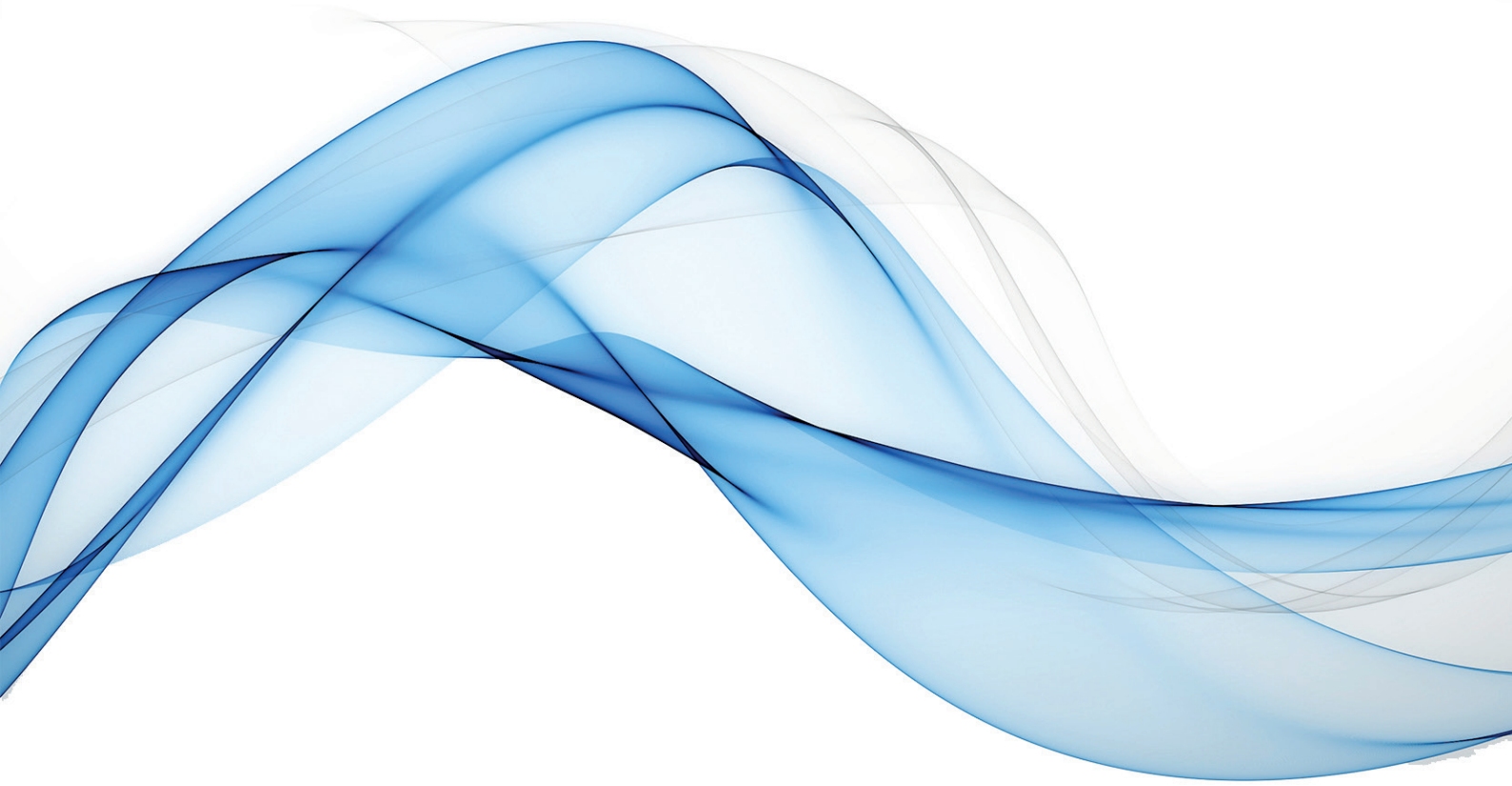
JOURNAL OF APPLIED MICROECONOMETRICS

HOLISTENCE publications

E-ISSN: 2791-7401
DOI: 10.53753/jame

International Peer-Reviewed and Open Access Electronic Journal



Volume: 1
Issue: 1
2021

E-ISSN: 2791-7401

DOI: 10.53753/jame

International Peer-Reviewed and Open Access Electronic Journal

Volume: 1

Issue: 1

Year: 2021

WEB: https://journals.gen.tr/jame

E-mail: jame@holistence.com

Address: Sarıcaeli Köyü ÇOMÜ Sarıcaeli Yerleşkesi No:29 D:119
Merkez-Çanakkale/TURKEY

ABOUT THE JOURNAL

The Journal of Applied Microeconometrics (JAME) is an internationally peer reviewed and periodical open access e-journal covering any issues in theoretical and applied microeconometrics. The journal also covers quantitative research in microeconomics. The main topicssubject areas of the journal are as follows:

- Estimation and Testing
- Robust Regression
- Discrete Choice Modelling
- Models for Count Data
- Duration Models
- Limited Dependent Variables
- Microeconometric Theory
- Bayesian Microeconometrics
- Big Data and Machine Learning in
- Microeconometrics
- Spatial Econometrics
- Non-Parametric Regression

JAME aims to serve as a platform for high quality research in applied microeconometrics. The scope of the Journal includes any papers dealing with identification, modelling, estimation, testing and prediction issues encountered in the analysis of individual-level data on the economic behavior of individuals or firms using regression methods for microeconometric data i.e. cross sectional data, repeated cross- sectional data, pool data, cohort and panel data etc. The journal also accepts case study articles written for both developing and developed countries.

We receive six types of submissions:

1- Regular submissions (open submissions)

2- A special issue

a) consisting of selected full length articles from national/international conferences, symposiums or workshops on microeconometrics.

b) by a guest editor invited by the editor in chief focusing on a specific area of research in microeconometrics.

3-Critiques of Previously Published Research Papers

4-Case Study Articles

5-Book Review

6-Letter to the Editor

The publication language of the journal is English.The Journal of Applied Microeconometrics is an open access electronic journal. Each paper published in the Journal is assigned a DOI® number, The journal is published 2 times a year as SUMMER (June) and WINTER (December) periods.

Send your manuscript to the editor at : https://journals.gen.tr/jame/user/register/

Owner

Holistence Publications

Contact
Adress: Sarıcaeli Köyü ÇOMÜ Sarıcaeli Yerleşkesi No:29, D.119, Merkez-Çanakkale / TURKEY

Tel: +90 530 638 7017

WEB : https://journals.gen.tr/jame

E-mail :jame@holistence.com
GSM 1: +1 (734) 747-4665 / WhatsApp
GSM 2: +90 530 638-7017 / WhatsApp

EDITORS

REFEREES IN THIS ISSUE

Hülya ÜNLÜ

Cankiri Karatekin University, TURKEY

Tuğba DAYIOĞLU

Nisantasi University, TURKEY

Tuğba SÖKÜT AÇAR

Canakkale Onsekiz Mart University /TURKEY

Özlem ERGÜT

Marmara University /TURKEY

Işın KIRIŞKAN

Giresun University / TURKEY

Muhammed Hanifi VAN

Van Yuzuncu Yıl University/TURKEY

Hale KIRER SILVA LECUNA

Bandirma Onyedi Eylul University /TURKEY

Nazan ŞAK

Osmaniye Korkut Ata University/TURKEY

Ayşe DEMİRHAN

Yildiz Technical University / TURKEY

Fatma ZEREN

Inonu University/ TURKEY

CONTENTS

"This page is left blank for typesetting"

JAME
E-ISSN:2791-7401
Volume : 1, Issue 1, 2021
URL: https://journals.gen.tr/jame

Received: 09.05.2021
Acccepted: 15.06.2021

RESEARCH ARTICLE

Has the gender wage gap narrowed or widened in a decade? Some recent evidence from the Turkish labor market

Fulden Kömüryakan[1] **Metehan Yılgör**[2]

1 Bandirma Onyedi Eylul University, Department of Econometrics, Turkey e-mail: fkomuryakan@bandirma.edu.tr

2 Bandirma Onyedi Eylul University, Department of Econometrics, Turkey e-mail: myilgor@bandirma.edu.tr

Abstract

The principal objective of this study is to determine the variation in the gender wage gap in the last decade of the Turkish labor market and to reveal possible factors driving the wage disparities. The data set covers the Household Budget Statistics surveys 2009 and 2018. In order to prevent biased results, the empirical strategy contains a two-stage model estimation and selectivity corrected decomposition approach. The findings claim a widening gender wage gap over the 10 years period. The portion of the gender wage gap resulting from the labor market discrimination tends to increase whereas the wage gap based on gender differences in characteristics decreases. Despite the decrease, if the female employees had the same characteristics as males, their mean wages would be higher. Moreover, the gender wage gap attributable to gender discrimination in the labor market continues to increase.

Keywords: *Mincer wage equation, gender wage gap, sample selection model, labor force participation, selectivity corrected decomposition, Turkey*

JEL Codes: *C31, C34, J16, J31*

Citation: KOMURYAKAN, F,. YILGOR, M., (2021). "Has the gender wage gap narrowed or widened in a decade? Some recent evidence from the Turkish labor market". Journal of Applied Microeconometrics (JAME). 1(1), 1-17, DOI: 10.53753/jame.1.1.01

Corresponding Author:
Fulden Kömüryakan
E-mail: fkomuryakan@bandirma.edu.tr

1. INTRODUCTION

The female labor force participation and the gender wage inequality are some of the most significant socio-economic issues that economies have to deal with. Females experience challenges to participate in the labor force for majority of labor markets. Even if they manage to participate, in spite of the efforts, females still struggle with a gender wage gap as a reason for occupational segregation, the patriarchal structure of society, gender norms set by the society, motherhood, long maternity leaves, and discrimination. Although most of the industrialized countries have acknowledged equal treatment for females in their labor markets, the gender wage gap, while declining in some countries, is still a longstanding issue of every labor market (Blau and Kahn 2003). Yet, policymakers may have concerns to confront the gender pay differences in favor of males (Kunze 2008).

Turkey is an emerging market and newly industrialized economy with a population of 83.4 million (International Monetary Fund 2021). According to the International Labour Force (ILO), the female employment rates (percent of female population ages 15-64) have increased between 2009 and 2018 from 27.81% to 38.15% in Turkey. Despite a slightly more than 10 percent increase in a decade, this rate has not still reached to any rates reported by the European Union (EU), Asia-Pacific, or the United States (US). The female labor force participation rate of Turkey is still at least 30% lower than the EU, Asia-Pacific, and the US. Nevertheless, surprisingly, male labor force participation which is reported as 78.39% in 2018 is almost similar or even greater than the EU, Asia-Pacific, the US economies (ILO 7.5.2021a).

Since the ratios and reports indicate a presence of wage disparity in the Turkish labor market, it becomes crucial to examine this socio-economic problem and provide information and policy recommendations to the policymakers. In this line, this study has two main objectives. The first main objective is to determine whether the gender wage gap tends to narrow or widen in a decade by employing cross-sectional data for 2009 and 2018. Determining the possible factors that drive the wage disparities in the Turkish labor market is the second main objective of this study. In line with these objectives, this study estimates an extended wage equation and decomposes the estimation results. In order to prevent biased results, the empirical strategy of this study takes into account the possible sample selection bias for both model estimation and decomposition approaches.

The contribution of this paper is to provide recent insights into the gender wage gap for the Turkish labor market and to determine the variation in the wage disparity by adopting a robust estimator in the case of sample selection bias. The empirical findings may help policymakers take precautions that are more accurate in order to reduce the gender wage gap in the Turkish labor market.

This paper is structured into six main sections. The next section provides an insight into the previous studies on the gender wage gap for both Turkish and other labor markets. The third section explains the data used in the study. The fourth section describes the methodological strategy. The fifth section presents the empirical findings. The last section concludes remarks and provides further discussion.

2. A BRIEF REVIEW OF THE PREVIOUS STUDIES

There exists a wide literature investigating the gender wage gap. This section contains a brief review of the studies that examine the gender wage disparity for both the Turkish and the other labor markets at different time periods via mostly decomposition methods.

2.1. The Gender Wage Gap in Labor Markets

Wright and Ermisch (1991) examine the gender wage gap using a decomposition method for the United Kingdom (UK) in 1980 and find that females would earn 20 percent more in the case of a non-discriminative labor market. Another study in the UK by Harkness (1996) reports that the gender wage gap has been narrowing by applying the decomposition method over the 1973-1993 period. Khitarishvili (2009) evaluates the gender wage gap in Georgia for the period of 2000 - 2004 by selectivity corrected model and decomposition methods. The results indicate a wider gender wage gap in Georgia than in other transition economies; however, the gender wage gap has narrowed in five years. Mysíková (2012) analyzes the gender wage disparity in some of the European labor markets. The study adopts the selectivity corrected model and decomposition methods and the findings indicate that the gender

wage gap is wider in Czech and Slovak Republics than in Hungary and Poland. Hinks (2002) is one of the first studies on the gender wage gap by race in South Africa. The selectivity corrected decomposition findings indicate a gender wage gap and the largest gap is between the Indian/Asian and colored and black population groups. Assaad and Arntz (2005) analyze the gender gap in wages for the Egyptian labor market and find that the gender gap is widened from 1988 to 1998.

Asian labor markets struggle with wage disparities just like the European and African labor markets. Liu (2004) examines the gender wage gap in the Vietnamese labor market over the period 1993 - 1998 and confirms a gender difference among wages in favor of males. Agrawal (2014) examines wage discrimination in India by adopting the selectivity corrected estimator and decomposition approaches. The results determine a gender wage gap based on the discrimination in the Indian labor market.

Fortin (2008) examines the gender wage gap in the context of psychological factors among young employees in the US via the corrected decomposition method. The findings show that the gender differences in the importance of money and work have a significant but modest effect on the gender wage gap. One study that performs a cross-country analysis for the gender wage gap and employment by considering sample selection bias and applying the decomposition method is by Olivetti and Petrongolo (2008). The results of the study determine a wider wage differential for several Organization for Economic Co-operation and Development (OECD) countries. However, the gender wage gap is smaller for economies where the gender employment gaps are low such as the UK, the US, and some EU countries.

The studies addressing the gender wage gap for European, African, Asian, and American labor markets provide evidence that most of the economies have wage differential in favor of males, and females still experience discrimination. Some of the labor markets, especially European, may have succeeded in narrowing the gender wage gap over time. However, the differences in wages by gender are still a serious socio-economic problem that economies have to deal with.

2.2. The Gender Wage Gap in the Turkish Labor Market

The gender wage gap is a serious socio-economic problem in the Turkish labor market as well. The findings of the limited studies addressing the gender wage gap in the Turkish labor force examined by mostly selectivity corrected model and decomposition methods are discussed below.

A study by Tansel (2005) is one of the first studies addressing the wage gap in the Turkish labor market. The study examines the public and private sector wages by gender using the 1994 Household Expenditure Survey data with selectivity corrected model and decomposition method. The findings claim a gender wage gap in the Turkish labor market and the wage gap is wider in the private sector, unlike the public administration. Kara (2006) analyses the occupational gender wage disparity by employing selectivity corrected estimator and decomposition methods using the 1994 Household Expenditure and Income Survey data. Parallel to the former study, the results show that the gender wage gap exists in the Turkish labor force, however, it tends to decrease with an increase in education. Another study that determines a gender wage gap in the Turkish labor market is by Ilkkaracan and Selim (2007). The study finds a wage disparity in favor of males by exploiting the decomposition method using the Employment and Wage Structure Survey data for 1995. Cudeville and Gurbuzer (2010) estimate the gender wage discrimination in the Turkish labor force by analyzing the Household Budget Statistics via selectivity corrected model and decomposition methods. The findings indicate a gender wage gap in the labor force and the gap is wider than the labor forces of Spain and Greece. One of the recent studies on the gender wage gap is by Tekguc et al. (2017). The study adopts a selectivity corrected model along with the decomposition and examines the Labor Force Survey for 2004 and 2011. The findings provide an insight into the presence of the gender wage gap and report that the gender wage gap becomes wider for 2011 compared to 2004.

3. DATA

Our study exploits the cross-sectional Turkish Household Budget Statistics (HBS) data conducted by the Turkish Statistical Institute (TURKSTAT). According to the TURKSTAT, the HBS data has been collected annually since

2002 in order to provide information about the socio-economic structures of households living in Turkey and to test the validity of the applied socio-economic policies. Therefore, the HBS data is one of the most important and reliable sources to examine socio-economic problems like the gender wage gap.

The households in HBS data are collected with a stratified two-stage cluster sampling method. In order to examine the variation of the gender wage gap for the recent decade in the Turkish labor force, this study employs 2009 and 2018 HBS data. The HBS data was conducted on 1,050 and 1,296 households that change every month between January 1st and December 31st in 2009 and 2018. 12,600 and 15,552 households have taken the surveys in 2009 and 2018, respectively. The HBS data contains 28,041 and 30,737 individuals aged above 15 in 2009 and 2018, respectively. Following the previous studies for both national and international labor markets, the data are restricted to those aged between 15 and 65 in order to take into account the retirement choices and to receive results that are more relevant. The restricted data contain 12,128 (48%) males and 13,125 (52%) females in 2009 whereas it contains 13,284 (49%) males and 13,872 (51%) females in 2018.

This study adopts secondary data containing no identifying information for any of the individuals in the surveys. The data that support the findings of this study are used under the license and not publicly available due to the privacy restrictions of TURKSTAT.

4. METHODOLOGICAL STRATEGY

Mincer wage equation, proposed in the pioneer study of Mincer (1974), is known as one of the most important equations to examine the wages in labor markets. The standard Mincer equation provides a better understanding of how wages are distributed across the population by explaining the logarithmic wages with human capital in terms of schooling, experience, and the square of the experience. This study examines the gender wage gap by estimating and decomposing the extended Mincer wage equation explained below for both genders in 2009 and 2018.

$$ln(W_i) = \alpha + \sum_{j=1}^{3} \beta_j EducationDummies_{ij} + \delta Experience_i + \gamma Experience_i^2 + \sum_{j=1}^{8} \emptyset_j OccupationGroups_{ij} + u_i \quad (1)$$

The dependent variable in Equation 1 is the logarithmic monthly wages and represented by . In order to analyze the effect of schooling more specifically, education is divided into four dummy variables as illiterate or compulsory education, high school, associate or bachelor's, and postgraduate degrees. Since the school starting age is 6 in Turkey, the experience variable is calculated by subtracting the age of 6 and the years of schooling from the age variable. The standard Mincer wage equation is extended with the occupation groups in this study. The occupation group variable set indicates the nine major occupation groups based on the International Standard Classification of Occupations (ISCO:08) of the ILO[1]. Lastly, and represent the constant and error terms respectively.

4.1. Heckman's Sample Selection Model

Mincer wage equation has been extended with different variables and estimated in literature via several different methods to examine different socio-economic structures. Some studies restrict their data to working employees who reported positive income in order to estimate the Mincer wage equation. However, Heckman (1979) states that this restriction may lead to sample selection bias. According to Heckman (1979), if the sample of working males and females who reported positive wages are not randomly selected, specification error occurs. In this case, the Ordinary Least Squares (hereafter OLS) estimator tends to be biased and inconsistent. In order to overcome this sample selection bias for non-randomly selected samples, Heckman (1979) proposed a selectivity corrected two-stage efficient estimator. The first stage of the sample-selection model contains an estimation of the probability of labor force participation via a probit model. The first stage also provides an estimation of another variable

1 For detailed occupation classification information, please see International Labour Force (7.5.2021b).

referred to as an inverse Mills ratio[2] (i.e., selectivity term). In the second stage, the extended Mincer wage equation is estimated with the Mills ratio estimated from the first stage of the sample-selection model. According to the Heckman procedure, the inverse Mills ratio is a proxy variable for Equation 2. Adding the inverse Mills ratio as another explanatory variable in Equation 1 makes it possible to measure the sample selection bias because of the individuals which are non-selected for the wage model (Dolton and Makepeace 1986). In the first stage, this study estimates the selection model below that indicates the probability of labor force participation

$$p_i = \Phi\Big(\alpha + \sum_{j=1}^{2} \gamma_j\, AgeCategories_{ij} + \sum_{j=1}^{3} \lambda_j\, EducationCategories_{ij} + \sum_{j=1}^{2} \delta_j MaritalStatus_{ij} + \beta_1 HouseholdSize_i + \beta_2 PresenceofPreschooler_i + \beta_3 ln(NonlaborIncome)_i\Big) \quad (2)$$

where denotes the conditional probability of labor force participation and denotes the cumulative standard normal distribution probability density function. The dependent variable of the model is the labor force participation that takes the values of 1 if the individual is employed. The explanatory variables are age categories, education dummies, marital status dummies, number of people in the household, the dummy variable that indicates the presence of preschoolers in the household, and logarithmic monthly nonlabor income. Monthly nonlabor income indicates income such as pension, interest yield, fund, scholarship, etc. other than wages. The selection of the explanatory variables included in the labor force participation model is based on previous studies (e.g., Korenman and Neumark 1992; motherhood, and wages. We find that heterogeneity leads to biased estimates of the "direct" effects of marriage and motherhood on wages (i.e., effects net of experience and tenureBudig et al. 2012; Mysíková 2012; Agrawal 2014; Glauber 2018).

4.2. Decomposition

As the decomposition method provides more detailed information about the differences and discrimination in different groups such as gender, race, occupation, age, etc., it has become one of the most preferred methods in economics to examine disparities. The very first studies introducing the decomposition method in economics are by Blinder (1973) and Oaxaca (1973). The standard decomposition method was proposed in the same year by Blinder and Oaxaca; therefore, it is referred to as O-B decomposition. Standard O-B decomposition[3] takes into account possible interactions and the relationships and measures the endowments and coefficients after estimating the models in this case, the extended Mincer wage equations for both males and females. Standard O-B decomposition decomposes the differences into two categories as endowment and coefficient effects. The endowment component indicates the portion of the differences attributable to the differences in characteristics (Blinder 1973). The coefficient effect is described as the quantity of the wage gap as a reason for gender differences in coefficients (Daymont and Andrisani 1984).

Standard O-B decomposition may also be affected by the sample selection bias explained by Heckman (1979). In order to overcome sample-selection bias in decomposition after the OLS, the decomposition method has been developed. After estimating the wage equations for both males and females , the selectivity corrected decomposition approach can be expressed as below

$$\left(\bar{W}_m - \bar{W}_f\right) = \hat{\beta}_m\left(\bar{X}_m - \bar{X}_f\right) + \bar{X}_f\left(\hat{\beta}_m - \hat{\beta}_f\right) + \left(\hat{\theta}_f\bar{\lambda}_f - \hat{\theta}_m\bar{\lambda}_m\right) \quad (3)$$

where represents the logarithmic monthly wages. contains the regressors explained in Equation 1, indicates the slope parameters and constant term, is the inverse Mills ratio explained above, and is an estimate of (Neuman and Oaxaca 2004)Heckit selectivity correction introduces some fundamental ambiguities in the context of wage decompositions. The ambiguities arise from group differences in the selection term which consists of a parameter

2 Mills ratio can be calculated as where indicates the matrix of the control variables in first stage model and and indicate the density and distribution functions. For more information, please see Heckman (1979).

3 For a detailed discussion for O-B decomposition, please see Appendix 1.

multiplied by the Inverse Mills Ratio (IMR. This study exploits the decomposition approach that takes into account the selection bias to overcome the bias that may occur in the standard O-B decomposition.

5. EMPIRICAL FINDINGS

This section is divided into three subsections. The first subsection contains the initial findings from the HBS data regarding the labor force participation and wage models explained above. The next subsections provide the estimation results of both OLS and Heckman's two-stage models. The last subsection contains the findings of selectivity corrected decomposition.

5.1. Initial Findings

Table 1 contains the descriptive statistics of the variables in the first and the second stage models explained in Equations 1 and 2.

Table 1. Initial Findings

	2009		2018	
	Male	**Female**	**Male**	**Female**
First stage model				
Labor force participation	0.5	0.16	0.52	0.2
Age categories				
Age < 30	0.35	0.38	0.32	0.31
30 ≤ Age < 45	0.34	0.33	0.33	0.34
45 ≤ Age ≤ 65	0.31	0.29	0.35	0.35
Education categories				
Illiterate or compulsory education	0.66	0.77	0.62	0.7
High school	0.22	0.15	0.22	0.17
Associate or bachelor's degree	0.11	0.07	0.15	0.12
Postgraduate degree	0.01	0.01	0.01	0.01
Marital status				
Married	0.69	0.69	0.66	0.68
Separated or widow	0.01	0.07	0.02	0.08
Single	0.3	0.24	0.32	0.24
Household size*	4.58	4.58	4.2	4.13
Presence of preschooler	0.3	0.3	0.26	0.26
Monthly nonlabor income* (*in Turkish Liras*)	376.03	148.13	512.29	216.04
No. of obs.	12,128	13,125	13,284	13,872
Second stage model				
Monthly wages* (*in Turkish Liras*)	930.82	645.03	2,452.39	1,978.89
Logarithmic monthly wages*	6.32	5.66	7.51	7.13
Human capital				
Illiterate or compulsory education	0.59	0.53	0.54	0.45
High school	0.25	0.22	0.24	0.2
Associate or bachelor's degree	0.15	0.23	0.2	0.32
Postgraduate degree	0.01	0.02	0.02	0.03
Experience	21.62	18.21	22.52	20.64

Occupation group				
ISCO1	0.06	0.03	0.04	0.04
ISCO2	0.08	0.14	0.09	0.19
ISCO3	0.08	0.1	0.07	0.06
ISCO4	0.06	0.14	0.06	0.12
ISCO5	0.15	0.15	0.19	0.25
ISCO6	0.07	0.07	0.05	0.03
ISCO7	0.19	0.08	0.19	0.06
ISCO8	0.15	0.05	0.15	0.04
ISCO9	0.16	0.24	0.16	0.21
No. of obs.	6,100	2,133	6,301	2,792

Notes: * denotes a continuous variable; other variables are dummies. The first stage model contains information on all individuals available in the data; the second stage model contains only employed individuals.

Table 1 shows that the labor force participation rates are 34 percent more for males in 2009 and the gap in the labor force participation has decreased only by 2 percent in a decade.

The common perception of society for females is such that they need to stay at home and take care of children. This perception may cause difficulties for females to be involved in the labor force. This perception may have been more common and acceptable in the past. However, it is expected to change over the years. Yet, interestingly, the labor force participation gap by gender has not been significantly changed over a decade in Turkey. According to the age categories, most of the individuals who take the survey are aged below 45. The education categories ratios of the first stage model (i.e., for both working and not working individuals) indicate that males are more educated than females for both 2009 and 2018. Most of the individuals are either illiterate or have only compulsory education set by the Turkish Ministry of National Education. However, this ratio is noticeably higher for females. Even though education degrees get higher in a decade, females are still less educated than males. Many papers have examined the girls' education in Turkey and most of them support that there is a significant gender gap in educational attainment because of the conservative attitudes in Turkey, especially in the east side (e.g., Tansel 2002; Rankin and Aytac 2006; Duman 2010; Caner et al. 2016). Marital status ratios show that most of the individuals are married and the ratios are similar for both genders. The ratios have modestly increased for the marital status of being separated or widow in a decade as a reason for the increase in divorce rates. The number of people in the household is around 4 and the rate of the presence of preschoolers is around 30 percent. Monthly nonlabor income has increased in a decade in Turkey and the monthly nonlabor income is higher for males in both 2009 and 2018.

According to the monthly wages, males earn more compared to females. In 2009, males earn more than the official minimum wage set by the Ministry of Labor and Social Security whereas females earn slightly less. Furthermore, the gap has not been narrowed in a decade and the females still earn slightly less than the minimum wage. Although the wage differences have been decreased over a decade, males still earn more compared to females. The ratios of education categories for the second stage model (i.e., only for employed individuals) indicate that employed individuals are more educated. Interestingly, employed females are more educated than employed males, especially in terms of associate, bachelors, and postgraduate degrees. As for the occupation groups, most of the female employees are employed in jobs without any qualification (ISCO9) in 2009. There has been a change in this ratio in a decade through an increase in education female employees are now mostly employed in the occupation group of service and sales workers (ISCO5) rather than the elementary jobs. Another interesting change that happened in a decade is that the ratio of the female managers (ISCO1) has become equal to the male managers and females are noticeably more employed in the occupation group of professionals (ISCO2) in 2018. Interestingly, the ratio of the male managers has decreased in a decade whereas the females' ratio has increased. Since some of the occupations are predominantly by males, males are significantly more employed in the occupation groups of craft and related trades workers and plant (ISCO7) and machine operators and assemblers (ISCO8), for both 2009 and 2018.

5.2. Heckman's Sample Selection Model Findings

Another advantage of the Heckman two-stage model is to provide information about wages and labor force participation decisions. Before analyzing the findings of the second stage Heckman's model, Table 2 provides information about the first stage labor force participation model of the Heckman procedure.

Table 2. Probit Model Findings

Dependent variable: labor force participation				
	2009		**2018**	
	Male	**Female**	**Male**	**Female**
Age categories				
Age < 30	0.4195*** (0.047)	0.3926*** (0.0506)	0.2424*** (0.0474)	0.1269*** (0.0469)
30 ≤ Age < 45	0.5545*** (0.0344)	0.6965*** (0.0406)	0.4782*** (0.1728)	0.4814*** (0.0341)
45 ≤ Age ≤ 65 (reference)				
Education categories				
Illiterate or compulsory education (reference)				
High school	0.2317*** (0.03)	0.4204*** (0.0376)	0.1728*** (0.0291)	0.4178*** (0.0354)
Associate or bachelor's degree	0.6182*** (0.042)	1.3166*** (0.0489)	0.5241*** (0.0353)	1.1982*** (0.0379)
Postgraduate degree	0.966*** (0.1695)	1.857*** (0.1957)	1.0992*** (0.1509)	1.6598*** (0.1244)
Marital status				
Married	0.9058*** (0.042)	-0.1902*** (0.0447)	0.8471*** (0.0422)	0.0885** (0.0443)
Separated or widow	0.8412*** (0.1055)	0.5271*** (0.0736)	0.5691*** (0.0837)	0.5621*** (0.0665)
Single (reference)				
Household size	-0.0593*** (0.0063)	-0.0749*** (0.0087)	-0.0933*** (0.0071)	-0.0449*** (0.0087)
Presence of preschooler	0.0609* (0.0321)	-0.1657*** (0.0378)	0.1139*** (0.0336)	-0.3362*** (0.0366)
Logarithmic nonlabor income (*in Turkish Liras*)	-0.1057*** (0.0049)	-0.0561*** (0.0071)	-0.1292*** (0.0044)	-0.066*** (0.0061)
Constant	-0.5746*** (0.0583)	-1.0955*** (0.0692)	-0.1628*** (0.0561)	-1.1092*** (0.0618)
Diagnostics statistics				
Mills ratio	0.6949*** (0.0547)	-0.4386*** (0.1141)	-0.4513*** (0.0331)	-0.4645*** (0.0825)
Wald Chi square	1675.98 [0.000]	552.66 [0.000]	2204.3 [0.000]	577.23 [0.000]
No. of obs selected	6,100	2,133	6,983	2,792
No. of obs nonselected	6,028	10,992	6,301	11,080

Notes: ***p<0.01, **p<0.05, *p<0.1. The numbers in brackets indicate the standard errors; the numbers in square brackets are the probabilities of the test statistics.

The labor force participation model of the Heckman two-stage estimator contains the aforementioned inverse Mills ratio and it is significant for both males and females for 2009 and 2018. This may indicate a sample-selection bias for both males and females in those years.

All age categories increase the probability of being involved in the labor force, according to those aged between 45 and 65. The individuals aged between 30 and 45 tend to take place in the labor force more for both males and females. The probability of labor force participation for males aged below 30 is more likely than females and this difference has increased in a decade. Females enter the labor market at a later age than males as a possible result of marriage and motherhood. Education is one of the most significant factors in obtaining jobs and individuals that are more educated are expected to have a higher probability to work. One of the most significant factors of females' labor force participation is education in Turkey and the labor force participation increases with the increase in educational attainment (Dayioglu and Kırdar 2010). In line with the other studies in the literature, all the education degrees have a positive impact on the labor force participation for both males and females, and education has noticeably more contribution on females' labor force participation. Marital status dummies show that married males tend to take place in the labor force more because of the idea of being the breadwinner of the household. Similar to the other labor markets (e.g. Sheran 2007; Lee et al. 2008), married females' probability of labor force participation is less. Although the impact of marriage is still significantly more for the likelihood of males' labor force participation, this structure has changed in a decade in favor of females and marriage does not decrease females' likelihood of working in 2018. According to the household structures, the increase in the number of people in the household has a negative contribution to the labor force participation for both genders.

The presence of children aged below 6 is one of the most significant factors in females' labor force participation. The presence of preschoolers dummy is significant for both males and females in 2009 and 2018. However, the contribution of it is positive on males' labor force participation whereas it is significantly negative on females'. Furthermore, these positive and negative effects have been increased in a decade instead of decreasing. Parallel to the literature, more monthly nonlabor income reduces the likelihood of working in 2009 and 2018 for males and females and this decrease is higher for the males.

Table 3 presents further estimation results of the extended Mincer wage equation via both OLS and Heckman's two-stage estimators.

Table 3. OLS and Heckman's Sample Selection Model Findings

Dependent variable: logarithmic monthly wages

	2009				2018			
	Male		Female		Male		Female	
	OLS	**Heckman**	**OLS**	**Heckman**	**OLS**	**Heckman**	**OLS**	**Heckman**
Human capital								
Illiterate or compulsory education (reference)								
High school	0.6363*** (0.0357)	0.5317*** (0.0382)	0.6328*** (0.0984)	0.4843*** (0.0967)	0.1964*** (0.0226)	0.1484*** (0.0236)	0.4626*** (0.0614)	0.3096*** (0.062)
Associate or bachelor's degree	1.0902*** (0.0524)	0.8523*** (0.0575)	1.3379*** (0.113)	0.9205*** (0.1597)	0.5382*** (0.032)	0.4023*** (0.0327)	0.8761*** (0.0703)	0.4806*** (0.0987)
Postgraduate degree	1.6841*** (0.1015)	1.3411*** (0.1466)	1.8068*** (0.1495)	1.2673*** (0.2714)	0.9695*** (0.0746)	0.7361*** (0.0784)	1.4491*** (0.0943)	0.9289*** (0.1552)
Experience	0.1077*** (0.0039)	0.081*** (0.0041)	0.0658*** (0.0073)	0.0588*** (0.0072)	0.0848*** (0.0029)	0.0676*** (0.0026)	0.0671*** (0.0047)	0.0582*** (0.0047)
Experience^2	-0.0017*** (0.00007)	-0.0012*** (0.0001)	-0.0011*** (0.0001)	-0.0009*** (0.0001)	-0.0015*** (0.00005)	-0.0012*** (0.00005)	-0.0011*** (0.0001)	-0.0009*** (0.00009)
Occupation group								
ISCO1	0.3908*** (0.0687)	0.4195*** (0.0639)	1.3167*** (0.1881)	1.31*** (0.1904)	0.6694*** (0.0558)	0.6886*** (0.0501)	1.0979*** (0.0869)	1.1016*** (0.1074)
ISCO2	0.6509*** (0.0701)	0.6477*** (0.0714)	1.0221*** (0.1342)	1.0286*** (0.1386)	0.6555*** (0.04459	0.6408*** (0.0427)	0.7253*** (0.0793)	0.7332*** (0.0776)
ISCO3	0.5096*** (0.0596)	0.4932*** (0.0603)	1.1055*** (0.1236)	1.1074*** (0.1199)	0.5054*** (0.0375)	0.4874*** (0.0406)	0.5806*** (0.0895)	0.5905*** (0.09)
ISCO4	0.4455*** (0.0599)	0.4533*** (0.0624)	0.9856*** (0.1128)	0.9731*** (0.1088)	0.3637*** (0.0368)	0.3535*** (0.04)	0.5446*** (0.0762)	0.5469*** (0.0737)
ISCO5	0.2718*** (0.0522)	0.2724*** (0.0468)	0.4455*** (0.0996)	0.4291*** (0.0909)	0.1586*** (0.0306)	0.1558*** (0.028)	0.1925*** (0.0574)	0.2116*** (0.0512)
ISCO6	-0.873*** (0.0662)	-0.8583*** (0.0572)	-1.0193*** (0.1278)	-1.007*** (0.1142)	-0.7249*** (0.0579)	-0.6914*** (0.0412)	-1.029*** (0.1244)	-0.9736*** (0.0981)
ISCO7	0.3288*** (0.0465)	0.3207*** (0.043)	-0.1541 (0.1424)	-0.1439 (0.1059)	0.2174*** (0.0297)	0.2015*** (0.0277)	-0.0047 (0.1084)	-0.0053 (0.0765)
ISCO8	0.4379*** (0.0472)	0.417*** (0.0456)	1.1181*** (0.1303)	1.1309*** (0.1341)	0.3145*** (0.0288)	0.2991*** (0.0295)	0.6536*** (0.0777)	0.6567*** (0.0919)
ISCO9 (reference)								
Constant	4.507*** (0.0594)	5.2661*** (0.0791)	4.04*** (0.1019)	4.7768*** (0.215)	6.2521*** (0.0421)	6.7333*** (0.0485)	5.6843*** (0.0802)	6.4738*** (0.1567)
R^2		-		-		-		-

Notes: ***p<0.01, **p<0.05, *p<0.1. The OLS results are only for the employed individuals. The numbers in brackets are the robust standard errors for the OLS.

The signs of the coefficients estimated with the OLS and two-stage estimators are the same; however, the impacts vary regarding the sample selection bias. In general, the OLS estimates the impacts of human capital and occupation groups higher than the selectivity corrected model. The OLS calculates the coefficients greater than they should be in the case of sample selection bias.

Concerning the education categories of Heckman's model, all education categories are significant and have a positive contribution to the wages of both males and females. Parallel to the previous studies in the literature (e.g., Kara 2006; Tansel and Daoud 2011; the returns to the education degrees on the wages are greater for the females whereas the returns to the experience are greater for males in 2018. Nevertheless, the contribution of high school and postgraduate degrees are slightly higher for males in 2009. The postgraduate degree has the highest incremental effect on the wages for both males and females. Table 3 also reveals that the returns to education have been decreased for both males and females over a decade. The coefficients of the occupation groups indicate that being employed as managers (ISCO1) has the most incremental effect on females' wages, according to being employed in elementary occupations (ISCO9), in 2009 and 2018 whereas males earn more as professionals (ISCO2) in 2009. Both male and female individuals employed as skilled agricultural, forestry, and fishery workers (ISCO6) earn less than the individuals employed in the elementary occupations.

5.3. Decomposition Findings

In order to examine the gender disparity in wages and its variation in a decade, this study adopts the selectivity corrected decomposition method, aforementioned above. Table 4 reveals the overall selectivity corrected decomposition results in 2009 and 2018. This study also calculates the standard O-B decomposition after the OLS for employed individuals to compare the results. The results are available in Appendix 2 and 3.

Table 4. Overall Selectivity Corrected Decomposition Results

	2009			2018		
	Coef.	**S.E.**	**%**	**Coef.**	**S.E.**	**%**
Male	6.7919***	0.0412		7.8027***	0.0238	
Female	0.9199***	0.0191		1.7066***	0.0257	
Difference	5.8718***	0.0454	100	6.0961***	0.0351	100
Endowments	1.199***	0.0458	20	0.7993***	0.0277	13
Coefficients	4.6728***	0.0712	80	5.2967***	0.0466	87

Notes: ***p<0.01, **p<0.05, *p<0.1. Coef. denotes the coefficients; S.E. denotes the standard errors.

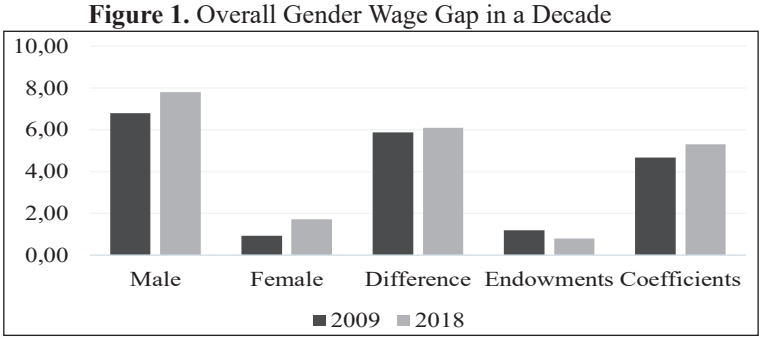
Table 4 shows that the mean logarithmic monthly wage is 6.79 and 0.92 for males and females in 2009 whereas it is 7.8 and 1.71 in 2018. Males' wages have increased by 1.01 in a decade whereas females' have increased only by 0.79. There is a gap of 0.22 in increase between males' and females' wages in the favor of males. The results provide evidence that there is a significant raw gender difference in wages for both 2009 and 2018. More interestingly, the gender wage gap has been increased from 5.87 to 6.1 in a decade, rather than decreasing.

One might compare the standard O-B decomposition results after OLS with the selectivity corrected decomposition results. Appendix 2 indicates that standard O-B decomposition measures the gender wage gap narrower than it actually is. After calculating the selectivity corrected decomposition, Table 4 indicates that the raw gender wage differential is wider in the Turkish labor market.

The raw differential in wages is divided into two categories as endowment and coefficient. All the categories are statistically significant and the coefficient component is larger than other categories in both 2009 and 2018. Endowment effects explain the 20 and 13 percent of the raw gender wage differentials whereas coefficient effects explain 80 and 87 percent of it. In other words, the gender wage gap attributable to the labor market discrimination has been increased in a decade whereas the wage gap attributable to the gender differences in human capital and occupation groups has been decreased.

Endowment effects have decreased from 1.2 to 0.8 in a decade and the difference in characteristics of employed males and females decreased by 0.4. Yet, employed males have better characteristics than employed females. If females had the same characteristics as males, their mean wages would increase by 1.2 and 0.8 in 2009 and 2018, respectively. Coefficients effects indicate the change in females' wages are 4.67 and 5.3 in 2009 and 2018 in the case of applying males' coefficients to females' coefficients. This result indicates that the gender wage gap arises from the discrimination in the labor market and the discrimination still increases. Figure 1 illustrates the statistics from Table 4.

Figure 1. Overall Gender Wage Gap in a Decade



Source: Authors' own calculation from the HBS data.

Figure 1 shows the mean wage differences among males and females and the widening gender wage gap in a decade. The presence of the gender wage gap is consistent with the results of other studies for both Turkish and other labor markets. However, there is a wider difference in wages in favor of males in the Turkish labor market than the others. Furthermore, while some of the developed and emerging economies may achieve to narrow the gender wage gap over the years, the gender wage gap has widened in a decade in the Turkish labor market. In order to evaluate the possible reasons for the determined gender wage gap, Table 5 presents the results of detailed selectivity corrected decomposition.

Table 5. Detailed Selectivity Corrected Decomposition Results

	2009		2018	
	Endowments	**Coefficients**	**Endowments**	**Coefficients**
Human capital				
Illiterate or compulsory education (reference)				
High school	0.0505*** (0.0049)	0.0609*** (0.0064)	0.0106*** (0.0019)	-0.0311*** (0.0069)
Associate or bachelor's degree	0.0692*** (0.0063)	0.0322*** (0.0045)	0.0308*** (0.0033)	-0.0471*** (0.0071)
Postgraduate degree	0.0094*** (0.0022)	0.0015** (0.0007)	0.0059*** (0.0014)	-0.0047*** (0.0015)
Experience	1.5127*** (0.0781)	-0.5563*** (0.0204)	1.2433*** (0.0498)	-1.0105*** (0.0325)
Experience^2	-0.6487*** (0.0468)	0.2496*** (0.0121)	-0.6544*** (0.0308)	0.4347*** (0.0205)
Occupation group				
ISCO1	0.0254*** (0.0041)	-0.0139*** (0.0019)	0.022*** (0.0023)	-0.0219*** (0.0024)
ISCO2	0.0348*** (0.0045)	-0.0863*** (0.0053)	0.0311*** (0.0032)	-0.1294*** (0.0066)
ISCO3	0.0289*** (0.0039)	-0.0578*** (0.0041)	0.0263*** (0.0027)	-0.0405*** (0.0035)
ISCO4	0.0189*** (0.003)	-0.0781*** (0.0048)	0.0146*** (0.002)	-0.0766*** (0.0048)
ISCO5	0.0329*** (0.0058)	-0.0552*** (0.0034)	0.0223*** (0.0041)	-0.1265*** (0.006)
ISCO6	-0.0513*** (0.0045)	-0.0154*** (0.0016)	-0.0307*** (0.0026)	-0.0113*** (0.0016)
ISCO7	0.0562*** (0.0077)	-0.0169*** (0.0016)	0.0356*** (0.0049)	-0.0259*** (0.0024)
ISCO8	0.0597*** (0.0068)	-0.0191*** (0.002)	0.0415*** (0.0043)	-0.0213*** (0.0023)
ISCO9 (reference)				
Constant	-	5.2277*** (0.0795)	-	6.4092*** (0.0509)

Notes: ***p<0.01, **p<0.05, *p<0.1. The numbers in brackets are the standard errors.

Endowment effects in Table 5 provide insights into how much the gender differences in characteristics of the individuals affect the gender wage gap. The gender differences in human capital explain 83 percent of the endowment effects whereas the gender differences in occupation groups explain 17 percent of it in 2009. As for 2018, the gender differences in human capital and occupation groups explain 80 and 20 percent of the endowment effects, respectively.

Males have better educational characteristics for both 2009 and 2018 and the gender differences in all education degrees contribute to the gender wage gap, as expected. Associate or bachelor's degree has the highest impact on widening gender wage gap. Moreover, the contributions of the education degrees to the gender wage gap have been decreased in a decade along with the increase in females' education. The gender differences in the experience of males and females wider the gender wage gap in 2009 and 2018, however, the contribution has slightly decreased in 2018. The endowment effects of occupation groups indicate that the gender differences in occupation groups wider the gender wage gap, except for the occupation group of skilled agricultural, forestry, and fishery workers (ISCO6). The gender differences in occupation groups that mostly male-dominated such as craft and related trades workers (ISCO7) and plant and machine operators and assemblers have (ISCO8) the highest contribution to the gender wage gap both for both 2009 and 2018.

6. CONCLUSION AND DISCUSSION

The principal aims of this study are to determine the variation in the wage disparities in the last decade of the Turkish economy and examine the possible reasons females earning less in the Turkish labor market. First stage results indicate that females need to be supported in participating in the labor market at early ages, especially before 30. Since the presence of preschoolers prevents females to work, nursery and financial supports may be provided. Education is one of the most crucial factors for females to be employed. Policymakers should specifically continue to attach more importance to girls' education in Turkey. The second stage findings determine that return to the education degrees on the wages have been decreased for both males and females in a decade. Recently, employers started to give less weight to education degrees with the noticeable increase in education degrees. Interestingly, the selectivity corrected decomposition results claim a widening the gender wage gap in favor of males around 0.23 over a decade, rather than a decrease. The findings indicate that the gender wage gap attributable to the labor market discrimination has been increased over a decade whereas the wage gap attributable to the gender differences in characteristics has been decreased. Another finding is that although the difference in characteristics of employed males and females is decreased by 0.4, male employees still have better characteristics than female employees. If the female employees had the same characteristics in human capital and occupation groups as males, their mean wages would be about 0.8 higher in 2018. Moreover, the gender wage gap attributable to the gender discrimination against females in the labor market continues to increase after a decade.

In order to prevent the wage disparities in terms of gender and to achieve equality in payments, new policies, campaigns, and acts may be organized by the authorities. This may help to reduce the effects of this serious socio-economic problem. Recently, some universities in Turkey have been started to add mandatory social gender equality lectures to the bachelor's degree curriculum. Expanding these kinds of lectures and educate the children and the younger individuals in a way of gender equality may help to lower the socio-economic disadvantages in the long term. Supporting the young females to take place in the labor force, ensuring the companies, especially in the private sector, hiring a similar rate of male and female employees may be another effective anti-discrimination policy recommendations. Some of the countries have been established specific ministries to help to reduce gender inequalities in terms of improving females' status across all social and economic areas. One of the first countries that established such a ministry is Indonesia. The Indonesia Ministry of Women Empowerment and Child Protection was established in 1978. South Korea established the Ministry of Gender Equality and Family in 2001. Another country that established a Gender Equality ministry is Spain. The Spain Ministry of Equality was established in 2010. Canada and Argentina established the Ministry of Women and Gender Equality and the Ministry of Women, Genders and Diversity in 2018 and 2019, respectively. The main objective of the Ministries is to improve the females' social and economic status through the enhancement of females' rights over the country. By following the leading countries, establishing a Ministry of Gender Equality may help to reduce the disadvantages that females are facing in Turkey.

This study could provide a clear understanding for policymakers by determining the change of the gender wage gap and the factors that wider it on how to design efficacious gender equality policies in terms of wages.

Funding

The authors declare that this study has no financial support.

Conflict of Interest

Metehan Yılgör and Fulden Kömüryakan declare that they have no conflict of interest.

Submission declaration statement

We confirm that this work is original and has not been published elsewhere, nor is it currently under consideration for publication elsewhere.

Data Availability Statement

This study adopts secondary data containing no identifying information for any of the individuals in the survey. The data that support the findings of this study are used under the license and not publicly available due to the privacy restrictions of TURKSTAT. However, data are available with the permission of the TURKSTAT.

REFERENCES

- AGRAWAL, T. (2014). Gender and Caste-Based Wage Discrimination in India: Some Recent Evidence. *Journal for Labour Market Research* 47(4), 329–340.
- ASSAAD, R., & ARNTZ, M. (2005). Constrained Geographical Mobility and Gendered Labor Market Outcomes under Structural Adjustment: Evidence from Egypt. *World Development.* 33(3), 431–454.
- BLAU, F.D., & KAHN L.M. (2003). Understanding International Differences in the Gender Pay Gap. *Journal of Labor Economics.* 21(1), 106–144.
- BLINDER, A.S. (1973). Wage Discrimination : Reduced Form and Structural Estimates. *The Journal of Human Resources.* 8(4), 436–455.
- BUDIG, M.J., MISRA J., & BOECKMANN I. (2012). The Motherhood Penalty in Cross-National Perspective: The Importance of Work-Family Policies and Cultural Attitudes. *Social Politics.* 19(2), 163–193.
- CANER, A., GUVEN, C., OKTEN, C., & SAKALLI, S.C. (2016). Gender Roles and the Education Gender Gap in Turkey. *Social Indicators Research.* 129(3), 1231–54.
- CUDEVILLE, E., & GURBUZER, L.Y. (2010). Gender Wage Discrimination in the Turkish Labor Market: Can Turkey Be Part of Europe. *Comparative Economic Studies.* 52(3), 429–463.
- DAYIOGLU, M., & KIRDAR, M.G. (2010). Determinants of and Trends in Labor Force Participation of Women in Turkey. Welfare and Social Policy Analytical Work Program Working Paper Number 5. State Planning Organization of the Republic of Turkey and World Bank.
- DAYMONT, T.N., & ANDRISANI, P.J. (1984). Job Preferences, College Major, and the Gender Gap in Earnings. *Journal of Human Resources.* 19(3), 408–428.
- DOLTON, P. J., & MAKEPEACE, G. H. (1986). Sample Selection and Male-Female Earnings Differentials in the Graduate Labour Market. *Oxford Economic Papers.* 38(2), 317–341.
- DUMAN, A. (2010). Female Education Inequality in Turkey: Factors Affecting Girls' Schooling Decisions. *International Journal of Education Economics and Development.* 1 (3), 243-258.
- FORTIN, N.M. (2008). The Gender Wage Gap among Young Adults in the United States: The Importance of Money versus People. *Journal of Human Resources.* 43(4), 884–918.
- GLAUBER, R. (2018). Trends in the Motherhood Wage Penalty and Fatherhood Wage Premium for Low, Middle, and High Earners. *Demography* 55(5), 1663–1680.
- HARKNESS, S. (1996). The Gender Earnings Gap: Evidence from the UK. *Fiscal Studies.* 17(2), 1–36.
- HECKMAN, J.J. (1979). Sample Selection Bias as a Specification Error. *Econometrica.* 47(1), 153–161.
- HINKS, T. (2002). Gender Wage Differentials and Discrimination in the New South Africa. *Applied Economics.* 34(16), 2043–2052.
- ILKKARACAN, I., & SELIM, R. (2007). The Gender Wage Gap in the Turkish Labor Market. *Labour.* 21(3), 563–593.
- INTERNATIONAL LABOUR FORCE (7.5.2021a). Labour force participation rate by sex and age [online], Web address: https://www.ilo.org/shinyapps/bulkexplorer4/?lang=en&segment=indicator&id=EAP_2WAP_SEX_AGE_RT_A [Date Accessed: 7 May 2021].
- INTERNATIONAL LABOUR FORCE (7.5.2021b). International Standard Classification of Occupations [online]. Web address: https://www.ilo.org/public/english/bureau/stat/isco/. [Date Accessed: 7 May 2021].
- INTERNATIONAL MONETARY FUND. (2021). *World Economic Outlook: Managing Divergent Recoveries*. Washington, DC: International Monetary Fund.

- KARA, O. (2006). Occupational Gender Wage Discrimination in Turkey. *Journal of Economic Studies.* 33(2), 130–143.
- KHITARISHVILI, T. (2009). Explaining the Gender Wage Gap in Georgia. Working Paper (577). *Levy Economics Institute*.
- KORENMAN, S., & NEUMARK, D. (1992). Marriage, Motherhood, and Wages. *The Journal of Human Resources.* 27(2), 233–255.
- KUNZE, A. (2008). Gender Wage Gap Studies: Consistency and Decomposition. *Empirical Economics.* 35(1), 63–76.
- LEE, B.S, JANG, S., & SARKAR, J. (2008). Women's Labor Force Participation and Marriage: The Case of Korea. *Journal of Asian Economics.* 19(2), 138–154.
- LIU, A.Y.C. (2004). Gender Wage Gap in Vietnam: 1993 to 1998. *Journal of Comparative Economics.* 32(3), 586–596.
- MINCER, J. (1974). *Schooling, experience and earnings*. New York: Columbia University Press.
- MYSÍKOVÁ, M. (2012). Gender Wage Gap in the Czech Republic and Central European Countries. *Prague Economic Papers*. 21(3), 328-346.
- NEUMAN, S., & OAXACA, R.L. (2004). Wage Decompositions with Selectivity-Corrected Wage Equations: A Methodological Note. *Journal of Economic Inequality.* 2(1), 3–10.
- OAXACA, R.L. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review.* 14(3), 693–709.
- OLIVETTI, C., & PETRONGOLO, B. (2008). Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps. *Journal of Labor Economics.* 26(4), 621–654.
- RANKIN, B.H., & AYTAC, I.A. (2006). Gender Inequality in Schooling: The Case of Turkey. *Sociology of Education.* 79(1), 25–43.
- SHERAN, M. (2007). The Career and Family Choices of Women: A Dynamic Analysis of Labor Force Participation, Schooling, Marriage, and Fertility Decisions. *Review of Economic Dynamics.* 10(3), 367–399.
- TANSEL, A. (2002). Determinants of School Attainment of Boys and Girls in Turkey: Individual, Household and Community Factors. *Economics of Education Review.* 21(5), 455–470.
- TANSEL, A. (2005). Public-Private Employment Choice, Wage Differentials, and Gender in Turkey. *Economic Development and Cultural Change.* 53(2), 453–477.
- TANSEL, A., & DAOUD, Y. (2011). Comparative Essay on Returns to Education in Palestine and Turkey. No. 5907. *IZA DP*.
- TEKGUC, H., ERYAR, D., & CINDOGLU, D. (2017). Women's Tertiary Education Masks the Gender Wage Gap in Turkey. *Journal of Labor Research.* 38(3), 360–386.
- WRIGHT, R.E., & ERMISCH, J.F. (1991). Gender Discrimination in the British Labour Market : A Reassessment. *The Economic Journal.* 101(406), 508–522.

APPENDIX

Appendix 1. Standard O-B decomposition

After estimating the wage equations for both males and females , standard O-B decomposition can be expressed below:

$$R = E(W_m) - E(W_f) = E(W_m)'\beta_m - E(W_f)'\beta_f \quad (1)$$

where represents the logarithmic monthly wages; represents the raw differential and indicates the slope parameters and constant term. In order to apply decomposition, O-B decomposition can be explained as below:

$$R = Endowment\ (E) + Coefficient\ (C) \quad (2)$$

$$E = \beta_f[E(X_m) - E(X_f)]' \quad (3)$$

$$C = E(X_f)'(\beta_m - \beta_f) \quad (4)$$

where contains the regressors explained in Equation 1 and also contains the unexplained portion of the difference (Blinder 1973).

Appendix 2. Overall Standard O-B Decomposition Findings Only for Employed Individuals

	2009		2018	
	Coef.	**S.E.**	**Coef.**	**S.E.**
Male	6.3157***	0.0158	7.5132***	0.0104
Female	5.6609***	0.0344	7.1315***	0.0215
Difference	0.6548***	0.0379	0.3817***	0.0239
Endowments	0.0399**	0.0201	-0.0661***	0.0132
Coefficients	0.6149***	0.0326	0.4479***	0.0201

Notes: ***p<0.01, **p<0.05, *p<0.1. Coef. denotes the coefficients; S.E. denotes the standard errors.

Appendix 3. Detailed Standard O-B Decomposition Findings Only for Employed Individuals

	2009		2018	
	Endowments	**Coefficients**	**Endowments**	**Coefficients**
Human capital				
Illiterate or compulsory education (reference)				
High school	0.0155** (0.0067)	0.0007 (0.0209)	0.008*** (0.002)	-0.0522*** (0.0118)
Associate or bachelor's degree	-0.0933*** (0.012)	-0.0579* (0.03)	-0.0655*** (0.0065)	-0.1083*** (0.0241)
Postgraduate degree	-0.0143** (0.0056)	-0.0024 (0.0052)	-0.0147*** (0.0037)	-0.0154*** (0.0047)
Experience	0.3674*** (0.0361)	0.7632*** (0.144)	0.1599*** (0.0259)	0.3639*** (0.1031)
Experience^2	-0.2371*** (0.0274)	-0.3356*** (0.0829)	-0.11*** (0.0227)	-0.2551*** (0.0616)
Occupation group				
ISCO1	0.0153*** (0.0031)	-0.0238*** (0.0061)	0.0016 (0.0028)	-0.0158*** (0.0046)
ISCO2	-0.0442*** (0.0073)	-0.0539** (0.0228)	-0.0698*** (0.0071)	-0.0135 (0.0172)
ISCO3	-0.0125*** (0.004)	-0.0592*** (0.0139)	0.0035 (0.0027)	-0.0044 (0.0058)
ISCO4	-0.0338*** (0.006)	-0.0759*** (0.0182)	-0.0186*** (0.0032)	-0.021** (0.0098)
ISCO5	-0.0005 (0.0024)	-0.0254* (0.0152)	-0.0089*** (0.0022)	-0.0084 (0.0148)
ISCO6	-0.0022 (0.0055)	0.01 (0.0089)	-0.0127*** (0.0032)	0.0102*** (0.0038)
ISCO7	0.0344*** (0.0052)	0.0407*** (0.0101)	0.0274*** (0.0038)	0.014*** (0.0053)
ISCO8	0.0453*** (0.0056)	-0.0322*** (0.0074)	0.0336*** (0.0036)	-0.0136*** (0.0041)
ISCO9 (reference)				
Constant	-	0.4667*** (0.1107)	-	0.5677*** (0.0767)

Notes: ***p<0.01, **p<0.05, *p<0.1. The numbers in brackets are the standard errors.

"This page is left blank for typesetting"

JAME
E-ISSN:2791-7401
Volume : 1, Issue 1, 2021
URL: https://journals.gen.tr/jame

Received: 22.05.2021
Acccepted: 17.06.2021

RESEARCH ARTICLE

Demographic variables affecting participation in mammography screening: a cross-sectional study

Hatice Özkoç[1]

1 Mugla Sitki Kocman University, Faculty of Economics and Administrative Sciences, Department of Business, Turkey
e-mail: haticeozkoc@gmail.com

Abstract

Breast cancer is one of the most common cancers among women. There are three commonly used diagnostic methods in this particular cancer type, for which the survival rate can be quite high with early diagnosis. Compared to breast self-examination and clinical examination, mammography stands out as a more accurate diagnostic method. Although there are health policies regarding mammography screening in many countries, mammography screening behavior has not reached to the desired level worldwide. In this study, we aim to reveal the demographic factors affecting the frequency of women getting mammograms in Turkey. The frequency of mammography screening was estimated with a generalized logit model using the microdata of TURKSTAT "Turkey Health Survey". As a result, we found that the mammography screening differentiated across different age and income groups. The disease status was reported to increase the frequency of mammography screening. However, marital status and employment status did not have any significant effect.

Keywords: *Breast Cancer, Mammography, Income Levels, Ordered Categorical Dependent Variable, Generalized Ordered Logit Model*

JEL Codes: *C02, I12*

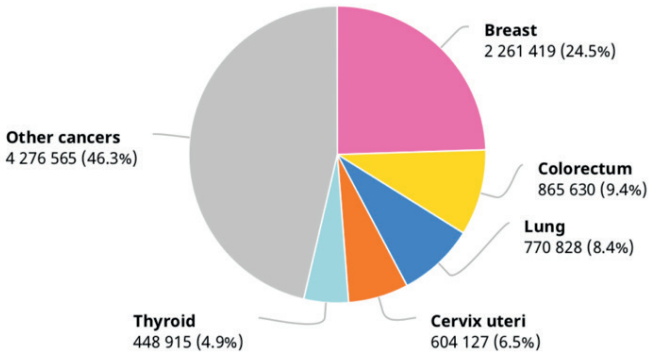
Citation: OZKOC, H. (2021). Demographic Variables Affecting Participation in Mammography Screening: A Cross-Sectional Study. Journal of Applied Microeconometrics (JAME). 1(1), 18-26, DOI:10.53753/jame.1.1.02

Corresponding Author/
Hatice Özkoç
E-mail: haticeozkoc@gmail.com

1. INTRODUCTION

Breast cancer, the most frequently diagnosis among women, is also a leading cause of cancer deaths (Bhandari et al. 2021). According to the GLOBACAN (2020) report published by The International Agency for Research on Cancer, breast cancer is one of the most common (24.5%), rapidly spreading and therefore important cancer types among women. In 2020, the female population in the world was 3,864,824,712, and in the same year, 9,227,484 women were newly diagnosed with cancer. 2,261,419 of them were reported to be diagnosed with breast cancer. Looking at women's deaths due to cancer worldwide in the same year, it is observed that deaths due to breast cancer ranked fourth (684,996) (GLOBACAN 2020). Bhandari et al. (2021) stated that breast cancer deaths differ from each other in low- and high-income countries. In the same study, it was stated that the death rate is three times higher in low-income countries. Figure 2 shows the percentages of the 10 most common cancer types among women in Turkey. As seen in the figure, breast cancer ranks first among other cancers with 24.8%.

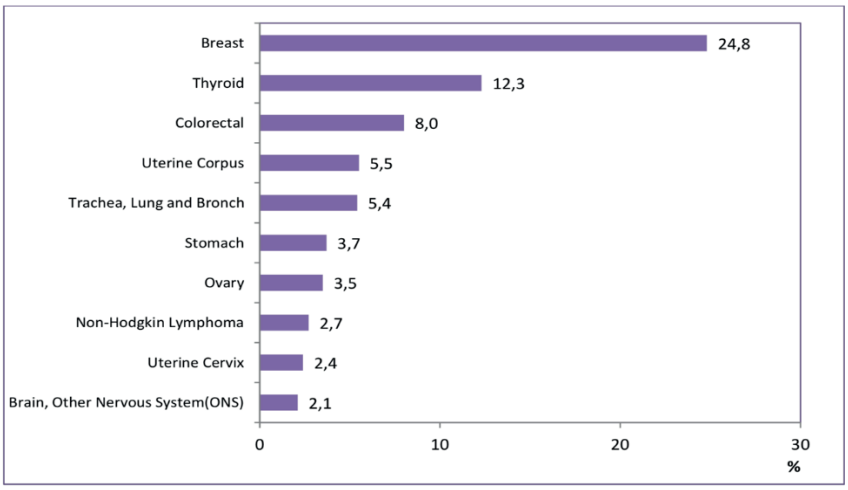
Figure 1. The number of new cancer cases in women in 2020



Source: The Global Cancer Observatory

While deaths associated with breast cancer rank first among the leading causes of death in females in underdeveloped countries, it comes after lung cancer in developed countries (Karaca et al. 2019). It was estimated that approximately 15% of cancer deaths among women were due to breast cancer (Torre et al. 2015). While 2% of total deaths in the EU in 2017 were due to breast cancer, this rate was 1% in Turkey during the same period (Health Statistics Yearbook 2019). Among the 20 most common causes of death in women in Turkey, breast cancer ranks the 8th with 2.1% (Akova, Hasdemir and Türkoğlu 2019: 90). While the incidence of breast cancer per 100,000 is 45.6 in Turkey, this value is 79 in the EU (Health Statistics Yearbook 2019).

Figure 2. Prevalence of the Most Common 10 Cancer Types Among Females in the Total Prevalence, Turkey, (%), 2016



Source: Ministry of Health

The GLOBACAN (2020) report states that 2.6 million women worldwide will have breast cancer in 2030 and 817 thousand women will die from breast cancer. Cancer screening is one of the most effective methods in the fight against cancer in order to create awareness in society regarding cancer. Screening methods with proven efficacy are used in order to detect breast cancer at an early stage. Breast cancer is a progressive disease and is more likely to be cured when diagnosed early. Hereby, life expectancy is high. It is stated that the 5-year survival rate in patients diagnosed with breast cancer with early diagnosis and treatment methods in developed countries is approximately 90-95% (Akova, Hasdemir and Türkoğlu 2019: 90; Aslaner 2019: 8).

There are three methods generally used for early diagnosis of breast cancer. These are Mammography, Clinical and breast self-examination (Bhandari et al. 2021). Compared to breast self-examination and physician examination, mammographic screening can help detect breast cancer earlier (Karaca et al. 2019) and is recommended as a standard screening test worldwide (Bhandari et al. 2021).

Turkey is one of those countries implementing a national breast cancer screening program (National Cancer Institute 2015). In Turkey, mammography is one of the most preferred screening methods for breast cancer screening due to its applicability and easy accessibility for women. According to the standards set by the Cancer Department of the General Directorate of Turkish Public Health, women aged 40-69 get mammograms every two years. In addition, clinical breast examination is recommended for women participating in screening to increase the effectiveness of mammography (Aşkın et al. 2019: 26).

In this study, the mammographic screening will be discussed and it will be examined whether there is a difference in mammography behaviors in the context of demographic characteristics. This study aims to examine the relationship between demographic characteristics and the frequency of mammography screening using the Generalized Ordered Logit model. In the following chapters, the studies in the literature that examine mammography behavior in the context of demographic characteristics will be included. Then, the qualitative preference model results will be included, based on a micro data set obtained at the national level.

2. LITERATURE

The common purpose of the studies in the literature regarding the acceptance and application of breast cancer screening methods is to increase the acceptance of these methods by women and thus to increase the early diagnosis rates.

In different studies, the rate of women getting mammograms varies between 3.4% and 85% (Bhandari et al. 2021; Ghanbari et al. 2020; Aşkın et al. 2019). This arising wide ranged ratio brings these questions along: “What is the reason for such a high difference?” and “How can this ratio be increased?”. In one of the studies conducted with the aim of finding answers to these questions, Aşkın et al. (2019) examined the relationship between mammography screening and sociodemographic variables using logistic regression. Their results suggested that age and education status had a statistical effect on mammography screening. Another study conducted in Malaysia, education, employment status, income, insurance and smoking were found to be significantly associated with the use of preventive medical care such as mammography (Cheah and Tang 2017). Aslaner (2019). On the other hand, this study revealed that while working status had a significant effect on mammography behaviors; marital status, educational status and income level did not make a difference in behaviors. Unlike the findings of Cheah and Tang (2017) and Aslaner (2019), Secginli and Nahcivan (2006) observed that employment status was not statistically significantly associated with the use of breast cancer screening.

Lack of education and lack of knowledge about mammography were found to have a significant adverse affect on the use of mammography among low-income women (Davis et al. 1996). In another recent study, it is determined that geographical inequalities and low socioeconomic status reduced the likelihood of early diagnosis in women (Shen, Chen and Hsieh 2020). Davis et al. (1996) determined that cost was one of the reasons for the decrease in the rate of mammography screening among low-income women.

What effect marital status has on breast cancer screening is a question of interest. Ghanbari et al. (2020) found that the breast cancer screening rate was low in married women. The same study determined that the social determinants

had an important role in breast cancer screening (Ghanbari et al. 2020). Mammography performance was associated with the health insurance and the family history. Moreover, clinical breast examination was associated with the age of women, and the breast self-examination was associated with the age and profession of women (Ghanbari et al. 2020). In another study, alcohol consumption was examined and it was found to be not associated with the breast cancer screening (Matsubara et al. 2013).

Unlike the studies in the literature examining the relationship between income and breast cancer screening rates, the number of studies on income groups is quite limited. Gathirua-Mwangi et al. (2018) conducted one of the pioneering studies examining the mammography screening rates for different income groups. Their studies reported that there were differences in predictors of mammography compliance for different income groups. Age was found to be an important determinant in both low-income and high-income groups. Doctor's recommendation was significant only in the low income group (Gathirua-Mwangi et al. 2018).

3. METHODS

3.1. Data

The data used in this study, in which the mammography behaviors of women were examined, was obtained from Turkey Health Survey micro-data set conducted by the TURKSTAT in 2016. Since women over the age of 40 are recommended to have mammograms, women aged 40 and over were included in the study. In total, the whole sample consists of 5266 women.

Explanations and descriptive statistics for the variables are presented in Table 1. Table 1 suggests that more than half of the women (55%) studied within the scope of the study had never had a mammogram in their lifetime. A very small portion of the women (19%) worked, whereas the vast majority (97%) were married (Table 1).

Table 1. The Variables Included in the Analysis

Variable Name	Explanations	N	Proportion
Mammography	When was the last time you had a mammogram/film mammography?		
	Never	2894	54.96%
	Over 5 years ago	443	8.41%
	More than 3 years but less than 4 years	314	5.96%
	More than 2 years but less than 3 years	295	5.60%
	More than 1 year but less than 2 years	552	10.48%
	In the last 12 months	768	14.58%
Age	Person's age		
	40-50	1918	36.42%
	51-60	1475	28.01%
	61-70	1057	20.07%
	70 and over	816	15.50%
Employment status	Employment status in the last week		
	Working	988	18.76%
	Not working	4278	81.24%
Marital status	Marital status		
	Never married	143	2.72%
	Married, divorced, spouse died	5123	97.28%
Income (TL)	Average monthly income of the household		
	0-1264	1428	27.12%
	1265-1814	1378	26.17%
	1815-2540	912	17.32%
	2541-3721	798	15.15%
	3722 and above	750	14.24%
Disease state	Having an illness or not		
	Yes	3910	74.25%
	No	1356	25.75%

This study examined the mammography behaviors of women in Turkey and the last date to undergo a mammography screening was considered as the dependent variable. Considering the values of this variable (Table 1), it has an ordered structure. If the dependent variable has an ordered structure, different estimation methods can be used. In the following section, the Generalized Ordered Logit Model is explained and the estimation results obtained with this model are given.

3.2. The Generalized Ordered Logit Model

Methods such as the Ordinary Least Squares Regression requires dependent variables to be on a range or a ratio scale. When the dependent variable is ordered, different estimation methods should be used. The most popular method used in the literature is the ordered logit model, also known as the proportional probability model (Williams 2016: 7). The basic assumption of ordered logit models is the parallel regression assumption. The parallel regression assumption known as the proportional odds assumption of the ordered logit model is important in order to see the applicability of these models. However, in practice, this assumption is mostly not met. If this assumption is not met, the ordered logit model results cannot be trusted and alternative models should be used. If the Multinomial Logit Model is used as an alternative model, there would be a loss of efficiency in the estimation results since the modeling is created in an unordered structure despite the ordered structure of the dependent variable. One of the less constrained alternative models that do not require parallel regression assumption among ordered logit models is the Generalized Ordered Logit Model. This model takes into account the ordered structure of the dependent variable and does not restrict the proportional odds assumption. The Generalized Order Logit Model includes linear constraints, alternative model parameterization, automatic model fitting, options for alternative connection functions, and the calculation of the estimated probabilities via the estimation command (Williams 2006: 58).

For the ordered dependent variable with M category, the Generalized Ordered Logit model can be written as follows (Williams 2006: 59; Abrudan, Pop and Lazăr 2020: 11):

$$P(Y_i > j) = g(X\beta_j) = \frac{exp(\alpha_j + X_j\beta_j)}{1+\{exp(\alpha_j + X_j\beta_j)\}} \quad , \; j = 1, 2, \dots, M-1 \tag{1}$$

Here M is the category number of the ordered dependent variable. In accordance with the equation numbered (1), each probability values for Y from 1 to M is calculated as follows:

$$\begin{aligned} &P(Y_i = 1) = 1 - g(X_i\beta_1) \\ &P(Y_i = j) = g(X_i\beta_{j-1}) - g(X_i\beta_j) \qquad j = 2, \dots, M-1 \\ &P(Y_i = M) = g(X_i\beta_{M-1}) \end{aligned} \tag{2}$$

When the category number of the dependent variable is 2 ($M = 2$), the generalized ordered logit model is equal to the logistic regression model. When $M > 2$, the Generalized Ordered Logit Model equates to a set of binary logit models in which the categories of the dependent variable are combined. For example; assuming that the dependent variable consists of 4 categories, for $M = 1$ category 1 is compared with the others. For $M = 2$ categories 3 and 4 are compared against categories 1 and 2. For $M = 3$ categories 1, 2, and 3 are compared with category 4. In this case, while the β coefficients are estimated the same for some values of j in the Generalized Ordered Logit Model, they can be estimated differently for others (Williams 2006: 59). In other words, it generates much more parameters in the Generalized Ordered Logit Model compared to the Ordered Logit Model. Thus, it provides an opportunity to examine and interpret the effects of the explanatory variables on the categories of the dependent variable in more detail.

In this study, when the structure of the dependent variable (last mammogram date) is considered, the Ordered Logit Model was applied first. In order to control for the parallel slopes assumption of the Ordered Logit Model, the necessary tests were carried out and the results obtained are given in Table 2. As the parallel slopes assumption

was not met, the Generalized Ordered Logit Model estimation results were obtained as an alternative method and these results are given in Table 3.

Table 2. Results of Testing the Parallel Slopes Assumption

Tests	Chi2	P> Chi2
Wolfe Gould	121.8	0.000
Brant	118.8	0.000
Score	124	0.000
Likelihood Ratio	119	0.000
Wald	123.9	0.000

Table 3. The Results of the Generalized Ordered Logit Model Regression

Last mammogram date	1: Never, 2: More than 5 years ago, 3: More than 3 years but less than 4 years, 4: More than 2 years but less than 3 years, 5: More than 1 year but less than 2 years, 6: Within the last 12 months								
	N=5266	Log Likelihood=-7049.3363		LR chi2 (50)=477.81		Prob>chi2=0.0000			
	Never		More than 5 years ago		More than 3 years but less than 4 years		More than 2 years but less than 3 years		More than 1 year but less than 2 years	
Variables		Odds Ratio		Odds Ratio		Odds Ratio		Odds Ratio		Odds Ratio
Age 40-50	0.469*** (0.097)	1.598	0.784*** (0.106)	2.191	0.920*** (0.116)	2.510	1.046*** (0.129)	2.848	1.178*** (0.174)	3.247
Age 51-60	0.943*** (0.096)	2.568	1.173*** (0.105)	3.232	1.179*** (0.115)	3.251	1.244*** (0.128)	3.471	1.359*** (0.172)	3.892
Age 61-70	0.637*** (0.100)	1.890	0.650*** (0.110)	1.916	0.714*** (0.121)	2.042	0.826*** (0.134)	2.285	0.960*** (0.181)	2.613
Working	-0.140* (0.076)	0.869	-0.106 (0.078)	0.899	-0.068 (0.080)	0.934	-0.058 (0.085)	0.944	-0.059 (0.102)	0.943
Married, divorced, spouse died	0.106 (0.178)	1.112	0.133 (0.184)	1.142	0.233 (0.196)	1.262	0.383 (0.219)	1.467	0.561* (0.295)	1.753
Income 1265-1814	0.444*** (0.080)	1.559	0.321*** (0.084)	1.378	0.310*** (0.089)	1.363	0.237** (0.095)	1.268	0.192 (0.121)	1.212
Income 1815-2540	0.588*** (0.089)	1.800	0.498*** (0.092)	1.645	0.448*** (0.097)	1.566	0.320*** (0.105)	1.377	0.323** (0.131)	1.381
Income 2541-3721	0.803*** (0.094)	2.231	0.675*** (0.096)	1.964	0.639 *** (0.100)	1.895	0.562*** (0.106)	1.753	0.693*** (0.128)	2.000
Income 3722+	1.066*** (0.096)	2.904	1.015*** (0.097)	2.761	0.965 *** (0.100)	2.624	0.084*** (0.085)	2.316	0.788*** (0.129)	2.198
Present Illness	0.482*** (0.068)	1.619	0.471*** (0.070)	1.602	0.449*** (0.073)	1.567	0.432*** (0.078)	1.540	0.530*** (0.100)	1.699
Constant	-1.703*** (0.203)	0.182	-2.212*** (0.213)	0.110	-2.624*** (0.228)	0.073	-3.068*** (0.255)	0.047	-4.138*** (0.349)	0.016

Notes: (i) values in parentheses are standard errors. (ii) ***, ** and * indicate the 1%, 5% and 10% significance levels, respectively. (iii) N is the number of sample units, shows the coefficients.

Using the data of Turkey Health Survey, the factors affecting the mammography screening time of women were estimated with the Generalized Ordered Logit Model and the results are given in Table 3. From Table 3, one can clearly see that the effect of the age and disease status of women on mammography behaviors is statistically significant. At the same time, it is determined that there is a statistically significant difference in the frequency of mammography screening, depending on the differences observed in income levels. On the other hand, the marital status and the employment status of a woman are found to affect the mammography behavior only in some categories.

Women between the ages of 40-50 are 3.25 times more likely to have a mammogram in more than 1 year but less than 2 years compared to the probability of having a mammogram in the last 12 months, compared to those over 70. The same probability is 3.90 times higher in women between the ages of 51-60, and 2.61 times higher between the ages of 61-70. Compared to those over the age of 70, the probability of never having a mammogram is approximately 1.60 times higher for women among the ages of 40 and 50. While it is 2.60 times higher in women aged 51-60 years, it is 1.89 times higher among the ages of 61-70. Moreover, our results suggest that women between the ages of 51-60 are more likely to have mammograms compared to other age groups.

The model results suggest that the probability of not having mammograms at all is reduced if among working women. In other words, if a woman works, her possibility of having a mammogram increases. Furthermore, married women are 1.75 times more likely to have a mammogram in more than 1 year but less than 2 years compared to the base category. We also found that the probability of women undergoing mammography increases in case of any diseases.

It is seen that the frequency of mammography increases with the increase in income level. For example, women in the highest income level are found to be approximately 2.2 times more likely to have a mammogram in more than 1 year but less than 2 years compared to women in the lowest income level.

4. CONCLUSION

This study focused on determining the demographic factors affecting the mammography behaviors of women over the age of 40. The data obtained from Turkey Health Survey 2016 questionnaire conducted by TURKSTAT were estimated with the generalized ordered logit model.

Our findings suggest that there was a difference in mammography behaviors depending on age and income. Healthcare providers should be aware of these differences observed in income and age and highlight strategies that increase mammography compliance for each income group.

Although marital status was not found to be a significantly effective factor, as far as is known from the literature, the breast cancer screening rate is low among married women. At this point, taking the cultural barriers into account, effective health education programs should be conducted and new ones should be implemented in order to increase the effectiveness of screening programs.

Lack of education and misconceptions are the main obstacles to cancer screening, so effective community-based health education is necessary to take structural and cultural barriers into account, reduce inequality in health care and increase the effectiveness of screening programs. Healthcare providers should take the time to educate women on mammography and help low-income women in particular determine how much screening costs. Providing efficient and timely screening services might increase the likelihood of receiving screening services.

Although this study fills an important gap in terms of examining mammography behaviors in different income and age groups in Turkey, it is very open to development. For example, family history should be taken into consideration, along with geographical inequalities and socioeconomic determinants. At the same time, breast cancer screening programs should be revised when necessary by observing the effect of training and health services provided with the help of longitudinal studies.

Funding

The author declares that this study has no financial support.

REFERENCES

- ABRUDAN, I. N., POP, C. M., LAZĂR, P. S. (2020). Using a General Ordered Logit Model to Explain the Influence of Hotel Facilities, General and Sustainability-Related, on Customer Ratings. *Sustainability,* 12(21), 9302.

- AKOVA, İ., HASDEMİR, Ö., TÜRKOĞLU, H. (2019). Bir İlde 40-69 Yaş Kadınlarda Meme Kanseri Taramalarının Değerlendirilmesi. *Bozok Tıp Dergisi*, 9(1), 89-92.

- ASLANER, Ç. (2019) Bir Aile Sağlığı Merkezine Başvuran Kadınların Meme Kanserinden Korunma Davranışlarının Karma Yöntemlerle Araştırılması, Uzmanlık Tezi, İstanbul Üniversitesi, İstanbul Tıp Fakültesi, İstanbul.

- AŞKIN, M., KOÇ, E. M., ARSLAN, M., ATEŞ, M. Y., DAĞ, G., SÖZMEN, M. K. (2019). Kadınların Meme Kanseri Taramalarına Katılımlarını Etkileyen Demografik Değişkenler. *İzmir Eğitim ve Araştırma Hastanesi Tıp Dergisi*, 23(1), 25-32.

- BHANDARI, D., SHIBANUMA, A., KIRIYA, J., HIRACHAN, S., ONG K. I. C., JIMBA, M. (2021). Factors Associated With Breast Cancer Screening Intention in Kathmandu Valley, Nepal. *PLoS ONE* 16(1), 1:15.

- CHEAH, Y. K., TANG, C. F. (2017). Factors Influencing The Use Of Preventive Medical Care in Malaysia: Evidence From National Health And Morbidity Survey Data. *Asian Economic Journal*, 31(2), 119-137.

- DAVIS, T. C., ARNOLD, C., BERKEL, H. J., NANDY, I., JACKSON, R. H., GLASS, J. (1996). Knowledge And Attitude On Screening Mammography Among Low-Literate, Low-Income Women. *Cancer: Interdisciplinary International Journal of the American Cancer Society,* 78(9), 1912-1920.

- GATHIRUA-MWANGI, W., COHEE, A., TARVER, W. L., MARLEY, A., BIEDERMAN, E., STUMP, T., RAWL, S. SKINNER, C.S. CHAMPION, V. L. (2018). Factors Associated With Adherence To Mammography Screening Among Insured Women Differ By Income Levels. *Women's Health Issues*, 28(5), 462-469.

- GHANBARI A., RAHMATPOUR P., HOSSEINI N., KHALILI M. (2020). Social Determinants of Breast Cancer Screening Among Married Women: A Cross-Sectional Study. *J Res Health Sci.* 20(1).

- GLOBOCAN. (2020) https://gco.iarc.fr/today/data/factsheets/populations/900-world-fact-sheets.pdf , [Data Accessed: 10/5/2021]

- HEALTH STATISTICS YEARBOOK, The Ministry of Health of Turkey. (2019) https://dosyasb.saglik.gov.tr/Eklenti/40566,health-statistics-yearbook-2019pdf.pdf?0, [Data Accessed: 9/5/2021]

- KARACA, M., BİLGETEKİN, İ., AVŞAR, E., ÖZTÜRK, B. (2019). Meme Kanserinin Erken Tanısında Klinik Bulgular mı? Mamografi mi?. *Akd Tıp D.* 5(3):510-5.

- MATSUBARA, H., HAYASHI, K., SOBUE, T., MIZUNUMA, H., SUZUKI, S. (2013). Association Between Cancer Screening Behavior And Family History Among Japanese Women. *Preventive Medicine*, 56, 293–298.

- NATIONAL CANCER INSTITUTE. (2015). Breast Cancer Screening Programs in 26 ICSN Countries, 2012: Organization, Policies, and Program Reach.

- SECGİNLİ, S., NAHCİVAN, N. O. (2006). Factors Associated With Breast Cancer Screening Behaviours in A Sample Of Turkish Women: A Questionnaire Survey. *International Journal of Nursing Studies*, 43, 161–171.

- SHEN, C. T., CHEN, F. M., HSIEH, H. M. (2020). Effect Of A National Population-Based Breast Cancer Screening Policy On Participation in Mammography And Stage At Breast Cancer Diagnosis in Taiwan. *Health Policy*, 124(4), 478-485.

- TORRE, L. A., BRAY, F., SIEGEL, R. L., FERLAY, J., LORTET-TIEULENT, J., JEMAL, A. (2015). Global cancer statistics, 2012. *CA: A Cancer Journal For Clinicians.* 65(2), 87-108.

- WILLIAMS, R. (2006). Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *The Stata Journal,* 6(1), 58-82.

- WILLIAMS, R. (2016). Understanding and interpreting generalized ordered logit models. *The Journal of Mathematical Sociology,* 40(1), 7-20.

"This page is left blank for typesetting"

JAME
E-ISSN:2791-7401
Volume : 1, Issue 1, 2021
URL: https://journals.gen.tr/jame

Received: 15.05.2021
Acccepted: 19.06.2021

RESEARCH ARTICLE

Main drivers of tobacco consumption among adolescents: the case of Kyrgyzstan

Zamira Oskonbaeva[1]

1 Kyrgyz-Turkish Manas University, Department of Economics, Kyrgyzstan, e-mail: zamira.oskonbaeva@manas.edu.kg

Abstract

Tobacco smoking among youth is a global concern. Tobacco's dangers and the probability of addiction are often underestimated by young people. Preventing tobacco consumption among youth is critical since tobacco-related illnesses will take away the lives of considerable number of young people who continue to smoke into adulthood. Youth are more vulnerable to nicotine (Goriounova and Mansvelder 2012) and may become addicted to it quicker than adults may. Even though they decide to quit in a few years, three out of four teen smokers continue to smoke into adulthood due to nicotine addiction.

The main purpose of this study is to investigate the tobacco consumption of adolescents in Kyrgyzstan. For this aim, a school-based survey of students aged 13-15 years was utilized. Global Youth Tobacco Survey was conducted in 2019 by Centres for Disease Control and Prevention. The questionnaire consists of 56 questions. This survey was conducted in order to get knowledge about adolescents' attitudes about smoking.

The Heckman selection model was employed for the empirical analysis. In the light of obtained results, it can be concluded that the main drivers of tobacco consumption among adolescents are age, gender, amount of spending money, parental smoking status, among others. The results indicate that tobacco consumption is higher among males. Parents' smoking habit leads to a higher probability of an adolescent to become a smoker in the future.

The outcomes from this study may provide some insights into the policies implemented to lessen smoking among adolescents not only for Kyrgyzstan but also for other countrie

Keywords: *Tobacco Consumption, Heckman Selection Model, Kyrgyzstan, Survey, Adolescents, Cross Section Data, Microeconometrics*

JEL Codes: *C1, C52, I12, C83, C81*

Citation: OSKONBAEVA Z, (2021). Main Drivers of Tobacco Consumption Among Adolescents: The Case of Kyrgyzstan Journal of Applied Microeconometrics (JAME). 1(1), 28-38 DOI: 10.53753/jame.1.1.03

Corresponding Author:
Zamira Oskonbaeva
E-mail: zamira.oskonbaeva@manas.edu.kg

1. INTRODUCTION

Each year eight million people lost their lives because of smoking. In other words, one death occurs every five seconds (Martin 2020). It costs the world 2 trillion dollars due to rising health related costs and productivity loss (UNDP 2019). It is considered that fifty percent of today's children and youth who consume tobacco will not live until the age of 70. Every consumed tobacco takes 11 minutes off your life (Martin 2020).

Tobacco smoking among young people is a serious problem both developed and developing countries have to manage. According to estimates, more than 10% of 13–15 years old students consume tobacco in some way. Furthermore, about 3 out of 4 school smokers will become adult smokers (US Department of Health and Human Services 2014).

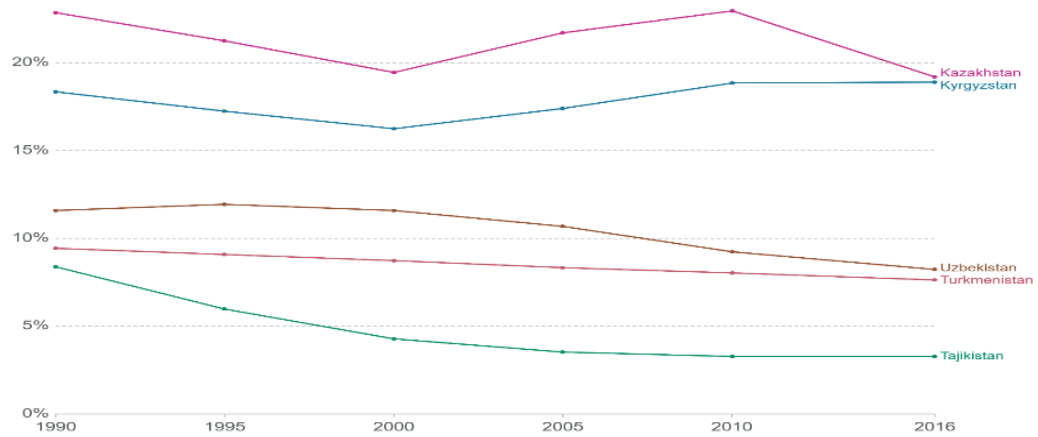
Borracci and Mulassi (2015) reported that smoking even one cigarette a month during adolescence was strongly linked to adult tobacco use, suggesting that many adult smokers could be prevented if smoking was not commenced during adolescence. It implies that preventing adolescents from smoking is critical since it has a long term impact. This finding emphasizes that by preventing the tobacco consumption of young people, the health status of the whole nations can be improved. Young people are known to be more likely to be influenced by those around them. Tobacco usage is portrayed as a societal norm in films, music videos, and advertisements leading younger generation to smoke.

Youth's smoking behavior is affected by the tobacco industry's advertising and marketing campaigns to a greater degree compared to the adult smoking behavior (Pollay et al 1996). Almost nine out of ten smokers that start smoking by the age of 18, and more than 80% of young smokers select tobacco brands which are the top three most heavily marketed brands (US Department of Health and Human Services 2017).

Additionally, they are more likely to become a smoker if their parents or peers do smoke. These findings underline once more that preventing tobacco consumption among adolescents is a worldwide concern requiring the attention from both families and governments.

As in other countries, youth smoking is widespread in Kyrgyzstan. Kyrgyzstan is a landlocked country in the Central Asia with a population of 6.5 million people. The prevalence of the current tobacco use (% of adults) is 27.9% (World Bank 2020). Figure 1 shows that the prevalence of daily smoking among people aged 10 years old and above is considerably high among the Central Asian countries.

Figure 1. Prevalence of Daily Smoking Aged 10 Years Old and Above (1990 to 2016)

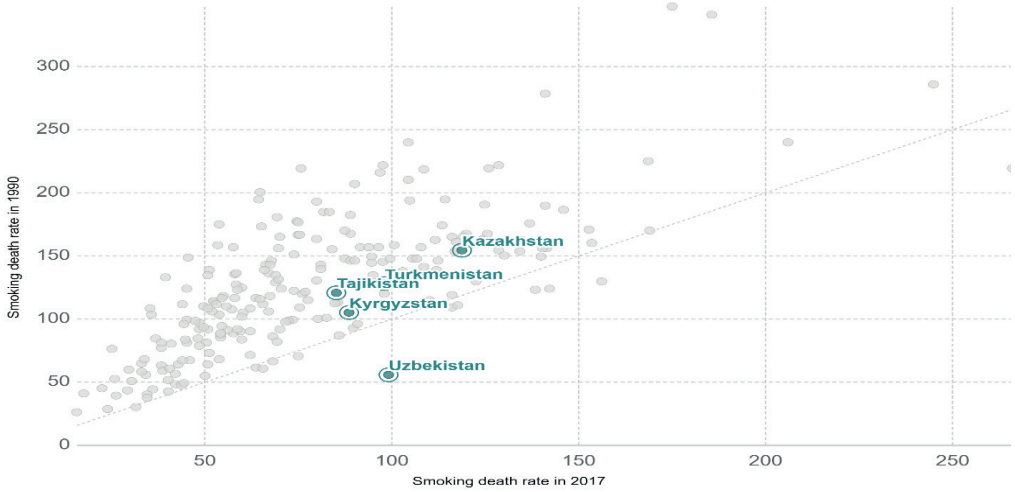


Source: IHME, Global Burden of Disease, Our World in Data.

There exist gender differences in tobacco consumption. The prevalence of tobacco consumption among men is 52.5 %, while this figure is approximately 3.4 % for females. More than 4500 people die as a result of tobacco-related

diseases annually in Kyrgyzstan (The Tobacco Atlas 2021). After heart disease, smoking-related lung disease comes as the second leading cause of death in Kyrgyzstan. Figure 2 below presents premature deaths attributed to smoking per 100,000 individuals. As can be observed from the figure, in all the Central Asian countries smoking related death rate is still high (except Uzbekistan).

Figure 2. Smoking Death Rate (1990-2017)



Source: IHME, Global Burden of Disease, Our World in Data.

In Kyrgyzstan 16 women lose their lives due to tobacco consumption every week (The Tobacco Atlas 2021). About 2500 boys and 700 girls aged 10-14 years old are using tobacco daily. This indicates that smoking consumption is widely spread among adolescents, too (The Tobacco Atlas 2021). In Kyrgyzstan the economic cost of smoking is 6450 million som including both direct and indirect costs associated with healthcare spending and reduced productivity as a result of premature mortality and morbidity (The Tobacco Atlas 2021).

Widespread tobacco consumption among adolescents is a sign of a great threat to the health of the total population. The findings of research conducted in 2017 demonstrate that there are many tobacco points of sale (POS) near the schools in Kyrgyzstan (Kyrgyzstan Tiny Targets Report 2018). The results indicate that in these POS 88.9% of all tobacco products were displayed in front of children; 61.1% of all tobacco were displayed near sweets or toys, and in 87.3% of POS cigarettes are sold by piece (Kyrgyzstan Tiny Targets Report 2018). All these factors may contribute to the spread of cigarette consumption among school pupils. Since children are our future, not only families but governments as a whole should care about children's health. All necessary steps should be taken in order to prevent the spread of smoking among adolescents.

The main purpose of this study is to explore the main drivers of tobacco consumption among adolescents in Kyrgyzstan. In the empirical analysis, GYTS (Global Youth Tobacco Survey 2019) data on tobacco consumption were employed. The Heckman (1979) selection model was employed to investigate the main drivers of smoking among adolescents. The main advantage of this specification is that it takes zero consumption observations into consideration. This is the case when the sample contains observations in which the dependent variable is zero and this feature is particularly notable among addictive goods like cigarettes (Yen 2005). Selection and outcome equations are estimated in the empirical analysis. Selection equation considers the main factors which affects adolescents' decision to smoke, while outcome equation determines the factors that influence the amount of tobacco used.

The remainder of this study is structured as follows. A brief review of the literature is presented in the second section. Third section is devoted to the dataset and variables utilized in the empirical analysis. Methodology and obtained findings are discussed in the fourth section. Discussion and conclusion are presented in the last section.

2. LITERATURE

The main contributing factors of smoking have been elaborated by researchers in the context of different countries. They focus on the factors associated with the current status of tobacco use among adolescents, adults, men and women separately. The findings indicate that the main drivers of tobacco consumption are age (Alasqah et al. 2021), gender (Granja et al. 2020), marital status (Alkan and Abar 2020), parental smoking status, education level (Klosterhalfen et al. 2021), socioeconomic status, living environment (Aho et al. 2018), social factors (Lazurenko et el. 2020), among others.

Below is a brief overview of previous studies grouped by factors that promote smoking.

The previous literature suggests that the social factors are crucial factors which have an impact on smoking behaviour of youths. Bonilha et al. (2013) concluded that social smoking is a more important motivator for adolescents compared to adult smokers. Lazurenko et al. (2020) investigated the main motives for smoking behaviour among students studying at college in Russia. According to their findings, smoking is considered as a tool of communication and integration for college students. In other words, social factors were found to be significant in explaining tobacco smoking.

Another group of studies concluded that lifestyle factors were main drivers of smoking. Mutlu Çamoğlu (2013) in case of Turkey revealed that the stress of urban life and working environment is important factor influencing smoking. Alkan and Abar (2020) investigated the factors of influencing tobacco smoking in Turkey. Logistic and probit regression results unveiled that age, gender, education, marital status, fruit consumption, exposure to tobacco smoke at home, and frequent of alcohol consumption were found as determining factors of smoking. La Fauci et al. (2021) investigated the impact of individual habits, lifestyles, the presence of smokers in the family and the use of electronic cigarettes on tobacco consumption among adults in Italy. The results indicate that if there is one smoker in the family it increases the likelihood of being a current smoker three times higher. Alcohol and coffee consumption, being unemployed and being a student were found to be contributing factors to smoking. Alkan and Ünver (2021) explored the main determinants of female's smoking behaviour. Logistic regression results indicated that age, education level, working conditions, income, general health status, psychosocial support conditions, and alcohol consumption were statistically significant variables in explaining smoking behaviour. Guimarães et al. (2021) revealed that the absence of religious beliefs, unhealthy eating habits, illicit drug abuse, and self-rated health were the main drivers of smoking behaviour of students in Brazil. Jain et al. (2021) reported that smokeless tobacco use (i.e. loose leaf, plug, twist, snus, or snuff) often leads to the increased consumption of cigarettes, e-cigarettes, and alcohol employing multivariable logistic regression models for the United States.

The results of other studies indicated that living environment was more critical in spreading tobacco consumption. De Vries et al. (2003) explored the impact of parents' and friends' smoking behaviours the on smoking status of adolescents in six European countries. By employing multiple regression model they investigated predictors of smoking. The results indicated that parents' and friends' smoking behaviours were found to be statistically significant in explaining the tobacco consumption of adolescents. Aho et al. (2018) investigated factors affecting smoking behaviour of Finnish vocational school students. Multinomial regression model results suggested that mother's smoking status was more influential on adolescent daily tobacco consumption.

The findings of other studies emphasized that age and gender were major contributing factors to smoking. Granja et al. (2020) reported that the habit of smoking was more prevalent among male students studying at a Brazilian private university. Alasgah et al. (2021) explored smoking predictors among adolescents using Saudi Arabian cross-sectional survey data. Their results revealed that age, gender, and academic performance were main drivers of smoking. Similarly, a study for Germany by Klosterhalfen et al. (2021) revealed that the water pipe consumption was associated with older age, male, migration background, lower educational level, and current smoking status.

A group of researches focusing on the ICT (Information and Communication Technologies) use and concluded that cell phone use and video games were the main motives for smoking among adolescents. Piola et al. (2021) explored the relationship between cell phone use and tobacco consumption among adolescents with different physical activity levels. Logistic regression results indicated that the cell phone use was a risk factor for alcohol and tobacco consumption. Garcia-Garcia et al. (2021) assessed the relationship between video games and tobacco

smoking among adolescents living in Spain. They concluded that there was a positive association between those two variables.

According to the results of other researches, it can be concluded that income was a major influencing factor of smoking. Donfouet et al. (2021) assessed the determinants of smoking use in Kenya. They concluded that poorer households were more impacted by tobacco use than richer households were. Using Korean data, Kim and Park (2021) investigated the impact of tobacco price on smoking behaviour. The findings indicated that cigarette price increases had no statistically significant impact on smoking cessation among smokers.

Advertising was found to be a contributing factor to smoking in studies conducted by Ali et al. (2020) and Ibarra-Salazar et al. (2020). Ali et al. (2020) explored the impact of e-cigarette advertisement exposure on e-cigarette and cigarette consumption among US adults. They found that e-cigarette advertising increased the amount of tobacco and e-cigarette used among adults of all ages. In the same way, Ibarra-Salazar et al. (2020) employing Dynamic Least Squares model (DOLS) with quarterly data (1994-2015), in the context of Mexico have revealed that health warnings reduced tobacco consumption.

According to the findings obtained by Escobedo et al. (1998), Fergusson et al. (2003), Rohde et al. (2004), McManus et al. (2010) and Nkomo et al. (2021) mental health behaviours predict the initiative to smoke.

In the case of Kyrgyzstan there existed a study conducted by Vinnikov et al. (2006) which were devoted to the analysing of the prevalence of smoking among medical students. Authors concluded that the prevalence were higher among men compared to women. Also the results unveiled that the likelihood of consuming alcohol was higher among smokers.

Literature review has shown that regression models (Aho et al. 2018; Ibarra-Salazar et al. 2020), logistic and probit regression models (Alkan and Abar 2020; Jain et al. 2021) and Heckman selection model (Madden 2008; Mutlu Çamoğlu 2013; Nkomo et al. 2021) are widely adopted in examining smoking behaviour of adolescents.

As can be seen from the above discussed literature, the main motives for smoking have been elaborated from different aspects. The findings differ due to the methodology used and the dataset utilized. In the context of Kyrgyzstan, smoking behaviour of adolescents have not been a focus of the researchers, yet. Thus, I hope that this study adds value to the growing body of literature through identifying the main motives for smoking among adolescents.

3. VARIABLES AND DATA

This paper employs the Global Youth Tobacco Survey (GYTS) dataset to investigate the major drivers of tobacco consumption among adolescents in Kyrgyzstan. The dataset is a cross-sectional survey of adolescents aged 13 to 15 years. This survey is conducted by the Centres for Disease Control and Prevention during January/August 2019. The funding for the GYTS was provided by the Canadian Public Health Association, National Cancer Institute, United Nations Children Emergency Fund, and the World Health Organization-Tobacco Free Initiative.

The variables incorporated in the model are age, gender, amount of spending money, parental smoking behaviour, experimentation with any form of smoked tobacco products other than cigarettes, own something with tobacco brand on it, student's opinion on the cost of cigarette, and using smokeless tobacco/nasway. Nasway can be defined as a smokeless tobacco for oral use, which is widely consumed mainly in Central Asian countries (WHO 2018).

In the first stage, the dependent variable is 1 if the adolescent spends on cigarettes, 0 otherwise. In the second stage, the dependent variable is a number of days spent smoking during the last month. The total sample size is 6145 observations (individuals) with an average age of approximately 14 years with a majority of males (50.7%). 14.1 % of adolescents reported that they have ever tried with cigarette smoking. The definitions of variables in the study are presented in Table 1.

Table 1. Definitions of Variables Used in the Model

Variables	Definitions
Dependent variables	
Smoking participation	1 if the youth spends on cigarettes, 0 otherwise.
Number of days spent smoking during past month	A categorical variable that is categorized into following groups; 1- 2 days, 3-5 days, 6-9 days, 10-19 days, 20-29 days, 30 days.
Explanatory variables	
Age	Years
Gender	1= male, 0 otherwise.
Amount of spending money	A categorical variable that is categorized into seven groups, do not have any spending money, less than 20 som, 20-50 som, 51-100 som, 101-200, 201-500, and more than 500 som.
Using smokeless tobacco/nasway	1= yes, 0 otherwise.
Own something with tobacco brand on it (t shirt, pen, backpack)	1= yes, 0 otherwise.
Student's opinion on the cost of cigarette	A pack of 20 cigarettes costs: 20-40 som, 41-50 som, 51-60 som, 61-70 som, 71-100 som, 101-150 som, more than 150 som.
Experimentation with any form of smoked tobacco products other than cigarettes	1= yes, 0 otherwise.
Parental smoking status	1 if fathers smoke, 0 otherwise.

4. METHODOLOGY AND FINDINGS

In order to determine the major drivers of tobacco consumption among adolescents Heckman selection model is applied. In case of the existence of sample selection, the observed data do not represent a random sample of the population. In other words, the data omitted in the sample does not respond to a random selection process (Zuzana et al. 2018). Therefore, a conventional analysis that only considers complete cases would produce biased results. This study employs the sample selection model developed by James Joseph Heckman (1979) is employed to address the potential sample selection bias and to obtain unbiased estimates of smoking behaviour.

Several studies elaborating smoking behavior of individuals employed this procedure (Madden 2008; Nkomo et al. 2021; Mutlu Çamoğlu 2013). The Heckman two-step procedure estimates two different models (probit and OLS regression). The main advantage of this procedure is that it allows for the possibility that factors affecting the probability of smoking and the amount of cigarette consumption are not the same. This procedure provides two estimated equations. The first one is a selection equation concentrating on the participation into the sample. The second one is an outcome equation which estimates smoking intensity.

In other words, the Heckman model estimates the determinants of adolescents being a smoker and the amounts of tobacco consumed. The selection equation should contain at least one variable that is not incorporated into the outcome equation (Wooldridge 2002). The main advantageous of this procedure is that it can be employed to correct for a sample selection bias (Heckman 1979).

This procedure has been applied in wage equation estimations for the first time. Then it has been employed in other fields effectively (Farrell and Walker 1999; Madden 2008). In this procedure all zeros are determined at the first stage. That is, once an adolescent decides to smoke, we will observe positive values.

Following Heckman (1979), the outcome equation can be specified as below:

$$Y_i^* = X_{1i}B_1 + u_i \quad (1)$$

and selection equation can be written as below:

$$Z_i^* = X_{2i}B_2 + \varepsilon_i \qquad (i=1, 2, \ldots\ldots l) \quad (2)$$

where Xji is a vector of exogenous variables, B_j is a vector of parameters, ε_i and u_i error terms of two equations. It is assumed that they are jointly normally distributed.

$$(u_i, \varepsilon_i) \sim NID(0, \begin{bmatrix} \sigma^2_{\varepsilon i} & \rho\sigma_{ui} \\ \rho\sigma_{ui} & 1 \end{bmatrix}) \quad (3)$$

u_i and $\varepsilon_i \sim N(0,0, \sigma^2 u_i, \sigma^2 \varepsilon_i, \rho\sigma_{ui})$, bivariate normal distribution with expectation zero, $\sigma^2 u_i$ and $\sigma^2 \varepsilon_i$ denote variances, respectively and $\rho\sigma_{ui}$ is a covariance. The term σ_{ui} is the standard deviation of u_i and ρ stands for the correlation between u_i and ε_i (Greene 2003).

The results of Heckman model are displayed in Table 3.

Table 3. Heckman Model

Heckman two-step model		
	Selection Equation	Outcome Equation
Age	0.0817*** (0.0417)	0.0363*** (0.0215)
Amount of spending money	0.0641** (0.0295)	0.0263*** (0.0153)
Using smokeless tobacco/nasway	-0.1778* (0.0441)	0.0942*** (0.0567)
Own something with tobacco brand on it (t-shirt, pen, backpack)	0.3381* (0.1256)	
Gender	0.2428* (0.0855)	
Pupil's opinion on the cost of cigarette	0.0314** (0.0153)	
Experimentation with any form of smoked tobacco products other than cigarettes	0.8276* (0.0932)	0.6450* (0.1071)
Parental smoking status	0.1628*** (0.0950)	1.6077* (0.4215)
Constant	-0.4822 (0.3784)	1.0000
λ (IMR)		1.6077
rho		10.53*
sigma		0.0146
Wald chi2	10.53*	
Prob > chi2	0.0146	

***, **, *denotes statistically significance at 10%, 5% and 1%, respectively. Standard errors are reported in parentheses.

The Wald test results show the goodness of fit for the model. First column indicates that age, gender, amount of spending money, parental smoking status, experimentation with any form of smoked tobacco products other than cigarettes, own something with tobacco brand on it, using smokeless tobacco/nasway, pupil's opinion on the cost of cigarette have significant effects on the selection equation; showing that they are important drivers of being a smoker.

Any experimentation with tobacco products leads to a higher probability of an adolescent to become a smoker. Table 3 shows that this factor is the biggest contributor to adolescents initiating smoking. Another influencing factor to smoking participation is parental smoking status. The results indicate that adolescents, whose fathers smoke, are more likely to become a smoker compared to adolescents, whose fathers do not smoke. The coefficient of age is positive and significant. It implies that the probability of adolescents to become a smoker is rising with age. An adolescents' likelihood of smoking increased in line with the level of pocket money received from parents. Having something with tobacco logo (t-shirt, pen, backpack) increases the probability of smoking participation. The coefficient of smokeless tobacco/nasway using is found to be negative in the selection equation. It implies that the consumption of smokeless tobacco nasway is decreasing with the probability of being a cigarette smoker.

The positive coefficient in the outcome equation signifies that after becoming a smoker consumption of nasway increases the amount of tobacco used. According to the GYTS, 5.1% of adolescents aged 13–15 (7.6% of boys and 2.9% of girls) in Kyrgyzstan consume smokeless tobacco (WHO 2018).

Moreover, the results indicate that boys are more likely to become a smoker compared to girls. In the outcome equation, all variables are found to be statistically significant. In other words, the coefficient of age indicates that the amount of tobacco consumption increases with age. The amount of pocket money increases the level of tobacco consumption. The coefficient of pupil's opinion on the cost of cigarette is found to be positive. This suggests that the cost of tobacco does not prevent pupils from being a smoker. This can be explained partly by the fact that there are many POS around the school in which cigarettes are sold by piece. Study conducted by Robertson et al. (2016) focusing on the impact of POS concluded that the chances of having attempted smoking were around 1.6 times higher among children and young people who were regularly exposed to point-of-sale cigarette promotion comparing to those who were less frequently exposed. Furthermore, chance of being vulnerable to potential smoking were 1.3 times higher among never smokers who were often exposed to point-of-sale tobacco compared with those who were less frequently exposed.

The inverse of Mills' ratio (IMR) is also found to be statistically significant. In other words, it indicates that applied procedure is appropriate for estimation. In order to control for sample selection, the ordinary-least-square regression adds on the inverse Mills' ratio (IMR) calculated from the linear predictions of the probit model as an additional explanatory variable. Rho is the correlation coefficient of the error terms from the selection and the outcome equation. Sigma is the estimator of the standard error of the residual in the regression equation. The coefficients are different from zero. It implies that OLS is biased and there is need to use a selection-bias correction model.

These results are parallel to the findings of previous studies conducted by Klosterhalfen et al. (2021), De Vries et al. (2003), Granja et al. (2020) and Alasqah et al. (2021).

5. CONCLUSION

The main objective of this study is to determine the main drivers of tobacco consumption among adolescents using the data collected from Kyrgyzstan. In the empirical analysis, the survey data provided by the GYTS was employed. In order to explore the factors that influence tobacco use, a procedure developed by Heckman (1979) was utilized. The empirical model is appropriate since it takes observations with zero tobacco consumption into the consideration.

In terms of the parental smoking status, results illustrate that fathers smoking behaviour contributes significantly to the decision to smoke or not (selection equation). Turning to the amount of pocket money, results indicate that it contributes significantly to both the likelihood of being a smoker and amount of tobacco used. Moreover, it can be concluded that there are gender differences in tobacco consumption. In other words, findings signify that male adolescents smoke more than female adolescents do. Outcomes report that there is a positive relationship between age and tobacco smoking. Age increases the likelihood of being a smoker (in selection equation) and the amount of tobacco used (in outcome equation).

The findings obtained from this study may provide some insights into the policies implemented to decrease smoking among adolescents in Kyrgyzstan. Smoking bans in public spaces, anti-smoking advertising campaigns being implemented in order to minimize tobacco consumption. Policies and interventions that aim to prevent and reduce adolescents' tobacco consumption should be developed and applied. Enormous efforts should be undertaken by not only families but also governments as a whole to reduce tobacco smoking among adolescents.

Funding

The author declares that this study has no financial support.

REFERENCES

- AHO, H., KOIVISTO, A.M., PAAVILAINEN, E. and JORONEN, K. (2018). Parental Involvement and Adolescent Smoking in Vocational Setting in Finland. *Health Promotion International*, 3(5), 846–857. https://doi.org/10.1093/heapro/dax027

- ALASQAH, I., MAHMUD, I., EAST, L. and USHER, K. (2021). Prevalence of Smoking and Smoking Predictors among Adolescents in Qassim, Saudi Arabia: Does the Healthy City Programme Make Any Difference? *Journal of Public Health*, 1-8.

- ALKAN, Ö. and ABAR, H. (2020). Determination of Factors Influencing Tobacco Consumption in Turkey Using Categorical Data Analyses. *Archives of Environmental & Occupational Health*. 75(1), 27-35.

- ALKAN, Ö. and ÜNVER, Ş. (2021). Tobacco Smoke Exposure among Women in Turkey and Determinants. *Journal of Substance Use*, 1-7.

- ALI, F. R. M., DAVE, D. M., COLMAN, G. J., WANG, X., SAFFER, H., MARYNAK, K. L., DENCH, D. and GROSSMAN, M. (2021). Association of e-Cigarette Advertising with e-Cigarette and Cigarette Use among US Adults. *Addiction*, 116, 1212–1223. https://doi.org/10.1111/add.15281.

- BONILHA, A.G., DE SOUZA, E.S.T., SICCHIERI, M.P., ACHCAR, J.A., CRIPPA, J.A.S. and BADDINI-MARTINEZ, J. (2013). A Motivational Profile for Smoking among Adolescents. *Journal of Addiction Medicine*, 7(6), 439-446. doi:10.1097/01.ADM.0000434987.76599.c0.

- BORRACCI, R.A. and MULASSI, A.H. (2015). Tobacco Use During Adolescence May Predict Smoking During Adulthood: Simulation-Based Research. *Arch Argent Pediatr*, 113, 106–12.

- CENTERS FOR DISEASE CONTROL AND PREVENTION (CDC) (2019). Ministry of Education and Science (Kyrgyzstan), Ministry of Health (Kyrgyzstan), World Health Organization (WHO). Kyrgyzstan Global Youth Tobacco Survey 2019. Atlanta, United States of America: Centers for Disease Control and Prevention (CDC).

- DE VRIES, H., ENGELS, R., KREMERS, S., WETZELS, J. and MUDDE, A. (2003). Parents' and Friends' Smoking Status as Predictors of Smoking Onset: Findings from Six European Countries. Health Education Research. 18(5), 627–636, https://doi.org/10.1093/her/cyg032

- DONFOUET, H.P.P., MOHAMED, S.F. and MALIN, E. (2021). Socioeconomic Inequality in Tobacco Use in Kenya: A Concentration Analysis. *Int J Health Econ Manag*. https://doi.org/10.1007/s10754-020-09292-0

- ESCOBEDO, L.G., REDDY, M. and GIOVINO, G.A. (1998). The relationship between Depressive Symptoms and Cigarette Smoking in US Adolescents. *Addiction*, 93, 433–40.

- FARRELL, L. and WALKER, I. (1999). The Welfare Effects of Lotto: Evidence from the UK. *Journal of Public Economics*, 72(1), 99–120.

- FERGUSSON, D.M., GOODWIN, R.D. and HORWOOD, L.J. (2003). Major Depression and Cigarette Smoking: Results of a 21-Year Longitudinal Study. Psychological Medicine, 33(8), 1357-1367.

- GARCÍA-GARCÍA, E., MARTÍNEZ-GIMENO, M. L., BENÍTEZ-ANDRADES, J. A., MIRANDA-GÓMEZ, J., ZAPATA-CORNEJO, E. D. and ESCOBAR-AGUILAR, G. (2021). Substance Use and Addictive Behavior in Spanish Adolescents in Secondary School. *Healthcare (*Basel, Switzerland), 9(2), 186. https://doi.org/10.3390/healthcare9020186

- GORIOUNOVA, N. A. and MANSVELDER, H. D. (2012). Short- and Long-Term Consequences of Nicotine Exposure During Adolescence for Prefrontal Cortex Neuronal Network Function. *Cold Spring Harbor Perspectives in Medicine*, 2(12), a012120. https://doi.org/10.1101/cshperspect.a012120

- GRANJA, G. L., LACERDA-SANTOS, J. T., DE MOURA BRILHANTE, D., DE SOUSA NÓBREGA, I., GRANVILLE-GARCIA, A. F., JUNIOR, A. D. F. C. and DOS SANTOS, J. A. (2020). Smoking and Alcohol Consumption among University Students of the Healthcare Area. *Journal of Public Health*. 28(1), 45-52.

- GREENE, W.H. (2003). *Econometric Analysis*. 4th edn. Upper Saddle River, NJ: Prentice Hall.

- GUIMARÃES, N. S., DE PAULA, W., DE AGUIAR, A. S. and MEIRELES, A. L. (2021). Absence of Religious Beliefs, Unhealthy Eating Habits, Illicit Drug Abuse, and Self-Rated Health is Associated with Alcohol and Tobacco Use among College Students—Padu Study. *Journal of Public Health*, 1-9.

- GUTEMA, B.T., CHUKA, A., AYELE, G., ESTIFAONS, W., MELKETSEDIK, Z.A., TARIKU, E.Z., ZERDO, Z., BAHARU, A. and MEGERSA, N.D. (2021). Tobacco Use and Associated Factors among Adults Reside in Arba Minch Health and Demographic Surveillance Site, Southern Ethiopia: A Cross-Sectional Study. *BMC Public Health,* 21, 441. https://doi.org/10.1186/s12889-021-10479-4

- HECKMAN, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica* 45, 153–161.

- IBARRA-SALAZAR, J., ROMERO-ROJAS, J. C. and AYALA-GAYTÁN, E. (2020). The Duration Effect of Pictorial Health Warnings on Tobacco Consumption in Mexico. *Global Public Health.* 1-19.

- IHME, (Institute for Health Metrics and Evaluation) Global Burden of Disease (2021). http://www.healthdata.org/gbd/2019 Accessed on April 2, 2021.

- JAIN, V., AL-RIFAI, M., NADERI, S., BAROLIA, R., IQBAL, S., TAJ, M. and VIRANI, S. S. (2021). Association of Smokeless Tobacco Use with the Use of Other Illicit Drugs in the United States. *The American Journal of Medicine*, 134(1), e15-e19.

- KIM, D. and PARK, H. (2021). Effects of a Cigarette Price Increase on the Smoking Behavior of Smokers and Non-smokers. *American Journal of Health Behavior*. 45(2), 205-215.

- KLOSTERHALFEN, S., KOTZ, D., KUNTZ, B., ZEIHER, J. and STARKER, A. (2020). Waterpipe Use among Adolescents in Germany: Prevalence, Associated Consumer Characteristics, and Trends (German Health Interview and Examination Survey for Children and Adolescents, KiGGS). *International Journal of Environmental Research and Public Health*, 17(21). DOI: 10.3390/ijerph17217740

- KYRGYZSTAN TINY TARGETS REPORT (2018). Report prepared for the Campaign for Tobacco-Free Kids by Grassriots Inc Kyrgyzstan Tiny Targets Report: Tobacco Advertisement, Sales, Product Displays, and Purchase Incentives Around Primary, and Secondary Schools. Toronto, ON; -Washington, DC 2017. Accessed on April 14, 2021.

- LA FAUCI, V., MONDELLO, S., SQUERI, R., ALESSI, V., GENOVESE, C., LAUDANI, N. and CATTARUZZA, M.S. (2021). Family, Lifestyles and New and Old Type of Smoking in Young Adults: Insights from an Italian Multiple-Center Study. *Ann Ig*. 33(2), 131-140. doi: 10.7416/ai.2021.2419. PMID: 33570085.

- LAZURENKO, N.V., PODPORINOVA, N.N., POLENOVA, M.E., KOVALCHUK, O.V. and KOROLEVA, K.I. (2020). Social Factors of Spreading of Tobacco Smoking Among College Students. *Revista San Gregorio*, 37, 174-181.

- MADDEN, D. (2008). Sample Selection Versus Two-Part Models Revisited: The Case of Female Smoking and Drinking. *Journal of Health Economics*, 27(2), 300-307.

- MARTIN, T. (2020). Smoking Statistics from Around the World. https://www.verywellmind.com/global-smoking-statistics-2824393. Accessed on June 8, 2021.

- MCMANUS, S., MELTZER, H. and CAMPION, J. (2010). Cigarette Smoking and Mental Health in England: Data from the Adult Psychiatric Morbidity Survey. National Centre for Social Research. Accessed on June 8, 2021.

- MUTLU ÇAMOĞLU, S. (2013). Türkiye'de Sigara Talebinin Analizi: İki aşamalı Heckman Süreci. *EY International Congress on Economics I "Europe and Global Economic Rebalancing"* October 24-25, 2013, Ankara/Turkey.

- NKOMO, Y.N., BEATRICE, D.S.K. and BIYASE, M. (2021). The Impact of Mental Health Behaviour on Tobacco Consumption in South Africa. EDWRG Working Paper Number 02-2021.

- OUR WORLD IN DATA (2021). https://ourworldindata.org/ Accessed on April 15, 2021.

- PIOLA, T.S., PACIFICO, A.B., CAMPOS, J.G., RIBEIRO, A.G., BACIL, E.D., SILVA, M.P. and CAMPOS, W. (2021). Cell Phone Use is Associated with Alcohol and Tobacco Consumption in Insufficiently Active Adolescents.*The Journal of Sports Medicine and Physical Fitness*, 61(3),444-451. doi: 10.23736/S0022-4707.20.11356-2.

- POLLAY, R. W., SIDDARTH, S., SIEGEL, M., HADDIX, A., MERRITT, R. K., GIOVINO, G. A. and ERIKSEN, M.

P. (1996). The Last Straw? Cigarette Advertising and Realized Market Shares among Youths and Adults, 1979–1993. *Journal of Marketing*, 60(2), 1–16. https://doi.org/10.1177/002224299606000201.

- ROBERTSON, L., CAMERON, C., MCGEE, R., MARSH, L., and HOEK, J. (2016). Point-of-Sale Tobacco Promotion and Youth Smoking: A Meta-Analysis. *Tobacco Control,* 25(e2):e83-e9. Available from: https://www.ncbi.nlm.nih.gov/pubmed/26728139.

- ROHDE, P., KAHLER, C.W., LEWINSOHN, P.M. and BROWN, R.A. (2004). Psychiatric Disorders, Familial Factors, and Cigarette Smoking: II. Associations with Progression to Daily Smoking. *Nicotine & Tobacco Research*, 6, 119–32.

- IBARRA-SALAZAR, J., ROMERO-ROJAS, J.C. and AYALA-GAYTÁN, E. (2020). The Duration Effect of Pictorial Health Warnings on Tobacco Consumption in Mexico. *Global Public Health*, DOI: 10.1080/17441692.2020.1839928

- THE TOBACCO ATLAS (2021). http://files.tobaccoatlas.org/wp-content/uploads/pdf/kyrgyzstan-country-facts-en.pdf. Accessed on April 25, 2021.

- UNDP (2019). National Tobacco Control Strategies, file:///C:/Users/usre/Downloads/National_Tobacco_Control_Strategies_Toolkit.pdf. Accessed on June 8, 2021.

- US DEPARTMENT OF HEALTH AND HUMAN SERVICES (2017). Preventing Tobacco Use Among Youths, Surgeon General fact sheet 2017. https://www.hhs.gov/surgeongeneral/reports-and-publications/tobacco/preventing-youth-tobacco-use-factsheet/index.html Accessed on April 25, 2021.

- US DEPARTMENT OF HEALTH AND HUMAN SERVICES (2014). The Health Consequences of Smoking --- 50 Years of Progress: A Report of the Surgeon General. 2014. Accessed on June 8, 2021.

- VINNIKOV, D., LAHDENSUO, A. and BRIMKULOV, N. (2006). Medical Students of Kyrgyzstan: Smoking Prevalence and Attitudes to Smoking Cessation Counseling. *Global Heart*, 2(1), 31–37. DOI: http://doi.org/10.1016/j.precon.2006.07.002

- WOOLDRIDGE, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

- WHO (2018). Consumption and Approaches to the Regulation of Nasvay in the Commonwealth of Independent States. https://www.euro.who.int/__data/assets/pdf_file/0010/391762/Consumption-and-approaches-to-the-regulation-of-nasvay-in-CIS_EN.pdf . Accessed on April 20, 2021.

- WORLD BANK (2020). https://data.worldbank.org/indicator/SH.PRV.SMOK?locations=KG. Accessed on April 15, 2021.

- YEN, S. T. (2005). Zero Observations and Gender Differences in Cigarette Consumption. *Applied Economics*, 37(16), 1839-1849, doi: 10.1080/00036840500214322

- ZUZANA, S., SONIA, I. and CRISTINA, S. (2018). Understanding the Drivers for Natura 2000 Payments in Forests: A Heckman Selection Analysis. *Journal for Nature Conservation,* 46, 28-37. https://doi.org/10.1016/j.jnc.2018.07.003.

"This page is left blank for typesetting"

JAME
E-ISSN:2791-7401
Volume : 1, Issue 1, 2021
URL: https://journals.gen.tr/jame

Received: 02.05.2021
Acccepted: 22.06.2021

RESEARCH ARTICLE

Variable selection via the adaptive elastic net: mathematics success of the students in Singapore and Turkey

Kadriye Hilal Topal[1]

1 Nisantasi University, Department of Computer Programming, Turkey e-mail: hilal.topal@nisantasi.edu.tr

Abstract

The quality of education is crucial for its competitiveness in the developing world. International tests are organized at regular intervals to measure the quality of education and to see the place in the ranking of countries. The surveys on these examinations have provided a large number of variables that can be effective on the test scores, including family, teacher, school and course equipment and information communication technologies, etc. The important question is which variables are relevant for student achievement in these tests. The barriers of mathematics success of Turkish students in the TIMSS (International Mathematics and Science Study) exam were investigated and compared their status with Singaporean students who took part in at the top of the ranking in the exam. To do this, the adaptive elastic net which is one of the regularized regression methods was applied and compared their prediction accuracy according to three different alpha levels [0.1; 0.5; 0.9]. The findings indicate that individual, institutional, socioeconomic factors such as talented teachers, less homework, less extensive textbooks, and lessons that improving cognitive ability; home education resources as well as technological factors are effective in the educational performances of countries. The findings suggest that, a technology-oriented education system and these individual, institutional, socioeconomic factors can help increase the success of students in Turkey along with other countries which are having similar experiences in international tests.

Keywords: *Adaptive Elastic Net, Education, Machine Learning, Oracle Properties, TIMSSs*

JEL Codes: *A20, I21, C55*

Citation: TOPAL, K. H. (2021). Variable Selection via the Adaptive Elastic Net: Mathematics Success of the Students in Singapore and Turkey Journal of Applied Microeconometrics (JAME). 1(1), 40-54, DOI: 10.53753/jame.1.1.04

Corresponding Author:
Kadriye Hilal Topal
E-mail: hilal.topal@nisantasi.edu.tr

1. INTRODUCTION

Education is a key factor in the development of individuals and communities. For this important reason, various tests are conducted to see the educational status of countries and their rankings among other countries. These tests are important resources for individuals and institutions in determining the factors affecting the educational performance of countries and formulating educational policies. The 3rd must be removed. Trends in International Mathematics and Science Study (TIMSS) is one of the several international tests which measures educational performance. Understanding this test is very important in order to improve education from various aspects.

The huge educational data and information about students' and families' backgrounds and school characteristics can be obtained from the TIMSS test for researchers. It is applied to the 4th grade students aged 9.5 and older and the 8th grade students aged 13.5 and older, according to the International Standard Classification of Education (ISCED) developed by United Nations Educational, Scientific and Cultural Organization (UNESCO) (Martin et al. 2016). This is a test where students are asked various questions about their social and school lives, basic demographic information, home and environmental questions, school atmosphere, self-understanding, mathematics, and scientific studies as well as mathematics and science curriculum questions. Many different educational inputs affect educational performances of the countries directly or indirectly. The variable selection is a very important issue for determining variables, which are important for prediction of educational performances.

In the machine learning literature, least absolute shrinkage and selection operator (LASSO) and elastic net regression are the most commonly utilized variable selection and regularization methods that detect the subset of the "true" variable set of model and the optimal estimation rate. However, such method does not have any oracle properties: a) the asymptotic normality and b) the consistency in the variable selection. Zou (2006) proved that once a variable is selected from a set of considered variables, all variables are automatically selected from this set while constructing the model because of oracle properties of the adaptive LASSO procedure. However, because of the drawback of the adaptive LASSO in the case of collinearity, the adaptive elastic net method was proposed by Zou and Zhang (2009). The adaptive elastic net method can be employed to overcome the collinearity problem that might occur in variables by eliminating the parameter inconsistency and the variable selection bias as result of its oracle properties.

The aim of the study is to determine the effective variables on the student achievement of TIMSS in Turkey and Singapore to reveal the similarities and differences of the factors of educational performances between these two countries. The LASSO regression is a good variable selection method. The Elastic net regression takes into account multicollinearity (it means that there are high linear correlations between two or more explanatory variables) while performing variable selection. The adaptive elastic net regression has both these properties of lasso and elastic net regressions as well as oracle properties. For these reasons, the LASSO elastic net and the adaptive elastic net regression methods are applied on large-scale TIMSS data and are employed to compare forecasting accuracies of these models.

Table 1. The 8th Grade TIMSS Mathematics Performance by Years: Turkey and Singapore

Year	Number of participant Countries	Mathematics Score of TR	Rank(TR)	Mathematics Score of SGP	Rank(SGP)
1995	41	-	-	643	1
1999	38	429	31	604	1
2003	46	-	-	605	1
2007	50	432	30	593	3
2011	45	452	24	611	2
2015	50	458	24	621	1

Table 1 shows the 8th grade TIMSS mathematics performance by years for Turkey (TR) and Singapore (SGP). Turkey participated in 1999 for the first time in the TIMSS test and took 429 points. The rank of success of Turkey in mathematics score was 31 out of 38 countries. While Singapore has the highest achieving students in international education rankings with its teenagers coming at the top in mathematics tests, Turkey is still not good enough.

There are a few studies in the literature analyzing the TIMSS exam results through machine learning methods. For example, Filiz and Öz (2020) applied educational data mining on TIMSS 2015 the 8th grade Turkish students' data. They aimed to find best performer algorithm to classifying students' mathematic success and extract important features on success. The best algorithms were found as logistic regression and support vector machines. The "home educational resources", "student confident in mathematics" and "mathematics achievement too low for estimation" were selected as the most effective features on success. Filiz and Öz (2019) were purposed to find the algorithms that are the most appropriate for classifying the successes of students, especially in science subjects, and to determine the significant factors on this success by using the results of TIMSS 2015 8th grade Turkish students' data. Logistic regression and support vector machines were found as the most appropriate methods. The "computer tablet shared", "extra lessons last 12 month", "extra lessons how many month", "how far in education do you expect to go", "home educational resources", and "student confident in science" were selected as the most significant features in science success. Yoo (2018) utilized an elastic net method with a logistic regression model to determine the effective factors on TIMSS 2011 Korean 4th grade students' mathematics achievement. Among 162 TIMSS variables, 12 students and 5 teacher variables were selected through the elastic net method. Depren et al. (2017) employed a decision tree, a bayesian network, a logistic regression and neural network methods to determinate the TIMSS 2011 mathematics achievement factors for the 8th grade Turkish students. They found that the logistic regression was the best algorithm for the selection of the factors. Yoo and Rho (2017) applied random forest which was a supervised machine learning algorithm to determine the significant factors for the TIMSS 2015 Korean 8th graders' student, teacher, and school datasets such as "students' extra lessons or tutoring the last 12 months", "gen\years been teaching" and "total instructional hours per year". The highest relevant variables of TIMSS were found in the variables obtained from students' datasets.

In Yoo and Rho (2017) study, a data set with a total of 413 variables, consisting of 147 students, 175 teachers and 91 school variables, was used to investigate the important variables that affect the mathematics achievement of Korean students. 17 student variables were found possessing more than 20 variable significance. These variables measured students' math self-efficacy, confidence, and interest, math extra lessons, science self-efficacy, self-confidence, time spent on science homework and desire for education, the amount of books in the house and the education level of the father.

This study differs from other studies in the literature in terms of using adaptive elastic net regression, which is an advanced method enjoys oracle properties, in investigating the factors that determine the success of TIMMS. In addition, Singapore, which is a very successful country in TIMSS exams, and Turkey, which has average score, have been compared and their similarities and differences have been revealed.

2. METHODOLOGY

2.1. Adaptive Elastic Net

The least squares (LS) is an estimation method that gives all nonzero coefficients under fundamental assumptions such as linearity, homoscedasticity, independence and normality. In the literature, the subset of significant effective variables on the response variable utilizes the best-subset selection and stepwise selection procedures. Unfortunately, in datasets with a large number of explanatory variables, the best-subset selection procedure is infeasible since there is a need for many models to be estimated (Breiman 1995). In stepwise selection, the number of potential models to be estimated is less than in the best subset selection. However, this does not indicate that it is a more suitable model selection method, as stochastic errors are ignored in the variable selection step (Fan and Li 2004). Moreover, this can result in a locally optimal solution instead of global optimal solution especially if explanatory variables are associated (Shen and Ye 2002). Hoerl and Kennard (1970) proposed the ridge regression which is a superior method compared to the LS in terms of variance reduction when the multicollinearity problem arises among variables of large datasets. However, the ridge regression coefficients are not equal to zero since the method is not a variable selection procedure.

The LASSO method developed by Tibshirani (1996) provides a consistent selection of a subset of the relevant variables and predicted of model coefficients in a model. The LASSO estimator is as follows;

$$\hat{\beta}_{LASSO} = arg\ min\left[\sum_{i=1}^{n}(Y_i - \beta_0 - \sum_{j=1}^{p} X_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p}|\beta_j|\right] \quad (1)$$

Elastic net equation is written instead of LASSO equation. The LASSO equation must be written. Where λ is a tuning parameter of penalty |βj|, is a l1 penalty that is obtained by penalizing the absolute coefficients of the variables included in the model. According to Fan and Li (2001), there were two main problems of LASSO. The first problem was the shrinkage procedure of LASSO causing biased estimates of the large coefficients with an unacceptable estimation risk. The second problem was the variable selection of LASSO being automatic because the l1 penalty is not differentiable since it is singular at the origin. Zou (2006) proved that the LASSO is not consistent in model selection under mild conditions. To overcome such problems of LASSO, Zou (2006) developed the adaptive LASSO. This type of LASSO enjoys the oracle properties. The LASSO gains the oracle properties thanks to the optimal weights.

Elastic net method proposed by Zou and Hastie (2005) to overcome the parameter bias of LASSO in the case of collinearity. The elastic net estimator is as follows;

$$\hat{\beta}_{(Elastic\ Net)} = arg\ min\left[\frac{1}{2n}\sum_{i=1}^{n}(Y_i - \beta_0 - \sum_{j=1}^{p} X_{ij}\beta_j)^2 + \lambda(\frac{1-\alpha}{2}\sum_{j=1}^{p}\beta_j^2 + \alpha\sum_{j=1}^{p}|\beta_j|)\right] \quad (2)$$

Where λ and α are the tuning parameter of ℓ_1 and ℓ_2 penalty, respectively. $|\beta_j|$ is ℓ_1 penalty, $|\beta^2_j|$ is the ℓ_2 penalty which is obtained by penalizing the squared coefficients of the variables included in the model. The elastic net method establish a regression model taking advantage of ℓ_2 and ℓ_1 penalty of ridge and LASSO. Thus, the method simultaneously can make automatic variable selection and continuous shrinkage, and it can select groups of correlated variables. Nevertheless, this method does not enjoy the oracle properties such as parameter selection stability and normality. As a result, a significant bias can occur in the elastic net estimators.

Zou and Zhang (2009) developed the adaptive elastic net method, which enjoys the oracle properties and calculates the adaptive weights utilizing the elastic net estimators to overcome the problems of the elastic net.

The adaptive elastic net is a combination of ℓ_2 and ℓ_1 penalty. When the estimator matrix is sparse or ill-conditioned, the adaptive elastic net regression provides variable selection consistency as well as parameter stability. In addition, the ℓ_2 penalty deals with the trouble effects of strong collinearity (Zou and Zhang, 2009). Therefore, this method outperforms the elastic net. The estimator of the adaptive elastic net regression is as follows;

$$\hat{\beta}_{\left(\substack{Adaptive\\ ElasticNet}\right)} = (1+\frac{\lambda_2}{n})[\ \underset{\beta}{\operatorname{argmin}}\left\|y - \sum_{j=1}^{p} x_j\beta_j\right\|^2 + \lambda_2\sum_{j=1}^{p}\beta_j^2 + \lambda_1^*\sum_{j=1}^{p}\hat{w}_j|\beta_j|]$$
$$\text{Subject to } \hat{w}_j = \frac{1}{\left(|\hat{\beta}_j^{(enet)}|\right)^{\gamma}},\ j = 1, \dots, p\ . \quad (3)$$

Where λ_2 represents the tuning parameter of ℓ_2 regulation, λ_1^* represents the tuning parameter of ℓ_1 regulation, $\hat{w}_j$ is adaptive weight obtained through elastic net coefficients, is a positive constant calculated by $\gamma = \frac{2v}{1-v}$ and $v = \lim_{n\to\infty}\frac{log(p)}{\log(n)}$. Tunning parameter of the model function is as follows;

$$P_\alpha = (1-\alpha)|\beta|_1 + \alpha|\beta|_2^2 = \lambda_2|\beta|_2^2 + \lambda_1\|\beta\|_1 \quad (4)$$

Here $\alpha = \frac{\lambda_2}{\lambda_1+\lambda_2}$ and $\alpha\in(0,1)$. After Zou and Zhang (2009)'s elastic net prediction weights, Jiratchayut and Bumrungsup (2015) set rescaled the elastic net estimator with to be three estimators. Here is a parameter given relations between shrinkage parameters. Adaptive elastic net regression, unlike conventional regression methods, enables superior results in the analysis as it applies a reduction procedure for both variable selection and minimum

variance estimators (Jiratchayut and Bumrungsup, 2015).

3. RESULTS

3.1. Data Set

The dataset was obtained from the TIMSS international database for 2015. The data used the 8th grade student background (BSGM6) questionnaires for Turkey and Singapore. According to the TIMSS 2015 survey, the number of participated students was 6079 and 6116 respectively for Turkey and Singapore. The irrelevant variables were excluded and applied the cleaning process through listwise deletion to handle with the missing data. The contents of TIMSS mathematics assessment are algebra, data and chance, geometry, number, and cognitive areas (knowing, applying, reasoning). The dependent variable was the average of these values.

3.2. Empirical Findings

The LASSO elastic net and adaptive elastic net regression methods (at three alpha level) which are the machine learning technique has been applied to Turkey and Singapore's 2015 TIMSS Survey datasets. All methods carried out in R software using packages msaenet for adaptive elastic net, glmnet for LASSO and elastic net. First, I divided the dataset into training and test datasets with a ratio of 80:20. The dataset of Turkey consists of 3340 students by 2672 training and 668 test data whereas Singapore has 3318 number of students with 2654 training and 664 test data. The training dataset was utilized to predict model parameters and hyperparameters. The test dataset was utilized to validate forecasting accuracy of the models with these parameters and hyperparameters. Prediction accuracy measures how close the forecasting values are to the data employed in the estimation except of sample datasets. It informs whether the parameters are optimally selected and whether the best model is constructed. The validation of the hyperparameters of this constructed best model is measured by using the test dataset. The error-based criteria which are the root mean square error (MSE), the mean absolute error (MAE), and R-squared were used for prediction of hyperparameter and validation. After the detection and validation of hyperparameters, the regression coefficients were estimated through the utilization of the adaptive elastic net methods for both data sets.

Following the study of Jiratchayut and Bumrungsup (2015), three different approaches were applied for the detection of the alpha values, [0.1, 0.5, 0.9], and then the weighting of the adaptive elastic net was utilized. Finally, the most appropriate alpha value was chosen based on the prediction accuracy.

Table 2. Prediction and Forecasting Accuracy Table

Country	*weight*	*Train Sample*			*Test Sample*		
		RMSE	*MAE*	*R-squared*	*RMSE*	*MAE*	*R-squared*
TR	Lasso	0.123201	0.098334	0.63	0.126347	0.100251	0.61
	Elastic Net	0.127758	0.102687	0.61	0.131681	0.104535	0.58
	Adaptive Elastic Net(0.1)	0.115433	0.092498	0.68	0.114224	0.091378	0.67
	Adaptive Elastic Net(0.5)	0.115394	0.092517	0.68	0.114215	0.091350	0.67
	Adaptive Elastic Net(0.9)	0.115388	0.092514	0.68	0.114209	0.091341	0.67
SGP	Lasso	0.093833	0.071983	0.44	0.093934	0.069866	0.46
	Elastic Net	0.097089	0.074330	0.40	0.097025	0.072925	0.42
	Adaptive Elastic Net(0.1)	0.082870	0.063779	0.57	0.088314	0.068796	0.47
	Adaptive Elastic Net(0.5)	0.082813	0.063776	0.58	0.088310	0.068748	0.47
	Adaptive Elastic Net(0.9)	0.082808	0.063775	0.58	0.088301	0.068748	0.47

Notes: (i) In the adaptive elastic net method, penalty parameters, λ and α were optained by k-fold cross validation; k was chosen as 10.

(ii) Adaptive elastic net results were obtained according to α=(0.1,0.5,0.9) values.

(iii) $RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_t-\hat{y}_t)^2}{n}}$, $MAE = \frac{\sum_{i=1}^{n}|y_t-\hat{y}_t|}{n}$, $R-squared = \frac{\sum_{i=1}^{n}(\hat{y}_t-\bar{y}_t)^2}{\sum_{i=1}^{n}(y_t-\bar{y}_t)^2}$

Table 2 shows the RMSE, MAE, and R-squared results based on errors obtained from training and test sample purposes for two countries and each methods. The RMSE, MAE and R-squared values of TR are [0.115388, 0.092514, 0.68] for training data and [0.11421, 0.09134, 0.67] for test data; the values of SGP are [0.082808, 0.063775, 0.58] for training data and [0.08830, 0.06875, 0.47] for test data. As presented in Table 2, the adaptive elastic net regression rescaled with an alpha level of [0.9] is the most accurate method for both countries.

According to the estimation results, for the model for Turkey, 60 non-zero coefficients were estimated from 98 variables by the adaptive elastic net method with the alpha level 0.9 whereas for the model for Singapore, 68 non-zero coefficients were estimated from the 100 variables by the method with the alpha level 0.9. The results of adaptive elastic net regression are presented in Table 3.

Table 3. Results of the Adaptive Elastic Net

			TR	SGP
Variable	**Definition**		**Adaptive Elastic Net(0.9)**	**Adaptive Elastic Net(0.9)**
ITSEX	Sex Of Students	(1 if female, otherwise 0)	-0.0116	-0.0090
BSBG03	Gen\Often Speak <Lang Of Test> At Home	1(always) to 4(Never)	-0.0266	0
BSBG04	Gen\Amount Of Books in Your Home	1(0-10) to 4(More than 200)	0.0149	0.0079
BSBG05	Gen\Digital information Devices	1(None) to 4(More than 10)	0.0050	0.0085
BSBG06A	Gen\Home Possess\Computer Tablet Own	(1 if yes, otherwise 0)	0	0.0048
BSBG06B	Gen\Home Possess\Computer Tablet Shared	(1 if yes, otherwise 0)	0.0304	0.0180
BSBG06C	Gen\Home Possess\Study Desk	(1 if yes, otherwise 0)	0.0058	0.0091
BSBG06E	Gen\Home Possess\internet Connection	(1 if yes, otherwise 0)	0	0.0291
BSBG06F	Gen\Home Possess\Own Mobile Phone	(1 if yes, otherwise 0)	0.0332	0
BSBG06G	Gen\Home Possess\Gaming System	(1 if yes, otherwise 0)	0	-0.0081
BSBG06H	Gen\Home Possess\<Country Specific>	(1 if yes, otherwise 0)	0.0074	0.0142
BSBG06I	Gen\Home Possess\<Country Specific>	(1 if yes, otherwise 0)	-0.006	0
BSBG06J	Gen\Home Possess\<Country Specific>	(1 if yes, otherwise 0)	0.0186	0.0221
BSBG06K	Gen\Home Possess\<Country Specific>	(1 if yes, otherwise 0)	0	-0.0162
BSBG07A	Gen\Highest Lvl Of Edu Of Mother	1(Primary/lowerSecondary/None) to 7(Postgraduate)	0.0043	0.0064
BSBG07B	Gen\Highest Lvl Of Edu Of Father	1(Primary/lowerSecondary/None) to 7(Postgraduate)	0.0072	0.0015
BSBG08	Gen\How Far İn Edu Do You Expect To Go	1(Lower secondary) to 6(Postgraduate)	0.0257	0.0209
BSBG09A	Gen\Mother Born İn <Country>	(1 if yes, otherwise 0)	0	-0.0080
BSBG09B	Gen\Father Born İn <Country>	(1 if yes, otherwise 0)	0.0282	-0.0028
BSBG10A	Gen\Born İn <Country>	(1 if yes, otherwise 0)	0	0.0045
BSBG11	Gen\About How Often Absent From School	1(Once a week or more) to 4(Never,Almost Never)	0.0235	0.03178
BSBG12	Gen\How Often Breakfast On School Days	1(Every Day) to 4(Never,Almost Never)	0	-0.0038
BSBG13A	Gen\How Often Use Computer Tablet\Home	1(Every day,almost every day) to 4(Never,Almost Never)	0.0089	0
BSBG13B	Gen\How Often Use Computer Tablet\School	1(Every day,almost every day) to 4(Never,Almost Never)	0.0044	0.0012
BSBG13C	Gen\How Often Use Computer Tablet\Other	1(Every day,almost every day) to 4(Never,Almost Never)	0.0033	0.0073
BSBG14A	Gen\internet Use\Access Textbooks	(1 if yes, otherwise 0)	-0.0020	0
BSBG14B	Gen\internet Use\Access Assignments	(1 if yes, otherwise 0)	-0.0180	0.0302
BSBG14C	Gen\internet Use\Collaborate With Classmates	(1 if yes, otherwise 0)	0	0.0338
BSBG14D	Gen\internet Use\Communicate With Teacher	(1 if yes, otherwise 0)	-0.0359	0.0039
BSBG14E	Gen\internet Use\Find info To Aid in Math	(1 if yes, otherwise 0)	-0.0055	-0.0108
BSBG14F	Gen\internet Use\Find info To Aid in Science	(1 if yes, otherwise 0)	-0.0127	0.0158
BSBG15A	Gen\Agree\Being in School	1(Agree a lot) to 4(Disagree a lot)	0.0314	0.0114
BSBG15B	Gen\Agree\Safe At School	1(Agree a lot) to 4(Disagree a lot)	0	-0.0066
BSBG15C	Gen\Agree\Belong At School	1(Agree a lot) to 4(Disagree a lot)	0.0033	0
BSBG15D	Gen\Agree\Like To See Classmates	1(Agree a lot) to 4(Disagree a lot)	0.0042	-0.0023
BSBG15E	Gen\Agree\Fair Teachers	1(Agree a lot) to 4(Disagree a lot)	-0.0194	0.0045
BSBG15F	Gen\Agree\Proud To Go To This School	1(Agree a lot) to 4(Disagree a lot)	0	-0.0085

BSBG15G	Gen\Agree\Learn A Lot	1(Agree a lot) to 4(Disagree a lot)	0	0.0019
BSBG16A	Gen\How Often\Made Fun Of	1(At least once a week) to 4(Never)	0	0.0022
BSBG16B	Gen\How Often\Left Out Of Games	1(At least once a week) to 4(Never)	0.0090	-0.0044
BSBG16C	Gen\How Often\Spread Lies About Me	1(At least once a week) to 4(Never)	0	0.0039
BSBG16D	Gen\How Often\Stole Sth From Me	1(At least once a week) to 4(Never)	0.0018	0.0091
BSBG16E	Gen\How Often\Hurt By Others	1(At least once a week) to 4(Never)	-0.0055	0
BSBG16H	Gen\How Often\Posted Embarrassing Things	1(At least once a week) to 4(Never)	0.0144	-0.0038
BSBG16I	Gen\How Often\Threatened	1(At least once a week) to 4(Never)	0	0.0070
BSBM17A	Math\Agree\Enjoy Learning Mathematics	1(Agree a lot) to 4(Disagree a lot)	0.0041	-0.0003
BSBM17B	Math\Agree\Wish Have Not To Study Math	1(Agree a lot) to 4(Disagree a lot)	-0.0095	0.0087
BSBM17C	Math\Agree\Math is Boring	1(Agree a lot) to 4(Disagree a lot)	0.0034	0
BSBM17D	Math\Agree\Learn interesting Things	1(Agree a lot) to 4(Disagree a lot)	0	1.41666E-05
BSBM17F	Math\Agree\Like Numbers	1(Agree a lot) to 4(Disagree a lot)	0.0107	0
BSBM17G	Math\Agree\Like Math Problems	1(Agree a lot) to 4(Disagree a lot)	-0.0081	-0.0100
BSBM17H	Math\Agree\Look Forward To Math Class	1(Agree a lot) to 4(Disagree a lot)	0.0198	0.0061
BSBM17I	Math\Agree\Favorite Subject	1(Agree a lot) to 4(Disagree a lot)	-0.0075	0
BSBM18A	Math\Agree\Teacher Expects To Do	1(Agree a lot) to 4(Disagree a lot)	-0.0139	0
BSBM18B	Math\Agree\Teacher is Easy To Understand	1(Agree a lot) to 4(Disagree a lot)	0	-0.0063
BSBM18C	Math\Agree\interested in What Tchr Says	1(Agree a lot) to 4(Disagree a lot)	-0.0065	0.0066
BSBM18D	Math\Agree\interesting Things To Do	1(Agree a lot) to 4(Disagree a lot)	0	0.0057
BSBM18E	Math\Agree\Teacher Clear Answers	1(Agree a lot) to 4(Disagree a lot)	-0.0048	0
BSBM18F	Math\Agree\Teacher Explains Good	1(Agree a lot) to 4(Disagree a lot)	0	-0.0016
BSBM18G	Math\Agree\Teacher Shows Learned	1(Agree a lot) to 4(Disagree a lot)	0.0098	0
BSBM18H	Math\Agree\Different Things To Help	1(Agree a lot) to 4(Disagree a lot)	0.0097	0.0032
BSBM18I	Math\Agree\Tells How To Do Better	1(Agree a lot) to 4(Disagree a lot)	0	0.0043
BSBM18J	Math\Agree\Teacher Listens	1(Agree a lot) to 4(Disagree a lot)	0	-0.0045
BSBM19A	Math\Agree\Usually Do Well in Math	1(Agree a lot) to 4(Disagree a lot)	-0.0314	-0.0092
BSBM19B	Math\Agree\Mathematics is More Difficult	1(Agree a lot) to 4(Disagree a lot)	0.0214	0
BSBM19C	Math\Agree\Mathematics Not My Strength	1(Agree a lot) to 4(Disagree a lot)	0.0147	0
BSBM19D	Math\Agree\Learn Quickly in Mathematics	1(Agree a lot) to 4(Disagree a lot)	-0.0081	-0.0060
BSBM19E	Math\Agree\Mat Makes Nervous	1(Agree a lot) to 4(Disagree a lot)	0.0119	0
BSBM19F	Math\Agree\Good At Working Out Problems	1(Agree a lot) to 4(Disagree a lot)	-0.0072	-0.007
BSBM19G	Math\Agree\i Am Good At Mathematics	1(Agree a lot) to 4(Disagree a lot)	-0.0046	0.0046
BSBM19H	Math\Agree\Mathematics Harder For Me	1(Agree a lot) to 4(Disagree a lot)	0.0086	0.0184
BSBM19I	Math\Agree\Mat Makes Confused	1(Agree a lot) to 4(Disagree a lot)	0.0048	0.0060
BSBM20A	Math\Agree\Mathematics Will Help Me	1(Agree a lot) to 4(Disagree a lot)	0.0079	0.0116
BSBM20B	Math\Agree\Need Mat To Learn Other Things	1(Agree a lot) to 4(Disagree a lot)	0.0066	-0.002
BSBM20D	Math\Agree\Need Mat To Get The Job i Want	1(Agree a lot) to 4(Disagree a lot)	0	0.0157
BSBM20F	Math\Agree\Get Ahead in The World	1(Agree a lot) to 4(Disagree a lot)	0	0.0069
BSBM20G	Math\Agree\More Job Opportunities	1(Agree a lot) to 4(Disagree a lot)	0	-0.0074
BSBM20H	Math\Agree\Parents Think Math important	1(Agree a lot) to 4(Disagree a lot)	-0.0149	-0.0092
BSBM20I	Math\Agree\important To Do Well in Math	1(Agree a lot) to 4(Disagree a lot)	-0.0055	-0.0059
BSBM25AA	Math\How Often Teacher Give YouHomework/Math	1(Every Day) to 5(Never)	0.01293	-0.0096
BSBM25BA	Math\How Many Minutes Spent On Homework/Math	1(No Math Homework) to 6(More than 90 minutes)	0.01032	0.0101
BSBM26AA	Math\Extra Lessons Last 12 Month\Mathematics	1(Yes, to excel in class) to 3(No)	0	-0.0187
BSBM26BA	Math\Extra Lessons How Many Month\Mathematics	1(Did not attend) to 4(More than 8 months)	0.01191	0
ITAC-COMM1	Special Accommodation\Achievement Session	(1 if yes, otherwise 0)		-0.0784
BSDM-LOWP	Mathematics Ach Too Low For Estimation	1(not too low), o(low)	0.19278	0.2607

4. DISCUSSION

The findings of the adaptive elastic net regression are classified as home resources, breakfast habits, gender, origin and language, information and communication technologies, communication with social environment, talented teachers, mathematics achievement, homeworks and extra lessons. These could be summarized as follows:

Home Resources

Amount of books at home (BSBG04), digital information devices (BSBG05) and study desk possession (BSBG06C) variables can be grouped as home education opportunities variables. According to the TIMSS 2015 national mathematics and science pre-report published by the Republic of Turkey Ministry of National Education, the 8th grade students who participated in the 2015 study, 7% of home education opportunities are "high", 54% are "middle" and 40% are low (Yıldırım et al. 2016). The results of this report also support the findings of this study. The success rate of 8th grade students increased as their home education opportunities increased. These variables are important factors on the TIMSS Mathematics achievement for both countries. Students with more home education resources (amount of books, digital information devices, computer tablet, and study desk and country specific home resources) performed better overall on TIMSS compared to others, and the correlations between TIMSS and the school achievement were stronger for students had much more home education resources possession. Moreover, living in special accommodation (ITACCOMM1) separately from the family had a negative effect on the TIMSS math success of Singaporean 8th grade students. This is an important result in terms of the effect of the family on student success in Singapore.

The findings for Turkey show some similarities to the results of the study by Kaleli-Yılmaz and Hanci (2016). They found a relationship with TIMSS and school grades and parents' educational level but they did not come up with a gender effect. However, my results show that the family education (BSBG07A and BSBG07B) were also found to have a great effect on student success for both countries. In fact, the children's success should depend on the country's education system, not on the family's education. As long as the Ministry of National Education could provide support for families who are not capable of educational support for their children, this leads to the success of generations ongoing.

Breakfast Habits

According to Adolphus et al. (2013) and Littlecott et al. (2015) eating breakfast has a positive effect on children's cognitive performance, especially in the areas of memory and attention. Lundqvist et al. (2019) found a positive association between breakfast consumption and academic achievement among children and adolescents. The breakfast (BSBG12) was also found to be a supporter factor on the TIMSS math scores of students in Singapore.

Gender

According to the adaptive elastic net results, the gender variable (ITSEX) plays an important role in determining student success for two countries. The girls' achievement scores are less than boys for both countries. Although girls have self-confidence, sense of belonging, motivation, and liking learning they have lagged behind in transforming them into success unlike boys (Polat and Madra, 2018). In addition, as mentioned in the UNESCO 2017's "Cracking the code: girls' and women's education in science, technology, engineering and mathematics (STEM)" book, girls STEM performance can be affected from the individual level (biological and psychological factors), family and peer-level, school level and societal level. The methods offered can be summarized as individual-level that differentiated perception between genders is biological factors such as brain structure and hormones, as well as psychological factors such as interest and motivation. As with other cognitive skills, these are changeable, can be affected by training and practices, and can be developed. If parents make enough effort for the progress of girls, it is easier for girls to overcome social, economic, and cultural barriers. The education policies can build girls' interest, self-confidence, and professional expectations in STEM. Media and social media supports are needed to challenge sexual discrimination in science and to direct girls to science. In addition, social media literacy lessons should be given for girls to better benefit from digital technology.

Origin and language

The birth place of mother and father (BSBG09A and BSBG09B) variables have positive sign contrast to student born in country (BSBG10A) variable in Singapore Model. Singapore is a multi-ethnic country. About 75% of Singapore's population is Chinese and nearly %25 of the population is Malay or Indian (Ginsburg et al. 2005). Although students were born in the country, their parents can be immigrants. Since Singapore has bilingual education strategy that includes mother tongue (Chinese, Malay, or Tamil), language differences are not a problem

on Singapore's TIMSS success. The father born in the country (BSBG09B) variable was selected in the model for Turkey. This finding was not surprising for Turkish students, because of majority (or almost all) of the students and their parents were born in Turkey. Although the students and their parents were born in Turkey, the country has multilingual and multicultural structure. This is because the country has been a migration route for centuries and has hosted many different civilizations. (Yılmaz and Şekerci 2016). In some parts of the country, the language spoken within a family is different from the official language of the country.

Information and Communication Technologies

The Singapore's technology-oriented education system, possession of information communication technology devices (BSBG06A, BSBG06B, BSBG06E, BSBG06F), play a very important role in increasing Mathematics TIMSS achievement of Singapore. While mobile phone does not affect student achievement in Singapore, it has been observed to be effective in Turkey. According to the results of 2015 Turkish Statistical Institute (TUİK) Household Information Technology Usage Survey, 96.8% of the households have mobile phones, 43.2% of the households have computer tablets or laptop, 25.2% of the households have desktop computer and regular internet user rate is 94.2% in the first quarter of 2015. These rates explain the reasons for the positive effect of mobile phones on the 8th grade TIMSS achievement in Turkey. Since internet usage is very common, internet connection (BSBG06E) has not been found to be a distinctive feature of success. The using the internet to collaboration with classmates (BSBG14C) is not a factor in mathematics 8th grade students' TIMSS achievement of Turkey but adjuvant for Singapore's. Güler et al. (2017) examined the internet usage aims of children between the ages of 6-15 are examined. They found that, the rate of "making voice or video calls was 7%. In contrast to this low rate, the 6-15 age children's rate of "using the internet for homework and learning" was 85%. I found that the TIMSS scores decrease as the usage of computer tablets increases at home, school or anywhere among the Turkish 8th grade students.

Accessing textbooks and assignments through the internet (BSBG14A, BSBG14B), communication with teacher (BSBG14D), finding info to aid in math and science (BSBG14E, BSBG14F) variables are the factors that decreased student success by the contrary of Singapore. In contrast, the using of the internet to find info to aid in math (BSBG14E) and to access textbooks (BSBG14A) are the factors that increased success. The using the internet to find info to aid in math (BSBG14E) variable has a negative effect on Singapore's mathematics TIMSS success like Turkey. Since the Singapore mathematics curriculum is not exhaustive, it allows the student to understand better in lessons, and students use technology to access textbooks. Students use ICT for practice because they get "math sense" at school. Besides, the using the internet to access textbooks (BSBG14A) is not a distinctive feature on success, because each student can access. Moreover, the TIMSS scores of students increase as the level of their education expectations (BSBG08) increase and absent days from the school (BSBG11) decrease.

Communication With Social Environment

The communication of students in school has an important role in their mental, social, and academic development. In contrast, when the results of "The sense of school belonging" items are examined, the results of my analysis support the results of the 2011 TIMSS study of Topçu et al. (2016). I found that there was a negative correlation between sense of school belonging (BSBG15C) and TIMSS mathematics achievement scores in Turkey. The like to see classmates (BSBG15D) variable also has a negative effect on TIMSS mathematics achievement for Turkey. Besides, I can conclude that students who like being at school (BSBG15A) perform worse than others in the TIMSS math exam for both countries. Students who do not like being in school have more time to study lessons instead of having fun. In order to give students "math sense" in Singapore, mathematics lesson contents are not prepared exhaustively for the basic concepts of mathematics are learned more easily (CIU 2008). This inference explains why the decrease in the opinion of "I learn a lot at school", which is among my results, causes an increase in Singapore's TIMSS math score.

According to TIMSS 2015 survey results, in Singapore, 36% of students are every month and 6% of students are every week exposed to student bullying. In Turkey, 26% of students are every month, 6% of them are every week suffered from student bullying (Mullis et al. 2016). Student bullying may affect the academic performance of bullied children. Akyüz (2014) has studied Mathematics Achievement in TIMSS 2011 for Turkey and Finland and has revealed that student bullying is a significant factor of the 2011 mathematics achievement of Turkey. In

my TIMSS models, made fun of (BSBG16A), spread lies about me (BSBG16C), and threatened (BSBG16I) are decreasing factor for only Singapore's TIMSS performance; stole sth. from me (BSBG16D) is also a decreasing factor for both countries. With the development of technology, cyberbullying has become very common among adolescents recently (Gimenez-Gualdo et al. 2018). The posted embarrassing things (BSBG16H) is an example of cyberbullying. The left out of games (BSBG16B) and the posted embarrassing things (BSBG16H) variables have a negative effect on Turkish students' scores unlike Singapore's.

Talented Teachers

As Nye et al. (2004) suggests, teachers play an important role in the students' success in academic life as well as in daily life. The teacher quality is high in Singapore. Teachers are provided with training, salary arrangement and reward support (Levent and Yazıcı 2014). All teachers are trained at Nanyang Technological University, National Institute of Education. All candidate teachers are selected through interview by participants of a responsible institution (OECD 2012). In Singapore, information communication technologies are integrated into the class atmosphere through pedagogic methods by these talented teachers. In addition, another purpose of teachers is to give students math sense and provide an easy understanding of education in the lessons. Extra lessons are also provided for students had low perception. Since the education system of Singapore is pedagogical and has structured on the technology equipment, the teachers may not need to do and say interesting things for students' success. The reason for the negative effect of variables on success maybe only the unsuccessful students need these methods in Singapore. Besides, "teachers whose expectations were not known by the students (BSBM18A)", "who were thought not to say interesting things by the students (BSBM18C)", "who cannot give clear answers to the student questions (BSBM18E)" had a negative effect on Turkish Students' TIMSS achievement.

Mathematics Achievement

In Singapore, teachers who can be understood easily (BSBM18B) and can listen to what the students say (BSBM18J) have a positive effect on student achievement. In contrast, teachers who do and say interesting things about lessons (BSBM18C, BSBM18D), do various things to help students learn (BSBM18H), give information about how the students should do better when they make a mistake (BSBM18I) have a significant negative effect. In the TIMSS 2015 8th grade survey, there are 9 variables for measuring student confidence in mathematics. These can be listed as usually do well in math (BSBM19A), mathematics is more difficult (BSBM19B), mathematics not my strength (BSBM19C), learn quickly in mathematics (BSBM19D), math makes me nervous (BSBM19E), good at working out problems (BSBM19F), I am good at mathematics (BSBM19G), mathematics harder for me (BSBM19H) and math makes confused (BSBM19I). Contrary to Çavdar (2015), the TIMSS 2011 study, according to the Wilson and Narayan (2016), Choi et al. (2012), Ertürk and Erdinç-Akan (2018) studies that have been investigated the continuous effect of mathematics on achievement, as students' confidence in mathematics increases, their success also increases. All variables except "I am good at mathematics (BSBM19G)" supported the result of the studies. The student's confidence in mathematics has an important positive effect on 8th grade TIMSS math achievement. However, if the opinion that the student is good at math is decreasing, Turkish TIMSS scores are also decreasing contrary to Singapore.

My finding supports the Lee's (2013) study results that have found the students who think that they are bad in mathematics in Asia are more successful, students who think that they are good in mathematics in Europe are more successful.

The value that the student gives to mathematics perception was evaluated by 7 selected variables in the Adaptive Elastic Net model. Arıkan et al. (2016), Yavuz et al. (2017) have not found any significant relationship between the value given by Turkish students to mathematics and students' mathematics achievement scores in the TIMSS exams survey in 2007 and 2011. When the effect of the variables was examined one by one, the similar results were also found in this study. On the other hand, in Singapore, the decreasing in the expectation that more job opportunities can be obtained utilizing mathematics (BSBM20G) causes a decreasing effect on student math achievement. The parents think Math is important (BSBM20H) and being good at math is important (BSBM20I) variables have an important role in TIMSS scores in both two countries. These findings also show the negative effect of decreasing family support and students' math self-confidence on students' achievement.

Compared to the total time, students in Singapore are getting more hours of Mathematics lessons than Turkish students (Bozkurt et al. 2019). The result is not surprising as trying out feedback or different learning methods during limited mathematics lesson hours can make students lagged behind the other 8th class students in basic mathematics topics.

Homework and Extra Lessons

In contrast to Singapore in Turkey, decreasing the frequency of math homework (BSBM25AA) given to the 8th grade students is a negative factor on achievement. Because of the homework given to students very often in Turkey, students may be bored, feel tired or exhausted. This may lead to diminishing the will to study, so their academic success may reduce. As stated by the Mullis et al. (2016) TIMSS 2015 report, when compared to the rates of the 8th grade students spending more than 3 hours per week, it was seen that Singapore (22% rate) overtaken Turkey (12% rate). The rates of both countries are above the OECD average. Although the increase in the hours spent on homework, (BSBM25BA) seems to increase success in both countries in my results, the reason for Singapore's superiority in TIMSS exams cannot be explained only with this. In Turkey, homework does not play an improvement role in students' cognitive abilities because these are often repetitions of knowledge and practices in the classroom in contrast to Singapore. The mathematics lessons are given theoretical in Singapore and lesson practices are given to students as homework. Students do their homework from electronic lesson contents all schools served that are accessible to every student. In this way, students improve their cognitive abilities through the homework that utilized the information given in the class.

The attendance frequency of the extra lessons (BSBM26BA) have been found a positive effective factor in success only in Turkey. In contrast, the extra lessons last 12 month (BSBM26AA) has a negative effect on TIMSS success of 8th grade students in Mathematics in Singapore. This result is not surprising for Singapore, where the mathematics is teaching in the class detailed and that the students only who have unsuccessful need extra lessons.

5. CONCLUSION AND SUGGESTIONS

The subset of the variables affecting the TIMSS scores was selected for Turkey and Singapore using adaptive elastic net regression, which is a machine learning method that enables superior results in the analysis as it applies a reduction procedure for both variable selection and minimum variance estimators. The findings show that there exists some differences regarding the selected variables affecting the TIMSS scores among two countries. These are reported below:

In Turkey, there is the children descent from different ethnic origins. Therefore, the difference between the language spoken at home and the language of the test is a problem for Turkey. As a solution, extra official language lessons can be given to students who have problems with Turkish or can be provided bilingual (English and mother tongue) education to students similar to Singapore.

Home education resources and especially books that are important for the education and development of children have also been an important factor in the TIMSS achievement of students. Encouraging students to gain reading habit can be an important strategy for increasing the TIMSS achievement of Turkey.

The high frequency of homework given to the students is not found as an effective way to reach high TIMSS scores in Turkey contrary to Singapore. As in Singapore, instead of given homework that repeats information in class, teaching contents in the class and giving homework for only practice can be a way to increase students' cognitive development and TIMSS success in Turkey.

Mathematics textbooks in Turkey are quite extensive. This reduces the student's interest in mathematics. The content of the lesson can be reduced and lesson content that improves students' cognitive abilities instead of lesson content based on memorizing can be prepared like the Singapore education system.

Singapore has a technology-oriented education system unlike Turkey, information-communication technologies play an important role in TIMSS success. Unfortunately, technology-oriented education is not enough in Turkey.

Training talented teachers also play an important role in Singapore's exam success. According to the 2015 TIMSS results, Singapore students are the world's best students in mathematics. The Ministry of Education Singapore trains educated teachers who can use the technology pedagogically in lessons. If Turkey wants to move to a technology-oriented education system, the government should firstly raise the quality of teachers. The quality of education is directly related to the qualifications of teachers.

Overall, the findings indicate that individual, institutional, socioeconomic factors as well as technological factors are effective in the educational performances of the countries. They also show that the technology-oriented education system with talented teachers, less homework, less extensive textbooks, and lessons that improving cognitive ability, home education resources are significant parameters to construct the education policies in Turkey for obtaining higher scores in the next TIMSS exams.

Funding

The author declares that this study has no financial support.

REFERENCES

- ADOLPHUS, K., LAWTONAND, C. L., and DYE, L. (2013). The Effects of Breakfast on Behavior and Academic Performance in Children and Adolescents. *Frontiers Human Neuroscience, 7*, 1-28.
- AKYÜZ, G. (2014). The Effects of Student and School Factors on Mathematics Achievement in TIMSS 2011. *Education and Science*. 39(172), 150-162.
- ARIKAN, S., VAN DE VIJVER, F. J. R. and YAGMUR, K. (2016). Factors Contributing to Mathematics Achievement Differences of Turkish and Australian Students in TIMSS 2007 and 2011. *Eurasia Journal of Mathematics, Science & Technology Education.* 12(8), 2039-2059.
- BOZKURT, A., ÇIRAK-KURT, S. and TEZCAN, Ş. (2019). A Comparison of the Middle School Mathematics Curricula of Turkey and Singapore with Respect to the Learning Domain of Algebra. *PAU Journal of Education.* 48(6), 152-173.
- BREIMAN, L. (1995). Better Subset Regression Using the Nonnegative Garrote. *Technometrics.* 37(4), 373-384.
- ÇAVDAR, D. (2015). *TIMSS 2011 Matematik Başarısının Öğrenci ve Öğretmen Özellikleri ile İlişkisi [The Relationship Between TIMSS 2011 Math Achievement and Characteristics of Students and Teachers]*. Thesis (PhD). Gazi University.
- CHOI, K., CHOI, T. and McANINCH, M. (2012). A Comparative Investigation of the Presence of Psychological Conditions in High Achieving Eighth Graders from TIMSS 2007 Mathematics. *ZDM.* 44 (4), 189–199.
- CIU. (2008). *Learning from Singapore*. North Carolina: The University of North Carolina.
- DEPREN, S. K., AŞKIN, Ö. E. and ÖZ, E. (2017). Identifying the Classification Performances of Educational Data Mining Methods: A Case Study for TIMSS. *Educational Sciences: Theory & Practice.* 5(10), 1605-1623.
- ERTÜRK, Z. and AKAN, O. E. (2018). TIMSS 2015 Matematik Başarısını Etkileyen Değişkenlerin Yapısal Eşitlik Modeli ile İncelenmesi [The Investigation of The Variables Effecting TIMSS 2015 Mathematics Achievement with SEM]. *Ulusal Eğitim Akademisi Dergisi (UEAD).* 2(2), 14-34.
- FAN, J. and LI, R. (2001). Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties. *Journal of the American Statistical Association.* 96(456), 1348-1360.
- FAN, J. and LI, R. (2004). New Estimation and Model Selection Procedures for Semiparametric Modeling in Longitudinal Data Analysis. *Journal of the American Statistical Association.* 99(467), 710-723.
- FİLİZ, E. and ÖZ, E. (2019). Finding the Best Algorithms and Effective Factors in Classification of Turkish Science Student Success. *Journal of Baltic Science Education*. 18, 239-253.
- FİLİZ, E. and ÖZ, E. (2020). Educational Data Mining Methods for TIMSS 2015 Mathematics Success: Turkey Case. *Sigma J Eng & Nat Science.* 38 (2), 963-977.
- GIMENEZ-GUALDO, A.M., ARNAIZ-SANCHEZ, P., CEREZO-RAMIREZ, F. and PRODOCIMO, E. (2018). Teachers' and Students' Perception About Cyberbullying. Intervention Strategies in Primary and Secondary Education. *Media Education Research Journal.* 56(26), 29-38.
- GINSBURG, A., LEINWAND, S., ANSTROM, T. and POLLOCK, E. (2005). *What the United States Can Learn from Singapore's World-Class Mathematics System (And What Singapore Can Learn from the United States): An Exploratory Study.* Ed: E.Witt. WASHINGTON, DC: American Institutes for Research.
- GÜLER, H., ŞAHİNKAYASI, Y. and ŞAHINKAYASI, H. (2017). İnternet ve Mobil Teknolojilerin Yaygınlaşması: Fırsatlar ve Sınırlılıklar [The Penetration of Internet and Mobile Information Technologies: Opportunities and Limitations]. *Sosyal Bilimler Dergisi.*7 (12), 186-207.
- HOERL, A. E. and KENNARD, R. W. (1970). Biased Estimation for Nonorthogonal Problems. *Technometrics.* 12(1), 55-67.
- JIRATCHAYUT, K. and BUMRUNGSUP, C. (2015). A study of Adaptive Elastic Net Estimators with Different Adaptive Weights. *Science & Technology Asia.* 20(3), 1-7.

- KALELİ-YILMAZ, G. and HANCI, A. (2016). Examination of the 8th Grade Students' TIMSS Mathematics Success in Terms of Different Variables. *International Journal of Mathematical Education in Science and Technology.* 47(5), 674-695.

- LEE, W. O. (2013). The Development of a Future-Oriented Citizenship Curriculum in Singapore: Convergence of Character and Citizenship Education and Curriculum 2015. *In Globalization and the Singapore Curriculum from Policy to Classroom*. Eds: Z. Deng, S. Gopinathan And C. Kim- Eng Lee. pp. 241-260. SINGAPORE: Springer, First Edition, ISBN: 978-981-4451-56-7.

- LEVENT, F. and YAZICI, E. (2014). Singapur Eğitim Sisteminin Başarısına Etki Eden Faktörlerin İncelenmesi [Examination of Factors Affecting Success of Singapore Education System]. *Journal of Educational Sciences.* 39(January), 121-143.

- LITTLECOTT, H. J., MOORE, G. F., MOORE, L., LYONS, R. A. and MURPHY, S. (2015). Association Between Breakfast Consumption and Educational Outcomes in 9–11-Years-old Children. *Public Health Nutrition.* 19(9), 1575–1582.

- LUNDQVIST, M., VOGEL, N. E. and LEVIN, LARS-A. (2019). Effects of Eating Breakfast on Children and Adolescents: A Systematic Review of Potentially Relevant Outcomes in Economic Evaluations. *Food & Nutrition Research*. 63(9), 1-15.

- MARTIN, M.O., MULLIS, I.V.S. and HOOPER, M. (ed.) (2016). *Methods and Procedures in TIMSS 2015.* C.A., Boston: Boston College, ISBN: 978-1-889938-29-5.

- MULLIS, INA V.S., MARTIN, M. O., FOY, P. and HOOPER, M. (ed.) (2016). *TIMSS 2015 International Results in Mathematics.* C.A: TIMSS & PIRLS International Study Center, Boston College.

- NYE, B., KONSTANTOPOULOS, S. and HEDGES, L., V. (2004). How Large are Teacher Effects? *Educational Evaluation and Policy Analysis.* 26(3), 237-257.

- OECD. (2012). *Lessons from PISA for Japan, Strong Performers and Successful Reformers in Education*. Paris: OECD.

- POLAT, E. and MADRA, A. (2018). *PISA 2015 ve TIMSS 2015 Işığında Türkiye'de Cinsiyete Dayalı Başarı Farkı* [The Gender-Based Differences in Achievement in Light of PISA 2015 and TIMSS 2015 in Turkey]. ADV and ERG. [Online] https://www.egitimreformugirisimi.org/pisa-ve-timss-2015-bulgulari-isiginda-turkiyede-cinsiyete-dayali-basari-farki-bilginotu/ [Date accessed:1/5/2020].

- SHEN, X. and YE, J. (2002). Adaptive Model Selection. *Journal of the American Statistical Association.* 97(457), 210-221.

- TIBSHIRANI, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of Royal Statistical Society.* 58(1), 267-288.

- TOPÇU, M.S., ERBILGIN, E. and ARIKAN, S. (2016). Factors Predicting Turkish and Korean Students' Science and Mathematics Achievement in TIMSS 2011. *Eurasia Journal of Mathematics, Science & Technology Education.* 12(7), 1711-1737.

- TUİK (2013). *Household Information Technology Usage Survey, 2013*. İstanbul,Turkey [online] from http://tuik.gov.tr/PreHaberBultenleri.do?id=13569, [Date Accessed: 4/1/2020].

- TUİK (2015). *Household Information Technology Usage Survey, 2015.* İstanbul, Turkey [online] http://tuik.gov.tr/PreHaberBultenleri.do?id=18660, [Date Accessed: 4/1/2020].

- UNESCO. (2017). *Cracking The Code: Girls' and Women's Education in Science, Technology, Engineering and Mathematics (STEM).* France, Paris: UNESCO.

- WILSON, K. and NARAYAN, A. (2016). Relationships Among Individual Task Self-Efficacy, Self-Regulated Learning Strategy Use and Academic Performance in A Computer-Supported Collaborative Learning Environment. *Educational Psychology.* 36(2), 236–253.

- YAVUZ, H. Ç., DEMIRTAŞLI, R. N., YALÇIN, S. and DIBEK, M. İ. (2017). The Effects of Student and Teacher Level Variables on TIMSS 2007 and 2011 Mathematics Achievement of Turkish Students. *Education and Science.* 42(189), 27-47.

- YILDIRIM, A., ÖZGÜRLÜK, B., PARLAK, B., GÖNEN, E. and POLAT, M. (2016). *TIMSS 2015 Ulusal Matematik ve Fen Bilimleri Ön Raporu: 4. ve 8. Sınıflar [The National Preliminary Report of TIMSS 2015 Mathematics and Science: 4th and 8th Grades]*. Ankara: MEB.

- YILMAZ, F. and ŞEKERCİ, H. (2016). The Issue of Mother Tongue: Problems of Students Based on the Experience of Primary School Teachers. *Journal of Qualitative Research in Education*. 4(1), 47-63.

- YOO, J. E. and RHO, M. (2017). TIMSS 2015 Korean Student, Teacher, and School Predictor Exploration and Identification via Random Forests. *The SNU Journal of Education Research.* 26 (12), 43-61.

- YOO, J. E. (2018). TIMSS 2011 Student and Teacher Predictors for Mathematics Achievement Explored and Identified via Elastic Net. *Frontiers in Psychology*. 9, 1-10.

- ZOU, H. and HASTIE, T. (2005). Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society*. Series(B) 67, 301–320.

- ZOU, H. (2006). The Adaptive Lasso and its Oracle Properties. *Journal of the American Statistical Association*. 101(476), 1418-1429.

- ZOU, H. and ZHANG, H. H. (2009). On the Adaptive Elastic-net with a Diverging Number of Parameters. *The Annals of Statistics*. 37(4), 1733–1751.

"This page is left blank for typesetting"

JAME
E-ISSN:2791-7401
Volume : 1, Issue 1, 2021
URL: https://journals.gen.tr/jame

Received: 28.05.2021
Acccepted: 22.06.2021

RESEARCH ARTICLE

The behavior of capital structure: evidence from fast calibrated additive quantile regression

Umut UYAR[1]

1 Pamukkale University, Department of Business Administration, Turkey e-mail: uuyar@pau.edu.tr

Abstract

In finance, capital structure decisions are crucial due to their impact on the value of a firm. Some theories assert that the value of a firm is irrelevant to those decisions. However, there is a growing literature that criticizes this idea. Those studies are constructed on some modern theories, which are called trade-off theory, agency cost theory, signaling theory, and pecking order theory. This paper investigates the relationship between optimal capital structure and capital structure components. The annual data gathered from 195 firms traded in Borsa Istanbul for the period 2011-2020 is used. The fast calibrated additive quantile regression approach is chosen because of its superior properties. In that method, there is not a strong assumption about the functional form of the relationships between the dependent variable and the explanatory variables. The results indicate that the relationships between the debt ratios and the capital structure components differ for each quantile and these relations are nonlinear. Furthermore, evidence is provided for the fact that the relationships might be explained with the modern theories of capital structure.

Keywords: *Capital structure, Additive quantile regression, Non-parametric regression, Borsa Istanbul.*

JEL Codes: *G30, G31, C50*

Citation: UYAR, U., (2021). The Behavior of Capital Structure: Evidence from Fast Calibrated Additive Quantile Regression, Journal of Applied Microeconometrics (JAME). 1(1), 55-69, DOI: 10.53753/jame.1.1.05

Corresponding Author:
Umut Uyar
E-mail: uuyar@pau.edu.tr

1. INTRODUCTION

Capital structure decisions of firms are a topic of discussion in the financial literature for a half a century. The modern theory of capital structure started with Modigliani and Miller (1958) and was studied by many researchers (Bradley et al. 1984; Albayrak and Akbulut 2008; Demirhan 2009; Mac an Bhaird 2010; Brusov et al. 2011; Aboura and Lépinette 2013; Ahmeti and Prenaj 2015; Chang 2015: 17; Jaros and Bartosova 2015; Krstevska, et al. 2017; Al-Kahtani and Al-Eraij 2018; Onyinyechi 2019; Sibarani 2020). The firm's capital structure decision is based on the proportion of the debt and equity mix used in financing the assets. Reaching an optimal capital structure that maximizes the value of the firm is the main purpose of that research. While some of the studies argue that there is no optimal capital structure that maximizes firm value, others emphasize that debt and equity mix is directly related to the firm value (Baker and Martin 2011: 2). The pioneering researchers Modigliani and Miller show that the value of a firm is irrelevant of its capital structure under stringent conditions of competitive, frictionless, and complete capital markets. Therefore, financial managers cannot maximize the firm value by the capital structures that they choose. The counter idea of Modigliani and Miller indicates that managers might decide and calculate a firm's optimal capital structure. The assumptions of Modigliani and Miller are criticized after they proposed their theory. Thus, researchers have relaxed the restrictive assumptions and proposed new theories: trade-off theory (Kraus and Litzenberger 1973), agency cost theory (Ross 1973; Mitnick 1974), signaling theory (Ross 1977), and pecking order theory (Myers 1984; Myers and Majluf 1984). These theories relate directly to taxes, asymmetric information, agency problems, and bankruptcy costs. These theories may fail to explain absolute facts about the capital structure. Even though the existence of extensive research into the area of capital structure, determining the accurate dept and equity mix that maximizes the firm's market value is still incomprehensible.

The trade-off theory indicates that there is an optimal debt and equity mix where firm value is maximized. This can be reached by identifying a balance between several benefits of issuing debt and equity. One of these benefits is lower issuance costs, another is the tax shield. The agency cost theory provides a further theoretical scheme that supports the influence of diversification strategy on capital structure (Kochhar and Hitt 1998). Based on this theory, debt has a consistent role in lightening the overinvestment behavior of financial managers. Therefore, this situation supports diversification on the debt and equity and leads the managers to an optimal capital structure. The signaling theory suggests that profitable firms should run into debt more to convince investors of how high the firm's future profits will be. The theory predicts that a firm's stock price should rise when it issues debt and fall when it issues equity (Gitman and Zutter 2012: 534). The pecking order theory propose a hierarchy of financing that begins with retained earnings, which is followed by debt financing and finally external equity financing. The theory posits that there is no optimal debt ratio, by contrast, firms will not use debt when there is still sufficient internal financing (Wei 2014).

Although the validity of those theories is tested many times in the literature, they cannot provide a clear relationship between the decisions and optimal capital structure. The studies, which tested the efficacy of those theories, have an important problem. They have a strong linearity assumption between optimal capital structure and factors that affect capital structure decisions. Yet, in nonparametric methods, there is not a strong assumption about the functional form of the relationships between the dependent variable and the explanatory variables. Especially, the fast-calibrated additive quantile regression approach avoids the model specifications errors arising from determining the wrong functional form.

The purpose of the study is to examine the capital structure of 195 firms traded in Borsa Istanbul for the period 2011-2020. The fast calibrated additive quantile regression approach is chosen due to its some superior properties against parametric approaches. We focus to explore whether the theories of firm financing (trade-off theory, agency cost theory, signaling, and pecking order theory) can explain the capital structures. We employ the total debt and long-term debt in our models as dependent variables because the theories have various empirical implications regarding various types of debt instruments.

This study is organized as follows: the literature review about the topic is given in Section 2. Section 3 presents the methodology. Data are introduced in Section 4. Section 5 presents the findings; and the conclusions are given in the last section.

2. LITERATURE REVIEW

Searching for the optimal capital structure on firm value is a challenge in finance literature. The topic is studied many times by researchers. The existing literature points out some capital structure components that might impact the value of firms. They are profitability, size, tangibility, and growth rate.

Profitability variable is a proxy for earning power of a firm. In the literature, return on assets is used as profitability variable. It can be calculated by dividing net income by total assets of a firm. While a positive relationship is expected between leverage and profitability according to the trade-off and signaling theories, a negative relationship is excepted according to the pecking order theory (Kester 1986; Friend and Lang 1988; Titman and Wessel 1988; Barton et al. 1989; Demirguc-Kunt and Maksimovic 1994; Rajan and Zingales 1995; Jordan et al. 1998; Booth et al. 2001; Al-Sakran 2001; Bevan and Danbolt 2002; Bauer 2004; Chen 2004; Huang and Song 2006; Allen and Powell 2013; Sakti et al. 2017; Al-Hunnayan 2020; Assfaw 2020; Harun et al. 2020).

Size variable is indicated as the natural logarithm of the total assets in finance. It is usually used to fit firms into a common size measure because firms in the different sectors can vary greatly in terms of size. Although a negative relationship is expected between debt ratios and size of a firm in term of pecking order theory, the expected relationship is positive according to the other theories (Kester 1986; Kim and Sorensen 1986; Titman and Wessels 1988; Friend and Lang 1988; Barton et al. 1989; MacKie-Mason 1990; Rajan and Zingales 1995; Barclay and Smith 1996; Kim et al.1998; Wiwattanakantang 1999; Booth et al. 2001; Al-Sakran 2001; Bevan and Danbolt 2002; Hovakimian et al. 2004; Huang and Song 2006; Al-Mutairi and Naser 2015; Sakti et al. 2017; Ghosh and Chatterjee 2018; Assfaw 2020; Harun et al. 2020).

Tangibility is calculated by dividing net fix assets by total assets of a firm. Net fix assets in the formula indicate the noncurrent assets minus depreciation. According to the trade-off and pecking order theory, the relationship between leverage and tangibility is positive. However, the agency cost theory states that the relationship might be positive or negative (Titman and Wessels 1988; Van der Wijst and Thurik 1993; Rajan and Zingales 1995; Wiwattanakantang 1999; Booth et. al. 2001; Drobetz and Fix 2003; Hall et. al. 2004; Huang and Song 2006; Heyman et. al. 2008; Al-Mutairi and Naser 2015; Alkhazaleh and Almsafir 2015; Sakti et al. 2017; Ghosh and Chatterjee 2018; Al-Hunnayan 2020; Assfaw 2020; Harun et al. 2020).

Growth rate is measured as the annual change of the last three years of a firm's total assets. It is a major indicator for characterizing a firm as aggressive or conservative. The relationship between debt ratios and growth rate is negative according to all theories except for the pecking order theory (Kim and Sorensen 1986; MacKie-Mason 1990; Barclay and Smith 1996; Kim et al. 1998; Friend and Lang 1998; Al-Sakran 2001; Bevan and Danbolt 2002; Cai et. al.2008; Allen and Powell 2013; Al-Mutairi and Naser 2015; Alkhazaleh and Almsafir 2015; Sakti et al. 2017; Al-Hunnayan 2020; Assfaw 2020; Harun et al. 2020).

Nearly all studies in the literature use parametric models in their analyses. Moreover, most of them set up their methodologies under the linearity assumption. However, the relationship between capital structure components and leverage might not be linear, especially for the financial data. In these circumstances, making an assumption about the functional form of the relationship between variables might not be correct. Nonparametric approaches do not make any assumption about the functional form and they can be useful to find the appropriate relationship.

3. METHODOLOGY

In this study, we use the fast calibrated additive quantile regression approach introduced by Fasiolo et al. (2020). The methodology grounds on the traditional quantile regression approach introduced by Koenker and Bassett (1978). Traditional quantile regression allows us to examine the relationship between the dependent variable (y) and k-dimensional vector of explanatory variables (x) for the different parts (quantiles,$\tau \in (0,1)$) of the dependent variable's conditional distribution. When $F(y|x)$ is the conditional cumulative distribution function (c.d.f.) of y, the τ^{th} quantile of the conditional distribution of y or τ^{th} conditional quantile is defined as. $\mu = F^{-1}(\tau|x) = inf\{y:F(y|x) \geq \tau\}$ The aim is to obtain the τ^{th} conditional quantile estimation which minimizing the following function called expected loss:

$$L(\mu|x) = E\{\rho_\tau * (y-\mu)|x\} = \int \rho_\tau * (y-\mu) dF(y|x) \quad (1)$$

where μ=μ(x) and is the control function or pinball loss that might be defined as follows:

$$\rho_\tau = (\tau - 1)(y - \mu(x))I(y - \mu(x) < 0) + \tau I(y - \mu(x) \geq 0) \quad (2)$$

In the context of the linear regression model, since μ̂(x) is equal to x'θ̂ , expected loss function is revised and the quantile estimator in Eq. 3 is obtained:

$$\hat{\theta} = \underset{\theta}{argmin} \frac{1}{n} \sum_{i=1}^{n} \rho_\tau \{y_i - x_i' \hat{\theta}\} \quad (3)$$

where x_i is the i^{th} vector of explanatory variables and θ is the vector of regression coefficients.

While traditional quantile regression assumes that the relationship between y and x is linear, the additive quantile regression does not make an assumption about the functional form of the relationship between variables. Thus μ(x), has an unknown functional form in the additive quantile regression. In the latter approach, inferences about the functional form are made from the data, that it provides a flexible approach about determining the functional form. Furthermore, μ(x) has an additive structure and so the effect of each explanatory variable on dependent variable for each quantile is assumed separate:

$$\mu(x) = \sum_{j=1}^{m} f_j(x) \quad (4)$$

In Eq. 4, the *f* function refers to the nonparametric functions of the explanatory variables. These nonparametric functions can be defined in terms of spline basis:

$$f_j(x) = \sum_{i=1}^{r} \beta_{ji} b_{ji}(x_j) \quad (5)$$

where, β_{ji} the coefficients to be estimated and $b_{ji}(x_j)$ are the spline basis functions. Spline is one of the nonparametric methods that consider nonlinear relationships between dependent variable and explanatory variables. It is based on a piecewise linear regression model. In this model, the regression line is estimated for each sample subgroup by dividing the sample into subgroups. The piecewise linear regression model is obtained by combining these lines. However, the first-order derivatives of functions used in definition of regression lines are not continuous since the junction points of combined lines, that is, the jumping points, are discrete. To eliminate this problem, spline basis functions are used.

In Eq. 5, is the spline basis dimension and chosen that we guarantee avoiding over-smoothing. f_j is controlled by penalizing the deviations from f_j and the penalty term is applied on β_{ji} Thereafter, penalized pinball loss can be defined as follows

$$V(\beta, \lambda, \sigma) = \sum_{i=1}^{n} \frac{1}{\sigma} \rho_\tau \{y_i - \mu(x_i)\} + \frac{1}{2} \sum_{j=1}^{m} \lambda_j \beta' S_j \beta \quad (6)$$

where $\lambda=\{\lambda_1,\lambda_2,\ldots,\lambda_m\}$ is the vector of smoothing parameters. $1/\sigma$ is the learning rate which balances between the loss and the penalty. S_j's are positive semidefinite matrices, and they penalize the oscillations of the corresponding effect . The minimization of Eq. 6 with respect to β for fixed λ and gives the maximum a posteriori (MAP) estimator, that is β Consequently, in the additive quantile regression approach, the estimation of the nonparametric functions or f_j for each quantile is obtained by minimizing the Eq. 6. The optimal selection of λ and is discussed detailed in Fasiolo et al. (2020).

One of the important problems that can be encountered in the studies on panel data is poolability problem. Poolability problem is related to the question of whether the relationships between variables change over time. Some parametric tests have been developed regarding whether the panel data can be pooled (Hsiao's F test (2007), etc.). However, the nonparametric poolability test developed by Baltagi et al. (1996) is defined as a robust test against model identification errors caused by a wrong functional form. In this study, the nonparametric poolability test developed by Baltagi applied to test whether the panel data can be pooled or not. The hypotheses of the nonparametric poolability test can be expressed as follows:

$$H_0: f_{it}(x) = f_i(x)\ ,\ H_1: f_{it}(x) \neq f_i(x) \tag{7}$$

where H_0 hypothesis states that the relationship between dependent and explanatory variables does not change with time. The test statistic has a standard normal distribution, N (0, 1), and the test is one-sided. According to the result of this test, when the null hypothesis cannot be rejected, the models to be estimated are shown in Eq. 8 and 9, respectively:

$$TD_{it} = \gamma + f_{5it}(PROFIT_{it}) + f_{6it}(SIZE_{it}) + f_{7it}(TANG_{it}) + f_{8it}(GROWTH_{it}) + u_{it} \tag{8}$$

$$LTD_{it} = \alpha + f_{1it}(PROFIT_{it}) + f_{2it}(SIZE_{it}) + f_{3it}(TANG_{it}) + f_{4it}(GROWTH_{it}) + \varepsilon_{it} \tag{9}$$

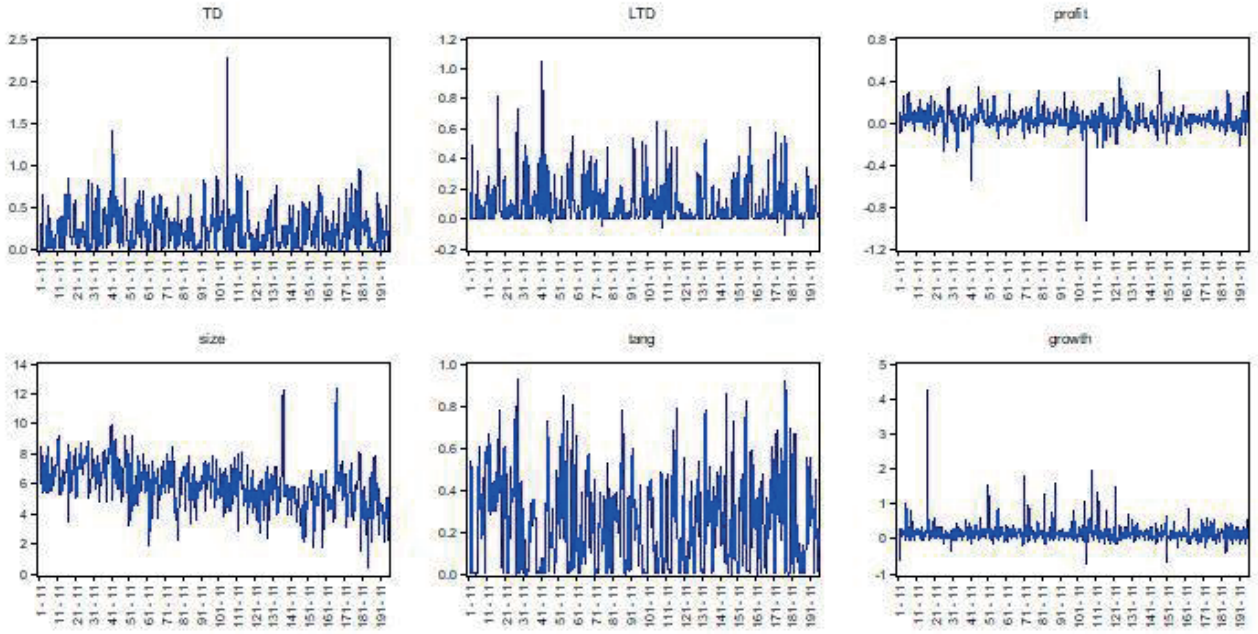
4. DATA

The panel dataset contains the financial information from 195 firms traded in Borsa Istanbul for the period 2011-2020. Financial and insurance sector firms are excluded from the data due to the different financial statement structures. To examine the capital structure, while the debt ratio and long-term debt ratio are used as dependent variables; profitability, size, tangibility, and growth rate are used as explanatory variables. Each variable gathered from the financial statements of firms. Calculations are done based on the previous three years' average amounts of each account. The descriptive statistics of variables are presented in Table 1. Also, the graphics are demonstrated in Figure 1.

Table 1. Descriptive Statistics

	Debt ratio (TD)	Long-term debt ratio (LTD)	Profitability (PROFIT)	Size (SIZE)	Tangibility (TANG)	Growth (GROWTH)
Mean	0.2320	0.0980	0.0339	5.8806	0.2419	0.1611
Median	0.1774	0.0383	0.0236	5.8523	0.2304	0.1318
Maximum	2.2721	1.0531	0.4909	12.329	0.9302	4.2683
Minimum	0.0000	-0.1126	-0.9386	0.4201	0.0000	-0.7016
Std. Dev.	0.2262	0.1340	0.0936	1.4456	0.2108	0.2317
Skewness	1.5311	1.8947	-1.0155	0.3549	0.5813	6.8150
Kurtosis	8.5150	7.3842	17.3355	4.5687	2.5780	95.7788
Jarque-Bera	3233.28	2728.52	17032.78	240.90	124.2982	714487.3
Observations	1950	1950	1950	1950	1950	1950

Figure 1. Graphics of All Variables



When summary statistics in Table 1 are examined, the maximum value of debt ratios is remarkable. If the debt ratio has a value over one, that means liabilities are exceeding the equities for some firms. It means that some of firms have overdose debt, so, those observations might create outliers in the sample. Moreover, high standard deviations indicate the existence of outliers for all variables. The outliers might be observed in graphics of all variables in Figure 1, as well. Another remarkable result is the maximum value of tangibility. The maximum value of the tangibility variable is 93% and it means that nearly all assets of some firms consist of fixed or noncurrent assets. Considering that the data contains firms from different sectors, this result might be evaluated as an expected situation. However, the summary statistics provide us some evidence that we should use quantile regression to estimate a model with our dataset. On the other hand, the calculation formulas of each variable are shown in Table 2.

Table 2. Formulas for Variables

Debt ratio (TD)	Long-Term Debt / Total Assets
Long-term debt ratio (LTD)	Total Debt / Total Assets
Profitability (PROFIT)	Net Income / Total Assets
Size (SIZE)	Logarithmic Total Assets
Tangibility (TANG)	Net Fixed Assets / Total Assets
Growth (GROWTH)	$Ln(Total\ Assets_{t}) - Ln(Total\ Assets_{t-1})$

5. FINDINGS

In this section, firstly the theoretical sign expectations related to the capital structure components are summarized on the grounds that the literature review section (Table 3). Then, we examine whether the panel data is poolable or not by applying the poolability test introduced by Baltagi et al. (1996). Finally, the nonparametric estimation results are depicted in Figure 1 to 4. Panel A in all Figures shows the nonparametric estimation graphs for each explanatory variable in Eq.8, while Panel B in all Figures demonstrates the nonparametric estimation graphs for each explanatory variable in Eq.9. Moreover, for different quantiles, the nonparametric estimation graphs of profitability, size, tangibility, and growth are given from Figures 1 to 4, respectively. In the Figures, while Q25 represents the firms in the 25% quantile with the lowest debt level, Q75 represents the firms in the 75% quantile with the highest debt level.

Table 3. The Theoretical Expectations of Variables

Variables	Trade-Off Theory	Agency Cost Theory	Signaling Theory	Pecking Order Theory
Profitability	+	?	+	-
Size	+	+	+	-
Tangibility	+	+/-	?	+
Growth	-	-	-	+

Note: (+) sign indicates that there is a positive relationship between debt ratios and variables, (-) sign indicates a negative relationship, and (?) indicates that there is no certainty about the direction of the relationship.

Source: Frank and Goyal 2009.

The nonparametric poolability test statistics for debt ratio and long-term debt ratio are 0.588 and 0.460, respectively. When these values are compared with the critical value of 1.645 in the standard normal distribution table, it can be suggested that H_0 hypothesis cannot be rejected. Therefore, panel data generated for both debt ratio and long-term debt ratio can be pooled. The models in Eq. 8 and 9 can be used for estimation.

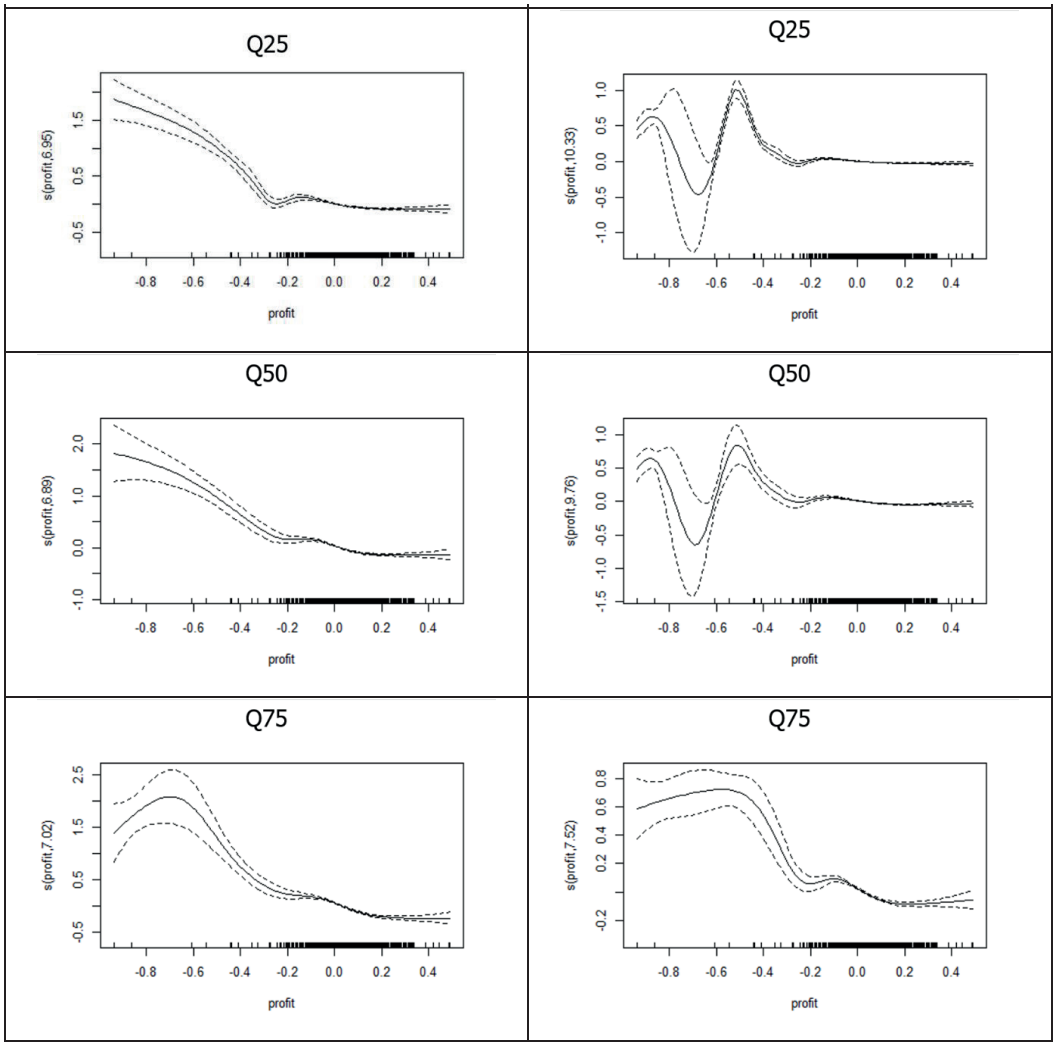
Figure 2. Estimation Results for Profitability

Panel A:

Dependent variable is Debt Ratio

Panel B:

Dependent variable is Long-Term Debt Ratio



In Figure 2, both for Panel A and B, there are negative and non-linear relationships between profitability and debt and long-term debt ratios for each quantile. In Panel A, the negative relationship is obvious and is consistent with the Pecking Order Theory. However, there is a sudden rising debt ratio in Q50 and Q25 while the lost level of firms increases. The same path can observed in Panel B and this finding is related to the Signaling Theory. There is a threshold in Q75 when the firms' profitability level reaches deep. The 75th quantile represents the firms that have very high-level debt ratios. Thus, firms with both high debt ratios and low profitability must reduce their debt level after a certain level due to the increased risk. This finding is related to the Trade-Off Theory. In Panel B, as the profitability level of firms increases, they choose the lower borrowing path. This finding is also consistent with the Pecking Order Theory. Moreover, the debt level is rising for the firms in the lower quantile (Q25). The reason for this is due to the desire of companies that want to give a strong company signal in accordance with the Signaling Theory, not to increase their total debt level, but to increase their long-term debt level. They are trying to send a signal to the market "We are a strong firm, and we can find easily debt". On the other hand, there are thresholds in each quantile when the firms' long-term debt level decreased. The reason for those findings is the Trade-Off Theory, as well.

Figure 3. Estimation Results for Size

Panel A:
Dependent variable is Debt Ratio

Panel B:
Dependent variable is Long-Term Debt Ratio

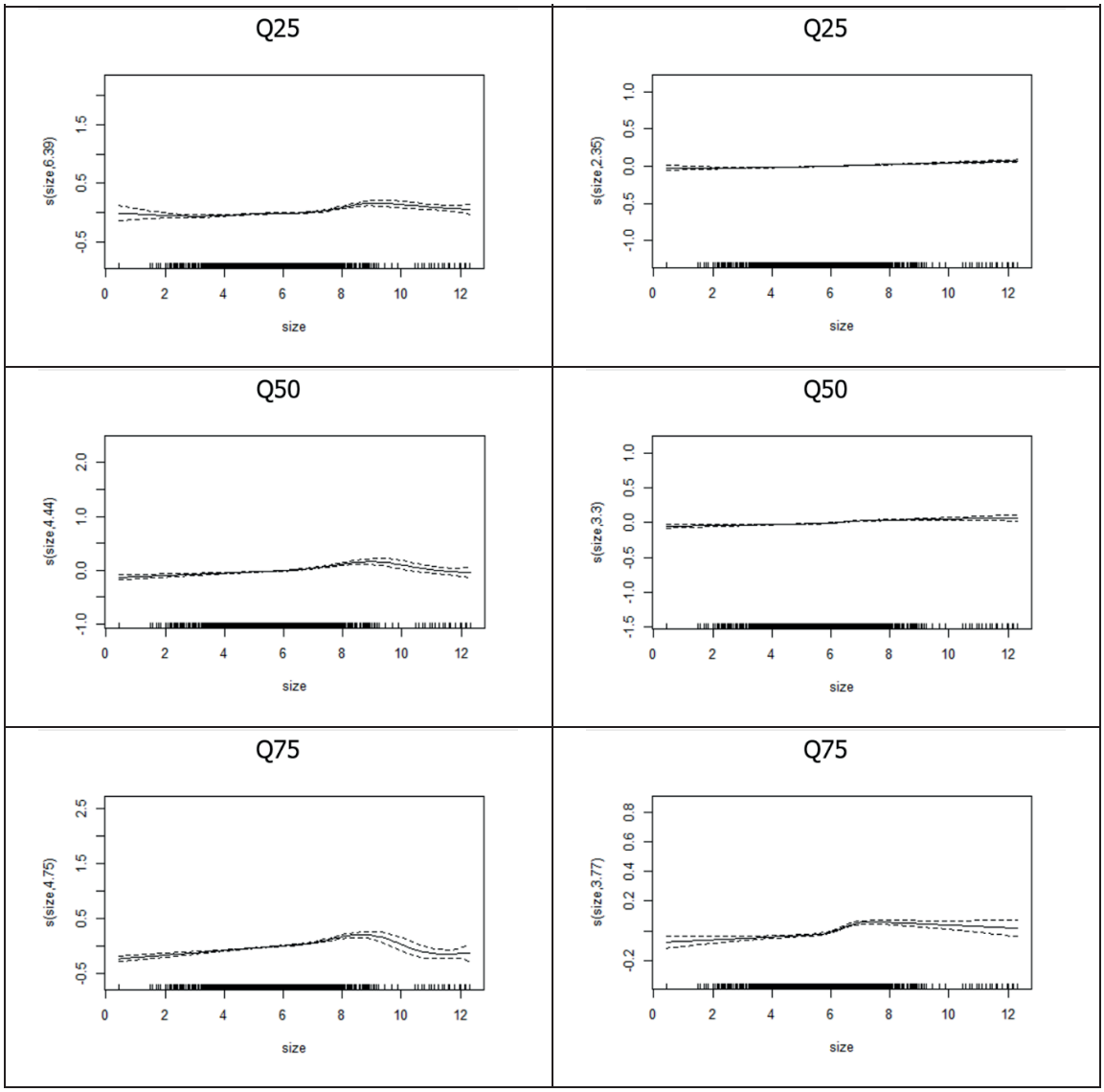
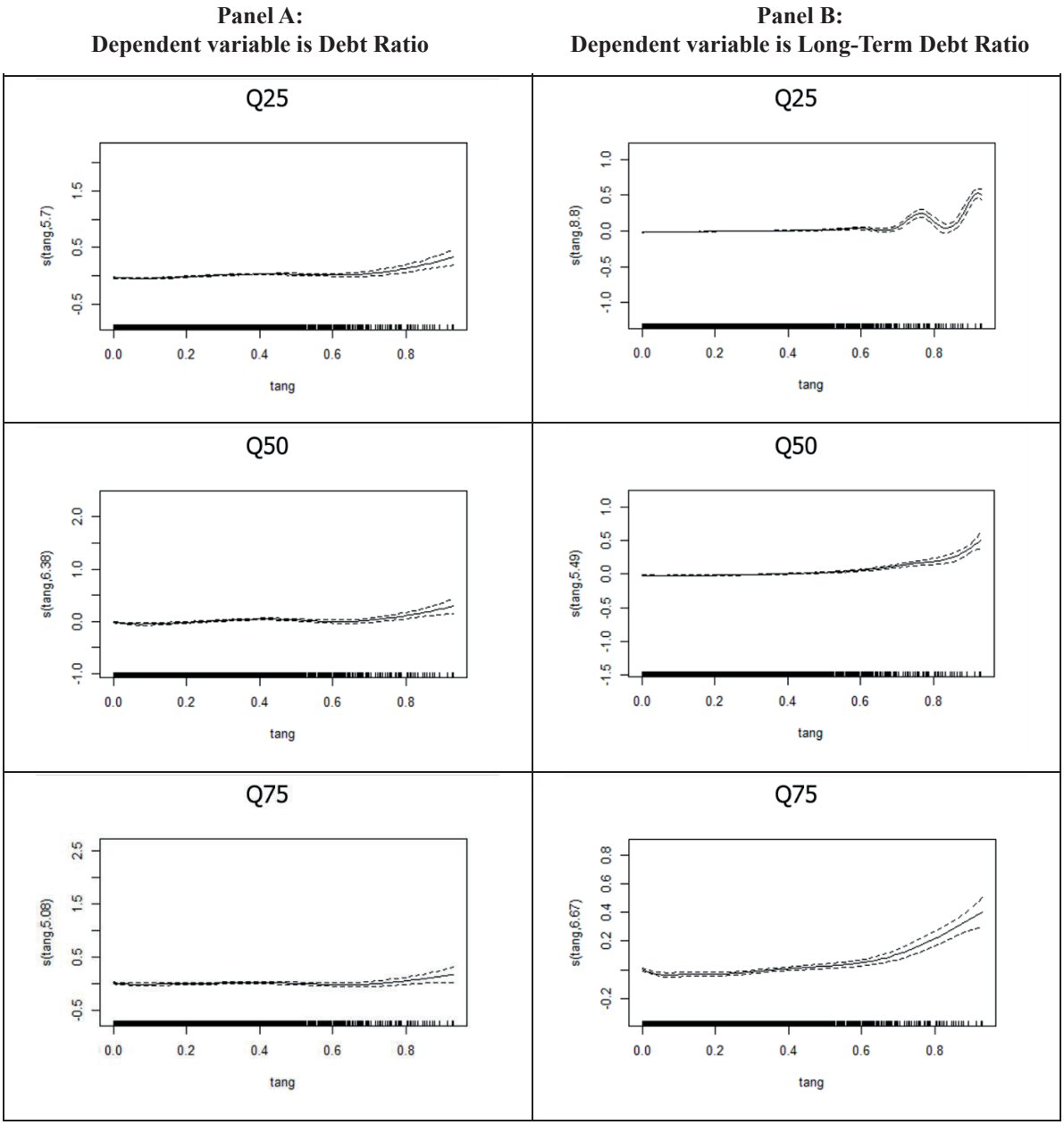


Figure 3 (both Panel A and B) suggests that there are slightly positive relationships between size and debt and long-term debt ratios in all quantiles. In Panel A, the positive relationship is clear. Growing firms increase their debt level in order to show themselves financially strong (related with the Signaling Theory) and gain the benefit of the tax advantage of debt (related with the Trade-off Theory). There are also thresholds in each quantile when the firms reached a certain size level. Those firms need a high debt ratio to finance their growth, however, the debt ratio is falling after a sufficient level because of increased risk. This finding is consistent with the Pecking Order Theory. In Panel B, the positive relationship between size and long-term debt ratio seems weak, except for Q75. One can easily observe the existence of the Pecking Order Theory in Q75.

Figure 4. Estimation Results for Tangibility



In Figure 4, there is a flat and nonlinear relationship between the tangibility and the debt ratio in all quantiles of Panel A. In lower tangibility levels, the magnitude is nearly zero for the Q50. However, the debt ratio is rising when the firms reached high tangibility levels in all quantiles. This finding is consistent with all theories in the literature (Table 3). In Panel B, there is a positive and nonlinear relationship between the tangibility and the long-term debt ratio in all quantiles. The positive relationship is very clear in high tangibility levels. The tangibility variable represents the proportion of fixed (long-term) assets in total assets. One might expect that the investment

of long-term assets is financed by long-term liabilities. Thus, the increase in long-term debts in proportion to the increase in tangibility level is interpreted as a situation which is suitable for the financing logic. Those findings are consistent with the Trade-Off, the Agency Cost, the Pecking Order Theories.

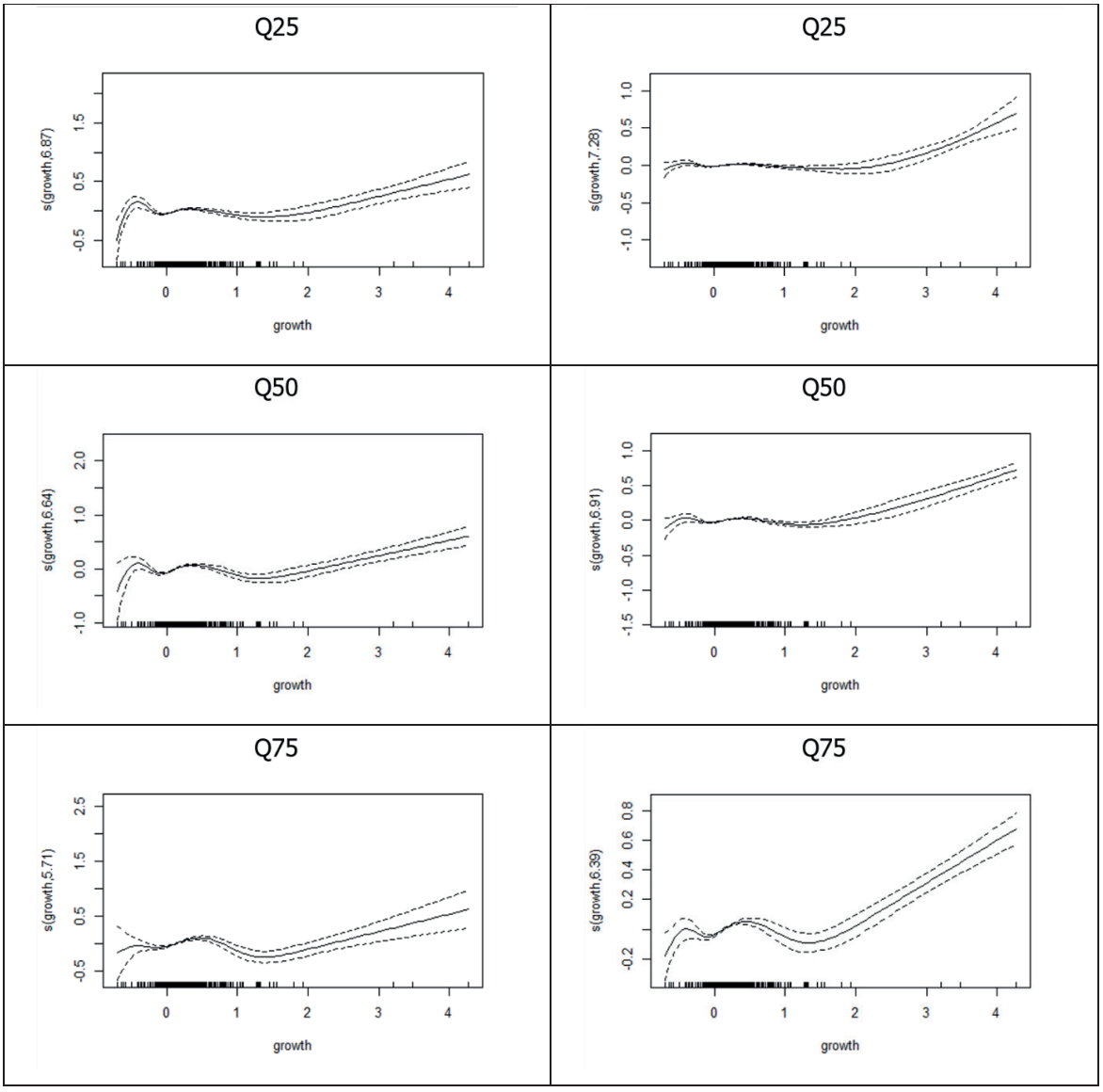
Figure 5. Estimation Results for Growth

Panel A:

Dependent variable is Debt Ratio

Panel B:

Dependent variable is Long-Term Debt Ratio



In Figure 5, there is a positive and nonlinear relationship between the growth rate and the debt and long-term debt ratio in all quantiles of Panel A and B. The relationship between debt and long-term ratio and growth, which is flat up to a certain level, becomes positive afterward. High growth rates have an increasing effect on debt ratios. The findings show that companies are ordered hierarchically from internal resources (retained earnings) to external resources (debts), according to the Pecking Order Theory.

6. CONCLUSION

One of the questions that have plagued the finance literature for a long time is whether there exists an optimal capital structure, or not. The pioneering research about the topic is by Modigliani and Miller (1958). They indicate

a new perspective on optimal capital structure in their paper. Using arbitrage arguments, they state that the capital structure decisions do not matter under very restrictive assumptions. However, many theories reveal that the opposite of this idea exists. The major ones are the trade-off theory, the agency cost theory, the signaling theory, and the pecking order theory. The validity of these theories has been tested many times in the financial literature. Yet, they fail to execute a certain and clear relationship between capital structure decisions and optimal capital structure. The methods testing the validity of these theories have a crucial problem. They have a strong assumption that the relationship between optimal capital structure and the factors that affect capital structure decisions is linear.

In nonparametric methods, there is no a priori assumption regarding the functional form of the relationships between the dependent variable and the explanatory variables. Functional form flexibility in the fast calibrated additive quantile regression approach averts model specification errors arising from the wrong functional form. In this regard, this approach has superior features to the traditional quantile regression approach.

The aim of the study is to investigate the optimal capital structure of 195 firms traded in Borsa Istanbul for the period 2011-2020. Particularly, we would like to examine whether the theories of firm financing (trade-off theory, agency cost theory, signaling, and pecking order theory) can explain the optimal capital structures or not. According to the existing literature, the total debt and long-term debt ratios are selected as dependent variables; profitability, size, tangibility, and growth rate (which called capital structure components) are used as explanatory variables.

The results show that the relationships between the debt ratios and the capital structure components differ for each quantile and these relations are obviously nonlinear. Moreover, the relationships might be explained with the modern theories of capital structure. While the behavior of growth rate variable is related to the Pecking Order Theory, the behavior of profitability and size variables are consistent with the Trade-Off, the Signaling, and the Pecking Order Theories. Furthermore, the behavior of the tangibility variable is consistent with the Trade-Off, the Agency Cost, the Pecking Order Theories.

Funding

The author declares that this study has no financial support

REFERENCES

- ABOURA, S. and LEPINETTE, E. (2013). Do Banks Satisfy the Modigliani-Miller Theorem? [online] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2348608 [Date Accessed: 5/15/2021]

- AHMETI, F. and PRENAJ, B. (2015). A Critical Review of Modigliani and Miller's Theorem of Capital Structure. *International Journal of Economics, Commerce and Management*. 3(6), 914-924.

- ALBAYRAK, A.S. and AKBULUT, R. (2008). Factors That Affect Capital Structure: The Analysis on the Firms Registered in Industrial and Service Sector of ISE. *Dumlupınar Universitesi Sosyal Bilimler Dergisi*, 22, 425-446.

- AL-HUNNAYAN, S. H. (2020). The Capital Structure Decisions of Islamic Banks in the GCC. *Journal of Islamic Accounting and Business Research*, 11(4), 745-764.

- AL-KAHTANI, N. and AL-ERAIJ, M. (2018). Does Capital Structure Matter? Reflection on Capital Structure Irrelevance Theory: Modigliani-Miller theorem (MM 1958). *International Journal of Financial Services Management*. 9(1), 39-46.

- ALKHAZALEH, A.M. and ALMSAFIR, M. K. (2015). Does Asymmetry of Information Drive Banks' Capital Structure? Empirical Evidence from Jordan. *International Journal of Economics and Finance*, 7(3), 86-97.

- ALLEN, D. E., and POWELL, R. J. (2013). The Determinants of Capital Structure: Empirical Evidence from Thai Banks. *Information Management and Business Review*, 5(8), 401-410.

- AL-MUTAIRI, A. and NASER, K. (2015). Determinants of Capital Structure of Banking Sector in GCC: An Empirical Investigation. *Asian Economic and Financial Review*, 5(7), 959-972.

- AL-SAKRAN, S. (2001). Leverage Determinants in The Absence of Corporate Tax System: The Case of Non-financial Publicly Traded Corporation in Saudi Arabia. *Managerial Finance*, 27(10-11), 58-86.

- ASSFAW, A.M. (2020). The Determinants of Capital structure in Ethiopian Private Commercial Banks: A Panel Data Approach. *Journal of Economics, Business, and Accountancy Ventura*, 23(1), 108-124.

- BAKER, H. and MARTIN, G. (2011). *Capital Structure and Corporate Financing Decisions: Theory, Evidence, and Practice - The Robert W. Kolb Series in Finance*. New York: John Wiley & Sons, Inc.

- BALTAGI, B.H., HIDALGO J. and LI, Q. (1996). A Nonparametric Test for Poolability Using Panel Data. *Journal of Econometrics*. 75, 345–67.

- BARCLAY, M.J. and SMITH, C.W. (1996). On Financial Architecture: Leverage, Maturity and Priority. *Journal of Applied Corporate Finance*, 8(4), 4-17.

- BARTON, S.L., NED, C.H. and SUNDARAM, S. (1989). An Empirical Test of Stakeholder Theory Predictions of Capital. *Financial Management*, 18(1), 36-44.

- BAUER, P. (2004). Determinants of Capital Structure Empirical Evidence from the Czech Republic. *Czech Journal of Economics and Finance (Finance a uver)*, 54(1-2), 2-21.

- BEVAN, A. and DANBOLT, J. (2002). Capital Structure and Its Determinants in the UK-a Decompositional Analysis. *Applied Financial Economics*, 12(3), 159-170.

- BOOTH, L., AIVAZIAN, V., DEMIRGUC-KUNT, A. and MAKSIMOVIC, V., (2001). Capital Structure in Developing Countries. *Journal of Finance*, 56(1), 87-130.

- BRADLEY, M., JARRELL, G.A. and KIM, E.H. (1984). On the Existence of an Optimal Capital Structure: Theory and Evidence. *Journal of Finance*, 39(3), 857-877.

- BRUSOV, P., FILATOVA, T., OREHOVA, N., BRUSOV, P. and BRUSOVA, N. (2011). Influence of Debt Financing on the Effectiveness of the Investment Project within the Modigliani–Miller Theory. *Research Journal of Economics, Business and ICT*. 2, 11-15.

- CAI, K., FAIRCHILD, R. and GUNEY, Y. (2008). Debt maturity structure of Chinese companies. Pacific-Basin Finance Journal, 16(3), 268-297.

- CHANG, K. P. (2015). *The Ownership of the Firm, Corporate Finance, and Derivatives*. Singapore: Springer, First Edition, ISBN: 978-981-287-353-8.

- CHEN, J. (2004). Determinants of Capital Structure of Chinese-listed Companies. *Journal of Business Research*, 57(12), 1341-1351.

- DEMIRGUC-KUNT, A. and MAKSIMOVIC, V. (1994). Capital Structures in Developing Countries Evidence from Ten Countries. Policy Research Working Paper, No: 1320, The World Bank.

- DEMIRHAN, D. (2009). Analysis of Firm Specific Factors Affecting Capital Structure: Application on ISE Service Firms. *Ege Academic Review, 9*(2), 677-697.

- DROBETZ, W. and FIX, R. (2003). What are the Determinants of the Capital Structure? Some Evidence for Switzerland. University of Basel. WWZ/Department of Finance, Working Paper, No: 4(03), 51-75.

- FASIOLO, M., WOOD, S. N., ZAFFRAN, M., NEDELLEC, R. and GOUDE, Y. (2020). Fast Calibrated Additive Quantile Regression. *Journal of the American Statistical Association.* 1-11.

- FRANK, M.Z. and GOYAL, V. K. (2009). Capital Structure Decisions: Which Factors are Reliably Important? *Financial Management*, 38(1), 1-37.

- FRIEND, I. and LANG, L.H.P. (1988). An Empirical Test of The Impact of Managerial Self-Interest on Corporate Capital Structure. *The Journal of Finance*, 43(2), 271-281.

- GHOSH, S. and CHATTERJEE, G. (2018). Capital Structure, Ownership and Crisis: How Different are Banks? *Journal of Financial Regulation and Compliance*, 26(2), 300-330.

- GITMAN, L.J. and ZUTTER, C. J. (2012). *Principles of Managerial Finance*. Boston: Pearson, 13th Edition, ISBN: 978-0-13-611946-3

- HALL G., HUTCHINSON P. and MICHAELAS N. (2004). Determinants of The Capital Structure of European SMEs. *Journal of Business and Accounting*, 31(5), 711-728.

- HARUN, T.W.R., KAMIL, N.K.M., HARON, R. and RAMLY, Z. (2020). Determinants of Banks' Capital Structure: A Review of Theoretical and Selected Empirical Research. *International Journal of Business and Social Science*, 11(12), 131-141.

- HEYMAN, D., DELOOF, M. and OOGHE, H. (2008). The Financial Structure of Private Held Belgian Firms. *Small business economics*, 30(3), 301-313.

- HOVAKIMIAN, A., HOVAKIMIAN, G. and TEHRANIAN H. (2004). Determinants of Target Capital Structure: The Case of Dual Debt and Equity Issues. *Journal of Financial Economics*, 71(3), 517-540.

- HSIAO, C. LI, Q. and RACINE, J. S. (2007). A Consistent Model Specification Test with Mixed Categorical and Continuous Data. *Journal of Econometrics*. 140(2), 802–26.

- HUANG, G. and SONG, F. M. (2006). The Determinants of Capital Structure: Evidence from China. *China Economic Review*, 17(1), 14-36.

- JAROS, J. and BARTOSOVA, V. (2015). To the Capital Structure Choice: Miller and Modigliani Model. *Procedia Economics and Finance*. *26*, 351-358.

- JORDAN, J., LOWE, J. and TAYLOR, P. (1998). Strategy and Financial Policy in UK Small Firms. *Journal of Business Finance and Accounting*, 25(1), 1-27.

- KESTER, C.W. (1986). Capital and Ownership Structure: A Comparison of United States and Japanese Manufacturing Corporations. *Financial Management*, 15(1), 5-16.

- KIM, C., MAUER, D.C. and SHERMAN, A.E. (1998). The Determinants of Corporate Liquidity: Theory and Evidence. *Journal of Financial and Quantitative Analysis*, 33(3), 335-359.

- KOCHHAR, R. and HITT, M. (1998). Linking Corporate Strategy to Capital Structure: Diversification Strategy, Type and Source of Financing. *Strategic Management Journal.* 19(6), 601–610.

- KOENKER, R. and BASSETT Jr, G. (1978). Regression Quantiles. *Econometrica: Journal of the Econometric Society*. 33-50.

- KRAUS, A. and LITZENBERGER, R.H. (1973). A State Preference Model of Optimal Financial Leverage. *Journal of Finance.* 38(4), 911–922.

- KRSTEVSKA, A., NENOVSKI, T. and POGACNIK KOSTOVSKA, K. (2017). Testing the Modigliani and Miller Theory in Practice: Evidence from the Macedonian Banking System. *Eastern European Economics*. 55(3), 277-289.

- MAC and BHAIRD, C. (2010). The Modigliani–Miller Proposition After Fifty Years and Its Relation to Entrepreneurial Finance. *Strategic Change.* 19(1-2), 9-28.

- MACKIE-MASON, J.K. (1990). Do Taxes Affect Corporate Financing Decisions? *The Journal of Finance*, 45(5), 1471-1493.

- MITNICK, B.M. (1974). *The theory of agency: The Concept of Fiduciary Rationality and Some Consequences*. Doctorate, Department of Political Science, University of Pennsylvania.

- MODIGLIANI, F. and MILLER, M.H. (1958). The Cost of Capital, Corporation Finance and the Theory of Investment. *The American Economic Review*, 48(3), 261-297.

- MYERS, S.C. (1984). The Capital Structure Puzzle. *Journal of Finance.* 39(3), 575–592.

- MYERS, S.C. and MAJLUF, N.S. (1984). Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have. *Journal of Financial Economics.* 13(2), 187–222.

- ONYINYECHI, U.G. (2019). Capital Structure and the Value of Quoted Firms in Nigeria: A Test of Miller and Modigliani Irrelevant Hypothesis. *Australian Finance & Banking Review*. 3(2), 1-15.

- RAJAN, R.G. and ZINGALES, L. (1995), What Do We Know about Capital Structure? Some Evidence from International Data, *The Journal of Finance*, 50(5), 1421-1460.

- ROSS, S.A. (1973). The Economic Theory of Agency: The Principal's Problem. *American Economic Review.* 62(2), 134-139.

- ROSS, S.A. (1977). The Determination of Financial Structure: The Incentive Signaling Approach. *Bell Journal of Economics.* 8(1), 23–40.

- SAKTI, M. R.P., TAREQ, M. A., SAITI, B. and AKHTAR, T. (2017). Capital Structure of Islamic Banks: A Critical Review of Theoretical and Empirical Research. *Qualitative Research in Financial Markets*, 9(3), 292-308.

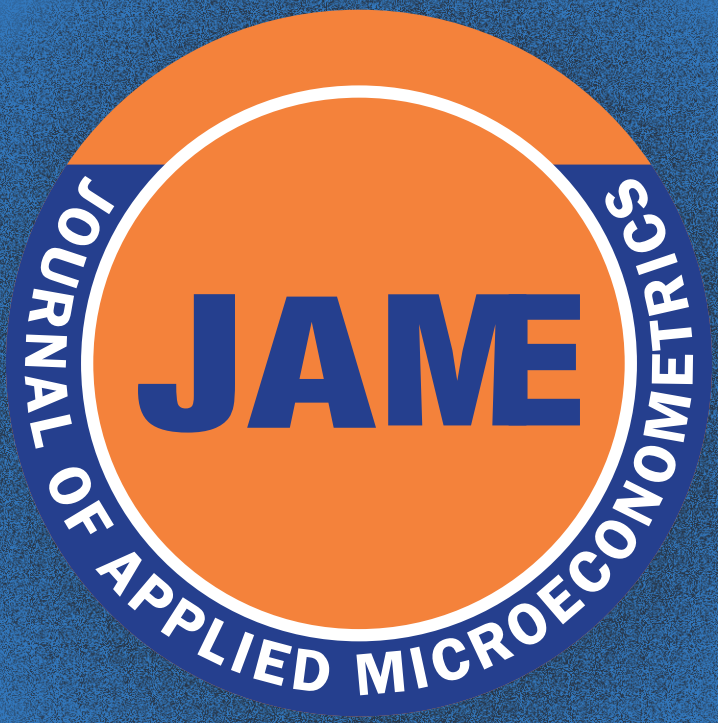
- SIBARANI, J.L., GINARTI, C., TAMBUN, R. and SURIANTI, M. (2020). Analysis of the Influence of Capital Structure and Profit Ability on Corporate Value: Empirical Study with Modigliani and Miller Theory of Investment Review. *European Journal of Business and Management*. 12(3), 164-170.

- TITMAN, S. and WESSELS, R. (1988). The Determinants of Capital Structure Choice. *The Journal of Finance*, 43(1), 1-19.

- VAN DER WIJST, N. and THURIK Y. (1993). Determinants of Small Firm Debt Ratios: An Analysis of Retail Panel Data. *Small Business Economics*, 5(1), 55-65.

- WEI, Y. (2014). The Determinants of Capital Structure: Evidence from Dutch Listed Firms, Master's Degree, University of Twente, the Netherlands.

- WIWATTANAKANTANG, Y. (1999). An Empirical Study on The Determinants of The Capital Structure of Thai Firms. *Pacific- Basin Finance Journal*, 7(3-4), 371-403.



HOLISTENCE
publications