JOURNAL OF APPLIED MICROECONOMETRICS

HOLISTENCE publications

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

International Peer-Reviewed and Open Access Electronic Journal

JAME
JOURNAL OF APPLIED MICROECONOMETRICS

Volume: 1
Issue: 2
2021

E-ISSN: 2791-7401

DOI: 10.53753/jame

International Peer-Reviewed and Open Access Electronic Journal

Volume: 1

Issue: 2

Year: 2021

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

E-mail: jame.editors@gmail.com

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

HOLISTENCE

publications

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 topics subject 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.editors@gmail.com
GSM 1: +1 (734) 747-4665 / WhatsApp
GSM 2: +90 530 638-7017 / WhatsApp

EDITORS

CONTENTS

"This page is left blank for typesetting"

JAME
E-ISSN:2791-7401
Volume : 1, Issue 2, 2021
URL: https://journals.gen.tr/jame

Received: 10.06.2021
Acccepted: 11.12.2021

RESEARCH ARTICLE

Returns to education: Empirical evidence from Kyrgyzstan

Burulcha Sulaimanova[1]

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

Abstract

The aim of this study is to identify the returns to education in Kyrgyzstan, with special reference to employment type and gender differences. The empirical analysis of this study is based on Life in Kyrgyzstan (LiK) survey data collected in 2016. The sample for analysis is constructed with employees and self-employed persons aged 18-65, who indicated their monthly income from employment. According to the empirical outputs, there is a wage premium for higher education such that the marginal return to education for women is higher than men.

Keywords: Education, Returns, Mincer Wage Model, Kyrgyzstan, Central Asia

JEL Codes: I26, I20, N35

Citation: SULAIMANOVA, B., (2021). Returns to education: Empirical evidence from Kyrgyzstan. *Journal of Applied Microeconometrics (JAME)*. 1(2), 73-84, DOI: 10.53753/jame.1.2.01

Corresponding Author:
Burulcha Sulaimanova
E-mail: burulcha.sulaymanova@manas.edu.kg

1. INTRODUCTION

The high level of education of the population, being the key to the development of human capital, makes a positive contribution to economic development of the country in the long run. However, the discrepancy between the acquired vocational and qualification skills and the needs of the labor market can have a negative impact on employment indicators and, as a consequence, on economic growth (Ryazantseva 2012; Allen and Van Der Velden 2001).

Kyrgyzstan has a fairly high level of access to education, both school and higher. In particular, over the past 25 years, the number of higher education institutions has rapidly increased. However, the increase in the number of higher education institutions does not meet the demand for skilled labor in the labor market. This problem causes various public discussions in order to reform the education system. This study aims to identify returns on education, whether obtained education pays off at the labor market of Kyrgyzstan. The study uses the nationally representative household survey "Life in Kyrgyzstan" for 2016, which is available from the International Data Service Center of the Institute for Study of Labour (IDSC IZA). Empirical analysis of the impact of education on income are based on the non-linear wage model of Mincer.

This study is structured as follows. Section 2 describes the state of the labor market and the education system in Kyrgyzstan. Section 3 presents the methodology for assessing the mismatch factors and the statistical data obtained. Sections 4 and 5 present the analysis results and conclusions respectively.

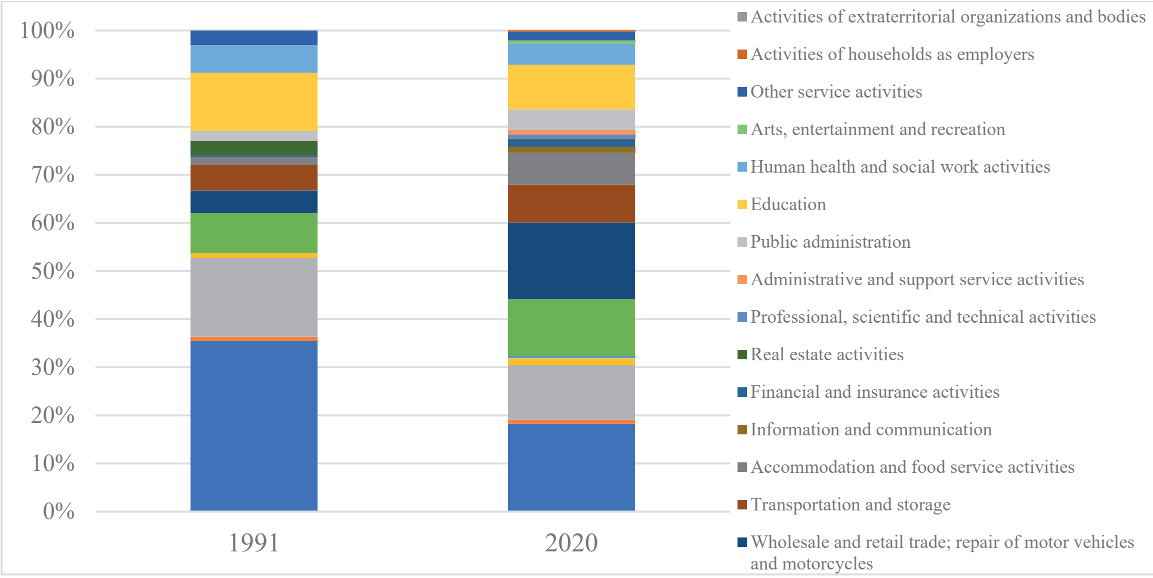
1.1. Education and Labor Market Trends in Kyrgyzstan

As a heritage of the Soviet Union, the employed population of the Kyrgyzstan has a high level of education. According to the National Statistics Committee of Kyrgyz Republic (NSCKR 2018) almost every fifth of the employed population has a higher or incomplete higher education, every tenth has a secondary vocational education. For 2017, the total number of employed women, the proportion of women who had higher professional education was 27% whereas; the share of men with higher professional education was 17%. Furthermore, 16% of women had secondary vocational and this ratio was 7% for men.

Over the years of independence, the number of higher educational institutions has rapidly increased from 9 in 1990 to 50 in 2016. Most of the higher educational institutions are located in Bishkek (about 64 %), which is due to the presence of a developed network of state educational institutions of higher professional education (NSCKR 2018).

According to the structure of student's specialties they receive, the most popular professions are economist, translator from a foreign language, doctor, lawyer, engineer, IT programmer, builder and manager. Thus, according to NSCKR the distribution of students of higher educational institutions by groups of specialties for 2017 is as follows: more than half of the students (53%) study in the humanities (i.e., the professions of economist, manager, lawyer, etc.), natural sciences -3.4%, education - 14.7%, healthcare - 6.7%, technical sciences - 19.6%, agricultural sciences - 1%, service - 0.3%, interdisciplinary sciences - 1.7%.

Figure 1. Employed Population by Type of Economic Activity



Source: NSCKR, (www.stat.kg/ru/statistics/zanyatost/ - accessed 15 July 2021)

However, obtaining a specialty is not always guaranteed by employment in the labor market. Thus, according to the NSCKR, in 2018 the total unemployment rate in the country was 6.2%, whereas about 40 % of those unemployed have tertiary or technical education. This fact indicates that it is rather difficult for graduates to get a job after receiving a diploma. This particularly shows that with high growth rates of tertiary education leads to an oversupply of highly skilled labor in Kyrgyzstan. This could be a result of the weaknesses of the education system and one of the labor market features of transition economies. Hence, in post-Soviet countries, the rapid job reallocation and slow creation of jobs in high productivity sectors, result in difficulty for individuals to join the labor market and put their skill to use (Kupets 2016). Consequently, in Kyrgyzstan in the years of independence, the structure of the employed population underwent significant changes, where the share of people employed in agriculture decreased, while in trade, services and construction increased. This could be seen in the Figure 1. In 1991, the economy of Kyrgyzstan prevailed with agriculture, manufacturing and education sectors, while in 2020 the services sectors, such as trade, construction, accommodation and food service activities prevailed. Also, another factor that led to such changes is labor migration, when migrants, leaving work in agriculture, move to trade and services or construction at the place of their new residence. (NSCKR 2016).

Over the past 5 years, the share of expenditures allocated to education in the expenditures of the state budget of Kyrgyzstan was quite significant (average from 21% to 24%). The bulk of education spending is on secondary education (57% on average) and tertiary vocational education (13%) (NSCKR 2018). Considering, the high cost of education as for individuals, and public investment in education in Kyrgyzstan, the analysis of return for education is very important.

1.2. Literature Review

The bulk of research that studies returns to education is mainly based on the theory of human capital. The theory stresses the importance of education for improving productivity. The main argument of human capital theory is that better educated people are generally more skilled and are expected to be more productive than people with lower levels of education, thus skilled workers will earn more (Joseph 2020; Wang et al. 2019). That is why the quantity of years of study is used as one of the key factors determining the level of earnings of employees (Rizk 2019). Even though the empirical literature has been dominantly presented with the interpretation within human capital theory, the alternative literature asserts that education may have effect over wages not because of productivity of worker, but for other reasons. For instance, education may act as an instrument signaling the ability or productivity of worker (Gunderson and Oreopolous 2020; Harmon et al. 2003). In this case one may see differences between education as productivity measurement or signaling instrument by comparing the returns to education of employees and self-employed workers. While education does not have any values as signal for self-employed individuals, since they know their own productivity and do not need to get more education; it will have significant effect on wages of employees (Harmon et al. 2003).

Analysis of the impact and establishment of return to education, shows that there are significant differences across wage distributions. Hence wage premium from higher education for individuals from low decile of income distribution is considerably less than for those from higher-income backgrounds (Bartik and Hershbein 2018; Harmon et al. 2003).

It should be noted that the marginal return on education among women is invariably higher than among men, which also makes it important to study the return on education separately by gender.

To our best knowledge, one of the first empirical studies of the impact of education on the wages in Kyrgyzstan is the study of Karymshakov and Sulaimanova (2019). They have investigated the impact of the education-job mismatch on the wages of youth in Kyrgyzstan, based on the International Labor Organization school-to-work transition survey for Kyrgyzstan. According to the empirical results, overeducated young men receives lower wage compared to their counterparts that are well matched with education and job position. The other study of Çağlayan Akay et al. (2019) used Mincer's earnings model to assess the impact of education on the wages of women working in the developed shopping and business center of Bishkek city (capital of Kyrgyzstan). The study sample consists of 675 employed women. The results of the study show that the return on education for female employees in the private sector is higher than in the public sector. Moreover, the total number of years of study has a strong causal effect on wages. The authors recommended paying sufficient attention to the education of women in Kyrgyzstan.

The main contribution of our study for the literature of analysis of returns to education will be in three ways. First of all, we will conduct a more detailed analysis of the impact of education on the income of the workers, with a nationally representative household survey for Kyrgyzstan. The use of representative data for Kyrgyzstan makes it possible to generalize the empirical results obtained for the entire population of the country. Also, a large number of observation units makes it possible to analyze the influence of education on the earnings by such subgroups as employment type and gender differences. Secondly, in contrast to the previous empirical works on Kyrgyzstan, our study aims to analyze returns by the level of education, rather than the total number of years spent on education. This may show a wage premium for fulfilling a particular year of education, like the last year of high school or high school (Harmon et al. 2003; Churchill and Mishra 2018), and make recommendations based on variations in the results. Thirdly, we investigate how particular field of education affects the earnings of workers.

3. METHODOLOGY

3.1. Data and descriptive statistics

This study uses the "Life in Kyrgyzstan" survey for 2016, which is a research-based, open access, multi-topic longitudinal survey of households and individuals in Kyrgyzstan. The survey is conducted by the German Institute for Economic Research, DIW Berlin and Stockholm International Peace Research Institute, and available from the International Data Service Center of the Institute for Study of Labour (IDSC IZA). This survey provides data at the national level as well as for urban and rural areas of Kyrgyzstan and contains a wide range of data including information on household characteristics (composition, dwelling, children, health etc.), assets, shocks, social networks, income, and expenditure of households. In addition, the survey allocates a particular section on employment and education of individuals.

To investigate the effect of education on income, the sample for analysis included employees and self-employed individuals aged 18-65, who indicated their monthly income from employment. The income variable consists of the monthly wages of employees and the monthly income of individual workers in KGS (national currency of Kyrgyzstan). The total sample for estimation consists of 3074 observations.

The average income in sample is 10588 KGS, and men workers earn on average for 2000 KGS more than women. The average age of the workers is 39 years, and most of them are married. While nearly half of workers are with secondary education, almost every fourth female worker holds tertiary education, and for male this share is 14.7%.

Men are more likely to have completed engineering, economics and law, while women are more likely to hold education, medicine and economics diploma.

Table 1. Descriptive Statistics

	TOTAL SAMPLE		MEN		WOMEN	
	N	Mean	N	Mean	N	Mean
Earnings	3,074	10588.39	1,928	11235.94	1,146	9498.967
Age	3,074	38.61451	1,928	38.2666	1,146	39.19983
Marital status (1=married)	3,074	0.715029	1,928	0.771266	1,146	0.620419
Education level:						
▪ Secondary	3,074	0.437215	1,928	0.458506	1,146	0.401396
▪ Technical	3,074	0.129148	1,928	0.121888	1,146	0.141361
▪ Tertiary	3,074	0.182173	1,928	0.147822	1,146	0.239965
Education field:						
▪ Natural science	3,074	0.009434	1,928	0.00778	1,146	0.012216
▪ Education	3,074	0.054001	1,928	0.025934	1,146	0.101222
▪ Medicine	3,074	0.031555	1,928	0.013486	1,146	0.061955
▪ Engineering	3,074	0.068966	1,928	0.088693	1,146	0.035777
▪ Computer	3,074	0.00553	1,928	0.007261	1,146	0.002618
▪ Agriculture	3,074	0.013663	1,928	0.01971	1,146	0.00349
▪ Economics	3,074	0.064737	1,928	0.042531	1,146	0.102094
▪ Law	3,074	0.018543	1,928	0.021784	1,146	0.013089
▪ International relations	3,074	0.004554	1,928	0.003631	1,146	0.006108
▪ Languages	3,074	0.009434	1,928	0.003631	1,146	0.019197
Residence (1=rural)	3,074	0.556279	1,928	0.60166	1,146	0.47993
Regions:						
▪ Issyk-Kul	3,074	0.091087	1,928	0.095954	1,146	0.082897
▪ Jalal-Abad	3,074	0.234548	1,928	0.231328	1,146	0.239965
▪ Naaryn	3,074	0.048146	1,928	0.056017	1,146	0.034904
▪ Batken	3,074	0.087508	1,928	0.091805	1,146	0.080279
▪ Osh	3,074	0.101496	1,928	0.111515	1,146	0.084642
▪ Talas	3,074	0.05823	1,928	0.063278	1,146	0.049738
▪ Chui	3,074	0.13175	1,928	0.133817	1,146	0.128272
▪ Bishkek city	3,074	0.182498	1,928	0.147822	1,146	0.240838
▪ Osh city	3,074	0.064737	1,928	0.068465	1,146	0.058464

Source*:* Authors calculations, LiK 2016

About 55% employees and employers are working in rural areas; this share is smaller for female sub-sample. According to regional distribution of workers, the concentration of women self-employed and employee is seen in the Bishkek city, the capital city of Kyrgyzstan.

3.2. Empirical Model

Analysis of the impact of education on income are based on the non-linear wage model of Mincer, which has the following form:

$$lnY_i = \alpha_i + \beta_i X_i + \gamma_i EDU_i + \varepsilon_i \qquad (1)$$

where the dependent variable is the logarithmic value of the employee's and employer's income, estimated by set of independent variables such as age, gender, marital status, and place of residence, occupation characteristics () (see detailed description of variables in Annex Table 1). The education variable - , represents the level of education and field of education. Education level variable denotes a certain level of education, such as: secondary general education, vocational, higher education and a group of young people with incomplete, primary education or without education (the last category of education in the model is the base for comparison). Using dummies for educational attainment has an advantage over using the total number of years spent in education, where the marginal return varies with educational attainment, and when the aim of the study is to examine the different influences of educational levels (Purnastuti et al. 2013). While the variable showing the field of education denotes: Natural science, Education, Medicine, Engineering, Computer, Agriculture, Economics, Law, International relations, Languages and other fields of study (the last category in the model is the reference group). Differentiating returns to education by field of study will give some insights on which educational programs pays-off most.

4. EMPIRICAL RESULTS

4.1. Returns to the Level of Education

The results of the returns to education presented in Table 2. These models are estimated using the least squares method, and corrected for heteroscedasticity of random residuals. The main variable of interest the educational attainment has a positive impact over earnings of workers. Hence individuals with tertiary education earn more than those who has an incomplete or primary education, or has no education at all. This result supports the college premium hypothesis and indicates that returns to tertiary education is higher than the lower levels of education (Wang et al. 2019; Mitra 2019). Whilst analyzing returns of education across gender, one may see that the tertiary education has a highly significantly impact on women earnings.

Further, when analyzing occupational differences in the return on education among workers, it can also be noted that employees with higher education demonstrate significantly higher returns, while education level does not have any impacts over earnings of self-employed. These results accept Signaling theory, and indicates that in Kyrgyzstan especially in cases where employers cannot easily observe the abilities or performance of workers, they rely on educational attainment as a signaling instrument in hiring decisions.

The other control variables have expected statistically significant signs. With increase of age, the earnings also increase with diminishing returns. The marital status has a statistically significant positive effect over earnings of men, supporting the specialization hypothesis stating marriage affects men's wages positively (Purnastuti et al. 2013). Since male workers who are married, can devote more time and effort to activities in the labor market and, as a result, this increases their earnings (Purnastuti et al. 2013). Kyrgyz workers on average earn less than other nationalities; this is particularly true for Kyrgyz self-employed men.

Estimates of the dummy variable for rural areas of residence show that, on average, rural residents earn significantly less than urban residents do. One may also see that there is a regional imbalance in the level of income. Thus, in regions located in north part of the country, this is Talas, Naryn and Issyk-Kul oblast, earnings are much lower than in Bishkek city or Chuy oblast. While senior official and manager earn much higher income than others; workers from agriculture, education and health sector earn less than in other economic sectors.

Table 2. Returns to the Level Education

	TOTAL SAMPLE			EMPLOYEE			SELF-EMPLOYED		
	Total	*Women*	*Men*	*Total*	*Women*	*Men*	*Total*	*Women*	*Men*
Education level:									
▪ *Secondary*	-0.0651**	0.00563	-0.0798**	-0.0341	0.00355	-0.0240	-0.145***	-0.0688	-0.156***
	(0.0277)	(0.0441)	(0.0349)	(0.0311)	(0.0453)	(0.0420)	(0.0506)	(0.117)	(0.0562)
▪ *Technical*	0.00205	0.0898	-0.0406	0.0369	0.0819	0.0161	-0.112	-0.0361	-0.136
	(0.0410)	(0.0681)	(0.0500)	(0.0473)	(0.0734)	(0.0608)	(0.0776)	(0.183)	(0.0847)
▪ *Tertiary*	0.165***	0.330***	0.0496	0.236***	0.327***	0.154***	-0.0854	0.232	-0.151
	(0.0331)	(0.0499)	(0.0450)	(0.0341)	(0.0515)	(0.0455)	(0.0818)	(0.173)	(0.0947)
Log of age	0.589***	0.818***	0.476***	0.524***	0.593***	0.497***	0.565***	1.316***	0.403*
	(0.107)	(0.145)	(0.150)	(0.120)	(0.150)	(0.182)	(0.196)	(0.434)	(0.230)
Log of squared age	-0.00230***	-0.00292***	-0.00200***	-0.00220***	-0.00237***	-0.00215***	-0.00200***	-0.00397***	-0.00166**
	(0.000383)	(0.000523)	(0.000519)	(0.000426)	(0.000544)	(0.000642)	(0.000678)	(0.00153)	(0.000765)
Marital status (1=married)	0.0903***	-0.0211	0.137***	0.0822***	-0.0265	0.140***	0.0722	0.0133	0.102
	(0.0270)	(0.0364)	(0.0418)	(0.0297)	(0.0386)	(0.0480)	(0.0550)	(0.0916)	(0.0731)
Ethnicity (1=Kyrgyz)	-0.0888***	-0.00814	-0.125***	-0.0208	-0.00210	-0.0232	-0.194***	-0.0423	-0.236***
	(0.0240)	(0.0394)	(0.0298)	(0.0276)	(0.0414)	(0.0356)	(0.0455)	(0.106)	(0.0501)
Residence (1=rural)	-0.227***	-0.190***	-0.268***	-0.166***	-0.198***	-0.159***	-0.427***	-0.286**	-0.470***
	(0.0235)	(0.0367)	(0.0297)	(0.0254)	(0.0375)	(0.0337)	(0.0505)	(0.118)	(0.0535)
Regions:									
▪ North	-0.319***	-0.281***	-0.363***	-0.205***	-0.206***	-0.217***	-0.480***	-0.573***	-0.453***
	(0.0354)	(0.0533)	(0.0458)	(0.0396)	(0.0557)	(0.0556)	(0.0679)	(0.144)	(0.0765)
▪ South	0.0219	0.0166	-0.000570	-0.0195	-0.0555	-0.0294	0.0100	0.115	-0.00826
	(0.0272)	(0.0431)	(0.0348)	(0.0300)	(0.0450)	(0.0396)	(0.0581)	(0.112)	(0.0671)
Economic sector:									
▪ Agriculture and fishing	-0.153***	-0.273***	-0.136***	-0.128	-0.311**	-0.0931	-0.116**	-0.311**	-0.0728
	(0.0396)	(0.0911)	(0.0449)	(0.119)	(0.151)	(0.145)	(0.0490)	(0.123)	(0.0537)
▪ Education	-0.230***	-0.155***	-0.159***	-0.217***	-0.0726*	-0.180***	-0.553	-0.469	-
	(0.0271)	(0.0371)	(0.0480)	(0.0286)	(0.0397)	(0.0492)	(0.520)	(0.500)	
▪ Health and social work	-0.149***	-0.0650	-0.0663	-0.119***	0.0165	-0.0867	-0.648**	-0.926***	-0.268***
	(0.0355)	(0.0472)	(0.0607)	(0.0365)	(0.0494)	(0.0617)	(0.328)	(0.341)	(0.0633)
Senior official and manager	0.419***	0.456***	0.400***	0.239**	0.389***	0.184	0.560***	0.501**	0.563***
	(0.0803)	(0.151)	(0.0919)	(0.103)	(0.135)	(0.127)	(0.105)	(0.250)	(0.117)
Constant	7.472***	6.525***	7.943***	7.567***	7.224***	7.703***	7.908***	5.108***	8.510***
	(0.337)	(0.458)	(0.467)	(0.374)	(0.471)	(0.567)	(0.623)	(1.422)	(0.723)
N	3074	1146	1928	1928	916	1012	1146	230	916
Log likelihood	-2764.9	-927.9	-1783.1	-1480.5	-667.4	-772.4	-1175.6	-223.2	-938.9

Note: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$.
Source: Authors calculations, LiK 2016

4.2. Returns to the Field of Education

This research studies the influence of education specialties on earnings, to determine which field of study pays off much in Kyrgyzstan. Table 3 reports the results of this analysis.

The greatest return on specialties relates to law and jurisprudence. On average, lawyers earn 29.2% more than other specialties. Gender differences are noticeable, for instance, female lawyers earn on average 42.5% more than other women, while male lawyers earn 20.8% more than other men. Both employees and self-employers with law are more likely to earn more than other specialties.

A significant high share of returns from the field-of-study was demonstrated in such areas as economics, management and banking. Consequently, on average, graduates of these areas earn more by 13.5%. It is noteworthy that the return on these specializations is significantly high among economist employees, while self-employed economists have no such causal relationships.

Graduates from education and pedagogy also earn comparatively more. However, statistically significant results were found only for employees. In other words, only wage employed graduates from education field earn on average 26% more than graduates in other specialties, and in the subsample of women these returns are much higher.

While in above given education fields the returns demonstrated evenly among gender, one may see that in some specialties the earnings with respect to education differ across gender. Thus, the highest female returns on specialties is demonstrated in the field of computer science for employee women. The same results found for women with medical education. Great returns have also been demonstrated among International relations graduates, which is significant only for men. That is, on average, a wage employed man with a specialization in the international relations earns 29.6% more than other men, while in self-employment this impact increase to 78.2%. Women graduate from languages field are more likely to earn than their counterparts.

This analysis of returns to specialties shows the relationship between the popularity of different specializations in Kyrgyzstan. The statistically significant differences in income by profession explains the choice of certain professions in Kyrgyzstan. Therefore, today the profession associated with computer skills is very popular, which (according to the analysis results) has a fairly high return. Also popular are the specialties of lawyer and economist, which have statically significant levels of return. International relation, medicine and education professions have gender differentiated returns. It can be concluded that the demand for the chosen profession in the labor market is positively correlated with the possible level of wages.

Table 3. Returns to the Field of Education

	TOTAL SAMPLE			EMPLOYEE			SELF-EMPLOYED		
	Total	*Women*	*Men*	*Total*	*Women*	*Men*	*Total*	*Women*	*Men*
Field of study:									
▪ Natural science	0.225**	0.128	0.278**	0.146	0.000232	0.239*	0.345***	0.485***	0.229
	(0.0946)	(0.140)	(0.115)	(0.108)	(0.138)	(0.143)	(0.119)	(0.155)	(0.171)
▪ Education	0.231***	0.319***	0.109	0.268***	0.327***	0.185**	0.0541	0.258*	-0.00950
	(0.0407)	(0.0454)	(0.0777)	(0.0416)	(0.0481)	(0.0796)	(0.105)	(0.135)	(0.154)
▪ Medicine	0.117*	0.172***	0.0617	0.139**	0.166**	0.0681	-0.0538	-0.0287	-0.00170
	(0.0645)	(0.0647)	(0.155)	(0.0641)	(0.0651)	(0.192)	(0.182)	(0.276)	(0.245)
▪ Engineering	0.0219	-0.0786	-0.00887	0.0636	0.00402	0.0183	-0.0852	-0.844**	-0.0621
	(0.0419)	(0.104)	(0.0458)	(0.0471)	(0.105)	(0.0530)	(0.0872)	(0.401)	(0.0840)
▪ Computer	0.246	1.168*	-0.0213	0.454*	1.730**	0.136	-0.467**	-0.151	-0.576**
	(0.209)	(0.655)	(0.170)	(0.245)	(0.702)	(0.190)	(0.226)	(0.139)	(0.257)
▪ Agriculture	0.0402	-0.190	0.00376	0.122	-0.665	0.176**	-0.0175	1.924***	-0.103
	(0.114)	(0.665)	(0.103)	(0.149)	(0.667)	(0.0772)	(0.154)	(0.174)	(0.154)
▪ Economics	0.135***	0.239***	0.0899	0.207***	0.257***	0.234***	-0.0828	0.0630	-0.134
	(0.0494)	(0.0619)	(0.0808)	(0.0471)	(0.0628)	(0.0690)	(0.132)	(0.228)	(0.168)
▪ Law	0.292***	0.425***	0.208***	0.269***	0.379**	0.170**	0.305**	0.432**	0.290*
	(0.0620)	(0.147)	(0.0671)	(0.0711)	(0.181)	(0.0729)	(0.125)	(0.202)	(0.156)
▪ International relations	0.308**	0.188	0.439***	0.253***	0.188	0.296**	0.854***	-	0.782***
	(0.121)	(0.129)	(0.160)	(0.0950)	(0.125)	(0.135)	(0.104)		(0.115)
▪ Languages	0.200**	0.355***	-0.140	0.203**	0.342***	-0.217**	0.227***	0.322**	0.177
	(0.0888)	(0.106)	(0.136)	(0.0981)	(0.119)	(0.0872)	(0.0725)	(0.162)	(0.129)
Log of age	0.583***	0.823***	0.484***	0.529***	0.624***	0.499***	0.582***	1.295***	0.446*
	(0.107)	(0.148)	(0.149)	(0.120)	(0.151)	(0.182)	(0.196)	(0.450)	(0.230)
Log of squared age	-0.0023***	-0.0029***	-0.00205***	-0.00223***	-0.00247***	-0.00212***	-0.00218***	-0.00411**	-0.00189**
	(0.000381)	(0.000530)	(0.000514)	(0.000431)	(0.000548)	(0.000648)	(0.000677)	(0.00160)	(0.000762)
Marital status (1=married)	0.0948***	-0.0212	0.139***	0.0820***	-0.0342	0.137***	0.0793	0.0547	0.0985
	(0.0271)	(0.0356)	(0.0419)	(0.0300)	(0.0370)	(0.0487)	(0.0554)	(0.0998)	(0.0741)
Ethnicity (1=Kyrgyz)	-0.0887***	-0.0136	-0.132***	-0.0156	-0.000258	-0.0268	-0.188***	-0.0746	-0.235***
	(0.0239)	(0.0392)	(0.0298)	(0.0275)	(0.0406)	(0.0354)	(0.0460)	(0.104)	(0.0510)
Residence (1=rural)	-0.232***	-0.193***	-0.271***	-0.170***	-0.189***	-0.166***	-0.433***	-0.317**	-0.465***
	(0.0236)	(0.0362)	(0.0299)	(0.0253)	(0.0358)	(0.0340)	(0.0510)	(0.123)	(0.0539)
Regions:									
▪ North	-0.336***	-0.301***	-0.378***	-0.220***	-0.228***	-0.226***	-0.497***	-0.541***	-0.476***
	(0.0358)	(0.0535)	(0.0463)	(0.0402)	(0.0561)	(0.0567)	(0.0691)	(0.144)	(0.0784)
▪ South	0.00903	-0.0189	-0.00722	-0.0337	-0.0870**	-0.0399	0.00487	0.103	-0.0137
	(0.0278)	(0.0424)	(0.0358)	(0.0306)	(0.0442)	(0.0415)	(0.0591)	(0.111)	(0.0688)
Economic sector:									
▪ Agriculture and fishing	-0.154***	-0.265***	-0.136***	-0.126	-0.293**	-0.103	-0.111**	-0.336***	-0.0666
	(0.0395)	(0.0916)	(0.0448)	(0.117)	(0.149)	(0.143)	(0.0493)	(0.125)	(0.0540)
▪ Education	-0.239***	-0.141***	-0.176***	-0.220***	-0.0683*	-0.196***	-0.471	-1.005***	-
	(0.0283)	(0.0370)	(0.0498)	(0.0301)	(0.0390)	(0.0526)	(0.499)	(0.148)	
▪ Health and social work	-0.154***	-0.0709	-0.0596	-0.127***	0.00572	-0.0770	-0.610*	-0.964***	-0.164***
	(0.0407)	(0.0501)	(0.0684)	(0.0419)	(0.0523)	(0.0729)	(0.365)	(0.313)	(0.0527)
Senior official and manager	0.450***	0.461***	0.419***	0.287***	0.399***	0.231*	0.569***	0.521*	0.562***
	(0.0815)	(0.149)	(0.0925)	(0.107)	(0.134)	(0.129)	(0.109)	(0.270)	(0.123)
Constant	7.469***	6.544***	7.885***	7.546***	7.152***	7.697***	7.773***	5.191***	8.284***
	(0.335)	(0.461)	(0.464)	(0.374)	(0.470)	(0.568)	(0.618)	(1.460)	(0.719)
N	3074	1146	1928	1928	916	1012	1146	230	916
Log likelihood	-2766.6	-922.1	-1781.6	-1483.8	-655.2	-769.5	-1174.6	-217.1	-939.0

Note: Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Source: Authors calculations, LiK 2016

5. CONCLUSION

Decent work and employment in Kyrgyzstan are one of the main socio-economic problems of the country. Against the background of the high level of unemployment among tertiary educated, the study of the impact of education on employment is becoming one of the most pressing topics. In recent years, there has been a tendency in the development of state policy in the field of reforming the higher education system in Kyrgyzstan. Policy measures are focused on improving the effectiveness of higher education, increasing youth participation in the vocational education process, updating curricula and integrating with the international education system. Nevertheless, the problems of unemployment, the transition from education to employment and the discrepancy between the skills acquired in educational institutions and the needs of the labor market remain unresolved. Taking this into account, this study is of great importance for studying the returns of education in Kyrgyzstan. In particular, this study has empirical results suggesting that the level of education affects the income of people significantly.

Based on the empirical analysis of the household survey data "Life in Kyrgyzstan" for 2016, the following conclusions can be drawn. First, the investment in education, in particular in higher education, pays off. Second, the marginal return on education for women is higher than for men. It can also be said that, given the higher return on education for women, efforts should be made to improve the educational level of women, since the participation of more educated women in the labor market is rewarded with a higher income. Third, the statistically significant and high rate of return on education for employees, compared to the self-employers, indicates the importance of higher education for entering labor market, or employee to be hired. It is also worth noting that this effect is significantly higher for wage employed women. It is appropriate here to recommend investing in education of women, by expanding access to education. Fourth, based on the results of the study of returns to specialization, it can be concluded that the popularity of certain specialties in Kyrgyzstan is closely related to the expected income from these specialties. Thus, according to the results of empirical analysis, the highest share of return on specialties falls on the most popular educational areas, such as computer science, economics and law.

Funding

The author declares that this study has no financial support.

REFERENCES

- ALLEN, J., & VAN DER VELDEN, R. (2001). Educational mismatches versus skill mismatches: effects on wages, job satisfaction, and on-the-job search. *Oxford economic papers*, *53*(3), 434-452.
- BARTIK, T., & HERSHBEIN, B. (2018). Degrees of poverty: The relationship between family income background and the returns to education.
- ÇAĞLAYAN AKAY E., OSKONBAEVA, Z., & SACAKLI-SACILDI, I. (2019). The causal effects of education on wages: evidence from Kyrgyzstan. *Economic Journal of Emerging Markets*, *11*(2), 183-194.
- CHURCHILL, S. A., & MISHRA, V. (2018). Returns to education in China: a meta-analysis. *Applied Economics*, *50*(54), 5903-5919.
- GUNDERSON, M., & OREOPOLOUS, P. (2020). Returns to education in developed countries. In The economics of education (pp. 39-51). Academic Press.
- HARMON, C., OOSTERBEEK, H., & WALKER, I. (2003). The returns to education: Microeconomics. *Journal of economic surveys*, *17*(2), 115-156.
- JOSEPH, C. (2020). Education and Labour Market Earnings in Low Income Countries: Empirical Evidence for Tanzania. *Tanzania Journal for Population studies and Development*, *26*(2).
- KARYMSHAKOV, K., & SULAIMANOVA, B. (2019). The school-to-work transition, overeducation and wages of youth in Kyrgyzstan. *International Journal of Manpower*.
- KLASEN, S. & LAMANNA, F. (2009). The impact of gender inequality in education and employment on economic growth: new evidence for a panel of countries. Feminist economics, 15(3), 91-132.
- KUPETS, O. (2016). Education-job mismatch in Ukraine: Too many people with tertiary education or too many jobs for low-skilled?. *Journal of comparative economics*, *44*(1), 125-147.
- MITRA, A. (2019). Returns to education in India: Capturing the heterogeneity. Asia & the Pacific Policy Studies, 6(2), 151-169.
- NATIONAL STATISTICAL COMMITTEE OF THE KYRGYZ REPUBLIC (NSCKR). (2016). 20 years of independence of the Kyrgyz Republic. Figures and facts. Bishkek
- NATIONAL STATISTICAL COMMITTEE OF THE KYRGYZ REPUBLIC (NSCKR). (2018). Education and Science in the Kyrgyz Republic. Statistical Book. Bishkek, 2018
- PURNASTUTI, L., MILLER, P.W., & SALIM, R. (2013). Declining rates of return to education: evidence for Indonesia. *Bulletin of Indonesian Economic Studies*, *49*(2), 213-236.
- RIZK, R. (2019). Returns to education in MENA countries: a continuing story of under-achievement. *International Journal of Education Economics and Development*, *10*(4), 427-448.
- RYAZANTSEVA, I. (2012) Organization and planning of reproduction of competitive specialists // Interexpo GEO-Siberia. 2012. T. 2, No. 2. P. 181–187. - Рязанцева И. В. Организация и планирование воспроизводства конкурентоспособных специалистов // Интерэкспо ГЕО-Сибирь. 2012. Т. 2, № 2. С. 181–187.
- WANG, W., DONG, Y., LUO, R., BAI, Y., & ZHANG, L. (2019). Changes in returns to education for off-farm wage employment: Evidence from rural China. China Agricultural Economic Review.

APPENDIX

Table A1. Description of Variables

DEPENDENT VARIABLE	
Earnings	Monthly income of employees, own-account workers in Soms (KGS)
INDEPENDENT VARIABLE	
Age	Full age in years
Married	1 = if individual is married, 0 = otherwise
Education level	
▪ Secondary education	1 = if individual obtained secondary level of education, 0 = otherwise
▪ Technical education	1 = if individual obtained technical level of education, 0 = otherwise
▪ Tertiary education	1 = if individual obtained tertiary level of education, 0 = otherwise
Field-of-study	
▪ Natural science	1 = if individual graduated from Natural science, 0 = otherwise
▪ Education	1 = if individual graduated from Education, 0 = otherwise
▪ Medicine	1 = if individual graduated from Medicine, 0 = otherwise
▪ Engineering	1 = if individual graduated from Engineering, 0 = otherwise
▪ Computer	1 = if individual graduated from Computer, 0 = otherwise
▪ Agriculture	1 = if individual graduated from Agriculture, 0 = otherwise
▪ Economics	1 = if individual graduated from Economics, 0 = otherwise
▪ Law	1 = if individual graduated from Law, 0 = otherwise
▪ International relations	1 = if individual graduated from International relations, 0 = otherwise
▪ Languages	1 = if individual graduated from Languages, 0 = otherwise
Ethnicity	
▪ Kyrgyz	1 = if individual is Kyrgyz, 0 = otherwise
Residence	1 = if individual reside in rural area, 0 = urban area
Regions	
▪ North	1 = if individual reside in Issyk-Kul, Naryn orTalas oblast, 0 = otherwise
▪ South	1 = if individual reside in Jalal-Abad, Batken or Osh oblast, 0 = otherwise
▪ Central	1 = if individual reside in Bishkek city or Chuy oblast, 0 = otherwise
Sector	
▪ Agriculture and fishing	1 = if individual employed in Agriculture and fishing sector, 0 = otherwise
▪ Education	1 = if individual employed in Education sector, 0 = otherwise
▪ Health and social work	1 = if individual employed in Health and social work sector, 0 = otherwise
Position	
▪ Senior official and manager	1 = if individual works as Senior official and manager, 0 = otherwise

JAME
E-ISSN:2791-7401
Volume : 1, Issue 2, 2021
URL: https://journals.gen.tr/jame

Received: 30.11.2021
Acccepted: 13.12.2021

RESEARCH ARTICLE

Customer churn analysis in banking sector: Evidence from explainable machine learning models

Hasraddin Guliyev[1] **Ferda Yerdelen Tatoğlu**[2]

1 The Economic Research Center of Turkish World, Azerbaijan State Economic University, Azerbaijan.
Ph.D. Student at Department of Econometrics, Istanbul University, Istanbul, e-mail: hasradding@unec.edu.az

2 Istanbul University, Faculty of Economics, Department of Econometrics, Istanbul/Turkey, e-mail: yerdelen@istanbul.edu.tr

Abstract

Although large companies try to gain new customers, they also want to retain their old customers. Therefore, customer churn analysis is important for identifying old customers without loss and developing new products and making new strategic decisions for retaining customers. This study focuses on the customer churn analysis, that is a significant topic in banks customer relationship management. Identifying customer churn in banks will helps the management to classification who are likely to churn early and target customers using promotions, as well as provide insight into which factors should be considered when retaining customers. Although different models are used for customer churn analysis in the literature, this study focuses on especially explainable Machine Learning models and uses SHapely Additive exPlanations (SHAP) values to support the machine learning model evaluation and interpretability for customer churn analysis. The goal of the research is to estimate the explainable machine learning model using real data from banking and to evaluate many machine learning models using test data. According to the results, the XgBoost model outperformed other machine learning methods in classifying churn customers.

Keywords: Customer Loyalty, Customer Retention, Customer Churn Analysis, Machine Learning Models, Tree-Based Predictive Models

JEL Codes: C53, C55, M31

Citation: GULIYEV, H. & YERDELEN TATOGLU, F., (2021). Customer churn analysis in banking sector: Evidence from explainable machine learning models. *Journal of Applied Microeconometrics (JAME).* 1(2), 85-99, DOI: 10.53753/jame.1.2.03

Corresponding Author:
Hasraddin Guliyev
E-mail: hasradding@unec.edu.az

1. INTRODUCTION

Customer churn is a serious issue in an era of increasingly crowded markets and increased competition amongst businesses (Colgate et al. 1996). According to much research, the cost of acquiring new consumers is 1/5 the cost of keeping existing ones (Athanassopoulos 2000). For this reason, firms prefer to retain existing customers than add new ones and apply policies in this direction. One of the most valuable qualities in strategies designed to decrease or prevent customer churn is customer behavior data in the current customer base (Ganesh et al. 2000). As a result, as part of a consumer strategic plan targeted at decreasing customer churn, the discovery, and exploration for customers with a strong desire to leave the organization, or customer churn prediction, is essential (Blattberg et al. 2008).

Customer churn is a well-known issue in most sectors (Saradhi and Palshikar 2011), hence it's critical to develop a perfect predictive model designed in support of customer churn that could be used to formulate customer retention strategy. This topic is much more important in markets where competition is high and acquiring new customers is more difficult than retaining existing customers. Banking is one of the sectors where analyzing customer behavior and estimating customer churn based on these behaviors is an essential topic of research. Customer churn analysis results have a large impact on the bank's policy. Because the results of churn analysis allow banks to develop new customer strategies or improve existing ones. In addition, banks are critical to a country's financial growth and development, so the banking sector is an essential factor to the country's and people's financial stability. Because it is not always possible to get new customers in the competitive banking market, banks' primary goal is to ensure that existing customers are retention. Because banks, like all companies in the service sector, are customer-oriented, customer relationships with banks are a priority to their long-term business achievement. Studies conducted for the banking sector of various countries have revealed that, due to the competitive and dynamic nature of the banking sector, ensuring customer satisfaction is an important policy to prevent customer churn. Developing strong relationships with their customers have an advantage to high customer satisfaction and thus customer loyalty and turns into the most important factor for the stability, growth, and profitability of businesses.

In customer churn prediction in banking or other sectors, for example, a scoring approach supports the calculation of a potential churn probability per customer is based on their past data. The need to retain existing customers to maintain market share has led to an increased need for the development of various machine learning techniques for churn customer analysis. In customer churn prediction logistic regression (LR) is a very widespread paradigm to predict a churn probability since of great comprehensibility (Verbeke et al. 2012). The LR model, however, has a weak classification performance. On the other hand, despite the high prediction performance of the machine learning models explanation is difficult. In this study, we suggest an explainable machine learning model that combines the comprehensibility of logistic regression with the high classification performance of machine learning models.

2. LITERATURE REVIEW

With the advancement of techniques in the last 10-15 years, customer churn analysis studies using machine learning have grown in popularity. In the literature, it's clear that customer churn analysis has been used in a variety of industries (Kawale et al. 2009). While most studies have been done for the telecommunications and communication sector, customer churn analysis has also found applications in a wide range of fields, including e-commerce, banking, insurance, retail trade, energy, games and entertainment, and the medical (see Ahn et al. 2006; Bose and Chen 2009; Khan et al. 2010; Soeini and Rodpysh 2012; Buettgens et al. 2012; Long et al. 2012; Abbasimehr et al. 2013). Because this study is a customer churn analysis application in banking, it will be focused on this sector. Although the factors affecting customer satisfaction and loyalty differ for all countries, there are common points in the studies carried out.

According to Chakiso (2015), to obtain valuable customers, satisfy customers, and ensure customer loyalty, relational marketing (such as trust, bonding, communication, and reciprocity) must be used in the banking sector. Customers are encouraged through strong customer relationships to be always more satisfied and loyalty of bank, according to many studies. Conforming to Ozatac et al. (2016) the most significant determinants of customer satisfaction in banking are the accuracy of the information, the responsiveness of employees, access to all services, ability of employees, reliability, security of financial transactions, personalization, and consistency. As determined by Singh et al. (2013) lead to customer satisfaction are punctuality, effective communication, direct and acceptable information, efficient employee services in banking. Pasha and Waleed (2016) supervised a study in Pakistan to determine the factors that control customer loyalty. Customer satisfaction, brand loyalty, pricing policy, and service quality were found to be the most important determinants of customer loyalty in their study. The development of relational marketing dimensions such as quality service, personalized products, reliability, personalized communication, problem management, customer education, customer engagement, and the use of new technology, have been improved customer empowerment, according to Chatterjee and Kamesh (2020).

Besides relationship marketing, studies show that more objective reasons play a significant role in customer loyalty for banking. Many customers are served by banks through a variety of channels, including ATMs, mobile applications, and internet banking. Customers who have become more aware of service quality could be moving their financial services from one bank to another for a variety of reasons, including technological advancements, customer-friendly service, low-interest rates, geographic closeness, and a variety of services offered. When customers' options for service expand, a competitive market emerges. As a result, the competition among banks improves bank reliability and service quality significantly however it also increases the risk of churn to customers. In banking, as in many other sectors, developing a model which predicts customer is churn based on demographic, psychological, and transaction data is critical and in machine learning models are possible to predict who is churn customer and why. These predictive models have the advantage to lead to the design of personalized service and products and encourage customer loyalty with resulting in increased customer high satisfaction.

Mutanen (2006) offered a logistics regression-based customer churn study of the retail banking sector. Naveen et al. (2009) conducted detailed research with data mining techniques for churn customers that use credit cards. Bilal (2016) used gender, age, average monthly income, consumer status (retired, student, employed, unemployed), and whether the customer uses two or more bank products as control variables in the neural network model. According to Bilal customers that use multiple banking products are less probability of churn. Keramati et al. (2016) used the decision tree (DT) model to investigate churn customers in electronic banking (internet bank, telephone bank, mobile bank, ATM). They discovered the customers' dissatisfaction (duration of customer engagement, number of customer complaints), service usage (total number of uses and transactional amounts), and demographic variables (age, gender, employment status, education level) are effective on customer churn. Brânduşoiu et al. (2016) used a big dataset that includes 21 control variables for an advanced data mining model that predict prepaid customer churn. He et al. (2009) utilized a prediction model based on the Artificial Neural Network(ANN) algorithm for the complication of customer churn in a big Chinese telecom corporation with roughly 5.2 million consumers. The overall accuracy rate for prediction was 91.1 percent in the study. Nie et al. (2011) applied LR and DT models to predict churn customers using credit cards belonging to a Chinese bank. They discovered that the LR model outperformed from DT model to predict churn customers in a large dataset containing financial data from 135 variables for 60 million customers. Rajamohamed and Manokaran (2018) compared different classification models such as the k-nearest neighbor, Support Vectors Machine, Random Forest, Decision Tree, and Naive Bayes to predict customer churn in banking and discovered the Support Vectors Machine model was the most accurate, followed by the Random Forest model. Lopez-Diaz et al. (2017) compared 7 classification models for their predicting customer churn in a Spanish bank with 823,985 customers and observed that logistic regression was the greatest performance used for customer churn prediction.

In this study, in parallel with the literature, the effects of various factors such as age, income, gender, credit card status, and discount opportunities offered by banks on customer churn were examined with LR, DT, RF and Xgboost classification models.

3. METHODOLOGY

3.1. Logistic Regression Model (LR)

Logistic regression models discover the relationship among qualitative and other variables. In most models established with logistic regression, dependent variable has only two results. Usually, the emphasized event that is being realized is indicated by 1 and the one which is not realized by 0. The scientific society in the domains of economics, financial sector, and other social and environmental sciences gets now incorporated these models (Jabeur 2017; Zheng et al. 2020). The LR model is used to estimate the likelihood of an occurring event based on a set of predictors. The following is the predicted output of the logistic regression:

$$L_i = ln\left(\frac{P_i}{1-P_i}\right) = Z_i = \beta X_i + u_i \tag{1}$$

In the above expression, is a linear representation of the input variables and takes a value between -∞ and +∞, while takes a value between 0 and 1. LR has several statistical flaws. Multicollinearity and decreased performance accuracy are two of them.

3.2. Decision Tree

Ross Quinlan developed the C4.5 Decision Tree (DT) Classification Method as an expansion of the ID3 algorithm, which he previously created. These classifiers use the data samples to build a decision tree as a machine learning technique. The edge-based segmentation strategy is used to build decision tree models with an information gain metric used to select an appropriate input variable from among all the tree's input variables. The study selects a test drive through n outcomes that splits the data set *N*, as well as, training data set into subsets (N_1, N_2, N_3,..,N_k). *(Ci,P)* is the total number of samples in *P* that belong to C_i, and |P| is the total number of samples in *P*. The entropy of the set P is given by;

$$info(P) = -\sum_{i=0}^{k} \frac{freq(C_i,P)}{|P|} log_2\left(\frac{freq(C_i,P)}{|P|}\right) \tag{2}$$

The overall knowledge subject of *N* may be calculated after *N* is split with regard to the outcomes of a given characteristic, about *z*. N's information content may be calculated using Info *(N)*. The entire information content of *N* is equal to the weighted sum of each subset's entropies.

$$info_z(N) = \sum_{i=0}^{n} \frac{|N_i|}{|N|} info(N_i) \tag{3}$$

The gain is given by:

$$Gain(z) = info(N) \; - \; info_z(N) \tag{4}$$

It divides *N* about the test on *z* to give information. This is completed to choose the attribute *z* with the greatest knowledge gain. (1) Condition all the samples in a dataset fit to the same class, the decision tree generates a leaf node to select that class. (2) Otherwise, any input variable provides any information gain, a decision node leading the tree with the class's expected value is produced. (3) If an unknown instance's class is confronted, a decision node is constructed leading the tree together with the class's expected value.

3.3. Random Forest

Random Forest (RF) is a famous ML model used for data classification (Çağlayan et al. 2020). This algorithm is frequently utilized in sectors such as investing (Jabeur 2017), customer management and marketing (Salminen et al. 2019). A group of trees underpins the RF. It is complemented with an aggregate of the prediction's mean value, which is produced at the conclusion of each of the trees, reducing the lack of robustness of a single tree. Each of the trees is created using a subset of input variables that are picked at random. The following is an expression for the estimated model:

$$\hat{y} = \frac{1}{n}\sum_{i=1}^{n} g_k(x) \tag{5}$$

The vector of input features is *x*, and *g(x)* is a collection of the k[th] learner random trees. The RF final estimate is the mean of the outcome of the whole trees. As a result, with such weights, each individual tree has an impact on the RF estimation. Corresponding to Yeşilkanat (2020), the Random Forest model is superior to other machine learning methods. This is due to the former' stability in the direction of acquiring training data from subsets automatically and shaping trees using random techniques. Furthermore, because the Random Forest model achieves training by applying bootstrapping on a randomly chosen independent subset of datasets, the overfitting quantity is preserved.

3.4. eXtreme Gradient Boosting

The eXtreme Gradient Boosting (XgBoost) is the model that implements Chen and Guestrin's (2016) gradient boosting technique. It is a widely utilized flexible tool on the way to tree boosting algorithm achieves cutting-edge classification and effectiveness (Mai et al. 2020). The result is generated by the XgBoost, which is a collection of regression trees. The following equation is used to arrive at the final score:

$$\hat{y} = \sum_{h=1}^{H} g_h(x) \tag{6}$$

The number of trees in this equation is H, and the score for each tree's leaf is K. Multicollinearity has no effect on the XgBoost, which is an additional benefit. In order to maximize the model performance, XgBoost involves the selection of certain parameters. Parameter tuning is essential for the XGBoost to get around overfitting and too much confusion of the model. But, because the XgBoost utilizes multiple settings, this can be difficult. On the way to maximize the hyper-parameter values, we applied the grid search method with cross-validation.

3.5. The Performance Metrics of Classification Models

In order to determine which of the applied machine learning classification models are more successful both individually and among themselves, some performance metrics must be examined (Çağlayan 2020). These metrics are used to assess the effectiveness of the classification method in use and to compare classification models. Multiple metrics of models should be considered because evaluating these values as a single success criterion would be incorrect. All observation in the test data set is replaced in the model created with the training data set in the classification models, and classification prediction scores are achieved. The results of comparing the predicted values with the actual values are used to determine how well this model predicts, as well as its success and performance. The confusion matrix summarizes the results of the model's accuracy in making a prediction, as well as the conclusions of the performance evaluation of the machine learning classification model.

Figure 1. Confusion Matrix

Confusion Matrix		**Actual Values**	
		0	**1**
Predicted Values	**0**	True Positive TP	False Positive FP
	1	False Negative FN	True Negative TN

Figure 1 shows the confusion matrix is explained as follows for a two-category classification model:

True Positive (TP); indicates that observations with a true class value of 1 are correctly predicted as 1.

True Negative (TN); indicates the situation where observations with a true class value of 0 are correctly predicted as 0.

False Negative (FN); shows that observations with a true class value of 1 are incorrectly evaluated as 0 as a result of the prediction.

False Positive (FP); shows that observations with a true class value of 0 are incorrectly evaluated as 1 as a result of the prediction (Deng et al. 2020).

The accuracy rate (ACC) is calculated by taking the ratio of the number of classified observations () to the total number of samples (). This enables the evaluation of the estimation of the value of the estimation result made with the classification model as 1 when the true value of a class is 1, and the case that the estimated value of the class is 0 when the true value of the class is 0. ACC can be calculated using the following formula:

$$ACC = \frac{TP + TN}{FN + FP + TP + TN} \tag{7}$$

With a confusion matrix, we can also calculate sensitivity and specificity rates. The sensitivity is the ratio of correctly classified (TP) positive input values to the total true positive values (TP + FN):

$$Sensitivity = \frac{TP}{TP + FN} \tag{8}$$

The specificity is the ratio of the correctly classified (TN) to the total positive values (TN+FP) of the number of observations:

$$Specificity = \frac{TN}{TN + FP} \tag{9}$$

The sensitivity, specificity rate and the ACC metrics have values ranging from 0 to 1, and if they are close to 1, the model's performance is exceptionally good. Furthermore, sensitivity and specificity are inversely related, which means that as sensitivity grows, specificity decreases and conversely (Lambert and Lipkovich 2018).

The other metric AUC is a measurement of the entire area under the Receiver Operating Characteristic (ROC) curve, and it is one of the metrics used to evaluate model performance, together with the ROC curve (Jabeur et al 2021). The AUC value ranges from 0 to 1, with a value near 1 indicating a more accurate model. The distributions of TN and TP do not intersect when the area under the ROC curve is large, indicating that the classes have been successfully separated (Mai et al 2019).

3.6. Imbalanced Classification Problems

For customer churn analysis, studies have identified an imbalanced class distribution on customer data sets. Because the sample size of churn customers is substantially less than that of non-churn customers, the following scenario might occur; the accuracy of the classification is high, while churn customer prediction accuracy is low. So, the problem with unbalanced datasets is that typical classification learning techniques are typically biased towards the majority classes (referred to as "negative"), resulting in a greater misclassification rate in minority class occurrences (referred to as "positive" class) (Chawla 2009). The most common approach to this problem is to use a resampling technique to balance the class distribution of the training set before training a classification model. Random oversampling (ROS) and random undersampling (RUS) are two approaches for resampling (RUS). ROS, which consists of decreasing the data by deleting instances belonging to the majority class with the goal of equalizing the number of examples of each class; and RUS, which intends to reproduce or generate new positive examples in order to acquire importance (Batista et al. 2004). The main disadvantage of random undersampling is that it might lose potentially relevant data that could be significant in the induction process. The elimination of data is an important decision to make, hence many undersampling proposals include heuristics to overcome the limits of non-heuristic decisions. Random oversampling, on the other hand, may increase the likelihood of overfitting since it duplicates the minority class instances exactly. In this manner, a symbolic classifier, for example, may generate rules that appear to be accurate but only cover one reproduced case.

Ensemble learning (tree-based) models are another option for improving the performance of a single classifier by training multiple separate classifiers and integrating their outputs to produce the final choice (Kuncheva 2004). Cost-sensitive ensembles, on the other hand, use the ensemble learning algorithm to lead cost reduction rather than altering the underlying classifier in order to accept costs in the learning process. Ensemble learning models include Random Forest, AdaBoost, and XgBoost. Ensemble Learning models are well-known in data mining and machine learning for their good performance in a wide range of applications, and it may be the better alternative for the class imbalance problem (Wozniak 2014). For example, Ahmad et. al (2019) discovered that tree-based models performed better from undersampling for unbalanced classification in their customer churn analysis in the telecommunications sector.

4. DATA AND ANALYSIS

4.1. Data and Variables

This paper aims on application of machine learning models for predicting the churn customers. The research is based on real data from a bank. Before customer churn analysis, we need to determine churn status of customers. Customers who close individual loans and do not apply for new loans despite 9 months after the close date of the loan are included in the churn customer category and take 1 value for the dependent variable of classification model, however, it takes 0 value when customer applying to loan during the 9 months from the close date of loan. According to the calculations, 91% of the customers generally applied for the second loan within 9 months, so we used the 9-month criterion to determine customer churn status.

The database of a bank was used in the data collecting process, and 274,542 observations were analyzed once all the pre-elimination processes are completed. The most appropriate input variables which according to the local market conditions are selected for predicting the customer's churn status and summarized in Table 1.

Table 1. Definition of Variables

Variable	Assigned Short Name	Definition
churn status	churn_status	shows customer's churn category. If the customer is churn, it takes a value of 1, if is non-churn, a value of 0.
Input Variables		
customer's age	age	Shows customer's age
average income	salary	shows the average income of the customer for the last 12 months.
gender	gender	if the customer is male, it takes a value of 1, if is female, a value of 0.
loan amount	amount	amount of customer's last used credit
interest rate	interest_rate	nominal interest rate calculated to the customer's last used loan
credit term	duration	duration of customer's last used credit (months)
credit closing and early payme nt (days)	closed	It gets negative values daily if the customer has paid the loan before the expiry date; positive values if the loan has been delayed.
interest rate discount	rate_discount	if the customer is offered an interest rate discount in last loan compared to previous loan, it gets a value of 1 and in other case a value of 0.
amount increase	amountup	if the customer is offered an increase in the loan amount in last loan compared to previous loan, it gets a value of 1 and in other case a value of 0.
competition region	competition	if customers live in competition region where more branch of other banks exist 1, if not, a value of 0.
credit card status	creditcard	if the customer has a credit card, it takes a value of 1, if not, a value of 0.
salary card	card_status	if the customer's salary card belongs to bank, it takes the value 0 and if it belongs to another bank, it takes the value of 1.
credit count	creditcount	it shows how many individual loans the customer has drawn to determine the relationship with the bank. since it is a categorical variable, customers with 1 individual loan were taken as a base and 3 dummy variables were

4.2. Describing of Variables and Correlation Analysis

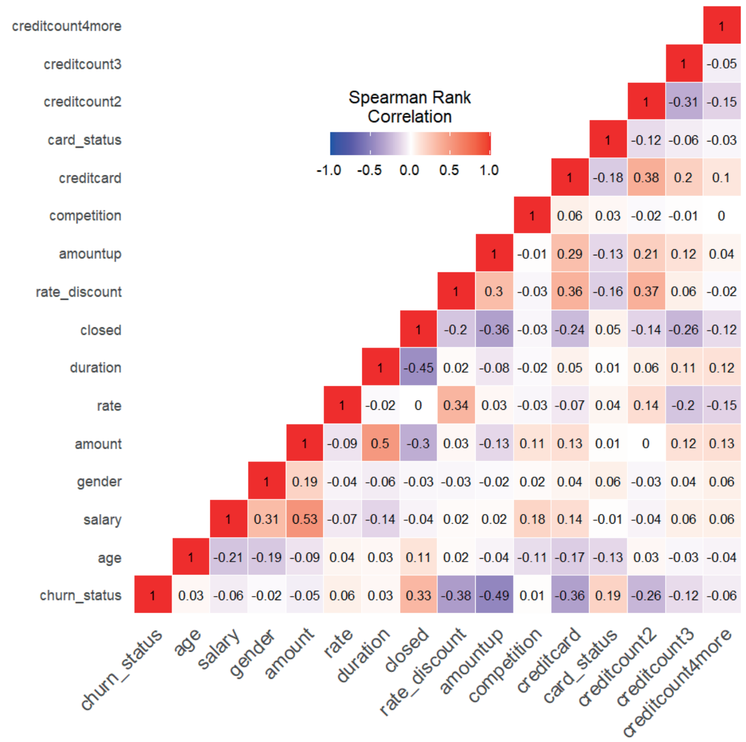
Table 2 summarizes the descriptive statistics for all variables examined in this study. When looking at the proportion of outcome variable that is churn status, we see that 88.5 percent of customers are no churner and 11.6 percent of customers are churner. The average age of the customers that input variable is 49.5, with a standard deviation of 12.8, and a range of 19 to 70. The mean annual salary is 309, with a low of 50 and a high of 11850. When we look at the categorical input variable like gender, we notice that males account for 52.3 percent of the customer while females account for 47.7%. In other input variables, it can be interpreted the same way.

Table 2. Descriptive Statistics of Variables

Variables	Statistics / Frequency	Variables	Statistics / Frequency	Variables	Statistics / Frequency
churn status	No-churn (88.5%) Churn (11.5%)	interest rate	Mean (sd) : 28.4 (2.2) min ≤ med ≤ max: 14 ≤ 28.2 ≤ 39	competition	No (73.0%) Yes (27.0%)
age	Mean (sd): 49.5 (12.8) min ≤ med ≤ max: 19 ≤ 52 ≤ 70	duration	Mean (sd): 29.2 (10.5) min ≤ med ≤ max: 3 ≤ 36 ≤ 156	credit card	No (31.5%) Yes (68.5%)
salary	Mean (sd): 309 (250.9) min ≤ med ≤ max: 50 ≤ 199.6 ≤ 11850	closed	Mean (sd): -370.6 (378.9) min ≤ med ≤ max: -1820 ≤ -264 ≤ 120	card status	No (92.7%) Yes (7.3%)
gender	Male (52.3%) Female (47.7%)	rate discount	No (39.7%) Yes (60.3%)	credit count	1 (41.2%) 2 (46.0%) 3 (10.3%) 4+ (2.5%)
credit amount	Mean (sd): 2558.4 (2132.8) min ≤ med ≤ max: 300 ≤ 2000 ≤ 20000	amount up	No (31.1%) Yes (68.9%)		

We have calculated the correlation coefficients before using machine-learning techniques to ensure that the input variable selection is accurate. Figure 2 shows the pairwise Spearman`s rank correlation among the variables in our study:

Figure 2. Spearman's Correlation Heatmap



Source: Authors 'own calculation from the dataset.

Spearman's correlation evaluates monotonic relationships, and it is a reliable tool for big datasets with outliers. Calculated Spearman's correlation coefficients indicate customer churn status has a moderate negative correlation with rate discount (ρ=-0.38) and the amount up (ρ=-0.49) in the next credit approval. Furthermore, customer churn has a moderate negative correlation with credit card status (ρ=-0.38) and credit counts variables (creditcount2, creditcount3 and creditcount4+), however a positive correlation with credit closed (ρ=0.33) and card status (ρ=0.19) that indicating the customer's salary card belongs to another bank. Many variables' correlation coefficients were calculated at a low level. As a result, using nonlinear machine learning models to better explain the relationship with churn status may be advantageous because of low-level correlation coefficients.

4.3. Machine Learning Models' Overall Performance and Variable Importance

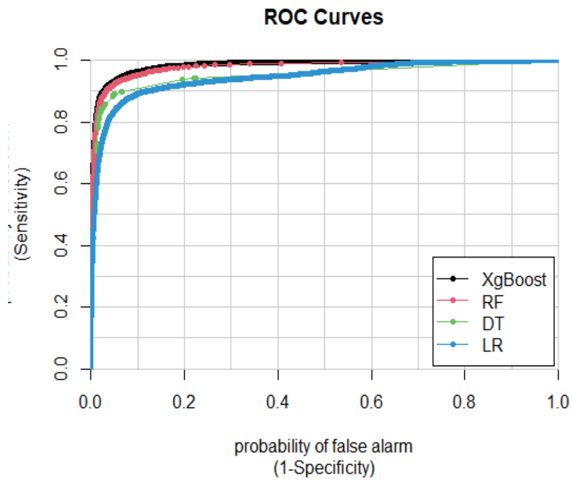
The performance of the indicated machine-learning models is compared in this section. The data from 274,542 customers was split into two parts; 80% of the dataset (219,634 customers) was used to train machine learning models, while 20% of the dataset (54,908 customers) was used to test them. The effectiveness of each individual model was determined using some performance metrics. We have explored the results of the proposed machine learning models and evaluated these results through the model performance metrics (i.e., sensitivity, specificity, accuracy, and Receiver Operating Characteristic- ROC curve).

Table 3. The Performance of Machine Learning Models in Testing Data

Performance Metrics	Machine Learning Models			
	LR	DT	RF	Xgboost
Sensitivity	0.9896	0.9819	0.9843	0.9854
Specificity	0.5693	0.8206	0.8387	0.8504
ACC	0.9410	0.9632	0.9675	0.9697
AUC	0.9464	0.9510	0.9797	0.9850

Table 3 expresses the results computed with the advanced machine-learning models such as Logistic regression (LR), Decision Tree (DT), Random Forest (RF), and XgBoost. We used the *caret* package in R for model estimates in the training dataset and parameter turning to deal with the overfitting problem and model performance boosting. To calculate hyperparameters, we used 5-cross validation and the grid search method. First, we used the Logistic regression model; while its sensitivity (0.9896) is high, its specificity (0.5693) is very low, so we cannot conclude that the LR model is very accuracy model in classification. For the Decision Tree model, the complexity parameter is estimated to be 0.0012. Although the DT model's sensitivity (0.9819), specificity (0.8296), ACC (0.9410) and AUC (0.9510) are acceptable, it cannot be considered the best model. The maximum depth parameter, which is the number of variables randomly sampled as candidates at each split, is calculated as 4 for the Random Forest model. Although the performance of the RF model is very good, it still cannot be said to be the best model. We used the XgBoost model and determined that the optimal number of tree sizes was 130. Because of the values of sensitivity (0.9854), specificity (0.8504), accuracy (0.9697), and AUC (0.9850), as well as the closeness of these metrics to 1, the XgBoost model had a higher predictive performance in the test data set when all models were compared. To emphasize, we would like to point out that a more accurate estimation of the churned customer (positive class) is more important for customer churn analysis, so having a higher specificity rate is more advantageous for us. Consequently, the area under of Receiver Operating Characteristic (ROC) curve plotted for XgBoost is higher compared to other machine learning models.

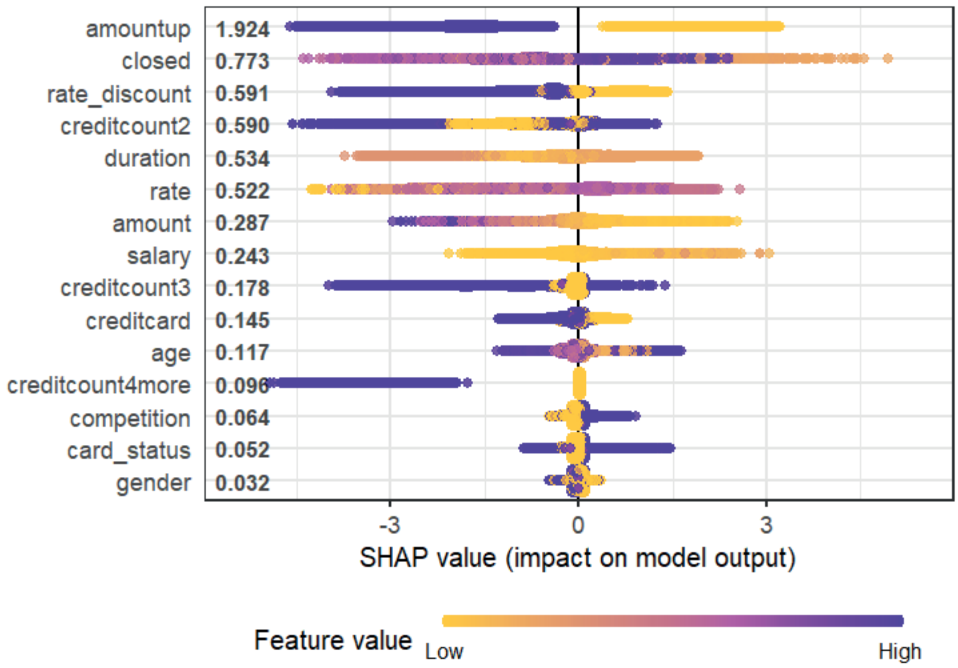
Figure 3. Compare of Receiver Operating Characteristic (ROC) Curves



Source: Authors 'own calculation from the dataset.

It is useful to know the proportional contributions of all factors on the final forecast outcome when predicting the churner. Lundberg et al. (2018) recently suggested the SHAP to assess the significance of specific characteristics. This can benefit in balancing the accuracy and interpretability of black-box machine-learning models. The importance of variables is shown in Figure 4.

Figure 4. The Shapley values of XgBoost Model.



Source: Authors 'own calculation from XgBoost Model.

Figure 3 shows the importance of the variables with the effects of the variables. A Shapley value for a feature and an instance is represented by each point on the summary plot. The feature is demonstrated on the y-axis position, while the Shapley value is demonstrated on the x-axis position. The value of the feature is represented by the color, which ranges from low to high. To get a meaning of the Shapley value per input variable, the overlapping points are jittered in the y-axis direction. The features are ranked from the most important to the least important one. The 5 most important variables to explain customer churn status are the customer is offered a higher loan amount compared to the previous loan (amount up), early payment of credit or delaying credit (closed), the customer is offered a lower interest rate compared to previously loan (rate discount), the number of credits (credit count 2) and the duration of customer's last used credit (duration). Shapley values show that when an interest rate discount (rate discount) and more loan amount (amount up) is determined is applied to a new loan, the probability of churn decreases. In addition, if the number of credits (credit number 2, credit number 3, credit number 4 +) increases, the probability of churn decreases. Interestingly, a decrease in the duration of the previous loan and having a credit card (card status) reduces the probability of customer churn. However, the probability of churn increases as the customer's age rises and another bank provides a monthly salary to customers with their own bank card (this is due to the change in the customer's workplace). Furthermore, the interest rate of credit and the customer to live in the competitive region increases the probability of churn.

5. CONCLUSION

In this paper, we proposed ML methods for predicting customer churn. The machine learning predictive models need to achieve high AUC values. Firstly, to test and train the model, the sample dataset is divided into 80% for training and 20% for testing. For validation and hyperparameter tuning, we selected to use 5-fold cross-validation. In addition, we contended with another problem: the data was not balanced and only about 11.5% of the dataset had included churn customers. To solve this problem, undersampling methods or tree base algorithms are suggested, so we applied the tree-based models such as Decision Tree, Random Forest, and XgBoost. The XgBoost outperformed for each metric, with an AUC of 96.97 percent and the Random Forest model comes in second.

Shapley values present the most important variable that explains the churn status of the customers and it indicated positive or negative effect of input variables for XgBoost model. In general, the important reasons that increase the risk of churn of the customer are the fact that the salary card belongs in another bank in the next period and that the customer inhabits in the competitive region where there are other alternatives to take credit. The most important reason that increases the probability of churn is that the applied credit interest is high, which in turn reduces the customer's ability to repay the loan and shifts the customer towards the other bank that offers more favorable interest rates for subsequent loans. The customer's relationship with the bank is a critical component in minimizing the risk of customer churn. If the customer has a longer-term relationship with the bank, then the customer will benefit from the advantages offered by the bank's loyalty program and will maintain the relationship with the bank for a long time. In addition, applying the interest rate discount and upping the amount of credit to the customer decreases the risk of churner.

Our results suggest that building a model that can accurately anticipate customer retention might have some management and financial consequences for banking in order to reduce the probability of churn. Firstly, correctly classifying a customer as a churner or non-churner helps decrease the expenses associated with misclassification. Second, our findings show that academics and practitioners do not have to rely exclusively on conventional methodology as logistic regression for predicting customer churn. Finally, our findings suggest management recommendations for improving the decision-making process in the context that customer churn prediction. Banks and financial institutions may use XgBoost models to correctly identify clients who are at risk of churn, focus their efforts on them, and potentially get profit. Companies should more focus on customer retention policies rather than concentrating on new target markets, which are generally difficult to gain. So, the findings of the machine-learning techniques of this research could have a variety of policy implications for customer relationship management and the marketing strategy of the company. In the future, more explainable machine learning methods should be used, and models with higher performance should be suggested for predicting customer churn.

Funding

The authors declare that this study has no financial support.

Conflict of Interest

The authors declare that they have no conflicts of interest.

REFERENCES

- ABBASIMEHR, H., SETAK, M., & SOROOR, J. (2013). A framework for identification of high-value customers by including social network-based variables for churn prediction using neuro-fuzzy techniques. *International Journal of Production Research*, *51*(4), 1279-1294.

- AHMAD, A. K., JAFAR, A., & ALJOUMAA, K. (2019). Customer churn prediction in telecom using machine learning in big data platform. *Journal of Big Data*, *6*(1), 1-24.

- AHN, J. H., HAN, S. P., & LEE, Y.S. (2006). Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry. *Telecommunications policy*, *30*(10-11), 552-568.

- AKAY, E.C., SOYDAN, N.T.Y. & GACAR, B.K. (2020). MAKİNE ÖĞRENMESİ VE EKONOMİ: BİBLİYOMETRİK ANALİZ. PressAcademia Procedia, 12 (1), 104-105.

- ATHANASSOPOULOS, A.D. (2000). Customer satisfaction cues to support market segmentation and explain switching behavior. *Journal of business research*, *47*(3), 191-207.

- BATISTA, G.E., PRATI, R.C., & MONARD, M.C. (2004). A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD explorations newsletter*, *6*(1), 20-29.

- BILAL ZORIĆ, A. (2016). Predicting customer churn in banking industry using neural networks. *Interdisciplinary Description of Complex Systems: INDECS*, *14*(2), 116-124.

- BLATTBERG, R. C., KIM, B.D., & NESLIN, S.A. (2008). Churn Management. In *Database Marketing* (pp. 607-633). Springer, New York, NY.

- BOSE, I., & CHEN, X. (2009). Quantitative models for direct marketing: A review from systems perspective. *European Journal of Operational Research*, *195*(1), 1-16.

- BRÂNDUŞOIU, I., TODEREAN, G., & BELEIU, H. (2016). Methods for churn prediction in the pre-paid mobile telecommunications industry. In *2016 International conference on communications (COMM)* (pp. 97-100). IEEE.

- BUETTGENS, M., NICHOLS, A., & DORN, S. (2012). Churning under the ACA and state policy options for mitigation. *Prepared for Robert Wood Johnson Foundation, Timely Analysis of Immediate Health Policy Issues, http://www.urban. org/UploadedPDF/412587-Churning-Under-the-ACA-and-State-Policy-Options-for-Mitigation. pdf.*

- ÇAĞLAYAN AKAY, E. (2018). Ekonometride Yeni Bir Ufuk: Büyük Veri ve Makine Öğrenmesi. Social Sciences Research Journal, 7(2): 41-53.

- ÇAĞLAYAN AKAY, E. (2020). Ekonometride Büyük Veri ve Makine Öğrenmesi: Temel Kavramlar, Der Yayınları, İstanbul.

- CHAKISO, C.B. (2015). The effect of relationship marketing on customers' loyalty (Evidence from Zemen Bank). *EMAJ: Emerging Markets Journal*, *5*(2), 58-70.

- CHATTERJEE, D., & KAMESH, A.V.S. (2020). Significance of Relationship marketing in banks in terms of Customer Empowerment and satisfaction. *European Journal of Molecular & Clinical Medicine*, *7*(4), 999-1009.

- CHATTERJEE, D., SEKHAR, S.C., & BABU, M.K. (2021). Customer Empowerment-A Way to Administer Customer Satisfaction in Indian Banking Sector. *NVEO-NATURAL VOLATILES & ESSENTIAL OILS Journal| NVEO*, 1621-1629.

- CHAWLA, N.V. (2009). Data mining for imbalanced datasets: An overview. *Data mining and knowledge discovery handbook*, 875-886.

- CHEN, T., & GUESTRIN, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).

- COLGATE, M., STEWART, K., & KINSELLA, R. (1996). Customer defection: a study of the student market in Ireland. *International journal of bank marketing*.

- DENG, X., LIU, Q., DENG, Y., & MAHADEVAN, S. (2016). An improved method to construct basic probability assignment based on the confusion matrix for classification problem. *Information Sciences*, *340*, 250-261.

- GANESH, J., ARNOLD, M.J., & REYNOLDS, K.E. (2000). Understanding the customer base of service providers: an examination of the differences between switchers and stayers. *Journal of marketing*, *64*(3), 65-87.

- HE, Y., HE, Z., & ZHANG, D. (2009). A study on prediction of customer churn in fixed communication network based on data mining. In *2009 sixth international conference on fuzzy systems and knowledge discovery* (Vol. 1, pp. 92-94). IEEE.

- JABEUR, S.B. (2017). Bankruptcy prediction using partial least squares logistic regression. *Journal of Retailing and Consumer Services*, *36*, 197-202.

- JABEUR, S. B., GHARIB, C., MEFTEH-WALI, S., & ARFI, W.B. (2021). CatBoost model and artificial intelligence techniques for corporate failure prediction. *Technological Forecasting and Social Change*, *166*, 120658.

- KAWALE, J., PAL, A., & SRIVASTAVA, J. (2009). Churn prediction in MMORPGs: A social influence based approach. In 2009 international conference on computational science and engineering (Vol. 4, pp. 423-428). IEEE.

- KERAMATI, A., GHANEEI, H., & MIRMOHAMMADI, S. M. (2016). Developing a prediction model for customer churn from electronic banking services using data mining. *Financial Innovation*, *2*(1), 1-13.

- KHAN, A.A., JAMWAL, S., & SEPEHRI, M.M. (2010). Applying data mining to customer churn prediction in an internet service provider. *International Journal of Computer Applications*, *9*(7), 8-14.

- KUNCHEVA, L. (2004). Combining pattern classifiers methods and algorithms. john wiley&sons. *Inc. Publication, Hoboken*.

- LAMBERT, J., & LIPKOVICH, I. (2008). A macro for getting more out of your ROC curve. In *SAS Global forum, paper* (Vol. 231).

- LONG, X., YIN, W., AN, L., NI, H., HUANG, L., LUO, Q., & CHEN, Y. (2012, March). Churn analysis of online social network users using data mining techniques. In *Proceedings of the international MultiConference of Engineers and Conputer Scientists* (Vol. 1).

- LÓPEZ-DÍAZ, M. C., LÓPEZ-DÍAZ, M., & MARTÍNEZ-FERNÁNDEZ, S. (2017). A stochastic comparison of customer classifiers with an application to customer attrition in commercial banking. *Scandinavian Actuarial Journal*, *2017*(7), 606-627.

- LUNDBERG, S. M., ERION, G. G., & LEE, S. I. (2018). Consistent individualized feature attribution for tree ensembles. *arXiv preprint arXiv:1802.03888*.

- MAI, F., TIAN, S., LEE, C., & MA, L. (2019). Deep learning models for bankruptcy prediction using textual disclosures. *European journal of operational research*, *274*(2), 743-758.

- MUTANEN, T. (2006). Customer churn analysis–a case study. *Journal of Product and Brand Management*, *14*(1), 4-13.

- Naveen, N., Ravi, V., & Kumar, D. A. (2009). Application of fuzzyARTMAP for churn prediction in bank credit cards. *International Journal of Information and Decision Sciences*, *1*(4), 428-444.

- NIE, G., ROWE, W., ZHANG, L., TIAN, Y., & SHI, Y. (2011). Credit card churn forecasting by logistic regression and decision tree. *Expert Systems with Applications*, *38*(12), 15273-15285.

- OZATAC, N., SANER, T., & SEN, Z.S. (2016). Customer satisfaction in the banking sector: the case of North Cyprus. *Procedia Economics and Finance*, *39*, 870-878.
- RAJAMOHAMED, R., & MANOKARAN, J. (2018). Improved credit card churn prediction based on rough clustering and supervised learning techniques. *Cluster Computing*, *21*(1), 65-77.
- RISSELADA, H., VERHOEF, P.C., & BIJMOLT, T.H. (2010). Staying power of churn prediction models. *Journal of Interactive Marketing*, *24*(3), 198-208.
- SALMINEN, J., YOGANATHAN, V., CORPORAN, J., JANSEN, B.J., & JUNG, S.G. (2019). Machine learning approach to auto-tagging online content for content marketing efficiency: A comparative analysis between methods and content type. *Journal of Business Research*, *101*, 203-217.
- SARADHI, V.V., & PALSHIKAR, G.K. (2011). Employee churn prediction. *Expert Systems with Applications*, *38*(3), 1999-2006.
- SINGH, S., ANUSHA, B., & RAGHUVARDHAN, M. (2013). Impact of Banking Services on Customer Empowerment, Overall Performance and Customer Satisfaction: Empirical Evidence. *Journal of Business and Management (IOSR-JBM)*, *16*(1), 17-24.
- SOEINI, R. A., & RODPYSH, K. V. (2012). Applying data mining to insurance customer churn management. *International Proceedings of Computer Science and Information Technology*, *30*, 82-92.
- VERBEKE, W., DEJAEGER, K., MARTENS, D., HUR, J., & BAESENS, B. (2012). New insights into churn prediction in the telecommunication sector: A profit driven data mining approach. *European journal of operational research*, *218*(1), 211-229.
- WALEED, A., PASHA, A., & AKHTAR, A. (2016). Exploring the impact of liquidity on profitability: Evidence from banking sector of Pakistan. *Journal of Internet Banking and Commerce*, *21*(3).
- WOŹNIAK, M., GRANA, M., & CORCHADO, E. (2014). A survey of multiple classifier systems as hybrid systems. *Information Fusion*, *16*, 3-17.
- YEŞILKANAT, C. M. (2020). Spatio-temporal estimation of the daily cases of COVID-19 in worldwide using random forest machine learning algorithm. *Chaos, Solitons & Fractals*, *140*, 110210.
- ZHENG, K., ZHANG, Z., & SONG, B. (2020). E-commerce logistics distribution mode in big-data context: A case analysis of JD. COM. *Industrial Marketing Management*, *86*, 154-162.
- ZORIĆ, A.B. (2016). Predicting customer churn in banking industry using neural networks. *Interdisciplinary Description of Complex Systems: INDECS*, *14*(2), 116-124.

"This page is left blank for typesetting"

JAME
E-ISSN:2791-7401
Volume : 1, Issue 2, 2021
URL: https://journals.gen.tr/jame

Received: 10.08.2021
Acccepted: 15.12.2021

RESEARCH ARTICLE

Income inequality-labor productivity relationship: CS-ARDL approach*

Hilal Bağlıtaş[1]

1 Erciyes University, Faculty of Economics and Administrative Sciences, Kayseri/Turkey, e-mail: hhilalbaglitas@erciyes.edu.tr

Abstract

With the acceleration of globalization, "Reducing Inequalities", which is the 10th of the sustainable development goals, has started to attract more attention in the world. Many factors lead to inequality. Therefore, inequality requires consensus and strength at the interdisciplinary, local, national, and international levels. The leading indicator of inequality is income inequality. Its measurability and widespread impact are sources of its importance and priority. Unfair income distribution might have unfavorable effects on employees such as being more reluctant to work and the well-being of workers. In addition, if workers believe they earn less than they deserve, this might negatively affect the labor productivity. Ultimately, this process may cause countries to reduce their production output.

This study aims to explore the link between income inequality and labor productivity among 31 countries in Europe with the period of 2005-2019. To do this, a cross-sectional auto-regressive distributed lag model (CS-ARDL) is employed. According to the results, wage inequality damages the productivity of labor. A 1% increase in the wage inequality reduces labor productivity by 0.16%. Moreover, the unequal income distribution has an explanatory power of approximately 33% on the decrease in productivity. This helps to determine the possible effects of the unequal income distribution leading towards two targets. These targets are to create an efficient wage structure and eliminate the destructive effects of inequality, respectively. In terms of the policy effectiveness, simultaneous application of tools may be more beneficial.

Keywords: Wage Inequality, Labor Productivity, Sustainable Development, CS-ARDL

JEL Codes: C23, D63, J24

* This paper is the extended version of "Investigation of the Relationship Between Income Inequality and Labor Productivity in the Context of Sustainable Development" presented at the International Conference on Eurasian Economies between 24-25 August 2021 – Istanbul, TURKEY.

Citation: BAĞLITAŞ, H., (2021). Income inequality-labor productivity relationship: CS-ARDL approach. *Journal of Applied Microeconometrics (JAME)*. 1(2), 101-111, DOI: 10.53753/jame.1.2.02

Corresponding Author:
Hilal Bağlıtaş
E-mail: hhilalbaglitas@erciyes.edu.tr

1. INTRODUCTION

In Brundtland Report (1987), the definition of sustainable development (SD) is "the needs of the present without compromising the ability of future generations to meet their own needs." The SD, led by the United Nations (UN), includes 17 objectives and many sub-objectives for achieving a sustainable future. The UN prioritized three SD goals from 2015 to 2030. One of the three main objectives is to "fight inequality & injustice." Its inclusion in this emergency plan demonstrates the importance of the subject (The Global Goals 2021).

The importance of inequality is related to its effects on various areas. In this respect, inequality negatively affects the well-being of individuals. It also affects the productivity level of the labor (hence, the economy). Moreover, inequality disrupts resource allocation. Therefore, it is a concern of sustainability because of the limited resources. One aspect of inequality is a waste of resources when the advantaged group has more resources than they need. On the other hand, disadvantaged groups have fewer resources, which they have to share.

Terminologically, not being equal in status, rights and opportunities is defined as inequality. This issue has different manifestations (Afonso et al. 2015: 1). Income or wage inequality is the most popular inequality measurement in social sciences. Uneven income distribution among the population is the definition of income inequality. Unequal wage distribution is a source of income inequality. On the opposite side, a fair wage allocation has the criteria of equal pay for equal work (Meidner and Rehn, 1957 cited in Policardo et al. 2018: 3). Income inequality sourced from wages occurs when work and wages are out of balance.

Usually, the formation mechanism of inequality is the focus of scientific research. However, if one would like to reduce income inequalities for a sustainable society, knowing about the consequences of inequalities can pave the way for determining for which, how much and how the necessary tools are applicable. In addition, income is one of the basic human needs for individuals and households. A sufficient and fair income level is a requirement to sustain the well-being. Moreover, equitable and effective wage level is also significant for the firms, industries, and economy. Unequal income distribution can cause employees to lose their enthusiasm and thus, to lower productivity and output per employee. Therefore, a sustainable economy might be exposed to unequal distribution negatively. On the productivity side, it is a rate between input and output. In addition, this ratio calculates the efficiency level of inputs such as capital and labor. Countries should measure and raise their productivity levels if they are enthusiastic about the high growth rates and competitiveness. Additionally, productivity data is beneficial to evaluate the performance of labor and product markets (Krugman 1994: 1). Labor productivity in the labor market has influences on the efficient wage level. More productive workers get higher wages than fewer ones. Ultimately, all workers get the wages they deserved.

Income inequality increases when income is distributed unjustly and disproportionately among workers. Rising inequality rates are mainly due to the financial liberalization and trade globalization (Stiglitz 2013: 59). Unequal income distribution is especially true for the manufacturing sector. Technological progress in this sector causes an increase in productivity and reduces the need for employment. At the same time, changes in the employment structure are favor more skilled workers and against unskilled workers. In addition, differences in skills generally lead to differences in wages (Stiglitz 2013: 61). Wages, which are especially optimal wages, differ based on industry, sector, and firm (Stiglitz 1982: 78).

Moreover, productivity and wage level have a bidirectional relationship. Productive workers may receive higher wages, or higher wage levels increase their productivity by leading workers to exert more effort. The structure is also significant for perceived fairness in wages. If workers believe that the wage structure is not equitable, belief in inequality is on hand. Akerlof (1984) and Akerlof & Yellen (1986) argue that employees put in less effort if they believe they are underpaid (cited in Liu 2002: 454). If workers believe that the wage structure is unjust, their belief would reinforce the reality of income inequality.

There is a wealth of literature related to inequality and productivity measurement focusing on the effect of productivity on wage or vice versa. There are only a few papers directly mentioning the link between wage inequality and productivity level. Moreover, inequality indicators generally use the Gini coefficient and firm or sector data but not country-level. Related studies do not examine the variables for the long-run relationship. In the light of this information, our contribution to the literature is in three aspects. The first contribution is using the income quintile share ratio as an inequality indicator. The latter is to use the country-level panel data and the last contribution is the estimation of long-run coefficients by employing the cross-sectional panel cointegration technique.

In this regard, the second section covers a literature review. The third section gives information about the data and the methodology. The last section reveals the empirical results.

2. LITERATURE

Factor productivity has been a topic of much thought and discussion within the field, especially for economics. There is a large and diverse literature on productivity. Some of them examine the determinants of productivity such as Isaksson (2007), Choudhry (2009), Islam (2008), Isaksson and Ng (2006), Khan et al. (2011), Loko and Diouf (2009), Kose et al. (2009), Aghion et al. (2009). Some papers also investigate regional or spatial differences in productivity or unequal productivity distribution (i.e. He et al. 2017; Ezcurra et al. 2007).

Another group argues the historical process of macroeconomic variables with inequality and productivity. Paul (2020) focuses on the historical process of inequality and productivity. He finds that rising inequality and low productivity are predictors of crises for 17 countries' different crisis dates. Meager and Speckesser (2011) also find that the growth of productivity and wages depicts simultaneous movement for 25 countries during 1995-2009.

Additionally, the related literature investigates the connection between productivity and income inequality. However, the direction of the effect is from productivity to inequality as in technical changes based on skill biases. The skill-biased shifts suggest that innovations in production technology are against the low-skilled labor but not the skilled ones. As a result, there might be a wage gap between more and low-skilled laborers. More skilled laborers benefit from a rise in total factor productivity, but inequality worsens. Some of the other studies in the literature could be listed as Gries and Naudé (2018), Hornbeck and Moretti (2018), Maoz and Sarid (2021), Fuentes et al. (2014), Kampelmann and Rycx (2012), Caroli and Van Reenen (2001) and Leung (2001).

Some papers investigate the wage effects on productivity level or wage efficiency theory arguing that an increase in (real) wages stimulates productivity in labor. Furthermore, we account literature such as Basril et al. (2018), Stansbury and Summers (2018), Trpeski, et al. (2016), Cohn et al. (2015), Feldstein (2008), Zhang and Liu (2013), Herman (2020), Levine (1992), Akerlof and Yellen (1990), Spitz (1989), Levine (1989), Rebitzer (1987), Katz (1986).

The last group of literature and the focus of this study are related to the income inequality effects on labor productivity. However, scholars rarely investigate the unequal determination of the wages's effects on productivity level (Policardo et al. 2018: 2-3; Espoir and Ngepah 2020: 2612). Thus, papers studying on the connection between inequality and productivity are limited. Here, some examples of the most relevant studies and their findings are provided.

Freeman and Medoff (1984) are pioneering works analyzing the connection between inequality and productivity. They focus on union and nonunion workers for selected sectors in the United States. They reveal that wages are more homogeneous in unionized firms. Hence, inequality is low in these firms and this reduction might exist because of union workers' preferences and ideas about fairness. They also state that workers in the unionized companies are more productive. Productivity growth might be due to "industrial relations climate" and "more rational, professional management" (cited in Liu 2002: 455).

Liu (2002) examines the manufacturing industries following wage inequality and industrial productivity. Liu investigates the effects of wage inequality in the context of relative deprivation and efficient wage. Sen's index is a measurement for determining the relative deprivation and efficient wage levels. Sample countries are Taiwan (1979-1996) and South Korea (1993-1996). In the regression analysis of this study, the hourly output of labor is a dependent variable, and the Sen index for aggregated relative deprivation is one of the explanatory variables. The Gini index is inside the Sen's index, which is a tool for measuring the extent of economic deprivation in society as aggregate or relative deprivation. Regression results revealed that workers are reluctant when payments are less than they deserve. The coefficients of relative deprivation are highly negative for the two countries. However, this result is not consistent with the literature in which efficient wage affects industrial productivity. He also found that relative deprivation and efficient wages are more important than wage inequality.

Kim and Sakamoto (2008) examine the American manufacturing industry ranging from 1979 to 1996. They use productivity data for 72 manufacturing industries and the Gini index. Other variables are real capital stock, material cost spent, and the number of workers. Their results are not proof of the skill-biased technological-change argument and they find that wage inequality negatively affects productivity. The Gini coefficient is – 0.15 and statistically valid for the second model. 1% increase in inequality causes 0.15 % decrease in productivity on average. The general interpretation of the coefficients is that the relationship between inequality and productivity is significantly negative if the model has fixed effects for industry and year.

Mahy et al. (2009) evaluate the effect of inequality on productivity level with wage dispersion for Belgium. Wage dispersion is a measure for wage differentials between similar workers. They calculate conditional wage inequality using Winter-Ebmer and Zweimüller (1999) methodology. They also use age, education, sex, gross income, working hours, and occupation and sector employers, number of employees, wage bargaining. The results of ordinary least squares demonstrate that there is a relationship between wage dispersion and productivity. Namely, a small wage dispersion might be detrimental to productivity. Moderate wage increases are beneficial to firm performance. Hibbs and Locking (2000) also analyze wage dispersion for Sweden from 1960 through 1980. Model's variables are hourly wage distribution and value-added per worker. They reveal that a reduction in interindustry wage differences might lead to a productivity improvement.

DiPietro (2014) uses productivity growth as a dependent variable, an average annual Gini coefficient from 2000 to 2010. Six control variables are included in the model such as the level of economic development, the amount of human capital, the size of the private sector, wage flexibility, and government waste. He employs regression analysis to examine the cross-country data. Unfortunately, he does not mention which countries are included in his work. He reports that the coefficient of Gini is approximately -0.78 if there is only one explanatory variable. If control variables are included to the models one by one, this value becomes – 0.99.

Policardo et al. (2018) investigate that wage inequality and labor productivity for 34 Organisation for Economic Co-operation and Development (OECD) countries. They use generalized methods of moments for the period of 1995-2007. Labor productivity per hour worked is used as the dependent variable; the Gini index is employed as the independent variable. Control variables are Gross Domestic Product per capita, fertility, life expectancy, annual hours per worker, and total employed population rate. They find that wage inequality harms labor productivity. The coefficient of this effect is about -0.06 in which a one-dollar increase in the inequality index causes 6 cent decrease in labor productivity.

Britton and Propper (2016) investigate the impact of teacher pay on school productivity in England. They collect cross-sectional data from more than 3000 schools and around 200.000 teachers . Variables used in the analysis are school performance and the wage gap. Moreover, the school efficiency is the added value of the school measures by national tests. They depict that the teachers respond to the low payments. An unforeseen 10% change in the wage gap worsens school performance by 2%. A larger wage gap between formal payments and non-labor market wage levels reduces the school productivity.

Espoir and Ngepah (2020) examine the effects of income inequality on total factor productivity based on location and distance for South Africa. For this purpose, they apply the spatial econometric technique and use municipal panel data from 1995 to 2015. Their findings conclude that there are positive spatial interactions in the effects of income inequality on total factor productivity. It means that there is a neighboring effect among the municipalities. They further reveal that the impact of income inequality on productivity is negative for the direct effect and it is positive for the indirect effect. Municipalities (with high inequality) transfer jobs, investments, and skilled labor to municipalities (with medium income inequality and high-income opportunities).

Da Silveira and Lima (2021) investigate the endogenous macroeconomic fluctuations with the effects of wage inequality. They employ the frequency distribution at the micro-dynamic base. Their results demonstrate that the labor productivity is changeable across workers depending on different levels of wages. They also provide empirical evidence for the endogeneity of labor productivity and the persistency of wage inequality.

3. DATA AND METHODOLOGY

This chapter gives information about the the data and the methodology

3.1. Data

Finding sufficient and powerful data is usually difficult if one would like to examine unequal income distribution. This is because some indicators could be available for restricted periods and countries. Missing observations might be notable. Collecting two variables (labor productivity and income inequality) from the same source contributes to the availability and robustness of the data. In this vein, Statistics of the European Commission (Eurostat) has various and qualified data related to income inequality and labor productivity. Inequality of income is the measurement of the S80/S20 income quintile share ratio. This ratio is the household income ratio obtained by dividing the top 20% by the bottom 20%. The second variable is labor productivity calculated per person employed and hour worked (EU27_2020=100). The logarithmic transformations of variables as *log_inc* and *log_pro* representing income inequality and labor productivity are employed for the analysis, respectively.

The sample period ranges from 2005 to 2019 for 31 countries in Europe. These countries are Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, and United Kingdom.

3.2. Methodology

This section provides information concerning the co-integration approach used to estimate the long-run coefficients in the study. The cross-sectional dependence and the stationary tests are necessary before the estimation of the long-run coefficients. The methodology of Frees (1995) to test the existence of the cross-sectional dependence is explained below:

$$R_{AVE} = \binom{n}{2}^{-1} \sum_{i<j} r_{ij} \tag{1}$$

where represents the Spearman rank correlation coefficient between the ith and jth units. The cross-section dependence tests of Pesaran (2004), Frees (1995), and Friedman (1937) are initially used in this study. However, test statistics come up with some contradictory results. Pesaran and Friedman tests confirm the null hypothesis suggesting that there exists no cross sectional dependence across units in the panel but Frees test does not. Therefore, we conclude that there is cross-sectional dependence across units according to DeHoyos and Sarafidis (2006: 494).

Having found that there is a cross-sectional dependence among units, we should use the unit root test allowing cross-sectional dependence Because of the unbalanced panel data, we can apply Pesaran's (2007) methodology. Pesaran developed the cross-sectionally augmented Dickey-Fuller (CADF) statistics. In the process, he calculated a general panel unit root statistics by using unit root statistics of each cross-section in a panel data (Koçbulut and Altıntaş 2016: 16). CIPS is a general test statistic for unit root (Pesaran 2007: 276):

$$CIPS(N,T) = N^{-1} \sum_{i=1}^{N} t_i(N,T) \tag{2}$$

where $t_i(N,T)$is the CADF statistic for the i. cross-section unit based on CADF regression ($\Delta y_{it} = \alpha_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_i \Delta \bar{y}_t + e_{i,t}$) (Pesaran 2007: 269). The null hypothesis has the statement of non-stationary series ($H_0: b_i = 0$ for all i) (Pesaran 2007: 268-69).

Because of the connected units, the existence of a small sample (micro panel), and having different integration levels for series, the panel auto-regressive distributed lag (ARDL) model is the most suitable for the long-run coefficient estimation. Deviations from long-run equilibrium are significant than short-run equilibrium. Therefore, the long-run relationship between variables and unbiasedness are favorable in the estimation process (Granger 1986: 213). We run the ARDL approach under the integration level of I(1) for dependent variable and I(0) for explanatory variable (Pesaran et al. 2001: 315). The fundamental model of ARDL (p_x, p_y) with respect of dependent and independent variables (Ditzen 2018: 6):

$$y_{i,t} = \sum_{I=1}^{p_y} \lambda_{I,i}\, y_{i,t-I} + \sum_{I=0}^{p_x} \beta_{I,i}\, x_{i,t-I} + e_{i,t} \tag{3}$$

where the lag length of y and x are p_x,p_y, respectively. If we calculate the coefficients of long-run β and average group (Ditzen 2018: 6):

$$\theta_i = \frac{\sum_{I=0}^{p_x} \beta_{I,i}}{1-\sum_{I=1}^{p_y} \lambda_{I,i}}, \qquad \bar{\theta}_{MG} = \sum_{i=1}^{N} \theta_i \tag{4}$$

In calculating θ_i and $\bar{\theta}_{MG}$ values, Chudik et al. (2016) suggest the cross-sectionally augmented ARDL (CS-ARDL) and the cross-sectionally augmented distributed lag (CS-DL) estimators. CS-ARDL technique calculates the long-run coefficients from short-run coefficients and includes cross-sectional average. Extended version of equation 3 is explained below (Ditzen 2018: 9):

$$y_{i,t} = \sum_{I=1}^{p_y} \lambda_{I,i}\, y_{i,t-I} + \sum_{I=0}^{p_x} \beta_{I,i}\, x_{i,t-I} + \sum_{I=0}^{p_T} \gamma'_{I,i}\, \bar{z}_{i,t-I} + e_{i,t} \tag{5}$$

where $\bar{z}_{t-I}$ consists of $\bar{y}_{i,t-I}$ and $\bar{x}_{i,t-I}$. $e_{i,t}$. is serially uncorrelated process across for all i. Adding the cross-sectional mean to the formula can remove cross-sectional dependence in the errors. Thus, the estimation fulfills the validity criteria (Erülgen et al. 2020: 9).

4. ANALYSIS

At the beginning of the analysis, a related cross-sectional dependence test is necessary to check the possible correlations among units. The null hypothesis consists of a cross-sectional independence argument (DeHoyos and Sarafidis 2006: 492). The test statistics of Frees is 5.589 and critical values are 0.2828 (0.10), 0.3826 (0.05), and 0.5811 (0.01). Because Frees' statistics are bigger than critical values, we reject the null hypothesis suggesting the cross-sectional independence. After that, we implement cross-sectional augmented Dickey Fuller method. The results are reported in Table 1.

Table 1. Stationary Test Results

cons.	cons. & trend	lag	cons.	cons. & trend	lag
	log_pro			log_inc	
0.873	0.837	0	-2.828 **	-2.135 **	0
- 0.300	- 0.449	1	-1.690 **	1.394	1
	Δlog_pro			Δlog_inc	
-7.310 ***	-4.674***	0	-12.002 ***	-10.302 ***	0
-4.243 ***	-1.067	1	-3.510 ***	-0.594	1

Note: ***, **, and * represent % 1, % 5 and % 10 respectively.

When we examine the stationary levels of the variables, log_pro and log_inc are integrated I(1) and I(0), respectively. Furthermore, cross-sectional dependent variables necessitate the CS-ARDL approach. Table 2 reports the analysis results with different lag lengths.

Table 2. Long-Run Coefficients

Coef.	z stat.	Lag(s)	F stat.	R-sqr.	CD-stat. (p)
-0.1302 (0.0616)	-2.12**	0	1.53***	0.21	0.32 (0.750)
-0.1609 (0.0764)	-2.11**	1	1.58***	0.33	-0.63 (0.532)
-0.1475 (0.052)	-1.73 *	2	1.68***	0.46	-0.67 (0.503)
-0.1258 (0.0895)	1.40	3	1.55***	0.57	-0.15 (0.884)

Note: ***, **, and * represent % 1, % 5 and % 10 respectively. Standard errors in parentheses.

All models are statistically valid according to the F statistics. The most suitable model is ARDL (1,1). One lagged values of variables have a long-run relationship. There is a negative relationship between wage inequality and labor productivity. 1 % increase in wage inequality results in a 0.16 % decrease in labor productivity. Wage inequality explains 33% of productivity declines. This percentage is high if we consider only one explanatory variable. Moreover, the CD test is a post-estimation test for validity and shows that errors are cross-sectionally independent.

In the analysis, the income quintile share ratio is used as an inequality indicator while other studies employ the Gini index. Furthermore, this study estimates the long-run coefficients, other studies make general estimations without any differentiation. Considering the results, the coefficient is close to the 0.15 value of Kim and Sakamoto (2008), smaller than 0.78 and 0.99 of DiPietro (2014). In addition, it is bigger than the 0.06 and 0.02 of Policardo et al. (2018) in absolute values. Our findings are consistent with the theory and the empirical studies except for DiPietro. Briefly, we should pay attention to the dynamics between wage and productivity for a stable improvement in labor productivity in the long run. In addition, control variables can clarify the impact of wage inequality on the productivity level.

5. CONCLUSION

Improving productivity and diminishing inequal income/wage distribution require long-term policy and strategies. Therefore, this paper investigates the long-term parameters. An inequal income ratio negatively affects labor productivity. Unfair wage determination is one of the sources of income inequality. Various factors such as gender discrimination, subjective assessment, and regional differences might be other sources. Therefore, penetrating or reducing the inequalities probably depict long-run rather than short-run solutions.

In this respect, sustainable development goals determine reductions in inequalities as one of the three urgent goals. Three goals are “fight inequality & injustice,” “end extreme poverty,” and “fix climate change” from 2015 to 2030. Alternative policies at local, regional, territorial, and global levels might be more beneficial and powerful for this purpose. Furthermore, policy recommendations may include some measures on how to reduce inequality. Remuneration determination criteria based on skills, abilities, education, and the perception of the right wage level can help to prevent wage inequality. In addition, the government can detect and control whether the companies determine fair wages.

Cross-sectional dependency and data size restrict possible analysis techniques. A larger and more comprehensive data set will allow for more appropriate analysis. Further research can analyze time series data with different methods considering structural breaks and nonlinear dynamics. Researchers might extent and develop cross-sectional or panel models with spatial effects.

Funding

This work was not funded by any institution.

REFERENCES

- AFONSO, H., LAFLEUR, M. and ALARCÓN, D. (2015). Concepts of Inequality: Development Issues No. 1. *Development Strategy and Policy Analysis Unit in Development Policy and Analysis Division*, New York, USA.

- AGHION, P., BACCHETTA, P., RANCIERE, R., and ROGOFF, K. (2009). Exchange rate volatility and productivity growth: The role of financial development. *Journal of monetary economics*, *56*(4), 494-513.

- AKERLOF, G. A. (1984). Gift Exchange and Efficiency-Wage Theory: Four Views, *American Economic Review*, *74*, 79-83

- AKERLOF, G. A. and YELLEN, J. L. (1990). The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics*, *105*(2), 255-283.

- BASRIL, N. M., KARIM, Z. A., ISMAIL, R., and SULAIMAN, N. (2018). The Effect of Wages and Industry-Specific Variables on Productivity of Manufacturing Industry in Malaysia: A Dynamic Heterogeneous Panel Evidence. *International Journal of Economics & Management*, *12*(2), 379-391.

- BRITTON, J. and PROPPER, C. (2016). Teachers pay and school productivity: Exploiting wage regulation, *Journal of Public Economics. 133*(2016), 75-89.

- BRUNDTLAND COMMISSION. (1987). *Our Common Future, Chapter 2: Towards Sustainable Development*. World Commission on Environment and Development (WCED). Geneva: United Nation.

- CAROLI, E. and VAN REENEN, J. (2001). Skill-biased organizational change? Evidence from a panel of British and French establishments. *The Quarterly Journal of Economics*, *116*(4), 1449-1492.

- CHOUDHRY, M. T. (2009). 1 Determinants of Labor Productivity: An Empirical Investigation of Productivity Divergence.

- http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.614.8306

- CHUDIK, A., MOHADDES, K., PESARAN, M. H., and RAISSI, M. (2016). *Long-run effects in large heterogeneous panel data models with cross-sectionally correlated errors*. Essays in Honor of man Ullah (Advances in Econometrics, Vol. 36), Emerald Group Publishing Limited, Bingley, 85-135. https://doi.org/10.1108/S0731-905320160000036013

- COHN, A., FEHR, E., and GOETTE, L. (2015). Fair wages and effort provision: Combining evidence from a choice experiment and a field experiment. *Management Science*, *61*(8), 1777-1794.

- DA SILVEIRA, J.J. and LIMA, G.T. (2021). Wage inequality as a source of endogenous macroeconomic fluctuations, *Structural Change and Economic Dynamics. 56*(2021), 35-52.

- DE HOYOS, R. E. and SARAFIDIS, V. (2006). Testing for cross-sectional dependence in panel-data models. *The Stata Journal*, *6*(4), 482-496.

- DIPIETRO, R. (2014). Productivity growth and income inequality. *Journal of Economics and Development Studies*, *2*(3), 01-08.

- DITZEN, J. (2018). Estimating long run effects in models with cross-sectional dependence using xtdcce2. Centre for Energy Economics Research and Policy (CEERP), *Working Paper No. 7*.

- ERÜLGEN, A., RJOUB, H. and ADALIER, A. (2020). Bank Characteristics Effect on Capital Structure: Evidence from PMG and CS-ARDL. *Journal of Risk and Financial Management*, *13*(12), 1-13.

- ESPOIR, D. K. and NGEPAH, N. (2021). The effects of inequality on total factor productivity across districts in South Africa: a spatial econometric analysis. *GeoJournal, 86*(6), 2607-2638.

- EUROSTAT, (2021). Products Datasets: Inequality of income distribution S80/S20 income quintile share ratio - EU-SILC and ECHP surveys, https://ec.europa.eu/eurostat/web/products-datasets/-/ilc_pns4, [Date Accessed: 30/04/2021].

- EUROSTAT (2021). Products Datasets: Labour productivity per person employed and hour worked (EU27_2020=100), https://ec.europa.eu/eurostat/web/products-datasets/-/tesem160, [Date Accessed: 30/04/2021].
- EZCURRA, R., PASCUAL, P. and RAPÚN, M. (2007). Spatial inequality in productivity in the European union: sectoral and regional factors. *International Regional Science Review*, *30*(4), 384-407.
- FELDSTEIN, M. (2008). Did wages reflect growth in productivity? *Journal of Policy Modeling*, *30*(4), 591-594.
- FREEMAN, R.B., and MEDOFF, J. L. (1984). *What Do Unions Do?* New York: Basic Books
- FREES, E. W. (1995). Assessing cross-sectional correlation in panel data. *Journal of Econometrics*, *69*, 393–414.
- FRIEDMAN, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, *32*, 675–701.
- FUENTES, R., MISHRA, T., SCAVIA, J., and PARHI, M. (2014). On optimal long-term relationship between TFP, institutions, and income inequality under embodied technical progress. *Structural Change and Economic Dynamics*, *31*, 89-100.
- GRANGER, C. J. (1986). Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics*, *48*(3), 213-228
- GRIES, T. and NAUDÉ, W. (2018). Artificial intelligence, jobs, inequality, and productivity: Does aggregate demand matter? *Discussion Paper Series IZA DP No. 12005*
- HE, C., ZHOU, Y. and ZHU, S. (2017). Firm dynamics, institutional context, and regional inequality of productivity in China. *Geographical Review*, *107*(2), 296-316.
- HERMAN, E. (2020). Labour Productivity and Wages in the Romanian Manufacturing Sector, *Procedia Manufacturing. 46*(2020), 313-321.
- HIBBS Jr, D.A. and LOCKING, H. (2000). Wage dispersion and productive efficiency: Evidence for Sweden. *Journal of Labor Economics*, *18*(4), 755-782.
- HORNBECK, R., and MORETTI, E. (2018). Who benefits from productivity growth? Direct and indirect effects of local TFP growth on wages, rents, and inequality. National Bureau of Economic Research. *Working Papers No. w24661*.
- ISAKSSON, A., and NG, T. H. (2006). Determinants of productivity: cross-country analysis and country case studies. *Research and Statistics Branch*, UNIDO.
- ISAKSSON, A. (2007). Determinants of total factor productivity: a literature review. *Research and Statistics Branch*, UNIDO.
- ISLAM, N. (2008). Determinants of productivity across countries: an exploratory analysis. *The Journal of Developing Areas*, 201-242.
- KAMPELMANN, S., and RYCX, F. (2012). Are occupations paid what they are worth? An econometric study of occupational wage inequality and productivity. *De Economist*, *160*(3), 257-287.
- KATZ, L.F. (1986). Efficiency wage theories: A partial evaluation. NBER macroeconomics annual, *Working Papers Series No 1*, 235-276.
- KHAN, M., LUINTEL, K. B., and THEODORIDIS, K. (2011). How robust is the R&D-productivity relationship? Evidence from OECD countries. *Economic Research Working Paper*, (Vol. 1). WIPO.
- KIM, C., and SAKAMOTO, A. (2008). Does inequality increase productivity? Evidence from US manufacturing industries, 1979 to 1996. *Work and Occupations*, *35*(1), 85-114.
- KOÇBULUT, Ö. and ALTINTAŞ, H. (2016). İkiz açıklar ve Feldstein-Horioka Hipotezi: OECD ülkeleri üzerine yatay kesit bağımlılığı altında yapısal kırılmalı panel eşbütünleşme analizi. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, (48), 145-174.

- KOSE, M. A., PRASAD, E. S., and TERRONES, M. E. (2009). Does openness to international financial flows raise productivity growth? *Journal of International Money and Finance*, *28*(4), 554-580.

- KRUGMAN, P. (1994). Defining and measuring productivity, Organisation for Economic Co-operation, and Development website, https://www.oecd.org/sdd/productivity-stats/40526851.pdf, [Date Accessed: 23/11/2021].

- LEVINE, D.I. (1989). Cohesiveness, productivity, and wage dispersion. *Journal of Economic Behavior & Organization*, *15*(2), 237-255.

- LEVINE, D.I. (1992). Can wage increases pay for themselves? Tests with a productive function. *The Economic Journal*, *102*(414), 1102-1115.

- LEUNG, C.K.Y. (2001). Productivity growth, increasing income inequality and social insurance: the case of China? Journal of Economic Behavior & Organization, *46*(4), 395-408.

- LOKO, M.B., and DIOUF, M. A. (2009). *Revisiting the Determinants of Productivity Growth-What's new?* International Monetary Fund.

- LIU, J. (2002). Does wage inequality affect labor productivity? Some evidence from manufacturing industries of Taiwan and South Korea, *Asia Pacific Management Review*, *7*(4), 449-76.

- MAHY, B., RYCX, F., and VOLRAL, M. (2011). Wage dispersion and firm productivity in different working environments. *British Journal of Industrial Relations*, *49*(3), 460-485.

- MAOZ, Y.D., and SARID, A. (2021). Sectoral Heterogeneity, Income Inequality and Productivity Dynamics. *The Journal of Income Distribution*.

- MEAGER, N., and SPECKESSER, S. (2011). Wages, productivity, and employment: A review of theory and international data. European Employment Observatory Thematic expert ad-hoc paper, *Institute for Employment Studies 1*, 1-73.

- MEIDNER, R. and REHN, G. (1952). Wages Policy under Full Employment. London: William Hodge; 1st Edition.

- PAUL, P. (2020, March). Historical patterns of inequality and productivity around financial crises. *Federal Reserve Bank of San Francisco Working Papers.*

- PESARAN, M.H. (2004). General diagnostic tests for cross section dependence in panels. University of Cambridge, *Cambridge Working Papers in Economics*, No. 0435.

- PESARAN, M.H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, *22*(2), 265-312.

- PESARAN, M.H., SHIN, Y., and SMITH, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, *16*(3), 289-326.

- POLICARDO, L., PUNZO, L. F. and CARRERA, E.J.S. (2018). Wage inequality and labor productivity in OECD countries, *Texto Para Discussão Discussion Paper*, *136*(2018).

- REBITZER, J.B. (1987). Unemployment, long-term employment relations, and productivity growth. *The Review of Economics and Statistics*, *69*(4), 627-635.

- SPITZ, J. (1989). An Empirical Test Efficiency Wage Theory. Graduate School of Business, Stanford University. *Working Paper No. 1053.*

- STANSBURY, A.M., & SUMMERS, L.H. (2018). Productivity and Pay: Is the link broken? (No. w24165). National Bureau of Economic Research. *Working Paper Series No.24165*.

- STIGLITZ, J.E. (1982). Alternative Theories of Wage Determination and Unemployment: The Efficient Wage Model. in *The Theory and Experience of Economic Development*. Eds. M. Gersovitz, C. F. Diaz-Alekandro, G. Ranis and M. R. Rosenzweig. pp. 78-106, London: George Allen & Unwin Ltd., First Edition, ISBN: 0-04-330323-4.

- STIGLITZ, J.E. (2013). *The Price of Inequality: How Today's Divided Society*. W.W. New York: Norton, First Edition, ISBN: 978-0-393-08869-4.

- THE GLOBAL GOALS, https://www.kureselamaclar.org/en/, [Date Accessed: 23/11/2021].

- TRPESKI, P., EFTIMOV, L., and CVETANOSKA, M. (2016). Labor productivity and real wages in Macedonia: An overview before and after the global economic crisis. *European Scientific Journal*. 12 (10), p.352-365.

- ZHANG, J. and LIU, X. (2013). The evolving pattern of the wage–labor productivity nexus in China: Evidence from manufacturing firm-level data, *Economic Systems*, *37*(2013), 354-36.

JAME

JOURNAL OF APPLIED MICROECONOMETRICS

HOLISTENCE publications