.081***	0.084***
Mother primary (8 years) school graduate	-0.002	-0.070	0.072*
Mother secondary school graduate	-0.002	-0.077*	0.080*
Mother university graduate	-0.004	-0.123*	0.128*
Father employed	0.003**	0.078***	-0.081***
Father self-employed	0.003**	0.082***	-0.086***
Mother's participation	-0.0004	-0.001	0.001
Mother engaged in agriculture	0.0009	0.025	-0.026
Father engaged in agriculture	0.001	0.030	-0.031
Household Characteristics			
(Ln) Household expenditure	-0.001**	-0.047***	0.049***
Own dwelling	-0.001	-0.032*	0.033*
Number of rooms	-0.0003	-0.009	0.009
Number of adults	-0.002***	-0.065***	0.068***
Extended family	0.005***	0.133***	-0.138***
Number of young (aged 0-5) siblings	0.002***	0.060***	-0.063***
Eldest resident son/daughter	-0.005***	-0.129***	0.134***
Community Characteristics			
Urban	0.00006	0.001	-0.001
Metro	0.00004	0.0001***	-0.0001***
Aegean	0.0006	0.015	-0.016
Mediterranean	-0.001	-0.034	0.036
Central Anatolia	0.001	0.030	-0.031
Black Sea	-0.0008	-0.023	0.024
Eastern	0.002*	0.073**	-0.076**
Southeastern	0.011***	0.287***	-0.298***

Notes: 1)***p<0.01, **p<0.05, *p<0.1.
2) The marginal effects are calculated at the mean values of the explanatory variables.

Table A4.7: Marginal Effects for the Random Effects Ordered Probit Model for Boys (*Primary Education*)

Variables	Illiterate	Literate	Primary -8 years-
Parental Characteristics			
Father primary (5 years) school graduate	-0.007**	-0.061***	0.062***
Father primary (8 years) school graduate	-0.001**	-0.084***	0.085***
Father secondary school graduate	-0.008*	-0.073**	0.074**
Father university graduate	-0.001**	-0.133***	0.135***
Mother primary (5 years) school graduate	-0.0004	-0.033*	0.033*
Mother primary (8 years) school graduate	-0.0007	-0.064*	0.064*
Mother secondary school graduate	-0.0009	-0.073*	0.074*
Mother university graduate	-0.001	-0.111*	0.112*
Father employed	0.001**	0.101***	-0.102***
Father self-employed	0.001**	0.094***	-0.095***
Mother's participation	0.0002	0.019	-0.019
Mother engaged in agriculture	0.0001	0.014	-0.014
Father engaged in agriculture	0.0006	0.050*	-0.050*
Household Characteristics			
(Ln) Household expenditure	-0.005***	-0.042***	0.043***
Own dwelling	-0.002	-0.021	0.021
Number of rooms	0.0001	0.008	-0.008
Number of adults	-0.001***	-0.090***	0.091***
Extended family	0.001***	0.145***	-0.146***
Number of young (aged 0-5) siblings	0.0004	0.037***	-0.037***
Eldest resident son/daughter	-0.001***	-0.147***	0.149***
Community Characteristics			
Urban	-0.0005	-0.004	0.004
Metro	0.0001*	0.0001***	-0.0001***
Aegean	0.0001	0.008	-0.008
Mediterranean	-0.00003	-0.002	0.002
Central Anatolia	-0.0001	-0.009	0.009
Black Sea	-0.0003	-0.026	0.026
Eastern	-0.0003	-0.030	0.030
Southeastern	0.002**	0.223***	-0.225***

Notes: 1)***p<0.01, **p<0.05, *p<0.1.

2) The marginal effects are calculated at the mean values of the explanatory variables.

Table A4.8: Estimation Results for the Ordered Probit Model
(*Secondary Education*)

Variables	GIRLS		BOYS	
	Estimate	St.Err.	Estimate	St.Err.
Parental Characteristics				
Father primary (5 years) school graduate	0.251***	0.066	0.275***	0.070
Father primary (8 years) school graduate	0.796***	0.103	0.749***	0.114
Father secondary school graduate	0.946***	0.108	0.782***	0.116
Father university graduate	1.150***	0.164	0.854***	0.174
Mother primary (5 years) school graduate	0.460***	0.057	0.305***	0.060
Mother primary (8 years) school graduate	0.703***	0.144	0.586***	0.162
Mother secondary school graduate	0.464***	0.157	0.707***	0.187
Mother university graduate	0.467*	0.275	0.663***	0.255
Father employed	-0.127**	0.065	-0.098	0.070
Father self-employed	-0.042	0.077	-0.060	0.080
Mother's participation in the labour market	-0.069	0.099	-0.052	0.106
Mother engaged in agriculture	-0.209	0.151	-0.215	0.179
Father engaged in agriculture	-0.390***	0.093	-0.298***	0.098
Household Characteristics				
(Ln) Household expenditure	0.281***	0.047	0.245***	0.051
Own dwelling	-0.001	0.064	0.174***	0.067
Number of rooms	0.047	0.031	0.062*	0.033
Number of adults	-0.112***	0.021	-0.061**	0.024
Extended family	-0.082	0.066	-0.033	0.071
Number of young (aged 0-5) siblings	-0.230***	0.051	-0.133***	0.040
Eldest resident son/daughter	-0.088	0.056	-0.077	0.059
Community Characteristics				
Urban	0.224***	0.067	0.242***	0.074
Metro	-0.0004***	0.0001	-0.0001	0.0001
Aegean	-0.242***	0.086	0.081	0.089
Mediterranean	0.431***	0.096	0.056	0.104
Central Anatolia	0.209***	0.081	0.067	0.083
Black Sea	0.276***	0.089	0.229**	0.101
Eastern	-0.029	0.099	0.484***	0.113
Southeastern	-0.618***	0.088	-0.248***	0.094
_cut1	3.448		2.992	
_cut2	3.951		3.493	
_cut3	5.385		4.770	
_cut4	5.988		5.572	
Log-likelihood	-2712.36		-2273.59	
Pseudo R Squared	0.1931		0.1293	
LR chi2 (28)	1298.08		675.35	
N	2610		2268	

Notes: 1)***p<0.01, **p<0.05, *p<0.1.

2) Omitted category for the education status of parents is 'no education', the omitted category for the employment status of the father is 'not employed' and the omitted category for regions is the Marmara region.

Table A4.9: Marginal Effects for the Ordered Probit Model for Girls (*Secondary Education*)

Variables	Illiterate	Literate	Primary (5 years)	Primary (8 years)	Secondary
Parental Chr.					
Father primary (5 years) school graduate	-0.006***	-0.011***	-0.073***	-0.007***	0.098***
Father primary (8 years) school graduate	-0.009***	-0.021***	-0.202***	-0.071***	0.305***
Father secondary school graduate	-0.010***	-0.024***	-0.233***	-0.088***	0.356***
Father university graduate	-0.010***	-0.023***	-0.256***	-0.121***	0.411***
Mother primary (5 years) school graduate	-0.011***	-0.021***	-0.133***	-0.014***	0.179***
Mother primary (8 years) school graduate	-0.008***	-0.018***	-0.180***	-0.064***	0.271***
Mother secondary school graduate	-0.006***	-0.014***	-0.127***	-0.034**	0.183***
Mother university graduate	-0.006***	-0.014***	-0.127**	-0.036	0.184*
Father employed	0.003*	0.005*	0.037**	0.003**	-0.050**
Father self-employed	0.001	0.001	0.012	0.001	-0.016
Mother's participation	0.001	0.003	0.020	0.001	-0.027
Mother engaged in agriculture	0.006	0.010	0.061	0.002	-0.080
Father engaged in agriculture	0.012***	0.021***	0.113***	0.002	-0.149***
Household Chr.					
(Ln) Household exp.	-0.006***	-0.012***	-0.082***	-0.009***	0.111***
Own dwelling	0.0003	0.0006	0.0004	0.0005	-0.0006
Number of rooms	-0.001	-0.002	-0.013	-0.001	0.018
Number of adults	0.002***	0.005***	0.032***	0.003***	-0.044***
Extended family	0.002	0.003	0.024	0.002	-0.032
Number of young (aged 0-5) siblings	0.005***	0.010***	0.067***	0.007***	-0.090***
Eldest resident child	0.002	0.003	0.025	0.002	-0.034
Community Chr.					
Urban	-0.005***	-0.010***	-0.065***	-0.005***	0.087***
Metro	0.0001***	0.002***	0.0001***	0.001***	-0.0001***
Aegean	0.007**	0.012**	0.071***	0.002*	-0.093***
Mediterranean	-0.007***	-0.014***	-0.120***	-0.028***	0.170***
Central Anatolia	-0.004***	-0.008***	-0.060***	-0.010*	0.083***
Black Sea	-0.005***	-0.010***	-0.079***	-0.014**	0.109***
Eastern	0.0007	0.001	0.008	0.0008	-0.011
Southeastern	0.024***	0.038***	0.173***	-0.009	-0.227***

Notes: 1)***p<0.01, **p<0.05, *p<0.1.

2) The marginal effects are calculated at the mean values of the explanatory variables.

Table A4.10: Marginal Effects for the Ordered Probit Model for Boys (*Secondary Education*)

Variables	Illiterate	Literate	Primary (5 years)	Primary (8 years)	Secondary
Parental Chr.					
Father primary (5 years) school graduate	-0.003***	-0.007***	-0.064***	-0.034***	0.109***
Father primary (8 years) school graduate	-0.004***	-0.011***	-0.133***	-0.122***	0.272***
Father secondary school graduate	-0.004***	-0.012***	-0.140***	-0.127***	0.284***
Father university graduate	-0.004***	-0.011***	-0.142***	-0.142***	0.300***
Mother primary (5 years) school graduate	-0.003***	-0.008***	-0.070***	-0.038***	0.120***
Mother primary (8 years) school graduate	-0.003***	-0.009***	-0.108***	-0.095***	0.217***
Mother secondary school graduate	-0.003***	-0.010***	-0.124***	-0.117***	0.256***
Mother university graduate	-0.003***	-0.009***	-0.117***	-0.110**	0.241***
Father employed	0.001	0.002	0.023	0.012	-0.039
Father self-employed	0.0007	0.001	0.014	0.007	-0.024
Mother's participation	0.0006	0.001	0.012	0.006	-0.021
Mother engaged in agriculture	0.003	0.007	0.052	0.022	-0.085
Father engaged in agriculture	0.004**	0.009**	0.073***	0.031***	-0.118***
Household Chr.					
(Ln) Household exp.	-0.002***	-0.006***	-0.057***	-0.031***	0.097***
Own dwelling	-0.002**	-0.005**	-0.041**	-0.019***	0.069***
Number of rooms	-0.0007*	-0.001*	-0.014*	-0.007*	0.024*
Number of adults	0.0007**	0.001**	0.014**	0.007**	-0.024**
Extended family	0.0004	0.0009	0.007	0.004	-0.013
Number of young (aged 0-5) siblings	0.001***	0.003***	0.030***	0.016***	-0.052***
Eldest resident child	0.0009	0.002	0.017	0.009	-0.030
Community Chr.					
Urban	-0.003**	-0.007***	-0.057***	-0.028***	0.096***
Metro	0.00001	0.00002	0.00002	0.00001	-0.00004
Aegean	-0.0008	-0.002	-0.018	-0.010	0.032
Mediterranean	-0.0006	-0.001	-0.012	-0.007	0.022
Central Anatolia	-0.0007	-0.001	-0.015	-0.008	0.026
Black Sea	-0.002**	-0.005**	-0.049**	-0.032**	0.089**
Eastern	-0.003***	-0.009***	-0.095***	-0.075***	0.184***
Southeastern	0.003**	0.008**	0.060**	0.026***	-0.099***

Notes: 1)***p<0.01, **p<0.05, *p<0.1.

2) The marginal effects are calculated at the mean values of the explanatory variables.

Table A4.11: Estimation Results for the Random Effects Ordered Probit Model (*Secondary Education*)

Variables	GIRLS		BOYS	
	Estimate	St.Err.	Estimate	St.Err.
Parental Characteristics				
Father primary (5 years) school graduate	0.455***	0.130	0.417***	0.116
Father primary (8 years) school graduate	1.417***	0.218	1.149***	0.204
Father secondary school graduate	1.703***	0.232	1.210***	0.212
Father university graduate	2.048***	0.336	1.286***	0.295
Mother primary (5 years) school graduate	0.860***	0.122	0.469***	0.104
Mother primary (8 years) school graduate	1.337***	0.288	0.882***	0.266
Mother secondary school graduate	0.902***	0.302	1.094***	0.308
Mother university graduate	0.865*	0.511	1.013**	0.407
Father employed	-0.281**	0.125	-0.144	0.112
Father self-employed	-0.097	0.147	-0.099	0.127
Mother's participation in the labour market	-0.079	0.190	-0.058	0.167
Mother engaged in agriculture	-0.443	0.290	-0.360	0.283
Father engaged in agriculture	-0.756***	0.185	-0.441***	0.160
Household Characteristics				
(Ln) Household expenditure	0.520***	0.096	0.375***	0.087
Own dwelling	0.013	0.122	0.251**	0.108
Number of rooms	0.079	0.060	0.115**	0.054
Number of adults	-0.217***	0.044	-0.100**	0.040
Extended family	-0.142	0.127	-0.053	0.113
Number of young (aged 0-5) siblings	-0.453***	0.104	-0.188***	0.065
Eldest resident son/daughter	-0.221**	0.098	-0.106	0.089
Community Characteristics				
Urban	0.390***	0.128	0.394***	0.122
Metro	-0.0008***	0.000	-0.0001	0.000
Aegean	-0.403**	0.165	0.122	0.141
Mediterranean	0.780***	0.189	0.092	0.164
Central Anatolia	0.376**	0.155	0.119	0.132
Black Sea	0.531***	0.174	0.349**	0.163
Eastern	-0.013	0.189	0.759***	0.191
Southeastern	-1.098***	0.183	-0.406***	0.154
_cut1	6.282***	1.805	4.568***	1.667
_cut2	7.211***	1.820	5.345***	1.683
_cut3	9.813***	1.876	7.326***	1.739
_cut4	10.920***	1.905	8.574***	1.782
Rho	0.702***	0.041	0.585***	0.073
Log-likelihood	-2264.27		-2257.15	
LR chi2 (28)	1181.00		644.91	
N	2610		2268	

Notes: 1)***p<0.01, **p<0.05, *p<0.

2) Omitted category for the education status of parents is 'no education', the omitted category for the employment status of the father is 'not employed' and the omitted category for regions is the Marmara region.

Table A4.12: Marginal Effects for the Random Effects Ordered Probit Model for Girls (*Secondary Education*)

Variables	Illiterate	Literate	Primary (5 years)	Primary (8 years)	Secondary
Parental Chr.					
Father primary (5 years) school graduate	-0.0001**	-0.004***	-0.128***	-0.046***	0.174***
Father primary (8 years) school graduate	-0.0004**	-0.001***	-0.398***	-0.143***	0.544***
Father secondary school graduate	-0.0005**	-0.001***	-0.479***	-0.172***	0.654***
Father university graduate	-0.0006**	-0.002***	-0.576***	-0.207***	0.786***
Mother primary (5 years) school graduate	-0.0002**	-0.009***	-0.242***	-0.087***	0.330***
Mother primary (8 years) school graduate	-0.0003**	-0.001***	-0.376***	-0.135***	0.513***
Mother secondary school graduate	-0.0002**	-0.009***	-0.254***	-0.091**	0.346***
Mother university graduate	-0.0002	-0.009**	-0.243*	-0.087	0.332*
Father employed	0.0008	0.003	0.079**	0.028**	-0.108**
Father self-employed	0.00002	0.001	0.027	0.009	-0.037
Mother's participation	0.00002	0.0008	0.022	0.008	-0.030
Mother engaged in agriculture	0.00001	0.004	0.124	0.044	-0.170
Father engaged in agriculture	0.0002**	0.0008**	0.212***	0.076***	-0.290***
Household Chr.					
(Ln) Household exp.	-0.001***	-0.005***	-0.146***	-0.052***	0.199***
Own dwelling	-0.0002	-0.002	-0.003	-0.001	0.005
Number of rooms	-0.002	-0.008	-0.022	-0.008	0.030
Number of adults	0.0006**	0.002***	0.061***	0.022***	-0.083***
Extended family	0.0004	0.001	0.040	0.014	-0.054
Number of young (aged 0-5) siblings	0.001***	0.004***	0.127***	0.045***	-0.174***
Eldest resident child	0.0006**	0.002*	0.062**	0.022**	-0.085**
Community Chr.					
Urban	-0.0001**	-0.004***	-0.110***	-0.039**	0.150***
Metro	0.0001**	0.001***	0.0002***	0.008***	-0.003***
Aegean	0.0001*	0.004**	0.113**	0.040**	-0.155**
Mediterranean	-0.0002**	-0.008***	-0.219***	-0.079***	0.299***
Central Anatolia	-0.0001*	-0.004**	-0.106**	-0.038**	0.144**
Black Sea	-0.0001*	-0.005**	-0.149***	-0.053**	0.204***
Eastern	0.00007	0.0001	0.003	0.001	-0.005
Southeastern	0.003***	0.001***	0.309***	0.111***	-0.421***

Notes: 1)***p<0.01, **p<0.05, *p<0.1.

2) The marginal effects are calculated at the mean values of the explanatory variables.

Table A4.13: Marginal Effects for the Random Effects Ordered Probit Model for Boys (*Secondary Education*)

Variables	Illiterate	Literate	Primary (5 years)	Primary (8 years)	Secondary
Parental Chr.					
Father primary (5 years) school graduate	-0.0003**	-0.005***	-0.064***	-0.100***	0.165***
Father primary (8 years) school graduate	-0.0009**	-0.001***	-0.177***	-0.275***	0.455***
Father secondary school graduate	-0.0009**	-0.001***	-0.187***	-0.290***	0.479***
Father university graduate	-0.0001**	-0.001***	-0.199***	-0.308***	0.509***
Mother primary (5 years) school graduate	-0.0003**	-0.006***	-0.072***	-0.112***	0.185***
Mother primary (8 years) school graduate	-0.0007**	-0.001***	-0.136***	-0.211***	0.349***
Mother secondary school graduate	-0.0008**	-0.001***	-0.169***	-0.262***	0.433***
Mother university graduate	-0.0008**	-0.001**	-0.156**	-0.242**	0.401**
Father employed	0.0001	0.001	0.022	0.034	-0.057
Father self-employed	0.00007	0.0001	0.015	0.023	-0.039
Mother's participation	0.00004	0.00008	0.009	0.013	-0.023
Mother engaged in agriculture	0.0002	0.004	0.055	0.086	-0.142
Father engaged in agriculture	0.0003**	0.0006**	0.068***	0.105**	-0.174***
Household Chr.					
(Ln) Household exp.	-0.003***	-0.005***	-0.058***	-0.090***	0.148***
Own dwelling	-0.0002*	-0.003**	-0.038**	-0.060**	0.099**
Number of rooms	-0.0009*	-0.001**	-0.017**	-0.027**	0.045**
Number of adults	0.0008*	0.001**	0.015**	0.024**	-0.039**
Extended family	0.0004	0.0007	0.008	0.012	-0.021
Number of young (aged 0-5) siblings	0.0001**	0.002*	0.029***	0.045***	-0.074***
Eldest resident child	0.0008	0.001	0.016	0.025	-0.042
Community Chr.					
Urban	-0.003**	-0.005***	-0.061***	-0.094***	0.156***
Metro	0.00001	0.00002	0.00002	0.00003	-0.00006
Aegean	-0.0009	-0.001	-0.018	-0.029	0.048
Mediterranean	-0.0007	-0.001	-0.014	-0.022	0.036
Central Anatolia	-0.0009	-0.001	-0.018	-0.028	0.047
Black Sea	-0.0002*	-0.004**	-0.054**	-0.083**	0.138**
Eastern	-0.0006**	-0.001***	-0.117***	-0.182***	0.300***
Southeastern	0.003**	0.005***	0.062***	0.097**	-0.160***

Notes: 1)***p<0.01, **p<0.05, *p<0.1.

2) The marginal effects are calculated at the mean values of the explanatory variables.

CHAPTER 5: CONCLUSION

Turkey has experienced major policy reforms in the health and education sectors over the recent decade. These reforms have particularly focused on issues such as financial protection against health care costs, decreasing the smoking prevalence rate and closing the gender gap in educational attainment. One of the main reasons of the major policy reforms in these areas may be attributed to the importance to the country of enhancing human capital. Additionally, Turkey, as a candidate country for full membership of the EU, has been trying to fulfil the EU requirements and these issues have been among the priority areas in the accession negotiations. In this context, the overall aim of this thesis was to examine and analyse these three specific issues in health and education, which have been attracted increasing attention on the policy agenda in Turkey, particularly over the recent decade.

The first empirical study presented in Chapter 2 examined the prevalence of 'catastrophic' out-of-pocket health care expenditure in Turkey and the factors associated with its risk, with particular focus placed upon the association between the poverty status of the household and the likelihood of incurring catastrophic health expenditure. The Household Budget Surveys from 2002 to 2008 were used in the analysis. A standard probit model was estimated for the pooled data set as a first step. In addition, the Sartori selection (Sartori, 2003) and the Heckman selection (Heckman, 1979) models were estimated to account for the potential selection problem which may arise from the fact that many households may choose not to seek health care at all due to high health care costs. There is no consensus, however, in terms of the definition of catastrophic health expenditure in the existing literature. Therefore, a number of threshold levels, ranging from 2.5% to 20% for the denominator of total household expenditure and from 10% to 40% for the denominator of total non-food expenditure, were used to provide a comprehensive picture of catastrophic health expenditure and to explore the sensitivity of the results. The distribution of catastrophic health expenditure by survey years indicated that the proportions of households incurring catastrophic health expenditure were quite similar across the years with the exception of 2008. The proportion of households experiencing catastrophic health expenditure decreased in 2008, which is the year that the Universal Health Insurance system was implemented.

From 2002 to 2008, nearly 5% of households spent at least 10% of their income on health care which is equivalent to more than 3.5 million individuals.

The empirical analysis presented in Chapter 2 sheds light on the household characteristics associated with catastrophic health expenditure and this may help policy-makers in policy design such as targeting vulnerable households, which have specific risk factors, with for example extension of insurance coverage or the provision of subsidised health care services. The main finding of this chapter suggests that poor households are less likely to seek health care relative to non-poor households. This finding highlights the vulnerability of poor households in terms of their health care seeking position and implies that special attention in Turkey should be devoted to overcome the health cost barriers for poor households. However, the results indicate that poor households are less likely to incur catastrophic health expenditure as compared to non-poor families even after accounting for the medical care seeking behaviour of households. This finding may be attributed to the fact that rich households in Turkey generally prefer to use private facilities since public health services are often criticised as being unsatisfactory, because of, for example, overcrowding in public hospitals and a perceived relatively low quality of health care. After the health reforms, patients from all insurance institutions gained access to outpatient and inpatient services in contracted private hospitals. However, in order to encourage private sector interest in contracting with the Social Security Institution, private hospitals were allowed to implement an extra charge (OECD, 2008). Since this extra cost was decreased as a result of the reforms, this improvement in access to private facilities may create demand inducement among non-poor households and may lead to a higher probability of incurring catastrophic health expenditure. However, this hypothesis cannot be explored in the empirical analysis due to the data limitations.

The results further indicate that the households with a disabled or ill member and households with more preschool or elderly members are more likely to seek health care and are more likely to incur catastrophic health expenditure. Thus, it can be argued that providing exemptions from health care costs particularly for households that have members belonging to the more risky groups in terms of health status would be a possible move towards preventing catastrophic health expenditure. Moreover, health insurance is found to be an important protection factor against the probability of catastrophic health expenditure, which implies that the universal health insurance in

Turkey should continue to expand its insurance coverage mechanisms, particularly for the groups which currently face a lack of formal coverage such as informal-sector workers. In addition, higher levels of education and living in an urban area are found to be inversely associated with experiencing catastrophic health expenditure. These findings are relatively robust to changes in the definition of catastrophic health expenditure (i.e., different threshold levels and denominators). The risk factors associated with catastrophic health expenditure are also consistent across the three estimation methods used in the analysis, namely the probit model, the Sartori selection model and the Heckman selection model.

It is important to note that there are some limitations of the empirical analysis presented in this chapter. In the selection models, namely the Sartori selection and Heckman selection models, the aim is to analyse the risk of catastrophic health expenditure while taking account of the households who needed health care but could not afford it. However, the samples used in the selection models cannot be limited to those who 'needed' health care because of the unavailability of a measure of the health status of individuals in the data set. In this respect, the empirical analysis presented in Chapter 2 could be potentially improved by restricting the samples to those who 'needed' health care while accounting for the potential selection problem once the necessary data becomes available. Furthermore, the 'having or not having positive health expenditure' dichotomy, which is the only information related to the medical care seeking behaviour of households in the data set, is modelled in the selection equation. Since, even under full public health insurance coverage, individuals must pay a co-payment for drug expenditure in Turkey, it can be argued that having positive health expenditure provides information on medical visits (Erus and Aktakke, 2012). This dichotomy does not allow for the exploration of some important dimensions of the health care seeking behaviour of households such as the type of provider (i.e., public or private) or the number of household members seeking health care. However, data capturing such dimensions of health care seeking behaviour are not available. Another limitation is that the indirect costs of illness such as lost working time and the financial coping strategies used by households to deal with high health care costs, such as using credit or savings, borrowing from friends or relatives, cannot be explored in the analysis again due to the lack of longitudinal data. If the relevant longitudinal data becomes available,

investigating the effects of catastrophic health expenditure over time may be an important contribution to the existing literature.

Despite these limitations, however, the empirical analysis presented in Chapter 2 is important in terms of providing an analysis of health care expenditure in excess of a substantial fraction of the household's budget in the context of the economic consequences of illness. Moreover, this study is the first attempt to investigate the determinants of catastrophic health expenditure in Turkey in the context of attempting to allow for the medical care seeking behaviour of households. This study has further extended the existing literature by using the Sartori selection model (Sartori, 2003) to investigate the implications of the potential selection bias arising from the health care seeking position of households.

The second empirical study presented in Chapter 3 investigated the determinants of adult smoking propensity and intensity from a gender perspective. The Global Adult Tobacco Survey for 2008 was used to estimate negative binomial, zero-inflated negative binomial and two-part/hurdle count data models. The main results of Chapter 3 are twofold. First, the factors distinguishing potential smokers from non-smokers were found to differ from the factors associated with the level of cigarette consumption. Second, the findings highlight the potential effectiveness of specific tobacco control policies for males and females since the determinants of cigarette demand were found to differ between genders. An increase in cigarette prices is found to be associated with a decrease in the probability of being a smoker for both genders while females are found to be more sensitive to changes in cigarette prices. This finding may also reflect the indirect effect of having a lower income for females compared to males. Furthermore, the findings indicate a negative relationship between cigarette prices and the level of cigarette consumption for females, whereas cigarette prices are not a statistically significant determinant of the level of cigarette consumption for males. In this respect, using taxes as a policy instrument may be particularly effective in reducing cigarette consumption among females.

The gender specific findings indicate arguably the most interesting difference between males and females in the relationship between education and smoking behaviour. There is an inverse relationship between education and smoking prevalence and intensity for males while the opposite relationship is found to exist for females. In a similar vein,

perception or knowledge of the health risks of smoking are found to be inversely associated with the level of cigarette consumption for males while they have no effect on the smoking behaviour of females. These findings imply that the anti-smoking policies using education as the principal instrument may not be effective in decreasing smoking prevalence among females. The findings also indicate that tobacco advertising is still noticed by individuals despite the fact that tobacco advertising was illegal at the time of the survey. In this regard, the findings reinforce the hypothesis that pro-cigarette marketing has no effect on the decision to start smoking since exposure to tobacco advertising is found to have no effect on smoking participation for both males and females. However, the findings indicate a positive association between exposure to tobacco advertising and smoking intensity for males, which highlights the need for better enforcement. In the context of these findings, it can be argued that promoting and improving public awareness programs on the adverse health effects of tobacco use and the tobacco advertising bans may be effective anti-smoking policies if the particular focus is to decrease cigarette consumption among males. The findings further indicate that younger male adults have a higher probability of smoking compared to older male adults whereas the opposite relationship is found to exist for females. This finding arguably implies that males are more likely to participate in cigarette smoking at an early age as compared to females. In this respect, educational campaigns about the health dangers of tobacco use should be disseminated to schools to prevent initiation of smoking for males, in particular.

Perhaps most importantly, the presence of other smokers in the household is found to be one of the main determinants of both smoking propensity and intensity for both genders. This indicates the importance of social interaction for smoking behaviour and supports the fact that Turkey should continue to create a tobacco-free environment by prohibiting smoking in all indoor public places, places of employment, public transportation and places of entertainment. This policy is important not only for preventing exposure to second-hand smoking but also for reducing the effect of social interaction on smoking behaviour. It should, however, be noted that the variable indicating the presence of other smokers in the household does not capture interaction outside the household and it may lead to a potential simultaneity problem as members of the same household may affect each other (Jones, 1994). Another limitation of the study is related to the cross-sectional nature of the survey used in this chapter. Using cross-sectional data does not allow for

accounting for the inter-temporal dependence of tobacco consumption, which is an important characteristic of such addictive behaviour. Furthermore, some of the variables that are commonly used in the existing literature such as the marital status and health status of individuals and alcohol consumption cannot be included in the analysis due to the unavailability of relevant data.

Notwithstanding these limitations, this study has provided empirical evidence for the existence of gender differences in cigarette consumption patterns in Turkey. Furthermore, the empirical study presented in Chapter 3 is the first attempt to focus on gender differences in cigarette consumption in Turkey as well as to examine cigarette consumption using individual level rather than household level or aggregate time-series data. Implementing a successful tobacco control program to decrease the rate of cigarette smoking and to curb the increase in female smoking is particularly important for Turkey. In this context, future studies might improve this study by focusing on the inter-temporal dependence of cigarette consumption from a gender perspective to provide further policy implications if the relevant longitudinal data becomes available.

The third empirical study presented in Chapter 4 examined the educational attainment of girls and boys at the primary and secondary education levels in Turkey focusing on the roles of parental, household and community characteristics. The Household Budget Survey for 2003 was used to estimate ordered probit and random effects ordered probit models. There were two main reasons for estimating the models for the primary and secondary school educational attainment, separately. The first one is to overcome the censoring problem, which arises from the fact that for children, who were enrolled at school at the time of the survey, final educational attainment is unknown. The samples were restricted to only include children above the graduation ages. Secondly, estimation using two different samples allows for investigating the hypothesis that the potential factors associated with gender inequality differ by the level of education. The main finding of this chapter indicates that the factors that are related to compulsory primary school educational attainment differ in some respects to those related to secondary school educational attainment and that the factors related to gender inequality differ by the level of education. The intergenerational aspect of educational attainment is confirmed by the findings since parental education is found to be one of the main determinants of the educational attainment of children. This finding highlights the importance of increasing the education level of the current generation. Additionally, the

results confirm the importance of maternal education particularly in the education of girls although the effect of father's education is greater for both genders, which suggests that public policies should devote attention to the education of girls in order to close the gender gap in the future.

The results presented in Chapter 4 also demonstrate the positive association between household income and the educational attainment of children at all levels of education and this relationship is found to be strongest for girls. In this respect, the assistance programs targeting poor families, such as cash transfer programs, may help to increase educational attainment by removing one of the most important barriers to the education of girls. Another barrier to the schooling of children in Turkey relates to the use of child labour in the family business, farm or with housework, which is a pervasive problem. The findings indicate that having a self-employed father or having a father engaged in agriculture is negatively associated with the primary school attainment of boys. Similarly, girls who have fathers engaged in agricultural work are found to be less likely to go beyond compulsory primary school educational attainment. These findings suggest that efforts should be directed towards curbing the use of child labour in agriculture or family businesses through, for example, an effective monitoring system for these children and comprehensive legal protection for child labour. In a similar vein, the number of very young siblings (aged 0-5) is found be negatively associated with the educational attainment of children, particularly in the case of the secondary school educational attainment of girls. This finding supports the fact that girls are more likely to be expected to perform household tasks such as taking care of younger siblings and this may increase the opportunity cost of their schooling time. Therefore, the government should focus not only on cash transfer programs but also on policy instruments that release time for girls' schooling such as establishing subsidised child care centres.

The findings from Chapter 4 further indicate that community characteristics are among the most important determinants of the secondary school educational attainment of children, particularly for girls. Living in a rural area and living in the eastern part of the country are both found to be associated with a lower probability of secondary school educational attainment. It is also found that, as the distance to a metropolitan centre increases, girls are less likely to attend secondary school. These findings suggest that the government can potentially increase secondary school educational attainment

through urbanisation and school-building programs in the economically disadvantaged areas such as in the Eastern and Southeastern regions, rural areas and at the critical distances from the metropolitan centres. Another main finding is related to the econometric methodology employed in this chapter. The results for the random effects ordered probit models are found to differ from the results for the ordered probit models in terms of different signs and statistical significance for some variables as well as greater marginal effects. This highlights the importance of accounting for unobserved family characteristics for children coming from the same household in the analysis.

Another important issue in educational attainment studies is related to the potential selection problem, which arises in the case of including only children who live with their parents in the analysis. In this respect, one of the limitations of this study is that the potential selection problem cannot be explicitly considered due to the unavailability of the relevant data. However, it can be argued that using a relatively low upper age bound for secondary school educational attainment may serve to minimise this kind of selection bias. Another limitation of this study is that supply side variables, such as the quality and quantity of schools in the regions and provinces as well as the academic performance of children, which are potential determinants of the demand for education, cannot be included in the analysis. Furthermore, the influences of parents' preferences such as favouring males for non-economic reasons or higher expected benefits from male education cannot be explored directly in the empirical analysis since, once again, information on these variables is unfortunately unavailable. For future studies, investigating the effects of such variables on the educational attainment of children may be an important research area.

Notwithstanding these limitations, the empirical study presented in Chapter 4 is important in terms of providing empirical evidence in support of the hypothesis that the potential factors associated with the educational attainment of girls and boys differ across the primary and secondary school education levels. Furthermore, this study differs from the existing studies on educational attainment for Turkey by accounting for the censoring problem and the influence of unobserved family characteristics, which are important econometric issues in modelling educational attainment. Moreover, this study extends the existing literature by using the Household Budget Survey for the year 2003, which provides information on the residence of the household at the regional and province level. Thus, this data set has enabled the exploration of the association

between educational attainment and a range of community characteristics such as urban/rural residence, region dummy variables and distance to a metropolitan centre in the empirical analysis.

In conclusion, the empirical studies presented in this thesis have provided a number of interesting insights into three specific issues related to health and education, which have been on and remain on the policy agenda in Turkey. Thus, the empirical findings presented in this thesis may be useful for policy making in these areas. Moreover, health and education have been the subject of many economic studies for both developing and developed countries because of the close relationship between economic growth and development and improvements in the health and education of the population. In this respect, research on health and education is of critical importance for Turkey since the government aims to improve the human capital of the country, and in so doing, increase economic growth and development. However, there are only a few studies that have focused on such issues in Turkey, particularly from a microeconometric perspective. In this context, this thesis plays an important role in filling this gap in the existing literature.

REFERENCES

Adaman, F. (2003), Study on the Social Protection Systems in Turkey. Entry for the report study on the social protection systems in the 13 applicant countries, prepared for European Commission 2003.

Alderman, H. and E.M. King (1998), "Gender Differences in Parental Investment in Education", *Structural Change and Economic Dynamics*, 9(4): 453-468.

Al-Samarrai, S. and B. Reilly (2000), "Urban and Rural Differences in Primary School Attendance: An Empirical Study for Tanzania", *Journal of African Economics*, 9(4): 430-474.

Al-Samarrai, S. and T. Peasgood (1998), "Educational Attainments and Household Characteristics in Tanzania", *Economics of Education Review*, 17(4): 395-417.

Arabmazard, A. and P. Schmidt (1982), "An Investigation of the Robustness of the Tobit Estimator to Non-Normality", *Econometrica*, 50(4): 1055-1063.

Aristei, D. and L. Pieroni (2008), "A Double-Hurdle Approach to Modelling Tobacco Consumption in Italy", *Applied Economics*, 40(19): 2463-2476.

Aristei, D. and L. Pieroni (2009), "Addiction, Social Interactions and Gender Differences in Cigarette Consumption", *Empirical Economics*, 36(3): 245-272.

Atkinson, A., Gomulka, J. and N. Stern (1984), "Household Expenditure on Tobacco 1970-1980: Evidence from the Family Expenditure Survey", Discussion Paper in Taxation Incentives and the Distribution of Income No: 57, London School of Economics.

Attanayake, N., Fox-Rushby, J. and A. Mills (2000), "Household Costs of 'Malaria' Morbidity: A Study in Matale District, Sri Lanka", *Tropical Medicine and International Health*, 5(9): 595-606.

Baeza, C. (2002), Towards decent work: Social protection in health for all workers and their families. Conceptual framework for the extension of social protection in health. Geneva: International Labour Organisation.

Baltagi, B.H. and D. Levin (1986), "Estimating Dynamic Demand for Cigarettes using Panel Data: The Effects of Bootlegging, Taxation and Advertising Reconsidered", *Review of Economics and Statistics*, 68(1): 148-155.

Baltagi, B.H. and J.M. Griffin (2001), "The Econometrics of Rational Addiction", *Journal of Business and Economic Statistics*, 19(4): 449-454.

Barros, R. and D. Lam (1996), Income Inequality, Inequality in Education, and Children's Schooling Attainment in Brazil, In *Opportunity Foregone: Education in Brazil*, Birdsall, N. and R. Sabot (eds.), The Inter-American Development Bank.

Bauer, T.K., Gohlmann, S. and M. Sinning (2007), "Gender Differences in Smoking Behaviour", *Health Economics*, 16(9): 895-909.

Becker, G. and K.M. Murphy (1988), "A Theory of Rational Addiction", *Journal of Political Economy*, 96(4): 675-700.

Becker, G., Grossman, M. and K.M. Murphy (1994), "An Empirical Analysis of Cigarette Addiction", *American Economic Review*, 84(3): 396-418.

Becker, G.S. (1960), An Economic Analysis of Fertility, In *Demographic and Economic Change in Developed Countries*, Becker, G.S. (eds.), Princeton University Press, Princeton.

Becker, G.S. (1962), "Investment in Human Capital: A Theoretical Analysis", *Journal of Political Economy*, 70(5): 9-49.

Becker, G.S. (1974), "A Theory of Social Interactions", *Journal of Political Economy*, 82(6): 1063-1093.

Becker, G.S. and G.H. Lewis (1973), "On the Interaction between Quantity and Quality of Children", *Journal of Political Economy*, 81(2): 279-288.

Behrman, J., Khan, S., Ross, D. and R. Sabot (1997), "School Quality and Cognitive Achievement Production: A Case Study for Rural Pakistan", *Economics of Education Review*, 16(2): 127-142.

Behrman, J.R. and B. Wolfe (1987), "Investment in Schooling in Two Generations in Pre-revolutionary Nicaragua: The Roles of Family Background and School Supply", *Journal of Development Economics*, 27(1-2): 395-419.

Behrman, J.R. and J.C. Knowles (1999), "Household Income and Child Schooling in Vietnam", *World Bank Economic Review*, 13(2): 211-256.

Behrman, J.R. and M.R. Rosenweig (1994), "Caveat Emptor: Cross-Country Data on Education and the Labour Force", *Journal of Development Economics*, 44(2): 147-171.

Berki, S.E. (1986), "A Look at Catastrophic Medical Expenses and the Poor", *Health Affairs*, 5(4): 139-145.

Bilgic, A., Florkowski, W.J. and C. Akbay (2009), "Demand for Cigarettes in Turkey: An Application of Count Data Models", *Empirical Economics*, 39(3): 733-765.

Bilir, N., Cakir, B., Dagli, E., Erguder, T. and Z. Onder (2009), *Tobacco Control in Turkey*. Copenhagen: WHO Regional Office for Europe.

Bilir, N., Gucis, B. and Y.N. Yildiz (1997), *Smoking Behaviours and Attitudes*. Ankara: Hacettepe Public Health Foundation, International Development Research Centre.

Binder, M. (1998), "Family Background, Gender and Schooling in Mexico", *Journal of Development Studies*, 35(2): 54-71.

Birdsall, N. (1980), *The Economics of Birth Order Effects in a Developing Country*. The World Bank.

Birdsall, N. (1982), Child Schooling and the Measurement of Living Standards. The Living Standard Measurement Study (LSMS) Working Paper, No: 14, The World Bank.

Birdsall, N. (1985), "Public Inputs and Child Schooling in Brazil", *Journal of Development Economics*, 18(1): 67-86.

Blaylock, J.R. and W.N Blisard (1992a), "US Cigarette Consumption: The Case of Low Income Women", *American Journal of Agricultural Economics*, 74(3): 698-705.

Blaylock, J.R. and W.N Blisard (1992b), "Self-evaluated Health Status and Smoking Behaviour", *Applied Economics*, 24(4): 429-435.

Blundell, R. (1988), "Consumer Behaviour Theory and Empirical Evidence-A Survey", *The Economics Journal*, 98(389): 16-65.

Blundell, R. and C. Meghir (1987), "Bivariate Alternatives to the Tobit Model", *Journal of Econometrics*, 34(1-2): 179-200.

Cameron, A. and P.K. Trivedi (1998), *Regression Analysis of Count Data*. Cambridge: University Press.

Cameron, A. and P.K. Trivedi (2005), *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.

Case, A. and A. Deaton (1996), *School Quality and Educational Outcomes in South Africa*. Mimeo.

Chaloupka, F.J. (1990), Men, Women and Addiction: The Case of Cigarette Smoking. National Bureau of Economic Research (NBER) Working Paper Series, No: 3267.

Chaloupka, F.J. and K.E. Warner (2000), "The Economics of Tobacco", In *Handbook of Health Economics*, Culyer, A.J. and J.P. Newhouse (eds.), Elsevier: Amsterdam.

Chernichovsky, D. (1985), "Socioeconomic and Demographic Aspects of School Enrolment and Attendance in Rural Botswana", *Economic Development and Cultural Change*, 33(2): 319-332.

Cragg, J. (1971), "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods", *Econometrica*, 39(5): 829-844.

Dayioglu, M., Kirdar, M.G. and A. Tansel (2009), "Impact of Sibship Size, Birth Order and Sex Composition on School Enrolment in Urban Turkey", *Oxford Bulletin of Economics and Statistics*, 71(3): 399-426.

De Tray, D. (1978), *Child Schooling and Family Size: An Economic Analysis*. Rand Corporation, R-2301-NICHD.

Deaton, A. (1997), *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore, MD: The John Hopkins University Press and The World Bank.

Demographic and Health Survey (DHS) (2003), Hacettepe University, Institute of Population Studies, Ministry of Health-Directorate of Mother and Child Health and Family Planning, State Planning Organisation, Ankara, Turkey.

Deolalikar, A. (1993), "Gender Differences in the Returns to Schooling and in School Enrolment Rates in Indonesia", *The Journal of Human Resources*, 28(4): 899-932.

Dorsett, R. (1999), "An Econometric Analysis of Smoking Prevalence among Lone Mothers", *Journal of Health Economics*, 18(4): 429-441.

Dreze, J. and G.G. Kingdon (2001), "School Participation in Rural India", *Review of Development Economics*, 5(1): 1-24.

Duman, A. (2010), "Female Education Inequality in Turkey: Factors Affecting Girls' Schooling Decisions", *International Journal of Education Economics and Development*, 1(3): 243-258.

Eiser, J.R., Morgan, M., Gammage, P. and E. Gray (1989), "Adolescent Smoking: Attitudes, Norms and Parental Influence", *British Journal of Psychology*, 28(3): 193-202.

Erguder, T., Cakir, B., Aslan, D., Warren, C.W., Jones, N.R. and S. Asma (2008), "Evaluation of the Use of Global Youth Tobacco Survey (GYTS) Data for Developing Evidence-based Tobacco Control Policies in Turkey" BioMed Central (*BMC) Public Health*, 8(1):S4.

Erten, B. and D. Aslan (2008), "Patterns and Dynamics of Female Tobacco Use: A Comprehensive Perspective of Turkey and the United States", *Turkish Journal of Cancer*, 38(3): 99-112.

Erus, B. and N. Aktakke (2012), "The Impact of Healthcare Reforms on Out-of-Pocket Health Expenditures in Turkey for Public Insurees", *The European Journal of Health Economics*, 13(3): 337-346.

Escardibul, J.O. (2005), "The Non-monetary Effects of Education on Smoking in Spain", *European Journal of Education*, 40(1): 45-58.

Euro Education (2011), Structure of the Education System in the United Kingdom. Available from: http://www.euroeducation.net/prof/ukco.htm. [Accessed: 26/06/2011].

European Commission (EC) (2000), Regular Report on Turkey's Progress towards Accession. Brussels.

European Commission (EC) (2008), Regular Report on Turkey's Progress towards Accession. Brussels.

European Commission (EC) (2010), Regular Report on Turkey's Progress towards Accession. Brussels.

Eurydice (2010), Organisation of the Education System in Turkey 2009/2010. European Commission.

Farrell, P. and V.R. Fuchs (1982), "Smoking and Health: The Cigarette Connection", *Journal of Health Economics*, 1(3): 217-230.

Filmer, D. and L.H. Pritchett (2001), "Estimating Wealth Effects without Expenditure Data –or Tears: An Application to Educational Enrolments in States of India", *Demography*, 38(1): 115-132.

Flandorfer, P., Wegner, C. and I. Buber (2010), Gender Roles and Smoking Behaviour, Vienna Institute of Demography Working Papers, No: 7.

Flores, G., Krishnakumar, J. and O. O'Donnell (2008), "Coping with Health-Care Costs: Implications for the Measurement of Catastrophic Expenditures and Poverty", *Health Economics*, 17(12): 1393-1412.

Foster, A. (1994), "Poverty and Illness in Low-Income Rural Areas", *American Economic Review*, 84(2): 216-220.

Frechette, G.R. (2001), Random-effects Ordered Probit, *Stata Technical Bulletin*, 59: 23-27.

Fry, V. and P. Pashardes (1988), *Changing Patterns of Smoking: Are there Economic Causes?* London: Institute of Fiscal Studies.

Fry, V. and P. Pashardes (1994), "Abstention and Aggregation in Consumer Demand: Zero Tobacco Expenditures", *Oxford Economic Papers*, 46(3): 502-518.

Fuchs, V.R. (1982), Time Preference and Health: An Exploratory Study, In *Economic Aspects of Health*, Fuchs, V.R. (eds.), Chicago: The University of Chicago Press.

Garcia, J. and J.M. Labeaga (1996), "Alternative Applications to Modelling Zero Expenditure: An Applications to Spanish Demand for Tobacco", *Oxford Bulletin of Economics and Statistics*, 58(3): 489-506.

Gertler, P. and P. Glewwe (1990), "The Willingness to Pay for Education in Developing Countries", *Journal of Public Economics*, 42(3): 251-275.

Gertler, P., and J. Gruber (2002), "Insuring Consumption against Illness", *American Economic Review*, 92(1): 51-70.

Glewwe, P. and H. Jacoby (1992), Estimating the Determinants of Cognitive Achievement in Low-income Countries: The Case of Ghana, LSMS, No: 91, The World Bank.

Glick, P. and D.E. Sahn (2000), "Schooling of Boys and Girls in a West African Country: The Effects of Parental Education, Income and Household Structure", *Economics of Education Review*, 19(1): 63-88.

Global Adult Tobacco Survey (GATS), Turkey Report (2010), Ministry of Health Publication.

Global Youth Tobacco Survey (GYTS) (2009), Ministry of National Education and WHO Turkey Office, unpublished data.

Goksen, F., Cemalciler, Z. and C.F. Gurlesel (2006), Drop Outs in Turkey's Basic Education Policies for Monitoring and Prevention, Executive Summary, Project: Raising Women: Reducing Gender Disparity in Education through Functional and Political Literacy, Parent Training, Collective Action and Advocacy.

Greene, W. (2003), *Econometric Analysis*, Fifth Edition, Upper Saddle River, NJ: Prentice Hall.

Greene, W.H. (1994), Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models, Department of Economics Working Paper EC-94-10, Stern School of Business, New York University.

Greenhalgh, S. (1985), "Sexual Stratification: The Other Side of 'Growth with Equity' in East Asia, *Population and Development Review*, 11(2): 265-314.

Grootendorst, P.V. (1995), "A Comparison of Alternative Models of Prescription Drug Utilization", *Health Economics,* 4(3): 183-198.

Guio, A.C. (2004), The Laeken Indicators: Some Results and Methodological Issues in Acceding and Candidate Countries, Background paper prepared for the workshop "Aligning the EU Social Inclusion Process and the Millennium Development Goals", April 26-27 2004, Vilnius, Lithuania.

Gurmu, S. and P.K. Trivedi (1996), "Excess Zeros in Count Models for Recreational Trips", *Journal of Business and Economic Statistics,* 14(4): 469-477.

Handa, S. (1996), "Maternal Education and Child Attainment in Jamaica: Testing the Bargaining Hypothesis", *Oxford Bulletin of Economics and Statistics*, 58(1): 119-137.

Harris, M.N. and X. Zhao (2007), "A Zero-inflated Ordered Probit Model, with an Application to Modelling Tobacco Consumption", *Journal of Econometrics,* 141(2): 1073-1099.

Hatt, L.E. (2006), Measuring Risk Factors for Catastrophic Health Expenditures in Peru and their Effects on Families over Time, PhD Dissertation, Department of International Health, Johns Hopkins University, Baltimore.

Hatziandreu, E.J., Pierce, J.P., Fiore, M.C. *et al*. (1989), "The Reliability of Self-Reported Cigarette Consumption in the United States", *American Journal of Public Health*, 79(8): 1020-1023.

Heckman, J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, 47(1): 153-162.

Hill, M.A. and M.E. King (1995), "Women's Education and Economic Well-being", *Feminist Economics*, 1(2): 21-47.

Hisarciklilar, M. (2002), A Censored Regression Model for the Educational Attainment of Boys and Girls in Turkey, Middle East Technical University (METU) International Conference in Economics.

Hisarciklilar, M., McKay, A. and P. Wright (2010), Gender Based Differences in Educational Achievement in Turkey: What Has Changed Over Time?, Presented at the Middle East Economic Association (MEEA) Conference, Atlanta.

Hjortsberg, C. (2003), "Why do the Sick not Utilise Health Care? The Case of Zambia", *Health Economics,* 12(9): 755-770.

Holmes, J. (2003), "Measuring the Determinants of School Completion in Pakistan: Analysis of Censoring and Selection Bias", *Economics of Education Review*, 22(3): 249-264.

Hu, T-W., Ren Q-F., Keeler, T.E. and J. Bartlett (1995), "The Demand for Cigarettes in California and Behavioural Risk Factors", *Health Economics,* 4(1): 7-14.

Human Development Report (2010), *20th Anniversary Edition, The Real Wealth of Nations: Pathways to Human Development*. The United Nations Development Programme.

Idris, B.I., Giskes, K., Borrell, C. *et al*. (2007), "Higher Smoking Prevalence in Urban compared to Non-urban Areas: Time Trends in Six European Countries", *Health and Place*, 13(3): 702-712.

Jamison, T. and M.E. Lockheed (1987), "Participation in Schooling: Determinants and Learning Outcomes in Nepal", *Economic Development and Cultural Change*, 35(2): 279-306.

Jha, P. and F.J. Chaloupka (1999), *Curbing the Epidemic: Governments and the Economics of Tobacco*, Washington, D.C.: World Bank.

Jones, A.M. (1989a), "A Double-Hurdle Model of Cigarette Consumption", *Journal of Applied Econometrics*, 4(1): 23-39.

Jones, A.M. (1989b), "The UK Demand for Cigarettes 1954-1986, A Double-Hurdle Approach", *Journal of Health Economics*, 8(1): 133-141.

Jones, A.M. (1995), A microeconometric Analysis of Smoking in the UK Health and Lifestyle Survey, Discussion Paper No: 139, The University of York.

Jones, A.M. (2000), Health Econometrics, In *Handbook of Health Economics*, Culyer, A.J. and J.P. Newhouse (eds.), Elsevier: Amsterdam.

Jones, A.M. and S.T. Yen (2000), "A Box-Cox Double-Hurdle Model", *The Manchester School*, 68(2): 203-221.

Kabubo-Mariara, J. and D. Mwabu (2007), "Determinants of School Enrolment and Educational Attainment: Empirical Evidence from Kenya", *South African Journal of Economics*, 75(3): 572-593.

Kalmijn, M. (1994), "Mother's Occupational Status and Children's Schooling", *American Sociological Review*, 59(2): 257-275.

Kawabata, K., Xu, K. and G. Carrin (2002), "Preventing Impoverishment through Protection against Catastrophic Health Expenditure", *Bulletin of the World Health Organization*, 80(8): 612.

Keeler, T.E., Hu, T.W., Barnett, P.G., Manning, W.G. and H.T. Sung (1996), "Do Cigarette Producers Price-Discriminate by State? An Empirical Analysis of Local Cigarette Pricing and Taxation", *Journal of Health Economics*, 15(4): 499-512.

Keeler, T.E., Hu, T.W., Manning, W.G. and Sung, H.Y. (2001), "State Tobacco Taxation, Education and Smoking: Controlling for the Effects of Omitted Variables", *National Tax Journal*, 54(1): 83-102.

Keng, C. (2004), "Household Determinants of Schooling Progression among Rural Children in Cambodia", *International Education Journal*, 5(4): 552-561.

Kenkel, D.S. (1991), "Health Behaviour, Health Knowledge and Schooling", *The Journal of Political Economy*, 99(2): 287-305.

King, E.M. and L.A. Lillard (1983), Determinants of Schooling Attainment and Enrolment Rates in the Philippines, RAND Note, N-1962-AID.

King, E.M. and L.A. Lillard (1987), "Education Policy and Schooling Attainment in Malaysia and the Philippines", *Economics of Education Review*, 6(2): 167-181.

King, E.M. and M.A. Hill (1993), *Women's Education in Developing Countries: Barriers, Benefits and Policies*. Baltimore: The John Hopkins University Press for the World Bank.

Kirdar, M. (2009), "Explaining Ethnic Disparities in School Enrolment in Turkey", *Economic Development and Cultural Change*, 57(2): 297-333.

Kisa, A., Yilmaz, F., Younis, M.Z., Kavuncubasi, S., Ersoy, K. and P.A. Rivers (2009), "Delayed Use of Healthcare Services among the Urban Poor in Turkey", *Education, Business and Society: Contemporary Middle Eastern Issues*, 3(3): 232-240.

Knaul, F. (2000), *The WHO Index of Fairness of Financial Contribution and Framework for Health System Performance Assessment: Equity and Efficiency of Health Finance.* Geneva: World Health Organisation.

Koopmenschap, M. and F. Rutten (1994), "The Impact of Indirect Costs on Outcomes of Health Care Programs", *Health Economics*, 3(6): 385-393.

Lambert, D. (1992), "Zero-inflated Poisson Regression with an Application to Defects in Manufacturing", *Technometrics*, 34(1):1-14.

Lazear, E. (1977), "Education: Consumption or Production", *Journal of Political Economy*, 85(3): 569-597.

Lewit, E.M. and D. Coate (1982), "The Potential for Using Excise Taxes to Reduce Smoking", *Journal of Health Economics*, 1(2): 121-145.

Lillard, E.M. and E.M. King (1984), "Methods for Analysing Schooling Choice with Household Survey Data", RAND Note, N-1963-AID.

Lillard, L.A. and R.J. Willis (1994), "Intergenerational Educational Mobility: Effects of Family and State in Malaysia", *Journal of Human Resources*, 29(4): 1126-1166.

Long, J.S. (1997), *Regression Models for Categorical and Limited Dependent Variables*, California: SAGE Publications.

Long, J.S. and J. Freese (2006), Regression Models for Categorical Dependent Variables Using Stata, STATA Press Second Edition.

Maddala, G. (1999), *Limited-dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.

Maitra, P. (2003), "Schooling and Educational Attainment: Evidence from Bangladesh", *Education Economics*, 11(2): 129-153.

Makinen, M., Waters, H., Rauch, M., Almagambetove, N., Bitran, R. *et al.* (2000), "Inequalities in Health Care Use and Expenditures: Empirical Data from Eight Developing Countries and Countries in Transition", *Bulletin of the World Health Organization*, 78(1): 55-65.

Marsh, A. and S. McKay (1994), *Poor Smokers*, London: Policy Studies Institution (PSI).

McIntyre, D., Theiede, M., Dahlgren, G. and M. Whitehead (2006), "What are the Economic Consequences for Households of Illness and Paying for Health Care in Low and Middle Income Country Contexts?", *Social Science and Medicine*, 62(4): 858-865.

McLoughney, E., Fornora, M., Zavarko, S. and K.L. Neal (2007), Eliminating Gender Disparity in Primary Education: The Case of Turkey, Division of Policy and Planning Working Papers, United Nations Children's Fund (UNICEF), New York.

Mincer, J. (1958), "Investment, Human Capital and Personal Income Distribution", *Journal of Political Economy*, 66(1): 281-321.

Mincer, J. (1974), *Schooling Experience and Earnings*. New York: National Bureau of Economic Research.

Mincer, J. (1984), "Human Capital and Economic Growth", *Economics of Education Review*, 3(3): 195-205.

Ministry of Education (2010), Education Indicators, Turkey.

Ministry of Health (2009), Prevention and Control Program for Cardiovascular Diseases: Strategic Plan for the Risk Factors, Directorate of General Primary Health Care Services, Turkey. Available from: http://www.tkd.org.tr/english.asp?pg=396. [Accessed: 11/02/2011].

Mullahy, J. (1986), "Specification and Testing of Some Modified Count Data Models", *Journal of Econometrics*, 33(3): 341-365.

Mullahy, J. (1997), "Instrumental-Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behaviour", *The Review of Economics and Statistics*, 79(4): 586-593.

Murray, C.J., Knaul, F., Musgrove, P., Xu, K. and K. Kawabata (2000), Defining and Measuring Fairness in Fairness in Financial Contribution to the Health System, Global Programme for Evidence in Health Policy Discussion Paper Series: No: 24. Geneva: World Health Organisation.

National Education Statistics (2009), Formal Education 1997-2009, Turkey.

O'Donnell, O. and E. Van Doorslaer (2005), "Explaining the Incidence of Catastrophic Expenditures on Health Care: Comparative Evidence from Asia", EQUITAP Project Working Paper 5, Thessaloniki, Greece: European Commission Cooperation for Development Program.

O'Donnell, O., Doorslaer, E.V., Wagstaff, A. and M. Lindelow (2008), *Analysing Health Equity Using Household Survey Data*: *A Guide to Techniques and Their Implementation*. World Bank Institute, WBI Learning Resources Series.

Onder, Z. (2002), The Economics of Tobacco in Turkey: New Evidence and Demand Estimates, Tobacco Control World Health Organization Tobacco Control Papers, University of California, San Francisco.

Organization for Economic Co-operation and Development (OECD) (2008), OECD Reviews of Health Systems: Turkey, OECD and the International Bank for Reconstruction and Development. The World Bank.

Organization for Economic Co-operation and Development (OECD) (2010), Turkey-Economic Outlook 88 Country Summary. Available from: http://www.oecd.org/document/34/0,3343,en_33873108_33873854_45270434_1_1_1_1,00.html?rssChId=33873854. [Accessed: 13/08/2011].

Organization for Economic Co-operation and Development (OECD) (2010), OECD Health Data 2010, Selected Data.

Pal, S. (2004), "How Much of the Gender Difference in Child School Enrolment can be Explained? Evidence from Rural India", *Bulletin of Economic Research*, 56(2): 133-158.

Parish, W.L and R.J. Willis (1993), "Daughters, Education and Family Budgets: Taiwan Experiences", *The Journal of Human Resources*, 28(4): 863-898.

Pradhan, M. and N. Prescott (2002), "Social Risk Management Options for Medical Care in Indonesia", *Health Economics*, 11(5): 431-446.

Rankin, B.H. and I.A. Aytac (2006), "Gender Inequality in Schooling: The Case of Turkey", *Sociology of Education*, 79(1): 25-43.

Ranson, M.K. (2002), "Reduction of Catastrophic Health Care Expenditures by a Community-based Health Insurance Scheme in Gujarat, India: Current Experiences and Challenges", *Bulletin of the World Health Organization*, 80(8): 613-621.

Rawlings, L.B. and G.M. Rubio (2005), "Evaluating the Impact of Conditional Cash Transfer Programs", *World Bank Research Observer*, 20(1): 29-55.

Rose, P. and S. Al-Samarrai (2001), "Household Constraints on Schooling by Gender: Empirical Evidence from Ethiopia", *Comparative Education Review*, 45(1): 36-63.

Rosenzweig, M. (1978), The Value of Children's Time, Family Size and Non-household Child Activities in a Developing Country: Evidence from Household Data, In *Research in Population Economics*, Simon, J.C. (eds.), Vol. 1: 331-347.

Ross, H. and F.J. Chaloupka (2003), "The Effect of Cigarette Prices on Youth Smoking", *Health Economics*, 12(3): 217-230.

Russell, S. (2004), "The Economic Burden of Illness for Households in Developing Countries: A Review of Studies Focusing on Malaria, Tuberculosis and Human Immunodeficiency Virus/Acquired Immunodeficiency Syndrome", *The American Journal of Tropical Medicine and Hygiene*, Connecticut, JAI Press.

Sartori, A.E. (2003), "An Estimator for Some Binary-Outcome Selection Models without Exclusion Restrictions", *Political Analysis*, 11(2): 111-138.

Sauerborn, R., Adams, A. and M. Hien (1996), "Household Strategies to Cope with the Economic Costs of Illness", *Social Science and Medicine*, 43(3): 291-301.

Savas, B.S., Karahan, O. And O. Saka (2002), Health Care Systems in Transition: Turkey, In *Copenhagen, European Observatory on Health Care Systems*, Thomson, S. and Mossialos, Copenhagen: WHO Regional Office.

Sawada, Y. and M. Lokshin (2001), Household Schooling Decisions in Rural Pakistan, Policy Research Working Paper, No: 2541, World Bank.

Schultz, T.P. (1993), Returns to Women's Education, In *Women's Education in Developing Countries: Barriers, Benefits and Policies*, King, E.M. and M.A. Hill (eds.), Baltimore, London: The John Hopkins University Press for the World Bank.

Schultz, T.W. (1960), "Capital Formation by Education", *Journal of Political Economy*, 68(12): 571-583.

Schultz, T.W. (1961), "Investment in Human Capital", *The American Economic Review*, 51(1): 1-17.

Schultz, T.W. (1963), *The Economic Value of Education*. New York: Columbia University Press.

Sepehri, A., Sarma, S. and W. Simpson (2006), "Does Non-profit Health Insurance Reduce Financial Burden? Evidence from the Vietnam Living Standards Survey Panel", *Health Economics*, 15(6): 603-616.

Sheu, M.L., Hu, T.W., Keeler, T.E., Ong, M. and H.Y. Sung (2004), "The Effect of a Major Cigarette Price Change on Smoking Behaviour in California: A Zero-Inflated Negative Binomial Model", *Health Economics*, 13(8): 781-791.

Smits, J. and A.G. Hosgor (2006), "Effects of Family Background Characteristics on Educational Participation in Turkey", *International Journal of Educational Development*, 26(5): 545-560.

State Planning Office (SPO) (2006), Ninth Five-Year Development Plan 2007-2013.

Strauss, J. and D. Thomas (1995), Human Resources: Empirical Modelling of Household and Family Decisions, In *Handbook of Developing Economics*, Behrman, J. and T.N. Srinivasan (eds.), Elsevier: Amsterdam.

Sulku, S.N. and D. Bernard (2009), Financial Burden of Health Care Expenditures in Turkey: 2002-2003, Agency for Healthcare Research and Quality (AHRQ) Working Paper Series, No: 09004.

Sung, H.Y., Hu, T.W. and T.E. Keeler (1994), "Cigarette Taxation and Demand: An Empirical Model", *Contemporary Economic Policy*, 12(3): 91-100.

Tansel, A. (1993), "Cigarette Demand, Health Scares and Education in Turkey", *Applied Economics*, 25(4): 521-529.

Tansel, A. (1997), "Schooling Attainment, Parental Education and Gender in Cote d'Ivoire and Ghana", *Economic Development and Cultural Change*, 45(4): 825-856.

Tansel, A. (2002), "Determinants of School Attainment of Boys and Girls in Turkey: Individual, Household and Community Factors", *Economics of Education Review*, 21(5): 455-470.

Tansel, A. (2005), "Public-Private Employment Choice, Wage Differential and Gender in Turkey", *Economic Development and Cultural Change*, 53(2): 453-477.

Tansel, A. and F. Bircan (2006), "Demand for Education in Turkey: A Tobit Analysis of Private Tutoring Expenditures", *Economics of Education Review*, 25(3): 303-313.

Tatar, M. and P. Kanavos (2006), "Health Care Reform in Turkey", *Eurohealth*, 12(1): 20-22.

Tatar, M., Ozgen H., Sahin, B., Belli, P. and P. Berman (2007), "Informal Payments in the Health Sector: A Case Study from Turkey", *Health Affairs*, 26(4): 1029-1039.

The General Directorate of Highways (2010), Map of Highways, Ankara. Available from http://www.kgm.gov.tr/Sayfalar/KGM/SiteEng/Root/Distances.aspx, [Accessed: 13/01/2011].

The World Factbook (2010), Life Expectancy at Birth and Infant Mortality Rate. https://www.cia.gov/library/publications/the-world-factbook/. [Accessed: 22/08/2011].

Thomas, D. (1994), "Like Father, Like Son, Like Mother, Like Daughter: Parental Resources and Child Height", *Journal of Human Resources*, 29(4): 950-988.

Tobin, J. (1958), "Estimation of Relationships for Limited Dependent Variables", *Econometrics*, 26(1): 24-36.

Toros, A. and Z. Oztek (1993), Health Services Utilization Survey in Turkey, Ministry of Health, Ankara.

Tunali, I. (1996), Education and Work: Experiences of 6-14 Year-Old Children in Turkey, In *Education and the Labour Market in Turkey*, Bulutay, T. (eds.), Proceedings of a Seminar Held in Ankara, Turkish Statistical Institute, Ankara.

Turkish Statistical Institute (2000), Provincial-in Migration, Provincial-out Migration, Net Migration and the Rate of Net Migration, 2000 Census of Population, Turkey.

Turkish Statistical Institute (2003), Cigarette Consumption Profile of Women in Turkey. Available from: www.tuik.gov.tr/Beslenme/excel/0800t05.xls. [Accessed: 05/11/2010].

Turkish Statistical Institute (TurkStat) (2006), Household Budget Survey, Methodology.

Turkish Statistical Institute (TurkStat) (2007), Public and Private Hospitals 2004-2007.

Turkish Statistical Institute (TurkStat) (2009), Poverty Analysis 2009.

Turkish Statistical Institute (TurkStat) (2010), 2004-2007 Health Expenditure Statistics.

Turkish Statistical Institute (TurkStat) (2010), Global Adult Tobacco Survey 2008.

Turkish Statistical Institute (TurkStat) (2010), Population Statistics and Projections. Available from: http://www.turkstat.gov.tr/PreTablo.do?tb_id=39&ust_id=11. [Accessed: 18/09/2011].

Turkish Statistical Institute (TurkStat) (2011), Labour Force Statistics. Available from: http://www.tuik.gov.tr/Gosterge.do?metod=GostergeListe&tb_id=25&ust_id=8. [Accessed: 26/08/2011].

Turkish Statistical Yearbook (TurkStat) (2010), Turkish Statistical Institute.

United Nations Development Programme (UNDP) (1995), *Human Development Report*, New York: Oxford University Press.

United States Department of Agriculture (USDA) (2007), U.S. Tobacco Import Update 2005/06, A Report from the Economic Research Service. Available from: http://www.ers.usda.gov/Publications/TBS/May07/TBS26201/. [Accessed: 13/12/2010].

Unuvar, N., *et al*., (2006), Turkey National Burden of Disease and Cost-Effectiveness Study, National Household Survey 2003, Basic Findings, Ministry of Health, Ankara.

Van de Ven, W.P.M.M. and B.M.S. Van Pragg (1981), "The Demand for Deductibles in Private Insurance: A Probit Model with Sample Selection", *Journal of Econometrics*, 17(2): 229-252.

Van Doorslaer, E., Wagstaff, A., Van der Burg, H. *et al.* (1999), "The Redistributive Effect of Health Care Finance in Twelve OECD Countries", *Journal of Health Economics*, 18(3): 291-313.

Vuong, Q. (1989), "Likelihood Ratio Tests for Model Selection and Non-nested Hypothesis", *Econometrica*, 57(2): 307-333.

Wagstaff, A. and E. Van Doorslaer (2003), "Catastrophe and Impoverishment in Paying for Healthcare: with Applications to Vietnam 1993-1998", *Health Economics*, 12(11): 921-933.

Wagstaff, A., M. Lindelow (2008), "Can Insurance Increase Financial Risk? The Curios Case of Health Insurance in China", *Journal of Health Economics*, 27(4): 990-1005.

Waldron, I. (1988), Gender and Health-Related Behaviour, In *Health Behaviour: Emerging Research Perspectives*, Gochman, D.S. (eds.), New York: Plenium Press.

Warner, K.E. (1978), "Possible Increases in the Underreporting of Cigarette Consumption", *Journal of the American Statistical Association*, 73(362): 314-318.

Wasserman, J., Manning, W., Newhouse, J., Winkler, J. (1991), "The Effects of Excise Taxes and Regulations on Cigarette Smoking", *Journal of Health Economics*, 10(1): 43-64.

Waters, H.R., Anderson, G.F. and J. Mays (2004), "Measuring Financial Protection in Health in the United States", *Health Policy*, 69(3): 339-349.

Winkelmann, R. (2008), *Econometric Analysis of Count Data*. Berlin, Heidelberg: Springer Verlog.

Wolfe, B. and J. Behrman (1984), "Who is Schooled in Developing Countries? The Roles of Income, Parental Schooling, Sex, Residence and Family Size", *Economics of Education Review*, 3(3): 231-245.

Wolfe, B. and J. Behrman (1986), "Child Quantity and Quality in a Developing Country: Family Background, Endogenous Tastes, and Biological Supply Factors", *Economic Development and Cultural Change*, 34(4): 703-718.

World Bank (2003), Turkey: Reforming the Health Sector for Improved Access and Efficiency, Volume 1 and 2, report no: 24358_TU.2003. Washington DC: The International Bank for Reconstruction and Development. The World Bank.

World Bank (2010), World Development Indicators. Available from: http://data.worldbank.org/indicator/NY.GNP.PCAP.PP.CD/countries/1W-TR?display=default. [Accessed: 22/08/2011].

World Health Organization (WHO) (2000), The World Health Report 2000, Health Systems: Improving Performance. Geneva: The World Health Organization.

World Health Organization (WHO) (2003), WHO Framework Convention on Tobacco Control. Geneva: The World Health Organization. Available from: http://whqlibdoc.who.int/publications/2003/9241591013.pdf. . [Accessed: 13/01/2011].

World Health Organization (WHO) (2009) Report on the Global Tobacco Epidemic: Implementing Smoke-free Environments. Geneva: The World Health Organization. Available from: http://whqlibdoc.who.int/publications/2009/9789241563918-eng-full.pdf. [Accessed: 25/02/2011].

Wyszewianski, L. (1986), "Families with Catastrophic Healthcare Expenditures", *Health Services Research*, 21(5): 617-634.

Xu, K., Evans, D.B., Kawabata, K., Zeramdini, R., Klavus, J. and C.J. Murray (2003), "Household Catastrophic Health Expenditure: A Multicountry Analysis", *Lancet*, 362(9378): 111-117.

Yardim, M.S., Cilingiroglu, N. and N. Yardim (2009), "Catastrophic Health Expenditure and Impoverishment in Turkey", *Health Policy*, 94(1): 26-33.

Yen, S.T and B.H. Lin (2006), "A Sample Selection Approach to Censored Demand Systems", *American Journal of Agricultural Economics*, 88(3): 742-749.

Yen, S.T. (2005), "Zero Observations and Gender Differences in Cigarette Consumption", *Applied Economics*, 37(16): 1839-1849.

Yen, S.T. and A. Jones (1996), "Individual Cigarette Consumption and Addiction: A Flexible Limited Dependent Variable Approach", *Health Economics*, 5(2): 105-117.

Yildirim, J., Yilmaz, E. and N. Korucu (2011), "The Determinants of Out-of-Pocket Payments: Evidence from Selected Hospitals in Ankara, Turkey", *Applied Economics Letters*, 18(12): 1159-1162.

Yurekli, A.A. and A.M. Perucic (2010), Economics of Tobacco Toolkit: Economic Analysis of Demand Using Data from the Global Adult Tobacco Survey (GATS), Geneva: World Health Organisation Press.

Yurekli, A.A. and P. Zhang (2000), "The Impact of Clean Indoor-Air Laws and Cigarette Smuggling on Demand for Cigarette: An Empirical Model", *Health Economics*, 9(2): 159-170.

Yurekli, A.A., Onder, Z., Elibol, M., Erk, N. Cabuk, A., Fisunoglu, M., Erk, S.F and F.J. Chaloupka (2010), *The Economics of Tobacco and Tobacco Taxation in Turkey*. Paris: International Union against Tuberculosis and Lung Disease.

Zorn, C.J.W. (1996), Evaluating Zero-Inflated and Hurdle Poisson Specifications, Midwest Political Science Association Political Methodology Working Papers.