EXPLAINING THE INEQUALITY PUZZLES IN THE TURKISH LABOUR MARKET

Okan Yılmaz

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

Date Submitted: 7 JULY 2015

ABSTRACT

Turkey constitutes an economic and cultural bridge between the Eastern and the Western world and through its development process it has become an industrial, urbanised country with stronger relations to the global economy. The overall purpose of this thesis is to analyse the Turkish labour market in order to understand the unusual inequality patterns that are observed in it over the post 2000 period.

The first empirical study presented in Chapter 2 investigates the underlying reasons for the decrease in wage inequality in Turkey. The results of the study show that the increasing supply of the educated labour, which is due to increases in the number of universities and the reform to the education system which increased the years of compulsory education, played a substantial role in the decreasing wage inequality in Turkey between 2002 and 2010. Chapter 3 analyses the relationship between domestic violence, traditional social norms, and female autonomy in the household which is one of the factors that plays an important role in the labour force participation decision of women in Turkey. The findings of the analyses suggest that: domestic violence has a negative effect on female autonomy; female autonomy in turn has a significant and negative effect on domestic violence; and finally the strength of the husband's social norms on traditional gender roles positively affects the risk of violence. The final empirical study of the thesis presented in chapter 4 focuses on the structure of the informal sector employment in Turkey. The results of the empirical analysis suggest a heterogeneous structure for the informal sector in Turkey in which the lower tier consists of workers who involuntarily participate in the informal sector and the higher tier includes the workers who voluntarily choose to be in the informal sector.

ACKNOWLEDGEMENTS

First and foremost, I owe my deepest gratitude to Dr Gurleen Popli for her invaluable advice, comments and encouragement. I am grateful to her for her inspiring and dedicated guidance in both theoretical, methodological, ethical and time management issues. Taking the time to help me whenever I needed, trusting me and providing me with continuous moral support throughout my time in Sheffield, she has been more than a supervisor to me. I am also grateful to Dr Jonathan Perraton for his valuable comments, suggestions, and his interest in this process. He was there to help me with understanding theoretical problems.

I would like to thank my family for their unconditional support and patience in this tough process. I would like express my thanks to my friends, Przemyslaw Antoni Wójcik, Pınar Öktem, Burçak Tatlı, Milad Karimi, Gökçe Kale and Odai Al Zoubi for their warm support they gave over this time. I also would like to thank Dr Gökçer Özgür who has always been an inspiration and mentor to me.

Most of all, I would like to thank Çiğdem Gedikli, whose support and belief in me made this all possible, for her encouragement, sacrifices and understanding.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

DHS:	Demographic and Health Survey
EU:	European Union
GDP:	Gross Domestic Product
GDSPW:	The General Directorate on the Status and Problems of Women
GLS:	Generalised Least Squares
HBS:	Household Budget Survey
HSS:	Household Structure Survey
IDB:	Inter-American Development Bank
ILO:	International Labour Organization
IRT:	Item Response Theory
JMP:	Juhn-Murphy-Pierce
LFS:	Labour Force Survey
ML:	Maximum Likelihood
OECD:	The Organisation for Economic Co-operation and Development
OLS:	Ordinary Least Squares
OSYM:	Student Selection and Placement Centre
RRR:	Relative Risk Ratios
SBTC:	Skill Biased Technological Change
SEM:	Structural Equation Modelling
SLS:	Semi-parametric Least Square
TurkStat:	Turkish National Statistics Institute
UN:	United Nations
UNICEF:	The United Nations Children's Fund
WHO:	World Health Organization
WLSMV:	Weighted Least Squares Means and Variance Adjusted

CHAPTER 1: INTRODUCTION

1.1 Aims and Motivation of the Thesis

Due to its geographical location, Turkey constitutes an economic and cultural bridge between the Eastern and the Western world and it has experienced major transformations in its process through industrialisation. While its economy was mainly based on agricultural production previously, the contribution of industrial production has begun to outweigh the agricultural sector starting from the 1980's. The rapid industrialisation has also been accompanied by urbanisation; only a third of the population was living in the urban areas in the 1960's and this ratio doubled in the 2000's. Starting from the early 1990's, the educational attainment of the population has increased dramatically. With the increasing number of universities, with 29 universities in 1991 to 165 in 2011, the number of students in university education increased by more than 300 per cent, rising from 1.2 million to 3.7 million over this period. The education reform which increased the compulsory education from 5 to 8 years came into effect in 1997 and played an important role in the schooling of the population. The total enrolment rate for primary education was around 84 per cent in 1997 and increased to 98 per cent in 2010. The education reform also closed the enrolment gap between boys and girls significantly; the enrolment rate for primary education was 90 per cent for boys and 78 per cent for girls in 1997, these rates equalised at 98 per cent in 2010 (Turkish National Statistics Institute).

The socio-economic transformation within Turkey was also accompanied by a stronger relationship with the global economy, particularly with the European Union (EU hereafter). With a gradual process, which started in 1980 and ended in 1989, the movement of international capital was freed. In 1997, Turkey became a member of the Customs Union after agreement with the EU. Turkey also became a member of the G-

20 organisation when it was established in 1999. In the same year, Turkey was officially recognised as a candidate for full membership to the EU and negotiations started in 2005.

Based on the socio-economic developments and the stronger relations with the global economy, it can be argued that Turkey has followed a path towards the Western world and has become a more industrial, more urbanised, and more educated country. Yet, the Turkish labour market has exhibited surprising characteristics which are not observed among developed and developing countries that experienced or are experiencing similar transformations. This thesis focuses on three exceptional patterns that are observed in the Turkish labour market and tries to explore the reasons for these patterns with empirical analyses in the light of existing economic theories.

The first interesting observation on the Turkish labour market is the decreasing wage inequality which is in contrast to increasing trends observed in many developed and developing countries. There has been considerable interest in studying the distribution of wages over the last three decades, following the dramatic increase in wage inequality observed across several countries. In particular, the steep increase in the wage gap between the college and the high school graduates in the United States (US) has been documented by many authors, such as Katz and Murphy (1992) and Bound and Johnson (1992). In line with these studies, most of the analysis in the literature has focused on wage inequality between workers with high and low educational qualifications. Besides the US, an increasing educational premium has also been documented for many OECD countries (Berman et al., 1997; Machin and Reenen 1998).

The Stolper-Samuelson theory argues that competition with labour abundant countries decreases the relative price of labour-intensive goods and accordingly reduces the real

wages of less educated workers in both relative and absolute terms which in turn leads to an increase in wage inequality between workers with different educational qualifications (Hanson and Harrison, 1999). The skill biased technological change theory argues that in accordance with the diffusion of higher technologies of information and communication, labour demand has shifted in favour of skilled workers (Acemoglu, 2002). Therefore, the technological developments have increased the skill premiums and caused increases in wage inequality.

Based on economic theories on changes in the wage distribution and the increasing trend in wage inequality observed in many countries, the first empirical contribution of the thesis is an attempt to explain the underlying reasons for the decreasing wage inequality observed in the Turkish labour market. The thesis focuses on the effect of increasing educational attainment on wage inequality.

The second interesting observation on the Turkish labour market is the low level of female labour participation. According to the Global Gender Gap Report 2013, among 136 countries, Turkey has the 120th rank in the global gender gap index[1] and 127th rank in the sub-index for economic participation and opportunity for women. The Report also shows that Turkey has the 123rd rank in the female labour force participation rate. Another interesting aspect of this observation is that female labour force participation has been decreasing over last the 60 years. In 1950, the female labour force participation in Turkey was 52 per cent and this rate decreased to 35 per cent in 1988 and to 25 per cent in 2006 (Fernandez and Fogli, 2005; Dayioglu and Kirdar, 2010). The decreasing female labour force participation in Turkey is surprising as decreases in

[1] The global gender gap index examines the gap between men and women in four sub-indexes: economic participation and opportunity, educational attainment, health and survival, and political empowerment. The overall index score is calculated as the un-weighted average of each sub-index score. The highest possible score 1 represents equality while the lowest possible score 0 indicates inequality. The value of the overall index score of Turkey for the year 2013 was 0.58.

fertility rates and increases in the educational attainment of women took place at the same period.

The second empirical contribution of this thesis focuses on female autonomy in the household which is one of the factors that plays an important role in the labour force participation decision of women in Turkey. According to the Demographic and Health Survey for the year 2005, apart from many factors such as household responsibilities and education, around 24 per cent of women in Turkey do not work or look for a job because their husbands or elders prevent them from doing so. In that sense, the lack of autonomy of women within the household lowers the labour force participation. Goksel (2012) analyses the effect of conservativeness on female labour force participation in Turkey. She shows that in a conservative and traditional environment, where men have higher decision power and stronger social norms on traditional gender roles exist, women tend to stay at home and do not participate in the labour force.[2] In accordance with these findings, this chapter focuses on the determinants of female autonomy in Turkey. The particular focus of the second empirical contribution is the relationship between female autonomy, domestic violence and social norms.

The third issue that this thesis explores is the substantial size of informal sector employment in the Turkish labour market. One of the common features of the labour markets in developing countries is the high rates of informal employment. According to the Global Employment Trends Report (ILO, 2013), between 40-50 per cent workers in countries in the Latin American and Caribbean region are employed in the informal sector, while this share is between 15-30 per cent in Central and South-eastern Europe

[2] Goksel (2012) creates an index for the husband's conservativeness using several questions provided in the Household Structure Survey for the year 2006. The same dataset and similar questions are used to capture social norms in the analysis of the third chapter. The details of the dataset and the questions used in the analysis are provided in Chapter 3.

(non-EU) countries. The size of informal employment in Turkey has been quite high compared to the neighbouring European and Middle Eastern countries. Although there has been a decreasing trend between 2000 and 2010, the share of informal sector workers in total employment in Turkey is still around 40 per cent, which is close to the values for the Latin American countries where the highest rates of informal employment in the world are observed.

Analysis of informal sector employment is important because, traditionally, the informal sector has been seen as inferior compared to the formal sector in terms of wages, security and protection from exploitation regarding labour standards. As the informal sector is free of any regulation, wage offers can be below the minimum wage[3] level, the standards regarding age, length and time of working hours can be ignored, and workers may have to work in risky environments with unsafe equipment. In addition, employment benefits such as social security and compensation do not apply to informal employment due to its unrecorded nature (Losby et al, 2002). Supporting these arguments, a report by the Turkish Central Bank (2015) shows that 23 per cent of employees in Turkey work for wages below the national minimum wage level. The third empirical contribution of the thesis examines the structure of the informal sector employment in Turkey.

1.2 Structure and Content of the Thesis

In accordance with the aims and motivations described above, this thesis looks at labour market inequalities in Turkey from three angles: earnings inequality between employees which show a decreasing trend over the last decade; female autonomy which is a

[3] Turkey introduced minimum wage laws in 1971. The minimum wage is set by the 'The Minimum Wage Determination Commission' which has representatives from the government, trade unions and employers. It should be noted that over the last few decades the power of unions has gone down considerably. Once the minimum wage is set, the Central Government (specifically, the Ministry of Labour and Social Security) monitors the implementation of it, but only in the formal registered sectors of the economy. For further details on the minimum wage in Turkey and how they are set see Koçer and Visser (2009).

potential reason for the low female labour force participation; and finally the structure of the informal sector employment which constitutes a significant share of the total employment in Turkey. Chapters 2, 3 and 4 provide the empirical studies that focus on these angles. Chapter 5, the final chapter of the thesis, summaries the findings of the analyses and presents the overall conclusion. The contents of Chapters 2, 3 and 4 are summarised below.

1.2.1 Chapter 2

Chapter 2 explores the underlying reasons behind the observed decrease in wage inequality in the Turkish labour market using data from the Labour Force Surveys (LFS). To understand the factors affecting wage inequality in Turkey, the change in the wage distribution between 2002 and 2010 is decomposed into three components, namely, the effect of returns to human capital (price of human capital), the effect of changes in the composition of human capital (changes in covariates) and changes in the residual distribution (price and composition of unmeasured human capital characteristics). The decomposition method used in this analysis is proposed by Lemieux (2002) and it is a unified form of the two commonly used decomposition techniques in the literature which were proposed by Juhn et al. (1993) and DiNardo et al. (1996). In order to correct for the possible selection bias related to participation in wage employment, a two-step procedure, proposed by Dubin and McFadden (1984), is used.

The decomposition results reveal that decreasing wage inequality was mainly due to decreasing returns to human capital or, in other words, decreasing skill premiums and decreasing residual wage dispersion. Using the human capital approach it is also found that the decrease in residual wage dispersion was mainly due to the decreasing prices of unmeasured human capital. Finally, the effect of changes in the composition of human

capital is found to be moderate. The results of the study suggest that the increasing supply of educated labour, which is due to increases in the number of universities and the reform to the education system which increased the years of compulsory education, played a substantial role in the decreasing wage inequality in Turkey between 2002 and 2010.

1.2.2 Chapter 3

In this chapter, using data from the Household Structure Survey (HSS) for the year 2006, the relationship between domestic violence, social norms, and female autonomy in the household is analysed. Dyson and Moore (1983) define autonomy as the capacity to manipulate one's personal environment and the technical, social and psychological abilities in order to obtain information and to use it as the basis for making decisions about one's private concerns and those of one's intimates. Empirically, it is found that higher levels of female autonomy within the household lead to better outcomes for the household members, especially for children. For instance, Quisumbing and Maluccio (2000) show that while husbands tend to spend money on tobacco and luxury consumption goods, women, when they have the control of assets, spend more on children's education, health and clothing. Mason (1987) and Eswaran (2002) argue that the empowerment of women may lead to lower fertility rates as women face higher costs of children than men. To analyse the relationship between domestic violence, social norms, and female autonomy in the household, an Item Response Theory (IRT) model is applied.

The key findings of the study suggest that, consistent with the arguments of feminist theory on domestic violence, violence has a significant and decreasing effect on female autonomy and the incidence of violence decreases with the level of female autonomy. It is also found that domestic violence is an increasing function of the strength of social

norms of a husband on traditional gender roles. Further, while education and income have an increasing effect on female autonomy, they have a decreasing effect on the strength of the social norms of the husband on traditional gender roles. Further analyses of the data show that education and having an income indirectly affect the level of the wife's autonomy through their social norms. Once the level of the wife's social norms is controlled for, there is no direct effect of these variables on female autonomy.

1.2.3 Chapter 4

This chapter investigates the heterogeneity of employment in the informal sector by looking at data from the Turkish labour market in which informal workers constitute more than 40 per cent of total employment. According to Stiglitz (1976), and Dickens and Lang (1985), being involved in the informal sector is not a voluntary decision, instead it is a survival strategy for people who are rationed out from the overly regulated formal sector. The neoclassical comparative market view, on the other hand, argues that informal employment can be the desirable option based on people's rational evaluation of the relative costs and benefits of entering the formal system (De Soto, 1989). Fields (2005) and Maloney (2004) argue that the different segments of the informal sector exhibit different properties. They assert that, in the "upper-tier" segment, or "voluntary" as Maloney (2004) puts it, given their individual characteristics, workers voluntarily choose the informal sector expecting relatively high earnings. On the other hand, the "lower-tier" or the "involuntary" is the segment which is seen as the last resort of employment by the workers who are rationed out from the formal sector.

In order to understand the structure of the informal sector in Turkey, the differences between the informal and formal sector wages are decomposed along the entire distribution. To do so, a decomposition technique, which not only allows for correction

of the potential selection bias problem due to self-selection into formal and informal sector in wage estimations, but which also decomposes the informal/formal sector wage gap at any quantile of interest, is applied. The results of the analysis suggest a heterogeneous structure for the informal sector in Turkey in which the lower tier consists of workers who involuntarily participate in the informal sector and the higher tier includes the workers who voluntarily choose to be in the informal sector. It is also found that the wage gap between the formal and informal sectors is higher and the segmentation at the bottom end of the wage distribution is more spread out (i.e. observed in a wider range of quantiles) for females than for males.

CHAPTER 2: EDUCATIONAL ATTAINMENT AND WAGE INQUALITY IN TURKEY

2.1 Introduction

There has been considerable interest in studying the distribution of wages over the last three decades, following the dramatic increase in wage inequality observed across several countries. In particular, the steep increase in the wage gap between college and the high school graduates in the United States has been documented by many authors, such as Katz and Murphy (1992) and Bound and Johnson (1992). In line with these studies, most of the analysis in the literature has focused on wage inequality between workers with high and low educational qualifications. Besides the US, an increasing educational premium has also been documented for many OECD countries (Berman, Bound, and Machin 1997; Machin and Reenen 1998). In these studies, the increasing wage inequality is mostly explained by the supply-demand approach and attributed to increasing skill demand.[4] In this chapter, the relationship between the wage distribution and educational attainment of the workforce is analysed. While a number of studies have looked at the link between wage inequality and the demand for skills, studies analysing the supply of skill and its impact on wage inequality are very limited.[5]

In this chapter, the effect of increasing supply of educated labour on decreasing wage inequality is explored. The focus of the analysis is the Turkish labour market which experienced a substantial decrease in wage inequality between 2002 and 2010. To understand the factors affecting wage inequality in Turkey over this period, the change in the wage distribution between 2002 and 2010 is decomposed into three components, namely, the effect of returns to human capital (price of human capital), the effect of

[4] There are also some authors who assert that institutional factors such as, unionization, minimum wages and collective bargaining practices have substantial effects on the wage distribution (e.g. Card, Lemieux, and Riddell 2003; Lee 1999).
[5] See Naticchioni et al. (2008) for an analysis of the Italian labour market.

changes in the composition of human capital (changes in covariates) and changes in the residual distribution (price and composition of unmeasured human capital characteristics). The decomposition method used in this analysis is proposed by Lemieux (2002) and it is a unified form of the decomposition techniques which were proposed by Juhn et al. (1993) and DiNardo et al. (1996). In order to correct the possible selection bias related to the participation in wage employment, a two-step procedure, proposed by Dubin and McFadden (1984), is also applied.

The results of the study show that the increasing supply of educated labour, which is due to increases in the number of universities and the reform to the education system, which increased the years of compulsory education, had a substantial effect on the wage distribution in Turkey between 2002 and 2010. As a result of the increase in supply, wage inequality decreased at both the top and the bottom of the wage distribution. The decomposition results reveal that decreasing wage inequality was mainly due to decreasing returns to human capital or, in other words, decreasing skill premiums and decreasing residual wage dispersion. Using the human capital approach, it is also found that the decrease in residual wage dispersion was mainly due to the decreasing prices of unmeasured human capital. Finally, the effect of changes in the composition of human capital is found to be moderate.

The plan of the chapter is as follows: Section 2 summarizes the existing literature on developed and developing countries together with other empirical studies on the Turkish labour market. Section 3 presents the underlying theories and macroeconomic indicators of the Turkish economy. Section 4 outlines the methodology. Section 5 presents the data and descriptive statistics. Section 6 shows the wage regression estimations and decomposition results, whilst Section 7 summarizes and concludes and finally Section 8 presents the policy implications of the study.

2.2 Literature Review

Several theories have been proposed to reveal the underlying reasons for the increasing demand for skills. Two of the most popular theories are the Stolper-Samuelson effects and SBTC. The first theory argues that competition with labour abundant countries decreases the relative price of labour-intensive goods and accordingly reduces the real wages of less educated workers in both relative and absolute terms (Hanson and Harrison, 1999). The SBTC theory argues that in accordance with the diffusion of higher technologies of information and communication, labour demand has shifted in favour of skilled workers and increased the skill premiums (Acemoglu, 2002).

Although there is no clear-cut agreement, it is generally believed that the SBTC is the main reason behind increasing wage inequality. Berman et al. (1997) summarized three factors leading to this belief. First, employment shifts to skill intensive sectors have been too small to be consistent with the arguments that are based on product demand shifts, such as sector biased technological changes and Stolper-Samuelson type effects. Second, even though there had been increases in the relative cost of skilled labour in the US, in most of the sectors, shifts in the composition of employment were within the sector rather than between the sectors. Finally, indicators of technological change and skill demand reflect a strong positive correlation within the sectors.[6]

Following the early paper of Katz and Murphy (1992), numerous studies have analysed the changes in skill demand and wage premia in the US. In support of the SBTC hypothesis, various economic models have been developed to explain the channels through which technology affects the demand for skills. Berman et al. (1994) suggested

[6] Krugman (2008), however, argued that most of the empirical evidence for these arguments came from studies which used the data until the mid-1990's. However, since then the trade between the US and developing countries like China and the imports of more sophisticated goods like computers and electronic products has increased. Therefore he asserted that the effect of trade on the wage distribution can be higher than it was in 1990's. However, this is not a crucial point for this chapter, as Krugman argued in the same paper, the Stolper-Samulson process is not expected to be effective in the short run.

that the skill-biased technological change is the most likely explanation for the increase in the demand for white collar workers and in their relative wages. They argued that shifts are mostly due to changes in labour demand within sectors rather than reallocation of employment among sectors. Galor and Tsiddon (1997) argued that the major technological inventions increase the concentration of high ability workers in the technologically advanced sectors while stimulating future technological progress and growth and rising wage inequality. Caselli (1999) asserted that in technological revolutions, learning investments required to use the new machines are higher than those required to use the old existing machines. Therefore, these revolutions cause reallocation of capital from slow to fast learning workers and reduce both the absolute and relative wages of slow learning workers. Krusell et al. (2000) proposed a theory of acceleration in skill-biased technical change. They argued that the elasticity of substitution between capital equipment and unskilled labour is higher than the elasticity of substitution between capital equipment and skilled labour. Based on this argument, they asserted that the increases in wage inequality result from the new and efficient technologies that are embodied in new capital equipment.

As the increasing trend in wage inequality was also observed in other OECD countries, many authors studied the changes in the wage distribution for these countries. One of the most comprehensive papers is Martins and Pereira (2004). In their paper, the authors analysed the effects of schooling on wage inequality for 16 European countries by looking at returns to schooling across quantiles. The main finding of their paper is that more skilled workers are associated with a stronger education-related earnings increment. Accordingly, they argued that schooling has a positive impact upon within-levels wage inequality. Arguably the closest paper to this study is Naticchioni et al. (2008). In their paper, they analysed wage inequality in Italy and found that the change

in inequality was mainly driven by the decreasing educational premiums and stable skill demand.

To a lesser extent, the evolution of the wage distribution has been analysed in developing countries. By using cross-sectional household data for Argentina, Chile, Costa Rica, Colombia, Malaysia, Mexico, Philippines, Chinese Taipei and Uruguay, Robbins (1996) found that trade liberalization was accompanied by a rise in relative wages and the demand for skilled workers. Hanson and Harrison (1995) examined the effects of Mexican trade reforms on wage inequality and they asserted that the rising wage gap was associated with changes internal to industries and that it could not be explained by the Stolper-Samuelson type effects. Galiani and Sanguinetti (2003) analysed the rapid trade liberalization process (which reduced the trade tariffs significantly and eliminated the import licences) that took place in Argentina during the 1990's and tested if the observed increase in wage inequality was triggered by the import penetrations. Their results showed that the trade openness explained only a small proportion of the increase in wage inequality. On the other hand, Berman and Machin (2000) analysed the SBTC in developing countries and found evidence that the demand for skilled labour increased in middle income countries and this increase was mainly due to skill-upgrading within industries rather than a reallocation of employment from low to high-skill industries.

Studies on wage inequality in Turkey are few and not up to date. Kizilirmak (2003) analysed wage inequality in the manufacturing sector for the 1988-2000 period and argued that the change in relative demand for skilled workers was primarily due to the within-industry skill upgrading. Elveren and Galbraith (2009) examined the sub-sectors of Turkish manufacturing by using the between-group component of Theil's T statistic and found an increase in the sector premiums between 1980 and 2001. Tansel (2005),

using the 1994 Household Expenditure Survey, investigated the gender wage gap and the wage gap between public and private sector workers. The results of her study indicated that for men, public sector wages were higher than private sector wages except at the university level, and state owned enterprise wages for men were higher than private sector wages. Also, while wages of men and women were at parity in the public administration, there was a large gender wage gap in the private sector in favour of men. Ozkoc et al. (2011) found an increase in the wage inequality by gender and occupations between the years 2002 and 2006.

Using the quantile regression technique, which makes it possible to analyse the changes in the wage inequality at different parts of the distribution, Tansel and Bodur (2012) analysed the evolution of male wage inequality over the 1994-2002 period. Using the Household Budget Survey (HBS) they found an overall decline in wage inequality in Turkey; wage inequality declined in this period at the lower end of wage distribution while it increased at the top end. Their results showed that education contributed to higher wage inequality through both within and between dimensions; the within-groups inequality increased and between-groups inequality decreased. Therefore, they argued that the latter factor might have dominated the former contributing to the observed decline in male wage inequality over the 1994–2002 period.

In contrast to the above studies, this study uses the most recent data from the Household Labour Force Survey (LFS), instead of the Household Budget Survey (HBS). The main aim of HBS is to collect consumption and expenditure information, although it has some information on the labour market. The LFS has numerous advantages over the HBS: in particular, it has a much wider coverage and more observations which allow us to look at women, who are underrepresented in the labour market; all the national

statistics are based on the LFS, and these are used for policy analysis; and lastly, the main aim of this survey, unlike the HBS, is to look at the labour market.

This study will contribute to the existing literature in two main ways. First, it will analyse the impact of the increase in the education of the labour force on wage inequality: in contrast, the existing literature mainly focuses on other aspects such as technology and labour market institutions. Second, it will be the first application of the Lemieux (2002) model in the context of a developing country.

2.3 Underlying Theories and the Turkish Labour Market

2.3.1 Underlying Theories

In his seminal paper, Mincer (1974) proposed his famous human capital earnings function in which wages are defined as a function of education and experience. Mincer asserted that individuals can accumulate human capital from two channels. First, they can increase their educational qualifications by increasing their schooling level and second they can increase their abilities by on-the-job training (experience). The wage function can accordingly be written as:

$$y_i = \beta X_i + \varepsilon_i \quad (1)$$

where y_i is log wages, X_i is the vector of individual characteristics that determine wages; β is the coefficient vector giving the marginal returns to the covariates in X_i; and ε_i is the random error term. The key individual characteristics that Mincer stressed were education (schooling) and experience (years in the labour market, including but not restricted to, on-the-job training). However, in subsequent empirical work, vector X_i has included a wide range of wage-determining characteristics. Estimating the wage equation, as given by equation (1) is often referred to as the human capital approach.

Most of the empirical literature uses Mincer's wage determination model or its extensions and finds that schooling and experience have significant and positive effects on wages. Accordingly, it is now well-accepted that individuals who make higher investments in human capital receive higher returns from the labour market.

The second important implication of the human capital approach is that if the demand for the skill increases, returns for the unmeasured human capital such as unobservable skills linked to school quality, intrinsic ability and effort, which are the main reasons why workers with the same level of education and experience have different wages, increase as well. In particular, the rate of increase in returns to unmeasured characteristics is greater for individuals who have more education. In econometric terms, residuals in Mincerian-type equations are empirically heteroskedastic. Lemieux (2002) provides a possible explanation for this heteroscedasticity issue: schooling is an imperfect proxy for true educational inputs and the error term includes unmeasured aspects of educational inputs such as school quality. He further argued that systematic differences in the residual variance across education groups arise if the residual dispersion in school quality is different for different levels of schooling. Martin and Pereira (2004) also note that school quality differences are more likely to be prevalent at higher schooling levels, because those are the stages that exhibit greater heterogeneity in schooling paths and school quality. Moreover, differences in school quality and the variance of residuals increase even more dramatically if admissions to schools get more selective at higher levels of education. Therefore changes in skill premiums affect wage inequality in two ways simultaneously; first it affects the wage gap between workers with lower and higher education levels which causes an increase in the between-group inequality component. Second, it affects the wage dispersion

within workers who have the same level of education but studied at schools with different qualities, which causes an increase in the within-group inequality component.[7]

While higher levels of education provide higher earnings at the individual level, an increase in the educational attainment of the whole population does not necessarily mean an increase in returns to education. In his very famous paper, Pritchett (2001) analysed the relationship between rising educational attainment of the labour force and the growth rate of output per worker. Using cross-national data, he found a negative relationship between these two indicators on average. He argued that the marginal return to adding an additional year of schooling in whole population can be substantially different from average returns estimated with a Mincerian regression at a single point in time depending on the shifts in skill demand.[8] For instance, he asserted that marginal returns to education decrease as the supply of educated labour expands if the demand remains stagnant.

One classic example of this condition is given by Schultz (1975). He asserted that rapid technological progresses require constant adaptations to technologically induced disequilibrium and education pays off only if new technologies and inputs are available. Therefore, he argued that in an agricultural environment where technological development is stagnant (so the requirement of adaptation to new disequilibrium is absent), production gains from education will be zero as even the least educated farmers can eventually reach the efficient allocation of factors.

[7] Card and Krueger (1992) provided evidence that men who are educated at higher quality schools have a higher return to an additional year of schooling and returns are also higher for individuals who studied with better educated teachers.

[8] In fact, Pritchett summarized some of the empirical studies showing that the median coefficient for years of schooling increased in the US, decreased in Egypt during the 1980s and fluctuated in Chile between 1960 and 1993.

In their seminal paper, Katz and Murphy (1992) analysed the changes in wage inequality in the US during the 1963-1987 period. Their results show that the college wage premium decreased in the 1971-1979 period in which there was a large increase in the supply of college graduates. On the other hand, the college wage premium increased in the 1979-1987 period in which the growth of the supply of graduates was very small. Accordingly, they argued that, combined with the smooth increase in skill demand, the fluctuations in the growth of the supply of college graduates as a fraction of the labour force played an important role in explaining the large differences in the relative wages of college graduates between these two decades. Based on the findings of Pritchett (2001) and Katz and Murphy (1992), it can be argued that the education premium and wage inequality can increase or decrease depending on the differences between the relative growth rate of the supply and demand for the skills.

In this chapter, the changes in the wage inequality in Turkey between 2002 and 2010 are analysed. Turkey is an interesting case as wage inequality has been showing a decreasing trend despite the rapid industrialisation. The next section documents the recent macroeconomic environment and descriptive statistics on the labour market in Turkey.

2.3.2 Macroeconomic Environment in Turkey

Turkey experienced several major economic crises starting from the 1990s. The first crisis occurred with the gulf war in 1991. The second crisis, which was triggered by the fiscal and external imbalances, occurred in 1994 and a GDP growth rate of negative 6.1 was experienced. After a short period of recovery, due to the adverse effects of the Asian, Russian and Brazilian crises, the Turkish economy experienced a slowdown in 1998 with a growth rate of 3.1 per cent, and then contracted in 1999 at the rate of negative 3.4 per cent. Even though the economy was in boom in 2000 with a 7.3 per

cent growth rate, the heaviest crisis of Turkey's recent history, mainly due to the major capital outflows, which occurred in November 2000 and February 2001 and the GDP declined by 5.7 per cent in 2001 in real terms (Tansel and Bodur, 2012).

In terms of growth rates, the post-2001 period can be defined as the recovery period for Turkey. The real growth rate was 6.2 per cent in 2002 and the economy grew by 6 per cent on average until 2007 when the recent global economic crises first showed its effects on Turkey. On the other hand, in contrast to the fast growth performance across sectors, additional employment could not be generated. The rate of unemployment was 6.5 per cent in 2000 and it increased to 10.4 per cent in 2002. The unemployment rate remained high and never fell below 10 per cent despite the rapid surges in GDP and exports (see Figures 2.A1 and 2.A2). This observation is defined as *jobless-growth* in the literature. Many authors argued that the IMF stabilization program which was signed in 2000 and implemented through the post-2000 period was responsible for the jobless growth. They argued that due to virtually unregulated capital accounts and given the high real rates of interest prevalent in the Turkish financial markets, Turkey experienced massive inflows of short term financial capital. As a consequence, the domestic currency appreciated and Turkey suffered from a growing current account deficit. The appreciated currency stimulated imports together with a contraction of labour intensive, traditional export industries such as textiles, clothing, and food processing. This led to a contraction of formal jobs and increased informalisation of economic activities (Yeldan 2006; Telli, Voyvoda, and Yeldan 2007). The analysis of driving forces of the jobless growth in Turkey is, however, beyond the scope of this chapter.

2.3.3 The Turkish Education System and Composition of the Labour Force

Before the reform in the compulsory education system which came into effect in 1997, the Turkish education system was organized as compulsory primary school (5 years), middle (or secondary) school (3 years), high school and vocational high school (3 years), and university education (two to six years). With the reform in the education system, compulsory primary education was extended to 8 years and the middle school was abolished.[9]

The second important development was the rapid increase in the number of universities. The establishment of new universities began with the second Five Year Development Program (FYDP hereafter) which was put in practice in 1968 (State Planning Organisation, 1967). In this program, it was planned to establish universities outside the three big cities (namely, Istanbul, Ankara and Izmir) where most of the existing (7 out of 9) universities were located. According to this plan, each of the new universities were going to be considered as a separate project and extensive planning of their needs was going to be made for each of them particularly *before* their establishment. Over the following years, the number of universities increased gradually with the establishment of 20 universities between 1971 and 1991. However, in 1991 the number of universities increased suddenly as 24 new universities were established in one year. In that sense, it can be argued that the principle of extensive planning *before* the establishment of new universities was violated. For instance, according to the report on higher education, which is a part of the 9th FYDP and was published in 2006 (State Planning Organisation, 2006), the physical completion rate of the universities that were established 1991 was less than 50 per cent.

[9] The Turkish education system was reformed again in 2012 and a more complex system known as the ‘’4+4+4 system’’ has been adapted.

In accordance with the 2nd FYDP, a report by the State Planning Organisation, which was published in 1970 (State Planning Organisation, 1970), suggested that the new universities should be established in the cities which could serve as development centres and spread out the development inducing effect of the new universities. Accordingly, nearly all of the universities established after 1972 were outside the three big cities. There were no references to a university in each city in any of the programs or reports from before 2000's; however the new universities became the subject of political campaigns starting from the 2000's. With the "a university in each city" project the number of universities increased substantially in the post 2000 period. The fast increase in the establishment of the universities has been heavily criticised by many authors. Kaynar and Parlak (2005) argue that while the impact of the new universities on the cities has been exploited by the politicians, these universities were established without proper planning and before minimum physical requirements were met. Arap (2010) notes that some of the universities in this period were established by simply changing the name of the faculties which were originally a part of an old university located in neighbouring cities.

As a result of the "a university in each city" policy and growing private sector participation, in 2011 the number of universities reached 165. As a consequence of these new universities, the number of university students also increased substantially (see Figures 2.A3 and 2.A4). For instance, from 1994 to 2011, the number of students in formal university education increased by more than 300 per cent and rose from 1.2 million to 3.7 million (Turkish National Statistics Institute; TurkStat hereafter).

The fast increase in the number of universities in Turkey has been accompanied by a rising difference between old and new universities in terms of equipment, funding and resources which can be expected to lead to differences in the quality of education. In

fact, Hatakenaka (2006) argued that the new universities which are not located in metropolitan centres find it difficult to recruit qualified staff because they are unwilling to relocate to these regions even though several measures were tried in the past to address these issues.[10] She asserted that deeper structural changes were needed before any further expansion as the student/staff ratio for Turkey is one of the highest even among comparator developing countries. In addition, due to the central university entrance examination system, more successful students have the chance of studying with other successful students in the universities with better resources and this leads to further skill differences between university graduates.[11]

Considering these two developments, it can be argued that the supply of educated labour continuously increased starting from the early 1990's. In fact, Figures 2.A5, 2.A6 and 2.A7 show that share of individuals with higher education in the labour force (employed and unemployed) increased. For example, according to the TurkStat web database, between 1988 and 2011 the share of individuals with university education in the labour force increased from 4 per cent to 21 per cent for women and 5 per cent to 15 per cent for men. For the same period, the share of individuals with education less than high school in the labour force decreased for both men and women. Similar patterns are also observed in the composition of the employed and unemployed. One crucial observation is that, for men, the share of university graduates among the unemployed has become even higher than the share of high school graduates since 2007. Based on these observations, a decrease in the educational premium can be expected: the increasing supply of labour with higher education is likely to decrease the upward

[10] For instance, there was a requirement for professors to 'serve' in outer areas before being promoted. Today, there is a salary supplement to provide incentives for people to work in these universities.

[11] There are also teaching quality differences between high schools in Turkey. Apart from vocational high schools, there are 3 different types of public high schools in Turkey, namely science high schools, Anatolian high schools and general high schools which have different levels of selectiveness in their admissions.

pressure on the wages of university graduates while the decreasing share of workers with low levels of education moderates the downward pressure on the wages of those workers.

2.4 Methodology

2.4.1 Decomposition Method

The literature on decomposition of wage inequality goes back to the seminal papers of Oaxaca (1973) and Blinder (1973). In their models they showed how to decompose the change in the mean of wages into two components; the effect of changes in the distribution of characteristics (changes in the distribution of human capital) and the effect of changes in the regression coefficients (changes in the market prices of human capital). The Blinder-Oaxaca decomposition can be shown as:

$$y_{it} = x_{it} b_{it} + u_{it} \tag{2}$$

$$y_{is} = x_{is} b_{is} + u_{is} \tag{3}$$

$$\bar{y}_t - \bar{y}_s = \bar{x}_t (b_t - b_s) + (\bar{x}_t - \bar{x}_s) b_s \tag{4}$$

Where y_i is the log wage for individual *i*, $\bar{y}$ is the mean wage for the periods t and s, b is the vector of OLS estimated coefficients for the returns to human capital, x_i is the vector of covariates representing the human capital characteristics, and u_i is the regression residuals. The first term on the right hand side of equation (4) shows the changes the mean of wages due to the changes in regression coefficients (price effect) while the second term shows the change in the mean of wages due to the changes in covariates (composition effect).

Following their seminal work, many other models have been developed over the last two decades. One of the most popular models came from Juhn et al. (1993) (JMP hereafter). The JMP method became popular as, unlike Oaxaca-Blinder method, it takes into account the distribution of residuals with the inclusion of the residual imputation procedure. However, the main problem with this model is that it does not account for changes in the distribution of covariates. Although in their original work JMP formally allow for the distribution of residuals to be dependent on covariates, they did not explicitly explain how to implement it empirically.[12] What they do, in their model, instead is to calculate the price effect as above (equation 4) and use their residual imputation method to account for the changes in the distribution of residuals and finally capture the effect of changes in characteristics as the part unexplained by the price and residual effect components. However, as mentioned in the previous section, error terms are not identically and independently distributed in Mincerian models. In other words, this type of modelling suffers from a heteroscedasticity problem. Therefore, the results of the JMP model can be misleading since this decomposition is efficient only in the case of homoscedastic residuals (Melly, 2005).

This study will make use of a decomposition method developed by Lemieux (2002). This model unifies the residual imputation method by JMP and the re-weighting factor method by DiNardo et al. (1996). The model has several advantages compared to the other decomposition methods. First, it allows for the decomposition of changes in the entire distribution of wages rather than the decomposition of the change in the mean wages only. In this model, it is possible to reveal the changes at different moments of the wage distribution. Second, unlike JMP it accounts for the distribution of covariates.

[12] In their original paper JMP defined the residual of the distribution as $u_{it} = F_t^{-1}(\eta_{it}|x_{it})$ although many applications of the JMP decomposition defined $u_{it} = F_t^{-1}(\eta_{it})$ where η_{it} is the percentile of an individual residual in the residual distribution and F_t^{-1} is the inverse of the cumulative distribution function.

Therefore, it also accounts for the problem of heteroskedasticity. Finally, it is also possible to decompose the changes in the residual distribution into the effect of the changes in unobservable characteristics and the effect of changes in returns to those characteristics. Therefore, it is possible to test the hypothesis of the human capital approach which asserts that a positive change in returns to observable skills exerts a positive impact on returns to unobservable skills.

The first step of the Lemieux method is to estimate separate Mincerian wage regressions for each year:

$$y_{i2010} = x_{i2010} b_{2010} + u_{i2010} \quad (5)$$

$$y_{i2002} = x_{i2002} b_{2002} + u_{i2002} \quad (6)$$

To compute the effect of changes in the prices of characteristics, following JMP, a counterfactual wage vector that would prevail in 2002 (the base year) if the price of human capital were the same as in 2010 is constructed. To get the counterfactual wage vector, the coefficients from the 2002 wage regression are replaced with the coefficients from 2010 such as[13]:

$$y^a{}_{i2002} = x_{i2002} b_{2010} + u_{i2002} \quad (7)$$

Once the counterfactual wages $y^a{}_{2002}$ are constructed, it is straightforward to obtain the share of the effect of changes in prices in the total change of the wage distribution by comparing any inequality measure (such as Gini, Theil index, variance or wage gaps) for $y^a{}_{2002}$ and empirical wages y_{2002}.

[13] Similar to the Blinder-Oaxaca and JMP methods, the Lemieux method is also path dependent. In other words, the results of the decomposition are sensitive to the choice of the base year which is 2002 here. However, there is no 'correct' choice of base year as the choice of either year is equally valid a priori.

The second step of the method is to create the re-weighting factor which will be used to modify the original sample weights and calculate the effect of changes in the covariates. The idea of modifying the sample weights, which are used to calculate sample statistics representative of the population, was originally proposed by DiNardo et al. (1996). This method basically attaches a new counterfactual weight to each individual to keep the distribution of characteristics at their 2010 level and thereby it makes it possible to account for changes in covariates. Their idea is to pool the samples of the two periods and estimate a logit or probit model for the probability of being in the base year.[14] Then the re-weighting factor is calculated by the propensity score which is the predicted probability that an individual in the pooled sample comes from the base year conditional on covariates:

$$P_{i2002} = \text{Prob}(period = 2002 | x_{i2002}) \tag{8}$$

And the re-weighting factor is defined as

$$\psi_i = \frac{1 - P_{i2002}}{P_{i2002}} \frac{P_{2002}}{1 - P_{2002}} \tag{9}$$

Where P_{2002} is the unconditional probability that an observation is in year 2002 (the weighted share of the pooled sample that is in 2002). Finally the new sample weights are computed by multiplying the original sample weights (w_i) with the re-weighting factor:

$$w^a_{i2002} = w_{i2002}\psi_i \tag{10}$$

Following the notation of Lemieux (2002), the counterfactual values of wages that will be generated by using the new sample weights are summarized in Table 2.1. The

[14] Both the OLS regression and the logit model use the same set of covariates.

difference between the distributional statistics (e.g. Gini, Theil index) of $y^{a}{}_{2002}$ and y_{2002} (using the original sample weights w_{i2002} from the 2002 data) gives the coefficient effect. This corresponds to the comparison of the distributions presented in the second and the first rows of the Table 2.1. On the other hand, the distribution statistics that are calculated by using $y^{a}{}_{2002}$ and the new sample weights w^{a}_{i2002} contain the covariates effect as well as the coefficient effect. Accordingly, the covariates effect is calculated as the difference between the distributional statistics that are obtained by using $y^{a}{}_{2002}$ with the original sample weights w_{i2002} and with the new sample weights w^{a}_{i2002}. This corresponds to the comparison of the distributions presented in the fourth and the second rows of Table 2.1.

The final stage of the decomposition is to calculate the effect of changes in the residuals on the change in the wage distribution between the two periods. As mentioned earlier, Lemieux (2002) uses the residual imputation method which is provided by JMP.

With a strict human capital interpretation of wages, JMP argue that residual wage dispersion is mainly due to the imperfect measurement of human capital in standard datasets. Accordingly, assuming a stable distribution of unmeasured skills over time, changes in residual wage inequality must be the result of a change in the returns to these unmeasured skills.

To summarize the residual imputation method, consider the following simple form of the model for wage residuals:

$$u_{i2002} = p_{2002}\eta_{i2002} + \varepsilon_{i2002} \qquad (11)$$

where η_{i2002} is unmeasured human capital, p_{2002} is the return to unmeasured human capital and ε_{i2002} is a random error term not linked with skills. Then the variance of residuals can be calculated as:

$$\sigma^2_{2002} = p^2_{2002}\sigma^2_{\eta,2002} + \sigma^2_{\varepsilon,2002} \quad (12)$$

Where $\sigma^2_{\eta,2002}$ = Var (η_{i2002}) and $\sigma^2_{\varepsilon,2002}$ = Var (ε_{i2002}). In this model, with the assumption that the distribution of unmeasured skills is constant ($\sigma^2_{\eta,2002} = \sigma^2_{\eta,2010} = \sigma^2_{\eta}$**)**, and $\sigma^2_{\varepsilon,2002}$ is zero or stable over time, changes in skill prices are the only source of any change in residual wage inequality.

$$\sigma^2_{2002} - \sigma^2_{2010} = (p^2_{2002} - p^2_{2010})\sigma^2_{\eta} \quad (13)$$

The main disadvantage of this form is that the residuals are assumed to be a linear function of unmeasured skills. To understand the effect changes in unmeasured human capital prices have on the wage distribution, JMP propose a more general setting in which a non-linear pricing scheme is applied:

$$u_{i2002} = p_{2002}(\eta_{i2002}) + \varepsilon_{i2002} \quad (14)$$

Where $p_{2002}(.)$ is a monotonic and continuous function; and for simplicity ε_{i2002} is assumed to be zero.

The JMP model, as given in equation (14), is more general compared to the form shown in equation (13) since it provides more flexibility by making it possible to generate any distribution of u_{i2002} from an arbitrary distribution of skills η_{i2002}. For simplicity, assume that the η_{i2002} follows a uniform distribution over the interval of [0, 1] since:

$$\eta_{i2002} = F_{2002}(u_{i2002}), \quad (15)$$

Where $F_{2002}(.)$ is the cumulative distribution function of u_{i2002}. Combining equations (14) and (15) we get:

$$u_{i2002} = p_{2002}(\eta_{i2002}) = F_{2002}^{-1}(\eta_{i2002}), \quad (16)$$

Where η_{i2002} can be interpreted as the rank of observation *i* in the distribution of residuals while the non-linear skill pricing function $p_{2002}(.)$ is the inverse cumulative distribution of u_{i2002}.

By using the skill pricing function, the counterfactual wages (equation (7)) can be rewritten as

$$y_{i2002}^{a} = x_{i2002}b_{2010} + u_{i2002} = x_{i2002}b_{2010} + p_{2002}(\eta_{i2002}) \quad (17)$$

The decomposition is finalized by replacing the residuals in 2002 with the residuals that would prevail if the skill pricing function was $p_{2010}(.)$ instead of $p_{2002}(.)$ such as

$$y_{i2002}^{a} = x_{i2002}b_{2010} + p_{2010}(\eta_{i2002}) = x_{i2002}b_{2010} + u_{i2002}^{b} \quad (18)$$

Where $u_{i2002}^{b} = p_{2010}(\eta_{i2002}) = F_{2010}^{-1}F_{2002}(u_{i2002})$ is the counterfactual residual for the observation *i.*

The imputation method used to compute counterfactual residuals may seem a bit complicated. However, JMP suggest a simple procedure. First, the rank $\eta_{i2002} = F_{2002}(u_{i2002})$ is computed from the empirical residual distribution in 2002 and then the residual at the same rank in the residual distribution in period 2010 is picked.[15]

[15] Unless the numbers of observations are equal in two periods, it is not possible to match residuals exactly. Lemieux (2002) suggested a simple solution to solve the problem. His idea is to discretize the distribution of residuals in *k* intervals which contain the same number of observations and replace the actual residuals by the average residual in each interval. In this analysis *k* is chosen as 500.

The obtained counterfactual wages now can be used to decompose the changes in wage inequality. Extending JMP, they can also be combined with counterfactual weights to control for the distribution of covariates. In addition, having the counterfactual wage vectors, it is possible to calculate any measure of inequality to see the effects of different factors on different parts of wage distribution.

The last but not the least important feature of the decomposition method is that it is also possible to decompose the changes in wage residuals. Since sample weights are used to calculate indices that are representing the population, comparison of variances of residual wages using the original sample weights and counterfactual weights provides information about how much of the change in residual distribution is due to the change in covariates and the change in the skill pricing function.

2.4.2 Selection Bias

A selection bias problem occurs in many applications of econometrics and a well-known example of it is observed in wage regression estimations. By definition, wage estimations are performed for the individuals who have reported their wages. Therefore, while estimating the returns to human capital, say schooling and labour market experience, selection bias may occur if some individuals who enrolled in schooling level do not work. The reason may be that these individuals who have relatively low levels of schooling will receive low wage offers which may be lower than their reservation wages. As a result, employment levels for these individuals will be low and selection bias problems will occur as one will only observe the wages of individuals with high levels of education and high wages. The important consequence of this problem is that simple OLS estimation of wages on human capital will be biased as the sample selection is not random and the sample is not representative of the whole population which is the subject of interest (Heckman, 1976, 1979).

According to TurkStat, nearly 30 per cent of working men were self-employed in the year 2000. This share decreased to 23 per cent in 2010. However, this considerable share of the self-employed is an indicator of non-random participation into the wage sector. The selection bias problem becomes even more crucial in wage estimations for women. Over the last 50 years, Turkey's female labour participation has been decreasing (Goksel, 2012). According to the Global Gender Gap Report 2013 (World Economic Forum, 2013), the female labour force participation rate in Turkey is 30 per cent and this ratio puts Turkey in the 123rd place out of 136 countries. Another problem that may cause selection bias is that there is a substantial share of unpaid family workers who traditionally work in the agricultural sector and are recorded as employed in employment statistics. According to TurkStat, for the year 2002, 49 per cent of total female employment consists of unpaid workers. However, these individuals do not report any form of labour income and so they are automatically omitted from the sample. Based on these issues, the selection bias problem should be considered while estimating wage equations in the case of Turkey.

Heckman (1979) proposed two estimation techniques to correct for the self-selection problem; one requires maximum likelihood (ML) estimation of a selection model assuming bivariate normality of the error terms in the wage and participation equations. The other one consist of two steps. In the first step of the process, a probit model is performed with the whole sample to estimate the inverse Mill ratio, λ, which is a monotone decreasing function of the probability of being in the subsample of participants. Then the predicted probabilities obtained from the probit model, λ, are included in the OLS (or GLS) estimation of the wage equation as an additional regressor in the second step. Due to its easy-to-implement feature this method has become the standard estimation procedure for empirical wage equations.

Due to the existence of unpaid family workers and the self-employed and the different characteristics of these individuals from those who simply do not participate in the work force, usage of a simple probit model in the first step of the correction process can be misleading. This is because, just like labour force participation, the decisions to be an unpaid family worker or to be self-employed are also not random. It is more accurate to account for the differences between these groups in the first step of the correction process.

Dubin and McFadden (1984) proposed an alternative two step method to account for selection bias for the case where the number of choices is greater than two[16]. They suggested using a multinomial logit model for the first step of the correction process. Following their method, the whole sample of woman is grouped into three categories. Individual *i* is categorized as $fw_i = 2$ if working in the wage sector, $fw_i = 1$ if she is an unpaid family worker and $fw_i = 0$ if she is economically inactive (a non-participant). Then, the probability of being in group *j* is obtained by estimating a multinomial logit model:

$$P_{ij} = P(fw_i = j \mid Z_i) = \frac{\exp(Z_i \alpha_j)}{1 + \sum_{j=0}^{2} \exp(Z_i \alpha_j)} \qquad \text{for j=0,1,2} \tag{19}$$

where the numerator is normalized to 1 for $j = 0$, $P_{ij} = P(fw_i = j|Z_i)$ is the conditional probability of individual *i* being in group *j*, conditional on Z_i, a vector of covariates and α_j is a column vector of coefficients corresponding to the *j*th sector.

[16] The Dubin and McFadden (1984) method is chosen over the alternative Lee (1983) method as the latter includes strong assumptions on the covariance between the residuals of the selection equation and the wage equation. Bourguignon et al. (2004) note that these assumptions are often violated in empirical applications, and using Monte-Carlo simulations they show that the Lee method performs worse relative to the Dubin and McFadden method.

Then the predicted probabilities from the multinomial logit model are used to construct a selection correction term, λ_{ij}, by using the following formula provided by Hill (1983);

$$\lambda_{ij} = \left(\frac{6}{\pi^2}\right)(-1)^{J+1}\left[\frac{J-1}{J}\ln\hat{P}_{ij} + \sum_{k\neq j}\left(\frac{\hat{P}_{ik}\ln\hat{P}_{ik}}{1-\hat{P}_{ik}}\right)\right] \qquad \text{for j=0,...J} \tag{20}$$

The same model is applied to the case of men and the only difference is that the $fw_i = 1$ represents self-employment instead of unpaid-family work.

In the second step of the model, augmented wage equations are estimated separately for men and women by including the sample selection correction term in the set of covariates as an additional regressor.

$$\log(Y_{ij}) = X_{ij}\beta_j + \theta_j\lambda_{ij} + \varepsilon_{ij}, \tag{21}$$

where Y_{ij} are the wages of individuals in group j, X_{ij} is the vector of individual characteristics, β_j is the estimated vector of returns to those characteristics in group j, θ_j is the unknown coefficient related to the selection correction term and ε_{ij} is the independent residual term. The augmented wage equation estimation results, then, are used in the decomposition of the change in the wage distributions of men and women between 2002 and 2010.[17]

Wooldridge (2010) notes that even if the selection equation and the wage equation have the identical regressors (i.e. $Z=X$), the coefficients in the augmented wage equation can be identified by the non-linearity of the selection correction term. However, he also

[17]When decomposing the wage differentials in the presence of selection bias, we need to make assumptions about the selection term ($\theta_j\lambda_{ij}$). One option is to decompose the selection corrected wages, defined as $[\log(Y_{ijt}) - \log(Y_{ijs})] - [\theta_{jt}\lambda_{ijt} - \theta_{js}\lambda_{ijs}]$; for details see Neuman and Oaxaca (2004). Another option, used in this chapter, is to treat the selection term as an additional explanatory variable; hence it is the change in the observed wage distributions which is decomposed.

notes that relying solely on non-linearity can introduce severe collinearity among the regressors in the regression of the augmented wage equation. For identification, one needs a situation where there is at least one element of *Z* which is excluded from *X*. In other words, there should be at least one variable that affects the labour market status but does not affect the wages. [18]

Considering this issue, three dummies are added to the covariates in the selection equation to capture the labour market status of other members of the household: presence of wage workers, presence of self-employed, and presence of unpaid family workers in the household. The justification for using the labour market status of other members of the household in the selection equation comes from the well documented evidence on the importance of social networks in determining the labour market outcomes. For instance, Calvo-Armengol and Jackson (2004) argue that each individual is connected to others through a network (e.g., kinship, friendship, and acquaintanceship). He also argues that, when information about jobs arrives, individuals, who are unemployed and directly hear of a job, use the information to obtain a job. On the other hand, individuals who are already employed, depending on whether the job is more attractive than their current job, might take the job or else might pass information to one (or more) of their direct connections in the network. In accordance with this argument, it is possible that a wage worker (or self-employee) may share the information about a waged employment (or self-employment) opportunity with his or her household members who are unemployed or self-employed.

In the selection equation for women, the presence of the grandmother in the household is also added to the regressors. The presence of other adults in the household can aid in

[18] Due to the lack of information about tenure and membership of the social security system for the self-employed, unpaid family workers, and economically inactive individuals, these variables are only included in the wage equations.

the labour force participation of women, by providing child care; however, it can also be detrimental, as other adults may be either considered as labour market substitutes for women or additional care responsibilities for women. In this case a dummy for the presence of grandmother in the household can aid in labour force participation of women by providing child care, but at the same time can be a deterrent for labour force participation as an elderly relative in the household may be an additional care responsibility for the women. For instance, using US data, Ettner (1996) shows that caregiving responsibilities for elders is a significant detriment to labour force participation for women. In case of China, on the other hand, Maurer-Fazio et al. (2011) found that the presence of co-residence with elders had a positive effect on the labour force participation of women, especially married women. Similarly Marenzi and Pagani (2005) found that in Italy presence of elderly parents in the house can be beneficial for the labour force participation of women, especially those with pre-school children.

2.5 Data and Descriptive Statistics

2.5.1 Data

The data sources are the two waves of the LFS which is conducted by TurkStat for the years 2002 and 2010. The first LFS was conducted in year 2000. However, to avoid the distortions in the Turkish labour market due to the economic crises which occurred in 2000 and 2001, the 2002 LFS data is used in the analysis as the starting period. At the time of the analysis, the last wave of the LFS that was available was 2010, which is used here. The LFS has the largest sample from the Turkish labour force and contains information on both the workers' demographic characteristics and the characteristics of their main job for each individual in the household. For the analysis, two cross sections

of workers aged between 20 and 64 are created.[19] Following that, individuals who are students, ill, disabled or retired are dropped from the sample. As the LFS does not provide any information on the wages that are earned from any additional jobs, individuals who have an extra job are dropped. Women who are self-employed and men who are unpaid family workers are also dropped from the sample (see Table 2A.1 for the percentage share of the omitted groups in the full sample). From the resulting sample, a new category, namely "economically inactive", is created for the individuals who do not participate in the labour force and those who are unemployed. The selection equation estimations are performed with 78,040 and 132,894 female, 58,204 and 104,650 male individuals for the years 2002 and 2010, respectively. The wage equation estimations are performed with 9,191 and 18,644 female workers and 32,226 and 59,696 male workers for the years 2002 and 2010, respectively.[20]

The educational qualifications of workers are grouped in 7 categories namely, illiterate, literate without formal education, primary, secondary, high school, vocational high school and university. In 1997, compulsory primary education in Turkey was increased from 5 to 8 years and so the "middle school" which corresponds to the 3 year education between primary and high school level was abolished. However, since there are still a substantial number of workers with 5 years of primary education, an additional category "secondary education" is created to aggregate the middle school and 8 year primary education categories. Both of these education levels correspond to the same years of

[19] The reason for choosing this age band is that the decision to pursue higher education is endogenous as it depends on the returns to higher education. However individuals who are aged 20 or older are the ones who have already made their decisions.

[20] The reason for the higher number of observations in the 2010 data is that the TurkStat increased the sample size of the LFS gradually from 2002 to 2010

schooling (8 years) and theoretically these workers can be expected to have the same qualifications as the investment they make in human capital is equivalent.[21]

The LFS contains information only on net monthly wages. This is not a big disadvantage since Tansel (2005) suggests that monthly wages convey fewer recording errors compared to annual wages. In order to get hourly wages, monthly wages are first divided by 4.3 and then divided by the usual hours of work per week. Finally, hourly wages are deflated by the consumer price index which is also provided by the TurkStat to obtain real hourly wages.

2.5.2 Descriptive Statistics

Table 2.2A presents the descriptive statistics of the variables for each sample year for women. Wage earning women are the most educated, with 40 per cent having a university education in 2010. Unpaid family workers are the least educated, with a third of them having no formal education. Between 2002 and 2010 there has been a shift in the distribution of education for the wage earners, with a fall in the wage earners with primary (high school) education by 6 percentage points (4 percentage points), and an increase in the university educated women by 8 percentage points. Wage earning women are more likely to be heads of their household, and are more likely to come from households where other members of the household are also wage earners. The mean values of the variables that represent presence of children in the household are smaller for wage earning women. Women in unpaid work tend to come from rural households where other members of the household are self-employed.

Table 2.2B presents the descriptive statistics of the variables for each sample year for men. Wage earning men are more educated relative to the other two groups

[21] Years of schooling is the most popular proxy that is used for capturing educational qualifications in the literature.

(economically inactive and self-employed). There has been a shift in the distribution of education between 2002 and 2010, towards university educated workers in the wage sector; the share of workers with primary education decreased by 8 percentage points while the share of workers with a university degree increased by 5 percentage points. There is not much difference in the mean values of the variables that represent presence of children in the household between self-employed men and wage earning men; for the economically inactive men on the other hand, the values of these variables are smaller. Self-employed men have a high proportion of other members of the household who do unpaid family work (possibly women in the family).

Looking at the descriptive statistics presented in Tables 2.2A and 2.2B, it can be seen that the change in the compulsory education system and the increasing number of universities are associated with a substantial change in the educational composition of wage workers (and the workforce in general). The new education system, which abolished the primary school and made secondary education compulsory, reduced the number of workers with primary education only for both man and women. With the increasing accessibility[22] of university education, the number of workers with a university degree increased for both genders. Based on these observations, it can be argued that the supply of educated labour increased between 2002 and 2010.

The mean wages for both men and women have increased over time; with a higher increase in the mean wages of women. There has been a fall in the variance of wages for both men and women over time. The mean wage for men is higher than the mean wage for women in 2002 and this relationship reverses in 2010. This observation does

[22] Before their abolishment in 2012, the tuition fees for university education in Turkey varied across different departments. For instance, the range of tuition fees in 2010 was 71-591 Turkish Liras per year (which corresponds to 47-394 US dollars with the exchange rate on January 1st 2010). Student loans were provided by the government and their interest rate was fixed to the wholesale price index.

not necessarily imply positive discrimination towards women. For instance, using the Structure of Earnings Survey 2006 conducted by the Turkish Statistical Institute, Kaya (2010) shows that on average women earn 2.5 percentage points more than men. She also argues that the higher average wage for women is a consequence of a composition effect as most of the female employees have a university degree while male employees mostly have only primary education. By exploiting a quantile decomposition method, she also shows that, while the composition effect had a narrowing effect on the gender wage gap (favouring women), the effect of differences in the returns to characteristics had an increasing effect on the gender wage gap (favouring men) showing that the human capital characteristics of women are rewarded less than the characteristics of their male counterparts. Considering Kaya's (2010) findings, the reversal of relative wages between men and women between 2002 and 2010 can be explained by the higher increase in the share of university graduates amongst female wage workers. As noted above, the share of university graduates amongst female wage workers increased by 8 percentage points while the size of the increase was 5 percentage points for males.

2.6 Empirical Results

The first step in the empirical investigation is to look at the multinomial logit model for labour market status. Tables 2.3A and 2.3B show the relative risk ratios (RRR) from the maximum likelihood estimation of the multinomial logit model for women and men, respectively. The base category of the model is the group 'economically inactive' for both men and women. The RRR tells us how the probability of choosing wage employment or unpaid family work (self-employment in the case of men) relative to being economically inactive changes if the independent variable increases by one unit. If the RRR is greater than 1, it means that the individual is more likely to be in wage

employment or unpaid family work (self-employment for men) and accordingly, if it is lower than 1, the individual is more likely to be economically inactive.

Table 2.3A shows that for women, education at all levels increases the probability of being in wage employment. Having other members of the household in the wage sector, being head of the household, and (for the year 2002) having a grandmother in the household (this will potentially capture help available to women, within the household, for childcare and other household responsibilities) also increases the probability of being in wage employment. Presence of children in the household aged 5 or less, being married, and presence of self-employed members in the household all have a decreasing effect on being in wage employment.

Table 2.3B shows that for men, education at all levels, presence of children in the household aged 5 or less, other members of the household in the wage sector, and being the head of the household increase the probability of being in wage employment. Living in a rural area, and the number of unpaid family workers in the household all increase the likelihood of being self-employed.

An interesting result that can be drawn from Tables 2.3A and 2.3B is that for both men and women, the RRR's are much smaller in 2010 than they were in 2002. This means that education does not increase the probability of being a wage worker (as opposed to being economically inactive) in 2010 as much as it did in 2002.

To see the changes in the returns to these human capital characteristics, OLS regressions are estimated for each year and also for men and women. The regression results (see Table 2.4) show that the coefficients for the education variables decreased from 2002 to 2010 for men and women. The percentage decrease is more at the lower education levels than at higher education levels. For example, while returns to

secondary education for women (men) decreased by 86 per cent (76 per cent) over time, considering 2002 as the base year, the returns to university education decreased by 43 per cent (28 per cent). The dramatic decreases in returns to education suggest that the market price adjustment mechanism gave a sharp response to the increase in the supply of educated labour by decreasing the returns to education.

According to the wage regression results, compared to the base category "never married", being married has a positive effect on wages. Also working in the formal sector (which is captured by membership of the social security system) has an increasing effect on wages. The coefficient for the selection term is positive for men and women and statistically significant in each year, which means that the unobservable factors that affect selection into wage employment are positively correlated with the unobservable factors that affect wages.

The descriptive statistics and the regression results show that the composition of the human capital of wage earners and returns to the human capital changed significantly between these two years. However, it is not possible to reveal to what extent these changes were responsible for the changes in the wage distribution without decomposing the changes in wage inequality. Tables 2.5A and 2.5B present the decomposition results for women and men, respectively.

The top two rows of Tables 2.5A and 2.5B give the measures of wage inequality for the two years. Wage inequality is higher for women than men. Whichever measure is looked at, wage inequality decreased for both men and women over this period; the decrease is greater at the bottom part of the distribution (p50/p10) than at the top half of the distribution (p90/p50).

The decomposition results for both women and men (Tables 2.5A and 2.5B) suggest that the decreases in the returns to observed characteristics (the effect of coefficients; change in the betas), have a substantial share in the decreasing inequality in this period. The effect of this change in decreasing inequality is very high in both the upper and lower tails of the wage distributions of men and women. For women, the fall in the returns to the unmeasured characteristics (pricing function) has a considerable share at the top end of the distribution while the size of this effect is found to be very small at the lower tail of the distribution. The effect of pricing in the wage distribution of men is also negative and relatively high at the upper tail.

The effect of changes in the composition of characteristics is negative and relatively small for women. The change in characteristics (an increase in education being one of them), holding returns same, decreases wage inequality for women. For men, the same holds true, with the exception of changes in inequality in the upper tail, where a change in characteristics, holding returns constant, leads to an increase in wage inequality. Part of the explanation for the differences in this finding for men and women could be due to the fact that the proportion of women with university education (who are most likely to be in the upper tail) working in the wage sector is much higher than the proportion of men in the wage sector with university education. A further increase in the supply of educated women dampens the returns to them, whereas for men it does not.[23]

One important observation is that the effect of changes in the composition of human capital is moderate compared to the size of other effects. Considering the rapid increase in the educational level of wage earners, this result may seem surprising. However, as

[23] Robinson (1976) presents a two sector model where he shows that there exists an inverted U-shaped relationship between wage inequality and the proportion of skilled labour; as the proportion of skilled labour increases, wage inequality first increases and then decreases. The turning point of the relationship is an empirical question.

Naticchioni et al. (2008) noted, use of the partial equilibrium assumption that aggregate quantities of covariates do not affect market prices may cause misleading interpretations and this assumption is a major drawback of the wage inequality literature. The market price adjustment mechanism can play quite an important role in wage distributions and the small effect of changes in the human capital composition is most likely to be the result of market price adjustment to these changes. In other words, the price of human capital is determined by supply and demand jointly. If there is an increase in the supply of human capital, which is the case here, the market price of human capital decreases if such an increase is not balanced with an increase in the demand. Therefore, even if the effect of changes in the covariates is found to be very small here, increasing supply of human capital had a crucial role by triggering the market price adjustment mechanism.

The last rows of Tables 2.5A and 2.5B present the residual change in the wage distribution which could not be explained by these three factors. Lemieux (2002) argued that when a random variable like y is the sum of two random variables like xb and u, knowing the distributions of these two variables is generally not enough to characterize the marginal distribution of y. Generally, an infinite number of marginal distributions of y is compatible with given marginal distributions of xb and u.

Looking at the last two columns of Tables 2.5A and 2.5B, the decomposition results indicate that the systematic part of the wage equation (the sum of the covariate and coefficient effects) could not explain around 40-50 per cent of the change in the Gini and Theil index for men. Although the size of this effect is smaller in the case of women, there is still a considerable part, around 25-30 per cent, which is explained by the skill pricing function of the unmeasured skills.

Tables 2.6A and 2.6B present the decomposition of changes in wage residuals for women and men, respectively. The reported results show that the variance of residuals

decreased by around 50 per cent for both men and women. Also the residual wage gaps decreased in accordance with the decrease in variance. The changes in the composition of unmeasured skills had an inequality decreasing effect for men while it was inequality increasing for women. Nevertheless, in both cases, the magnitude of this effect is once again negligible. The skill pricing function also explained almost all of the variation in the change in the dispersion of the wage residuals. This is evidence that decreasing returns to human capital were accompanied by a decrease in the pricing function of unmeasured characteristics.

2.7 Conclusion

In this study, the relationship between the human capital composition of employment and wage inequality is investigated using data on the Turkish labour market. Using the decomposition methodology proposed by Lemieux (2002), it is found that the decreasing wage inequality observed in Turkey between 2002 and 2010 was mainly due to two factors: The decreasing between group inequality which is related to the decreasing coefficients for education and the decreasing within group inequality due to the decreasing skill pricing function of unmeasured skills. On the other hand, the role of the change in the distribution of the covariates is found to be negative and small.

The results of the study show that the increasing supply of educated labour, which is attributed to the increase in the number of universities and the reform in the education system which increased the years of compulsory education, had a substantial effect on the wage distribution in Turkey between 2002 and 2010. As a result of the increase in supply, wage inequality decreased in both the top and the bottom half of the wage distribution. It is also found that the decrease in wage inequality was relatively low in the top half of the wage distribution. A potential explanation for the lower decrease in wage inequality in the top half of the distribution could be the SBTC effect which

possibly had a positive effect on the skill demand. However, the decreases in returns to education and the decreasing wage inequality indicate that the skill demand increasing effect of SBTC fell behind the effect of large increase in the supply of educated labour.

The main limitation of the analysis is the lack of information in the data about the changes in skill demand. A possible strategy to understand the changes in skill demand could have been to look at the changes in the relative share of occupations with different skill requirements. Unfortunately, the LFS data includes only the one-digit-level of International Standard Classification of Occupations (ISCO88) classification which has nine categories. Due to the one-digit-level classification, each of these categories contains a very wide range of occupations which makes it difficult to compare these categories in terms of skill requirements. Nevertheless, as summarised in Section 2.2, SBTC is found to have a positive effect on the skill demand in both developed and developing countries. Accordingly, it is reasonable to expect an increase in the skill demand in Turkey as well. However, as noted above, the effect of an increase in skill demand seems to be overwhelmed by the substantial increase in the supply of educated labour.

2.8 Policy Implications

Following the increase in the number of universities and reforms in the educational system, the supply of educated labour rapidly increased in Turkey. However, steep decreases in the returns to education and the increasing share of university graduates in unemployment suggest that the labour market could not accommodate this increasing supply. Accordingly, wage inequality decreased as a result of relatively low demand for the educated workers. As noted in Section 2.3.3, the share of university graduates among the male unemployed has become higher than that for high school graduates since 2008. Moreover, according to the TurkStat employment statistics (TurkStat web

database), the unemployment rate for university graduates which was around 6-8 per cent in 1990's became 11 per cent in 2010. This indicates that a large number of university graduates are facing the problem of unemployment. This leads to inefficiency since skilled labour, which is a valuable resource, remains idle. To overcome the problem, policy makers could consider encouraging investments that create new employment opportunities and also R&D investments even though it should be acknowledged that such investments may lead to labour saving technologies.

The results of the study indicate that the decreasing skill premium led to a substantial change in the skill pricing function of unmeasured characteristics. This is in accordance with the skill price theory of Lemieux (2002) who argued that the negative changes in the coefficients component exert a negative impact on the residual component along the wage distribution, providing a measure for unmeasured skills pricing. A substantial change in the skill pricing function can be interpreted as sign of high quality differences between universities. This is not a surprising result since the new universities have encountered financial difficulties and a lack of facilities such as laboratories and buildings. For instance, Cukadar et al. (2011) argue that the new universities face problems regarding libraries in terms of physical space, as well as a lack of personal and academic resources. They also show that the budget that old universities possess is greater than the new universities. Unfortunately, this hypothesis is not testable since there is no information in the dataset about the universities that the individuals graduated from. However, there is evidence that the quality of education is decreasing in general. For instance, according to the Student Selection and Placement Centre (OSYM) higher education statistics, the number of students per teaching staff increased from 13.8 to 18.8 from 1984 to 2011. In fact, the establishment of new universities under the "a university in each city'' project has been criticized by many authors (e.g.

Dortlemez, 1995; Kaynar and Parlak, 2005). It is argued that the scientific requirements and maintenance of minimum standards have been ignored as politicians used the stimulating effects of universities on local economies in their election campaigns. Decreasing quality of university education has two crucial economic outcomes; it reduces the productivity of the educated labour below its potential and causes an inefficient allocation of resources - resources which are quite limited in many developing countries. Therefore, in addition to issues regarding accessibility, policymakers should give importance also to the quality of education when decisions are made on education policies.

TABLES OF CHAPTER 2

Table 2.1: Counterfactual distributions

Variable	Weight	Resulting Distribution
$\boldsymbol{y_{i2002}}$	w_{i2002}	Distribution at 2002
$\boldsymbol{y^{a}_{i2002}}$	w_{i2002}	2002 distribution with $\boldsymbol{b}$ of 2010
$\boldsymbol{y_{i2002}}$	${w^{a}}_{i2002}$	2002 distribution with covariates of 2010
$\boldsymbol{y^{a}_{i2002}}$	${w^{a}}_{i2002}$	2002 distribution with covariates and $\boldsymbol{b}$ of 2010

Note: The base year is 2002. w_{i2002} is the original sample weights from the year 2002. ${w^{a}}_{i2002} = w_i \Psi_i$ is the counterfactual weights and Ψ_i is the re-weighting factor.

Table 2.2A: Sample Averages for Women

	2002			2010		
	E. inactive	U. family worker	Wage worker	E. inactive	U. family worker	Wage worker
Mean log real wages			5.38			5.65
			(0.91)			(0.78)
Education Dummies						
No formal education	0.22	0.33	0.04	0.25	0.32	0.06
Primary school	0.58	0.63	0.28	0.47	0.55	0.22
Secondary school	0.07	0.02	0.07	0.10	0.07	0.08
High school	0.08	0.02	0.17	0.08	0.03	0.13
Vocational high school	0.04	0.01	0.11	0.05	0.02	0.11
University	0.02	0.00	0.32	0.04	0.01	0.40
Tenure	-	19.55	6.42	-	15.27	5.07
	-	(0.15)	(0.07)	-	(0.11)	(0.05)
Presence of children						
Age ≤ 4 years	0.33	0.31	0.18	0.33	0.27	0.18
5 ≤ Age < 11 years	0.41	0.42	0.29	0.39	0.38	0.27
11 ≤ Age < 15 years	0.22	0.28	0.17	0.21	0.25	0.16
Other members of the household						
Presence of wage workers	0.50	0.12	0.64	0.57	0.21	0.64
Presence of self-employed	0.21	0.93	0.10	0.16	0.87	0.10
Presence of unpaid family workers	0.05	0.55	0.02	0.04	0.42	0.02
Marital status dummy						
Never married	0.10	0.13	0.38	0.10	0.09	0.35
Married	0.83	0.85	0.56	0.83	0.89	0.57
Divorced	0.01	0.002	0.04	0.02	0.01	0.06
Widowed	0.05	0.01	0.02	0.05	0.01	0.02
Age						
20-24	0.16	0.16	0.24	0.13	0.10	0.18
25-34	0.32	0.26	0.43	0.30	0.23	0.42
35-44	0.24	0.22	0.24	0.24	0.26	0.28
45 and above	0.27	0.35	0.08	0.32	0.41	0.11
Member of social security system	-	0.01	0.73	-	0.06	0.77
Head of household dummy	0.07	0.0003	0.10	0.08	0.00	0.10
Grandmother	0.05	0.13	0.05	0.06	0.13	0.05
Rural	0.30	0.92	0.18	0.23	0.86	0.13
Observations	60,811	7,787	9,442	97,594	15,220	20,080

Note: standard errors for continuous variables in parenthesis. Sample weights are used

Table 2.2B: Sample Averages for Men

	2002			2010		
	E. inactive	Self-employed	Wage worker	E. inactive	Self-employed	Wage worker
Mean log real wages			5.42			5.61
			(0. 798)			(0. 672)
Education Dummies						
No formal education	0.06	0.10	0.02	0.09	0.09	0.03
Primary school	0.52	0.70	0.43	0.39	0.62	0.35
Secondary school	0.12	0.09	0.14	0.18	0.13	0.16
High school	0.14	0.06	0.15	0.13	0.07	0.12
Vocational high school	0.09	0.03	0.12	0.09	0.06	0.13
University	0.08	0.02	0.15	0.11	0.04	0.20
Tenure	-	17.81	8.00	-	15.07	6.20
	-	(0.12)	(0.04)	-	(0.08)	(0.03)
Presence of children						
Age ≤ 4 years	0.27	0.32	0.35	0.25	0.28	0.34
5 ≤ Age < 11 years	0.35	0.43	0.42	0.33	0.39	0.38
11 ≤ Age < 15 years	0.22	0.27	0.21	0.21	0.25	0.19
Other members of the household						
Presence of wage workers	0.30	0.12	0.29	0.34	0.18	0.34
Presence of self-employed	0.18	0.07	0.08	0.13	0.08	0.08
Presence of unpaid family workers	0.06	0.48	0.03	0.04	0.44	0.03
Marital status						
Never married	0.40	0.05	0.18	0.39	0.06	0.22
Married	0.58	0.93	0.81	0.58	0.92	0.77
Divorced	0.01	0.01	0.01	0.02	0.01	0.01
Widowed	0.01	0.01	0.002	0.003	0.01	0.001
Age						
20-24	0.30	0.03	0.12	0.22	0.02	0.11
25-34	0.33	0.24	0.43	0.33	0.19	0.40
35-44	0.20	0.29	0.29	0.21	0.29	0.30
45 and above	0.17	0.43	0.15	0.23	0.49	0.18
Member of social security system	-	0.42	0.73	-	0.39	0.78
Head of household dummy	0.50	0.89	0.76	0.51	0.85	0.71
Rural	0.36	0.66	0.24	0.27	0.56	0.17
Observations	11,426	14,007	32,771	17,704	23,869	63,077

Note: standard errors for continuous variables in parenthesis. Sample weights are used.

Table 2.3A: Relative Risk Ratios (Women)

VARIABLES	2002		2010	
	Unpaid family worker	Wage worker	Unpaid family worker	Wage worker
Primary School	1.10**	1.74***	1.30***	1.56***
	(0.05)	(0.11)	(0.04)	(0.06)
Secondary school	0.68***	3.01***	1.20***	2.42***
	(0.09)	(0.23)	(0.07)	(0.11)
High school	0.66***	5.54***	0.92	4.21***
	(0.08)	(0.38)	(0.08)	(0.18)
Vocational high school	0.69*	8.13***	0.92	5.41***
	(0.13)	(0.60)	(0.09)	(0.24)
University	1.13	33.58***	0.85	21.26***
	(0.26)	(2.38)	(0.10)	(0.90)
Presence of children in the household;				
Age≤4	0.69***	0.52***	0.65***	0.46***
	(0.03)	(0.02)	(0.02)	(0.01)
5≤Age<11	0.97	0.81***	0.98	0.76***
	(0.04)	(0.03)	(0.03)	(0.02)
11≤Age<15	0.88***	0.99	0.92**	1.02
	(0.04)	(0.04)	(0.03)	(0.03)
Other members of the household;				
Presence of wage workers	0.72***	2.08***	0.82***	1.63***
	(0.04)	(0.07)	(0.03)	(0.04)
Presence of self-employed	14.23***	0.85***	12.21***	0.93**
	(0.80)	(0.05)	(0.42)	(0.03)
Presence of unpaid family workers	4.91***	0.74***	4.88***	0.64***
	(0.23)	(0.08)	(0.18)	(0.05)
Marital status;				
Married	1.31***	0.36***	1.70***	0.43***
	(0.10)	(0.02)	(0.10)	(0.01)
Divorced	0.37***	0.99	0.55***	1.39***
	(0.13)	(0.09)	(0.09)	(0.08)
Widowed	0.47***	0.34***	0.85	0.40***
	(0.08)	(0.04)	(0.11)	(0.03)
Age				
25-35	1.57***	1.53***	1.82***	1.55***
	(0.11)	(0.07)	(0.11)	(0.05)
35-45	1.81***	1.44***	2.74***	1.72***
	(0.14)	(0.07)	(0.17)	(0.07)
45-65	1.80***	0.60***	2.51***	0.62***
	(0.14)	(0.04)	(0.16)	(0.03)
Head of household dummy	0.02***	2.02***	0.21***	1.56***
	(0.01)	(0.13)	(0.03)	(0.07)
Grandmother	1.67***	1.20***	1.26***	1.07
	(0.11)	(0.08)	(0.06)	(0.05)
Rural	9.70***	1.00	8.16***	0.88***
	(0.41)	(0.04)	(0.24)	(0.03)
Constant	0.00***	0.07***	0.00***	0.11***
	(0.00)	(0.01)	(0.00)	(0.01)
Observations	7,787	9,442	15,220	20,080

Robust standard deviations in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Base category: economically inactive

Table 2.3B: Relative Risk Ratios (Men)

VARIABLES	2002		2010	
	Self-employed	Wage worker	Self-employed	Wage worker
Primary School	1.35***	2.10***	1.19***	1.58***
	(0.09)	(0.14)	(0.06)	(0.07)
Secondary school	1.39***	3.23***	1.29***	2.22***
	(0.11)	(0.23)	(0.08)	(0.10)
High school	1.19**	3.67***	1.09	2.29***
	(0.10)	(0.26)	(0.07)	(0.11)
Vocational high school	0.90	4.27***	1.15**	3.31***
	(0.08)	(0.32)	(0.08)	(0.17)
University	0.63***	4.96***	0.73***	3.96***
	(0.06)	(0.36)	(0.05)	(0.19)
Presence of children in the household;				
Age≤4	0.99	0.96	1.00	1.02
	(0.04)	(0.03)	(0.03)	(0.03)
5≤Age<11	0.90***	0.93**	0.92***	1.00
	(0.03)	(0.03)	(0.03)	(0.02)
11≤Age<15	0.90**	0.95	0.92***	0.96
	(0.04)	(0.03)	(0.03)	(0.02)
Other members of the household;				
Presence of wage workers	0.68***	1.43***	0.73***	1.30***
	(0.03)	(0.04)	(0.02)	(0.03)
Presence of self-employed	0.17***	0.83***	0.35***	0.99
	(0.02)	(0.04)	(0.02)	(0.03)
Presence of unpaid family workers	23.91***	1.06	25.43***	1.35***
	(1.74)	(0.08)	(1.31)	(0.07)
Marital status;				
Married	2.19***	2.06***	2.29***	2.02***
	(0.19)	(0.10)	(0.14)	(0.08)
Divorced	1.32	1.13	1.28**	1.01
	(0.24)	(0.15)	(0.14)	(0.08)
Widowed	1.62**	0.96	2.42***	1.05
	(0.30)	(0.18)	(0.41)	(0.18)
Age				
25-35	2.33***	1.62***	2.72***	1.46***
	(0.20)	(0.07)	(0.20)	(0.05)
35-45	2.85***	1.35***	3.69***	1.20***
	(0.27)	(0.07)	(0.29)	(0.05)
45-65	3.03***	0.75***	4.00***	0.62***
	(0.29)	(0.04)	(0.31)	(0.03)
Head of household dummy	2.71***	2.53***	2.39***	2.13***
	(0.19)	(0.11)	(0.11)	(0.07)
Rural	1.66***	0.67***	1.63***	0.66***
	(0.06)	(0.02)	(0.05)	(0.02)
Constant	0.07***	0.29***	0.06***	0.61***
	(0.01)	(0.02)	(0.00)	(0.03)
Observations	14,007	32,771	23,869	63,007

Robust standard errors in parentheses*** p<0.01, ** p<0.05, * p<0.1 Base category: economically inactive

Table 2.4: Wage Regression Estimations

	Women		Men	
Dependent Variable: Log real wages	2002	2010	2002	2010
Education;				
Primary School	0.09*	-0.06**	0.20***	0.01
	(0.05)	(0.03)	(0.02)	(0.01)
Secondary school	0.21***	0.03	0.37***	0.09***
	(0.06)	(0.03)	(0.03)	(0.01)
High school	0.48***	0.14***	0.54***	0.21***
	(0.06)	(0.03)	(0.03)	(0.01)
Vocational high school	0.46***	0.12***	0.59***	0.20***
	(0.06)	(0.03)	(0.03)	(0.01)
University	0.96***	0.55***	1.09***	0.78***
	(0.07)	(0.04)	(0.03)	(0.02)
Tenure	0.03***	0.04***	0.04***	0.03***
	(0.00)	(0.00)	(0.00)	(0.00)
Tenure squared (10^{-3})	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Presence of children in the household;				
Age≤4	0.06***	0.10***	-0.01	-0.00
	(0.02)	(0.01)	(0.01)	(0.00)
5≤Age<11	-0.02	0.02*	-0.01*	-0.02***
	(0.02)	(0.01)	(0.01)	(0.00)
11≤Age<15	-0.05**	-0.04***	-0.02	-0.02***
	(0.02)	(0.01)	(0.01)	(0.01)
Marital status dummies;				
Married	0.16***	0.20***	0.08***	0.06***
	(0.02)	(0.01)	(0.02)	(0.01)
Divorced	0.13***	0.06***	-0.02	0.06***
	(0.04)	(0.02)	(0.05)	(0.02)
Widowed	0.15***	0.07**	-0.19	0.11*
	(0.06)	(0.03)	(0.14)	(0.06)
Age	0.15***	0.11***	0.12***	0.11***
25-35	(0.02)	(0.01)	(0.01)	(0.01)
	0.21***	0.17***	0.23***	0.21***
35-45	(0.03)	(0.02)	(0.02)	(0.01)
	0.32***	0.29***	0.29***	0.29***
45-65	(0.04)	(0.02)	(0.02)	(0.01)
	0.15***	0.11***	0.12***	0.11***
Member of social security system	0.63***	0.35***	0.49***	0.29***
	(0.02)	(0.01)	(0.01)	(0.01)
Λ (selection term)	0.21***	0.35***	0.18***	0.27***
	(0.04)	(0.03)	(0.03)	(0.02)
Constant	4.26***	5.02***	4.27***	4.93***
	(0.09)	(0.05)	(0.04)	(0.02)
Observations	9,191	18,644	32,226	59,696
R-squared	0.56	0.58	0.44	0.49

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 2.5A: Decomposition of Changes in the Wage distribution for Women

	p90/p10	p90/p50	p50/p10	Gini coefficient	Theil index
2002	1.507	1.216	1.239	0.096	0.015
2010	1.371	1.199	1.143	0.074	0.009
Total change	-0.136	-0.017	-0.096	-0.021	-0.006
%	(-9.9)	(-1.4)	(-8.4)	(-28.9)	(-70.9)
Effect of					
Coefficients	-0.077	-0.006	-0.059	-0.010	-0.003
	(-56.6)	(-35.3)	(-61.3)	(-45.4)	(-40.9)
Pricing function	-0.021	-0.015	-0.001	-0.005	-0.002
	(-15.4)	(-88.2)	(-1.5)	(-24.6)	(28.2)
Covariates	-0.018	-0.003	-0.012	-0.002	-0.001
	(-13.2)	(-17.6)	(-12.1)	(-7.3)	(-7.2)
Unexplained	-0.020	0.007	-0.024	-0.005	-0.002
	(-14.7)	(41.2)	(-25.0)	(-22.7)	(-23.7)

Note: percentage shares of each effect in total change are shown in parentheses

Table 2.5B: Decomposition of Changes in the Wage distribution for Men

	p90/p10	p90/p50	p50/p10	Gini coefficient	Theil index
2002	1.422	1.206	1.179	0.081	0.011
2010	1.331	1.188	1.120	0.064	0.007
Total change	-0.091	-0.018	-0.059	-0.017	-0.004
%	(-6.8)	(-1.5)	(-5.3)	(-25.9)	(-59.0)
Effect of					
Coefficients	-0.051	-0.016	-0.028	-0.008	-0.002
	(-56.0)	(-88.9)	(-47.6)	(-48.8)	(-44.2)
Pricing function	-0.034	-0.014	-0.015	-0.007	-0.002
	(-37.4)	(-77.8)	(-24.7)	(-41.2)	(-48.7)
Covariates	-0.017	0.001	-0.016	-0.002	-0.001
	(-18.7)	(5.6)	(-27.8)	(-14.6)	(-17.5)
Unexplained	0.011	0.011	-0.001	0.001	0.001
	(12.1)	(61.1)	(-2.3)	(4.6)	(10.4)

Note: percentage shares of each effect in total change are shown in parentheses

Table 2.6A: Decomposition of Changes in the Wage Residuals (Women)

	Variance	90-10	90-50	50-10
2002	0.363	1.350	0.681	0.669
2010	0.236	1.122	0.562	0.560
Total change	-0.127	-0.228	-0.119	-0.109
%	(-53.7)	(-20.3)	(-21.2)	(-19.4)
Effect of				
Covariates	0.0057	0.019	0.014	0.005
	(4.5)	(8.3)	(12.1)	(4.2)
Pricing function	-0.133	-0.247	-0.133	-0.113
	(-104.5)	(-108.3)	(-112.1)	(-104.2)

Note: percentage shares of each effect in total change are shown in parentheses

Table 2.6B: Decomposition of Changes in the Wage Residuals (Men)

	Variance	90-10	90-50	50-10
2002	0.331	1.340	0.686	0.654
2010	0.212	1.113	0.571	0.543
Total change	-0.119	-0.227	-0.115	-0.111
%	(-56.0)	(-20.4)	(-20.2)	(-20.5)
Effect of				
Covariates	-0.008	-0.022	0.003	-0.025
	(-6.7)	(-9.6)	(2.7)	(-22.4)
Pricing function	-0.111	-0.205	-0.119	-0.086
	(-93.3)	(-90.4)	(-102.07)	(-77.6)

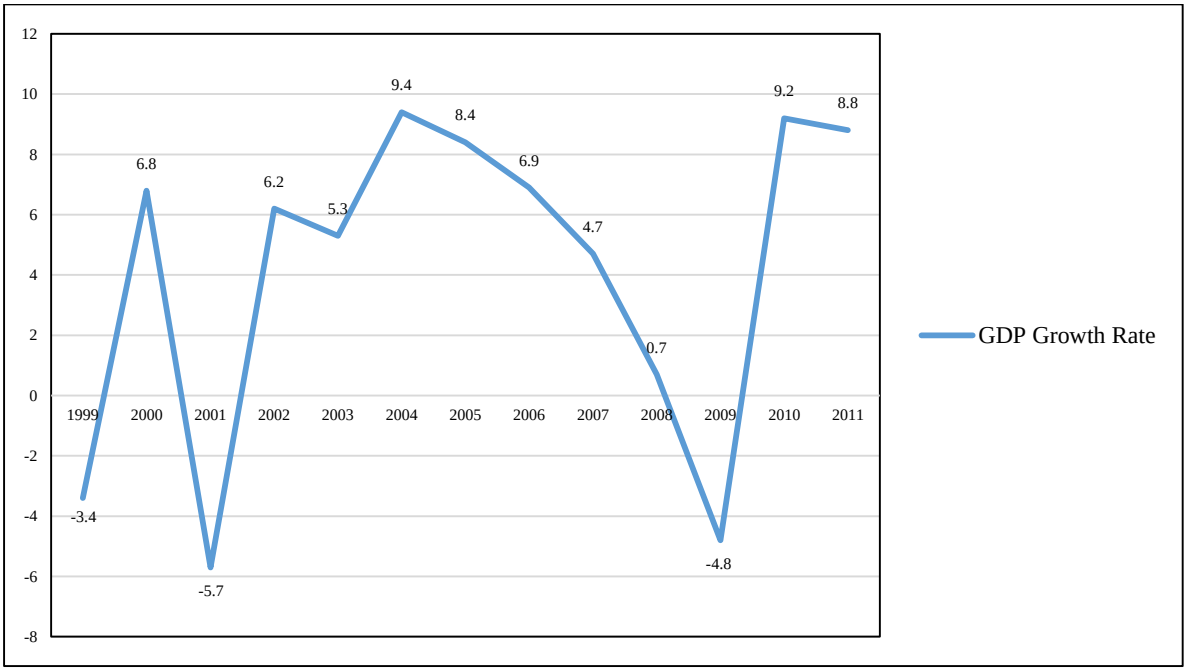
Note: percentage shares of each effect in total change are shown in parentheses

2.9 APPENDIX TO CHAPTER 2

Table 2.A1: Exclusion Criteria for the sample

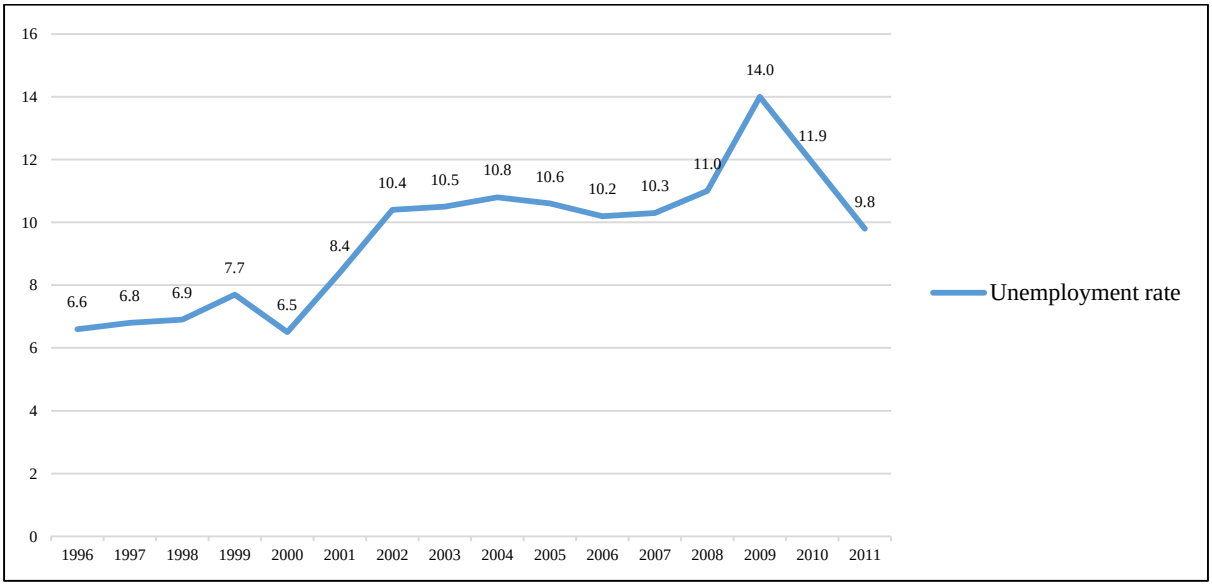
	2002		2010	
	Frequency	%	Frequency	%
Full Sample	300689		522171	
Dropped observations				
Age over 65 or less 20	135422	45	227112	43.5
Students	3252	1.1	6190	1.2
Disabled or ill	2838	0.9	10691	2
Retired	10987	3.7	18426	3.5
Has more than one job	1259	0.4	4747	0.9
Employers (women)	288	0.1	570	0.1
Self-employed (women)	2552	0.8	5835	1.1
Employers (men)	5082	1.6	7032	1.3
Unpaid family workers (men)	2765	0.9	4024	0.8
Resulting sample for selection equation	136244	45.3	237,544	45.5
Non-wage earners	94031	31.3	154387	29.6
Reported extreme values of experience (inconsistent with age)	4	0.0	0	0.0
Reported zero wage	792	0.3	4817	0.9
Resulting sample for wage equations	41,417	13.8	78,340	15.0

Figure 2.A1: GDP Growth Rate (1987=100)



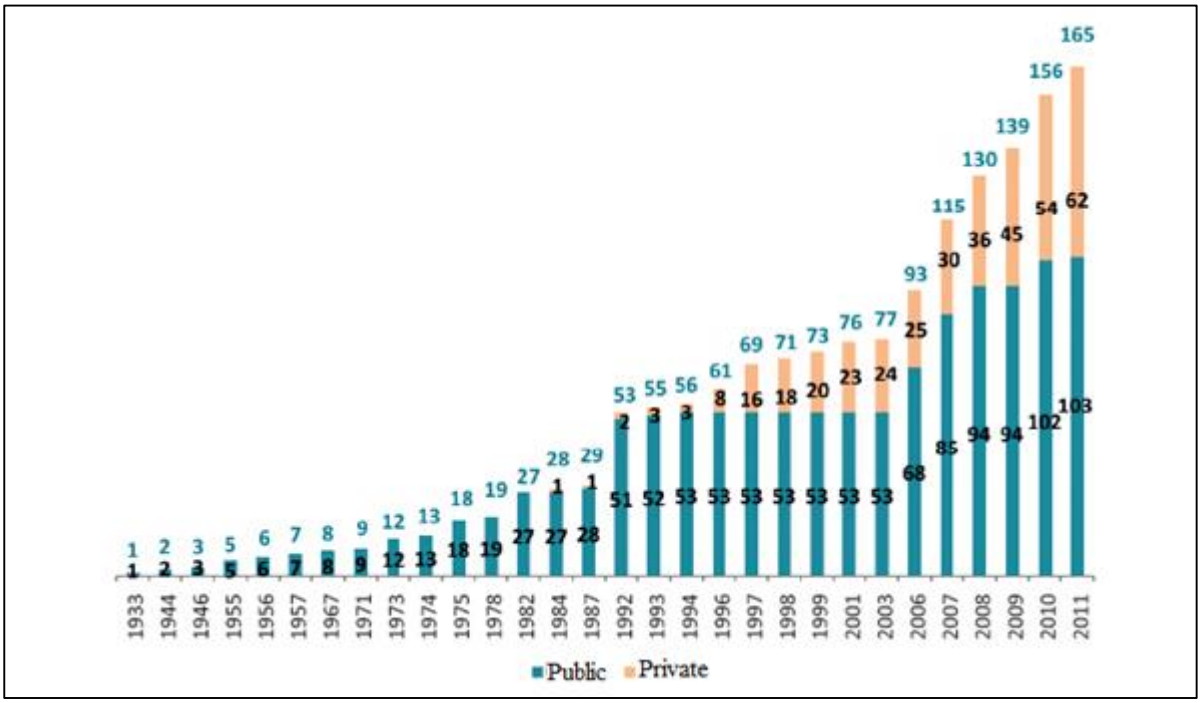
Source: TurkStat Web Database

Figure 2.A2: Unemployment rate



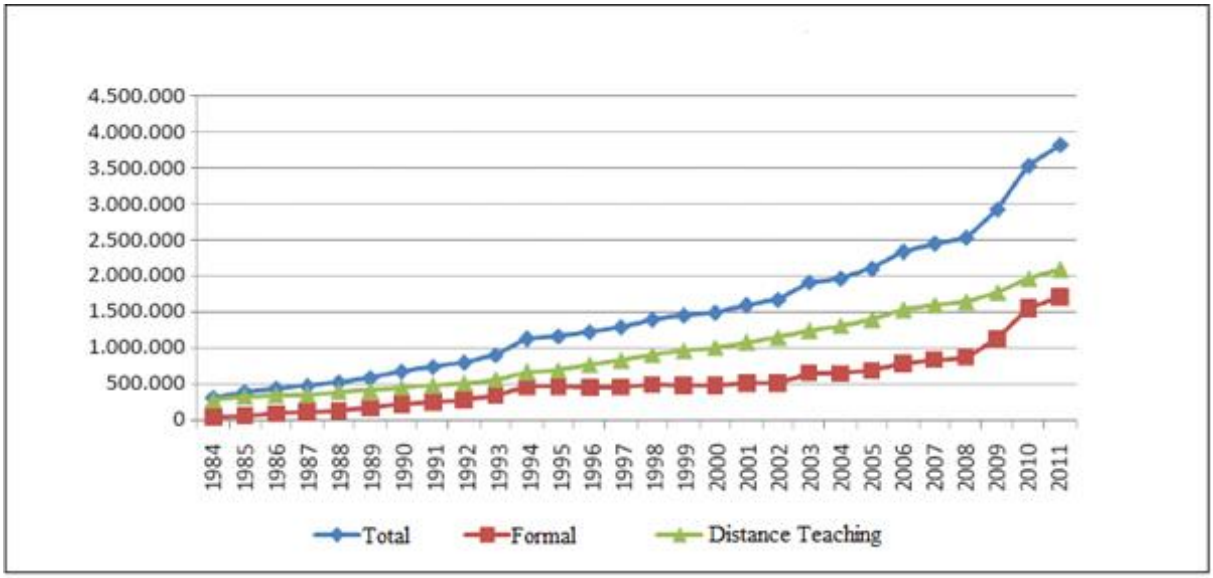
Source: TurkStat Web Database

Figure 2.A3: Number of universities by year



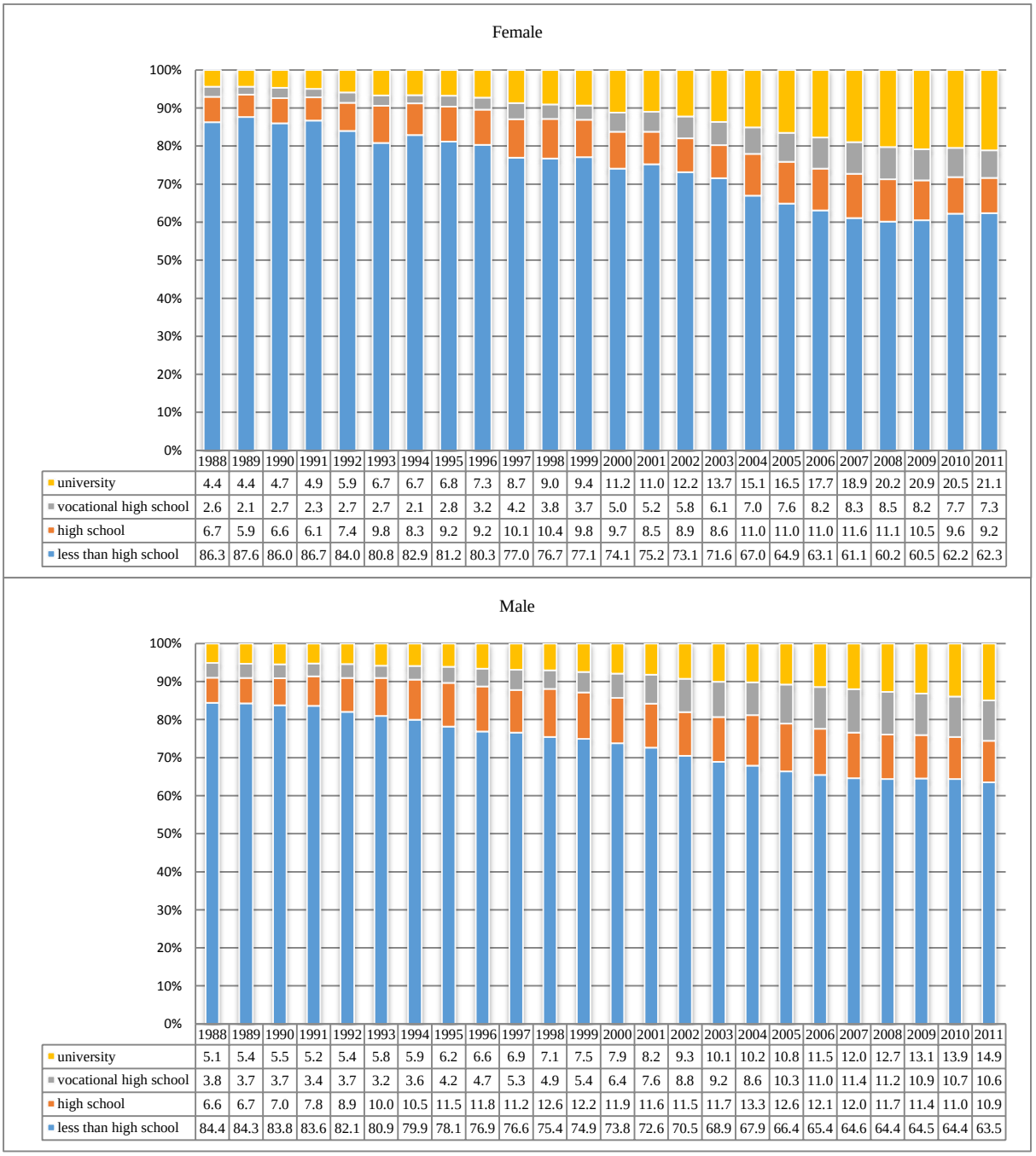
Source: TurkStat Web Database

Figure 2.A4: Number of University students by year



Source: TurkStat Web Database

Figure 2.A5: Labour force by education



	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
university	4.4	4.4	4.7	4.9	5.9	6.7	6.7	6.8	7.3	8.7	9.0	9.4	11.2	11.0	12.2	13.7	15.1	16.5	17.7	18.9	20.2	20.9	20.5	21.1
vocational high school	2.6	2.1	2.7	2.3	2.7	2.7	2.1	2.8	3.2	4.2	3.8	3.7	5.0	5.2	5.8	6.1	7.0	7.6	8.2	8.3	8.5	8.2	7.7	7.3
high school	6.7	5.9	6.6	6.1	7.4	9.8	8.3	9.2	9.2	10.1	10.4	9.8	9.7	8.5	8.9	8.6	11.0	11.0	11.0	11.6	11.1	10.5	9.6	9.2
less than high school	86.3	87.6	86.0	86.7	84.0	80.8	82.9	81.2	80.3	77.0	76.7	77.1	74.1	75.2	73.1	71.6	67.0	64.9	63.1	61.1	60.2	60.5	62.2	62.3

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
university	5.1	5.4	5.5	5.2	5.4	5.8	5.9	6.2	6.6	6.9	7.1	7.5	7.9	8.2	9.3	10.1	10.2	10.8	11.5	12.0	12.7	13.1	13.9	14.9
vocational high school	3.8	3.7	3.7	3.4	3.7	3.2	3.6	4.2	4.7	5.3	4.9	5.4	6.4	7.6	8.8	9.2	8.6	10.3	11.0	11.4	11.2	10.9	10.7	10.6
high school	6.6	6.7	7.0	7.8	8.9	10.0	10.5	11.5	11.8	11.2	12.6	12.2	11.9	11.6	11.5	11.7	13.3	12.6	12.1	12.0	11.7	11.4	11.0	10.9
less than high school	84.4	84.3	83.8	83.6	82.1	80.9	79.9	78.1	76.9	76.6	75.4	74.9	73.8	72.6	70.5	68.9	67.9	66.4	65.4	64.6	64.4	64.5	64.4	63.5

Source: TurkStat Web Database

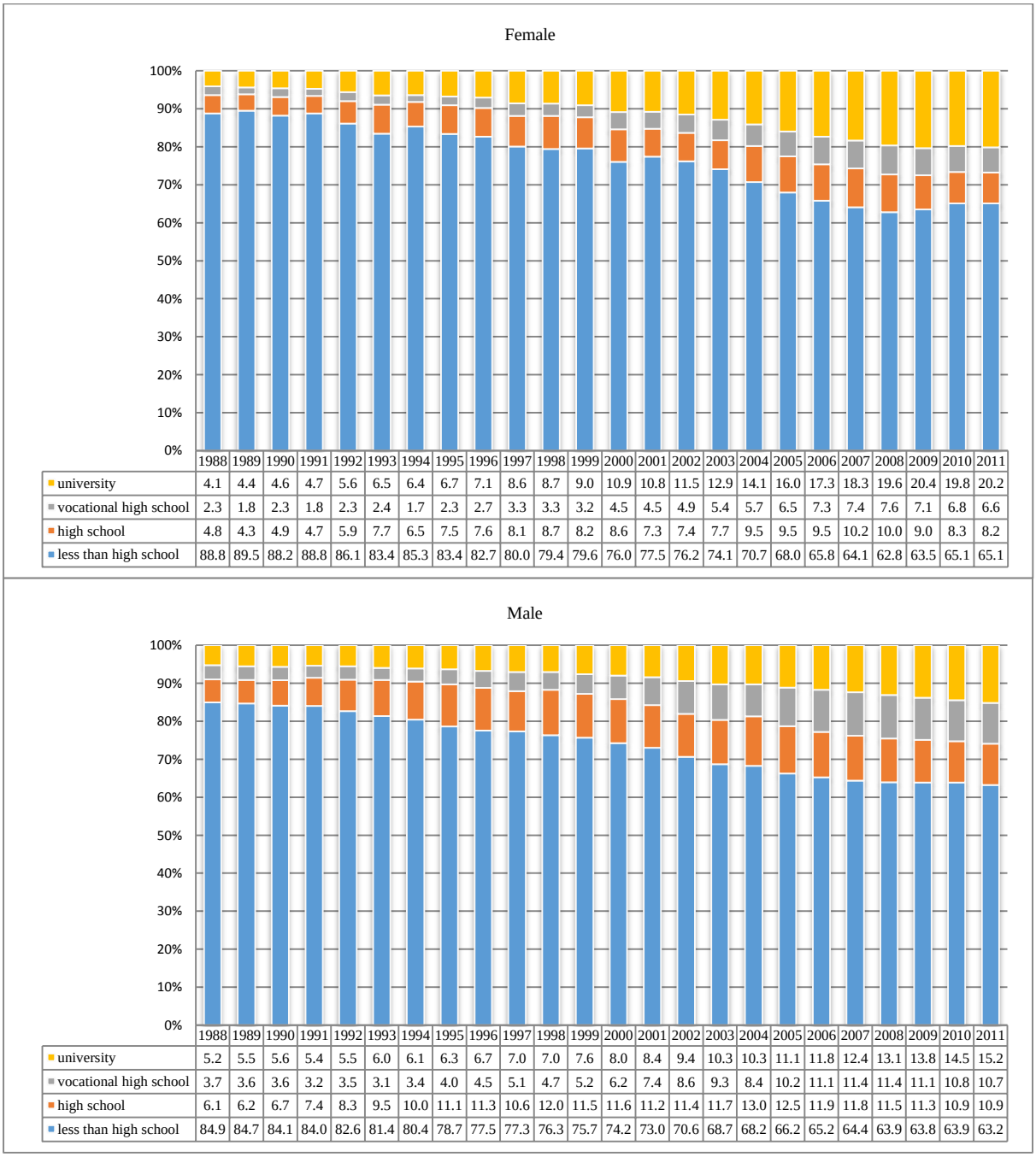
Figure 2.A6: Unemployment by education

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
university	4.0	3.6	3.9	3.8	4.2	4.5	4.1	4.5	5.6	5.7	7.4	6.6	7.2	6.3	7.9	8.4	9.3	8.2	8.9	8.9	9.5	9.1	9.8	12.4
vocational high school	5.2	5.1	4.7	5.1	5.3	3.9	5.3	6.7	7.6	8.3	7.4	7.7	9.1	9.7	10.2	8.7	10.4	10.8	10.4	11.0	9.6	10.0	9.8	9.8
high school	12.7	11.5	11.4	11.5	14.5	15.0	15.9	16.9	19.3	19.7	21.4	19.6	15.7	15.3	12.5	12.2	15.2	13.3	13.6	13.1	12.9	12.4	12.2	11.3
less than high school	78.1	79.8	80.0	79.6	76.0	76.7	74.7	72.0	67.5	66.3	63.7	66.2	68.0	68.7	69.3	70.7	65.2	67.7	67.1	66.9	67.9	68.5	68.2	66.4

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
university	7.4	4.6	5.9	8.0	9.1	8.8	9.4	7.4	10.6	10.9	14.4	13.3	16.3	14.7	19.0	20.7	23.3	20.9	20.7	23.9	24.9	23.8	25.1	28.5
vocational high school	5.3	4.9	6.5	7.7	7.2	5.9	6.5	9.1	11.4	15.6	11.9	10.8	12.9	14.5	14.6	12.5	16.9	16.4	15.3	15.9	15.1	14.8	13.2	13.0
high school	22.2	20.7	24.9	24.8	25.4	30.3	29.1	30.0	35.7	33.7	34.2	29.3	25.6	22.8	22.6	17.3	22.8	22.7	22.6	23.5	19.8	19.2	18.6	17.5
less than high school	65.1	69.6	62.7	59.7	58.3	55.2	55.0	53.2	42.6	39.8	39.6	46.7	44.7	48.1	43.6	49.5	37.0	40.0	41.2	36.8	40.2	42.3	43.0	41.0

Source: TurkStat Web Database

Figure 2.A7: Employment by education



	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
university	4.1	4.4	4.6	4.7	5.6	6.5	6.4	6.7	7.1	8.6	8.7	9.0	10.9	10.8	11.5	12.9	14.1	16.0	17.3	18.3	19.6	20.4	19.8	20.2
vocational high school	2.3	1.8	2.3	1.8	2.3	2.4	1.7	2.3	2.7	3.3	3.3	3.2	4.5	4.5	4.9	5.4	5.7	6.5	7.3	7.4	7.6	7.1	6.8	6.6
high school	4.8	4.3	4.9	4.7	5.9	7.7	6.5	7.5	7.6	8.1	8.7	8.2	8.6	7.3	7.4	7.7	9.5	9.5	9.5	10.2	10.0	9.0	8.3	8.2
less than high school	88.8	89.5	88.2	88.8	86.1	83.4	85.3	83.4	82.7	80.0	79.4	79.6	76.0	77.5	76.2	74.1	70.7	68.0	65.8	64.1	62.8	63.5	65.1	65.1

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
university	5.2	5.5	5.6	5.4	5.5	6.0	6.1	6.3	6.7	7.0	7.0	7.6	8.0	8.4	9.4	10.3	10.3	11.1	11.8	12.4	13.1	13.8	14.5	15.2
vocational high school	3.7	3.6	3.6	3.2	3.5	3.1	3.4	4.0	4.5	5.1	4.7	5.2	6.2	7.4	8.6	9.3	8.4	10.2	11.1	11.4	11.4	11.1	10.8	10.7
high school	6.1	6.2	6.7	7.4	8.3	9.5	10.0	11.1	11.3	10.6	12.0	11.5	11.6	11.2	11.4	11.7	13.0	12.5	11.9	11.8	11.5	11.3	10.9	10.9
less than high school	84.9	84.7	84.1	84.0	82.6	81.4	80.4	78.7	77.5	77.3	76.3	75.7	74.2	73.0	70.6	68.7	68.2	66.2	65.2	64.4	63.9	63.8	63.9	63.2

Source: TurkStat Web Database

CHAPTER 3: FEMALE AUTONOMY, SOCIAL NORMS AND DOMESTIC VIOLENCE AGAINST WOMEN IN TURKEY

3.1 Introduction

As a global epidemic, domestic violence undermines the equality, dignity and the fundamental freedoms of women. The ever-present threat of violence makes women suffer physically, psychologically, sexually and economically. It denies their rights of protection, making their own choices and voicing opinions. The most pervasive form of violence against women is perpetrated by the intimate partner and takes place at home. This chapter investigates the relationship between intimate partner violence and women's autonomy in the household to understand how intimate partner violence impairs women's autonomy and how incidence of violence changes with the level of women's autonomy.

Over the past few decades, recognition of domestic violence and the importance of its consequences have globally increased. However, the domestic violence issue was not the centre of academic attention until the mid-1970's. An interesting example is that until 1969, "The Journal of Marriage and the Family" did not have any papers with the word violence in its title (O'Brien 1971, cited Gelles, 1980 p.873). The studies on violence in the family were primarily focused on child abuse and the rare studies conducted on intimate partner violence portrayed wife-battering husbands and victims as people who suffer from personality disorders (Schultz, 1960; Snell et al., 1964).[24] Along with the growth of women's movements in the US in the mid-1970's, as Straus (1974) puts it, the domestic violence issue shifted from "selective inattention" to becoming a "high priority social issue". The first examples of academic works on domestic violence were written in the second half of the 1970's. One of the first major books on the topic of domestic violence against women was written by Martin (1976)

[24]See Gelles (1980) for a review of the domestic violence literature in the 1970's.

who was also the organizer and chair of the National Organization for Women task force dealing with wife battering in the US. In the same period, with the spread of feminist ideas, early classics of feminist theory were written by Walker (1979) and Dobash and Dobash (1979).

In 1980, for the first time, domestic violence against women was explicitly addressed in an official meeting of the United Nations (UN) in Copenhagen. In the final report of the 1985 UN conference held in Nairobi, the governments were strongly advised to undertake an increase in the public awareness of violence against women as a societal problem (UN, 1985). The following decades witnessed increasing awareness of domestic violence with the help of emerging civil movements and the efforts of international organizations such as the World Health Organization (WHO) and United Nations Children's Fund (UNICEF). The feminist movement especially played a crucial role in increasing the recognition of the problem. A recent study by Htun and Weldon (2012) showed that it is the feminist mobilization in civil society that accounts for the variation in policy developments in different countries rather than the intra-legislative political phenomena such as leftist parties or women in government or economic factors like national wealth.

Beside domestic violence, the other important concept at the centre of this study is female autonomy. Dyson and Moore (1983) define autonomy as the capacity to manipulate one's personal environment and the technical, social and psychological abilities in order to obtain information and to use it as the basis for making decisions about one's private concerns and those of one's intimates. Empirically, it is found that higher levels of female autonomy within the household lead to lower levels of fertility, child mortality and better levels of health status for the household members (Jejeebhoy, 1999; Kishor and Johnson, 2004; Presser and Sen, 2000).

In this chapter, using data from Turkey, the relationship between domestic violence and female autonomy in the household is analysed. The key findings of the study suggest that, consistent with the arguments of feminist theory on domestic violence, violence has a significant and decreasing effect on female autonomy and the incidence of violence decreases with the level of female autonomy. It is also found that domestic violence is an increasing function of the strength of social norms of a husband on traditional gender roles.

This study makes several contributions to the existing literature. It is the first study that estimates the two way relationship between domestic violence and female autonomy at the same time. Second, due to the unique property of the dataset, to the author's knowledge it will be the first study that analyses the effect of social norms on domestic violence at the individual level. Finally, to the author's knowledge it is the first extensive analyses of female autonomy in Turkey.

The chapter is structured as follows: Section 2 provides the background information and discussed the implications of female autonomy together with the measurement methods. Section 3 presents the definition, magnitude and consequences of domestic violence. Section 4 summarizes the existing theoretical and empirical literature on female autonomy and domestic violence, followed by Section 5 which describes the data and provides descriptive statistics. Section 6 outlines the estimation method. Sections 7 and 8 respectively present the estimation results and conclude the chapter.

3.2 Female autonomy

Autonomy is defined as the capacity to manipulate one's personal environment and ability to make decisions about one's private concerns and those of one's intimates. Female autonomy, in particular, is an important concept as empirically it has been

found that higher female autonomy provides better outcomes for the household members, especially for the children. For instance, using the data from Bangladesh, Indonesia, Ethiopia, and South Africa, Quisumbing and Maluccio (2000) show that while husbands tend to spend money on tobacco and luxury consumption goods, women, when they have the control of assets, spend more on children's education, health and clothing. Similarly, Arulampalam et al. (2015) show that maternal autonomy has a positive impact on the long term nutritional status of children in rural India. In accordance with these findings, Mason (1987) and Eswaran (2002) argue that the empowerment of women may lead to lower fertility rates as women face higher costs of children than men. In fact, Gudbrandsen (2013) shows that, in Nepal, the families where wives have high levels of autonomy have fewer children relative to other families.

A fundamental assumption in economics is that individuals make their choices by comparing the costs and benefits of all the alternative options available to them. Arulampalam et al. (2012) argue that autonomy can help women to make the best use of economic opportunities. However, lack of female autonomy prevents women from exercising their preferences regardless of the evaluation that they make of alternative options. Studies that analyse the wage determination model of women almost always take into account the self-selection of women into the labour force (Chzhen and Mumford, 2010; Buchinsky, 1998). In these models, it is assumed that women prefer to participate in the labour force if the utility of participation is higher than the utility of non-participation (i.e. if the wage offer is higher than the women's reservation wage); this assumption however may be incorrect, at least in the case of Turkey. According to the Demographic and Health Survey (DHS) for the year 2005, apart from factors such as household responsibilities and education, around 24 per cent of women in Turkey do not work or look for a job because their husbands or elders prevent them from doing so

(see Table 3.A1 in the appendix). This means that a considerable number of women in Turkey cannot work even if they would choose to do so. In that sense, lack of autonomy of women within the household may lower their labour force participation. This observation is quite crucial considering the fact that female labour force participation in Turkey has decreased from 35 per cent to 25 per cent from 1988 to 2006 (Dayioglu and Kirdar, 2010).

Although there is a large body of studies looking at the effects of female autonomy on various outcomes, the ways in which female autonomy is measured in these studies are very inconsistent. In the early examples of the literature, education, employment, and asset ownership have been used as the proxy measures for female autonomy (Bradley and Khor, 1993; Mason 1986). However, Malhotra and Mather (1997) argue that these measures can be considered as indicators of access to resources and they do not indicate the control over resources. They further argue that even if these variables can possibly serve as good proxies, the pathway between female autonomy and these proxies should be established rather than simply assumed.

Another way to measure female autonomy is to use women's ability to make decisions as an indicator. In most of the studies that used this method, indicators are created by using women's responses to questions about their having a say on specific decisions. After assigning a value to the answer for each question (indicator), the index scores for female autonomy are created by taking the average of the values of the indicators (see, for example, Eswaran and Malhotra, 2011). Yet, this method is also argued to be potentially problematic. Arulampalam et al. (2012) argue that not all questions that are asked to women are equally persuasive as indicators of women's autonomy because not all have the same consequential significance on women's lives. Kabeer (1999) also notes that in most societies, the decision making responsibilities are often segregated in

accordance with gender roles in the household. Therefore, the measure of female autonomy can be problematic if one uses the average value of indicator variables (which are categorical variables in most cases) as this method assigns equal weight to each indicator and assumes that they contain an equal amount of information on female autonomy.

In this chapter, Item Response Theory (IRT) is used to measure female autonomy. Female autonomy is treated as a latent variable as its true value is not observable. IRT allows female autonomy (as a latent variable) to affect different indicators in different ways as reflected by the factor loadings. The factor loadings are determined by the data instead of assigning equal weights to each of the indicators. In this method, it is also possible to model measurement error in indicator variables. Further details on Item Response Theory are discussed in Section 3.6.

3.3 Domestic Violence

3.3.1 Definition

As domestic violence has become a globally recognised problem, the definition of the concept has started to be the subject of debate. While there is no widely accepted definition of domestic violence, the United Nations Declaration on the Elimination of Violence against Women (UN, 1993) defines violence against women as "any act of gender-based violence that results in, or is likely to result in, physical, sexual or psychological harm or suffering to women, including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or in private life". Although there are arguments supporting the inclusion of poverty and unequal access to health and education, some authors like Heise et al. (1994) emphasize the importance of a limited definition in maintaining the descriptive power of the term. The Declaration provides detailed definitions of physical, sexual, psychological and economic abuse and

recognizes violence as a crucial social mechanism that forces women into a subordinate position relative to men. The Declaration also points out that although violence against women can be perpetrated by the general community or by the state, the most common form of violence against women is committed by the husband or an intimate male partner (WHO, 1997).

3.3.2 Magnitude and Consequences of Domestic Violence

Due to the intimate nature of domestic violence, many of the incidents remain unreported. For example, Greenfeld et al. (1998) estimated that only about half of domestic violence incidents in the US are reported to the police. Therefore the police force and the justice department records are likely to underestimate the magnitude of the problem. Accordingly, most of the academic analyses are based on regional or countrywide surveys (see Kapoor, 2000). Even though these surveys can potentially provide better information, measurement issues still exist due to the differences in sampling techniques and the way domestic violence is defined in these surveys. For instance, using the 2006 Demographic and Health Survey (DHS) for Cambodia, Eng et al. (2010) show that 22 per cent of ever-married women reported that they experienced physical violence at least once since the age of 15. They capture the incidence of violence by using the detailed information about the type of violence that is perpetrated by the husband such as pushing, hitting, and kicking. However, in many datasets, there is no detailed information about the type of violence. Hence, incidence of violence is mostly captured by one single question. For example, Koenig et al. (2003) used 1994 DHS data for Bangladesh and showed that 42 per cent of married women experienced domestic violence by their husbands. In the survey that they used, domestic violence was captured by the question of "Do you encounter the problem of physical beating from your husband or from his family?" The authors acknowledged that the use of a

single question to capture violence can be problematic as it lacks the specificity concerning which particular actions constitute physical violence, leaving the interpretation largely up to the individual respondent.[25]

Apart from the differences between the specificity of the questions, another measurement issue arises from the differences between the time periods that the questions cover. In some cases, such as the studies mentioned in the preceding paragraph, the questions ask about whether they have ever experienced violence (or experience since marriage), and in other cases questions ask about the experience of violence over one year immediately preceding the survey date (see, for example Eswaran and Malhotra, 2011). As a result of the differences between these country specific surveys, it is not possible to compare the statistics on domestic violence by looking at the studies based on these surveys.[26]

Nevertheless, a recent WHO report on violence against women documented the severity of the problem on a global scale (WHO, 2013). According to this report, 35 per cent of women worldwide have experienced either physical and/or sexual forms of intimate partner violence or non-partner sexual violence. Almost 30 per cent of all women who have been in a relationship have suffered physical and/or sexual violence by their intimate partner and this ratio goes up to almost 40 per cent in the regions such as South Asia and Latin America.

Obviously, the most crucial consequence of violence against women is the violation of the most fundamental human rights. However, this is not the only cost. The Inter-American Development Bank (IDB) (Buvinic et al., 1999) carried out a study in six

[25] Analysis in this chapter will also have similar problems, as domestic violence is captured using a single question; this is discussed further in Section 3.5.
[26] Most of the studies on domestic violence use the DHS surveys. However, the DHS surveys also differ from country to country in terms of how domestic violence is captured. For instance, the DHS surveys on Turkey do not have any information on domestic violence at all.

Latin American countries – Brazil, Colombia, El Salvador, Mexico, Peru and Venezuela – taking a more holistic look at the socio-economic costs of domestic violence. For analytical purposes, the IDB study divided the costs of domestic and social violence into four categories using the following framework.

Direct costs: these include the value of goods and services used in treating or preventing violence such as psychological counselling and medical treatment of victims, housing and shelters for women and their children, and costs imposed on police services and the criminal justice system on capturing and punishing the abusers. For instance, Greaves (1995) analysed the case of Canada and found that the total direct cost of violence against women exceeds one billion Canadian dollars. According to research conducted by the National Center for Injury Prevention and Control (2003), the estimated cost of intimate partner violence in the US is around 4.1 billion dollar per year.

Non-monetary costs: these include the health hazards that do not necessarily require a demand for medical services, but nevertheless heavily affect the victims by way of increased morbidity and mortality through homicide and suicide, increased dependence on drugs and alcohol and other depressive disorders. Lozano (1999) estimated that rape and intimate partner violence against women were the third most important causes of the disability-adjusted life years lost in Mexico City after diabetes and prenatal conditions, but ahead of auto accidents, congenital anomalies, rheumatoid and osteo-arthritis, cardiovascular disease, stroke and pneumonia. Heise et al. (1994) estimated that more than nine million disability-adjusted years of life are lost each year worldwide as a result of rape and family violence, more than that from all types of cancer and more than twice that lost by women in motor vehicle accidents.

Social multiplier effects: the costs that include the inter-generational impact of violence on children, decreasing social capital, reduced quality of life and participation in democratic processes. These effects are not easy to quantify by their nature. However, there are some empirical studies documenting these effects. For instance, Straus et al. (1980) noted that the rate of partner violence was substantially higher for men who observed domestic violence in childhood as compared to the men who had not witnessed such violence.

Economic multiplier effects: these include the reduced participation of women in the labour force and increased absenteeism at work, and lower productivity and earnings for women compared to men. These effects in turn are argued to lead to lower levels of savings and investments at the macro level. For the US, Stanley (1992) reported that 30 per cent of abused women lost their jobs as a direct result of the abuse. Greaves (1995) reported that, in Canada, 34 per cent of battered women and 11 per cent of sexual assault victims were not able to work on the following day of the assault. He also estimated that the value of lost earnings was more than 7 million Canadian dollars per year. Staggs and Riger (2005) analysed the effect of intimate partner violence on women's employment using data from Illinois and concluded that violence has a negative effect on the stability of the women's employment.

3.4 Literature Review

3.4.1 Theories of Domestic Violence

Starting with the pioneering work of Becker (1965, 1973, 1981), economists have analysed family relations and resource allocation within the household, using a utility maximising framework. The early examples of this literature were based on cooperative models which assume a household to be a single unit. According to the cooperative models, the household acts like one economic agent which, given a budget constraint,

tries to maximize a single utility function that is based on a common set of preferences. However, as Rode (2011) notes, since the 1980's, the cooperative model of the household came under attack on both theoretical and empirical grounds. Firstly, it was argued that the cooperative models do not take into account the intra-household inequality, the possibility of individual household members having different and may be even conflicting preferences, and having different levels of access to resources. Secondly, empirical evidence did not support the key predictions of the cooperative household modelling (see Lundberg et al., 1997). Accordingly, recent studies by economists are based on non-cooperative models in which the individual preferences of the household members matter.

Tauchen et al. (1991) present a non-cooperative bargaining model for the determinants of domestic violence. In their model, they assume that a relationship must provide each individual with a level of utility that is at least as great as the utility that the individual would obtain were the relationship to dissolve. Physical violence is both a source of satisfaction (such as relief of frustration) and an instrument for controlling wife's behaviour. The authors assert that the husband may resort to violence towards his partner as well as possibly transfer income to her. In that sense, men "purchase" violence from women with income transfers until a threat point level of utility that must be maintained for woman. According to their model, if the victim's reservation utility constraint is binding, violence is solely a source of direct gratification; an increase in the husband's income increases violence and redistributes welfare towards him leaving no benefit for the wife; similarly, increases in the victim's income can be generally expected to decrease violence. On the other hand, if both individuals gain from the relationship, then violence can be an instrument to obtain desired behaviours as well as a source of direct gratification. Tauchen et al. (1991) empirically tested their model

using the data from California. Their results showed that for the low income couples in which both the male and female work, increases in the husband's income increase violence while increases in the wife's income have a negative, but generally insignificant effect on violence. For high income couples in which the husband provides a bigger share of the family income, an increase in the income of either member decreases the violence. Their results also show that in high income families in which the wife provides most of the income, increases in her income serve to increase violence.

Farmer and Thiefenthaler (1997) also model the determinants of violence within a non-cooperative game setting. In their model, each spouse, with independent preferences and threat points, maximizes his/her utility given the behaviour and the threat point of the other. The wife's threat point determines the level of violence she will tolerate for a given transfer from husband. With the increases in her income (and consumption), the marginal utility she receives from an additional unit of consumption declines. Accordingly, the man's ability to "purchase" violence from her decreases and the violence decreases. They also argued that anything that increases the wife's utility outside of the marriage, such as the availability of shelters or extended family resources, will increase her threat point and the probability that she will leave. Therefore, presence of these alternatives lowers the levels of violence if she stays.

While studies by Tauchen et al. (1991) and Farmer and Thiefenthaler (1997) suggest that women's income has a protective effect from domestic violence, there is also a body of empirical research whose results do not support this argument. For instance, Jejeebhoy (1998) analysed the determinants of domestic violence in Uttar Pradesh and Tamil Nadu in India with a logistic regression (where domestic violence is a binary dependent variable). His results show that wage earning women do not enjoy the protective effect of income, as the coefficients for wage employment were insignificant

in both settings. Schuler et al. (1998) show that, in rural Bangladesh, the incidence of violence was higher for the women who contributed to family income than for those who do not make any contribution. Vyas and Watts (2009) investigate the published data on domestic violence from 11 countries to analyse the effect of women's involvement in income generating activities on domestic violence. They find a negative association in five cases and a positive association in six cases.

A possible explanation for these inconsistent results is provided by Aizer (2010). She argues that most previous studies of the relationship between women's income and domestic violence fail to establish a causal relationship by failing to account for the potential reverse causality. Moreover, she asserts that even the studies that account for endogeneity focus largely on a woman's own actual wage. She suggests that, in a household bargaining model, the woman's relative wage matters and potential, not actual, wages determine the bargaining power and the level of violence. To overcome the endogeneity of individual wages, she looks at the demand for female and male labour and captures the wage gap by the relative wage growth in industries in which men and women are traditionally concentrated. Finally, using administrative data for California, she finds that decreases in the wage gap reduce violence against women.

Gibson-Davis et al. (2005) also highlight the inconsistent results of the studies that analyse the effect of women's employment on domestic violence. They argue that this inconsistency is due to the potential reverse causality between women's employment and domestic violence. Using the data from the Minnesota Family Investment Program, they analysed this relationship with a two stage least square method. In the Minnesota Family Investment Program, women were randomly assigned into two groups; and while one group was provided with only financial incentives, the other group received both financial incentives and was subject to employment mandates. They use random

assignment status as an instrument to predict employment, and other economic variables, which are then included in a second-stage model predicting domestic abuse. Their results show that maternal employment decreases subsequent reports of domestic abuse.

Panda and Agarwal (2005) criticized the studies which only focus on the effect of employment status of women. They asserted that to examine the effect of the economic status of women, one should go beyond the employment status and consider the effect of women's property status since having land or a house can act as a tangible exit-option for women while employment status is subject to the unpredictability of the labour market. Using a household survey from Kerala in India they found a negative and significant relationship between the property status of women and experience of domestic violence.

Beside economics, various disciplines have proposed theories explaining domestic violence. For example, evolutionary psychologists argue that because the paternity of children was never certain in the evolutionary past, natural selection would have favoured proprietary behaviour by males with regard to sexual access to their mates. Spousal violence, in this view, stems from the insecurity and jealousy that males feel when their partners are exposed to the possibility of sexual encounters with other males (Wilson and Daly, 1993). Biological Theory explains violent behaviour by the changes in brain development due to trauma. Researchers in this group link the trauma of early exposure to chronic violence to changes in a child's brain functioning that lead to violent behaviour as an adult (Perry, 1997). Framed around the sociological perspective of social exchange theory, resource theory asserts that the family is a power system and that men with low levels of economic resources (earnings, social status, and educational attainment) may use violence as an alternative form of resource to control their partners.

This theory sees violence as an additional resource that men can use to maintain dominance within the family and suggests a correlation between poverty and intimate partner violence (Goode, 1971).

In his seminal paper, Heise (1998) provides a broad conceptual model for the determinants of domestic violence. In this model, he presents domestic violence as a function of contextual and community-level factors, household and individual-level factors, women's status and autonomy and interrelated effects of these factors. Household and individual-level factors include socioeconomic status, intergenerational exposure to violence, life cycle factors such as age and risk behaviours such as alcohol and substance abuse. Contextual and community-level factors, closely related to feminist arguments, include social norms on domestic violence and gender inequality. Heise argues that these factors not only have a direct effect on domestic violence but also have an indirect effect through their influence on female autonomy.

A large and growing body of literature has investigated the relationship between domestic violence and household and individual level factors outlined in Heise's framework. Kim and Cho (1992), in the case of Korea, found a decreasing risk of violence with the age of partners. Jejeebhoy and Cook (1997) provided evidence that violence is negatively related to the women's age at marriage. Family related factors such as the number of living children, especially sons, (Schuler et al., 1996) and support from women's natal family (Rao, 1997) are found to have a decreasing effect on domestic violence. Supporting the Biological theory, which emphasises the role of intergenerational transmission of violence, Straus and Gelles (1990) and Martin et al. (2002) show that experiencing or witnessing domestic violence in childhood has an increasing effect on the probability of resorting to violence or being a victim of it as an adult. There are also some studies documenting the link between violence and risk

behaviours such as substance and/or alcohol abuse that are noted in Heise's framework (Koenig et al., 2003; Rao, 1997).

Feminist theory further asserts that intimate partner violence originates from inequalities within the marriage. It strengthens male power and control and increases female subordination. Although there are different approaches to conducting research on intimate partner violence, most studies look at power imbalances that cause and perpetuate domestic violence against women. It is argued that these imbalances exist in patriarchal societies where structural factors hinder equal participation of women in the social, economic and political systems and they are reproduced when men exercise their power and control over women (Yllo 1994). Dobash and Dobash (1979) assert that patriarchy is fostered by the economic and social system and it contributes to domestic violence. They argue that the patriarchal system has defined the husband as the dominant, strong, authoritarian, aggressive and rational provider for the family, while the wife is traditionally accepted to be dependent, passive, submissive and soft. Therefore, typically, violence against women is explained in terms of a power struggle and it is argued that in a patriarchal society, men, with all the power, must resort to violence when their position of dominance is threatened. In that sense, it is argued that domestic violence is a consequence of patriarchy, and part of a systematic attempt to maintain male dominance in the home and in the society (Knickmeyer et al. 2004).

Both feminist theory and Heise (1998) emphasise the importance of social norms when analysing domestic violence and female autonomy. According to these arguments, patriarchy and conservative attitudes on gender equality lead to stronger social norms on traditional gender roles. Stronger social norms, in turn, lead to higher levels of domestic violence against women. Therefore, strength of the social norms is argued to be a key determinant of domestic violence.

3.4.2 Female autonomy and Domestic violence –the causality issue

With the growing recognition of the importance of female autonomy and awareness of the domestic violence problem, numerous studies have attempted to explain the link between these two concepts. The majority of the studies on this relationship focus on one direction of this link: the effect of female autonomy on domestic violence. However, the empirical studies do not provide consistent results; some studies find decreasing levels of domestic violence with higher female autonomy while others, in contrast, document a higher risk of violence with increasing female autonomy. Lamichhane et al (2011), in the case of Nepal, found that low autonomy of women significantly increases the odds of experiencing violence among married women. In the example of India, Jejeebhoy and Cook (1997) found lower levels of domestic violence when women have a greater control over the resources. On the other hand, Hindin and Adair (2002), using data from the Philippines, found that risk of violence was the highest when the major decisions related to the household were dominated by the women.[27] Menon and Johnson (2007) also provided evidence that women with higher autonomy in South India experience higher levels of violence.

These conflicting results can be found even amongst the studies focusing on the same country. For instance, the impact of micro saving and credit programs in Bangladesh on the level of domestic violence has been the subject of many studies. Some studies argued that membership to the microcredit systems was associated with lower levels of domestic violence; these programs not only increased the contribution of women to household income but also facilitated visibility and public exposure of women (Schuler et al., 1996; Kabeer 2001). In contrast, Rahman (1999) found that while only one-fifth

[27] Alternative scenarios were the husband's domination in decision making or joint decision making.

of the women saw a decrease in the violence, 70 per cent of the women reported an increase in violence from the time their membership in the program began.

Koenig et al. (2003), using data from two rural areas in Bangladesh, found that in the culturally more conservative area Sirajgonj, where the social norms related to the status of women are more traditional and rigid and social mobility of women is more restricted, higher individual-level women's autonomy and short-term membership in savings and credit groups were both associated with significantly higher risk of violence. However, in the less culturally conservative area Jessore, individual-level women's status indicators were unrelated to the risk of violence, but the community-level measures of women's status were associated with significantly lower risk of violence.[28]

The major problem that is common across these studies is that none of the studies consider possible reverse causality and only focus on one direction of this relationship; ignoring the potential reverse effect of violence on female autonomy. As summarised in the previous section, feminist theory defines domestic violence as a means that is used by the husband to control women and strengthen their subordination. In that sense, it is argued that domestic violence hinders female autonomy. Indeed this argument was empirically tested by Eswaran and Malhotra (2011) using the National Family Health Survey data of India. Their paper makes an important contribution to the literature for two reasons: it is the first study that uses a household bargaining model to analyse the relationship between domestic violence and female autonomy; also it is the first study that acknowledges the endogeneity problem due to the potential two-way relationship

[28] In order to capture individual level women's status, Koenig et al. (2003) used variables such as women's freedom to talk to strangers, freedom to seek medical care, freedom to possess any cash. To capture the community level women's status, they used percentage of married women in a community with any formal education, percentage of married women in a community belonging to a savings and credit group, and the mean value of the individual level women's status in each community.

between these two variables. In their household bargaining model, they choose a non-cooperative model due to the problems with the cooperative models (summarised in the previous subsection).

They designed their model in a way that domestic violence is a means to ensuring that the victim allocates resources more in line with the preferences of the abuser which is similar to previous household bargaining models. However, unlike the previous examples, domestic violence does not arise from a taste for violence (in fact the husband's utility decreases with the violence that he inflicts because this leads to strained relations, and loss of intimacy). The Eswaran and Malhotra model also takes into account that if a husband acquires more liberal views this may affect his decision to commit violence. In their model, female autonomy, which is captured by her decision making ability, among other things such as her education and income, is determined by the domestic violence that she confronts. Then main difference between the Eswaran and Malhotra model and the pervious household bargaining models is that they argue that factors that increase the reservation utility of women such as education and income cause an increase in the risk of violence instead of a decrease which is suggested by the previous models.

Based on the design of their model, they underline the potential reverse causality between domestic violence and female autonomy, and argue that such reverse causality can cause endogeneity problems in simple regression estimations. To overcome this endogeneity problem, they exploit a two stage least squares regression model and use the height index, calculated by the deviation of the woman's height from the mean value of height of women for the region that she belongs to, as an instrument for domestic violence. They assert that a husband is more likely to engage in spousal violence if he feels he can physically overpower his wife. They defend their choice of

instrument by arguing that the height of a woman is determined prior to marriage and is likely to be exogenous. They also note that their height index is a more appropriate indicator than the actual height of wife since the average height of women varies significantly across different states in India.

The results of their analysis show that domestic violence has a highly significant and negative effect on the autonomy of the women in the household. Although Eswaran and Malhotra argue that higher female autonomy increases the domestic violence that the women face, they do not empirically test this argument in their analysis. Instead, they used women's employment status in the first stage of their 2SLS estimation where the determinants of domestic violence are estimated. They find a positive effect of women's employment on domestic violence. However, they do not take into account the potential endogenity of employment status which is noted by Gibson-Davis et al. (2005).

Eng et al. (2010) used Structural Equation Modelling (SEM hereafter) to analyse the determinants of domestic violence in Cambodia focusing on the effect of the husband's control and spousal discussion on emotional and physical violence.[29] They captured the latent husband's control variable using indicators such as "Husband jealous if talking with other men," "Husband accuses her of unfaithfulness," and "Husband does not permit her to meet her female friends". They defined the spousal discussion as an opportunity for women to voice their opinion and engage in discussion and created a latent variable for the frequency of discussion between the spouses regarding daily life topics such as home, money, community happenings, and work. Their results showed that husband's control has a positive effect on both physical and emotional violence. The spousal discussion variable, on the other hand, has a positive effect only on emotional violence. Conceptually, spousal discussion, even the husband's control to

[29] This chapter also uses SEM. The details of the model are discussed in section 3.6.

some extent, is similar to female autonomy. However, their study seems to ignore the potential reverse causality and so the effect of violence on female autonomy.

3.4.3 Domestic violence in Turkey

The first nationwide research on domestic violence in Turkey was conducted in the Causes and Consequences of Domestic Violence Report (Turkish Prime Ministry, 1995). An interesting finding of this study was the mismatch between the responses of wives and husbands regarding the incidence of violence: while 34 per cent of men reported that, in the case of a disagreement with their partners they resort to physical violence, only 29 per cent of women reported that they were exposed to violence. A more recent survey, namely, Domestic Violence against Women in Turkey (The General Directorate on the Status and Problems of Women, 2008, GDSPW hereafter), reveals the magnitude of this problem in Turkey. According to the survey results, 40 per cent of women in Turkey were exposed to physical violence from their partners in their life time; 10 per cent were subject to violence in the last 12 months. Moreover, 15 per cent of women were sexually abused at least once in their life time. Using the data from this survey, Yüksel-Kaptanoğlu et al. (2012) analysed the determinants of domestic violence in Turkey. Their results show that the level of violence was much higher when one or both of the partners was/were exposed to violence in childhood. Women who had support of family and friends were less likely to experience violence. A surprising result of their study is that women's ownership of property and their ability to raise money were not significantly associated with partner violence. Altinay and Arat (2009) conducted a survey with 1,520 women and they found similar results as Yüksel-Kaptanoğlu et al. (2012) in terms of the magnitude of domestic violence and factors that affect violence. In addition, they compared the results of their research with the results of Causes and Consequences of Domestic Violence Report (Turkish Prime Ministry,

1995) and concluded that the magnitude of violence had not changed between the mid-1990's and mid-2000's.

Kocacik et al. (2007) conducted a study that covers four cities in Turkey namely, Sivas, Adiyaman, Denizli and Kirklareli in order to explore the determinants of domestic violence which is captured by acts such as battering, sexual abuse and wounding. Using the Heise framework, they included individual and social factors in a logistic regression in which a dummy variable representing domestic violence was the dependent variable. In terms of the effects of the individual characteristics of women such as income and education on violence, they found consistent results with the previous studies. The significant contribution of their study is that they also analysed the effect of household decision making and the husband's control over the wife's behaviours on domestic violence. Their results showed that the risk of violence was lower when the decisions are collectively made by husband and wife. Nevertheless, their study suffers from three major shortcomings. First, the cities in which the surveys were conducted do not provide a nationally representative sample as admitted by the authors. Second, they ignore the potential reverse causality between domestic violence and the decision making process. Finally, it is arguably not convincing to use only one question, "who makes the decisions in the household" to capture the wife's bargaining power or her autonomy. This is because, as Ballon (2011) and the results which will be presented below show, not all the decisions made by wife (or collectively made with husband) serve as a good indicator of the autonomy of the wife.

3.5 Data and Descriptive Statistics

This study will make use of the 2006 Household Structure Survey (HSS hereafter). This survey is the product of joint research by TurkStat and the General Directorate of Family and Social Studies on the household structure of Turkish families and provides a

nationally representative sample. 48,235 individuals above the age of 18 were interviewed from 11,854 households. The questions in the survey are grouped into two categories. The first category contains questions related to socio-economic factors, such as: age, marital status, household size, income level, education. The second group of questions are related to perceptions such as opinions on marriage and roles and duties of family members in the household. In total, this survey provides numerous questions about the household structure, perceptions and habits of Turkish families. One of the important properties of the dataset is that interviews are conducted separately with each member of the household to assure that the answers are not heard by the other household members. The individuals who are single, divorced or widowed are dropped from the sample. Then, only the heads of households and their partners are kept as some of the key variables such as the female autonomy indicators (which will be explained below) are relevant to them only. Finally, couples are omitted if either of the partners did not respond to the survey. The final sample that is used in the estimation consists of 6,435 couples (see Table 3.A2 for the exclusion criteria for the sample).

The three main variables that will be used in the analyses are: domestic violence, female autonomy, and social norms. To capture domestic violence, answers to the question of "what is your partner/husband's reaction if you cannot agree on an issue?'' will be used. Participants are asked to choose between (not exclusively) 5 options: does not do anything, sulks, leaves the home, shouts, resorts to physical violence.

Table 3.1 provides the responses from both husbands and wives about their and their partner's reaction together with the frequency of these reactions. In almost all of the scenarios, the answers given by partners closely match each other especially the one for physical violence occurring "Never". Table 3.1 shows that, around 92 per cent of the women report no violence in the 12 months previous to the interview which means in

around 8 per cent of the marriages violence occurred at least once in the previous 12 months. Using the women's answers to this question, a dummy variable that represents domestic violence which takes the value of 1 if the woman reports that she was physically assaulted even once within the last 12 months and 0 otherwise, is created. To check for robustness of the variable, an alternative dummy variable which takes the value of 1 if the wife reports experience of violence or the husband admits resorting to violence is also created. The values of the original violence variable and the alternative variable were very similar showing that the responses of husbands and wives match.

Table 3.2, column 1 provides the descriptive statistics for the sample. More than half of the women in the sample are between 25 and 44 years old. While the majority of the women were between 18-25 years old when they got married, almost 30 per cent of women reported that they were younger than 18 when they got married.[30] Like the women in the sample, more than half of the men are between 25 and 44 years old. However, marriage before the age of 18 is not as common among men as it is in the case of women. The HSS includes a categorical variable providing income bands. As most of the women in the sample report having zero or low levels of income, a dummy variable to represent women who have a positive income is created. According to Table 3.2, only 20 per cent of the women in the sample reported having an income. This is consistent with the low level of female labour force participation in Turkey. Furthermore, only 17 per cent of women reported that they have any form of asset ownership (such as house, land, automobile, or other). The majority of men in the sample reported having an income. 25 per cent of the women do not have any form of formal education and nearly half of them attended only primary education.

[30] Since 2001, the legal minimum age to enter into a marriage in Turkey is 17 although this requires parental consent if a participant is younger than 18. Before the year 2001, with the consent of parents, women were allowed to marry at the age of 15 and men were allowed to marry at the age of 17.

The analyses also look at variables relating to the family setup, namely arranged marriage and bride price. Arranged marriage is a type of marital union in which partners are chosen by the family or other elders in the family. In the HSS, women are asked to report if the marriage was an arranged one and furthermore, they are asked if the arranged marriage was with or against their will. The arranged marriage category is created only for those whose marriage was against their will. Even after considering this difference, 33 per cent of women in the sample stay in this category. Bride price can be defined as the amount of money, property or wealth paid by the groom or his family to the parents of the bride in return for the consent of marriage.[31] In around 16 per cent of the marriages, bride money is reported to have been paid.

Table 3.2, column 2 shows the prevalence of violence against women by the characteristics of women and men. Although it is not possible to find a clear-cut relationship between age and violence, the incidence of violence is slightly higher at higher ages for both men and women. Lower levels of violence are observed in the marriages that start at early ages. The incidence of violence for the women who have an income or for those who have asset ownership is lower. The incidence of violence gets smaller at the higher levels of husband's income. Similarly, the incidence of violence decreases with the education level of men and women. Table 3.2 also shows that, on average, the incidence of violence is higher in arranged marriages and the marriages in which the bride price is paid.

Female autonomy is captured by the questions asking who in the household makes the final decision regarding the following: choice of house type, choice of house setup, issues about children, shopping, relations with relatives, relations with neighbours and issues about holiday and entertainment. For each question participants can choose

[31]See Anderson (2007) for further details on bride price and its differences from dowry.

between ‘husband alone’, ‘wife and husband together’ and ‘wife alone’. In some studies in the female autonomy literature, for example Eswaran and Malhotra (2011), a woman’s individual decision making is accepted as an indicator of autonomy. However the questions that are asked to women in these studies are related to individual behaviours such as buying jewellery or if she can access health care for herself. On the other hand, Koenig et al. (2003) used the woman having a say on family decision making as an indicator of autonomy. As the questions in the HSS are related to issues that are relevant to the whole household, the indicator variables will take the value of 0 if the man makes the final decision individually (for women this will be response ‘husband’) and 1 otherwise (for women this collapses the two categories ‘wife’ and ‘wife and husband together’). As a robustness check an alternative specification of the female autonomy indicator variables in which each indicator takes three different values for “wife alone”, “husband alone” and “wife and husband together” is also used. The results were qualitatively the same.

Table 3.3 shows the responses given by the women about the decision making process in the household. Based on the table, the husband’s dominance is highest for decisions on choice of house, shopping and entertainment and holiday which are all related to household finances. Compared to these three decisions, women have more say on house setup and children related issues (only asked to couples who have any children) which are traditionally accepted as the responsibility of women in the household.

As noted in Section 3, social norms are argued to be a key determinant of domestic violence. To capture the effect of social norms, the social norms of the husband variable which represents his perception of traditional gender roles, is created. The indicators for this variable will be created using the husband’s agreement with the following statements; “Having a male child increases the prestige of the wife”, “Not doing

housework is a sole reason for divorce", "The best age of marriage for women is between 15-19", "Continuation of the family is guaranteed only by a son", "It is appropriate for close relatives (first cousins) to marry" and "It is not appropriate for women to work".[32]

Table 3.4 provides the mean values of the dummy variables that are used as the indicators of the husband's social norms. A considerable number of husbands value sons more than daughters; 33 per cent believe that having a male child increases the prestige of the wife and, for 43 per cent, continuation of the family is guaranteed only by a son. Around 18 per cent of husbands agree that husbands can divorce if the women do not do the housework. One interesting observation is that even though the legal minimum marriage age is 18, 16 per cent of the men assert that the best age of marriage for women is from 15 to 19. Around 13 per cent of the men in the sample reported that they approve of the marriage between close relatives (defined as first cousins in the HSS). The respondents are also asked the reason of their approval of marriage between close relatives. Further analysis of the data showed that the main reasons of their approval were better knowledge of the family roots (37 per cent), having more things in common (28 per cent) and protection of traditions (15 per cent).

As noted in Section 3, around 24 per cent of women in Turkey do not work or look for a job since their husbands or elders prevent them from doing so. Consistent with this observation, 23 per cent of husbands agree that working is not appropriate for women. Further investigation of the data shows that, among those who agree with this statement, 61 per cent assert that the main duty of women is to take care of children at home and

[32] Using the same data, Goksel (2012) exploited some of these questions to capture social norms in a different setting where she analysed the effect of conservativeness on labour force participation of women. Her findings suggest a significant negative effect of conservativeness on female labour force participation.

14 per cent argue that it is against tradition. Table 3.4 also presents the mean values of the dummy variables that are created using the answers given by the wives. For all six indicators the magnitude of the mean values is smaller for women. The values of indicators regarding women doing housework, women's increasing prestige with having a male child, and close relative marriage are quite close to each other. On the other hand, the indicators regarding the continuation of family with a male child and employment of women have much greater mean values for men than women.[33]

3.6 Methodology

To analyse the relationship between domestic violence and female autonomy in the household, an Item Response Theory (IRT hereafter) model will be applied. IRT models are a member of SEM family in which the observed (indicator) variables are discrete and the latent variables are assumed to be continuous. In that sense, IRT models can be considered as factor models for non-continuous (including binary) indicators. SEM is used in many applications of economics, sociology, environment, psychology and health related studies. For example, Roth et al. (1989) used SEM to understand the relationship between exercise, fitness, stress, hardiness (resiliency) and illness. Grandjean et al. (1997) used SEM to analyse the effect of in-utero methylmercury exposure on neurodevelopment. Savage et al. (2013) used SEM to analyse the new class structure in UK. There are two main benefits of choosing SEM over the standard multiple regression technique. First, it allows for the measurement error related to the indicator variables. For instance, in this study social norm and female autonomy variables are assumed to be latent variables as their true values are unobservable. However, there are some indicators of these variables which include a

[33] A t-test is performed in order to see if the mean values of the indicator variables are significantly different for men and women. The results of the test suggest a significant difference at the 1 per cent level.

measurement error component in their absolute values. Using SEM, it is possible to separate the variation in indicator variables into two components: measurement error in the indicator and the variation due to the change in the latent variables. Second, in the multiple regression technique, in order to estimate the coefficients associated with a reverse causality relationship between two variables, one needs to run two separate regressions for each direction of the relationship. SEM, on the other hand, allows association between latent variables in a single model specification.

The conventional way of model presentation in SEM is to use path diagrams. Path diagrams contain two components, namely a structural model, which shows the causal hypothesis between the variables of interest, and measurement models which show the relationship between latent variables and indicators of these variables. Figure 3.1 shows the path diagram of the measurement model (the dashed rectangles) and the structural model (the solid rectangle). Latent variables, in this study social norm and female autonomy, are shown as ellipses. The exogenous observed variables are presented by dashed rectangles and domestic violence - the endogenous observed variable- is represented by a solid rectangle. The single headed arrows represent hypothesized direct effects of one variable on another. Two headed arrows represent the covariance between the variables. ε(.)'s are the error terms corresponding to the relevant variables and in the measurement models, they represent measurement errors in the indicator variables.

Model specification

The model specification that is used in the empirical analysis of the chapter is designed in accordance with the literature especially with the household bargaining model of Eswaran and Malhotra (2011) and the feminist theory of domestic violence. As

explained in Section 3, feminist theory argues that the patriarchal system has defined the husband as the dominant, strong, authoritarian, aggressive and rational provider for the family, while the wife traditionally is accepted to be dependent, passive, submissive and soft. In the line with this argument, Eswaran and Malhotra (2011) also assert that husbands with more liberal views can see wife beating as being more reprehensible. Based on these arguments, the structural model is designed in a way that the social norms of the husband have a direct effect on the violence. In other words, a husband with strong social norms can be expected to resort to violence more than one with relatively weak norms.

Feminist theory, also asserts that men resort to violence to sustain their dominant position. This argument has two implications: violence can affect the level of female autonomy in the household and the level of female autonomy can affect the incidence of violence. It means that there is a potential endogeneity problem. Eswaran and Malhotra (2011) used a two-stage instrumental variable model in their study to deal with endogeneity, which is a common way to solve this problem. Nevertheless, their analysis suffers from measurement error as they simply take the average of binary indicator variables to capture female autonomy. In order to allow for endogeneity and measurement error problems at the same time, a non-recursive version of SEM is used here. A non-recursive model can be defined as a SEM model in which two variables in the structural model have a feedback loop between them (Kline, 2011). As can be seen from Figure 3.1, the causality between violence and female autonomy is represented by a direct feedback loop showing the bilateral relationship between violence and female autonomy. Also, the disturbances of female autonomy and violence are allowed to be correlated to account for the unobservable variables that affect both female autonomy and domestic violence.

Finally, there are two sets of observable variables represented by the dashed rectangles. The first set of observable variables (X^{FA}) has an effect only on female autonomy. This set contains: age, age at marriage, education level and income status of the wife; arranged marriage and the number of male and female children. The second set of observable variables (X^{SN}) affects the social norms of the husband, it consists of age, age at marriage, education level and income of husband.

Structural Model:

The structural part of the model defines a linear relationship between the latent variables and the fixed covariates. In matrix form,

$$\theta = \delta + B\theta + \Gamma X + \varepsilon \qquad (1)$$

Where, $\theta = (\theta^V, \theta^{FA}, \theta^{SN})$ is the vector of dependent variables including violence θ^V, female autonomy θ^{FA}, and the husband's social norms θ^{SN}; δ is the vector of constants; X is the vector of covariates affecting the latent variables; B and Γ are the matrices of coefficients to be estimated and ε is the vector of independent and identically distributed error terms corresponding to the relevant dependent variables. The coefficient matrices in (1) are:

$$\begin{bmatrix} \theta^V \\ \theta^{FA} \\ \theta^{SN} \end{bmatrix} = \begin{bmatrix} \delta_V \\ \delta_{FA} \\ \delta_{SN} \end{bmatrix} + \begin{bmatrix} 0 & \beta_{12} & \beta_{13} \\ \beta_{21} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta^V \\ \theta^{FA} \\ \theta^{SN} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \gamma_{22} & 0 \\ 0 & 0 & \gamma_{33} \end{bmatrix} \begin{bmatrix} 0 \\ X^{FA} \\ X^{SN} \end{bmatrix} + \begin{bmatrix} \varepsilon_V \\ \varepsilon_{FA} \\ \varepsilon_{SN} \end{bmatrix}$$

Female autonomy and the husband's social norms are assumed to be continuous latent variables. As it can be seen from Figure 3.1 and the matrix representation above, there is a reciprocal relationship between the continuous latent female autonomy variable and the violence variables. As Maddala (1983, pp. 117-118) notes, a reciprocal relationship between a continuous and a binary variable is logically inconsistent and it is not

possible to perform estimations with standard maximum likelihood methods. To overcome this problem, the weighted least squares means and variance adjusted (WLSMV) methodology proposed by Muthén et al. (1997) is adopted. In this methodology, instead of the observed value of the binary violence variable, the underlying continuous response variable (θ^V) behind the observed violence variable acts as the predictor.[34]

Measurement Model:

For a formal representation, assume that v_i, a binary variable that represents the incidence of violence for the individual *i*, is an indicator of the event that some underlying unobserved continuous variable, say θ_i^V, exceeds a certain threshold (without loss of generality, it can be taken as zero) such that,

$$v_i = \begin{cases} 1 \ if \ \ \theta_i^V > 0 \\ 0 \ if \ \ \theta_i^V \leq 0 \end{cases} \qquad i = 1....\text{N} \tag{2}$$

The underlying response variable is defined by a linear in parameters specification (suppressing *i*);

$$\theta^V = \delta_V + \beta_{12}^{'}\theta^{FA} + \beta_{13}^{'}\theta^{SN} + \varepsilon_V \qquad \text{and } \varepsilon_V \sim N(0, \sigma^2) \tag{3}$$

Then, the marginal probability of a positive response is given by

$$\Pr(v = 1) = \Pr(\delta_V + \beta_{12}^{'}\theta^{FA} + \beta_{13}^{'}\theta^{SN} + \varepsilon_V > 0) = 1 - \Phi(-\delta_V - \beta_{12}^{'}\theta^{FA} - \beta_{13}^{'}\theta^{SN}) \tag{4}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution.

By definition, the latent variables θ^{FA} and θ^{SN} are not observed. Instead, there are *m* number of indicators assumed to measure female autonomy and the husband's social

[34] In contrast, with maximum likelihood methods it is the observed binary variable itself that is the predictor. The inconsistency stems from the fact that in maximum likelihood methods each direction of the relationship between female autonomy and domestic violence refers to a different regression: a linear regression when the female autonomy is the dependent variable and a logit regression when the binary domestic violence is the dependent variable.

norms. Let Z_j^q be the measure of θ^q, where q=(*FA*, *SN*) and j=(1, ...,m^q). Accordingly, the measurement equations for the latent variables are given as:

$$Z_j^q = \tau_j + \alpha_j^q{}'\theta^q + u_j^q \quad (5)$$

In SEM terminology, τ_j are defined as the thresholds, the coefficients α_j^q are referred to as the factor loadings and they represent the information of the latent variables that is contained in the indicators. The disturbance terms u_j^q represent the measurement errors of the indicator variables.

Model identification and assumptions:

For the identification of the model, there are four requirements:

1) $m^q \geq 2$, a minimum of two measures for each latent variable.[35]

2) The location of the latent variables should be defined. Therefore, the location of the latent variables are centered by setting constants in the structural model to zero, ($\delta = 0$).

3) $\alpha_1^q = 1$, the factor loadings are identified up to a scale, so the factor loading of one of the measures is normalised to 1.

4) $Cov(u^{FA}, u^{SN})=0$, the measurement errors in one measurement model are uncorrelated with the measurement errors in the other measurement model.

The requirements that are noted above are standard for any IRT model applications. As the model in this study specifies a reciprocal relationship due to the potential reverse causality between violence and female autonomy, three additional assumptions are made:

[35] SEM allows usage of more observations as one does not need to drop observations which have missing values for some variables. For instance, couples without children have missing values for the female autonomy indicator related to the decisions regarding the children. As long as there are at least two indicators, one can identify the latent variable.

5) $\mathrm{Cov}(\varepsilon_{\mathrm{V}}, \varepsilon_{\mathrm{FA}}) \neq 0$, the error terms of the reciprocal variables are correlated. As Schaubroeck (1990) notes, much of the error in predicting a reciprocally related variable (say θ^{FA}) will be due to its corresponding variable (θ^{V}). Because θ^{FA} also causes θ^{V} in turn, the errors in predicting θ^{FA} will become part of the estimator of θ^{V}. Schaubroeck also argues that failure to estimate this correlation may bias the analysis and the extent of the bias increases with the size of the correlation.

6) There should be at least one exogenous variable that affects one of the endogenous variables in the reciprocal relationship but does not affect the other (exclusion criteria). The model in this study satisfies this requirement as in its specification, as it can be seen in the path diagram (see figure 3.1) while the husband's social norms variable only affects violence, the individual characteristics of the wife and arranged marriage, bride price variables affect only female autonomy.

7) Any changes in the system underlying the feedback relationship between female autonomy and domestic violence have already manifested their effects and therefore the system is in a steady state.

3.7 Results

Table 3.5 shows the results of the measurement models for female autonomy and the social norm latent variables. The estimated parameters show the strength of the relationship between the latent variable and the observed indicator (item) variable. More precisely, they show the discriminatory power of items between the individuals with similar scores on the latent variable. All of the parameter estimates in the measurement models are statistically significant and positive which means the female autonomy and the husband's social norm latent variables have a positive and significant effect on their indicators.

As the latent variables and the underlying continuous response behind the indicator variables do not have a unit of measure, one can use the standardised estimates of coefficients to interpret the effect of latent variables on their indicators. The standardised coefficients show the mean change in indicators (in standard deviation units) for a 1 standard deviation change in the latent variables. According to the results of the measurement model, women having a say on choice of entertainment/holiday, relationships with the relatives and neighbours, in other words people outside the household, are the best indicators of female autonomy in the household. On the other hand, having a say on the house setup, is the weakest among the indicators of the female autonomy. This result is not surprising since house setup, or housework, is traditionally accepted as the duty of women. In terms of husband's social norms, indicators showing husbands' agreement on "having a male child increases the prestige of the wife" and "continuation of the family is guaranteed only by a son" have greater factor loadings. This means that superiority attached to a male child (or may be male as gender in general) by the husbands seems to be the strongest indicator of social norms of husbands.

Table 3.6 presents the results of the structural model. The female autonomy in the household is found to be positively and significantly affected by the wife's education. The coefficients of the age and age at marriage on female autonomy are positive: however, they are statistically insignificant. Having a personal income has an increasing effect on female autonomy. The positive coefficient for asset ownership is consistent with the arguments of Panda and Agarwal (2005) who assert that property ownership increases the bargaining power of women by acting as a tangible exit-option from the marriage. However, the coefficient for this variable fails to attain statistical significance. As noted by Malhotra and Mather (1997) access to assets does not

necessarily mean that women have control over these resources. Accordingly, the insignificant coefficient for asset ownership can be due to the women's lack of control over these assets. Arranged marriage and bride price has a negative effect on female autonomy. This is not a surprising result since the women who are forced to enter into a marriage are less likely to have a say on decision making processes within the household. Considering the results of the measurement model of the husband's social norms, which show the importance given to male child by husbands, two separate variables to control for the number of male and female children are included. The estimated coefficients for these variables are the same in terms of both size and magnitude, indicating a negative effect of having children on female autonomy.

The strength of the husband's social norms is found to be negatively related to the socio-economic status. The effect of education is negative and strongly significant. Similarly, the level of income has a negative effect on the husband's social norms. The effect of age is negative. The age at marriage is not significantly related to social norms.

Table 3.6 also reports the relationship between key variables of interest. First, domestic violence has a negative effect on female autonomy. This finding is evidence suggesting that women who are victims of domestic violence have less autonomy since, in the decision making process, they are forced to be in line with the preferences of the abuser. This result is consistent with the results of Eswaran and Malhotra (2011) who proposed a non-cooperative model of spousal violence in which women are, in the presence of violence, forced to allocate the resources in accordance with the husbands' choice. Second, female autonomy has a significant and negative effect on domestic violence. In other words, the incidence of violence is less likely to occur once the autonomy of women, which is promoted by the income and the education they possess and reflected by their presence in decision making process, gains acceptance. This result is in

accordance with the feminist theory on violence which asserts that intimate violence stems from the inequalities within the household and it is used by husbands to maintain subordination of women. In this sense, husbands' resort to violence becomes less likely in the absence of these inequalities and male dominance.

Finally, confirming the importance of conservatism, the strength of the husband's social norms on traditional gender roles positively affects the risk of violence. This result supports Heise (1998) who argues that adherence to rigid gender roles increases likelihood of violence against women. Koenig et al. (2003) find supporting results for this argument at the community level. In contrast to the existing literature this chapter examines the effect of social norms on violence at the individual level. This study contributes to the literature as its key findings support the literature on the relationship between domestic violence, female autonomy and social norms. More importantly, it makes a contribution to the literature since the model used in this study allows analysis of the relationship between these three concepts in a single framework that accounts for the potential reverse causality, and so endogeneity, between domestic violence and female autonomy.[36]

3.7.1 Robustness check- Alternative specification using the wife's social norms

The results of the model suggest that the strength of the husband's social norms positively affect violence and are negatively affected by education. As mentioned in

[36] As noted in the methodology section, exclusion criteria which are necessary to deal with endogeneity require at least one variable that affects domestic violence but does not affect female autonomy. In that sense, the husband's social norms act as an instrument for domestic violence. Since it is not possible to do traditional tests on validity of instruments in structural equation modelling, an alternative specification was estimated, where the husband's social norms were allowed to have a direct effect on both female autonomy and domestic violence, and the husband's characteristics such as education and income level were allowed to affect domestic violence directly to be able to identify and estimate the model. However, it was found that the effect of the husband's characteristics on domestic violence was insignificant; and the effect of the husband's social norms on female autonomy was also found to be insignificant Additionally, many other variables such as financial stress, alcohol usage and family support were also considered but their effect on violence was found to be insignificant so they could not be used as instruments.

Section 3.4.1, Heise (1998) argues that social norms affect domestic violence also indirectly through its effect on female autonomy. Including an additional path from the husband's social norms to female autonomy in Figure 3.1 would cause identification problems in the model (see assumption 6 in section 3.6). In order to overcome this problem and to see how female autonomy is affected by women's own social norms, an alternative model specification which includes the social norms of the wife is estimated.[37] By doing so, the effect of education and income on women's social norms is also analysed.

Figure 3.2 presents the path diagram of the alternative model specification. The same indicators of social norms are used for the husband and wife to test if the indicators show a different pattern for men and women. Finally, all the covariates that affect female autonomy are also allowed to affect wife's social norms.

Tables 3.7 and 3.8 provide the measurement model of the wife's social norms latent variable and the estimation results of the structural model of the alternative specification, respectively.[38] Looking at the results of the measurement model in Table 3.7, it can be seen that most of the factor loadings have greater values for women than men. In other words, these indicators contain more information on the strength of the social norms for women than they do for men. It is also seen that the superiority that is attached to a male child is once again the strongest indicator of social norms. One important observation is that the factor loading regarding the employment of women is much greater for women than it is for men. In that sense, being against women's

[37] A strong correlation between the husband's and wife's social norms can cause further endogeneity problems. However, the results of the model show that the correlation between these two latent variables is very low with the value of 0.05 (for comparison the correlation between female autonomy and wife's social norms is 0.32). Also as noted in footnote 30 above, the values of indicators of social norms are significantly different for men and women.

[38] We do not report the results of the measurement models of female autonomy and the husband's social norms and the effect of covariates on the husband's social norms. This is due to very small changes in these estimation results which leave the results qualitatively the same.

employment is a very informative indicator of women's acceptance of traditional gender roles.

The structural model results in Table 3.8 show that, compared to the main model, while their sign and significance level remain the same, there is an increase in the magnitudes of the effects of female autonomy and husband's social norms on violence. As expected, the wife's social norms have a negative and significant effect on her autonomy in the household. A possible explanation can be that a woman who internalises the traditional social norms and accepts the subordination will be less likely to participate in decision making processes than a woman with more egalitarian attitudes. The main difference between the alternative specification (Table 3.8) and the main model (Table 3.6) is that the effects of variables such as education, income, bride money, number of children on female autonomy lose their statistical significance in the alternative specification. On the other hand, these variables have a significant effect on wife's social norms. Similar to the case of men, while education, having an income, and age have a negative effect on the wife's social norms, the effects of arranged marriage, bride money and the number of children are positive and highly significant. One conclusion that can be drawn from the differences between the main model and the alternative specification is that education and having an income indirectly affect the level of the wife's autonomy through her social norms. Once the level of wife's social norms is controlled for, there is not an observed direct effect of these variables on female autonomy.

3.8 Conclusion

This study makes three important methodological contributions to the existing literature on domestic violence and female autonomy. First, it analyses the two-way relationship between female autonomy and domestic violence within the same model. Second,

exploiting the advantages of SEM it accounts for potential endogeneity and measurement error problems and finally, it tests the importance of social norms at the individual level.

The results of the study provide many suggestions for policy makers. For example, education and income not only weaken the strength of traditional social norms for both men and women but also promote a higher level of female autonomy. Therefore, policy makers should consider focusing on education, particularly promoting the educational attainment of women. According to the Global Gender Gap Report 2013, the female labour force participation rate in Turkey is 30 per cent and this ratio puts Turkey in the 123rd place out of 136 countries. As mentioned in Section 3, around 24 per cent of women who do not participate in the labour force report that they do not have autonomy on their participation decision. Therefore, policies targeting new employment opportunities for women should be accompanied by campaigns promoting female autonomy in cooperation with the civil movements and organisations that support empowerment of women's socioeconomic status. Finally, the results of the analyses show that women who do not have any source of income and have a low level of education suffer oppression in the household which in turn makes them the most vulnerable group to domestic violence.

There are some potential limitations of the study, which should be acknowledged. The first limitation is the cross-sectional nature of the data. As Kaplan et al. (2001) note, a cross-sectional design of a model with a feedback loop gives only a snapshot of an on-going dynamic process. In that sense, the model used in this study, controlling for husband's social norms, analyses the effect of female autonomy on violence by comparing the incidence of violence for women with different levels of autonomy. Therefore it is not possible to analyse the effect of female autonomy on violence over

time for the same women. In other words, the change in incidence of violence due to the increasing autonomy of an individual woman can only be analysed with panel data which is not available here. For instance, Koenig et al. (2003) argue that "more autonomous women, at least initially, are likely to violate established norms concerning gender roles and call into question the larger family's honor and prestige and, as a consequence, to incur a higher risk of domestic violence … it is only after women's individual and collective empowerment and autonomy gain acceptance and become commonplace that reductions in the risks of domestic violence are likely to be observed" (p.285). Therefore, the positive effect that is found in this study should be interpreted as the effect of female autonomy on domestic violence after the change in female autonomy fully manifests its effect.

Second, although the SEM model accounts for measurement errors in the indicators of the latent variables, the incidence of violence is likely to be underreported by the respondents given the intimate nature of the subject. Also, the question of "what is your partner/husband's reaction if you cannot agree on an issue?'' may not be the best question to capture incidence of violence as domestic violence can also occur in the absence of a disagreement since a husband can resort to violence due to psychological problems or substance abuse. Moreover, the dataset only contains information about physical violence. However, violence can take a sexual or an emotional form as well. It should also be noted that the type of physical abuse is not specifically addressed in the question that captures domestic violence. This leaves the interpretation of what violence is entirely to women. It is possible that some women may in fact simply identify physical violence only when a severe injury occurs. Additionally, the dataset provides information about violence that occurred only in the last 12 months. For example, Domestic Violence against Women in Turkey (GDSPW, 2008) documents that around

40 per cent of women in Turkey suffered from physical violence and around 15 per cent were sexually assaulted in their life time. Finally, the dataset does not contain any information on exposure to violence in childhood so it is not possible to test the intergenerational transmission of violence.

Despite these limitations, the findings of the chapter make an important contribution to the literature on domestic violence and female autonomy in Turkey. The rigorous analysis is based on looking at the two way relationship between the two concepts; and the findings suggest presence of significant reverse causality where domestic violence has a negative effect on female autonomy, but an increase in female autonomy helps to decrease the incidence of domestic violence. This study also establishes, for the first time, the importance of social norms at the individual level, when analysing the relationship between domestic violence and female autonomy. The empirical strategy that is adopted in the analysis and the findings presented here will hopefully serve to stimulate further research in this area especially for the case of Turkey.

TABLES OF CHAPTER 3

Table 3.1: "What is your and your spouse's reaction if you cannot solve a problem by talking?''(Per cent)

	Men				**Women**			
	Your Reaction				Your Reaction			
	Usually	Sometimes	Rarely	Never	Usually	Sometimes	Rarely	Never
Shout	27.0	38.6	11.7	22.4	15.9	32.3	12.8	38.7
Sulk	5.4	19.5	11.3	63.6	10.2	27.9	11.8	49.9
Leave the House	1.9	4.7	4.0	89.2	0.9	2.1	2.5	94.3
Keep Silent	22.4	34.3	13	30.1	41.7	29.1	8.5	20.5
Physical Violence	0.5	3.1	4.0	92.2	0.2	0.5	1.4	97.7
	Wife's Reaction				Husband's Reaction			
	Usually	Sometimes	Rarely	Never	Usually	Sometimes	Rarely	Never
Shout	13.7	27.9	12.2	46.1	29.8	35.6	10.5	23.9
Sulk	8.5	25.9	11	54.4	6.5	18.5	10.3	64.5
Leave the House	0.5	1.9	2.1	95.3	1.6	3.6	3.8	90.8
Keep Silent	39.1	31.5	8.2	21.0	22.4	33.3	13.3	30.9
Physical Violence	0.2	0.7	1.5	97.4	0.9	3.0	3.6	92.2

Note: Sample weights are used.

Table 3.2: Descriptive Statistics for Explanatory Variables

	Mean value of the dummy variable	Mean value of violence dummy variable
Wife's age		
18-24	0.07	0.07***
25-34	0.29	0.05***
35-44	0.26	0.07***
45-54	0.20	0.07***
55-64	0.11	0.08***
65+	0.07	0.07***
Wife's age at marriage		
<18	0.28	0.03***
18-24	0.62	0.07***
25-29	0.08	0.07***
>30	0.01	0.09*
Husband's age		
18-24	0.01	0.06**
25-34	0.23	0.06***
35-44	0.29	0.06***
45-54	0.22	0.09***
55-64	0.13	0.08***
65+	0.11	0.08***
Husband's age at marriage		
<18	0.06	0.06***
18-24	0.58	0.06***
25-29	0.29	0.08***
>30	0.07	0.07***
Does wife have an income		
No		0.08***
Yes	0.20	0.06***
Wife's asset ownership		
No		0.08***
Yes	0.17	0.06***
Husband's monthly income [1]		
No income	0.06	0.12***
Less than 400	0.18	0.10***
401 - 600	0.32	0.08***
601 - 800	0.20	0.06***
801 - 1 200	0.16	0.05***
1 201 - 2 500	0.07	0.03***
More than 2 501	0.02	0.05*
Wife's education		
No formal education	0.24	0.10***
Primary	0.53	0.07***
Secondary	0.07	0.07***
High school or above	0.16	0.04***
Husband's education		
No formal education	0.09	0.12***
Primary	0.52	0.08***
Secondary	0.12	0.07***
High school or above	0.28	0.05***
Arranged marriage		
No		0.06***
Yes	0.33	0.10***
Bride price		
No		0.06***
Yes	0.16	0.12***
Number of daughters [2]	1.33 (0.02)	
Number of sons [2]	1.43 (0.02)	
Observations	6,435	

Note: Sample weights are used. 1) In terms of Turkish Liras. 2) Number of daughters and sons enters the analysis as continuous variables. Standard deviations are given in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table 3.3: Mean Values of Dummy Variables for Female Autonomy Indicators

"Who makes the final decisions on following issues?"

	Husband	Husband-Wife	Wife
Choice of House	0.29	0.61	0.11
House Setup	0.12	0.46	0.42
Children Related Issues	0.14	0.68	0.13
Shopping	0.19	0.65	0.16
Relationship With Relatives	0.16	0.76	0.08
Relationship With Neighbours	0.14	0.72	0.14
Entertainment and Holiday	0.17	0.78	0.05

Note: Sample weights are used.

Table 3.4: Mean Values of Dummy Variables for Social Norm Indicators

"Do you agree with following statements?"

	Husband	Wife
Having a male child increases the prestige of the wife	0.33	0.30
Not doing housework is a sole reason for divorce	0.18	0.16
The best age of marriage for women is between 15-19	0.16	0.11
Continuation of the family is guaranteed only by a son	0.43	0.35
It is appropriate for close relatives (first cousins) to marry	0.13	0.12
It is not appropriate for women to work	0.23	0.11

Note: Sample weights are used.

Table 3.5: Estimation Results of the Measurement Model

	Coefficient	Std. Err	Standardised coefficient	R-square
Latent variable: Female Autonomy (FA)				
Choice of House	1.00	Na	0.83	0.70
House Setup	0.96***	0.01	0.80	0.64
Children Related Issues	1.04***	0.01	0.87	0.75
Shopping	1.04***	0.01	0.87	0.75
Relationship With Relatives	1.19***	0.01	0.98	0.96
Relationship With Neighbours	1.15***	0.01	0.95	0.90
Entertainment and Holiday	1.13***	0.01	0.94	0.87
Latent Variable: Husband's Social Norm (SN)				
Having a male child increases the prestige of the wife	1.00	Na	0.58	0.34
Not doing housework is a sole reason for divorce	0.48***	0.07	0.28	0.08
The best age of marriage for women is between 15-19	0.60***	0.07	0.35	0.12
Continuation of the family is guaranteed only by a son	0.87***	0.09	0.51	0.25
It is appropriate for close relatives (first cousins) to marry	0.60***	0.08	0.35	0.12
It is not appropriate for women to work	0.63***	0.08	0.37	0.13

Notes: Sample weights are used. *** p<0.01, ** p<0.05, * p<0.1

Table 3.6: Estimation Results of the Structural Model

	Coefficient	Std. Err	Standardised coefficient
Dependent Variable: Violence			
Female Autonomy	-0.79***	0.00	-0.67
Husband's Social Norms	0.16*	0.09	0.09
Dependent Variable: Female Autonomy			
Violence	-0.77***	0.16	-0.91
Education level [1]			
Primary	0.06*	0.03	0.03
Secondary	0.06	0.04	0.03
High school or above	0.12*	0.07	0.05
Age [2]			
25-34	0.00	0.03	0.00
35-44	0.04	0.04	0.02
45-54	0.04	0.05	0.02
55-64	0.03	0.05	0.01
65+	0.02	0.06	0.01
Age at Marriage [3]			
18-24	0.00	0.01	0.00
25-29	-0.03	0.03	-0.01
>30	0.00	0.05	0.00
Economic status and family variables			
Having income	0.05*	0.03	0.02
Wife's asset ownership	0.03	0.02	0.01
Arranged marriage	-0.06*	0.03	-0.03
Bride price	-0.05*	0.03	-0.02
Number of daughters	-0.03*	0.01	-0.04
Number of sons	-0.03*	0.01	-0.04
Dependent Variable: Husband's Social Norms			
Education level [1]			
Primary	-0.19***	0.05	-0.16
Secondary	-0.29***	0.07	-0.16
High school or above	-0.37***	0.07	-0.28
Age [2]			
25-34	-0.24*	0.14	-0.17
35-44	-0.18	0.15	-0.14
45-54	-0.31*	0.16	-0.22
55-64	-0.37**	0.18	-0.22
65+	-0.43**	0.19	-0.23
Age at marriage [3]			
18-24	0.03	0.06	0.03
25-29	-0.03	0.06	-0.02
>30	0.05	0.09	0.02
Income Level [4]			
Income Less than 400	-0.05	0.07	-0.03
401 - 600 TL	-0.19***	0.06	-0.15
601 - 800 TL	-0.14**	0.07	-0.09
801 - 1 200 TL	-0.27***	0.07	-0.17
1 201 - 2 500 TL	-0.29***	0.09	-0.13
More than 2 501 TL	-0.28**	0.14	-0.06
Disturbance Covariance (Female Autonomy, Violence)	0.94***	0.10	

Notes: N=6435, RMSEA=0.015; CFI=0.984. Sample weights are used. Base categories: (1) no formal education, (2) age 18-24, (3) marriage before the age of 18, (4) no income. *** p<0.01, ** p<0.05, * p<0.1

Table 3.7: Estimation Results of the Measurement Model of Wife's Social Norms

	Coefficient	Std. Err	Standardised coefficient	R-square
Latent Variable: Wife's Social Norm (SN)				
Having a male child increases the prestige of the wife	1.00	Na	0.66	0.44
Not doing housework is a sole reason for divorce	0.35***	0.05	0.24	0.06
The best age of marriage for women is between 15-19	0.82***	0.07	0.56	0.31
Continuation of the family is guaranteed only by a son	0.92***	0.06	0.62	0.38
It is appropriate for close relatives (first cousins) to marry	0.62***	0.06	0.43	0.18
It is not appropriate for women to work	0.87***	0.07	0.60	0.35

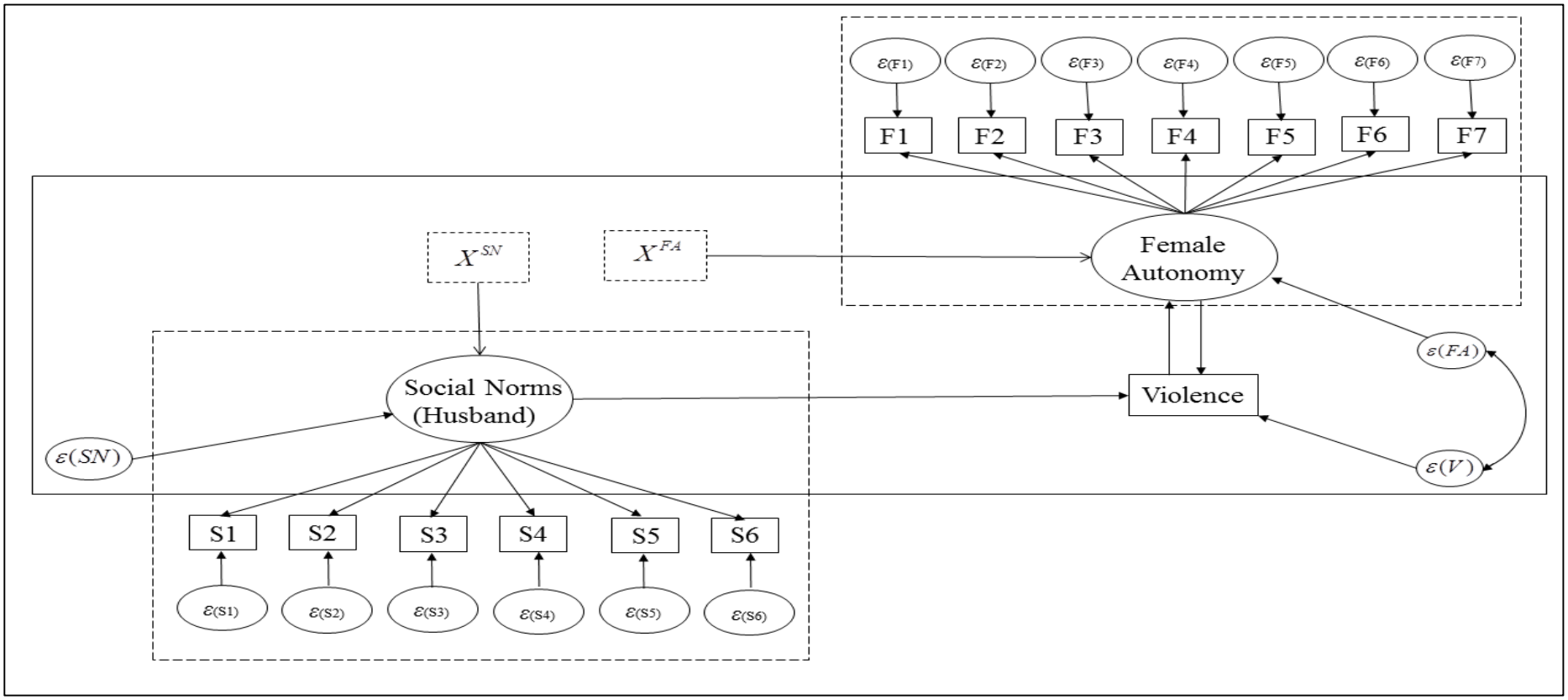
Notes: Sample weights are used. *** p<0.01, ** p<0.05, * p<0.1

Table 3.8: Estimation Results of the Structural Model of Alternative Specification

	Coefficient	Std. Err	Standardised coefficient
Dependent Variable: Violence			
Female Autonomy	-0. 84***	0.14	-0.71
Husband's Social Norms	0. 14*	0.08	0.08
Dependent Variable: Female Autonomy			
Violence	-0.77**	0.16	-0.91
Wife's Social Norms	-0.10**	0.04	-0.08
Education level			
Primary	0.02	0.01	0.01
Secondary	0.01	0.02	0.01
High school or above	0.04	0.03	0.02
Age			
25-34	-0.02	0.02	-0.01
35-44	0.01	0.03	0.01
45-54	0.02	0.04	0.01
55-64	0.01	0.04	0.00
65+	0.01	0.05	0.00
Age at Marriage			
18-24	0.00	0.01	0.00
25-29	-0.01	0.02	-0.01
>30	0.01	0.05	0.00
Economic status and family variables			
Having income	0.03	0.02	0.01
Wife's asset ownership	0.02	0.02	0.01
Arranged marriage	-0.04*	0.02	-0.02
Bride price	-0.02	0.02	-0.01
Number of daughters	-0.01	0.01	-0.02
Number of sons	-0.01	0.01	-0.01
Dependent Variable: Wife's Social Norms			
Education level			
Primary	-0.27***	0.04	-0.18
Secondary	-0.42***	0.07	-0.14
High school or above	-0.68***	0.07	-0.35
Age			
25-34	-0.19***	0.07	-0.12
35-44	-0.25***	0.09	-0.15
45-54	-0.17*	0.10	-0.10
55-64	-0.17	0.12	-0.10
65+	-0.20	0.13	-0.07
Age at marriage			
18-24	0.02	0.03	0.03
25-29	0.09	0.07	0.03
>30	0.09	0.15	0.01
Economic status and family variables			
Having income	-0.13***	0.04	-0.07
Wife's asset ownership	-0.01	0.04	-0.01
Arranged marriage	0.14***	0.03	-0.09
Bride price	0.17***	0.04	-0.09
Number of daughters	0.07***	0.01	-0.12
Number of sons	0.14***	0.01	-0.21
Disturbance Covariance (Female Autonomy, Violence)	0.94***	0.09	

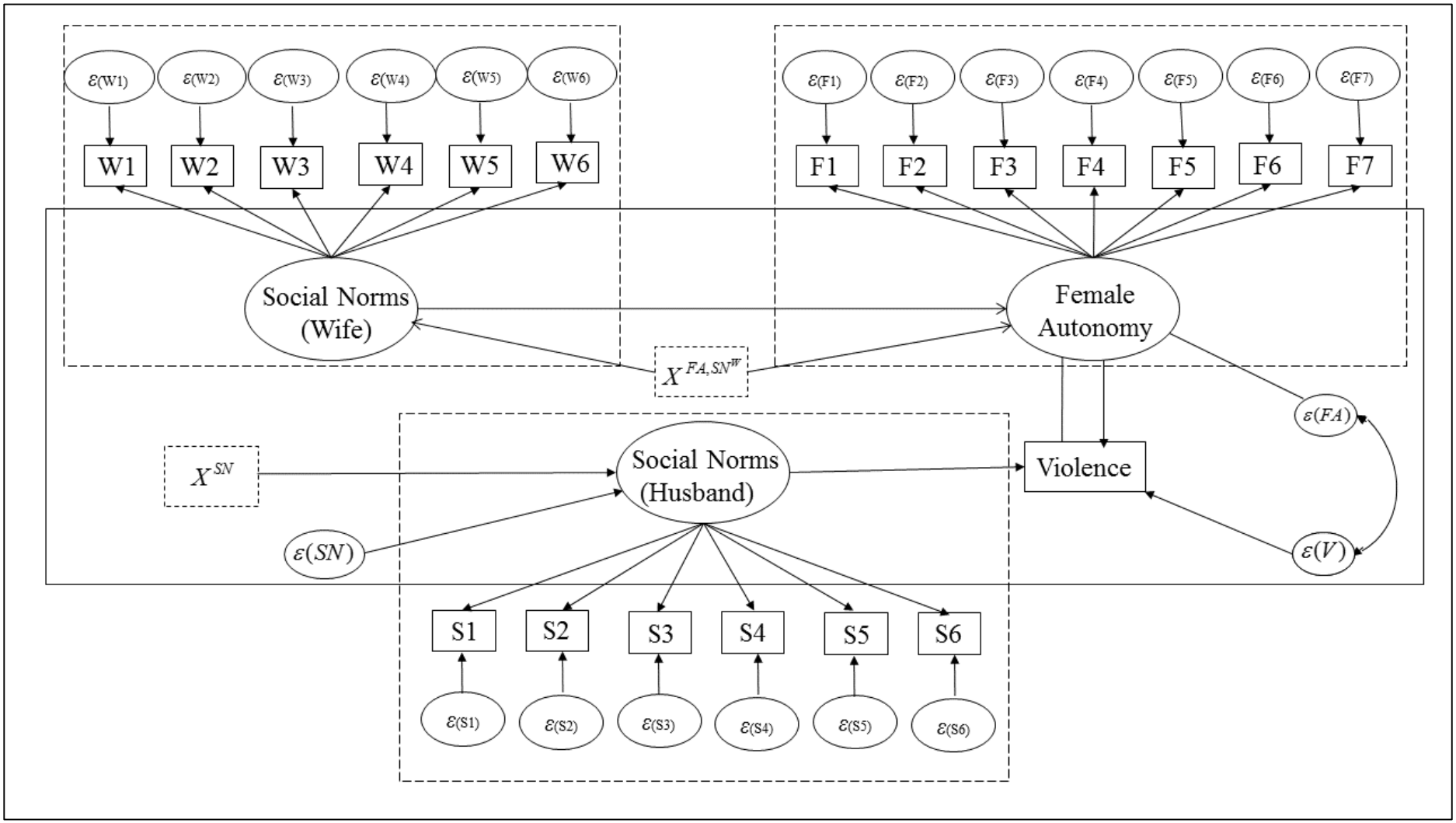
Notes: N=6435, RMSEA=0.03; CFI=0.91. Base categories are the same as Table 3.6. Sample weights are used.
***p<0.01,**p<0.05,*p<0.1

Figure 3.1: Path Diagram of the Model



Notes: The measurement models are the dashed rectangles; the structural model is the solid rectangle. Latent variables are shown as ellipses. The exogenous observed variables (X's) are presented by dashed rectangles and the endogenous observed variable is represented by a solid rectangle. The single headed arrows represent the hypothesized direct effect of one variable on another. Two headed arrows represent the covariance between the variables. ε(.)'s are the error terms corresponding the relevant variables and in the measurement models, they represent measurement errors in the indicator variables.

Figure 3.2: Path Diagram of the Alternative Model Specification



3.9 APPENDIX TO CHAPTER 3

Table 3.A1: Main reasons why women are not currently participating in the labour force (Turkey)

Reason	Per cent	N
Takes care of children or household	29.74	1616
Husband or elder does not want	24.77	1346
No need for working	11.12	604
Sick or handicapped	11.67	634
No talent/education	4.56	248
Continues education	6.73	343
Other	11.41	643
Total	100	5434

Source: Based on author's calculations using Demographic and Health Survey 2005 datasets.

Table 3.A2: Exclusion criteria for the sample

	N
Full sample	48235
Dropped observations	
Single, divorced, widowed	30367
Individuals who are not head of household or his/her spouse	1446
Individuals how or whose partner did not respond the survey	3552
Resulting sample	12,870

CHAPTER 4: HETEROGENEITY IN INFORMAL SECTOR EMPLOYMENT IN THE TURKISH LABOUR MARKET

4.1 Introduction

One of the common features of labour markets in developing countries is the considerable size of the informal sector employment. According to the Global Employment Trends Report (ILO, 2013), between 40-50 per cent workers in countries in the Latin American and Caribbean region are employed in the informal sector, while this share is between 15-30 per cent in Central and South-Eastern Europe (non-EU) countries. Beside its substantial volume, it is also observed that the informal sector is persistent in the developing world. According to the same report, between 2000 and 2010, the share of workers in informal employment declined in only 26 out of 49 countries for which estimates are available.

Traditionally, the informal sector has been seen as inferior compared to the formal sector in terms of wages, security and protection from exploitation regarding labour standards. For instance, as the informal sector is free of regulation, wage offers can be below the minimum wage level, the standards regarding age, length and time of working hours can be ignored, and workers may have to work in risky environments with unsafe equipment. Moreover, employment benefits such as social security and compensation do not apply to informal employment due to its unrecorded nature (Losby et al., 2002).

There are several questions that need to be addressed in order to understand the nature of the informal sector and to provide meaningful policy prescriptions. Such questions include: What determines the sector of employment? Is the sector of employment a matter of choice or is it a result of entry barriers between the sectors? If workers in the

informal sector had the same characteristics as those in the formal sector, would they earn the same wages?

The traditional segmented market theory asserts that the informal sector is nothing but a safety net for those excluded from the formal economy (Ferman et al., 1987). According to Stiglitz (1976), and Dickens and Lang (1985), involvement in the informal sector is not a voluntary decision, instead it is a survival strategy for people who are rationed out from the overly regulated formal sector. Due to the entry barriers that exist between these two sectors, informal sector workers earn less than those who have the same characteristics but work in the formal sector (Magnac, 1991). The neoclassical comparative market view, on the other hand, argues that informal employment can be the desirable option based on people's rational evaluation of the relative costs and benefits of entering the formal system (De Soto, 1989). In that sense, given their characteristics, people may voluntarily choose informal employment if they think that they have a comparative advantage in the informal sector (Maloney, 1999).

The most recent hypotheses combine these two alternative hypotheses by suggesting that the informal sector has a more complex structure that represents features of both segmentation and competitiveness. Fields (2005) and Maloney (2004) argue that the different segments of the informal sector exhibit different properties. They assert that, in the "upper-tier" segment, or the "voluntary" segment as Maloney (2004) puts it, given their individual characteristics, workers voluntarily choose the informal sector expecting relatively high earnings. On the other hand, the "lower-tier" or the "involuntary" segment is seen as the last resort of employment by the workers who are rationed out from the formal sector.

This chapter investigates the heterogeneity of the informal sector by looking at data from the Turkish labour market in which informal workers constitute more than 40 per cent of total employment. To do so, differences between the informal and formal sector wages are decomposed along the entire distribution. A decomposition procedure, which not only allows correction of the potential self-selection problem in wage estimations, but which also decomposes the informal/formal sector wage gap at any quantile of interest, is used. Using this method, the wage differentials between these two sectors are decomposed into two factors, namely, the characteristic effect and the coefficient effect. The characteristic effect is the share of the wage gap that is explained by the differences between the characteristics of formal and informal sector workers. The coefficient effect is the share of wage gap that is related to the differences in returns to those characteristics. The advantage of this method is that it allows decomposition of wages at any quantile. So it is possible to analyse the heterogeneity of the informal sector, which is not possible with a simple mean-wage decomposition.

In the context of the formal/informal sector wage gap, a relatively strong coefficient effect is an indicator of market segmentation as it shows that workers in the informal sector earn less even after controlling for differences in their characteristics. In contrast, if the characteristic effect is stronger, it means that the wage gap is mainly due to differences between human capital endowments, not because the two sectors have different wage determination processes. In this case, the informal labour market for relatively unskilled workers may be seen as integrated with the formal sector and entry to the informal sector is not an involuntary choice induced by a segmented market (Maloney, 1999; Carneiro and Henley, 2002).

The results of the study suggest a heterogeneous structure for the informal sector in Turkey in which the lower tier consists of workers who involuntarily participate in the

informal sector and the higher tier includes the workers who voluntarily choose to be in the informal sector. It is also found that the wage gap between the formal and the informal sectors is higher and the segmentation at the bottom end of the wage distribution is more spread out (i.e. observed in a wider range of quantiles) for females than it is for males.

The rest of the chapter is structured as follows: Section 2 provides a literature review and discusses the empirical evidence from other developing countries as well as Turkey. Section 3 outlines the estimation method and section 4 describes the data used in this chapter. Section 5 presents the wage regression estimations and decomposition results. Section 6 concludes.

4.2 Literature review

The substantial size of informal employment in developing economies and its persistence has drawn a lot of attention from academics and policy-makers. Accordingly, a considerable amount of empirical evidence about the structure of informal employment is provided in the existing literature.[39]

Several studies have documented the evidence that supports the segmented market theory. Knight and Song (2005) argue that the Chinese labour market consist of two segments: the urban-born workers employed in the state and collective formal sector, and the rural workers and rural to urban migrants employed in the informal sector. They assert that this segmentation is maintained by the residence restriction which limits the rights of the migrants in the cities providing protection for the urban workers. Hofmeyr (2002) analyses the labour market in South Africa and finds evidence of a segmented market. He asserts that the segmentation in South Africa is due to powerful trade

[39] See Ruffer and Knight (2007) for a survey of the empirical literature.

unions; while the unionised formal segment has been able to protect its wages, in the non-unionised segment wages have declined with the fast growing workforce and slow creation of employment. Gindling (1991), accounting for the potential selection bias due to the non-random selection into formal and informal sector employment, examines the labour market in Costa Rica. His results also suggest segmentation between formal and informal workers in the private sector.

There is also a body of literature that supports the competitive market theory. For instance, Pratap and Quintin (2006) test segmentation in the Argentinian labour market. Using standard OLS estimation, they find an earnings premium for the formal sector workers. However, once they use a propensity score matching technique to address self-selection into sectors, the earning premium disappears. Magnac (1991) uses multivariate probit and tobit techniques in an extended Roy model framework to analyse the Colombian labour market. His results suggest a competitive structure for the Colombian labour market. Carneiro and Henley (2002) used selection corrected wage equations to examine the informal sector in Brazil and concluded that, consistent with selection into the informal sector being a rational choice, informal sector workers appear to have a comparative earnings advantage in their selected sector.

Most of the empirical literature that tests the segmented or competitive market theories, some of which is summarised above, assumes a homogeneous informal sector. Accordingly, these studies mainly focus on the wage differences between the mean wages of the formal and informal sectors. However, as noted previously, most recent theories suggest a more complex structure for the informal sector (Maloney, 2004; Fields, 2005). According to these arguments, upper tier workers can voluntarily choose informal employment because, given their characteristics, they may think they are better off in the informal sector. On the other hand, lower tier workers may involuntarily end

up in the informal sector to avoid unemployment. Recent empirical evidence suggests that voluntary and involuntary employment can co-exist in the informal sector.

Bargain and Kwenda (2010) analyse the labour markets in Brazil, Mexico and South Africa by using fixed effect quantile regressions. For South Africa, their results show a substantial earnings penalty for informal sector workers along the wage distribution, suggesting segmentation. On the other hand, in Brazil and particularly in Mexico, the informal sector is found to be heterogeneous as the upper tier segment of self-employment is found to be the most rewarding type of employment in terms of wages, dominating both formal and informal waged employment. Tannuri-Pianto and Pianto (2002) use selection corrected quantile regressions to analyse the Brazilian labour market.[40] They find that, while the earnings gap at the high quantiles is completely explained by the differences in characteristics, at low quantiles the differences in returns to characteristics explain a considerable share of the gap. Accordingly, they conclude that informal employment is voluntary at the top end and involuntary at the bottom end of the earnings distribution. Using a finite mixture model, Gunther and Launov (2006) analyse the informal sector in the Ivory Coast. Their results show the existence of both segmented and competitive employment in the informal sector.

4.2.1 Review of the literature on Turkey

This chapter focuses on the Turkish labour market in which informal employment constitutes a substantial share of the total employment. The size of informal employment in Turkey has been quite high compared to the neighbouring European and Middle Eastern countries (ILO, 2013). Although there has been a decreasing trend between 2000 and 2010, the share of informal sector workers in total employment in

[40] This chapter uses a methodology similar to the Tannuri-Pianto and Pianto (2002). The details of the methodology will be explained in Section 4.

Turkey is still around 40 per cent, which is close to the values for the Latin American countries where the highest informality rates in the world are observed. The first empirical analyses on the informal sector in the Turkish labour market comes from Tansel's (1999, 2000) work. Tansel (1999), using the 1994 Turkish Household Expenditure Survey, analyses earning gaps between the formal and informal sectors by estimating selection corrected wage equations for each group. Her results show a substantial difference between the wages of these groups for both men and women. Tansel (2000), then, extends her previous work by including the self-employed in the analysis. Using a similar methodology, she finds that while male formal sector wage workers earn more than their informal sector and self-employed counterparts, the wages of women are similar across sectors.

Taymaz (2009), using the Labour Force Survey (LFS) for the year 2006, estimates Mincerian wage equations and accounts for selection bias using a multinomial logit model. In the multinomial model for selection, six types of jobs available for an individual are assumed: formal sector jobs in manufacturing, formal sector jobs in services, informal sector jobs in manufacturing, informal sector jobs in services, entrepreneur (employer) and self-employed. If an individual stays at home they are classified as economically inactive. Supporting the assertions of segmented market theory, his results show that, with the exception of female workers in the manufacturing sector, there is a significant wage gap between informal and formal sector workers even after controlling for selection and individual characteristics.

Baskaya and Hulagu (2011) use five waves of the LFS for 2005-2009 and analyse the differences in wages of formal and informal sector workers. They first estimate a simple Mincerian wage regression including a dummy variable for the formal sector. Then, they go beyond simple mean wage estimation and use a propensity score matching

technique to control for the potential selection bias due to self-selection into the formal or the informal sector. Accordingly, they estimate a probit model for selection in which the dependent variable is a dummy variable that takes value 1 to represent formal sector employment, and 0 otherwise. Their results show that, even though both models provide evidence that there is a wage gap between these sectors, propensity score matching models suggest that the wage gap between the formal and informal sectors is higher for females than it is for males.

Tansel and Kan (2012) use data from the longitudinal Income and Living Conditions Survey for the period of 2006-2009 to analyse the structure of informal employment in Turkey. They estimate both OLS and quantile wage equations to examine the informal/formal wage gap at the mean and along on the entire earnings distribution for both men and women separately. In their analysis, they include formal salaried, informal salaried, formal self-employed and informal self-employed workers. They also estimate a fixed effect model in order to help to overcome the potential self-selection into formal and informal sector. Their results suggest that there is an overall wage penalty for informal sector workers. However, this penalty disappears at the top end of the wage distribution, suggesting heterogeneity in the structure of the informal sector. Although they account for the selection bias in the fixed effects estimation, their quantile regression analysis still has the potential selection bias problem due to self-selection into formal and informal sector employment.

This study contributes to the previous literature in two ways: First, similar to Tansel and Kan (2012), it analyses the informal/formal sector wage gap using quantile regression analyses. However, this study corrects for the potential selection bias, resulting from non-random selection into the formal and the informal sector, by estimating a semi-parametric selection equation (Buchinsky, 1998). Second, unlike the previous studies

which focus on the value of the coefficient of the informal sector dummies in wage regressions, this study decomposes the differences in the wage distributions between the two sectors into the effects of differences in characteristics and the differences in returns to those characteristics. Finally, the study examines the relative shares of these effects at different points of the wage distribution to understand whether the Turkish labour market has a competitive or a segmented structure.

4.3 Methodology

The aim of this chapter is to examine the wage gap between the formal and the informal sector workers. To do so, the quantile decomposition methodology is used. This methodology allows the decomposition of wage differentials between the two groups across all quantiles of the wage distributions, not just the difference between the mean wages of each group. Additionally, unlike Juhn et al. (1993) and traditional Oaxaca (1973) and Blinder (1973) decompositions which assume independent and normally distributed error terms, the quantile regression method takes the heterogeneity of workers along the distribution into account. A potential selection bias problem in wage estimations that may result due to non-random selection into the formal or the informal sector is also considered.[41] In order to overcome this problem, a procedure that combines the selection corrected quantile regression method of Buchinsky (1998) and the quantile decomposition method of Melly (2006) is exploited.[42]

[41] A potential selection bias can also occur due to self-selection into employment. However, the analysis in this chapter does not control for this type of self-selection due to the nature of semi-parametric estimation of selection equation in which the dependent variable has to be a binary variable. The details of the semi-parametric estimation is explained below.

[42] The idea of combining the Buchinsky (1998) method and the quantile decomposition technique was originally proposed by Albrecht et al. (2009). However, Albrecht et al. use the Machado and Mata (2006) decomposition method instead of that of Melly (2006). The motivation for the choice of the Melly (2006) method is explained below.

The Melly (2006) decomposition method consists of three steps (the description below drawsheavily from the Melly paper). In the first step, the conditional quantile functions for wages are estimated for each specific quantile, $\theta \in (0,1)$,

$$F_{Y|X}^{-1}(\theta \mid X) = X' \beta(\theta) \qquad (1)$$

where $F_{Y|X}^{-1}(\theta \mid X)$ is the θ^{th} quantile of the log wage distribution Y conditional on a set of covariates X, and $\beta(\theta)$ is the vector of coefficients to be estimated in the corresponding quantile θ. In the quantile regression setting, the coefficients from the wage regressions represent the returns to the characteristics at the quantiles of interest. Koenker and Bassett (1978) show that $\beta(\theta)$ can be estimated separately for each quantile by,

$$\hat{\beta}(\theta) = \arg\min_{b \in \Re} \frac{1}{N} \sum (Y - Xb)(\theta - 1(Y \leq Xb)) \qquad (2)$$

Where $1(\cdot)$ is the control function. By definition, an infinite number of quantile regressions can be estimated along the wage distribution. However, estimation of the whole quantile regression becomes computationally demanding, and time consuming, as the number of observations increase. Melly (2006) suggests estimating a specific number of quantile regressions which are uniformly distributed over the wage distribution. He noted that the asymptotic results will still be valid if the quantile regression coefficients are estimated along a grid of quantiles. Accordingly, the coefficients that are estimated at a given point $\hat{\beta}(\theta_j)$ are presumed to stay unchanged

from θ_{j-1} to θ_j for j=1,…,J. The outcome of this procedure is a vector $\hat{\beta}$ which contains a finite number of quantile regression coefficients $\hat{\beta}(\theta_1),\ldots,\hat{\beta}(\theta_j),\ldots,\hat{\beta}(\theta_J)$.[43]

In the second step of the method, the conditional quantiles of wages are turned into estimates of unconditional quantiles of wages. This is done by integrating the conditional wage distribution over the whole range of the distribution of covariates included in the quantile regressions. This procedure, replaces each quantile estimate $F_{Y|X}^{-1}(\theta \mid X)$ by its consistent estimate $X_i\hat{\beta}(\theta_j)$. Formally, the θ^{th} quantile of wages, $\hat{Q}(\hat{\beta}, X)$, can be estimated by

$$\hat{Q}(\hat{\beta}, X) = \inf\left\{Q : \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{J}(\theta_j - \theta_{j-1})1(X_i\hat{\beta}(\theta_j) \le Q) \ge \theta\right\} \quad (3)$$

The first two steps of the method are applied separately to the wages of both the formal and the informal sector workers which provides $\hat{Q}_\theta(X_{for}, \hat{\beta}^{for}(\theta))$ and $\hat{Q}_\theta(X_{\inf}, \hat{\beta}^{\inf}(\theta))$ respectively. The final stage of the method is to decompose the wage gap between the formal and the informal sector into the characteristics effect and the coefficient effect. To do so, one needs to simulate the whole counterfactual distribution. The counterfactual distribution represents the distribution that would have prevailed in the informal sector if the returns to the characteristics in the informal sector were equal to the returns in the formal sector. The counterfactual wages $\hat{Q}_\theta(X'_{\inf}\hat{\beta}^{for}(\theta))$ are created by re-estimating equation (3) using the covariates of the informal sector workers and estimated coefficients from the formal sector.

[43] When deciding on the number of quantiles, at which to estimate the coefficients, there is a trade-off between computation time and precision. The estimation of conditional distribution will be more precise, the higher the number of quantile regressions estimated. In this chapter, initial results were estimated with 10 quantile regressions. Then, considering the precision issue, 100 quantile regressions from θ=0.01 to θ=0.99 were estimated. The results that are presented in the chapter are from 100 quantile regressions. Also, 1000 bootstrap iterations were performed to obtain the standard errors. The computation time increases with both the number of observations and the number of bootstrap iterations.

Once the counterfactual wage distribution is generated, the wage gap between the two sectors at quantile θ can be decomposed into two effects such that:

$$\begin{aligned}\hat{Q}_\theta(X_{for},\hat{\beta}^{for}(\theta))-\hat{Q}_\theta(X_{\inf},\hat{\beta}^{\inf}(\theta)) = & \left[\hat{Q}_\theta(X_{for},\hat{\beta}^{for}(\theta))-\hat{Q}_\theta(X_{\inf},\hat{\beta}^{for}(\theta))\right] \\ & +\left[\hat{Q}_\theta(X_{\inf},\hat{\beta}^{for}(\theta))-\hat{Q}_\theta(X_{\inf},\hat{\beta}^{\inf}(\theta))\right]\end{aligned} \tag{4}$$

The first term on the right hand side of equation (4) gives the effect of the characteristic at quantile θ. It is basically the gap between the wage distribution in the formal sector and the counterfactual wage distribution, i.e. the wage distribution that would have prevailed in the informal sector if the returns to characteristics had been the same as in the informal sector. The second term on the right hand side of equation (4) captures the effect of the coefficients and it is the difference between the counterfactual wage distribution and the wage distribution in the informal sector.

The Melly (2006) procedure has two important advantages over the Machado and Mata (2005) procedure. First, as Albrecht et al. (2003) note, by estimating coefficients for a grid of θ's, it eliminates the sampling error that is inherent in the Machado and Mata approach which samples θ's from a standard uniform distribution. Second, it reduces the computation time significantly as, unlike the Machado and Mata procedure, it does not rely on simulations. Melly (2006) also shows that if the number of simulations in the Machado-Mata procedure goes to infinity the procedures are numerically identical.

If selection into the informal sector was random, the procedure that is summarised so far would be appropriate to decompose the wage gap between the two sectors. However, it has been shown in many applications, such as Marcouiller et al. (1997) and Tannuri-Pianto and Pianto (2002), that there is a selection bias issue arising from the

self-selection into the formal or the informal sector. To obtain unbiased estimators in wage regressions, Buchinksy (1998) suggests adding a correction term to the quantile wage equations.[44]

In order to construct the correction term, *d*, an indicator variable for each worker *i* is defined such that

$$d_i \equiv \begin{cases} 1 & : \text{ informal} \\ 0 & : \text{ formal} \end{cases} \tag{4}$$

Further, an index variable, g, is assumed which is an underlying continuous response variable behind the indicator variable *d*:

$$d_i = \begin{cases} 1 : & g_i > 0 \\ 0 : & g_i \leq 0. \end{cases} \tag{5}$$

The underlying response variable can be interpreted as the utility that individual *i* gets when he or she chooses to be in the informal sector. If the utility of being in the informal sector for a worker is greater than the level of utility obtained from working in the formal sector (without losing generality, this level can be assumed to be zero), then the worker choses to be in the informal sector, and the formal sector otherwise. The equation for the index variable is given as:

$$g_i = z_i'\gamma + \upsilon_i, \tag{6}$$

Where z is the realisation of *Z*, the stochastic vector of characteristics that determine the probability of worker *i* being in the informal sector; γ is the vector of coefficients to be estimated; and υ is the vector of error. Similar to the standard Heckman procedure, Z must contain one variable that is not included in X, the vector of characteristics in the

[44] The idea of adding a selection term to the quantile wage regression is similar to the Heckman selection procedure. The difference of this methodology is in the way the selection term is estimated, which will be explained below.

wage equation.[45] In the standard Heckman procedure, equation (6) is estimated by using a probit model in which errors are assumed to be normally distributed. While such a methodology fits well with mean wage gap analyses, decomposing wage gaps at different quantiles along the earnings distribution demands an alternative methodology which does not require normality assumption for the distribution of errors in the selection equation (Buchinsky, 1998). Therefore, as a first step, the coefficient vector γ in equation (6) is estimated by the semi-parametric least squares (SLS hereafter) method of Ichimura (1993):

$$\hat{\gamma} \equiv \arg\min_{\gamma \in \Re} \frac{1}{n} \sum_{i=1}^{n} \left(d_i - \hat{E}(d_i \mid z_i, \gamma) \right)^2, \tag{7}$$

where

$$\hat{E}(d_i \mid z_i, \gamma) = \frac{\sum_{j \neq i} d_j \kappa((z_i^{'}, \gamma - z_j^{'}, \gamma)/h)}{\sum_{j \neq i} \kappa((z_i^{'}, \gamma - z_j^{'}, \gamma)/h)} \tag{8}$$

is the nonparametric kernel density estimate of d_i, with kernel density function κ(·) and its bandwidth *h*.[46]

The probability of participating in the informal sector is thus given by

$$P(d_i = 1) = \Psi(z_i^{'}\gamma) \tag{9}$$

where $z_i^{'}\gamma$ is the single index with an unknown functional form Ψ. In the second step, the single index (as a correction term) is introduced in the wage equations such that

$$F^{-1}_{Y_{for}|Z}(\theta \mid Z) = x^{'}_{for}\beta^{for}(\theta) + \Psi_{\theta}(z^{'}\gamma) \tag{10}$$

and

[45] As noted in the second chapter of the thesis, the coefficients in the wage equations can be identified from the non-linearity of the selection term. However, to avoid potential multi-collinearity problems, an exclusion restriction is applied to the analysis in this chapter. One additional restriction of the single index estimation is that at least one of the excluded variables has to be continuous. The choice of excluded variables is discussed in the next section.

[46] The single index estimation is not a part of the bootstrap iterations that are performed for the quantile estimations. See Buchinsky (1998) for statistical properties of the estimator in detail.

$$F^{-1}_{Y_{\inf}|Z}(\theta \mid Z\ ) = x^{'}_{\inf}\beta^{\inf}(\theta) + \Psi_{\theta}(z^{'}\gamma)\,. \qquad (11)$$

The form of single index, Ψ, is unknown. Therefore, the true value of the term $\Psi_{\theta}(z^{'}\gamma)$ is not known. To solve this problem, Buchinsky (1998) suggests a power series approximation to the term by using the predicted values of index $z^{'}\hat{\gamma}$, from equation (6). With the inclusion of the power series approximation, equations (10) and (11) can be rewritten as

$$F^{-1}_{Y_{for}|Z}(\theta \mid Z\ ) = x^{'}_{for}\beta^{for}(\theta) + \sum_{k=1}^{K}(\lambda(\mu + \sigma z^{'}\hat{\gamma}))^{k-1}\delta_k(\theta) \qquad (12)$$

$$F^{-1}_{Y_{\inf}|Z}(\theta \mid Z\ ) = x^{'}_{\inf}\beta^{\inf}(\theta) + \sum_{k=1}^{K}(\lambda(\mu + \sigma z^{'}\hat{\gamma}))^{k-1}\delta_k(\theta)\,. \qquad (13)$$

where μ and σ are the location and scaling parameters that are obtained from the probit regression of indicator variable d on the predicted values of index, $z^{'}\hat{\gamma}$; where $\lambda(\cdot) = \dfrac{\phi(\cdot)}{\Phi(\cdot)}$ denotes the usual (Heckman's) inverse Mill's ratio; K is the number of power series expansions[47]; and $\delta_k(\theta)$ is the coefficient of the kth power series at the θth quantile in the augmented wage regression. The advantage of using the single index method compared to the traditional Heckman (1979) approach is that it does not assume normality for the distribution of the errors. Moreover, it is more quantile specific as δ's vary with each value of θ. As Buchinsky notes, the SLS estimate is consistent, and independent of the distribution of errors, while the probit estimate is the efficient estimate under normally distributed errors.

In equations (12) and (13), when k=1, the term whose coefficient is $\hat{\delta}_1(\theta)$ will be equal to one. Therefore, it is not possible to identify the first series expansion separately from

[47] K=3 is chosen for the power expansion as a collinearity problem occurred with the inclusion of higher power series.

the constant term in the wage regression. To identify the constant term for the wage regressions, the identification at infinity method proposed by Andrews and Schafgans (1998) is applied. Accordingly, a sub-sample of workers whose probability of working in the informal sector is arbitrarily close to one is chosen and this sub-sample is used to estimate the constant term without adjusting for selection.

4.4 Data and descriptive statistics

The data used in the analysis is from the LFS which is conducted by TurkStat; the estimations are performed for women and men separately to determine the level of the wage gap between the formal and the informal sectors for each gender. Also, the analyses are repeated with both 2002 and 2010 data to explore the change in the wage gap between these two years. The first LFS was conducted in year the 2000. However, to avoid the distortions in the Turkish labour market due to the economic crises which occurred in 2000 and 2001, the 2002 LFS data is used in the analysis as the starting period. At the time of the analysis the last wave of LFS that was available was 2010, which is used here. The LFS has the largest sample of the Turkish labour force and contains information on both the workers' demographic characteristics and the characteristics of their main job; the information is available for each individual in the household. For the analysis, two cross sections of employees aged between 20 and 64 are used.[48]

According to the International Labour Organisation (ILO) classification, there are two definitions of labour market informality. The first definition, which is referred to as the productivity view, focuses on informality by looking at the characteristics of the enterprises, such as the firm size. In this approach, informal employment is captured by

[48] The reason of choosing this age band is that decision to pursue higher education is endogenous as it depends on the returns to higher education. However, individuals who are at the age of 20 or older are likely to have already made their educational decisions.

employment in the small scale firms which are not registered under specific norms of national legislation. The second definition, which is referred to as the legalistic view or the protectionist view, defines informal employees as the individuals who are employed by the formal or the informal enterprises but are not subject to the national labour legislation, income taxation, social protection or entitlement to certain employment benefits such as advance notice of dismissal and severance pay (for details, see Hussmanns, 2004).

In this chapter to capture the informality of employment, the protectionist definition of informality is used, i.e. the lack of social security protection. In the studies on informality in the Turkish labour market, this is the definition that is commonly chosen. Taymaz (2009) notes that this choice can be justified by two reasons: first, the definition of informality based on the firm size is arbitrary in practice. Second, most of the enterprises in Turkey are registered but they report their business revenue and employment only partially. The LFS contains information on membership of the social security system. Accordingly, using this information a dummy variable is created that takes the value of 1 if the individual is not registered for the social security system and hence classified as in the informal sector, and 0 otherwise.

All employers in Turkey are legally obligated to register their employees for the social security system. The membership of the social security system has implications for both employers and employees. For instance, several deductions are made from the earnings of the registered employees such as, the income tax, stamp tax, contribution to the social security system and to the unemployment benefit fund. A study on the Turkish labour market by Korkmaz (2001) shows that around 26 per cent of the gross earnings of a minimum wage worker is deducted for the taxes and contributions. The employers who register their employees are also subject to contributions to the social security

system and to the unemployment benefit fund. Moreover, the employers are obligated to pay severance pay and maintain certain standards (such as providing health and safety services) which increases with the number of employees.

The LFS contains information on self-reported after-tax monthly wages. Therefore, it does not require any adjustments for taxation. In order to get hourly wages, monthly wages are first divided by 4.3 and then divided by usual hours of work per week reported by the interviewee. Finally, hourly wages are deflated by the consumer price index which is also provided by the TurkStat to obtain real hourly wages.[49] Analyses are performed with 31,596 (34,946) male workers, and 9,170 (18,598) female workers for the year 2002 (2010). The actual size of the 2010 male sample after the exclusion criteria noted above was 58,243. However, the computation of the semi-parametric estimation becomes very time consuming with a large number of observations. Considering this computational restriction, a random 60 per cent from the 2010 male sample (after exclusion criteria) is selected to perform the analysis. The comparison of the full sample and the selected sample is shown in Table 4.1.[50]

Following the work of Mincer (1974), human capital is represented by education and age. In the LFS, the educational qualifications of workers are grouped into six categories, namely, no formal education, primary, secondary, high school, vocational high school and university which correspond to 0, 5, 8, 11, 11, and 15 years of schooling, respectively. However, due to the very small proportions in the first two

[49] The labour market status of the individuals in the LFS is categorised into 4 categories: employee, self-employed, employer, and unpaid family worker. As the dataset does not provide their wages, employer, unpaid family workers and self-employed individuals are dropped from the sample. The share of the self-employed in total employment is around 25 per cent for men and 13 per cent for women in both 2002 and 2010 datasets. Also the individuals who hold more than one job are dropped from the sample due to the lack of information about their earnings from the second job. The exclusion criteria of the sample can be found in Table 4.A1 in the appendix.

[50] With the randomly selected 60 per cent sample it took around a week to estimate the selection equation.

categories at the higher quantiles of the wage distribution of formal sector workers, a dummy variable is created to represent workers with less than secondary school education by combining the first two categories of no formal education and primary education. Tenure is obtained by using the question of "year that you started your latest job/occupation". Control variables such as, age, marital status, presence of children in the household, rural residence and being the head of the household are also included to the wage regressions as explanatory variables.

In the selection equation, education, marital status, age, firm size, and the number of children are included as determinants of the informal sector employment. As noted in the previous section, to identify the coefficients in the wage regressions, exclusion restrictions are applied in the analysis. Two variables namely, presence of a formal sector employee in the household and mean years of education of the household adults (excluding the respondent), are included in the selection equation. The theoretical justification for the use of the presence of a formal sector worker in the household variable comes from the literature on social networks. As summarised in the second chapter, Calvo-Armengol and Jackson (2004) argue that each individual is connected to others through a network (e.g., kinship, friendship, and acquaintanceship). He also argues that, when information about jobs arrives, individuals, who are unemployed and directly hear of a job, use the information to obtain a job. On the other hand, individuals who are already employed, depending on whether the job is more attractive than their current job, might take the job or else might pass information to one (or more) of their direct connections in the network. In that sense, it is possible that a formal sector worker may share the information about a formal sector employment opportunity with his or her household members who are employed in the informal sector.

As noted above, Korkmaz (2001) shows that a great amount of the earnings of a formal sector worker is deducted from her or his wages for taxes and for contributions to the social security system. In that sense, the household income can be argued to affect the decision of participating in the formal sector. For instance, if the total household income level is high, an individual may choose to work in the formal sector in order to enjoy the benefits of membership to social security system. On the other hand, if the total household income level is low, an individual may decide to stay in the informal sector to avoid the substantial deductions from his or her wages. Unfortunately, the LFS does not provide any information about the earnings of employers, self-employed. Therefore, it is not possible to calculate and include household income in the selection equation. In order to solve this problem, mean years of education of the household adults (excluding the respondent) is used as a proxy for the household income.

Table 4.1 provides the mean values of the variables that are included in the wage and selection equations. According to the table, in 2002, 25 per cent of male and female workers were employed in the informal sector. By 2010, the size of the informal sector reduced by 6 percentage points for both genders. Similar to the findings of Tansel and Kan (2012) who used Income and Living Conditions Survey for the year 2006, descriptive analysis show that, on average, regardless of gender, formal sector workers earn around three times more than the informal sector workers.[51] However, this observation on its own is not strong evidence that the labour market in Turkey has a segmented structure. In order to understand the structure of the informal sector, one has to consider the wage gap at the different parts of the wage distribution and take into account the characteristics effect of the workers in each sector and the returns to these characteristics.

[51] Baskaya and Hulagu (2011) also find a similar result using the LFS data for the 2005-2009 period.

The gap between the mean wages of the formal and the informal sector workers decreased between 2002 and 2010 for both men and women. A possible explanation for the decreasing wage gap between the two sectors could be the decreasing share of informal employment in total employment between 2002 and 2010. A lower number of workers in the informal sector can decrease the competition for the jobs in the informal sector. This in turn can reduce the downward pressure on the wages of the informal sector workers. Similarly, a larger number of workers in the formal sector will increase the competition for the jobs in the formal sector which can potentially decrease upward pressure on formal sector wages.

While the average wages for women, relative to men, is lower in the informal sector, in the formal sector women on average earn more than men. Also, the mean value of the wages has increased over time for both men and women across the two sectors. As noted in the second chapter of the thesis, Kaya (2010) shows that in the Turkish labour market, on average, women earn 2.5 percentage points more than men. She also argues that the higher average wages for women is a consequence of a composition effect, as most of the female employees have a university degree while male employees mostly have only primary education. Accordingly, a possible explanation for the higher wages for women in the formal sector can be that women who work in the formal sector are more educated than men who work in the formal sector.

Looking at the mean values of the dummies that represent worker characteristics for both years and genders, it can be seen that, formal sector workers are more educated. Around half of the female formal sector workers are university graduates and the majority of female informal sector workers have less than secondary education. Although the share of male university graduates is relatively high in the formal sector, men with less than secondary education still form the highest share of workers in the

formal waged employment. One important observation is that from 2002 to 2010, employees as a whole became more educated. For instance, the share of male workers with a university degree increased by 2 and 6 percentage points in the informal and the formal sector respectively, and similar increases are also observed for females by 2 and 8 percentage points, respectively. This change is most likely to be the result of the rapid increase in the number of universities in the country.[52]

Table 4.1 also shows that, for both genders, the vast majority of workers in each sector are married and the share of married workers in the formal employment is greater than the share of married workers in the informal employment. Due to their very small share, divorced and widowed male workers are dropped from the male dataset and the analysis of the male sample is performed with only married and never married workers (see Table 4.A1 for the share of omitted observations). It can also be seen that, compared to the informal sector, a greater share of male formal sector workers report themselves as the head of household, while for females there is no difference across the two sectors.

The share of rural residents is lower in the formal sector for each gender and year. The mean value of the dummy variable that represents the presence of at least one formal sector worker in the household (excluding the respondent) is higher for the formal sector workers. Also, the mean value of years of education of household adults (excluding the respondent) is higher for the formal sector workers. Finally, as expected, a substantial share of informal sector workers are employed in the firms in which there are less than 10 workers while around half of the formal sector workers are employed in firms that employ more than 50 workers. When the full and the random 60 per cent of

[52] In 1991, there were only 29 universities in Turkey. In 1992, 24 new universities were established in one year. With the "a university in each city" policy and growing private sector participation, the number of universities reached 165 by 2011. As a consequence of these new universities, the number of university students also increased substantially; for instance, from 1994 to 2011, the number of students in formal university education rose from 1.2 million to 3.7 million (TurkStat web database), a more than 300 per cent increase.

the 2010 male sample are compared, it can be seen that the distribution of characteristics in random 60 per cent sample is almost the same as of full sample. Therefore, there is no loss of any information due to the size restriction for this sample.

4.5 Results

4.5.1 Selection Equation

As mentioned in the previous section, the semi-parametric single index estimation methodology of Ichimura (1993) is used in the selection equation. The advantage of this methodology is that it does not require the normality assumptions for residuals. As Buchinsky (1998) suggests, a Hausman test under the null hypothesis of normally distributed residuals is performed. If the test result fails to reject the null hypothesis, one should use the probit estimates to create the inverse Mill's ratio. The results of the Hausman test presented in Table 4.2 reject the null hypothesis at the 1 per cent significance level. Therefore, the single index estimation results are used in the analysis.

In order to create the selection term to be used in the wage regressions, the selection equation should include at least one variable that is not included in the wage equation. Therefore, the mean years of education of the household adults and a dummy variable that represents the presence of a formal sector worker in the household are used for identification.[53] Education, marital status, age, presence of children, firm size and a dummy representing if the individual is the head of household are also included in the selection equation.

[53] Similar to the instrumental variables case, the literature on selection models has slowly advanced to develop tests for the validity of the exclusion variables. For example, very recently, Huber and Mellace (2014) provided a test for the assumption that exclusion variables are not related to the error term in the main equation of the selection model. However, this test can only be used if the excluded variable is a discrete variable. It is not possible to perform this test in the analysis done here because, as Buchinsky (1998) notes, at least one of the excluded variables has to be continuous.

According to Table 4.2, education has a negative and significant effect on the probability of being in the informal sector. Also the education of other household adults has a negative effect. The presence of a household member who is employed in the formal sector decreases the probability of being in the informal sector. Marriage has opposite effects for males and females; a negative coefficient for males and a positive coefficient for females. So, married women are more likely to work in the informal sector, whereas married men are more likely to work in the formal sector. Being the head of household has a negative effect for both genders. This is not a surprising result if it is assumed that the majority of heads of household are also the main income providers for the household. In the Turkish social security system the spouse and children (daughters until marriage, and sons until the age of 18, or 25 if they continue higher education) benefit from the employee's social security coverage. Therefore, the head of household may prefer to work (if possible) in the formal sector to be able to provide free health coverage for the household members. Presence of children, especially young ones, in the household increases the probability of being in the informal sector. This may be due to preference towards flexible working hours in the informal sector in order to be able to take care of children.

The coefficients of the dummy variables that represent the firm size categories are negative as expected. It is reasonable to expect bigger firms to be in the formal sector. Accordingly, workers who are employed in the bigger firms are more likely to be registered with the social security system. Also, registration with the social security system is legally mandatory for all firms and the bigger firms are more likely to be detected if they do not register their employees in the system.

Looking at the coefficients of the age categories, it is observed that, compared to the base age group 20-25, the probability of being in the informal sector is higher for the

oldest age group of 45 and above. On the other hand, the age groups 25-35 and 35-45 are less likely to be in the informal sector. Based on these findings, one can argue that there is a U-shaped relationship between age and the probability of being in the informal sector in which young and old people are more likely to be in the informal sector compared to the middle age groups. Similar results are found for the Turkish labour market by Taymaz (2009) who argued that workers tend to move to the formal sector over time, but after a certain age (the late 40s and early 50s), they tend to have informal jobs possible for their secondary activities. A possible explanation for this observation is that older workers who receive retirement earnings from their previous formal employment may choose to work in informal sector after their retirement to keep their retirement earnings secure.[54]

4.5.2 Wage Regressions

Female sample

Table 4.3A and Table 4.3B show the wage regression results for the 2002 female sample with and without the selection correction respectively.[55] Looking at the results for the formal sector; education and tenure have a positive effect on the wages of women. The effect of education increases towards the high end of the wage distribution. On the other hand, the returns to tenure are relatively high at the lower quantiles. The differences between the results with and without selection correction are that the magnitudes of the education and tenure coefficients (i.e. returns) get smaller with the inclusion of the correction term.[56]

[54] Turkish labour Law requires reductions in the retirement earnings if a retired individual goes back to employment.
[55] The interpretation of the results in the rest of the chapter are based on the selection corrected results. The difference between the results with and without selection correction are noted where necessary.
[56] The decrease in returns to education with the control for self-selection is a well-documented finding in the literature. One example for the US can be found in Dahl (2002).

The coefficients for the age group dummies are positive and mostly significant. While there is no consistent pattern across the genders and years with respect to the age dummies, broadly what does appears is the following: workers in the '45 and over' category are more likely to be in the informal sector (other than females in 2010); returns for the '45 and over' workers (both females and males) tend to higher in the informal sector (except the 90th percentile) consistently, especially once we control for selection. Marriage has a positive effect on the wages of the female formal sector workers at each quantile. The inclusion of the correction term decreases the size of the coefficient for this variable in the lower half of the wage distribution and increases it at the 90^{th} quantile. Being the head of the household has a positive effect on wages and the size of this effect is smaller after the inclusion of the selection term.

The coefficient for the dummy variable that represents rural residence is mostly insignificant except at the 10^{th} and 90^{th} quantiles where it has a moderate size and a negative sign. A potential underlying reason for the insignificant coefficient for this variable is that most of the formal sector workers in rural areas are likely to be public sector workers for whom the wages are insensitive to the area of residence as they are set by the central government.

Finally, the coefficient of the correction term is negative and significant at each quantile (the squared term coefficient is positive). The negative effect indicates that women in the formal sector receive lower earnings than one would expect based on their observable characteristics. As noted above, Tannuri-Pianto and Pianto (2002) used the same methodology to analyse the structure of informal sector employment in Brazil. Similar to the results of this study, they also found a negative coefficient for the selection term at each quantile in the formal sector wage regressions. They argued that this may be an indication that these workers choose to be in the formal sector because

they benefit from the safety of other forms of legal compensation, such as advanced notification of termination or the employer's social security contributions. They asserted that such an interpretation is consistent with the increasing probability of informal sector employment at older ages (they also found a U-shaped relation between age and probability of informal sector employment) as older workers may choose informal sector employment after securing retirement benefits in the formal sector.

The wage regression results for the female informal sector workers are quite different from those for formal sector workers particularly for education and tenure. The coefficient for high school education becomes significant only after the 50^{th} quantile. For vocational high school education the coefficients are not significant at any quantile. The coefficient for university education is insignificant at the bottom end of the informal sector wage distribution and this is most likely due to the very low number of university graduates at this level of the wage distribution. The coefficient of the tenure variable is insignificant at the median, negative and significant at lower quantiles, and positive and significant at higher quantiles. Education and tenure do not have positive contributions to the wages of the female informal sector workers at lower quantiles in contrast to the formal sector wages where these two characteristics have a positive and significant effect. However, without the selection correction the coefficients for the educational categories are significant at the lower quantiles.

Looking at the coefficients of the age categories, it can be seen that the coefficient for '45 and over' is greater than the other age category coefficients except at the 90^{th} quantile. The effect of marriage is quite different for informal sector wages; a negative effect at the 10^{th} quantile and a positive effect at the top quantiles, and no significant effect at median wages. The coefficient for the rural residence dummy is negative and significant at every quantile. Finally, the effect of the selection term is positive and

significant at the lower end and at the median of the wage distribution. This means the individuals who are at the bottom half of the wage distribution receive higher earnings than one would expect based on their observable characteristics.

Tables 4.4A and 4.4B show the wage regression results estimated using the 2010 data for females. Looking at the coefficients of the education and tenure variables for formal sector workers, there is no difference between the 2002 and 2010 results in terms of sign and significance. The sizes of the education coefficients get larger at the 90th quantile and get smaller on the rest of the distribution. Returns to tenure increase from 2002 to 2010. Except for at the 10th and 25th quantiles, the '45 and over' age group still has the biggest coefficient among the age groups. Similar to the 2002 results, marriage and being the head of household have a positive effect on the wages of female formal sector workers while the effect of rural residence is insignificant. The coefficients of the correction terms are negative and significant similar to the results for the 2002 female sample.

Focusing on the informal sector wage estimations, it can be seen that the coefficients for university education become significant only after the 75th quantile. Interestingly, the coefficient of high school education is negative and significant at the bottom half of the distribution and becomes significant and positive at the 90th quantile. So, once again, it is found that education does not seem to have a positive contribution to the wages of female informal sector workers at the lower half of the distribution. The tenure variable, on the other hand, has a positive and significant coefficient at both the low and high end of the distribution. The age group coefficients are insignificant below the median and, once again, the '45 and over' age group has the biggest coefficient except at the 90th quantile. The signs of the marriage dummy variable coefficients are the same as those in 2002, but they are insignificant at the lower tail. The coefficients for being the head of

household and the rural residence dummies are qualitatively the same as the 2002 estimation. The first term of the selection term power expansion is positive and significant until the 75th quantile where it becomes negative afterwards. The negative effect at higher quantiles shows that women in the high-wage informal sector earn less than what would be expected considering their observable characteristics. These workers may choose the informal sector for the non-wage benefits such as flexible working hours.

Male sample

Tables 4.5A and 4.5B show the wage estimation results for the 2002 male sample with and without correction results, respectively. As expected, the coefficients of the education level dummies and tenure are all significant and positive in the formal sector wage estimation. The age category coefficients are significant and positive at each quantile among which the "45 and over" age group has the largest coefficient. Interestingly, marriage has a negative effect on the wages of males in the formal sector, which is in contrast to the results for females. Similar to the female sample results, being head of the household increases the male formal sector wages and rural residence has no significant effect except at the bottom end of the distribution where it has a negative effect. The selection term has significant and negative coefficients at each quantile.

The wage estimation results for the informal sector are quite different from those for the formal sector in terms of significance and the sign of the coefficients. For the educational levels, only the coefficient of the university degree is significant at all quantiles. The coefficients for the other education groups are insignificant at quantiles below the median. Tenure has a positive and significant effect on wages except at the

10th quantile, where it is insignificant. The age group coefficients are significant only at the top half of the distribution. While marriage does not have a significant effect, rural residence has a decreasing effect on wages. The coefficient for the head of the household dummy is positive but significant except at the 75th quantile. The coefficient of the first selection term is positive and significant at the bottom half of the wage distribution while it is insignificant at the top half of the distribution.

Similar to the results of the wage estimation for the informal sector female workers, the inclusion of the selection term makes a great difference to the coefficients of the characteristics, particularly for the educational level coefficients. The reward for having a higher level of education seems to disappear at the bottom half of the informal sector wage distribution when corrected for selection bias.

Tables 4.6A and 4.6B present the wage regression results estimated by using the 2010 male sample. The coefficients of the education and tenure variables in the formal sector wage estimations are qualitatively similar to the 2002 results. One notable change is that, compared to the 2002 results, returns to a university degree are lower at the bottom half and higher at the top half of the distribution. As noted above, a similar change is observed between the 2002 and 2010 wage estimations of female formal sector workers. This is an indicator of an increase in inequality between university graduates. The number of universities in Turkey has increased dramatically over the last two decades. Therefore, a possible explanation for this result can be the increasing difference in the quality of education between the old universities and the new universities that are noted to be lacking adequate teaching staff and equipment (Tansel, 2012). The size of the tenure coefficient is smaller than the 2002 results at all quantiles. The coefficients of the age groups, marriage and the head of household dummy variables are also similar to the 2002 results in terms of sign and significance. The rural residence dummy has a positive

and significant coefficient above the median of the distribution yet the size of the coefficient is very small. Once again the coefficients of the selection terms are negative and significant along the distribution of formal sector wages.

Looking at the regression results of the informal sector wages for the year 2010, the same differences between the results with and without the sample selection correction can be found. Once the selection terms are introduced to the model, the coefficients of the educational categories lose their significance. Tenure has a positive effect on the informal sector wages, but it is significant only at the bottom end of the distribution. The age group dummies have positive and significant coefficients. While marriage has a positive and significant effect at the 50^{th} and 75^{th} quantiles, the coefficient of the head of the household dummy is significant and positive at the 90^{th} quantile. Rural residence has a negative effect on the informal sector wages which is consistent with the 2002 results. The coefficient of the selection term is similar to the 2002 result; positive and significant at the bottom of the wage distribution.

4.5.3 Decomposition results

The segmented market theory, as noted in Section 1, suggests that due to strong entry barriers between the formal and informal sectors, workers with the same characteristics may have different earnings as some of them involuntarily end up in the informal sector where the returns to characteristics are lower. An observed wage gap between formal and informal sector wages does not necessarily prove the existence of a segmented market. In order to understand the structure of the Turkish labour market, the observed wage gap between the formal and informal sector wages is decomposed into two components: the effect of characteristics and the effect of coefficients (returns to characteristics). The magnitude of the coefficient effect is quite crucial because it reveals the effect of the difference between the returns to worker characteristics (such as

education) after controlling for the differences in the distribution of these characteristics between the formal and the informal sectors. In that sense, if the segmented market theory holds, the coefficient effect should explain a substantial part of the wage gap between the workers in the formal and the informal sectors. However, if the wage gap between these sectors is mainly explained by the characteristic effect, it means that the informal sector workers earn less simply because they are less skilled than formal sector workers.

Figure 4.1 presents the decomposition results of the informal/formal sector wage gap for the 2002 and 2010 female samples. In Figure 4.1, the decomposition results without the selection term are also provided to be able to see the effect of selection into the sectors. Looking at the wage gap between the formal and informal sectors in 2002 at different quantiles, it can be seen that the wage gap is highest at the bottom end of the distribution where it approaches 2 log wage points. Then, the value of the gap decreases until the 30^{th} quantile, increases very slowly until the 80^{th} quantile and decreases after this point until its lowest value of 0.8 at the top end of the distribution. Comparing the values of the wage gap in 2002 and 2010, a general decrease in the wage gap between formal and informal sector wages across the quantiles is observed. The highest decrease in the gap is observed between the 5^{th} and the 15^{th} quantiles where the size of the decrease is around 0.4 log wage points. On the other hand, there is only a minor difference between these two years after the 90^{th} quantile.

The decomposition results for the 2002 female sample show that, while the characteristic effect explains the majority of the wage gap at higher quantiles, the coefficient effect explains a greater share at the bottom of the wage distribution. The sizes of these effects equalise below the median at around the 35^{th} quantile. However,

when the selection term is removed from the analysis, the shares of the effect equalise at the median level. Moving to the 2010 female sample results, a similar pattern in terms of the share of the effects is seen at the bottom and top parts of the distribution. The main difference between the results for the two years is that in 2010 the equalisation level, the 20th quantile, is much lower than in 2002. Even though the difference is not as visible as it is in 2002, it is found that the equalisation occurs at a higher point of the distribution when the selection term is not included in the analysis.

Figure 4.2 presents the decomposition results for the male sample for 2002 and 2010. Comparing the 2002 female and male sample results, it is observed that the wage gap between the two sectors is lower for males than females. The highest wage gap in the 2002 male data (which occurs at the bottom and the top end of the distribution) is 1 log wage points which is half of the highest gap in the 2002 female data (which occurs at the bottom end). At the median, the level of the wage gap for males is 0.3 log wage points (around 30 per cent) lower than it is for females. Similar to the case of females, a decrease in the wage gap between formal and informal sector male workers from 2002 to 2010 is observed. At the median, the wage gap in 2002 is 0.25 log wage points higher than the 2010 value, which corresponds to around a 32 per cent decrease between these two years.

Comparing the results with and without the correction for 2002, a major difference in the relative shares of coefficient and characteristic effects is observed. Before the correction for selection, the coefficient effect explains a greater share of the wage gap along the distribution (particularly at the bottom end) and the equalisation occurs at the top end of the distribution. On the other hand, results with the correction are closer to those for the 2002 female sample and the equalisation point is at a much lower level (at

the 10th quantile). The characteristics effect explains most of the total wage differential above this point.

The 2010 male sample results also show a similar pattern with a greater share of the coefficient effect at the bottom of the distribution and a greater effect of characteristics at the top end. Once again, the size of the wage gap for the male sample is smaller than for the 2010 female sample. Also, the difference between results with and without correction is similar to the results for the female samples and the 2002 male sample; the equalisation level is lower in the results with correction.

The results of the study show that the coefficient effect explains a greater share of the wage gap between the formal and the informal sectors at the bottom of the wage distributions for males and females. In contrast, the characteristic effect dominates the wage gap at the higher quantiles. Therefore, the results are inconsistent with the segmented market theory and the competitive market theory which assume a homogeneous structure for the informal sector. However, they support the "upper tier" and "lower tier" hypothesis that suggests a heterogeneous structure for the informal sector (Fields, 2005; Maloney, 2004). The higher share of the coefficient effect in the total wage gap is an indicator of segregation at the bottom of the wage distribution. Therefore, informal sector employment may not be a voluntary choice at the lower tier. In the upper tier, there is an indication of a competitive market as the wage gap is mostly explained by the differences between the characteristics of formal and informal sector workers, rather than differences between the remuneration of these characteristics. In that sense, given their characteristics, participation in the informal sector at the top end of the distribution is likely to be a voluntary choice of the workers.

The results of the analyses are consistent with the findings of Tansel and Kan (2012) who also find a heterogeneous structure in the informal sector. However, the results found in this thesis show that taking selection into account substantially changes the results of the wage regressions and where in the wage distribution the segmentation between the formal and informal sector occurs. It is found that returns to higher levels of education, especially at the lower quantiles, disappear in the informal sector once the selection term is introduced in the wage estimation. It is also found that the decomposition results overestimate the spread of segmentation over the higher quantiles when the selection term is omitted. On the other hand, when the correction term is included, it is seen that the spread of segmentation, particularly for males, is much narrower. The results of the study are also consistent with the findings of Baskaya and Hulagu (2011), who document a higher formal-informal wage gap for females than for males. In this study, it is also shown that the gap for females is not only higher at the mean wages, but also at all quantiles of the wage distribution. In that sense, this research contributes to the previous literature by providing a more detailed analysis of segmentation in the Turkish labour market.

4.6 Conclusion

This study analyses the wage gap between formal and informal sector workers to see whether there is segmentation between these two sectors in the Turkish labour market. To do so, a quantile regression decomposition procedure is performed on LFS data for the years 2002 and 2010. The results of the study suggest a heterogeneous structure for the informal sector in Turkey in which the lower tier consists of workers who involuntarily participate in the informal sector and the higher tier includes workers who voluntarily choose to be in the informal sector. It is also found that the wage gap

between the two sectors is higher and the segmentation is more spread out along the wage distribution for females than it is for males.

The main limitation of the analysis is that, since there is no data on their earnings, the self-employed are not included in the estimation. Yet, it is not expected that this exclusion will affect the generality of the results (especially for females since the share of self-employed is very low in total female employment) as the findings of the study are consistent with previous works on the Turkish labour market, such as Tansel and Kan (2012), which included the self-employed in their analysis.

The findings of the chapter provide useful information on informal sector employment in Turkey for policy makers. For instance, the results of the analysis show that informal sector workers who are at the bottom of the wage distribution involuntarily participate in the informal sector. Due to the unregulated nature of informal employment, these workers do not enjoy the protection from exploitation such as wages below the minimum wage, working in risky environments, and working with unsafe equipment. Moreover, these individuals do not have employment benefits such as health insurance. To eliminate such problems, policy makers should consider more efficient ways to detect the firms which avoid registering their employees. On the other hand, it is also found that some of the workers in the informal sector voluntarily choose to be in the informal sector. Therefore, before formulating a policy which targets elimination of informality, it is important to understand the motivations of these individuals. For instance, the results of the selection equation show that individuals, women in particular, who have children (especially those who have young children) are more likely to be in the informal sector. A possible explanation is that these individuals may choose to work in the informal sector to avoid the rigid working hours in the formal sector in order to spare more time to take care of their children. Accordingly, policy

makers could focus on increasing availability of childcare assistance for working parents.

TABLES OF CHAPTER 4

Table 4.1: Descriptive Statistics

	2002				2010					
	Female		Male		Female		Male (60%)		Male (full)	
VARIABLES	Informal	Formal	Informal	Formal	Informal	Formal	Informal	Formal	Informal	Formal
%	25.1	74.9	25.3	74.7	19.9	80.1	19.3	80.7	19.3	80.7
Real hourly wages	127.9	375.3	152.6	366.3	175.2	435.1	193.2	383.4	191.2	374.0
	(3.95)	(4.67)	(2.67)	(4.56)	(3.75)	(3.37)	(2.02)	(2.13)	(2.12)	(1.65)
Log of real wages	4.466	5.657	4.786	5.612	4.926	5.843	5.117	5.736	5.121	5.735
	(0.019)	(0.009)	(0.008)	(0.005)	(0.012)	(0.006)	(0.007)	(0.004)	(0.005)	(0.003)
Less than secondary school	0.725	0.180	0.713	0.357	0.624	0.165	0.588	0.317	0.588	0.318
Secondary school	0.077	0.063	0.122	0.138	0.193	0.066	0.194	0.145	0.195	0.147
High school	0.099	0.202	0.090	0.166	0.101	0.146	0.095	0.136	0.098	0.134
Vocational school	0.055	0.135	0.049	0.146	0.073	0.122	0.082	0.152	0.077	0.152
University	0.045	0.419	0.025	0.192	0.062	0.499	0.040	0.250	0.043	0.249
Tenure	5.393	6.891	6.640	8.615	1.906	6.243	3.246	7.288	3.239	7.317
	(0.174)	(0.080)	(0.094)	(0.050)	(0.062)	(0.057)	(0.073)	(0.046)	(0.056)	(0.036)
Age										
20-25	0.277	0.229	0.201	0.090	0.237	0.169	0.183	0.084	0.186	0.087
25-35	0.339	0.460	0.415	0.437	0.296	0.452	0.349	0.430	0.347	0.426
35-45	0.246	0.246	0.208	0.324	0.289	0.286	0.224	0.319	0.224	0.319
45-	0.138	0.065	0.175	0.149	0.178	0.093	0.245	0.167	0.244	0.168
Marital Status										
Single	0.316	0.399	0.243	0.156	0.313	0.373	0.267	0.206	0.266	0.206
Married	0.593	0.550	0.757	0.844	0.584	0.554	0.733	0.794	0.734	0.794
Divorced/widowed	0.091	0.051			0.103	0.073				
Household characteristics										
Head of household dummy	0.095	0.097	0.673	0.799	0.103	0.104	0.631	0.744	0.634	0.743
Rural	0.297	0.138	0.278	0.221	0.205	0.097	0.246	0.150	0.245	0.149
Mean of years of education of household adults	5.314	8.756	4.505	6.598	5.785	9.010	4.729	7.095	4.705	7.082
	(0.074)	(0.060)	(0.036)	(0.027)	(0.064)	(0.042)	(0.045)	(0.027)	(0.035)	(0.021)
Presence of a formal sector worker in the household	0.371	0.638	0.156	0.233	0.451	0.646	0.211	0.286	0.205	0.289
Presence of a child in the household										
Age≤4	0.187	0.176	0.381	0.346	0.169	0.181	0.349	0.337	0.349	0.340
5≤Age<11	0.358	0.266	0.415	0.427	0.341	0.242	0.404	0.383	0.407	0.382
11≤Age<15	0.255	0.137	0.221	0.204	0.250	0.133	0.232	0.183	0.230	0.183
Firm size										
Less than 10	0.717	0.134	0.774	0.205	0.639	0.172	0.729	0.220	0.727	0.220
10-24	0.122	0.112	0.106	0.111	0.148	0.130	0.102	0.118	0.103	0.120
25-49	0.075	0.202	0.062	0.168	0.138	0.237	0.100	0.210	0.102	0.210
50-	0.086	0.553	0.058	0.516	0.075	0.460	0.069	0.452	0.069	0.451
Observations	2,303	6,867	8,036	23,623	3,703	14,895	6,745	28,201	11,290	46,953

Note: Standard errors in parenthesis. Sample weights are used.

Table 4.2: Selection Equation

	2002				2010			
	Female		Male		Female		Male	
VARIABLES	Probit	Single Index	Probit	Single Index	Probit	Single Index	Probit	Single index
Mean of years of education of household adults	-0.018***	-0.018	-0.031***	-0.031	-0.025***	-0.025	-0.034***	-0.034
	(0.006)	.	(0.003)	.	(0.004)	.	(0.003)	.
Secondary school	-0.554***	-1.189***	-0.352***	-1.521***	-0.476***	-0.459***	-0.259***	-0.439***
	(0.069)	(0.126)	(0.028)	(0.061)	(0.046)	(0.057)	(0.026)	(0.032)
High school	-1.096***	-2.204***	-0.590***	-2.218***	-1.095***	-1.084***	-0.501***	-0.970***
	(0.059)	(0.224)	(0.031)	(0.0857)	(0.045)	(0.124)	(0.031)	(0.063)
Vocational high school	-1.125***	-2.174***	-0.752***	-3.009***	-1.075***	-1.059***	-0.547***	-0.974***
	(0.067)	(0.223)	(0.036)	(0.115)	(0.048)	(0.121)	(0.031)	(0.063)
University	-1.395***	-2.742***	-0.926***	-4.189***	-1.605***	-1.602***	-0.948***	-2.073***
	(0.066)	(0.284)	(0.041)	(0.162)	(0.046)	(0.180)	(0.037)	(0.120)
Presence of a formal sector worker in the household	-0.379***	-0.788***	-0.192***	-0.813***	-0.318***	-0.358***	-0.109***	-0.432***
	(0.042)	(0.084)	(0.027)	(0.041)	(0.030)	(0.047)	(0.024)	(0.031)
Marital Status								
Married	0.095	0.122***	-0.180***	-0.717***	-0.007	0.084***	-0.178***	-0.423***
	(0.060)	(0.026)	(0.040)	(0.032)	(0.042)	(0.028)	(0.037)	(0.029)
Divorced-widowed	0.163*	0.295***	.	.	-0.133**	-0.016	.	.
	(0.096)	(0.042)	.	.	(0.061)	(0.041)	.	.
Head of household dummy	-0.200**	-0.463***	-0.224***	-0.868***	-0.125**	-0.198***	-0.227***	-0.431***
	(0.082)	(0.043)	(0.034)	(0.042)	(0.054)	(0.041)	(0.031)	(0.030)
Presence of a child in the household								
Age≤4	0.180***	0.368***	0.155***	0.693***	0.184***	0.226***	0.165***	0.390***
	(0.057)	(0.048)	(0.023)	(0.026)	(0.041)	(0.034)	(0.023)	(0.026)
5≤Age<11	0.107**	0.352***	0.010	0.029*	0.225***	0.224***	0.031	0.011
	(0.047)	(0.042)	(0.021)	(0.015)	(0.032)	(0.033)	(0.021)	(0.017)
11≤Age<15	0.194***	0.416***	0.024	0.077***	0.133***	0.151***	0.068***	0.058***
	(0.049)	(0.047)	(0.024)	(0.016)	(0.035)	(0.027)	(0.023)	(0.018)
Firm Size								
10-24	-0.826***	-1.326***	-0.727***	-2.330***	-0.698***	-0.525***	-0.762***	-1.364***
	(0.056)	(0.133)	(0.028)	(0.087)	(0.039)	(0.058)	(0.028)	(0.074)
25-49	-1.327***	-2.623***	-1.213***	-4.685***	-1.018***	-1.103***	-1.078***	-2.123***
	(0.057)	(0.261)	(0.029)	(0.174)	(0.037)	(0.117)	(0.026)	(0.113)
More than 50	-1.869***	-3.748***	-1.823***	-7.003***	-1.636***	-1.603***	-1.512***	-2.943***
	(0.049)	(0.366)	(0.026)	(0.260)	(0.038)	(0.169)	(0.025)	(0.156)
Age								
25-35	-0.408***	-0.917***	-0.394***	-1.323***	-0.440***	-0.451***	-0.439***	-0.913***
	(0.054)	(0.093)	(0.035)	(0.041)	(0.042)	(0.055)	(0.033)	(0.061)
35-45	-0.375***	-0.570***	-0.485***	-1.944***	-0.419***	-0.464***	-0.490***	-0.945***
	(0.067)	(0.062)	(0.041)	(0.070)	(0.049)	(0.059)	(0.040)	(0.067)
45-	0.270***	0.509***	0.040	0.163***	-0.005	-0.131***	0.154***	0.322***
	(0.082)	(0.063)	(0.043)	(0.020)	(0.058)	(0.041)	(0.042)	(0.012)
Constant	1.326***	1.326	1.129***	1.129	1.166***	1.166	0.790***	0.790
	(0.058)	.	(0.034)	.	(0.047)	.	(0.034)	.
Hausman test statistics	20.67		198.17		105.91		63.249	
Hausman test p-value	0.00		0.00		0.00		0.00	
Observations	9,170		31,659		18,598		34,946	

Dependent variable takes value 1 if the worker is in the informal sector.

Base categories: less than secondary school (education), never married (marital status), less than 10 workers (firm size), 20-25 (age). The same base categories are used in the wage regressions. The constant and the coefficient of a continuous independent variable in single index selection equations cannot be identified. So, their values are normalised to probit estimates.

Table 4.3A: 2002 Wage regression for females (with correction)

			Informal					Formal		
VARIABLES	q10	q25	q50	q75	q90	q10	q25	q50	q75	q90
Secondary school	0.009	0.089	0.062	0.088	0.148	0.019	0.027	0.107***	0.155***	0.167**
	(0.150)	(0.077)	(0.049)	(0.055)	(0.097)	(0.039)	(0.028)	(0.033)	(0.029)	(0.066)
High school	0.057	0.069	0.082*	0.162***	0.230*	0.184***	0.300***	0.361***	0.465***	0.480***
	(0.133)	(0.075)	(0.049)	(0.056)	(0.124)	(0.033)	(0.026)	(0.027)	(0.033)	(0.057)
Vocational school	-0.088	0.034	0.036	0.038	0.044	0.253***	0.334***	0.416***	0.519***	0.515***
	(0.139)	(0.075)	(0.062)	(0.064)	(0.079)	(0.034)	(0.027)	(0.029)	(0.033)	(0.048)
University	0.234	0.353***	0.624***	1.001***	1.305***	0.713***	0.812***	0.902***	0.981***	0.953***
	(0.188)	(0.091)	(0.133)	(0.130)	(0.213)	(0.040)	(0.025)	(0.026)	(0.032)	(0.053)
Tenure (10^{-2})	-2.851**	-2.678***	0.468	2.120***	2.458**	4.140***	3.915***	3.765***	3.025***	2.812***
	(1.274)	(0.934)	(0.751)	(0.629)	(1.065)	(0.468)	(0.328)	(0.360)	(0.377)	(0.594)
Tenure squared	2.950	3.447	-4.528*	-8.639***	-8.460**	-5.978***	-6.760***	-6.879***	-4.572***	-4.386**
	(4.513)	(3.269)	(2.582)	(2.310)	(3.352)	(1.709)	(1.209)	(1.352)	(1.415)	(2.124)
Age										
25-35	-0.031	0.011	0.102**	0.110**	0.200**	0.043	0.133***	0.144***	0.127***	0.152***
	(0.095)	(0.066)	(0.043)	(0.048)	(0.079)	(0.034)	(0.024)	(0.022)	(0.024)	(0.035)
35-45	0.171	0.199**	0.192***	0.136**	0.254**	0.046	0.113***	0.118***	0.068**	0.081*
	(0.114)	(0.084)	(0.055)	(0.057)	(0.099)	(0.041)	(0.029)	(0.027)	(0.031)	(0.043)
45-	0.454***	0.334***	0.362***	0.260***	0.187*	0.100**	0.164***	0.130***	0.096**	0.141**
	(0.142)	(0.095)	(0.066)	(0.076)	(0.105)	(0.048)	(0.037)	(0.035)	(0.042)	(0.062)
Married	-0.346***	-0.098	0.030	0.125***	0.169**	0.183***	0.156***	0.113***	0.125***	0.131***
	(0.111)	(0.068)	(0.042)	(0.046)	(0.083)	(0.027)	(0.020)	(0.016)	(0.020)	(0.027)
Divorced-widowed	-0.062	0.062	0.077	0.156*	0.175	0.108*	0.054	-0.005	0.044	-0.006
	(0.152)	(0.118)	(0.087)	(0.094)	(0.152)	(0.057)	(0.039)	(0.040)	(0.036)	(0.070)
Head of household dummy	0.166	0.102	0.091	0.170*	0.213	0.109***	0.129***	0.171***	0.211***	0.308***
	(0.137)	(0.115)	(0.071)	(0.099)	(0.139)	(0.039)	(0.026)	(0.028)	(0.028)	(0.047)
Rural	-0.289***	-0.267***	-0.301***	-0.329***	-0.363***	-0.095**	-0.036	-0.013	-0.020	-0.065*
	(0.084)	(0.058)	(0.047)	(0.046)	(0.059)	(0.042)	(0.028)	(0.026)	(0.028)	(0.034)
λ	0.730***	0.284**	0.193**	-0.043	-0.105	-0.715***	-0.683***	-0.639***	-0.642***	-0.648***
	(0.245)	(0.138)	(0.092)	(0.103)	(0.204)	(0.106)	(0.071)	(0.068)	(0.082)	(0.119)
λ^2	-0.163*	-0.006	0.012	0.073	0.071	0.319***	0.314***	0.294***	0.304***	0.269***
	(0.095)	(0.053)	(0.039)	(0.046)	(0.109)	(0.068)	(0.045)	(0.045)	(0.061)	(0.080)
Constant	3.600***	4.008***	4.335***	4.604***	4.956***	4.443***	4.596***	4.845***	5.117***	5.427***
	(0.128)	(0.078)	(0.069)	(0.071)	(0.100)	(0.050)	(0.031)	(0.031)	(0.035)	(0.063)
Observations	2,303	2,303	2,303	2,303	2,303	6,867	6,867	6,867	6,867	6,867

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4.3B: 2002 Wage regression for females (without correction)

			Informal					Formal		
VARIABLES	q10	q25	q50	q75	q90	q10	q25	q50	q75	q90
Secondary school	0.235	0.180**	0.116	0.107	0.127	0.103***	0.106**	0.142***	0.208***	0.223**
	(0.143)	(0.090)	(0.132)	(0.384)	(0.092)	(0.035)	(0.048)	(0.034)	(0.038)	(0.099)
High school	0.341***	0.249***	0.223	0.168	0.275**	0.330***	0.409***	0.474***	0.579***	0.638***
	(0.103)	(0.068)	(0.160)	(0.388)	(0.129)	(0.032)	(0.045)	(0.024)	(0.036)	(0.079)
Vocational school	0.280**	0.265***	0.199***	0.076	0.090	0.364***	0.441***	0.510***	0.624***	0.646***
	(0.120)	(0.088)	(0.069)	(0.060)	(0.070)	(0.032)	(0.040)	(0.029)	(0.034)	(0.158)
University	0.643***	0.608***	0.953***	1.145***	1.322***	0.911***	1.001***	1.061***	1.138***	1.147***
	(0.143)	(0.111)	(0.210)	(0.436)	(0.200)	(0.031)	(0.048)	(0.020)	(0.028)	(0.048)
Tenure (10^{-2})	-2.889**	-3.151**	0.210	1.781*	2.345**	3.896***	3.899	4.019***	3.457***	3.158**
	(1.367)	(1.429)	(3.567)	(0.994)	(0.997)	(0.418)	(2.600)	(0.326)	(0.422)	(1.601)
Tenure squared	2.292	4.823	-4.053	-7.402***	-8.258***	-4.552***	-6.362	-7.665***	-5.418***	-4.797
	(4.852)	(4.152)	(10.251)	(2.674)	(3.145)	(1.589)	(8.692)	(1.245)	(1.653)	(6.641)
Age										
25-35	0.105	0.059	0.160	0.125	0.223***	0.101***	0.161***	0.180***	0.172***	0.193
	(0.100)	(0.085)	(0.197)	(0.321)	(0.071)	(0.032)	(0.059)	(0.023)	(0.025)	(0.142)
35-45	0.224*	0.259***	0.255	0.163	0.262***	0.078*	0.137**	0.149***	0.090***	0.100
	(0.129)	(0.092)	(0.329)	(0.313)	(0.097)	(0.042)	(0.060)	(0.028)	(0.031)	(0.080)
45-	0.447***	0.329***	0.366***	0.268***	0.218**	0.091	0.142	0.149***	0.079*	0.129
	(0.151)	(0.109)	(0.095)	(0.095)	(0.094)	(0.055)	(0.138)	(0.041)	(0.047)	(0.091)
Married	-0.401***	-0.144*	-0.011	0.110	0.169**	0.256***	0.180***	0.116***	0.128***	0.124
	(0.109)	(0.077)	(0.056)	(0.082)	(0.078)	(0.028)	(0.059)	(0.019)	(0.021)	(0.104)
Divorced-widowed	-0.032	0.029	0.086	0.110	0.249*	0.129**	0.068	-0.006	0.033	-0.052
	(0.150)	(0.202)	(0.170)	(0.404)	(0.145)	(0.052)	(0.098)	(0.040)	(0.043)	(0.187)
Head of household dummy	0.021	0.135	0.088	0.188	0.177	0.126***	0.157***	0.177***	0.226***	0.301***
	(0.148)	(0.118)	(0.068)	(0.407)	(0.141)	(0.042)	(0.051)	(0.030)	(0.031)	(0.055)
Rural	-0.244***	-0.290***	-0.315**	-0.337***	-0.354***	-0.083**	-0.018	-0.033*	-0.035	-0.110
	(0.093)	(0.075)	(0.126)	(0.095)	(0.056)	(0.038)	(0.052)	(0.018)	(0.032)	(0.127)
Constant	3.719***	4.099***	4.368***	4.683***	4.920***	4.107***	4.328***	4.608***	4.868***	5.163***
	(0.085)	(0.072)	(0.063)	(0.390)	(0.046)	(0.034)	(0.111)	(0.023)	(0.025)	(0.053)
Observations	2,303	2,303	2,303	2,303	2,303	6,867	6,867	6,867	6,867	6,867

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4.4A: 2010 Wage regression for females (with correction)

			Informal					Formal		
VARIABLES	q10	q25	q50	q75	q90	q10	q25	q50	q75	q90
Secondary school	-0.013	-0.003	-0.015	0.007	0.103	0.078***	0.111***	0.123***	0.139***	0.208***
	(0.085)	(0.048)	(0.026)	(0.036)	(0.066)	(0.022)	(0.016)	(0.015)	(0.020)	(0.030)
High school	-0.135*	-0.174***	-0.080**	0.020	0.178**	0.123***	0.185***	0.249***	0.362***	0.498***
	(0.077)	(0.046)	(0.035)	(0.044)	(0.086)	(0.021)	(0.016)	(0.014)	(0.020)	(0.029)
Vocational school	-0.203**	-0.134**	-0.051	0.009	0.133**	0.123***	0.190***	0.295***	0.443***	0.573***
	(0.087)	(0.068)	(0.034)	(0.037)	(0.061)	(0.022)	(0.018)	(0.018)	(0.020)	(0.026)
University	-0.112	-0.029	0.091	0.386***	0.554***	0.476***	0.613***	0.820***	0.974***	1.130***
	(0.109)	(0.079)	(0.065)	(0.105)	(0.106)	(0.021)	(0.018)	(0.016)	(0.019)	(0.025)
Tenure (10^{-2})	8.312***	5.203***	4.451***	4.714***	4.191***	4.221***	4.566***	4.237***	3.336***	3.101***
	(1.413)	(0.831)	(0.682)	(0.990)	(1.171)	(0.198)	(0.184)	(0.169)	(0.226)	(0.288)
Tenure squared	-38.460***	-22.979***	-18.353***	-12.674	-5.420	-3.844***	-6.802***	-8.126***	-6.324***	-6.310***
	(9.351)	(5.091)	(4.532)	(8.017)	(6.191)	(0.765)	(0.692)	(0.563)	(0.854)	(0.994)
Age										
25-35	-0.087	0.014	0.072**	0.059*	0.173***	0.009	0.059***	0.112***	0.155***	0.151***
	(0.071)	(0.047)	(0.031)	(0.033)	(0.060)	(0.016)	(0.016)	(0.014)	(0.016)	(0.021)
35-45	0.001	0.060	0.083**	0.073*	0.201***	0.070***	0.082***	0.129***	0.173***	0.198***
	(0.094)	(0.053)	(0.035)	(0.040)	(0.063)	(0.020)	(0.019)	(0.016)	(0.017)	(0.025)
45-	0.062	0.100*	0.125***	0.106**	0.142*	-0.002	0.046**	0.115***	0.174***	0.261***
	(0.095)	(0.054)	(0.038)	(0.045)	(0.074)	(0.028)	(0.021)	(0.019)	(0.024)	(0.036)
Married	-0.040	0.038	0.064**	0.114***	0.114**	0.186***	0.196***	0.164***	0.153***	0.131***
	(0.079)	(0.040)	(0.028)	(0.030)	(0.052)	(0.016)	(0.012)	(0.011)	(0.012)	(0.015)
Divorced-widowed	0.084	0.033	0.015	0.048	0.056	0.068***	0.068***	0.018	0.020	0.011
	(0.080)	(0.060)	(0.042)	(0.048)	(0.083)	(0.026)	(0.022)	(0.018)	(0.022)	(0.030)
Head of household dummy	0.020	0.061	0.042	0.100**	0.114	0.179***	0.198***	0.217***	0.215***	0.217***
	(0.086)	(0.047)	(0.040)	(0.039)	(0.081)	(0.026)	(0.018)	(0.014)	(0.018)	(0.023)
Rural	-0.148**	-0.162***	-0.156***	-0.167***	-0.212***	-0.014	-0.040***	0.001	0.030*	-0.005
	(0.067)	(0.033)	(0.023)	(0.022)	(0.038)	(0.018)	(0.014)	(0.015)	(0.017)	(0.014)
λ	1.101***	0.506***	0.180***	-0.114	-0.448***	-0.766***	-0.676***	-0.530***	-0.537***	-0.394***
	(0.186)	(0.079)	(0.068)	(0.070)	(0.126)	(0.054)	(0.048)	(0.049)	(0.059)	(0.075)
λ^2	-0.235***	-0.046	0.035	0.128***	0.210***	0.405***	0.354***	0.301***	0.316***	0.226***
	(0.068)	(0.030)	(0.029)	(0.028)	(0.051)	(0.041)	(0.037)	(0.040)	(0.044)	(0.058)
Constant	3.853***	4.203***	4.611***	4.885***	5.275***	4.752***	4.868***	4.996***	5.152***	5.291***
	(0.161)	(0.104)	(0.045)	(0.061)	(0.077)	(0.027)	(0.023)	(0.021)	(0.025)	(0.030)
Observations	3,703	3,703	3,703	3,703	3,703	14,895	14,895	14,895	14,895	14,895

Standard errors in parentheses*** p<0.01, ** p<0.05, * p<0.1

Table 4.4B: 2010 Wage regression for females (without correction)

			Informal					Formal		
VARIABLES	q10	q25	q50	q75	q90	q10	q25	q50	q75	q90
Secondary school	0.219*	0.154***	0.050	0.045	0.053	0.121***	0.149***	0.150***	0.177***	0.232***
	(0.122)	(0.054)	(0.060)	(0.037)	(0.107)	(0.023)	(0.016)	(0.015)	(0.019)	(0.025)
High school	0.288	0.146***	0.085	0.143***	0.205*	0.235***	0.285***	0.319***	0.446***	0.569***
	(0.176)	(0.045)	(0.132)	(0.044)	(0.109)	(0.020)	(0.015)	(0.017)	(0.017)	(0.022)
Vocational school	0.253**	0.223***	0.143***	0.109***	0.122	0.224***	0.282***	0.372***	0.546***	0.647***
	(0.103)	(0.060)	(0.047)	(0.039)	(0.093)	(0.019)	(0.015)	(0.017)	(0.020)	(0.024)
University	0.509**	0.501***	0.497***	0.735***	0.981***	0.652***	0.783***	0.949***	1.108***	1.234***
	(0.201)	(0.070)	(0.081)	(0.095)	(0.176)	(0.019)	(0.014)	(0.013)	(0.014)	(0.017)
Tenure (10^{-2})	7.815**	4.784***	5.048***	4.621***	4.977**	4.518***	4.867***	4.340***	3.538***	3.181***
	(3.408)	(0.954)	(1.541)	(0.956)	(2.161)	(0.234)	(0.207)	(0.173)	(0.201)	(0.248)
Tenure squared	-36.778**	-21.906***	-19.891***	-14.688*	-8.570	-3.988***	-7.539***	-8.213***	-6.689***	-6.405***
	(16.397)	(5.157)	(6.638)	(7.542)	(7.348)	(0.814)	(0.751)	(0.581)	(0.804)	(0.845)
Age										
25-35	0.056	0.135***	0.148**	0.116***	0.135	0.060***	0.110***	0.147***	0.181***	0.173***
	(0.056)	(0.048)	(0.060)	(0.033)	(0.121)	(0.020)	(0.015)	(0.014)	(0.015)	(0.018)
35-45	0.219	0.182***	0.157	0.150***	0.172**	0.125***	0.128***	0.160***	0.197***	0.216***
	(0.143)	(0.057)	(0.141)	(0.036)	(0.075)	(0.023)	(0.018)	(0.016)	(0.020)	(0.024)
45-	0.179	0.164***	0.157**	0.174***	0.129	-0.006	0.075***	0.136***	0.185***	0.272***
	(0.136)	(0.059)	(0.062)	(0.041)	(0.091)	(0.030)	(0.023)	(0.019)	(0.022)	(0.030)
Married	-0.223	-0.010	0.035	0.080***	0.140***	0.189***	0.197***	0.165***	0.155***	0.133***
	(0.138)	(0.047)	(0.089)	(0.027)	(0.053)	(0.016)	(0.013)	(0.012)	(0.012)	(0.016)
Divorced-widowed	-0.079	-0.015	-0.014	0.007	0.048	0.053*	0.052**	0.005	0.012	0.006
	(0.097)	(0.073)	(0.083)	(0.050)	(0.100)	(0.029)	(0.022)	(0.015)	(0.022)	(0.027)
Head of household dummy	0.148	0.092	0.058	0.079*	0.189**	0.183***	0.203***	0.222***	0.224***	0.208***
	(0.101)	(0.058)	(0.084)	(0.041)	(0.087)	(0.024)	(0.019)	(0.015)	(0.020)	(0.023)
Rural	-0.125**	-0.182***	-0.164***	-0.178***	-0.209***	-0.042*	-0.041***	0.001	0.017	-0.015
	(0.058)	(0.036)	(0.059)	(0.024)	(0.045)	(0.022)	(0.015)	(0.013)	(0.015)	(0.015)
Constant	3.922***	4.333***	4.691***	4.965***	5.202***	4.465***	4.605***	4.808***	4.963***	5.152***
	(0.142)	(0.046)	(0.044)	(0.032)	(0.117)	(0.024)	(0.016)	(0.018)	(0.016)	(0.018)
Observations	3,703	3,703	3,703	3,703	3,703	14,895	14,895	14,895	14,895	14,895

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4.5A: 2002 Wage regression for males (with correction)

	Informal					Formal				
VARIABLES	q10	q25	q50	q75	q90	q10	q25	q50	q75	q90
Secondary school	0.021	-0.018	0.018	0.104***	0.124***	0.117***	0.128***	0.127***	0.083***	0.106***
	(0.043)	(0.025)	(0.023)	(0.022)	(0.030)	(0.016)	(0.014)	(0.013)	(0.015)	(0.022)
High school	0.060	0.036	0.062***	0.135***	0.235***	0.261***	0.265***	0.244***	0.226***	0.227***
	(0.037)	(0.029)	(0.023)	(0.032)	(0.051)	(0.016)	(0.014)	(0.011)	(0.016)	(0.020)
Vocational school	0.021	-0.055	0.045	0.098***	0.117*	0.227***	0.260***	0.279***	0.262***	0.254***
	(0.045)	(0.043)	(0.033)	(0.032)	(0.061)	(0.018)	(0.015)	(0.013)	(0.016)	(0.021)
University	0.280***	0.222***	0.387***	0.651***	0.916***	0.672***	0.685***	0.670***	0.633***	0.687***
	(0.085)	(0.060)	(0.066)	(0.092)	(0.142)	(0.018)	(0.016)	(0.013)	(0.016)	(0.024)
Tenure (10^{-2})	0.575	0.661**	1.163***	1.735***	1.326***	4.430***	4.747***	5.292***	4.893***	4.563***
	(0.438)	(0.326)	(0.253)	(0.283)	(0.347)	(0.237)	(0.197)	(0.173)	(0.208)	(0.278)
Tenure squared	-3.836**	-2.431*	-3.148***	-4.253***	-2.879**	-7.336***	-8.675***	-10.881***	-10.432***	-10.405***
	(1.842)	(1.285)	(0.916)	(1.072)	(1.333)	(0.890)	(0.776)	(0.578)	(0.722)	(0.915)
Age										
25-35	-0.003	0.016	0.015	0.054**	0.153***	0.091***	0.106***	0.123***	0.130***	0.117***
	(0.045)	(0.029)	(0.022)	(0.024)	(0.044)	(0.020)	(0.021)	(0.016)	(0.020)	(0.028)
35-45	-0.025	0.018	0.043*	0.072**	0.121**	0.081***	0.094***	0.086***	0.102***	0.126***
	(0.053)	(0.031)	(0.025)	(0.028)	(0.048)	(0.023)	(0.023)	(0.019)	(0.022)	(0.032)
45-	0.047	0.050	0.076***	0.144***	0.300***	0.139***	0.174***	0.197***	0.247***	0.322***
	(0.051)	(0.033)	(0.027)	(0.032)	(0.054)	(0.026)	(0.024)	(0.023)	(0.028)	(0.037)
Married	0.023	-0.003	0.014	0.051*	0.005	-0.032	-0.055**	-0.096***	-0.089***	-0.073***
	(0.056)	(0.029)	(0.023)	(0.028)	(0.042)	(0.025)	(0.022)	(0.017)	(0.019)	(0.025)
Head of household dummy	0.084*	0.069***	0.061***	0.036	0.061*	0.141***	0.153***	0.186***	0.198***	0.177***
	(0.049)	(0.025)	(0.020)	(0.024)	(0.036)	(0.021)	(0.018)	(0.015)	(0.017)	(0.024)
Rural	-0.222***	-0.171***	-0.128***	-0.094***	-0.080***	-0.055***	-0.009	0.003	0.013	0.010
	(0.038)	(0.024)	(0.017)	(0.021)	(0.030)	(0.021)	(0.013)	(0.013)	(0.013)	(0.021)
λ	0.432***	0.385***	0.215***	0.110	0.129	-1.136***	-1.151***	-1.135***	-1.205***	-1.079***
	(0.102)	(0.083)	(0.067)	(0.077)	(0.131)	(0.057)	(0.050)	(0.044)	(0.053)	(0.072)
λ^2	-0.059	-0.045	0.011	0.061*	0.077	0.541***	0.567***	0.541***	0.549***	0.413***
	(0.036)	(0.033)	(0.026)	(0.033)	(0.061)	(0.051)	(0.044)	(0.038)	(0.044)	(0.059)
Constant	3.920***	4.230***	4.538***	4.785***	5.002***	4.547***	4.788***	5.088***	5.451***	5.770***
	(0.040)	(0.028)	(0.030)	(0.019)	(0.038)	(0.025)	(0.022)	(0.019)	(0.025)	(0.034)
Observations	8,036	8,036	8,036	8,036	8,036	23,623	23,623	23,623	23,623	23,623

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4.5B: 2002 Wage regression for males (without correction)

	Informal					Formal				
VARIABLES	q10	q25	q50	q75	q90	q10	q25	q50	q75	q90
Secondary school	0.121***	0.075	0.088***	0.152***	0.180***	0.214*	0.224***	0.204***	0.184***	0.159***
	(0.034)	(0.096)	(0.029)	(0.045)	(0.031)	(0.113)	(0.015)	(0.012)	(0.017)	(0.021)
High school	0.183***	0.143***	0.167***	0.225**	0.351***	0.432***	0.443***	0.404***	0.392***	0.370***
	(0.041)	(0.028)	(0.030)	(0.107)	(0.042)	(0.031)	(0.014)	(0.012)	(0.015)	(0.016)
Vocational school	0.180***	0.124***	0.194***	0.191*	0.328***	0.435***	0.474***	0.484***	0.484***	0.448***
	(0.047)	(0.041)	(0.033)	(0.115)	(0.076)	(0.029)	(0.184)	(0.015)	(0.016)	(0.018)
University	0.514***	0.448***	0.671***	1.033***	1.297***	0.968***	0.950***	0.920***	0.889***	0.923***
	(0.057)	(0.064)	(0.069)	(0.156)	(0.112)	(0.031)	(0.014)	(0.012)	(0.014)	(0.019)
Tenure (10^{-2})	0.401	0.195	0.735***	1.380*	0.839**	3.694***	4.729***	5.705***	5.920***	5.702***
	(0.769)	(0.306)	(0.270)	(0.815)	(0.386)	(0.848)	(0.459)	(0.172)	(0.236)	(0.324)
Tenure squared	-3.392	-1.335	-1.917*	-3.730	-1.644	-4.851	-8.306***	-11.661***	-12.841***	-12.911***
	(3.620)	(1.318)	(1.082)	(3.216)	(1.600)	(3.853)	(1.890)	(0.568)	(0.746)	(1.132)
Age										
25-35	0.071*	0.082***	0.043**	0.068	0.146***	0.171***	0.175***	0.199***	0.198***	0.193***
	(0.040)	(0.030)	(0.020)	(0.150)	(0.039)	(0.058)	(0.018)	(0.018)	(0.022)	(0.031)
35-45	0.102**	0.110***	0.084***	0.122**	0.174***	0.233***	0.212***	0.222***	0.228***	0.243***
	(0.049)	(0.034)	(0.025)	(0.060)	(0.043)	(0.048)	(0.021)	(0.020)	(0.025)	(0.036)
45-	0.073	0.082**	0.069***	0.170**	0.301***	0.202***	0.215	0.261***	0.279***	0.343***
	(0.047)	(0.034)	(0.026)	(0.079)	(0.052)	(0.044)	(0.183)	(0.023)	(0.029)	(0.042)
Married	0.064	0.041	0.041*	0.081	0.043	-0.016	-0.021	-0.064***	-0.064***	-0.087***
	(0.049)	(0.032)	(0.024)	(0.108)	(0.042)	(0.079)	(0.026)	(0.019)	(0.021)	(0.025)
Head of household dummy	0.107**	0.096***	0.099***	0.058**	0.093**	0.177***	0.192***	0.205***	0.231***	0.268***
	(0.051)	(0.033)	(0.019)	(0.027)	(0.038)	(0.055)	(0.022)	(0.015)	(0.019)	(0.024)
Rural	-0.228***	-0.182***	-0.143***	-0.099	-0.098***	-0.070*	-0.034**	-0.032***	-0.039**	-0.024
	(0.034)	(0.021)	(0.018)	(0.135)	(0.030)	(0.036)	(0.016)	(0.011)	(0.016)	(0.023)
Constant	3.917***	4.277***	4.581***	4.808***	5.059***	4.061***	4.301***	4.605***	4.934***	5.280***
	(0.031)	(0.023)	(0.018)	(0.105)	(0.026)	(0.049)	(0.018)	(0.017)	(0.015)	(0.025)
Observations	8,036	8,036	8,036	8,036	8,036	23,623	23,623	23,623	23,623	23,623

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4.6A: 2010 Wage regression for males (with correction)

	Informal					Formal				
VARIABLES	q10	q25	q50	q75	q90	q10	q25	q50	q75	q90
Secondary school	0.029	0.012	0.019	0.014	0.045	0.042***	0.054***	0.063	0.059***	0.024
	(0.030)	(0.019)	(0.017)	(0.022)	(0.029)	(0.013)	(0.011)	(0.051)	(0.011)	(0.016)
High school	-0.031	-0.041*	-0.045*	-0.017	0.081*	0.087***	0.134***	0.189***	0.219***	0.237***
	(0.038)	(0.025)	(0.024)	(0.030)	(0.047)	(0.016)	(0.012)	(0.039)	(0.012)	(0.017)
Vocational school	-0.069*	-0.025	-0.045*	-0.041	0.045	0.099***	0.122***	0.175***	0.220***	0.232***
	(0.039)	(0.028)	(0.027)	(0.032)	(0.044)	(0.014)	(0.011)	(0.025)	(0.014)	(0.016)
University	-0.149**	-0.068	0.078	0.269***	0.479***	0.477***	0.626***	0.693***	0.763***	0.803***
	(0.062)	(0.050)	(0.060)	(0.066)	(0.142)	(0.019)	(0.013)	(0.080)	(0.012)	(0.019)
Tenure (10^{-2})	1.439**	0.469	0.212	0.013	-0.410	2.329***	2.839***	3.605***	3.416***	2.851***
	(0.582)	(0.329)	(0.336)	(0.379)	(0.528)	(0.152)	(0.133)	(0.562)	(0.149)	(0.205)
Tenure squared	-2.881	-0.502	1.135	1.871	3.374	0.743	-1.499***	-4.660***	-4.879***	-4.558***
	(2.425)	(1.428)	(1.618)	(1.699)	(2.877)	(0.538)	(0.456)	(1.314)	(0.550)	(0.768)
Age										
25-35	0.105***	0.093***	0.075***	0.064**	0.114***	-0.016	-0.003	0.028	0.068***	0.099***
	(0.037)	(0.024)	(0.021)	(0.031)	(0.035)	(0.020)	(0.014)	(0.022)	(0.015)	(0.021)
35-45	0.104**	0.105***	0.068***	0.067*	0.160***	0.012	0.042***	0.061	0.134***	0.198***
	(0.048)	(0.029)	(0.023)	(0.036)	(0.040)	(0.023)	(0.015)	(0.037)	(0.018)	(0.023)
45-	0.176***	0.130***	0.096***	0.141***	0.251***	0.087***	0.115***	0.162***	0.248***	0.323***
	(0.050)	(0.030)	(0.023)	(0.035)	(0.045)	(0.025)	(0.017)	(0.062)	(0.021)	(0.027)
Married	0.014	0.036	0.051**	0.071**	0.026	0.001	-0.020	-0.054*	-0.054***	-0.057***
	(0.038)	(0.027)	(0.022)	(0.028)	(0.038)	(0.019)	(0.017)	(0.033)	(0.012)	(0.020)
Head of household dummy	-0.024	-0.012	0.028	0.024	0.067*	0.061***	0.081***	0.120***	0.142***	0.154***
	(0.039)	(0.025)	(0.019)	(0.024)	(0.037)	(0.016)	(0.014)	(0.011)	(0.010)	(0.017)
Rural	-0.154***	-0.122***	-0.091***	-0.090***	-0.041	-0.004	0.004	0.011	0.020**	0.030**
	(0.027)	(0.018)	(0.014)	(0.019)	(0.028)	(0.014)	(0.009)	(0.013)	(0.010)	(0.013)
λ	0.381***	0.219***	0.092	0.016	-0.071	-1.086***	-0.980***	-0.981***	-0.888***	-0.900***
	(0.118)	(0.074)	(0.090)	(0.086)	(0.111)	(0.063)	(0.044)	(0.154)	(0.048)	(0.076)
λ^2	-0.032	0.007	0.044	0.086**	0.120***	0.653***	0.614***	0.630***	0.557***	0.553***
	(0.042)	(0.025)	(0.035)	(0.033)	(0.045)	(0.065)	(0.044)	(0.089)	(0.046)	(0.078)
Constant	4.246***	4.559***	4.874***	5.150***	5.363***	4.988***	5.129***	5.317***	5.493***	5.736***
	(0.056)	(0.034)	(0.031)	(0.034)	(0.037)	(0.027)	(0.016)	(0.080)	(0.019)	(0.029)
Observations	6,745	6,745	6,745	6,745	6,745	28,201	28,201	28,201	28,201	28,201

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4.6B: 2010 Wage regression for males (without correction)

	Informal					Formal				
VARIABLES	q10	q25	q50	q75	q90	q10	q25	q50	q75	q90
Secondary school	0.087***	0.055***	0.049***	0.070***	0.069***	0.122***	0.114***	0.114***	0.107	0.076
	(0.030)	(0.020)	(0.017)	(0.023)	(0.026)	(0.013)	(0.012)	(0.010)	(0.104)	(0.236)
High school	0.099**	0.049	0.042*	0.070**	0.158***	0.217***	0.235***	0.274***	0.302***	0.333***
	(0.039)	(0.066)	(0.025)	(0.030)	(0.047)	(0.014)	(0.027)	(0.012)	(0.037)	(0.045)
Vocational school	0.056	0.067	0.042*	0.091**	0.112**	0.241***	0.242***	0.278***	0.321*	0.343*
	(0.041)	(0.516)	(0.023)	(0.037)	(0.044)	(0.013)	(0.046)	(0.010)	(0.166)	(0.201)
University	0.095	0.164**	0.283***	0.555***	0.879***	0.694***	0.801***	0.858***	0.912***	0.964**
	(0.069)	(0.080)	(0.040)	(0.072)	(0.107)	(0.015)	(0.016)	(0.011)	(0.073)	(0.417)
Tenure (10^{-2})	1.353**	0.524	0.054	0.517	-0.346	1.992***	2.906***	3.768***	3.995**	3.343
	(0.564)	(1.064)	(0.309)	(0.409)	(0.442)	(0.152)	(0.854)	(0.137)	(1.756)	(2.715)
Tenure squared	-3.215	-0.461	1.572	0.047	4.327*	1.880***	-1.636	-4.635***	-6.329	-5.812
	(2.916)	(5.051)	(1.511)	(2.460)	(2.353)	(0.534)	(3.050)	(0.528)	(7.015)	(14.130)
Age										
25-35	0.161***	0.146***	0.128***	0.123***	0.134***	0.073***	0.068***	0.088***	0.116**	0.145
	(0.041)	(0.052)	(0.021)	(0.025)	(0.035)	(0.019)	(0.016)	(0.011)	(0.047)	(0.440)
35-45	0.196***	0.176**	0.133***	0.128***	0.198***	0.117***	0.108***	0.137***	0.183	0.247
	(0.048)	(0.077)	(0.024)	(0.034)	(0.039)	(0.022)	(0.020)	(0.014)	(0.130)	(0.463)
45-	0.152***	0.133	0.121***	0.157***	0.242***	0.127***	0.142***	0.187***	0.248***	0.324
	(0.049)	(0.104)	(0.025)	(0.033)	(0.046)	(0.023)	(0.033)	(0.017)	(0.048)	(0.460)
Married	0.036	0.048	0.068***	0.089***	0.044	0.021	-0.009	-0.035***	-0.055	-0.059
	(0.038)	(0.044)	(0.021)	(0.027)	(0.037)	(0.021)	(0.050)	(0.011)	(0.073)	(0.364)
Head of household dummy	0.035	0.036	0.045**	0.049*	0.115***	0.130***	0.136***	0.145***	0.181***	0.194
	(0.034)	(0.039)	(0.018)	(0.027)	(0.029)	(0.016)	(0.034)	(0.011)	(0.032)	(0.197)
Rural	-0.191***	-0.139***	-0.110***	-0.108***	-0.065**	-0.019	-0.002	-0.001	-0.009	0.001
	(0.026)	(0.037)	(0.014)	(0.021)	(0.027)	(0.012)	(0.018)	(0.009)	(0.049)	(0.209)
Constant	4.280***	4.602***	4.887***	5.127***	5.353***	4.556***	4.776***	4.988***	5.197***	5.440***
	(0.041)	(0.069)	(0.022)	(0.025)	(0.029)	(0.017)	(0.034)	(0.013)	(0.102)	(0.036)
Observations	6,745	6,745	6,745	6,745	6,745	28,201	28,201	28,201	28,201	28,201

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Figure 4.1: Decomposition results for female sample

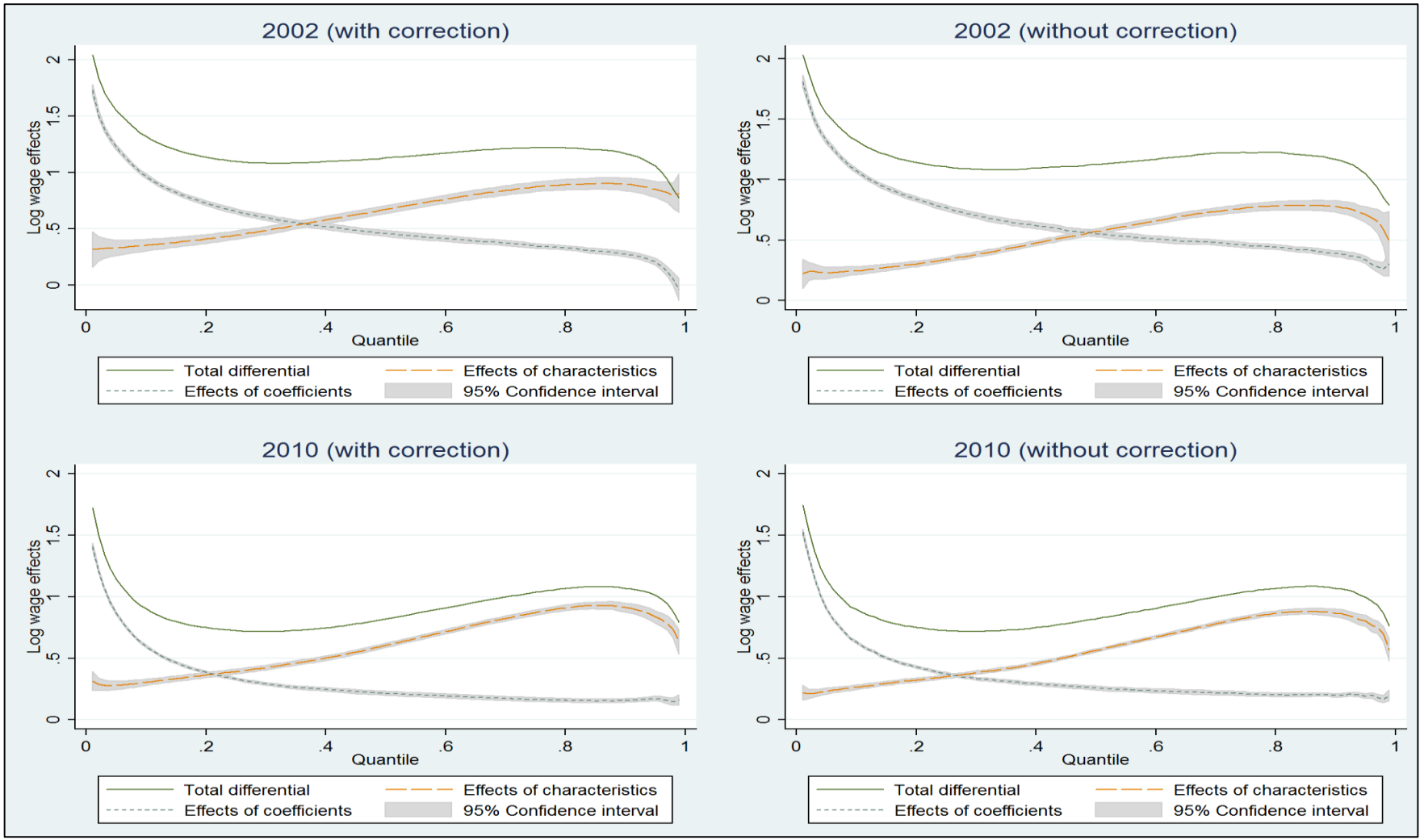
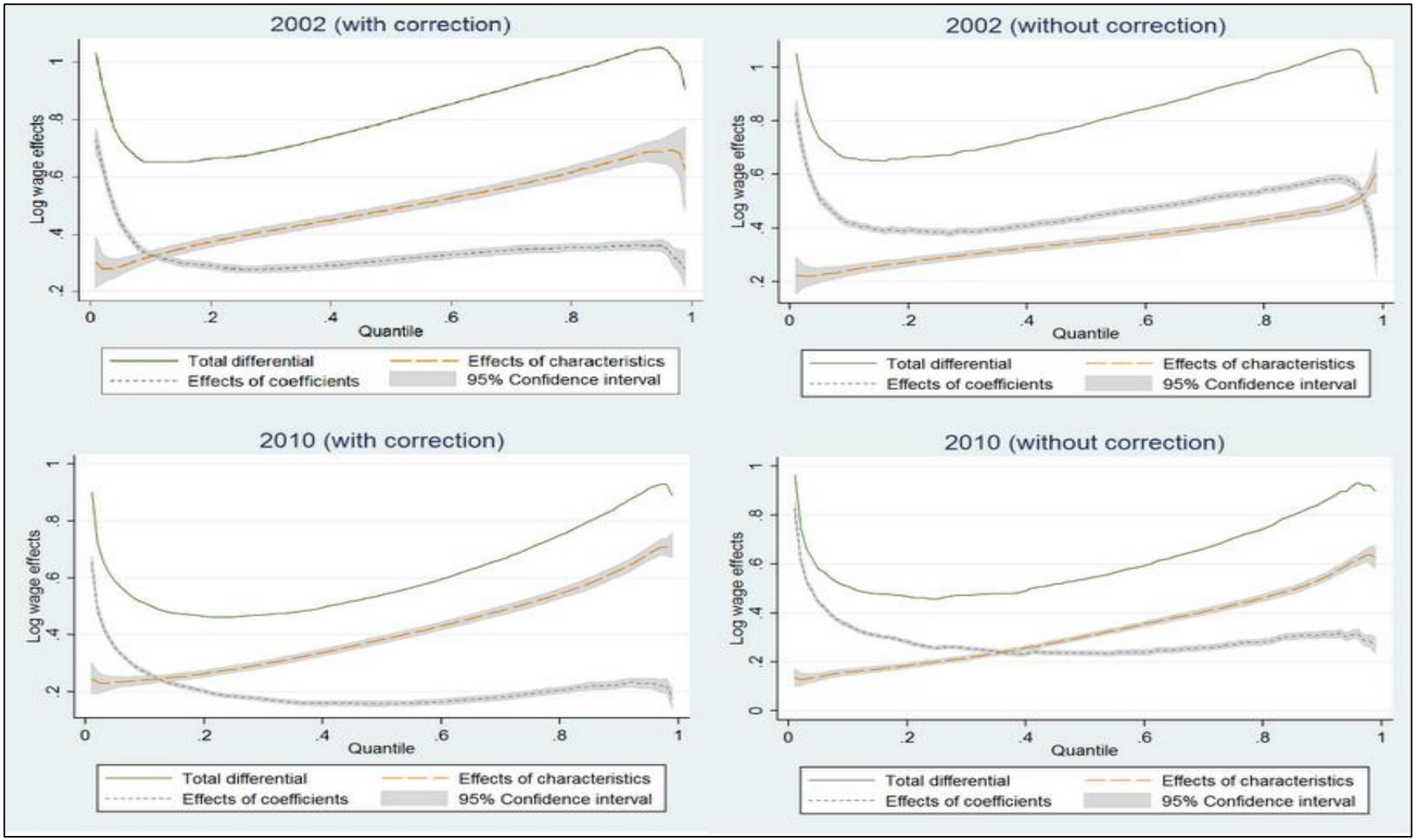


Figure 4.2: Decomposition results for male sample



APPENDIX TO CHAPTER 4

Table 4.A1: Exclusion criteria for the sample

	2002		**2010**	
	Frequency	%	Frequency	%
Full sample	300689		522171	
Dropped observations				
Non-wage earners	253589	84.3	429932	82.3
Age over 65 or less 20	4198	1.4	6480	1.2
Extreme values of experience (inconsistent with the age)	13	0.0	0	0
Has more than one job	688	0.2	2585	0.4
Extreme working hours (over 84 hours in a week)	276	0.1	754	0.1
Reported zero wages	787	0.3	4781	0.9
Divorced-widowed male	309	0.1	788	0.1
Random 60 sample (only for men)			23297	
Final sample	40829	13.6	53544	10.3

CHAPTER 5: CONCLUSION

Turkey has experienced major transformations in its course towards industrialisation. During the development process, it became more industrial and urbanised with increasing levels of education across the population. Over this time period, it also became more integrated with the world economy and created a stronger relationship with the western world both culturally and economically. On the other hand, the labour market in Turkey exhibits some interesting characteristics that are not observed in many other countries with the same income levels. In this context, the overall aim of this thesis was to examine and analyse three specific issues over the post 2000 period in which household level datasets have become available.

The first empirical study presented in Chapter 2 examined the decreasing wage inequality observed in the Turkish labour market with a particular focus on the effect of increasing educational attainment on the wage distribution. Two waves of the LFS for the years 2002 and 2010 were used in the analyses. To explore the underlying reasons behind the decreasing wage inequality, a two-step procedure was applied. In the first step, a selection equation was estimated to account for the possible selection bias due to non-random selection into wage employment. Due to the existence of unpaid family workers and the self-employed and the different characteristics of these individuals from those who simply do not participate in the work force, use of a simple probit model in the first step of the correction process can be misleading; just like labour force participation, the decisions about being in unpaid family work and self-employment are not random. Therefore, the

methodology proposed by Dubin and McFadden (1984) was exploited as this methodology accounts for the selection bias where the number of choices is more than two.

In the second step of the procedure, a decomposition methodology proposed by Lemieux (2002), which is a unified form of two commonly used decomposition techniques in the literature which were proposed by Juhn et al. (1993) and DiNardo et al. (1996), was applied. In this methodology, changes in the wage distributions are decomposed into three components, namely, the effect of returns to human capital (price of human capital), the effect of changes in the composition of human capital (changes in covariates) and changes in residual distribution (price and composition of unmeasured human capital characteristics). The changes in the residual wage inequality are further decomposed into two factors: the effect of changes in the distribution of unobservable characteristics and the effect of the pricing function of unobservable characteristics. Considering the differences in the wage determination processes, the analyses were performed separately for men and women.

The results of both the selection equations and the decomposition provide very important findings for understanding the changes in the labour market in Turkey. Consistent with the findings in the literature from other countries, the results of the selection equation showed that education increases the probability of choosing wage employment. Having other members of the household in the wage sector, and being the head of the household increases the probability of being in wage employment. Presence of children, being

married, and presence of self-employed members in the household all have a decreasing effect on wage employment. The results of the selection equation for men revealed that education, presence of children in the household, having other members of the household in the wage sector, and being head of the household increase the probability of choosing wage employment. Living in a rural area, and the number of unpaid family workers in the household all increase the likelihood of being self-employed. Arguably the most important finding of the selection equations is that, when the effect of education on the probability of being in wage employment for the years 2002 and 2010 is compared, it can be seen that education does not increase the probability of being a wage worker (as opposed to being economically inactive) in 2010 as much as it did in 2002.

The decomposition results that are obtained in the second step give the relative shares of the factors that affect the wage distributions. According to the decomposition results, the decreasing wage inequality observed in Turkey between 2002 and 2010 was mainly due to two factors: the decreasing between group inequality which is related to decreasing coefficients for education and the decreasing within group inequality due to the decreasing skill pricing function of unmeasured skills. On the other hand, the role of changes in the distribution of covariates is found to be negative and small. The results of the study suggest that increases in number of universities and the reform to the education system which increased the years of compulsory education, had a substantial effect on the wage distribution in Turkey between 2002 and 2010. Due to the increasing number of universities and reforms in the educational system, the supply of educated labour rapidly

increased in Turkey. However, steep decreases in the returns to education and an increasing share of university graduates in unemployment suggests that the labour market could not accommodate the increasing supply.

This study contributes to the existing literature two main ways. First, it analyses the impact of an increase in the 'education' of the labour force on wage inequality, as the literature mainly focuses on other aspects such as technology and labour market institutions. Second, it is the first application of the Lemieux (2002) model in a context of a middle income country.

The findings of the empirical analyses presented in Chapter 2 provide crucial suggestions for policy makers. Firstly, following the increase in the number of universities and reforms in the educational system, the supply of educated labour rapidly increased in Turkey. However, steep decreases in the returns to education and the increasing share of university graduates in unemployment are evidence that the labour market could not accommodate this increasing supply. Accordingly, wage inequality decreased as a result of relatively low demand for the educated workers. As noted in Chapter 2, the share of university graduates among the male unemployed has become higher than that for high school graduates since 2008. Moreover, according to the TurkStat employment statistics (TurkStat web database), the unemployment rate for university graduates which was around 6-8 per cent in 1990's became 11 per cent in 2010. This indicates that a large number of university graduates are facing the problem of unemployment. This leads to inefficiency since skilled labour, which

is a valuable resource, remains idle. To overcome the problem, policy makers could consider encouraging investments that create new employment opportunities and also R&D investments even though it should be acknowledged that such investments may lead to labour saving technologies.

Secondly, substantial change in the skill pricing function can be interpreted as a sign of quality differences between universities. The establishment of new universities under the "a university in each city'' project has been criticized by many authors (e.g. Dortlemez, 1995; Kaynar, 2005). It is argued that the scientific requirements and maintenance of minimum standards have been ignored as politicians used the stimulating effects of universities on local economies in their election campaigns. Decreasing quality of university education has two crucial economic outcomes; it reduces the productivity of the educated labour below its potential and causes an inefficient allocation of resources - resources which are quite limited in many developing countries. Therefore, in addition to issues regarding accessibility, policymakers should give importance also to the quality of education when decisions are made on education policies.

In the second empirical study presented in Chapter 3, using data from the Household Structure Survey (HSS) for the year 2006, the relationship between domestic violence, social norms, and female autonomy in the household is analysed. Female autonomy has been empirically found to affect the labour force participation of women in Turkey. Domestic violence is included in the analyses not only because it is argued to have an

effect on female autonomy but also because of its monetary and non-monetary costs on the economy. Finally, social norms are included as they affect both female autonomy and domestic violence. To analyse the relationship between domestic violence and female autonomy in the household, an Item Response Theory (IRT) model is applied.

In this setting, female autonomy and social norms are assumed to be latent variables as their true values are unknown yet there are some indicators which measure them with some error. Accordingly, the model design consists of two parts, namely, the measurement model in which the effect of the latent variables on their indicators is modelled, and the structural model in which the relationship between the observed and latent variables is examined. According to the results of the measurement model, women having a say on the choice of entertainment/holiday, relationships with the relatives and neighbours, in other words people outside the household, are the best indicators of female autonomy in the household. On the other hand, having a say in the house setup, is the weakest among the indicators of the female autonomy. This result is not surprising since house setup, or housework, is traditionally accepted as the duty of women. In terms of the husband's social norms, indicators showing husbands' agreement on "Having a male child increases the prestige of the wife" and "Continuation of the family is guaranteed only by a son" have greater factor loadings. This means that superiority attached to a male child (or may be male as a gender in general) by the husbands seems to be the strongest indicator of social norms of husbands.

The results of the structural model explain the key relationships that are the focus of the analyses. First, domestic violence has a negative effect on female autonomy. This finding suggests that women who are victims of domestic violence have less autonomy in the decision making process, they are forced to be 'in line' with the preferences of the husband. This result is consistent with the findings of Eswaran and Malhotra (2011) who proposed a non-cooperative model of spousal violence in which women are, in the presence of violence, forced to allocate resources in accordance with the husbands' choice. Second, female autonomy has a significant and negative effect on domestic violence. In other words, the incidence of violence is less likely to occur once the autonomy of women, which is promoted by the income and the education they possess and reflected by their presence in the decision making process, gains acceptance. This result is in accordance with the feminist theory on violence which asserts that intimate violence stems from the inequalities within the household and it is used by husbands to maintain subordination of women. Finally, the strength of the husband's social norms on traditional gender roles positively affects the risk of violence. This result supports Heise (1998) who argues that adherence to rigid gender roles increases the likelihood of violence against women.

There are some potential limitations of the study. The first limitation is the cross-sectional nature of the data. As Kaplan et al. (2001) note, a cross-sectional design of a model with a feedback loop gives only a snapshot of an on-going dynamic process. In that sense, the model used in this study, controlling for husband's social norms, analyses the effect of female autonomy on violence by comparing the incidence of violence for women with

different levels of autonomy. Therefore it is not possible to analyse the effect of female autonomy on violence over time for the same women. In other words, the change in incidence of violence due to the increasing autonomy of an individual woman can only be analysed with panel data which is not available here. Second, although the SEM model accounts for the measurement errors of the indicators of the latent variables, the incidence of violence is likely to be underreported by the respondents given the intimate nature of the subject. Moreover, the dataset only contains information about physical violence; however, violence can take a sexual or an emotional form as well. Additionally, the dataset provides information about violence that occurred only in last 12 months.

Nevertheless, this study makes significant contributions to the existing literature on domestic violence and female autonomy. First, it analyses the two-way relationship between female autonomy and domestic violence within the same model. Second, exploiting the advantages of SEM it accounts for the endogeneity and measurement error problems and finally, it tests the importance of social norms at the individual level.

The results of the analyses presented in Chapter 3 provide many suggestions for policy makers. For example, education and income not only weaken the strength of traditional social norms for both men and women but also promote a higher level of female autonomy. Therefore, policy makers should consider focusing on education, particularly promoting the educational attainment of women. According to the Global Gender Gap Report 2013, the female labour force participation rate in Turkey is 30 per cent and this ratio puts Turkey in

the 123rd place out of 136 countries. As already mentioned in the first section, around 24 per cent of women, who do not participate in the labour force, report that they do not have autonomy regarding their participation decision. Therefore, policies targeting new employment opportunities for women should be accompanied by campaigns promoting female autonomy in cooperation with the civil movements and organisations that support empowerment of women's socioeconomic status.

The third empirical analysis of the thesis presented in Chapter 4 examines the structure of the informal sector employment in Turkey. To do so, the differences between the informal and formal sector wages are decomposed along the entire distribution. A decomposition procedure which not only allows for correction of the potential self-selection problem in wage estimations but also decomposes the informal/formal sector wage gap at any quantile of interest, is applied. Using this method, the wage differentials between these two sectors are decomposed into two factors, namely, the characteristic effect and the coefficient effect. In the context of the formal/informal sector wage gap, a relatively strong coefficient effect is an indicator of market segmentation as it shows that workers in the informal sector earn less even after controlling for differences in their characteristics. In contrast, if the characteristic effect is stronger, it means that the wage gap is mainly due to differences between human capital endowments, not because the two sectors have different wage determination processes.

The results of the empirical analysis suggest a heterogeneous structure for the informal sector in Turkey in which the lower tier consists of workers who involuntarily participate in the informal sector and the higher tier includes the workers who voluntarily choose to be in the informal sector. It is also found that the wage gap between the formal and informal sectors is higher and the segmentation at the bottom end of the wage distribution is more spread out (i.e. observed in a wider range of quantiles) for females than it is for males.

The main limitation of the analysis is that, since there is no data on their earnings, the self-employed are not included in the estimation. Nevertheless, it is not expected that this exclusion will affect the generality of the results, especially for females since the share of self-employed is very low in total female employment. For instance, Tansel and Kan (2012) include the self-employed in their analysis as well as employees and find a heterogeneous structure for the informal sector employment in Turkey. Their results, both for men and women, are consistent with the findings of this chapter. In that sense, exclusion of the self-employed, at least qualitatively, does not seem likely to have affected the findings of the chapter.

Notwithstanding this limitation, this research contributes to the previous literature by providing a more detailed analysis of segmentation in the Turkish labour market. The results of the analyses are consistent with the findings of Tansel and Kan (2012) who also find a heterogeneous structure in the informal sector. However, the results of the analyses show that taking the selection into account substantially changes the results of the wage

regressions and where in the wage distribution the segmentation between the formal and informal sectors occurs. It is found that returns to higher levels of education, especially at the lower quantiles, disappear in the informal sector once the selection term is introduced into the wage estimation. It is also found that the decomposition results overestimate the spread of segmentation over the higher quantiles when the selection term is omitted. On the other hand, when the correction term is included, it is seen that the spread of segmentation, particularly for males, is much narrower. The results are also consistent with the findings of Baskaya and Hulagu (2011), who document a higher formal and informal wage gap for females than for males. In this study, it is also shown that the gap for females is not only higher at the mean wages, but also at all quantiles of the wage distribution.

The findings of the chapter provide useful information on informal sector employment in Turkey for policy makers. For instance, the results of the analysis show that the informal sector workers who are at the bottom of the wage distribution involuntarily participate in the informal sector. Due to the unregulated nature of the informal employment, these workers do not enjoy the protection from exploitation such as wages below the minimum wage, working in risky environments, and working with unsafe equipment. Moreover, these individuals do not have employment benefits such as health insurance. To eliminate such problems, policy makers should consider more efficient ways to detect the firms which avoid registering their employees. On the other hand, it is also found that some of the workers in the informal sector voluntarily choose to be in the informal sector. Therefore, before formulating a policy which targets elimination of informality, it is

important to understand the motivations of these individuals. For instance, the results of the selection equation show that individuals, women in particular, who have children (especially those who have young children) are more likely to be in the informal sector. A possible explanation is that these individuals may choose to work in the informal sector to avoid the rigid working hours in the formal sector in order to spare more time to take care of their children. Accordingly, policy makers could focus on increasing availability of childcare assistance for working parents.

In conclusion, the empirical studies presented in this thesis provide several interesting insights into three specific issues related to problems in the Turkish labour market. In that sense, the results of the thesis can be useful for policy makers as they provide information on the changes in the Turkish labour market and the underlying reasons behind the problems that are observed in it. Moreover, this thesis contributes to the literature on wage inequality, female autonomy and domestic violence, and the structure of informal sector employment by providing empirical evidence from an upper middle income country.

REFERENCES

Acemoglu, D. (2002). "Technical Change, Inequality, and the Labor Market", *Journal of Economic Literature*, 40: 7-72.

Aizer, A. (2010). "The Gender Wage Gap and Domestic Violence", *American Economic Review*, 100(4) (September): 1847-1859.

Albrecht, J., van Vuuren, A. and Vroman, S. (2009). "Counterfactual Distributions with Sample Selection Adjustments: Econometric Theory and an Application to the Netherlands", *Labour Economics*, 16(4): 383-396.

Altinay, A., and Arat, Y. (2009). *Violence against Women in Turkey.* İstanbul, Turkey: Punto.

Anderson, S. (2007). "The Economics of Dowry and Bride Price", *Journal of Economic Perspectives*, American Economic Association, 21(4):151-174.

Andrews, D. W., and Schafgans, M. M. (1998). "Semiparametric Estimation of the Intercept of a Sample Selection Model." *The Review of Economic Studies, 65*(3): 497-517.

Arap, K. S. (2010). "Türkiye Yeni Üniversitelerine Kavuşurken: Türkiye'de Yeni Üniversiteler ve Kuruluş Gerekçeleri", *SBF Dergisi* 65(1): 1-29 (in Turkish).

Arulampalam, W., Bhaskar, A., and Srivastava, N. (2012). *Maternal Autonomy: A New Approach to Measuring an Elusive Concept* Warwick: University of Warwick.

Arulampalam, W., Bhaskar, A., and Srivastava, N. (2015). *Does Greater Autonomy among Women Provide the Key to Better Child Nutrition*? Warwick: University of Warwick.

Ballon, P. (2011). "A Structural Model of Female Empowerment", in P. Ballon, *Model-based Multidimensional Poverty Indices: Theoretical Construction and Statistical Properties*. Doctoral Dissertation, Thesis No. 759, University of Geneva, ch. 3.

Bargain, O., and Kwenda, P. (2010). "Is Informality Bad? Evidence from Brazil, Mexico and South Africa", *IZA Discussion Paper*, No. 4711, IZA, Bonn, January.

Baskaya, Y. S., and Hulagu, T. (2011). "Informal-Formal Worker Wage Gap in Turkey: Evidence from a Semi-Parametric Approach", Central Bank of the Republic of Turkey Working Papers No.1115.

Becker, G. S. (1965). "A Theory of the Allocation of Time", *Economic Journal*, 75: 493-519.

Becker, G. S. (1973). "A Theory of Marriage: Part I", *Journal of Political Economy*, 81(4): 813-846.

Becker, G. S. (1981). *A Treatise on the Family*. Cambridge, MA: Harvard University Press.

Berman, E., and Machin, S. (2000). "Skill-Biased Technology Transfer around The World", *Oxford Review of Economic Policy*, 16(3): 12-22.

Berman, E., Bound, J., and Griliches, Z. (1994). "Changes in the Demand for Skilled Labor within U.S. Manufacturing: Evidence from the Annual Survey of Manufactures", *The Quarterly Journal of Economics,* 109 (2) (May): 367–397.

Berman, E., Bound, J., and Machin, S. (1997). *Implications of Skill-Biased Technological Change: International Evidence*. National Bureau of Economic Research No. w6166.

Blinder, A. S. (1973). "Wage Discrimination: Reduced Form and Structural Estimates", *The Journal of Human Resources,* 8 (4): 436–455.

Bound, J., and Johnson, G. E. (1992). "Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations." *American Economic Review,* 82(3) (June): 371–392.

Bourguignon, F., Fournier, M., and Gurgand, M. (2007). "Selection Bias Corrections Based on the Multinomial Logit Model: Monte Carlo Comparisons", *Journal of Economic Surveys*, 21(1): 174-205.

Bradley, K., and Khor, D. (1993). "Toward an Integration of Theory and Research on the Status of Women", *Gender and Society*, 7:347-378.

Buchinsky, M. (1998). "The Dynamics of Changes in the Female Wage Distribution in the USA: A Quantile Regression Approach", *Journal of Applied Econometrics,* 13(1): 1-30.

Buvinic M., Morrison A.R. and Shifter M. (1999). *Violence in the Americas: A Framework for Action in Too Close to Home: Domestic Violence in the Americas,* Washington DC: Inter-American Development Bank, Morrison A.R. Biehl M. L. (eds.)

Calvo-Armengol, A., and Jackson, M. O. (2004). "The Effects of Social Networks on Employment and Inequality", *American Economic Review*, 94(3):426-454.

Card, D., and Krueger, A. B. (1992). "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States", *Journal of Political Economy*, 100(1) (February): 1–40.

Card, D., Lemieux, T., and Riddell, W. C. (2003). "Unionization and Wage Inequality: A Comparative Study of the US, The UK, And Canada", National Bureau of Economic Research, No. w9473.

Carneiro, F. G., and Henley, A. (2001). "Modelling Formal vs. Informal Employment and Earnings: Micro-Econometric Evidence for Brazil", Annals of the XXIX National Meeting of Economics of ANPEC.

Caselli, F. (1999). "Technological Revolutions", *American Economic Review*, 89(1) (March): 78–102.

Chzhen, Y., and Mumford, K. (2011). "Gender Gaps Across the Earnings Distribution for Full-Time Employees in Britain: Allowing for Sample Selection", *Labour Economics*, *18*(6): 837-844.

Cukadar, S., Gultekin, G., Celik S. and Kahvecioglu, K. (2011). "Türkiye'de Üniversite Kütüphaneleri: Mevcut Durum ve Gelecek". *Uluslararasi Yükseköğretim Kongresi: Yeni Yönelisler ve Sorunlar*. 3: 2426-2439 (in Turkish).

Dahl, G. B. (2002). "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets", *Econometrica*, Econometric Society, 70(6): 2367-2420.

Dave, A., and Solanki, G. (2000). "Special Cell for Women and Children: A Research Study on Domestic Violence". In *A summary report of four records studies* (pp. 25–33). Washington, DC: ICRW.

David, H., Katz, L. F., and Kearney, M. S. (2005). *Rising Wage Inequality: The Role Of Composition and Prices*. National Bureau of Economic Research, No. w11628.

Dayioglu, M., and Kirdar, G.M. (2010). "A Cohort Analysis of Women's Labor Market Participation in Turkey", paper presented at the 17th Annual Conference of ERF, Antalya, Turkey.

De Soto, H. (1989). *The Other Path*, New York: Basic Books.

Dickens, W.T., and Lang, L. (1985). "A Test of Dual Labor Market Theory", *American Economic Review*, 75(4): 792-805.

DiNardo, J., Fortin, N. M., and Lemieux, T. (1996). "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica,* 64(5) (September): 1001-1044.

Dobash, R. E., and Dobash, R. (1979). *Violence against Wives*. New York: The Free Press.

Dörtlemez, Ö. (1995). *Yeni açılan üniversiteler*. İstanbul: Sarmal Yayınevi

Dubin, J. A., and McFadden, D. L. (1984). "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption", *Econometrica,* 52: 345-362

Dyson, T., and Moore, M. (1983). "On Kinship Structure, Female Autonomy, and Demographic Behaviour in India", *Population and Development Review,* 9(1): 35-60.

Elveren, A. Y., and Galbraith, J. K. (2008). "Pay Inequality in Turkey in the Neo-Liberal Era, 1980-2001", *European Journal of Comparative Economics,* 6(2) (December): 177–206.

Eng, S., Li, Y., Mulsow, M., and Fischer, J. (2010). "Domestic Violence against Women in Cambodia: Husband's Control, Frequency of Spousal Discussion, and Domestic Violence Reported by Cambodian Women", *Journal of Family Violence,* 25:237-246.

Eswaran, M. (2002). "The Empowerment of Women, Fertility and Child Mortality: Towards a Theoretical Analysis", *Journal of Population Economics*, 15(3): 433–545.

Eswaran, M., and Malhotra, N. (2011). "Domestic Violence and Women's Autonomy in Developing Countries: Theory and Evidence", *Canadian Journal of Economics/Revue canadienned'économique*, 44(4): 1222-1263.

Ettner, S. L. (1996). "The Opportunity Costs of Elder Care", *Journal of Human Resources,* 31(1):189-205.

Farmer, A., and Tiefenthaler, J. (1997). "An Economic Analysis of Domestic Violence", *Review of Social Economy,* 55(3), 337-358.

Ferman, L. A., Henry, S., and Hoyman, M. (1987). "Issues and Prospects for the Study of Informal Economies: Concepts, Research Strategies, and Policy", *The ANNALS of the American Academy of Political and Social Science*, 493(1): 154-172.

Fernandez, R., and Fogli, A. (2005). *Culture: An Empirical Investigation of Beliefs, Work, and Fertility*. National Bureau of Economic Research, No. w11268.

Fields, G. (2005). "A Guide to Multisector Labor Market Models", *Social Protection Discussion Paper Series*, No 0505, World Bank, Washington, D.C., April.

Galiani, S., and Sanguinetti, P. (2003). "The Impact of Trade Liberalization on Wage Inequality: Evidence From Argentina", *Journal of Development Economics*, 72(2): 497-513.

Galor, O., and Tsiddon, D. (1997). "Technological Progress, Mobility, and Economic Growth." *American Economic Review*, 87 (3) (June): 363–382.

Gelles, R. J. (1980). "Violence in the Family: A Review of Research in the Seventies", *Journal of Marriage and Family*, 42(4): 873-885

Gibson-Davis, C. M., Magnuson, K., Gennetian, L. A., and Duncan, G. J. (2005). "Employment and the Risk of Domestic Abuse among Low-Income Women", *Journal of Marriage and Family*, 67(5): 1149-1168

Gindling, T. H. (1991). "Labor Market Segmentation and the Determinants of Wages in the Public, Private-Formal and Informal Sectors in San-Jose, Costa Rica", *Economic Development and Cultural Change*, 39(3): 585

Goksel, İ. (2013). "Female Labor Force Participation in Turkey: The Role of Conservatism", *Women's Studies International Forum*, 41:45-54

Goode, W. J. (1971). "Force and Violence in the Family." *Journal of Marriage and Family* 33 (4):624–636.

Grandjean, P., Weihe, P., White, R. F., Debes, F., Araki, S., Yokoyama, K. and Jørgensen, P. J. (1997). "Cognitive Deficit in 7-Year-Old Children with Prenatal Exposure to Methylmercury", *Neurotoxicology and Teratology*, 19(6): 417-428,

Greaves, L. (1995). *Selected Estimates of the Costs of Violence Against Women*. Centre for Research on Violence against Women and Children. London

Greenfeld et al. (1998). "Violence by Intimates: Analysis of Data on Crimes by Current or Former Spouses, Boyfriends, and Girlfriends", *Bureau of Justice Statistics Factbook*. Washington DC: U.S. Department of Justice.

Gudbrandsen, N. H. (2013). "Female Autonomy and Fertility in Nepal", *South Asia Economic Journal*, 14(1): 157-173.

Günther, I. and Launov, A. (2006). "Competitive and Segmented Informal Labor Markets", IZA Discussion Papers No. 2349.

Hanson, G. H., and Harrison, A. (1995). *Trade, technology, and wage inequality*. National Bureau of Economic Research, No. w5110.

Hanson, G. H., and Harrison, A. (1999). "Trade Liberalization and Wage Inequality in Mexico", *Industrial & Labor Relations Review*, 52(2): 271-288.

Hatakenaka, Sachi. (2006). *Higher Education in Turkey for 21st Century: Size and Composition*. World Bank. Retrived from: http://siteresources.worldbank.org/EXTECAREGTOPEDUCATION/Resources/444607-1192636551820/S._Hatakenakas_report_on_Higher_Education_in_Turkey_for_21st_Century_Nov_2006.pdf.

Heckman, J. J. (1976). "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models", *Annals of Economic and Social Measurement*, 5(4): 475-492.

Heckman, J. J. (1979). "Sample Selection Bias as a Specification Error", *Econometrica*, 47(l): 53-161.

Heise L. L, Pitanguy J. and Germaine A. (1994). "Violence against Women; The Hidden Health Burden", Discussion papers No 225 Washington DC: The World Bank.

Heise, L. L. (1998). "Violence against Women an Integrated, Ecological Framework", *Violence against Women*, 4(3): 262-290.

Hill, M. A., (1983). "Female Labor Force Participation in Developing and Developed Countries-Consideration of the Informal Sector", *The Review of Economics and Statistics*, 65(3):459-68.

Hindin, M. J., and Adair, L.S. (2002). "Who's at Risk? Factors Associated with Intimate Partner Violence in the Philippines", *Social Science and Medicine,* 55:1385–99.

Hofmeyr, J. (2002). "The Importance of Segmentation in the South African Labour Market", In *DPRU/FES Second Annual Conference on Labour Markets and Poverty in South Africa*.

Htun, M., and Weldon, S. L. (2012). "The Civic Origins of Progressive Policy Change: Combating Violence against Women in Global Perspective, 1975–2005", *American Political Science Review,* 106 (03): 548–569.

Huber, M., and Mellace, G. (2014). "Testing Exclusion Restrictions and Additive Separability in Sample Selection Models", *Empirical Economics,* 47(1): 75-92.

Hussmanns, R. (2004), "Measuring the Informal Economy: From employment in the Informal Sector to Informal Employment", ILO Working Paper No. 53, ILO, Geneva.

Ichimura, H. (1993). "Semiparametric Least Squares (SLS) and Weighted SLS Estimation of Single Index Models", *Journal of Econometrics*, 58: 71–120.

International Labour Organization, (2013). "Global Employment Trends 2013: Recovering From a Second Jobs Dip", Global Employment Trends Reports, International Labour Office, Economic and Labour Market Analysis Department.

Jejeebhoy, S. J. (1998). "Wife-Beating in Rural India: A Husband's Right? Evidence from Survey Data." *Economic and Political Weekly,* 33:855–62.

Jejeebhoy, S. J. (1999). *Women's Education, Autonomy, and Reproductive Behaviour: Experience from Developing Countries*. Oxford: Clarendon Press.

Jejeebhoy, S. J. and Cook, R.J. (1997). "State Accountability for Wife-Beating: The Indian Challenge", *Lancet,* 349:110–112.

Juhn, C., Murphy, K. M., and Pierce, B. (1993). "Wage Inequality and the Rise in Returns to Skill", *Journal of Political Economy,* 101:410-442.

Kabeer, N. (1999). "The Conditions and Consequences of Choice: Reflections on the Measurement of Women's Empowerment", UNRISD Discussion Paper No 108.

Kabeer, N. (2001). "Conflicts Over Credit: Re-evaluating the Empowerment Potential of Loans to Women in Rural Bangladesh", *World Development,* 29:63–84.

Kaplan, D., Harik, P. and Hotchkiss, L. (2001). "Cross-Sectional Estimation of Dynamic Structural Equation Models in Disequilibrium", in: R. Cudeck, S. du Toit and S. Sörbom (Eds.): *Structural Equation Modeling: Present and Future.* A Festschrift in honor of Karl Jöreskog. Lincolnwood.

Kapoor, S. (2000). *Domestic Violence against Women and Girls*. Innocenti Research Centre.

Katz, L. F., and Murphy, K. M. (1992). "Changes in Relative Wages, 1963-1987: Supply and Demand Factors", *The Quarterly Journal of Economics*, 107(1):35–78.

Kaya, E. (2010). *Quantile Regression and the Gender Wage Gap: Is there a Glass Ceiling in the Turkish Labor Market?* Retrieved from: SSRN:http://ssrn.com/abstract=2320411

Kaynar, M. and Parlak, I. (2005). *Her İle Bir Üniversite Türkiye'de Yükseköğretim Sisteminin Çöküşü*. Ankara: Paragraf Yayinevi (in Turkish).

Kim, K. and Cho, Y. (1992). "Epidemiological Survey of Spousal Abuse in Korea", 277–82 in *Intimate Violence: Interdisciplinary Perspectives*, ed. E.C. Viano. Washington, DC: Hemisphere.

Kishor, S., and Johnson, K. (2004). *Profiling Domestic Violence: a Multi-country Study*. MEASURE DHS+, ORC Macro.

Kizilirmak, B. (2003). "Explaining Wage Inequality: Evidence from Turkey", *Ankara University Working paper series* no.57. Retrieved from: www.politics.ankara.edu.tr/eski/dosyalar/tm/SBF_WP_55.pdf.

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling*. Guilford Press.

Knickmeyer, N., Levitt, H. M., Horne, S. G., and Bayer, G. (2004). "Responding to Mixed Messages and Double Binds: Religious Oriented Coping Strategies of Christian Battered Women", *Journal of Religion and Abuse*, 5(2): 29-53.

Knight, J., and Song, L. (2005). *Towards A Labour Market in China*, Oxford: Oxford University Press.

Kocacik, F., Kutlar, A., and Erselcan, F. (2007). "Domestic Violence Against Women: A Field Study in Turkey", *Social Science Journal,* 44: 698-720.

Koçer, R. G., and Visser, J. (2009). "The Role of the State in Balancing the Minimum Wage in Turkey and the USA", *British Journal of Industrial Relations*, *47*(2): 349-370.

Koenig, M. A., Ahmed, S., Hossain, M. B., and Mozumder, A. K. A. (2003). "Women's Status and Domestic Violence in Rural Bangladesh: Individual-And Community-Level Effects", *Demography*, 40(2): 269-288.

Korkmaz, E. (2001) "Kayıtdışı İstihdam", *Ankara: İşveren Dergisi*. TİSK Yayını, Şubat Sayısı (in Turkish).

Krugman, P.R. (2008). "Trade and Wages, Reconsidered", *Brookings Papers on Economic Activity,* 39(1): 103-154.

Krusell, P., Ohanian, L. E., Ríos-Rull, J. V., and Violante, G. L. (2000). "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis", *Econometrica,* 68(5): 1029–1053.

Lamichhane, P., Puri, M., Tamang, J., and Dulal, B. (2011). "Women's Status and Violence against Young Married Women in Rural Nepal", *BMC Women's Health*, 11:19.

Lee, D. S. (1999). "Wage Inequality in the U.S. During the 1980s: Rising Dispersion or Falling Minimum Wage?", *The Quarterly Journal of Economics*, 114(3): 977-1023.

Lee, L. F. (1983). "Generalized Econometric Models with Selectivity", *Econometrica*, 51: 507-512.

Lemieux, T. (2002). "Decomposing Changes in Wage Distributions: a Unified Approach." *Canadian Journal of Economics,* 35(4): 646-688.

Losby, J. L., Else, J. F., Kingslow, M. E., Edgcomb, E. L., Malm, E. T., and Kao, V. (2002). "Informal Economy Literature Review", *ISED Consulting and Research*.

Lozano, R. (1999). "No Title." In *Too Close to Home: Domestic Violence in the Americas*. Morrison a. Washington, D.C: Inter-American Development Bank.

Lundberg S., R. A. Pollak and Wales, T. (1997). "Do Husbands and Wives Pool their Resources? Evidence from the UK Child Benefit", *Journal of Human Resources*, 32: 463-480.

Machado, J., and Mata, J. (2005). "Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression", *Journal of Applied Econometrics*, 20: 445-465.

Machin, S., and Van Reenen, J. (1998). "Technology and Changes in Skill Structure: Evidence From Seven Oecd Countries." *The Quarterly Journal of Economics*, 113 (4): 1215–1244.

Maddala, G. S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge: Cambridge university press.

Magnac, T. (1991). "Segmented or Competitive labor Markets?", *Econometrica*, 59(1): 165-187.

Mahajan, A. (1990). "Instigators of Wife Battering.", 1–10 in *Violence against Women*, edited by S. Sood. Jaipur, India: Arihant.

Malhotra, A. and Mather, M. (1997). "Do Schooling And Work Empower Women in Developing Countries? Gender and Domestic Decisions in Sri Lanka". *Sociological Forum*, 12(4): 599-630.

Maloney, W. F. (1999). "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico", *The World Bank Economic Review*, 13(2): 275-302.

Maloney, W. F. (2004). "Informality Revisited", *World Development*, 32(7): 1159-1178.

Marcouiller, D., De Castilla, V. R., and Woodruff, C. (1997). "Formal Measures of the Informal-Sector Wage Gap in Mexico, El Salvador, and Peru". *Economic Development and Cultural Change*, 45(2): 367-392.

Marenzi, A., and Pagani, L. (2005). "The Impact of Elderly Parents on Labour Market Participation of Italian Women", *Rivista di Politica Economica*, 95(2): 155-190.

Martin, D. (1976). *Battered Wives*. San Francisco: Glide Publications.

Martin, S. L., Moracco, K. E., Garro, J., Tsui, A. O., Kupper, L. L., Chase, J. L., and Campbell, J. C. (2002). "Domestic Violence Across Generations: Findings from Northern India", *International Journal of Epidemiology*, 31(3): 560-572.

Martins, P. S., and Pereira, P. T. (2004). "Does Education Reduce Wage Inequality? Quantile Regression Evidence from 16 Countries", *Labour Economics*, 11(3): 355–371.

Mason, K. O. (1986). "The Status Of Women: Conceptual and Methodological Issues in Demographic Studies", *Sociological Forum*, 1(2): 284-300.

Mason, K. O. (1987). "The Impact of Women's Social Position on Fertility in Developing Countries", *Sociological Forum*, 2(4): 718-745

Maurer-Fazio, M., Connelly, R., Chen, L., and Tang, L. (2011). "Childcare, Eldercare, and Labor Force Participation of Married Women in Urban China, 1982–2000", *Journal of Human Resources*, 46(2): 261-294.

Melly, B. (2005). "Decomposition of Differences in Distribution Using Quantile Regression", *Labour Economics* 12 (4): 577–590.

Melly, B. (2006). "Estimation of Counterfactual Distributions Using Quantile Regression", *Review of Labor Economics*, *68*(4):543-572.

Menon, N., and Johnson, M.P. (2007). "Patriarchy and Paternalism in Intimate Partner Violence: A Study of Domestic Violence in Rural India" in *Recent Studies on Indian Women: Empirical Work of Social Scientists*, ed. K.K. Misra, and J.H. Lowry (Jaipur,India: Rawat Publications).

Mincer, J. A. (1974). *Schooling, Experience, and Earnings*. New York. National Bureau of Economic Research.

Muthén, B., du Toit, S.H.C. and Spisic, D. (1997). "Robust Inference Using Weighted Least Squares and Quadratic Estimating Equations in Latent Variable Modeling with Categorical and Continuous Outcomes". Retrieved from: http://www.statmodel.com/bmuthen/articles/Article_075.pdf

Naticchioni, P., Ricci, A., and Rustichelli, E. (2008). "Wage Inequality, Employment Structure and Skill-biased Change in Italy", *LABOUR*, 22(1): 27-51.

National Center for Injury Prevention and Control. (2003). *Costs of Intimate Partner Violence against Women in the United States.* Atlanta (GA): Centers for Disease Control and Prevention.

Neuman, S., and Oaxaca, R. L. (2004). "Wage Decompositions with Selectivity-Corrected Wage Equations: A Methodological Note", *The Journal of Economic Inequality*, 2(1): 3-10.

Oaxaca, R. (1973). "Male-Female Wage Differentials in Urban Labor Markets", *International Economic Review* 14 (3): 693-709.

O'Brien, J. E. (1971) "Violence in Divorce Prone Families", *Journal of Marriage and the Family,* 33(November): 692-698.

Ozkoc, H., Kiren, O.G. and Ucdogruk, S. (2011). "Decomposition of Income Inequality in Turkey", *Current Research Journal of Social Sciences* 3(2): 96-103.

Panda, P. and Agarwal, B. (2005). "Marital Violence, Human Development and Women's Property Status in India", *World Development*, Elsevier, 33(5): 823-850.

Perry, D. B. (1997). "Incubated in Terror: Neurodevelopmental Factors in the 'Cycle of Violence'" 124-148. In: *Children, Youth and Violence: The Search for Solutions* (J Osofsky, Ed.). Guilford Press, New York.

Pratap, S. and Quintin, E. (2006). "Are Labor Markets Segmented in Developing Countries? A Semiparametric Approach". *European Economic Review*, 50(7): 1817-1841.

Presser, H., and Sen, G. (2000). *Women's Empowerment and Demographic Processes: Moving beyond Cairo*. Oxford University Press.

Pritchett, L. (2001). "Where Has All the Education Gone?", *The World Bank Economic Review,* 15(3): 367–391.

Quisumbing, A. R., and Maluccio, J. A. (2000). *Intrahousehold Allocation and Gender Relations: New Empirical Evidence from Four Developing Countries*. Washington, DC: International Food Policy Research Institute.

Rahman, A. (1999). "Micro-Credit Initiatives for Equitable and Sustainable Development: Who Pays?", *World Development*, 27:67–82.

Rao, V. (1997). "Wife-Beating in Rural South India: A Qualitative and Econometric Analysis", *Social Science and Medicine,* 44:1169–80.

Robbins, D. J. (1996). "Evidence on Trade and Wages in the Developing World", OECD Development Centre Working Papers no.119.

Robinson S. (1976). "A Note on the U Hypothesis Relating Income. Inequality and Economic Development", *American Economic Review*, 66(3):437-440.

Rode, A. (2011) *Literature Review: Non-Unitary Models of the Household (Theory and Evidence)*: retrieved from: www.econ.ucsb.edu/~pjkuhn/Ec250A/StudentsPapers/RodeNonUnitaryModels.pdf

Roth, D. L., Wiebe, D. J., Fillingim, R. B., and Shay, K. A. (1989). "Life Events, Fitness, Hardiness, And Health: A Simultaneous Analysis of Proposed Stress-Resistance Effects", *Journal of Personality and Social Psychology*, 57(1): 136-142.

Ruffer, T. and J. Knight. (2007). *Informal Sector Labor Markets in Developing Countries*. University of Oxford.

Savage, M., Devine, F., Cunningham, N., Taylor, M., Li, Y., Hjellbrekke, J., and Miles, A. (2013). "A New Model of Social Class? Findings from the BBC's Great British Class Survey Experiment", *Sociology*, 47(2):219-250.

Schaubroeck, J. (1990). "Investigating Reciprocal Causation in Organizational Behavior Research", *Journal of Organizational Behavior,* 11: 17-28.

Schuler, S. R., Hashemi, S. M., and Badal, S. H. (1998). "Men's Violence against Women in Rural Bangladesh: Undermined or Exacerbated by Microcredit Programmes?", *Development in practice,* 8(2): 148-157

Schuler, S. R., Hashemi, S. M., Riley, A. P., and Akhter, S. (1996). "Credit Programs, Patriarchy and Men's Violence against Women in Rural Bangladesh", *Social Science and Medicine,* 43:1729-42.

Schultz, L. G. (1960). "The Wife Assaulter", *Journal of Social Therapy,* 6(2):103-111.

Schultz, T. W. (1975). "The Value of the Ability to Deal with Disequilibria", *Journal of Economic Literature,* 13(3): 827–846.

Snell, J. E., Rosenwald, R. J., and Robey, A. (1964). "The Wife Beater's Wife: A Study of Family Inter-Action", *Archives of General Psychiatry*, 11: 107-113.

Staggs, S. L., and Riger, S. (2005). "Effects of Intimate Partner Violence on Low-Income Women's Health and Employment", *American journal of community psychology*, 36(1-2), 133-145.

Stanley, C. (1992). *Domestic Violence: An Occupational Impact Study*. *Too Close to Home: Domestic Violence in the Americas*. Tulsa, OK: Domestic Violence Intervention Services.

State Planning Organization. (1967). "2nd Five Year Development Plan". Retrieved from http://www.kalkinma.gov.tr/Lists/Kalknma%20Planlar/Attachments/8/plan2.pdf (in Turkish).

State Planning Organization. (1970). *Research on Higher Education Report.* Ankara. (In Turkish).

State Planning Organization. (2006). "9th Five Year Development Plan". Retrieved from http://www.kalkinma.gov.tr/Lists/Kalknma%20Planlar/Attachments/1/Ninth%20Development%20Plan%202007-2013.pdf (in Turkish).

Stiglitz, J. E. (1976). "The Efficiency Wage Hypothesis, Surplus Labor, and the Distribution of Income in LDCs", *Oxford Economic Papers*, 28(2): 185-207.

Straus, M. A., Gelles, R. J., and Steinmetz, S. K. (1980). *Behind Closed Doors: Violence in the American Family*. Transaction Pub.

Straus, M. and Gelles, R. (1990). *Physical Violence in American Families*. New Brunswick: Transaction Press.

Straus, M. A. (1974). "Forward". In Richard J. Gelles, (Ed.) *The Violent Home: A Study of Physical Aggression Between Husbands and Wives*. Beverly Hills: Sage Publications.

Tannuri-Pianto, M. E., and Pianto, D. M. (2002). "Informal Employment in Brazil .A Choice at the Top and Segmentation at the Bottom: A Quantile Regression Approach", Department of Economics Working Paper no.236, University of Brasilia.

Tansel, A. (1999). "Formal versus Informal Sector Choice of Wage Earners and Their Wages in Turkey", Economic Research Forum Working Paper No. 9927.

Tansel, A. (2000). "Wage Earners, Self Employed and Gender in the Informal Sector in Turkey", Policy Research Report on Gender and Development No.24, The World Bank.

Tansel, A. (2005). "Public-Private Employment Choice, Wage Differentials and Gender in Turkey", *Economic Development and Cultural Change*, 53(2): 453-477.

Tansel, A. and Bodur, F. (2012) "Wage Inequality and Returns to Education in Turkey: A Quantile Regression Analysis", *Review of Economic Development*, 16(1): 107-121.

Tansel, A. and Kan, E.O. (2012). "The Formal/Informal Employment Earnings Gap: Evidence from Turkey", IZA Discussion Papers 6556, Institute for the Study of Labor (IZA).

Tauchen, H. V., Witte, A.D. and Long, S.K. (1991). "Domestic Violence: A Non-Random Affair," *International Economic Review*, 32(2): 491-511.

Taymaz, E. (2009). "Informality and productivity: Productivity differentials between formal and informal firms in Turkey", ERC Working Papers no.901, ERC -Economic Research Center, Middle East Technical University.

Telli, Ç., Voyvoda, E., and Yeldan, E. (2007). "Modeling General Equilibrium for Socially Responsible Macroeconomics: Seeking for the Alternatives to Fight Jobless Growth in Turkey." *METU Studies in Development,* 33(2): 321-340.

The General Directorate on the Status and Problems of Women. (2008). *Domestic Violence against Women in Turkey*. Ankara.

Turkish Central Bank. (2015). *Enflasyon Raporu.* Turkish Central Bank, retrieved from http://www.tcmb.gov.tr/wps/wcm/connect/TCMB+TR/TCMB+TR/Main+Menu/Yayinlar/Raporlar/Enflasyon+Raporu/2015/Enflasyon+Raporu+2015-I/ (in Turkish).

Turkish Prime Ministry. (1995). *Causes and Consequences of Domestic Violence Report*. Ankara (in Turkish).

UN. (1985). *World Conference to review and appraise the achievements of the United Nations Decade for Women: Equality, Development and Peace Nairobi.* Retrieved from http://www.un.org/womenwatch/daw/beijing/nairobi.html.

UN. (1993). *Declaration on the Elimination of Violence Against Women* retrieved from http://www.un.org/documents/ga/res/48/a48r104.htm.

Vyas, S., and Watts, C. (2009). "How Does Economic Empowerment Affect Women's Risk of Intimate Partner Violence in Low and Middle Income Countries? A Systematic Review of Published Evidence", *Journal of International Development*, 21(5): 577-602.

Walker, L. E. (1979). *The Battered Woman*. New York: Harper and Row, Publishers.

WHO. (1997). *Violence against Women: A Priority Health Issue*. Geneva. Retrieved from: http://www.who.int/gender/violence/vawpriority/en/

WHO. (2013). *Global and Regional Estimates of Violence against Women*. Retrieved from: http://www.who.int/reproductivehealth/publications/violence/9789241564625/en/.

Wilson, M., and Daly, M. (1993). "An Evolutionary Psychological Perspective on Male Sexual Proprietariness and Violence against Wives", *Violence and Victims,* 89: 271–294.

Wooldridge, J.M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT press.

World Economic Forum. (2013). "Global Gender Gap Report 2013". Retrieved from http://www3.weforum.org/docs/WEF_GenderGap_Report_2013.pdf

Yeldan, E. (2006). "Neoliberal Global Remedies: From Speculative-Led Growth to IMF-Led Crisis in Turkey", *Review of Radical Political Economics*, 38: 193–213.

Yllo, K. A. (1994). "Through a Feminist Lens: Gender, Power and Violence." In *Controversies in Family Violence*, ed. R. Gelles and D. Loseke. Sage Publications.

Yüksel-Kaptanoğlu, İ., Türkyılmaz, A. S., and Heise, L. (2012). "What Puts Women at Risk of Violence from Their Husbands? Findings from a Large, Nationally Representative Survey in Turkey", *Journal of Interpersonal Violence*, 27(14): 2743-2769.