The Effect of Scholarship Programs on Performance and Success: The case of Roma students in Macedonia

By

Nadire Redjepi

Submitted to

Central European University

Department of Economics

In partial fulfillment of the requirements for the degree of Master of Arts in Economic Policy in Global Markets

Supervisor: Professor Andrea Weber

Budapest, Hungary

2017

ABSTRACT

This thesis evaluates the impact of the governments' scholarship program on the performance and success of the Roma students enrolled in secondary school in Macedonia. For the purpose of the analysis I used the collected database from the Ministry of Education for the period 2009-2016. Taking into consideration similar studies related to this topic, first, I used a linear regression Ordinary Least Squared analysis, with Fixed Effects and control variables to estimate the change of GPA from one year to another one. Second, to estimate the effects of the program I used Fuzzy Regression Discontinuity analysis to show the average effect of the treatment group at the cut-off. The results showed no effect at the local level, meaning that the scholarship program did not affect the Final GPA of students around the threshold.

Keywords: impact, Roma, linear regression, Fuzzy Regression Discontinuity, change of GPA, treatment group, local level, threshold

ACKNOWLEDGMENTS

First of all, I would like to express my biggest gratitude to my friend and tutor of Roma Graduate Preparation Program, Sergiu Burlacu who supported me fully throughout the process of writing the thesis. Thank you Sergiu for devoting special attention, time, and efforts on helping me with the design of the methodology, without your help I would not be able to do it.

Second, I thank to my supervisors Andrea Weber and Gabor Kezdi for their constant feedback and critical arguments; especially to Gabor Kezdi who supported me fully during my pregnancy and helped me to improve my work.

I also express my gratitude to my husband Elvis Shakjiri and my 14 months old daughter Aleda Shakjiri for their patience and support in the process of writing the thesis. Thank you for your understanding, patience, and love, I really need it.

Most importantly, I thank to my parents who made me decide to come to CEU. Thank you for constantly being with me and supporting me all the time.

Lastly, I would like to thank to my CEU friends, to the Economics department, to Roma Graduation Program, and to Romalitico for their strong influence into my academic life.

ABSTRACT	i
ACKNOWLEDGMENTS	ii
List of tables	iv
List of Figures	iv
List of Abbreviations	v
Introduction	1
Chapter 1: Education trends among Roma population in Macedonia	5
Chapter 2: Evaluation Study	8
2.1. Background of the evaluated program	8
Chapter 3: Data and Methodology	11
3.1. Data Description and Statistics	11
3.2. Methodology	23
Chapter 4: Ordinary Least Squared (OLS) Analysis	25
4.1. Ordinary Least Squares (OLS) Model and Results	25
Chapter 5: Regression Discontinuity Analysis	32
5.1. Regression Discontinuity Basics	32
5.2. Regression Discontinuity Model and Results	34
Conclusion and Policy Recommendation	40
Appendices	43
Bibliography	

Table of contents

List of tables

Table 1: Definition of the new variables	12
Table 2: Descriptive Statistics	14
Table 3 : Cross Tabulation Category - Previous AY got the stipend	18
Table 4: Proportion of students across different categories in the AY of 2009/10	20
Table 5: Proportion of students across different categories in the AY of 2010/11	21
Table 6: Proportion of students across different categories in the AY of 2011/12, 2012/13, and	
2013/14	21
Table 7: Proportion of students across different categories in the AY of 2014/15 and 2015/16	22
Table 8: Estimated results for change in GPA by academic year	27
Table 9: Estimated results for change in GPA using OLS and FE for 2010-2013	28
Table 10: Estimated results for change in GPA using FE with control variables (2010-2013)	30
Table 11: Estimated Results of Local Linear and Quadratic Regression	38
Table 12 : Estimated results of 2SLS Regression	39

List of Figures

Figure 1: Yearly Averages for Initial GPA (when students applied for scholarship) and Final	
GPA (at the end of the AY)	16
Figure 2: Histogram Change in GPA	17
Figure 3: Mean values of Final GPA for each student	35
Figure 4: Nonparametric Density Function, Scholarship Beneficiary-GPA	36
Figure 5 : Estimated Linear, Non-Linear, and Non-Parametric models	37

List of Abbreviations

OLS	Ordinary Least Squared
RD	Regression Discontinuity
AY	Academic Year
GPA	Grade Point Average
RCT	Randomized Control Trials
IV	Instrumental Variable
FE	Fixed Effects
LHS	Left Hand Side
RHS	Right Hand Side
LATE	Local Average Treatment Effect
ITTE	Intent To Treatment Effect
ATE	Average Treatment Effect

2SLS Two Stage Least Squared

Introduction

Education is one of the most important and valuable assets that an individual can possess in its lifetime process. It is a long-term investment process which by itself brings the highest returns to both, individual and collective level. As many other investment forms (financial, intellectual, business), the investment in education has also its own costs and benefits that are visible after a certain time period. In practice, the individual benefits and returns are seen in various forms such as: success, assets (money), knowledge, profession, critical thinking, well-being, social and economic position, and other forms. These forms have their own positive impacts that strongly influence the collective level and bring changes to the whole society. Moreover, the higher the quantity and quality of education, the higher is the social, economic, and political performance of a country or community.

Historically, education is seen as one of the key elements for alleviating poverty and for producing economic prosperity and growth in the country. The government can powerfully influence the economic growth of its national economy by: (1) *designing policies on "building up a larger stock of productive assets and human skills", and (2) by "increasing the productivity of these assets, skills and countries natural resources"* (World Bank, 1980). Low-income countries trying to catch up with the other developed countries tend to have high demand for well-educated and knowledgeable people. The quality of education is highly important because as people get educated they generate more income, spend more, and influence the overall GDP of the country.

In order to achieve a higher educational success, many scientists and researchers try to identify the main factors that determine individual's educational achievement. Stephen J. Caldas & Carl Bankston (1997), found a strong correlation between individual family poverty status and

individual minority race (Caldas & Bankston, 1997). Heyneman and Loxely found that depending on the country at which children attend the primary school, the educational achievements vary from low to high-income countries. According to their evidence, the influence of the socioeconomic status varies significantly with the national economic development of the country (Heyneman & Loxley, 1989).

In addition to the minority race and family poverty status, many other factors directly or indirectly influence the educational achievement of individuals. Segregation is one of the important pigments that directly influence the educational achievement of disadvantaged groups. The segregation in education is often done in the rural and poor areas where pupils do not have the same equal access as those living in urban and reach areas. The Organization for Economic Cooperation and Development, in the report for Equity and Quality in Education, noted that: "*To ensure equity and quality across education systems, funding strategies should: guarantee access to quality early childhood education and care (ECEC), especially for disadvantaged families; use funding strategies, such as weighted funding formula, that take into consideration that the instructional costs of disadvantaged students may be higher" (OECD, 2012).*

Throughout the centuries socially disadvantaged groups such as the Roma population fight against poverty issues (Ivanov & Kagin, 2014). These issues are closely related to the poor socioeconomic factors such as the lack of educational background, lack of specific knowledge and skills, segregation, discrimination, family status and many other problems which result with mismatch in the labor market, causing the well-known unemployment phenomena.

In the last decade different countries and institutions are trying to solve the socioeconomic and political issues of Roma by designing specific policies and programs for them. According to the

National Decade and Strategy for Roma (2014-2020), the Ministry for Labor and Social Affairs in Macedonia has been implementing various programs related to the social inclusion of Roma in five different fields: Employment, Education, Housing, Health, and Culture (Ministry for Labor and Socail Affairs, 2014). The evaluation of these programs is highly demanded because their impact has been hardly measured in practice. For example: In Macedonia the Ministry of Education and Science and the Roma Education Fund are the two major institutions which provide financial support to Roma students from different categories for almost a decade. The program is limited with evidence and research on the impact of the scholarship awards.

Therefore, with regard to the scholarship programs in my thesis, I am interested to analyze the effect of the financial award on the performance of the Roma students enrolled in the secondary education. More precisely, in my thesis I will analyze the impact of the governments' scholarship program on the performance and success of Roma students in Macedonia. For the purpose of my analysis I will use the collected data set from the Ministry of Education and Science– Macedonia, for the period between the academic year of 2009/10 and the academic year of 2015/16. The data set has two main parts:

- Data on students that applied for scholarship and received the scholarship
- And data on students that applied for scholarship and did not receive the scholarship

The thesis aims to evaluate the impact of the governments' scholarship program on the performance and success of the Roma students in Macedonia. Moreover, the goal of the thesis is:

- To evaluate and give policy recommendation on the scholarship program for Roma
- To identify the factors that play crucial role in the educational achievements of Roma

The structure of the thesis is organized in five chapters as stated below:

In the first chapter I give a short overview with some statistical information on educational trends among Roma in Macedonia. The aim of this chapter is to familiarize the reader with the basic issues and trends appearing in the Roma communities in Macedonia. The second chapter is devoted on the evaluation study, explaining the background information of the evaluated program. More precisely, in this chapter I provide basic information on the program's goals, selection criteria and some definitions of the categories. The third chapter is about the descriptive statistics and methodology. In this chapter I start the discussion by looking the average values and performances of the observations. The fourth and the fifth chapters are the most important ones; they provide all the findings and discussions. In the fourth chapter I show the results of Ordinary Least Squared (OLS) and pooled-regression, while in the next chapter I discuss the results of the Regression Discontinuity (RD) model with a causal interpretation. In the last chapter I summarize the findings and provide policy recommendations.

Chapter 1: Education trends among Roma population in Macedonia

According to the last census from 2002 in Macedonia there are 2,022,547 inhabitants. The official number of Roma is declared to be 53,879 or 2,66% of the total population (Statistical office, 2002). The unofficial data shows that the number is much higher and varies from 134.000 to 260.000 Roma (Council of Europe). It is important to note that in the last 15 years no census was held because of political disagreements and as a result there is no recent data on the population and households.

The poor condition of Roma in the field of education could be noted in many reports (Roma Education Fund, 2007) (Helsinki Committee, 2014). The most common factors for the poor education of Roma are: low awareness among the parents about the need to educate their children, material conditions, language barriers and discrimination. Nowadays the Roma still face with the same problems but not in the same scale as 15 years ago.

In order to improve the situation of Roma many Balkan and EU countries¹ joined the Decade of Roma Inclusion 2005-2015 which aim was to eliminate discrimination and close the gaps between Roma and non-Roma in the four fields: housing, education, employment and health. After introducing the Roma Decade 2005-2015, the situation of Roma in the field of education has significantly improved compared with the other priorities such as employment, housing and health (Helsinki Committee, 2014). The number of students enrolled in primary, secondary and tertiary education has increased. One of the reasons for improving the situation in the field of education are the affirmative action policies that the country introduced (Roma Education Fund, 2007). With

¹ Macedonia, Albania, Hungary, Bosnia and Hezegovina, Romania, Serbia, Montenegro, Czech Republic, Spain, Slovakia, Bulgaria and Croatia.

the affirmative action for secondary education the Ministry of education decreased the criteria's for Roma students in the process of application for secondary school. With the criteria Roma students who have 10% less points than the number of points stipulated in the call for application, are eligible to apply (Ministry of Education, Ministry of Education , 2015). The aim of this measure is to help Roma pupils to enroll in better schools where the competition among students with higher Grade Point Average (GPA) decreases the chances of Roma to be enrolled. The affirmative action for tertiary education is the reserved quotas for minorities.

Starting from 2008/09 the country introduced a new Law on secondary education. The secondary school in Macedonia became mandatory for everybody and the country provides free books and transportation to the place where students decide to study. The mandatory schooling in Macedonia for primary education is nine years (age 6 - 14) and for secondary education is four years (age 15-18).

Academic year	Roma students
2006/2007	1279
2007/2008	1304
2008/2009	1461
2009/2010	1628
2010/2011	1636
2011/2012	1654
2012/2013	1717
2013/2014	1688
2014/2015	1560
2015/2016	1420
(0)	

source: (State statistical office, 2017)

One of the policy measures the government introduced for increasing the number of students in secondary education is the scholarships for Roma students.

Academic year	Number of scholarships
2009/10	444
2010/11	611
2011/12	591
2012/13	593
2013/14	623
2014/15	671
2015/16	671
2016/17	670
2017/18	752

source: (Ministry of Education, 2017)

Today in Macedonia the issue of inclusion of Roma is targeted through the Strategy of Roma 2014-2020 and the Integration of Roma 2020 (within the Regional Cooperation Council), and represents a regional initiative in the Western Balkans and Turkey. This is considered as a post-regional.

Chapter 2: Evaluation Study

The second chapter provides basic information on the program's goals, selection criteria, and definition of the categories. The goal of this chapter is to familiarize the reader with the background information of the evaluated program. In this chapter, the reader can find information on the way the program has evolved and defined new categories with different amounts of the scholarship awards.

2.1. Background of the evaluated program

Since the academic year of 2009/2010, the Ministry of Education and Science together with the support of the Roma Education Fund started to provide financial aid to Roma students from various categories enrolled in the secondary education. One of the important goals of the program was to increase the enrollment rate among Roma students. In order to achieve the stated purpose the program successfully increased the proportion of the distributed scholarships, and in the academic year of 2015/2016 it reached to 671 scholarships.

In the first years, when the program started to provide financial aid to Roma students, from the Academic Year (AY) of 2009/2010 till the AY of 2013/2014, it supported Roma students with higher GPAs from two categories, while from the AY of 2014/15 it started supporting students with lower GPAs. Therefore, the definition of the categories and the criteria differed from year to year.

In the first academic year (2009/10) all applicants with the GPA of 3.00 and above were defined as "scholarship recipients", and all applicants below 3.00 were defined as "rejected applicants".

In the following four academic years (2010/11, 2011/12, 2012/13, and 2013/14) all applicants with the GPA between 3.50 and 5.00 were defined as "scholarship recipients in the first category", all applicants with GPA between 3.00 and 3.49 were defined as "scholarship recipients in the second category", and all applicants below 3.00 were "rejected applicants".

The most obvious changes in the definition of categories appeared from the AY of 2014/15 in which the program defined five different categories. Moreover, in the AY of 2014/15 and 2015/16, the program supported Roma students from five different categories of students enrolled in first, second, third, and fourth year in public or private secondary education. Based on the category to which the applicants belonged to, the criteria and the amount of scholarship varied significantly.

The scholarship awards were defined in the following way:

First category: Stipends for talented Roma students

The Roma students which succeeded to keep the highest GPA in the previous academic year, (between 4.50 and 5.00) belonged to the first category of scholarship recipients. All students enrolled in to the first, second, third, and fourth year which achieved high GPA were in this category. The monthly amount of scholarship for this category was 2.200 MKD or 35 Euro per month (Ministry of Education, Calls: Ministry of Education and Science, 2016).

Second category: Students with GPA between 3.50 and 4.49

The students with the GPA of 3.50 and 4.49 were categorized into the second category, and were awarded with 1.500 MKD or 25 Euro per month (Ministry of Education, Calls: Ministry of Education and Science, 2016).

Third category: Students with GPA between 3.00 and 3.49

The students with the GPA of 3.00 and 3.49 were into the third category, and were awarded with 1.000 MKD or 16 Euro per month (Ministry of Education, Calls: Ministry of Education and Science, 2016).

Fourth category: Scholarship for students with special educational needs

The students in this category were awarded with 1.000 MKD or 16 Euro per month (Ministry of Education, Calls: Ministry of Education and Science, 2016).

Fifth category: Motivation scholarship for students with GPA between 2.00 and 2.99

The students in this category were awarded with 600 MKD or 10 Euro per month. Students enrolled in the first and second year had a priority in receiving of this type of scholarship (Ministry of Education, Calls: Ministry of Education and Science, 2016).

- Eligibility criteria for Roma students belonging to the first, second, third, and fourth category which achieved at least 3.00 GPA (Ministry of Education, 2016):
- 1. To be enrolled in first, second, third, or fourth year of studies into private or public secondary education;
- 2. To not repeat the same year;
- 3. To not receive scholarship from other institutions;
- 4. To belong to the Roma community and;
- 5. To be citizens of the Republic of Macedonia.

Chapter 3: Data and Methodology

The third chapter is devoted on describing the data and methodology. In this chapter, I show the descriptive statistics together with the definition of the variables. I start the analysis by looking at the average values of the Initial GPA and Final GPA and discuss their trends. Additionally, by using cross-tabulations I analyze the potential relationship between Initial GPA and the binary variable on whether the student received scholarship in the previous AY or not. Thus, in this chapter I provide the statistical information on the way the program has evolved from year to year. Furthermore, I explain the methodology together with its initial assumptions and its causes.

3.1. Data Description and Statistics

To estimate the impact of the scholarship program on performance and success of the Roma students, I used the collected data set from the Ministry of Education and Science from Macedonia. The Ministry of Education and Science, throughout the years collected different information on individual applicants from which I selected the most important and common variables that play a crucial role in the selection criteria.

In Table 2, we can see the descriptive statistics of different variables. In total, we have 5507 observations for seven consecutive years, starting from the academic year of 2009/10 and ending in the AY of 2015/16. The original data set contains information on individual's name, surname, gender, enrolled year, category, GPA, city, and whether the person received a scholarship or not. For the purpose of my analysis, I created few dummy variables that take the value of 1 or 0, and three more outcome variables that measure the performance and success of the students. The definition of the variables is presented in the following table:

Table 1: Definition of the new variables

Variable	Definition				
Previous AY got the	Refers to the applicants that received a scholarship in the				
stipend	previous AY. If the applicant received a scholarship in the				
	previous AY, takes the value of 1 and 0 otherwise.				
Next AY got the stipend	Refers to the applicants that received the award in the next AY. If				
	the applicant received a scholarship in the following AY, takes				
	the value of 1 and 0 otherwise. The aim of this variable is to keep				
	track for all scholarship beneficiaries throughout the years				
	because according to the observations some students do not				
	frequently apply for each AY.				
Applied next AY	Refers to all applicants that applied in the following AY. The aim				
	of this variable is to measure the frequency of applications that				
	come from the same person no matter whether that person				
	previously received or did not received the scholarship.				
Times received	It counts how many times an applicant received an award in the				
scholarship in the	previous AY.				
previous AY					
Times received	It counts how many times an applicant received a scholarship in				
scholarship in the next	the following AY.				
AY					
Times applied in the next	It counts how many times an applicant applied for the next				
years	academic years.				
Total number of	It counts how many times an applicant received an award in total,				
received scholarship	no matter whether the person received the award in the previous				
	AY or the next AY.				
Completed previous AY	This variable is an outcome variable, which measures the				
	performance of the applicants. It refers to the students that are				
	enrolled in the second, third, fourth and fifth year. (Example: if a				

	student is enrolled in the third year, it takes a value of 1 meaning					
	that the student completed the previous AY)					
Final GPA (at the end of	It is another outcome variable which measures the success of the					
the AY)	applicants by looking the recorded GPA when the person applies					
	for a scholarship.					
Delta GPA	It calculates the difference in the GPA from year to year. The					
	change in the GPA is another outcome variable identifying the					
	performance and success of students.					

In addition to the newly created variables, I used Stata to code the student's name and surname, the city, and the school for which Stata gave an ID number to each individual, city, and school. According to the descriptive statistics, in Table 2, the maximum coded number for the Student ID shows that there are in total 3258 applicants applying various times across the observed years. In total there are 30 cities and 138 schools; 5507 observations for Gender (0 refers to boys, and 1 refers to girls); 5474 observations for the grade of the applicants; 5485 observations for the GPA; 5507 observations for the category to which the candidates belonged to; and 5507 observations on whether the student received the scholarship or not. We can notice that the total number of observations is not same for all the variables. There are 33 missing values for enrolled year, and 22 missing values for the GPA of the students.

Similarly, depending on the definition of the newly created variables the total number of observations is not the same. In total, there are 3056 observations on whether the student received a scholarship in the previous AY or not; 4191 observations on whether the student received a scholarship in the next AY or not; 4193 on whether the student applied in the next AY or not; 3056

observation on how many times an applicant received an award in the previous AY; 4191 observations on times received a scholarship in the next AY and times applied in the next years. However, the statistics on completed previous AY, GPA at the end of the AY, and delta GPA is more interesting for the analysis. We have 2035 missing observations for completed previous AY, implying that in total 2035 students were enrolled in the first grade. Similarly, there are 3258 missing observations on the GPA at the end of the AY, and 3259 missing observations on delta GPA.

Furthermore, according to the descriptive statistics, all of the applicants applying for the scholarship program were enrolled in four years of studies while only three candidates appeared to be enrolled in five years of studies. The average GPA for the observed period of the AY 2009/10 till the AY of 2015/16 is 3.5, the minimum GPA is 1.94, and the maximum GPA is 5.

Variable	Obs	Mean	Std. Dev.	Min	Max
Student ID	5507	1633.609	934.5296	1	3258
Academic Year	5507	2011.884	1.996823	2009	2015
City ID	5507	17.68458	7.294181	1	30
School ID	5507	64.89395	35.30353	1	138
Gender	5507	0.5075359	0.4999886	0	1
Grade	5474	2.129887	1.06148	1	5
GPA	5485	3.583621	0.8081404	1.94	5
Category	5507	1.214999	1.109141	0	5
Scholarship received	5507	0.7261667	0.445965	0	1
Previous AY got the stipend	3056	0.6276178	0.4835186	0	1
Next AY got the stipend	4191	0.4600334	0.4984596	0	1

Table 2: Descriptive Statistics

Applied next AY	4193	0.5363701	0.4987349	0	1
Times received scholarship in the previous AY	3056	0.9626963	0.8925492	0	3
Times received scholarship in the next AY	4191	0.7122405	0.8862463	0	3
Times applied in the next years	4191	0.8153185	0.9122681	0	3
Total number of received scholarship	5507	1.802433	1.323394	0	4
Completed previous AY	3472	1	0	1	1
GPA at the end of AY	2249	3.775447	0.7021427	1.94	5
Delta GPA	2248	-0.0466161	0.6178436	-2.88	2.86

However, if we look at the average GPA for each year, separately we can notice that there are ups and downs in the mean values. Figure 1, below, presents the average values of GPA for each AY when students apply for a scholarship (Initial GPA) and when they finish the AY (Final GPA).

Figure 1: Yearly Averages for Initial GPA (when students applied for scholarship) and Final GPA (at the end of the AY)



If we look at the blue line, representing the GPA of students when they apply for a scholarship, we can notice that the average GPA has an increasing tendency that starts from 3.61 in the AY of 2009/10 and reaches its highest point at 4.05 in the AY of 2012/13. After the AY of 2012/13, the average GPA when students apply for scholarship slightly decreases reaching to 3.73 points. The decline in the average GPA is one of the indicators showing that the scholarship program in the last two academic years targeted different categories that include students with lower GPA. In contrast to the last two academic years (2014/15 and 2015/16), in the previous years the scholarship program targeted students with at least 3.00 GPA. (See definition of categories in Chapter 2). However, when we compare the first academic year of 2009/10 with the academic year of 2014/15 we can notice that the average value of GPA in the academic year of 2014/15 is higher by 0.12 points.

Similarly, the GPA at the end of the observed AY has an increasing tendency in the first three years, reaching to maximum 4 points in the AY of 2011/12. When we compare the initial GPA with the final one, we can notice that there is an inverse relationship between the two variables. The inverse relationship between the two variables appears between the AY of 2011/12 and the AY of 2012/13 showing that the performance of some students worsened compared to the previous academic years when they applied for the scholarship. After the academic year of 2012/13, we can see a similar pattern between the two variables which seems to have a growing tendency for the following years.

In order to check the performance of the students in more details, in Figure 2 below, we can see the distribution of the change in GPA for each year. According to the graph, in the first three years, students have increased their GPA on average, while in the following three academic years (from 2013/14 till 2015/16) the performance of the students in GPA has worsened on average. The academic year of 2013/14 implies that most of the students have decreased their GPA by 0.25 points on average while in the previous academic year of 2012/13 most of the students have increased their GPA by 0.13 points on average.



Figure 2: Histogram Change in GPA

Based on the practical experiences, it is natural to have heterogeneity across the mean values of GPA in different years. For instance, there might be many factors that influence the performance of the students. Depending on the program, students might have difficulties to adopt and evolve with their studies. In some programs, students might easily get higher grades while in others not. Some studies might require previous knowledge on math, physics, and chemistry while others might not require that. In general, at this point we cannot say that the scholarship program had a positive or negative impact on the performance of the students. We need a deeper investigation with additional variables that would explain the performance and success of the students in more details.

To see if there is some relationship between the performance of GPA and whether the person received a scholarship in the previous academic year, I used a cross tabulation table. Table 3, below, presents the cross tabulation results between category and whether the person received a scholarship in the previous AY or not.

Previous AY got the stipend			Categor	·y			
	0	1	2	3	4	5	Total
No	405	275	328	67	17	46	1,138
Yes	218	878	576	146	55	45	1,918
Missing data	885	1,061	344	91	24	46	2,451
Total	1,508	2,214	1,248	304	96	137	5,507

Table 3 : Cross	Tabulation	Category -	Previous	AY got	the stipend
-----------------	-------------------	-------------------	-----------------	--------	-------------

Zero category refers to the rejected applicants, counting total 1,508 students that belonged to this group. 218 students from the rejected applicants were previous scholarship recipients, meaning that out of 1,508 students 218 students changed their status and became non-beneficiaries while 405 students were not beneficiaries in the previous AY. The missing data refers to students that were enrolled in the first year of their studies, and to students that applied in the first AY of the implementation of the scholarship program 2009/10. The students enrolled in the first year previously were enrolled in primary education, therefore we have missing data on them. According to the definition of the missing data, most of the rejected students or 885 students were either enrolled in the first year of their studies, or were applicants in the first year of the implementation of the program in 2009/10.

Depending on the definition of the categories the first category represents the students with the highest GPA in different years. In total 2,214 students belonged to the first category of which 878 students were previously scholarship recipients, meaning that out of 2,214 students 878 students successfully kept their status as "beneficiary." Additionally, 275 students that were not previously scholarship recipients succeeded to enter into the first category and receive the award. According to the definition of missing data, 1,061 students were either enrolled in the first year of their studies or were applicants in the first year of the program 2009/10.

Similarly, the other categories (category 2, 3, 4, and 5) refer to various groups that are defined differently. The number of scholarship recipients in the fourth and fifth category is much lower compared to the previous categories because both of them were recently defined and implemented. However, out of the total 5,507 applicants, 1,918 students were beneficiaries in the previous AY and 1,138 students were not beneficiaries in the previous year.

In addition to the cross tabulation table between the two variables discussed above, with the aim to see the evolution of the program, I also checked the simple statistics for to the proportion of students applying for scholarship in each AY.

According to the results of Table 4, below, out of the total 917 applicants, 472 students were rejected from the scholarship while 445 students were beneficiaries of the program. It seems that in the first year the program has poorly defined its target because around 51% from the total pool of applicants were rejected from the scholarship (zero category). This result indicates that most of the rejected applicants did not fulfill the criteria because their GPA was below 3.00.

Table 4: Proportion of students across different categories in the AY of 2009/10

Scholarship	Cate		
Beneficiary	0	1	Total
0 1	472 0	0 445	472 445
Total	472	445	917

In the AY of 2010/11, the total number of applicants decreased by 22% compared to the previous AY (from 917 students to 707 students). According to Table 5, below, 94 students did not receive scholarship and 613 students received the scholarship. Compared to the previous AY of 2009/10, in the AY of 2010/11, the program successfully increased the number of beneficiaries from 445 to 613 beneficiaries. From the results, it seems that the program succeeded to improve and define its new targets.

AY	2010/11				
Scholarship		Category	,		
Beneficiary	0	1	2	Total	
0 1	94 0	0 412	0 201	94 613	
TOTAL	94	412	201	707	

Table 5: Proportion of students across different categories in the AY of 2010/11

In the AY of 2011/12, the total number of applicants increased by 16.1% compared to the previous academic year (from 707 students to 821 students). According to additional tabulation (See, A 3: Cross Tabulation in AY of 2011/12), out of the total 591 beneficiaries, 368 students were beneficiaries from the previous AY, meaning that students kept continuity in performing good and remained in the category of beneficiaries throughout the years. In Table 6, below, we can notice similar patterns in each AY.

Table 6: Proportion of	students across different	categories in the A	AY of 2011/12 ,	, 2012/13, and
	2013	3/14		

AY	2011/12		2012/13				201	13/14				
Scholarship	C	ategoi	ry		(Categor	y		(Catego	ry	
Beneficiary	0	1	2	Total	0	1	2	Total	0	1	2	Total
0	230	0	0	230	284	0	0	284	103	0	0	103
1	0	395	196	591	0	400	200	600	0	411	213	624
TOTAL	230	395	196	821	284	400	200	884	103	411	213	727

Similarly, in the following two years, the program succeeded to increase the number of beneficiaries from 591 to 600 in 2012/13 and to 624 in 2013/14. Starting from the AY of 2014/15, the program defined new categories to increase the number of the scholarship recipients. In the AY of 2014/15, out of the total 732 applicants, 183 students previously did not receive an award while 332 students were beneficiaries from the previous AY of 2013/14 (See Table 7). Similarly, in the AY of 2015/16, out of total 719 applicants, 161 students previously did not receive a scholarship while 297 students were beneficiaries in the previous AY of 2014/15. (See, A 7: Cross Tabulation in the AY of 2015/16)

Scholarshi	i Category					TOTAL by AY		
p Beneficiary	0	1	2	3	4	5	2014/1 5	2015/1 6
0	325	0	0	0	0	0	191	134
1	0	151	438	304	96	137	541	585
TOTAL	325	151	438	304	96	137	732	719

|--|

From the results, we can notice that from the beginning of the program till the AY of 2013/14 the first category had the highest proportion of students, implying that the program successfully kept the significant proportion of students with higher GPA between 3.50 and 5. However, with the new definition of categories in the AY of 2014/15 and 2015/16 the proportion of the beneficiaries shifted to the second category students with the GPA between 3.50 and 4.49. The histograms confirm that the change in GPA is shifting differently. More precisely, in most of the years, the

distribution of the change in GPA is shifting to the left side meaning that most of the students have decreased their GPA, while in some of the years they are close to the normal distribution indicating that some students increased their GDP and some have reduced it.

3.2. Methodology

Most of the empirical studies focusing on analyzing the effect of the financial aid on students' performance face with various methodological issues. These issues are mostly caused by the lack of Randomized Control Trials (RCT) in the implementation process of the programs. By lacking the RCT, it is always difficult to design a proper methodology that would take care of underestimation and overestimation of the effect of the program.

According to the pre-selected criteria, the program obviously did not use randomization as preliminary methodology in the distribution process of the scholarships, but indeed it strongly used self-selection during the implementation process. The lack of RCT in the program implies that we have to find alternative ways to measure the true effect of the scholarship program on the performance and success of the students. For example: the study published by the Higher Education Quality Council of Ontario tried to control for omitted variable biases through using Regression Discontinuity analysis with Instrumental Variable (IV) (Fuzzy RD with IV). The study analyzed the relationship between entrance financial aid awards and success in university in Ontario (Dooley, Payne, & Robb, 2013). First, by using OLS regression it estimated the relationship without controlling for omitted variable biases and second, by using Regression Discontinuity analysis it controlled for omitted variables. In addition, Dynarsky (2002) used a quasi-experimental methodology to determine the exogenous sources and find the true effect between the schooling costs and schooling decision. The author measured the effect of the Hope

Scholarship program by using Difference-in-Difference analysis, by comparing the relative changes in the attendance rates before and after the program (Dynarski, 2002).

Taking into consideration similar studies related to this topic, first, I used a linear regression OLS model with Fixed Effects (FE) and control variables to estimate the change of GPA from one year to another one. Second, to estimate the effects of the program I used Regression Discontinuity analysis to show the average effect of treated and control that are around the threshold. The following variables are the key measures in determining the outcome of the program:

- Final GPA
- Change in GPA

The data set is consisted of two parts representing the total number of observations for seven academic years (2009/10, 2010/11, 2011/12, 2012/13, 2013/14, 2014/15, and 2015/16). In total, there are:

- 3999 observations for the scholarship recipients and;
- 1508 observations for the rejected applicants

My expectation from this analysis is to see positive relationship between the financial aid and success of the students. According to my view, in the last decade the number of Roma students enrolled in higher education has increased significantly. I assume that this increase was partially due to the fact that many Roma faced with different financial difficulties and the scholarship program had an important role in determining their needs. However, there are other factors which in the same time influenced the behavior of Roma students such as: different public and private policies targeting Roma inclusion and education, Roma NGO's, positive spillover effects, etc.

Chapter 4: Ordinary Least Squared (OLS) Analysis

The fourth chapter discusses the estimations of the OLS analysis. In this chapter, I show the basic OLS equation used in estimating the coefficients and I discuss them. Additionally, by using robust and clustered standard errors I show the difference in estimated coefficients controlling for the error term.

4.1. Ordinary Least Squares (OLS) Model and Results

I used the following equation as a baseline model to estimate the change in GPA from year to year:

Model: $\Delta \Upsilon_i \Box = \alpha + \beta x_i \Box + \gamma c_i \Box + u_i \Box$

equation (1)

Δ GPAit= α + β Scholarship Beneficiary it + γ Academic Year t + δ Control it + uit

 $\Delta \mathbf{y}_i \Box = \mathbf{y}_i \Box - \mathbf{y}_i \Box_{-1}$

Y: Change in GPA

i: individual

t:time

 α : intercept (constant), measures the average change in Y in the control group

 β : Slope, showing the effect of the program on the treated or it shows the average difference between then change in Y in the treatment group versus the control group.

x: Dummy variable on Scholarship Beneficiary, if a person is in the treatment group it is equal to one, and if a person is in the control group it is equal to zero.

c: Control variables (GPA, Gender, Category, Enrolled Year, and Total number of received scholarship)

uit: Error term

In Table 8, we can see the estimated coefficient results on delta GPA for all academic years excluding the base year of 2009/10. The Left Hand Side (LHS) variable is change in GPA and the Right Hand Side (RHS) "scholarship beneficiary" is the main explanatory variable of interest. In this regression we compare the change in GPA for treated and control group. The non-beneficiaries belong to the control group. When we look at the results, the significance level of the coefficients is the first thing to notice. Second, throughout the years we have different numbers of total observations, and third, the R-squared explaining the variation in Y is very low. The interpretation of the coefficients is the following:

The constant, shows the average change in GPA for the non-beneficiaries (or those belonging to the zero category). Throughout the years the non-beneficiaries negatively changed their GPA, meaning that they worsened their performance on average. It is noticeable that the coefficients on the non-beneficiaries are negatively performing. For example: In the academic year of 2015/16, the GPA of non-beneficiaries decreased by 0.62 points on average. Similarly, in the academic year of 2011/12, the GPA of non-beneficiaries decreased by 0.29 points.

In contrast to the non-beneficiaries, the coefficients on scholarship beneficiaries are positive. These coefficients show the average difference between the changes in GPA in the treatment group versus the control group. In the academic year of 2015/16, the GPA of the scholarship beneficiaries increased more than the GPA of non-beneficiaries, by 0.63 points on average. Similarly, in the academic year of 2011/12, the GPA of the scholarship beneficiaries increased more than the average GPA of non-beneficiaries, by 0.38 points.

Academic Year	2010/11	2011/12	2012/13	2013/14	2014/15	2015/16	
VARIABLES	delta_gp	delta_gp	delta_gp	delta_gp	delta_gp	delta_gp	
	а	а	а	а	а	а	
	(1)	(2)	(3)	(4)	(5)	(6)	
Scholarship_Beneficar	0.17	0.383***	0.440***	0.470***	0.540***	0.639***	
У							
	(0.1590)	(0.0731)	(0.0600)	(0.1220)	(0.0772)	(0.1300)	
Constant	-0.111	-	-	-	-	-	
		0.290***	0.213***	0.715***	0.597***	0.622***	
	(0.1520)	(0.0663)	(0.0538)	(0.1190)	(0.0695)	(0.1260)	
Observations	279	409	429	390	384	357	
R-squared	0.003	0.068	0.12	0.044	0.118	0.07	
Robust standard errors in pa	conthococ **	* n<0.01 **	n<0.05 * n<	0.1			

Table 8: Estimated results for change in GPA by academic year

Robust standard errors in parentheses p<0.01. p<0.05, r p<∪.1

According to the results, we cannot say much about the effect of the program. Indeed, we must improve the regression model because the change in the GPA might be affected by other factors that are not in the regression, meaning that the one explanatory variable in the RHS might bias the estimates. Moreover, if we want to capture the heterogeneity in the individual characteristics we need to control for the individual FE. Under the FE, the error term of the individual and the constant should not be correlated with others (Torres-Reyna, 2007).

Since the program defined various categories as it evolved, I decided to restrict my sample to the years when it used the same selection criteria. I used the same equation (1) in a pooled-regression using a subsample of 1591 observations for the period between the academic year of 2010/11 and 2013/14. The missing values on Final GPA and enrolled year were dropped from the analysis.

With attention to the individual FE, I compared the regression coefficients that control and that do not control for the individual heterogeneity. For an illustration, let look at Table 9 which compares the estimated results for the change in GPA with and without the FE.

The first regression does not control for the individual FE while the second one controls for the individual FE. It is noticeable that most of the results have changed and remained significant. The main coefficient of interest, scholarship beneficiary, increased in the second model and remained positive. In the first model, we compare two students with different treatment status, one which receives the scholarship and one which does not receive the scholarship. The GPA is expected to increase by 0.39 points for students that received the scholarship. Similarly, in the second model, we compare two different years for the same individual. In one of the years the person receives a scholarship while in the other year the person does not receive a scholarship. The GPA is expected to increase by 0.52 points in the year when the person receives a scholarship.

Furthermore, the two regressions use different standard errors which control for the error term. In the second model, by using clustered standard errors we control for the error term at the individual level. The second model is statistically more convenient one because it controls for the heterogeneity and selection bias of the individuals.

VARIABLES	delta_gpa	delta_gpa
	(1)	(2)
Scholarship_Beneficary	0.398***	0.528***
	(0.0427)	(0.0898)
2011.AcademicYear	0.0224	0.139*
	(0.0515)	(0.0772)
2012.AcademicYear	0.145***	0.336***
	(0.0496)	(0.0890)
2013.AcademicYear	-0.323***	-0.0824
	(0.0513)	(0.1000)
Constant	-0.325***	-0.584***
	(0.0584)	(0.1090)
Standard Errors	Robust	Cluster

 Table 9: Estimated results for change in GPA using OLS and FE for 2010-2013

Individual FE	No	Yes
Observations	1,507	1,507
R-squared	0.1170	0.1240
Number of Student_id		965

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

Although we control for the FE, we still have to add more explanatory variables that would capture the effect of the program. In Table 10, we can see the estimated results for the change in GPA with some control variables.

According to the regression results, by using FE, the main coefficients of interest on scholarship beneficiaries are not significant. In the first model, if we compare people with the same GPA but different treatment status, the treated group (scholarship beneficiary) decreased their GPA more than the control group, by 0.053 points on average. However, this result is not significant.

In contrast to the scholarship beneficiary, the estimates of GPA are all significant. If we compare students with the same category but with different GPA, the students with a point higher GPA increased their performance by 1.3 points on average. Similarly, in the other regression models we can see the similar pattern on the coefficient of GPA, meaning that category and grade are not influencing the estimates. The R-squared is also the same across models, implying that around 44% of the variations in the change of GPA are explained by the regression results.

VARIABLES	delta_gpa	delta_gpa	delta_gpa
	(1)	(2)	(3)
Scholarship_Beneficary	-0.0532	0.0575	0.0576
	(0.0805)	(0.1910)	(0.1920)
GPA	1.301***	1.267***	1.267***
	(0.0705)	(0.0865)	(0.0866)
Category		-0.0536	-0.0537
		(0.0834)	(0.0834)
Enrolledyear			-0.0135
			(0.0460)
2011.AcademicYear	-0.055	-0.0477	-0.0342
	(0.0635)	(0.0652)	(0.0824)
2012.AcademicYear	-0.0923	-0.0749	-0.048
	(0.0770)	(0.0844)	(0.1370)
2013.AcademicYear	-0.253***	-0.246***	-0.206
	(0.0819)	(0.0839)	(0.1960)
Constant	-4.876***	-4.787***	-4.770***
	(0.2450)	(0.2720)	(0.3040)
Individual FE	Yes	Yes	Yes
Observations	1,507	1,507	1,507
R-squared	0.438	0.439	0.439
Number of Student_id	965	965	965

Table 10: Estimated results for change in GPA using FE with control variables (2010-2013)

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

The estimated results of the pooled regression with individual FE indicate that by using only one explanatory variable we get positive and significant results, while by adding more variables they are getting close to nil. In practice various factors other than the scholarship award have important role in determining the success of the students at the end of the year. We have limited set of control variables that measure the correlation between the scholarship beneficiary and change in GPA. Therefore, according to the results, it seems that the program did not have influence on the change of GPA of students. Depending on the type of the students the effect of the program might be

bigger, smaller, positive, and negative or with zero effect. At this point, we cannot argue much because we need further investigation to see what is true and what is not true. In the next section, I use Fuzzy Regression Discontinuity design to investigate the Local Average Treatment Effect where I estimate the effect of the program for those students that are around the threshold.

Chapter 5: Regression Discontinuity Analysis

Chapter five is devoted on explaining the RD analysis and its results. It starts by explaining the theoretical background of RD analysis including the main assumptions and limitations of the model. In addition, it also explains the baseline equation of the regression analysis and provides all the findings and results. Chapter five is the main body of the thesis, because based on the results of the RD model I interpret the results with causality and estimate the effect of the program only for the students that are around the threshold (Local Average Treatment Effect), and provide policy recommendations to the government authorities.

5.1. Regression Discontinuity Basics

In the absence of randomized assignment, identifying the average treatment effect of an intervention raises challenges due to self-selection into the program. The limitations are usually caused by the "self-selection criteria" which bias the final results. It is always difficult to identify the true effect in absence of observable variables that influence the main outcome of an intervention. Even if we measure everything there can still be issues driven by self-selection, such as lack of common-support where the individuals in the treatment group are fundamentally different from the control group. For example, in the case of identifying the true effect of the scholarship program on the performance and success of the Roma students there is no qualitative data on the level of motivation of the students, their goals, activity and presence in the class, cognitive capacities, enrollment in university and other socio-economic background characteristics that play an important role in interpreting the results. However, in practice there are various methods used in trying to identify the causal effects of an intervention or program.

Regression Discontinuity (RD) is one of the powerful methods to identify the causal effects of an intervention on the treatment group at the threshold. The RD framework works in non-experimental researches where it compares people that are right at the threshold. In the absence of RCT, it assumes that people around the threshold are similar to each other. When comparing people around the threshold if close to each other there outcomes should not be different in the absence of the program; if a significant jump in average outcomes at the threshold then that must be the result of the program (Kezdi, 2016). However, the main assumption of RD is that there is discontinuous jump treatment as a result of an intervention.

We distinguish two types of RD designs:

- Sharp RD: the probability of treatment of the two sides (treated and control) is equal to one, meaning that all people assigned to the treatment group received the treatment and those assigned to the control group did not receive the treatment (analogy to perfect compliance) (Kezdi, 2016). Sharp RD identifies the Average Treatment Effect at the threshold, often called as "marginal" treatment effect.
- Fuzzy RD: the probability of treatment jumps from zero to something less than one, meaning that some of the people assigned to the treatment group received the treatment and some did not receive the treatment (analogy to imperfect compliance) (Kezdi, 2016). Fuzzy RD identifies the Intent to Treatment Effect (ITTE) for those at the threshold. It can also identify the Local Average Treatment Effect (LATE) for compliers, but not the "marginal" Average Treatment Effect (ATE) for all at the threshold.

5.2. Regression Discontinuity Model and Results

In my analyses I use the notation of *potential outcome approach* for causal inference where the treatment status is determined by the eligibility criteria. Let E be the binary variable denoting the eligibility status with E=1 for those eligible to receive the scholarship and E=0 otherwise, and (Y_1, Y_0) be the potential outcomes that would be realized if students eligible and not eligible. In our case, we are interested to see the potential outcome of Final GPA (at the end of the AY), as a measure of success for the treated and control group. The causal effect of eligibility on increasing or decreasing the Final GPA is defined as the difference between these outcomes $Y_1 - Y_0$ but in practice we observe only one of these outcomes, either Y_1 (if student eligible for scholarship Y_1 will be realized and Y_0 will be a counterfactual outcome) or Y_0 . Therefore, what we identify in an RD framework is Average Treatment Effect (ATE), average of Yi(1) – Yi(0) over a subpopulation.

Moreover, to find the Intent-to-Treatment Effect of the scholarship program, I used Linear Effects Model, Non-Linear Quadratic Model, Non-Parametric Linear Regression Model where:

 $Y_i = B_0 + B_1 X_i + B_2 Z_i + B_3 X_i Z_i + e_i$ equation (2)

 $Y_i = Final GPA$

 $X_i = Initial GPA$

 Z_i = Dummy on Eligibility (E=1 if GPA > 3 and GPA=3, or E=0 if GPA< 3)

X_iZ_i= Interaction term between Initial GPA and Eligibility

Taking into consideration that the treatment status of the scholarship beneficiaries is determined by the criteria of GPA, I decided to restrict my sample to the period of 2009 and 2013 when the minimum criteria of GPA is 3.00, same for all years (see Chapter 3, Background of the evaluated program). According to this criterion, all students with the GPA below 3.00 are in the control group (not eligible) and students with the GPA of 3.00 and above are in the treatment group (eligible). According to the obtained data, not all students with the GPA of 3.00 and above received the scholarship meaning that we have imperfect compliance (some received the treatment and some did not receive the treatment). Moreover, based on the stated criteria I defined a cut-off point at the GPA of 3.00.

If we look at Figure 3, below, we can notice that we have certain number of students that apply for scholarship with a GPA below 3.00.





In Addition, Figure 4, below, clearly shows that we have a Fuzzy RD design where the probability of receiving the treatment jumps discontinuously at the cut-off point from 0 to something less than

1. Hence, it shows that the likelihood of receiving the scholarship jumps at the GPA of 3.00 from around 50% to something less than 100%.



Figure 4: Nonparametric Density Function, Scholarship Beneficiary-GPA

Hence, the key assumption of RD is that there is only discontinues jump for scholarship beneficiaries at the threshold, GPA=3.00. In the absence of the program it assumes that there will be continuity at the threshold, which cannot be tested. We can test whether there is discontinues jump at the outcome variable GPA at the end of the AY (Final GPA) for the treatment group. According to Figure 5, below, we can see that there is a discontinuous jump between the treated and control group indicating that the average effect of the program for the treated group around the threshold is negative with no significance level. From the graph, it seems that the students who did not receive the scholarship around the threshold, were highly motivated to get the scholarship in the next year, meaning that they worked harder to increase their GPA at the end of the year. This is one of the possible explanations that we cannot prove it. Another explanation is that since the minimum requirement for scholarship award is 3.00 students with slightly higher GPAs did

not put more efforts to increase their GPAs at the end of the year. Similarly, the non-linear quadrated fitting shows the same discontinuous negative jump for the treated group but with higher confidence intervals (the bluish gray color) confirming our statement about the motivation of the non-beneficiaries (control group). Additionally, according to the Non-parametric model (Local Linear regression), we can conclude that there is a small negative discontinuous jump for the treated group at the threshold.



Moreover, in Table 11, below, we can see the estimated results of Local Linear and Quadratic Regression model. In the first model (Local Linear Regression), the regression coefficient of the variable *eligible* is negative meaning that the effect of eliglibility on Final GPA is negative (-1.49 + 0.22 * 0 = -1.49) at the cut-off GPA=3.00. However, the positive result on the interaction term

(0.444) implies that if we try to extrapolate for higher initial GPAs for these guys the net effect of being eligible is positive but they do not really have a control group.

This result indicates that as we move further from the threshold of 3.00, the program might have positive effect on the students with higher GPA. Similarly, in the second model (Quadratic Regression) we get negative but not significant results with higher R-squered.

Model	(1)	(2)
VARIABLES	gpa1	gpa1
eligible	-1.459***	-3.113
	(0.4240)	(4.1100)
gpa0	0.227	-3.043
	(0.1590)	(3.2090)
gpa02		0.65
		(0.6270)
gpa0_eligible	0.444***	2.484
	(0.1600)	(3.2220)
gpa0_eligible2		-0.497
		(0.6280)
Constant	2.642***	6.707*
	(0.4160)	(4.0720)
Observations	1,927	1,927
R-squared	0.409	0.415

Table 11: Estimated	I Results of Local	Linear and O	Luadratic Regression
----------------------------	---------------------------	--------------	-----------------------------

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

Furthermore, since the model is analogues to imprefect compliance I estimated Fuzzy RD model by using Two Stage Least Squared (2SLS) regression model where in the first stage I used elibility as an instrumental variable (IV) for the treatment group, and in the second stage I estimated the Local Average Treatment Effect (LATE).

Variable	first	IV
treated		-0.4636
		0.0622
		0.0000
eligible	0.8036	
	0.0176	
	0.0000	
Ν	1927	1927

Table 12 : Estimated results of 2SLS Regression

According to the results in Table 12, above, in the first stage the coefficient value on eligibility is 0.8, indicating that 80% of the eligible students receive the treatment (this is similar to compliance). In the second stage, our main coefficient of interest on the instrumental variable is negative, indicating negative impact of treatment effect on the treated. However, this results are not significant implying that the on the local level the program did not have impact on the final GPA of the students. The estimated result at the local level is an expected and logical one, because the final GPA of the students is influenced by many other factors that are not observed by the program.

Conclusion and Policy Recommendation

The main goal of the thesis is to evaluate the impact of the governments' scholarship program on the performance and success of the Roma students in Macedonia. For the purpose of my analysis I used the collected data base from the Ministry of Education for the period 2009-2016. In the phase of the analysis I faced with methodological challenges in identifying the right model of analysis as the data collection was not uniformed across the years. Thus, due to the lack of RCT I used Fuzzy RD analysis to investigate the effect of the program at the local level.

The first part of the analysis, estimated by the regression coefficients of OLS model, imply that GPA is not the only factor influencing the performance of students at the end of the year. Depending on the type of their studies students might face difficulties in adapting and evolving to the new environment. It is expected to see different patterns for students that have higher GPAs. The GPA of 5.00 is the maximum and for these students we expect to see similar patterns from year to year, but in practice we know that some of the students successfully keep the highest GPA, some lower it more, and some less. So, depending on the type of the students the effect of the program might be bigger, smaller, positive, and negative or with zero effect.

To investigate the effect of the program and interpret the results with causality I used Regression Discontinuity Analysis. By using RD model, I estimated the effect for the students that are around the threshold of 3.00 GPA (Local Average Treatment Effect, and Intent to Treatment Effect).

According to the RD design, there is a discontinuous jump between the treated and control group indicating that the average effect of the program for the treated group around the threshold is negative with no significance level. One possible explanation for this result is that students who did not receive the scholarship around the threshold, were highly motivated to get the scholarship in the next year, meaning that they worked harder to increase their GPA at the end of the year. Another explanation is that since the minimum requirement for scholarship award is 3.00, students with slightly higher GPAs did not put more efforts on increasing their GPAs at the end of the year. So, maybe the program focused on producing quality by influencing the performance of those students with lower GPAs and not of the ones with high GPA, we do not know that. All these arguments sound logical in explaining the negative jump at the threshold but we cannot prove them.

Furthermore, the results of the IV estimates show that the program did not affect the GPA of the students that are around the threshold of GPA= 3.00. This means that at the local level the performance of the students was depending on other factors. However, the estimates with IV regression are true only for the students around the threshold and not for the others that are far from the threshold. Therefore, the estimation with IV has high internal validity and weak external validity.

Since the results show no effect at local level around the threshold, the program should take the following actions:

- Re-design the scholarship awards: as the program is evolving and growing, it should design two types of awards:
 - <u>Merit-based-awards:</u> should be designed for students with higher GPAs. Here the target should be focused on quality and not quantity meaning that best students should be awarded with higher awards.
 - 2. <u>Needs-based-awards:</u> designed for students with lower GPAs and for students with low economic background. The target here should be focused on quantity

meaning that the higher the number of needs-based-awards the higher is the likelihood to create quality through positive spillover.

- Collection of data: the program must improve its capacities in collecting the data base. It is important that each year the program is homogenously collecting and coding the information. The mistakes in the row data are causing many troubles when trying to analyze the outcomes of the program. Also, in addition to the basic information (gender, city, school, and name) it should collect data on other individual and social characteristics (enrollment in university, type of studies, social and economic background).
- Evaluation of the program: the program must critically evaluate its goals and achievements. At the end of each academic year it should have an internal evaluation, evaluating the process of distributing the awards and identifying the gaps and needs of the program. In addition, it should also evaluate the effect of the program and hire external evaluator to measure their impact.
- Create a networking community: one of the ways to keep in touch with the beneficiaries is to create a networking community where students can share valuable information related to their professional development careers and help to each other. Also, through networking they can provide feedback information on the design and on the program facilities.

Appendices

A 1: Cross Tabulation in AY of 2009/10

Previous	Cate		
AY got the stipend	0	1	Total
Missing data	472	445	917
Total	472	445	917

A 2: Cross Tabulation in AY of 2010/11

Previous				
AY got the stipend	0	1	2	Total
No	35	97	84	216
Yes	6	155	62	223
Missing data	53	160	55	268
Total	94	412	201	707

A 3: Cross Tabulation in AY of 2011/12

Previous AY				
got the stipend	0	1	2	Total
No	100	54	56	210
Yes	63	213	92	368
Missing data	67	128	48	243
Total	230	395	196	821

A 4: Cross Tabulation in the AY of 2012/13

Previous				
AY got the stipend	0	1	2	Total
No	102	51	61	214
Yes	67	213	90	370
Missing data	115	136	49	300
Total	284	400	200	884

A 5: Cross Tabulation in the AY of 2013/14

Previous		Category		
AY got the stipend	0	1	2	Total
No	33	57	64	154
Yes	15	225	88	328
Missing data	55	129	61	245
Total	103	411	213	727

A 6: Cross Tabulation in the AY of 2014/15

Previous		Category					
AY got	0	1	2	3	4	5	Total
the							
stipend							
No	74	10	36	34	7	22	183
Yes	52	33	131	63	29	24	332
Missing	65	31	58	28	14	21	217
data							
Total	191	74	225	125	50	67	732

Previous		Category					
AY got the stipend	0	1	2	3	4	5	Total
No	61	6	27	33	10	24	161
Yes	15	39	113	83	26	21	297
Missing data	58	32	73	63	10	25	261
Total	134	77	213	179	46	70	719

A 7: Cross Tabulation in the AY of 2015/16

A 8: Estimated results for change in GPA using OLS and FE for 2009-2015

VARIABLES	delta_gpa	delta_gpa
	(1)	(2)
Scholarship_Beneficary	0.453***	0.533***
	(0.0362)	(0.0653)
2011.AcademicYear	0.0296	0.138*
	(0.0516)	(0.0767)
2012.AcademicYear	0.153***	0.340***
	(0.0496)	(0.0883)
2013.AcademicYear	-0.323***	-0.102
	(0.0514)	(0.0997)
2014.AcademicYear	-0.152***	0.199*
	(0.0529)	(0.1100)
2015.AcademicYear	-0.0734	0.361***
	(0.0535)	(0.1170)
Constant	-0.376***	-0.667***
	(0.0548)	(0.1000)
Standard errors	Robust	Cluster
Individual FE	No	Yes
Observations	2,248	2,248
R-squared	0.114	0.143
Number of Student_id		1,337

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

VARIABLES	delta_gpa	delta_gpa	delta_gpa
	(1)	(2)	(3)
Scholarship_Beneficary	-0.0322	0.0416	0.0417
	(0.0549)	(0.1060)	(0.1060)
GPA	1.224***	1.196***	1.196***
	(0.0494)	(0.0582)	(0.0582)
Category		-0.0283	-0.0283
		(0.0333)	(0.0333)
Grade			-0.0292
			(0.0428)
2011.AcademicYear	-0.0451	-0.0388	-0.0095
	(0.0624)	(0.0634)	(0.0773)
2012.AcademicYear	-0.068	-0.054	0.00448
	(0.0734)	(0.0766)	(0.1210)
2013.AcademicYear	-0.248***	-0.243***	-0.155
	(0.0808)	(0.0818)	(0.1840)
2014.AcademicYear	0.094	0.127	0.244
	(0.0882)	(0.0984)	(0.2320)
2015.AcademicYear	0.258***	0.294***	0.441
	(0.0948)	(0.1050)	(0.2770)
Constant	-4.632***	-4.564***	-4.553***
	(0.1750)	(0.1860)	(0.1910)
Individual FE	Yes	Yes	Yes
Observations	2,248	2,248	2,247
R-squared	0.466	0.466	0.466
Number of Student_id	1,337	1,337	1,336

A 9: Estimated results for change in GPA using FE with control variables for the entire sample

Cluster standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.

A 10: Histogram Initial GPA



A 11: Histogram Final GPA



Bibliography

Caldas, S., & Bankston, C. (1997). Effect of School Population Socioeconomic Status on Individual Academic Achievement. *The Journal of Educational Research*, 269-277.

Council of Europe. (n.d.). Council of Europe. Retrieved from https://www.coe.int/en/web/portal/roma

- Dooley, M., Payne, A., & Robb, A. (2013). The Impact of Scholarships and Bursaries on Persistence and Academic Success in University. *The Higher Education Quality Council of Ontario*, 2.
- Dynarski, S. (2002). The Behavioral and Distributional Implications of Aid for College . *American Economic Review*, 92(2): 279-285.
- Gertler, P., Martinez, S., Premand, P., Rawlings, L., & Vermeersch, C. (2011). *Impact Evaluation in Practice.* Washington DC.: The World Bank.
- Helsinki Committee. (2014). Analysis between the implementation and the reality of the Strategy for Inclusion of Roma.
- Heyneman, S., & Loxley, W. (1989). The effect of primary school quality on academic achievment across twenty-nine high and low income countries. *American Journal of Sociology*, 1162-1194.
- Ivanov, A., & Kagin, J. (2014). *Roma poverty from a human development perspective*. Istanbul:UNDP: Roma Inclusion Working Papers.
- Kezdi, G. (2016). Regression Discontinuity, Program Evalluation Slides. Budapest, Hungary, Hungary.
- Ministry for Labor and Socail Affairs. (2014). *National Decade and Strategy on Roma (2014-2020).* Ministry for Labor and Social Affairs, Republic of Macedonia.
- Ministry of Education. (2015). *Ministry of Education*. Retrieved from http://www.mon.gov.mk/index.php/konkursi/569-konkurs-upis-sredno-15-16
- Ministry of Education. (2016, October 03). *Calls: Ministry of Education and Science*. Retrieved October 03, 2016, from http://www.mon.gov.mk/index.php/konkursi/1507-romi-2016-2017-konkursi

Ministry of Education. (2017). Ministry of Education. Retrieved from https://goo.gl/ovuALV

OECD. (2012). Equity and Quality in Education: Supporting Disadvantaged Students and Schools. OECD Publishing.

Roma Education Fund. (2007). Advancing Education of Roma in Macedonia.

State statistical office. (2017). *State statistical office*. Retrieved from http://www.stat.gov.mk/PublikaciiPoOblast.aspx?id=41&rbrObl=5 Statistical office. (2002). Census of population, households and dwellings in the Republic of Macedonia.

Torres-Reyna, O. (2007, December). Panel Data Analysis Fixed and Random Effects using Stata (v.4.2.). Princton, New Jersey, United States of America.

World Bank. (1980). World Development Report 1980. Washington D.C : The World Bank.