EDUCATION INSTITUTIONS AND SCHOOL OUTCOMES OF IMMIGRANTS: A CROSS-COUNTRY ANALYSIS

by

YAŞAR ERSAN

Submitted to the Graduate School of Arts and Social Sciences in partial fulfillment of the requirements for the degree Master of Arts

Sabanci University

July, 2013

EDUCATION INSTITUTIONS AND SCHOOL OUTCOMES OF IMMIGRANTS: A CROSS-COUNTRY ANALYSIS

APPROVED BY:

Abdurrahman B. Aydemir …….

(Thesis Supervisor)

Esra Durceylan Kaygusuz ……..

Deniz Yükseker ……

DATE OF APPROVAL: 22.07.2013

Acknowledgements

I would like to start by expressing my gratitude to my thesis supervisor, Associate Professor Abdurrahman B. Aydemir, for his valuable support and encouragement for me to complete this study. This thesis could not be materialized without his guidance and persistent help. Throughout the process, he gives me constructive comments and warm encouragement.

Special thanks to my thesis jury members, Assistant Professor Esra Durceylan Kaygusuz and Assistant Professor Deniz Yukseker for their meticulous comments and criticisms about my study.

I owe very important debt to my primary school teacher Gulnur Ayoz for her all generous support over the years. I could not be here without her guidance and enormous help.

All my current and previous teachers deserve my thanks for their contribution to my intellectual life.

I want to offer my gratitude to my family for their all love and support over the years.

I also thank to the ones that produce the life everyday.

Lastly, I would like to thank to "TÜBİTAK", The Scientific and Technological Research Council of Turkey, for their financial support by providing scholarship.

EDUCATION INSTITUTIONS AND SCHOOL OUTCOMES OF IMMIGRANTS: A CROSS-COUNTRY ANALYSIS

Yaşar Ersan

Economics, MA Thesis, 2013

Supervisor: Abdurrahman B. Aydemir

Keywords: Immigrant students, education institutions, preprimary education, primary school starting age, tracking

Abstract

This paper is an attempt to analyze the effects of education institutions on the achievement of immigrant students in 31 high immigration countries. The role of education institutions on immigrant population has not been examined systematically in the literature. This paper aims at filling out this gap by assessing the effect of institutional arrangements in the educational system of host countries. Using OLS procedures with the 2003, 2006 and 2009 PISA datasets, we show that achievement gaps are wider for the immigrant students in the host countries where the school starting age is late. Likewise, the expected duration of preprimary education is a key determinant on the scores of the immigrant students since it can increase the gap between the natives and immigrants if the immigrants are deprived of attending to preprimary education. The tracking system definitely results in the increase of inequality between immigrant and native students. We also figure out that the time spent in the educational system of a host country is a crucial determinant on the achievement of immigrant students and it affects their performance positively. Moreover, the language spoken at home and the age of arrival to a host country have also notable effect on the immigrant performance. In addition, country specific effects such as HDI, income inequality, education spending, pupil-teacher ratio and teacher salary rates are also important factors for immigrant educational outcomes. The inequality between the immigrant and native students is reproduced through the institutional arrangements that are designed without sufficient consideration.

EĞİTİM KURUMLARI VE GÖÇMENLERİN OKUL BAŞARISI: ÜLKELERARASI BİR ANALİZ

Yaşar Ersan

Ekonomi, Yüksek Lisans Tezi, 2013

Tez Danışmanı: Abdurrahman B. Aydemir

Anahtar Kelimeler: Göçmen öğrenciler, eğitim kurumları, ilköğretim, okul öncesi eğitim, ilköğretim başlangıç yaşı, yönlendirme (seçilim)

Özet

Bu çalışma 31 ülkede göçmenlerin eğitim başarısının eğitim kurumlarıyla ilişkisini analizine yönelik bir çabadır. Eğitim kurumlarının göçmen nüfusu üzerindeki rolü literatürde sistematik bir şekilde incelenmemiştir. Bu çalışma, ülkelerdeki eğitim düzenlemelerinin göçmenler üzerindeki etkisini araştırarak bu boşluğu doldurmaya çalışmaktadır. 2003, 2006 ve 2009 PISA veri setleri yardımı ile yapılan Sıradan En Küçük Kareler Yöntemi analizi , okula başlama yaşının geç olduğu ülkelerde göçmen ve yerli öğrenciler arasındaki farkın göçmen öğrenciler aleyhine artmakta olduğunu göstermektedir. Buna ek olarak, göçmen öğrenciler okul öncesi eğitimden yoksun kalırlarsa beklenen okul öncesi eğitim süresi göçmen ve yerli öğrenciler arasındaki farkı artırmaktadır. Yönlendirme ya da seçilim mekanizması kesin bir şekilde aradaki farkı göçmenler aleyhine artırmaktadır. Ayrıca, eğitim sistemi içinde geçirilen zaman süresi önemli bir faktör ve göçmen öğrencileri ekstra pozitif etkilemektedir. Evde konuşulan dil, ülkeye kaç yaşında gelindiği göçmen öğrencilerin başarısı üzerinde kayda değer bir etkiye sahiptir. HDI seviyesi, gelir dağılımındaki eşitsizlik, öğrenci-öğretmen oranı, eğitim harcamaları gibi ülkeye özgü değişkenler göçmen öğrencilerin başarısı için ayrıca önemli faktörlerdir. Göçmen ve yerli öğrenciler arasındaki eşitsizlikler yeterince önem verilmeden yapılan kurumsal düzenlemelerle yeniden üretilmektedir.

CONTENTS

LIST OF TABLES

1. Introduction

Recent decades have witnessed high levels of immigration to industrialized countries therefore the educational achievement of immigrant population in host countries is an increasingly important issue. The inequality in education between immigrants and natives influences labor market outcomes and the integration of immigrants within host countries. Moreover, the integration of immigrant populations is crucial for ensuring social cohesion in the host countries. The immigrants carry their wealth of human capital to host countries however if this wealth is not utilized efficiently it may negatively contribute to the economic welfare and cultural diversity of the host countries. As a result, the integration of immigrants to the host country is a major issue for policy makers.

Although immigrant students often underperform at significant levels compared with their native counterparts, the immigrants are better than their native peers in some host countries. Thus, the educational outcome of immigrant students is varying between the host countries. For instance, immigrants that emigrate from the same country have different educational outcomes in the different host countries and these differences are significantly great (Stanat et al. (2006)). The educational achievement of immigrant students has been discussed extensively so far. Studies on immigrants in the economics literature are recently increasing and drawing attention (Aydemir et al. (2008), Gang and Zimmermann (2000), and Bauer and Riphahn (2007)). However, we know little about the main determinants of the facts above. The educational outcome of immigrants appears to be determined by the factors related to family background and school characteristics. In addition to family background and school characteristics, differences in the educational outcomes of immigrant students may be connected to institutional structures of the host countries in education.

This article focuses on 31 OECD countries that contain significant heterogeneity in the institutional structures related to education. In particular, we study the effect of 1) the primary school starting age, 2) preprimary education, 3) the time spent in a tracking system 4) the time spent in the education system of a host country, and investigate whether these institutional arrangements have distinct effects on the natives and immigrant students.

There may be various reasons behind the differences between the immigrant and native students as well as the differences among the immigrant students in different host countries. Firstly, the effect of family background may be a reason for this variation since

students from high socioeconomic status families outperform, through their greater access to educational resources, relative to their low family background peers (Schuetz et al.(2008), Ludemann and Schwerdt(2010), Schnepf(2007), and Brunello and Checchi(2007)). Secondly, the distinction between the immigrant and native students appears to be partly driven by the skill differences in the language of a host country owing to the fact that the low language ability is decreasing the cognitive capacities of immigrant students (Schnepf (2007)). Apart from these facts, the institutional arrangements in education are key mechanisms that determine the educational achievement of both immigrant and native peers.

The variation in the duration of preprimary education in the host countries is the first institutional characteristic that significantly influences the educational outcomes of the students including immigrant students (Schuetz et al. (2008), Schneeweis (2011), Datar (2006), Lubotsky (2009) and Deming et al. (2008)). The duration of preprimary education may improve the educational outcomes of both natives and immigrants for several reasons. For instance, the language ability of immigrant students may improve and the future negative effect of the obstacle of language may diminish (Dustmann, Frattini and Lanzar (2011), Schneeweis (2011)). Consequently, we expect that the longer period of preprimary education improves the educational achievement of both native and immigrant students. Similarly, entering school early may influence the students positively, especially those from low family background. The negative effect of family background may be eliminated if they have the access to educational resources and environment early (Cobb-Clark et al. (2012)). As a result, the early school starting age is expected to advance the educational outcomes of immigrant students due to providing familiarity with the language of the host country and diminishing the negative effect of family background. Furthermore, the existence of any kind of tracking i.e. selection of students[1], can affect the educational outcomes of the students. The tracking may increase the inequality in score distribution between immigrants and natives. In addition it may have detrimental effect on educational outcomes in terms of family background including the fact that the lower family background students are affected more negatively from the tracking system (Hanushek and Woessman (2006), Schuetz et al. (2008), Cobb-Clark et al. (2012), Dronkers and Velden (2012), Horn (2012), Korthals (2012), and Brunello and Checchi (2007)). In contrast to the claim of negative effect of a tracking system on

[1] For more details for conceptualization of tracking, see the part of the Previous Studies.

inequality in educational outcomes, Waldinger (2006) figures out that the tracking has no effect on inequality. Consequently, the effect of tracking system is ambiguous however we expect that it has a negative effect on the educational achievement of the immigrant students. Thus, the application process of tracking may not take into account that the immigrant students have both socioeconomic and cultural disadvantages. For instance, the immigrant students may drop into lower track because of their language and low family background conditions. Apart from these factors, we control for the time a student spent in the educational system of host country and we expect that as the time spent in education system increases the student perform better. As an example, the negative effects of cultural adaptation and institutional unfamiliarity may diminish as the time passes in the host country. On the whole, it seems plausible that each of institutions determine the educational achievement of both native and immigrant students. In this study, we test whether these educational institutions have distinct effects for the natives and the immigrants. Moreover, we test whether they are efficient[2] mechanisms to increase the performance of students.

This article contributes to the literature in several ways: firstly, the previous literature that investigates the effect of institutions conducts analysis focuses on single characteristics of an educational system. For instance, Dronkers and Velden (2012), Horn (2012), Brunello and Checchi (2007), and Hanushek and Woessman (2006) investigate the effect of tracking without considering other institutional mechanisms and the status of immigration. By conducting an analysis covering all potential institutional mechanisms, we go beyond these studies. Secondly, while similar to our study Cobb-Clark et al. (2012) investigates the effects of educational institutions on immigrant youth, due to model specification this study can measure effect of institutional characteristics on immigrants only but not natives. Our study focuses on the effects of institutional characteristics on both natives and immigrants. Thirdly, the previous studies are not considering the fixed effects of region of origin, i.e. the region of emigration, (Cobb-Clark et al. (2012), Horn (2012), Schnepf (2007), Schneeweis (2011)). This study takes into account differences in source country composition of immigrants while investigating the effects of institutions on immigrant students. These differences are important since, for instance, the region of origin may accelerate the adaptation of immigrant students to

[2] Efficiency is capturing the fact that whether a mechanism consistently increases the performance of students. If the tracking mechanism increases the scores of students then we call tracking "an efficient" mechanism.

the host country for several reasons (e.g. ethnic discrimination). Fourthly, we investigate the effect of time spent in the educational system of a host country, which is a crucial determinant of success for the immigrant students. Moreover, we also contribute to the literature by using a different construction for tracking. We use the time spent in a track to investigate the effect of tracking on the students since the effect of tracking occurs in time. Finally, this article takes the advantage of three waves of PISA (2003, 2006, and 2009) by pooling data across years by using the institutional heterogeneity in 31 OECD countries.

Our results confirm that the institutional mechanisms in the educational system of a host country have a remarkable effect on the educational achievements of immigrant students in addition to playing a key role in the variation of educational outcomes of OECD countries when considering the achievements of the native population of each country. Particularly, the preprimary education, tracking, school starting age and the time spent in the education system of a host country are decisive factors on the educational integration of an immigrant population.

The next section reviews the previous literature regarding the immigrants and their educational achievements, the institutional arrangements in the host countries. Section 3 describes the data and presents graphical evidences and sketches our empirical approach. Results and robustness test are given in section 4 and finally section 5 concludes the article.

2. Previous Studies

When considering the literature on the relationship between school tracking and educational achievement, one has to remember that the word 'tracking' refers to the presence of different curricula with an academic and a vocational emphasis, and that students are assigned into schools that specialize in each curriculum in Europe (Brunello and Checchi (2007), p.787). However, in the USA tracking represents ability streaming within a comprehensive school system. In the US, the students are not streamed before their stage of high school in fact they are separated into groups during the stage of high school in terms of ability level of a student. In Europe, the tracking almost refers to existence of a system that separates the students into different high schools such as assigning some into vocational schools or some into general high school.

The current empirical literature covers both country specific and cross-country analysis of the relationship between school tracking and educational achievement of individuals. The results on this relationship have not produced a consensus to date regarding the effects of tracking on educational achievement. The current literature tries to identify the effect of tracking systems on the educational achievement of students by considering all students, immigrant students and both, dividing them into separate samples. The studies covering the effects of tracking system on immigrant students in high immigration countries will be briefly described in the subsection below, as well as, the studies conducting an analysis that only covers immigrants.

The disadvantage of immigrant students in the education systems of host countries leads to detrimental influence for the future labor market decisions of immigrants. Kahn (2004) inquiries into the topic of skills of immigrants and employment by using a sample consisting of OECD countries like New Zealand, Canada, Switzerland and US. The measure of labor-market outcomes is the employment probability of immigrant individual through a cross-section analysis. Controlling for skills he reaches the conclusion that immigrants have lower cognitive skills than natives in each country, with the largest gaps in the US, and the smaller gaps in Canada and New Zealand. Hence, the examination of tracking and its effects on immigrants' educational success -which is the sign of skill characteristic for the worker in the labor-market according to economic theory-, is an inspiring topic of investigation given the potential negative effect of tracking on immigrant students.

Firstly, we will consider the studies which do not consider the status of immigration between immigrant and native students. By assuming tracking is implemented efficiently, Korthals (2012) finds evidence that equality of opportunity is best provided in a system with many tracks. This study carries out a random effect model analysis using the PISA 2009 dataset including 185000 students in 31 comparable countries. This paper contributes to the literature of tracking as follows. First, tracking is defined as number of school types or distinct educational programs available to 15 years old students; however tracking at earlier stages of education is not considered. Second, unlike the previous studies the school characteristics and country level variables are controlled by using a three-level random effect model which is based on individual, school and country level. Finally, the study is considering both the issue of performance and the equality of opportunity by using the test score in reading, math and science of students as the measure of achievement.

In terms of the equity effects, Waldinger (2006) uses a differences-in-differences approach to study the effects of tracking on family background[3] by using the datasets PISA 2003, TIMMS (1999) and PIRLS (2001). He identifies the tracking as the grade of first tracking in his student level analysis by considering 27 OECD countries. He tries to understand the relation between family background and tracking using math and reading scores as measures of achievement. As a result, he finds the following two results: First, he concludes that family background is more important in the countries where students are tracked in at an early age. Second, he finds evidence that family background has no importance after actual tracking has been implemented. Hanushek and Woessmann (2006) investigate the topic of early tracking and inequality of opportunity by using the datasets of PISA 2000 and 2003, PIRLS (2001) and TIMMS (1995, 1999, and 2003) for several years including 45 countries. They use a pooled data and carry out differences-in-differences method including a country level analysis. Standard deviation in math, science and reading

[3] In general, selection is influenced, directly or indirectly, by family background. For example, better-educated parents are more likely to send their children in a general track, which results in university participation. On the other hand, blue collar families send theirs to vocational school. Even the allocation requires a formal test, students from better educated families are more likely to enter the academic track, either because of their cognitive ability or genes or because of the results of their environment.

test scores is used as a measure of equity and age of first tracking as measure of institutions for tracking. They conclude that early tracking has significant effect on inequality and no clear effect on efficiency.

Brunello and Checchi (2007) investigate the topic of school tracking and equality of opportunity using a student level analysis through the datasets of PISA 2003, IALS (1994, 1996, and 1998), ISSP (1999) and ECHP (1995-2000) including nearly 30 countries. Even though their scope of research does not cover immigrants and their achievements, they estimate the OLS, the probit and the multinomial logit models and thus they contribute to the literature that tracking strengthens family background effects for formal education but weakens them for on the job learning. Their paper is important in two respects: first, tracking may have negative effect on individuals with poor family background. Secondly, while considering the overall living conditions of immigrant populations in OECD countries relative to the natives we can predict that the family background effect probably leads to detrimental results for the immigrant students. We try to address this issue in this paper. In addition to the study of Brunello and Checchi (2007) including the interaction between tracking and family background, Schuetz et al. (2008) also investigate the topic of equality of opportunity in the scope of tracking and its interaction with family background effect. Conducting a weighted clustering-robust linear regression (WCRLR) and a country fixed effect analysis on the datasets of TIMMS and TIMMS-R with 54 countries- they conclude that late tracking and pre-school duration reduce the impact of family background, which is a significant result for our study since one of our hypotheses is that early tracking increases the negative effect of poor family living conditions and of social class effect for immigrant students.

In addition to the studies above, there are additional studies whose scope specifically considers the immigrants' success and the effect of tracking systems in host countries. Cobb-Clark et al. (2012) use the PISA 2009 dataset, where they consider the effect of institutions in host countries on the immigrants. Their sample is composed of 34 OECD countries. Hence, they use an OLS method including country fixed effects. They find that achievement gaps are larger for immigrant students that arrive at older ages and those who do not speak the test language at home by conducting a student level analysis based on reading, math and science scores as the measures of achievement. They conclude that early age of starting school help immigrant youths in some cases but not for all. In addition to their consideration of

immigrants, their paper is unique to my knowledge as it considers tracking and institutional determinants. They consider the current ability grouping as the track measure although they control for the age first selection. However, in our study unlike the concern of Cobb-Clark et al. (2012) we are focusing on the effect of tracking by considering the previous track history of a student. In particular, we are investigating the fact that whether previous selection or tracking affects the performance of a student positively or negatively. Their result about the issue of tracking is that limited tracking (i. e. ability grouping in current school) on ability is advantageous for the immigrant students although complete tracking is detrimental.

Compared to previous research in the area of immigrant's educational achievement, the contribution of Dronkers and Velden (2012) is that they explicitly include track level and school level as independent units of analyses, which leads to more accurate results of the effects of characteristics of the educational system. Their sample based on PISA 2006 consists of 15 countries, 8251 student level observations including only immigrant students and their measure of tracking is the age of selection. The previous studies lack a sufficient design to investigate the effects of immigrants' countries of origin and destination as these relate to their final educational achievement. They use a hierarchical linear mode including three levels- students, country and track level- they conclude that first generation immigrants are more successful in comprehensive education systems relative to the systems based on tracking.

Horn (2012) uses a mixed random effect model (two-level Hierarchical Linear Model) using the PISA 2003 dataset and his measure of achievement is the literacy score of students. The level of analysis is based on student and country level as two separate levels. Moreover this study includes an immigrant dummy regarding first and second generation. The measure of institutions for tracking is defined as the age of first selection and this study unlike the previous ones concentrate on both performance and inequality of opportunity. Thus, this paper contributes to the literature that early age of selection links closely with high inequality of opportunity while the standardization enhances equality. Although this paper includes both immigrant and native students in the sample, it is not particularly related to the issue of immigrant disadvantage in OECD countries.

The study of Ludemann and Schwerdt (2010) shows that second generation immigrants face additional disadvantages with respect to grades and teacher recommendations

during the transition to secondary school in the German education system that involves tracks. They use the micro data, German extension of the Progress in International Reading Literacy Study (PIRLS-E) 2001, by carrying out an analysis based on multinomial logit model. This econometric model is taking into account a student level analysis which is the probability of receiving the teacher recommendation in the transition from primary school to secondary school. By testing whether there is an extra inequality of opportunity for 2nd generation immigrants in Germany in the education system, they conclude that the socioeconomic background of students, which may be called family background effect (FBE), play an important role in the transition as the worse conditions of the students' social class detrimentally result in worse track recommendations. Although the study claims the equality of opportunity in the regards of 2nd generation immigrants, the paper ignores the country of origin of immigrants and focuses on Germany as opposed to other studies that carry out analysis across countries with and without a tracking structure in their education systems.

Entorf and Miniou (2005) study the effect of immigrant status on the reading score of immigrant students using PISA 2000 dataset including 10 OECD countries mainly investigating the difference in the reading scores of immigrant students between the Continental European countries and New Zealand and Scandinavian countries. By using a cross-section OLS method, they find the socioeconomic effect to be highest in Germany, the UK and the US in spite of lowest effect in Canada and Scandinavia. According to their study language spoken at home is a key factor in explaining this variation.

Schnepf (2007) studies the immigrants' disadvantage in high immigration countries using the datasets PISA 2003, TIMMS (1995 and 1999), and PIRLS (2001). This study, as in the case of Entorf and Miniou (2005), is based on cross-section OLS by including a sample consisting of ten OECD countries. By using math scores of immigrant students, the paper contributes to the literature that immigrants are more successful compared to natives in English-speaking countries unlike the relative lower performance of immigrants in Continental Europe. The determinants of the above result are summarized as language skills, socio-economic background, and school segregation. However, this paper does not take into account institutional differences between the host countries as well as the country of origin of immigrant students. Although this paper includes the segregation as a determinant for the gap

between the natives and the immigrants, the segregation variable is based on residential segregation rather than the segregation due to institutional differences.

In addition to the studies examining the effects of segregation such as Entorf and Miniou (2005), and Schnepf (2007), Schneeweis (2010) analyses the topic of educational institutions and the integration of immigrants using the datasets of TIMMS, TIMMS-R 2003 and PISA 2000, 2003. The paper investigates the effects of ethnic segregation in schools, pre-primary enrollment, and school starting age, instruction time and external exams as the measure of institutions. Moreover the unexplained test score gap of immigrants is the measure of equity. As a result, this paper comes to the conclusion that the institutions are responsible for 20% of immigrant disadvantage, particularly in pre-primary education, young school starting age, and low classroom segregation and instruction time.

It is important to note that host country educational systems and social policy institutions may either accentuate or mitigate the effects of tracking systems on the educational success of immigrant students emigrating from the same country. The current literature generally makes no distinction between immigrant students and the native ones. For example Korthals (2012) focuses on all students by creating a sample that makes no distinction between the immigrant students and the natives and does not consider the effects of school characteristics on outcomes. Therefore the analysis of school level characteristics may be a crucial instrument to explain the variation between the score differences of students. In addition some studies[4] find that tracking has no effect on the educational success of students regarding equity and efficiency analysis. However, the findings can reverse when considering immigrant students. The study of Cobb-Clark (2012) et al. is significant regarding their consideration of immigrant students and their achievement however they are not taking into account the country of origin of the immigrant students, ethnic diversity within schools or the way that children are allocated to schools. They are not considering the features of schools allocated to immigrant students therefore they do not explain the variation of the success of same ethnic origin immigrants between different host countries.

The analysis of Dronkers and Velden (2012) is a remarkable paper as they include the region of arrival of immigrants similar to our study. Even though their analysis is unique in terms of adding country of origin to their estimation, they are not measuring for the relative

[4] Waldinger (2006)

success of immigrants to the natives. As a result, our paper contributes different viewpoints than the study of Dronkers et al. (2012) to literature as follows. First, we are focusing on success of immigrant students relative to the natives across different host countries taking into account ethnic origin of immigrants. Dronkers (2012) conduct their analysis using five broad regions for immigrants' origins as opposed to this study that considers country of origin. . This allows a better control of immigrant heterogeneity and reduces resulting biases. Secondly, our dataset is more plentiful relative to their dataset as our sample is composed of three PISA datasets including 2003, 2006 and 2009 years. Moreover, our sample employs both the data of natives and immigrants to conduct an analysis investigating the relative differences between natives and immigrants. Bedard and Dhuey(2006) investigate the topic of effects of relative school starting age by using an instrumental variable (IV) method[5] based on the datasets of TIMMS and TIMMS-R. The math and science grade are used as measure of achievement in their study and they conclude that the effects of relative school starting age is remarkable and sizeable on performance at ages 9 and 13.

In addition to the study of Dronkers and Velden (2012), the study of Horn (2012) does not consider including the variables of language and the country of origin for the same immigrant groups. Finally, the exclusion of school starting age and its relation with tracking effect is a major obstacle for the result of this paper. The findings of Entorf and Miniou (2005) are crucial for our study since our main aim is to explain this variation by considering institutional and family level variables. Furthermore, the paper of Schneeweis (2010) is crucial in terms of investigating the effects of institutional differences between countries. These all studies are important to some extent and our aim is to explain the variation in educational achievement of immigrant students across countries. The above studies consider generally single or some institutional factors and do not include the institutional factors whole. We consider all potential institutional factors in our study and investigate their effects on both native and immigrant students. In addition, while doing these we control for the source country of immigrants (i.e. region of origin).

[5] Instrumental variable in their study is age assigned by the cutoff date.

3. Data and Estimation Method

The student and school level data used in this study come from the 2003, 2006 and 2009 waves of the Programme for International Student Assessment (PISA), conducted by the Organization for Economic Co-operation and Development (OECD). The datasets contain internationally comparable test scores in reading, math and science as well as information on students' family background and their schools. The country level datasets are obtained from the OECD, Education at a Glance (2005), United Nations Development Program Database and the World Bank.

The first wave of PISA was conducted in 2000, followed by other waves every three years with a representative sample from all participating countries covering tests on reading, math and science. The results of the tests are standardized to a mean of 500 and a standard deviation of 100 for the OECD countries. Moreover, the students and school administrators (in some countries also the parents) are surveyed. The number of participating countries varies in terms of the test year. For example, 34 OECD countries and 41 partner countries are included in the PISA 2009. The economic, institutional and social diversity in the participating countries result in heterogeneity that may affect the test results. This study covers 31 countries included in the tests in 2003, 2006 and 2009. As a result, the datasets in this study is composed of three waves of the PISA tests as a pooled dataset.

A representative sample[6] from each participating country is obtained by the OECD in two stages: first, schools are selected and, then, students with the target age are selected in the chosen schools. The age of students is set to a range of 15 years and three months to 16 years and two months. The OECD uses weights to ensure sample representation because not all school and students selected were willing to participate and some schools and students were oversampled.

3.1. Analysis Sample

The sample is pooled and composed of three waves of PISA tests. The first wave is the 2003 PISA test that includes 131,789 observations after dropping missing observations. Secondly, the PISA 2006 part contains 183,056 observations after eliminating missing ones.

6 A two stage stratified sampling design is used by OECD. First a random sample of schools is selected and then a random selection of students is chosen from each school.

Finally, the PISA 2009 part contains 200,242 observations which results in a total of 515,087 observations. We eliminated observations that lack information on age, gender, and immigrant status or origin country of emigration. The resulting sample of 515,087 students comes from nearly 26000 schools in 31 countries.

The sample includes both the natives and the immigrant students. There are 17,381 immigrant students, 10,636 of 2nd generation immigrant and 6,745 of 1st generation ones. We classify the students into the following three groups: 1) native born: those born in the country of test; 2) first generation immigrant: those not born in the country of test; and 3) second generation immigrant: those born in the country of test but whose parent(s) (i.e. both parents) were born in another country. By considering the migration status of students, we generated a categorical variable in order to conduct an empirical analysis taking into account the country of origin for the immigrant student. This categorical variable includes the country of origin for the immigrant students in terms of both the country of self-birth for 1st generation immigrants and the country of birth of parents for the 2nd generation immigrants. This kind of categorization is used in Dronkers et al. (2012) however our categorical variable has more regions, which has 15 categories as presented in Table 1 below. This variable helps us control the variation in the scores related to the region of origin by including region of origin fixed effects. As a result our regions of origin controls are richer than Dronkers et al. (2012).

Region of Origin	Status of Immigration			
	Native	**Second Generation**	**First Generation**	Total
Central Africa	0	412	176	588
Central America	0	80	3	83
Eastern Asia	0	195	417	612
Eastern Europe	0	1,331	611	1,942
Middle East	0	147	99	246
Native	497,706	0	0	497,706
North America	0	733	53	786
Northern Africa	0	124	20	144
Northern Europe	0	832	517	1,349
Oceania	0	349	178	527
South America	0	18	83	101
Southeast Asia	0	129	151	280
Southern Africa	0	24	157	181
Southern Europe	0	3,572	3,373	6,945
Western Africa	0	20	22	42
Western Europe	0	2,670	885	3,555
Total	497,706	10,636	6,745	515,087

Table 1: Region of Migration and Immigrant Status

The final sample contains observations from 31 OECD countries and the observation number per country is summarized in Table 2 below. Some of the countries have no immigrant observation; however we include them in the analysis since we are curious about the effect of institutions on both native and immigrant populations in terms of equity and efficiency[7] in the educational system.

[7] The meaning of efficiency in this study is explained in footnote 1. Equity is capturing the fact that whether the institutional arrangements increase the gap between the immigrant and the natives.

	Status of Immigration			
COUNTRY	Native	Second Generation	First Generation	Total
Australia	8,603	527	633	9,763
Austria	11,545	642	579	12,766
Belgium	17,472	696	568	18,736
Canada	46,745	709	1	47,455
Czech Rep.	14,843	76	52	14,971
Denmark	9,504	502	134	10,140
Estonia	7,452	539	12	8,003
Finland	14,179	24	45	14,248
Germany	9,013	608	376	9,997
Greece	10,066	118	466	10,650
Hungary	11,233	14	0	11,247
Iceland	8,697	9	0	8,706
Ireland	8,784	75	132	8,991
Italy	44,844	84	4	44,932
Japan	13,462	2	0	13,464
Korea	13,509	1	0	13,510
Luxembourg	6,645	1,543	1,121	9,309
Mexico	48,073	85	26	48,184
Netherlands	9,808	418	65	10,291
Norway	10,170	112	38	10,320
New Zealand	8,467	441	634	9,542
Poland	12,973	2	0	12,975
Portugal	12,657	165	126	12,948
Spain	42,679	57	0	42,736
Slovakia	13,848	12	1	13,861
Slovenia	9,619	436	0	10,055
Sweden	10,062	143	0	10,205
Switzerland	20,453	2,375	1,698	24,526
Turkey	11,081	23	3	11,107
United Kingdom	24,858	198	31	25,087
USA	6,362	0	0	6,362
Total	497,706	10,636	6,745	515,087

Table 2: The distribution in terms of immigration status in the host countries

3.2. Measurement of PISA test Scores

PISA tests in all three years include five plausible values for each subject -i.e. (science, reading and mathematics) - by assigning random numbers drawn from the distribution of scores that could be attributed to each individual. The marginal posterior distribution is the statistical method that makes a students' outcome on any individual test random to some extent (see OECD (2012) for more detailed explanation). We construct for each individual a test score by averaging five plausible values for each section. The test scores have a distribution with a mean of 500 and a standard deviation of 100 across countries for each section of subject. In our analysis, we restandardize the test scores with mean of 0 and standard deviation of 1 for each year for an easier interpretation of the regression results. Moreover this strategy represents the standard deviation changes in the measure of interest which are science, math and reading scores in the PISA tests. Table 3 represents the distribution of test scores for each section across countries in terms of the immigrant status. Generally, the natives have higher score than the immigrants for both the 1^{st} and 2^{nd} generation across countries. The difference in percentage varies between -30% and 10%, which is sometimes detrimental results for the immigrants. By looking at the statistics in the Table 3, it is verified that there are only 3 countries in reading, 5 in science and 3 in math where both the first and the second generation do not underperform the native-born peers. Therefore, it is provided that in most of the host countries the immigrants have underperformed native-born peers and the gap between them is sometimes detrimental as shown in the Table 3.

Status of Immigration									
	Native			Second Generation			First Generation		
COUNTRY	Science	Reading	Math	Science	Reading	Math	Science	Reading	Math
Australia	529.0707	514.7701	514.4282	557.6208	548.7158	551.0454	537.5703	529.9637	529.1539
Austria	518.7251	501.4095	519.2193	438.8152	434.754	455.1906	436.0874	443.0395	458.1405
Belgium	530.9559	527.2654	544.8482	451.3014	453.3051	466.2678	472.7928	475.8748	478.6935
Canada	521.4896	518.7188	522.2563	518.3485	521.9035	521.1726			
Czech Rep.	538.8475	510.7919	534.6298	480.2895	467.9475	474.1963	520.887	494.6599	524.8113
Denmark	498.6391	501.0141	516.9456	414.8737	434.064	440.2667	417.9086	428.5616	427.5016
Estonia	538.8628	511.0438	521.48	497.032	466.0294	487.4174	476.935	463.0798	463.5472
Finland	554.5359	542.7051	545.398	519.1559	523.9542	509.5469	513.2484	511.0511	521.8719
Germany	539.7464	520.4307	529.8539	456.6592	452.8629	464.2767	480.3079	471.6949	481.7543
Greece	480.927	479.4456	462.5436	453.3514	462.0357	447.7028	435.3997	434.6281	418.5837
Hungary	510.031	493.0313	497.3506	471.7893	453.9712	495.569			
Iceland	495.7345	495.4228	511.6528	446.278	448.28	433.445			
Ireland	512.8064	515.6909	502.1712	513.0041	509.1223	492.0631	526.7518	522.6788	506.6129
Italy	507.3181	498.396	495.1515	473.8419	507.2343	456.1623	522.301	519.7915	509.978
Japan	547.1414	514.4233	535.4727	497.5123	390.5603	475.7673			
Korea	534.758	545.7567	547.7531	592.7178	370.9926	549.6884			
Luxembourg	510.6211	505.0631	512.109	454.083	450.7097	468.2529	443.9742	438.5972	458.3484
Mexico	425.7108	435.2762	427.5205	363.7979	338.3837	350.5495	365.1613	356.4399	363.1196
Netherlands	543.379	526.0165	548.4172	484.0852	485.9688	497.3166	466.2225	477.1948	489.2777
New Zealand	540.0096	531.3583	529.5857	515.074	521.5224	513.2062	545.5873	530.1414	546.9432
Norway	498.4614	502.7281	500.6373	437.6373	454.3889	448.7032	484.9822	490.215	491.9868
Poland	507.3069	508.3778	500.1919	552.2485	521.3035	514.5583			
Portuguese	483.4008	484.6802	478.9733	466.362	469.6228	451.8481	460.8593	460.8647	453.5796
Slovakia	501.5997	480.1168	504.9322	442.8848	420.957	449.6079	410.0462	417.9449	386.1895
Slovenia	501.7382	472.7467	488.7823	453.7461	451.3518	461.0852			
Spain	503.4785	490.8269	503.0411	472.558	456.8621	456.2802			
Sweden	513.7136	517.5438	512.8409	463.1522	484.2052	469.782			
Switzerland	527.4826	512.1606	545.2465	461.8872	462.722	486.4751	435.3148	432.0031	460.911
Turkey	440.9531	455.5552	435.8326	428.3527	425.1341	443.0265	460.8248	465.0619	445.8171
United Kingdom	521.2183	506.3815	505.6511	502.5926	503.1451	486.9138	520.2256	501.7725	509.2015
USA	507.5963	508.1812	495.42						

Table 3: Score distribution for each immigrant group

3.3. Measurement of Educational Institutions

The main concern of this study is to understand the role of educational institutions on the variation in the test scores of both the native and immigrant individuals. We are not only interested in the role of educational institutions in the host countries on the outcome of immigrants, but also we want to investigate the effect of institutions on the native born peers. Hence, we can carry out an empirical analysis concerning both efficiency and equity analysis since by adding the natives into our analysis we can extract the effect of institutions on them. As a result, the analysis may inform policy makers about which institutional arrangements are efficient for any country. Cobb-Clark et al. (2012) also explore the role of institutions on the outcome of migrant youth. Unlike their empirical strategy focusing only on immigrants we have the possibility of investigating the effect of institutions for both the natives and the immigrants separately. In our analysis, the measurement of institutional arrangement in a quantitative way is based on the inclusion of a series of country level variables that are represented in our empirical model. The data for these variables are obtained from OECD and EAG (2005), as well as, the World Bank Database.

Precisely, we generate variables for the primary school-starting age that has three categories using the data from the World Bank Database, and the expected total duration of pre-primary education in years from OECD Education Database. Moreover, considering the age of first selection in country level from EAG (2005) we generate a variable which gives the total time a student spends in a specific track. For instance, if a country has the age of selection as 10 then the total time a student spends is equal to the age of this student-10. The main idea of this construction is based on the fact that the first age of selection endures its positive or negative effects on the performance of the student in the whole education program after the selection. Besides generating a track variable, we generate a variable -especially crucial for the immigrant population- that takes into account the time in years a student spends in the education system of a country. For instance, suppose in a country the primary school starting age is 5 and the immigrant student comes to this host country at the age of 7 and he/she is 15 years old while taking the PISA test. Then it is assumed that she has spent 8 years in the education system of the host country. The adaptation of an immigrant student is presumptively increasing as long as the number of years spent in the education system of the host country increases. Another equally important institutional variable used in our analysis is

the expected pre-primary education in years in a host country. This variable captures exposure to school system before the start of primary school. Thus, through above variables we try to account for to what extent the student is familiar with the system of the host country as well as its culture, including language issues and ability generating processes.

These variables are not the only characteristics of educational systems that might be of interest. Therefore we include variables that try to account for diversity within schools, the features of schools plus the method that students are allocated to schools. For this purpose, we control the school average family background, teacher shortage, existence of ability grouping, school type and the selection method of schools etc.

In addition to institutional variables describing educational systems, we also control for a country's economic features including the human development index, Gini coefficient on income and employment ratio. Moreover, we also control for the expenditure on education such as teacher salary in purchasing power parity term, the ratio of educational expenditure in the government budget as well as pupil teacher ratio in order to extract the variation based on country specific items. Through these variables, we are trying to account for country specific characteristics that may influence student outcomes. The Appendix A shows how the institutional structures differ by country.

3.4. The Method of Estimation

In our analysis, we are using a linear regression model of the following form:

$$S_{isct} = \beta_0 + P_{isct}\beta_1 + I_{isct}\beta_2 + R_{isct}\beta_3 + INS_{isct}\beta_4 + INS_c * P_{isct}\beta_5 + A_{isct}\beta_6 + SCH_{cst}\beta_7 + C_c\beta_8 + Y_t\beta_9 + \varepsilon_{icst}$$

$$i = 1, \dots \dots N \qquad s = 1, \dots \dots, S \qquad c = 1, \dots \dots ., C \qquad t = 2003,2006,2009$$

where **S** is the reading, math or science test of score of a student *i* in school *s* of country *c.* **P** includes three population indicators that identify the immigrant status i.e. whether the student is native-born, 1st generation immigrant or 2nd generation one. Furthermore, **I** includes individual characteristics such as the age, the gender, family background index[8] ESCS , the highest employment category of either parent as well as whether the student is in the upper

[8] ESCS is the index on economic, social and cultural status. The index of ESCS is composed of five criteria: highest occupational status of parents (HISEI), highest educational level of parents (in years of education according to ISCED), family wealth, cultural possessions and home educational resources.

secondary education. Our estimation also contains source region of immigration for the immigrant population which is indicated by **R**. Moreover, we account for the effects of institutions on the students' scores, and therefore we include the institutional arrangements in **INS**. We interact **P** with **INS** in order to observe the effect of institutional arrangements on immigrant population, which is the main interest of the paper. As Entorf et al. (2005), Schnepf (2007) and Cobb-Clark et al. (2012) suggest the language spoken at home and the age of arrival to host country is a key factor to estimate the determinants of immigrant success, thus we are controlling these potential determinants in the variable **A**.

Our regression analysis also controls for the school characteristics of the students like teacher shortage, school average ESCS and the existence of ability grouping etc. therefore the vector of **SCH** is generated. Note that while carrying out this analysis it is impossible to control country specific factors due to multicollinearity between institutional characteristics and fixed effects. Nevertheless, we aim to control country specific variation by including some country level characteristics such as teacher salary in purchasing power parity term, the ratio of educational expenditure in the government budget in **C**. Furthermore, we control for the fixed effects of any year by adding dummies for each specific year. Finally, we have the error term ε.

Our baseline model starts with investigating the nativity gap in test scores and the following regression results are also summarized in the section of results. As an illustration, the first regression includes only immigrant characteristics of the individual as shown in Table 4. In spite of our effort in this study, Hanushek and Woessmann (2011) argue that the main problem in identifying the causal relationship between host country institutions on education success is the possible endogeneity problem due to unobserved country specific effects that are correlated with student achievement. Instead of including country specific effect as in Cobb-Clark et al. (2012), we include country specific variables in our regression analysis in order to identify the effect of institutions on immigrant population relative to native ones as well as to make an efficiency analysis across countries in terms of institutions.

4. Results

4.1. The gap between the immigrants and the natives

In the first regression reported in Table 4, we are interested in the nativity gap, and therefore we only add country specific variables and the population indicators that have the reference group of native-born peers. There is a large nativity gap for the immigrants in OECD countries for both 1st and 2nd generation immigrants with reference to the native born peers. By controlling gender, age, status of secondary education, family background, and the highest employment status of family[9] we find that the 1st generation peers perform 0.265 standard deviation lower in reading score (see column 2 in Table 4), 0.243 lower in math (see column 8 in Table 4) and finally 0.335 in science (see column 5 in Table 4) than the native ones and all coefficients are statistically significant. These findings are similar to previous studies (Cobb-Clark et al. (2012), Horn (2012), Entorf and Miniou (2005) and Brunello and Checchi (2007)).

Unlike the previous literature (Cobb-Clark et al. (2012), Horn (2012), Ludemann and Schwerdt (2010), Entorf and Miniou (2005), Schnepf (2007) and Dronkers and Velden (2012)) we carry out an analysis using the region of origin of immigrants in our OLS estimation. Needless to say, there is an additional variation in the outcome of immigrants due to their region of origin. Thus we are able to extract this variation with the help of this estimation strategy that distinguishes this paper from the rest of the literature. One advantage of this strategy is that we can explain variations due to the regional and maybe cultural factors (such as being from a Christian or Islamic region). Results from this third specification as a result of controlling for the region of origin for immigrants shows firstly that in math scores in the column 9 in the Table 4 the 1st generation immigrants perform lower than natives by an amount of a 0.455 standard deviation, which is lower than the situation of not controlling for region of origins. Secondly this trend for the 1st generation applies to the science scores and the difference between the natives and 1st generation immigrants increases to a 0.505 standard deviation as well as a 0.454 standard deviation for the reading scores (see column 6 and 3 in the Table 4). Not only this consistent trend applies to the 1st generation immigrants, it is also

[9] We do not need to use highest educational status of parents since the ESCS includes the effect of the educational status of family on the ground that the ESCS and educational variable are perfectly correlated. For this reason, we only use ESCS in our regression model.

valid for the 2nd generation immigrants. The gap in math scores between the natives and the 2nd generation immigrants increases to 0.378 standard deviation from 0.268 standard deviation after controlling the region of origins (see the column 8 and 9 in the Table 4). Besides this, the gap in science and reading scores increases to 0.476 and 0.381 standard deviation respectively (see Table 4 for details). To the best of my knowledge, there is no study considering regions of origins and its fixed effects for the immigrant students. Therefore our study displays that the gap between the natives and the immigrants on average is dramatically higher for particular immigrant groups unlike the previous studies Cobb-Clark et al. (2012), Horn (2012), Ludemann and Schwerdt (2010), Entorf and Miniou (2005), Schnepf (2007), Schneeweis (2010), and Dronkers and Velden (2012). All in all, the gap between the immigrants and the natives in education is detrimentally great regardless of the fact that we account for the variation due to the region of origins.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Reading Score	Reading Score	Reading Score	Science Score	Science Score	Science Score	Math Score	Math Score	Math Score
Second Generation Immigrant	-0.536***	-0.275***	-0.381***	-0.637***	-0.375***	-0.476***	-0.536***	-0.268***	-0.378***
	(0.0129)	(0.0117)	(0.0222)	(0.0127)	(0.0117)	(0.0222)	(0.0127)	(0.0116)	(0.0220)
First Generation Immigrant	-0.501***	-0.265***	-0.454***	-0.568***	-0.335***	-0.505***	-0.481***	-0.243***	-0.455***
	(0.0173)	(0.0156)	(0.0291)	(0.0170)	(0.0156)	(0.0290)	(0.0169)	(0.0155)	(0.0288)
Gender (Female)		0.347***	0.347***		-0.0636***	-0.0638***		-0.156***	-0.157***
		(0.00245)	(0.00245)		(0.00245)	(0.00244)		(0.00242)	(0.00242)
ESCS		0.358***	0.357***		0.372***	0.371***		0.382***	0.381***
		(0.00180)	(0.00180)		(0.00180)	(0.00180)		(0.00178)	(0.00178)
AGE		0.138***	0.139***		0.111***	0.112***		0.136***	0.136***
		(0.00422)	(0.00422)		(0.00422)	(0.00422)		(0.00418)	(0.00418)
Upper Secondary		0.306***	0.307***		0.257***	0.259***		0.267***	0.268***
		(0.00324)	(0.00324)		(0.00324)	(0.00324)		(0.00321)	(0.00321)
White collar low skilled parent		-0.0610***	-0.0611***		-0.0386***	-0.0387***		-0.0298***	-0.0298***
		(0.00339)	(0.00339)		(0.00339)	(0.00338)		(0.00336)	(0.00335)
Blue collar high skilled parent		-0.0780***	-0.0764***		-0.0314***	-0.0294***		-0.0223***	-0.0209***
		(0.00473)	(0.00473)		(0.00473)	(0.00473)		(0.00469)	(0.00468)
White collar low skilled parent		-0.0696***	-0.0672***		-0.0242***	-0.0215***		-0.0165***	-0.0144***
		(0.00555)	(0.00555)		(0.00555)	(0.00555)		(0.00550)	(0.00550)
Human Development Index	0.0382***	0.0815***	0.0835***	0.0697***	0.106***	0.109***	0.0895***	0.128***	0.129***
	(0.00125)	(0.00123)	(0.00124)	(0.00122)	(0.00123)	(0.00124)	(0.00122)	(0.00122)	(0.00122)
Gini	-0.175***	-0.219***	-0.220***	-0.0967***	-0.134***	-0.135***	-0.151***	-0.189***	-0.190***
	(0.00232)	(0.00211)	(0.00211)	(0.00228)	(0.00210)	(0.00210)	(0.00227)	(0.00208)	(0.00209)
Pupil teacher ratio	0.00402***	-0.0142***	-0.0151***	-0.000987	-0.0175***	-0.0187***	0.00402***	-0.0132***	-0.0141***
	(0.000628)	(0.000587)	(0.000589)	(0.000615)	(0.000587)	(0.000588)	(0.000615)	(0.000581)	(0.000583)
Education share in government expenditure	-0.0981***	-0.0666***	-0.0670***	-0.111***	-0.0814***	-0.0821***	-0.122***	-0.0911***	-0.0918***
	(0.00116)	(0.00107)	(0.00107)	(0.00114)	(0.00107)	(0.00107)	(0.00114)	(0.00106)	(0.00106)
Public expenditure education per GDP	0.176***	0.115***	0.115***	0.0699***	0.00456*	0.00453*	0.121***	0.0545***	0.0552***
	(0.00301)	(0.00273)	(0.00274)	(0.00295)	(0.00273)	(0.00274)	(0.00295)	(0.00270)	(0.00271)
Teacher salary	1.48e-05***	1.55e-05***	1.58e-05***	1.08e-05***	1.10e-05***	1.13e-05***	1.43e-05***	1.45e-05***	1.47e-05***
	(2.22e-07)	(2.03e-07)	(2.03e-07)	(2.18e-07)	(2.03e-07)	(2.03e-07)	(2.17e-07)	(2.01e-07)	(2.01e-07)
Region of origin (Central Africa)			0.0620			-0.00407			0.00838
			(0.0559)			(0.0558)			(0.0553)
Region of origin (Central America)			-0.600***			0.236***			-0.0737
			(0.0896)			(0.0895)			(0.0887)
Region of origin (Eastern Asia)			0.759***			0.990***			1.064***
			(0.0605)			(0.0604)			(0.0599)
Region of origin (Eastern Europe)			0.348***			0.304***			0.348***
			(0.0311)			(0.0311)			(0.0308)
Region of origin (Middle East)			-0.136			-0.219*			-0.101
			(0.130)			(0.130)			(0.129)
Region of origin (North America)			0.460***			0.397***			0.322***
			(0.0436)			(0.0436)			(0.0432)
Region of origin (Northern Africa)			-0.199*			-0.0850			0.00765
			(0.119)			(0.119)			(0.118)
Region of origin (Northern Europe)			0.388***			0.404***			0.246***
			(0.0371)			(0.0370)			(0.0367)
Region of origin (Oceania)			0.447***			0.465***			0.365***
			(0.0692)			(0.0691)			(0.0685)
Region of origin (South America)			0.202			0.0380			0.0185

			(0.132)			(0.132)			(0.131)
Region of origin (Southeast Asia)			0.605***			0.704***			0.666***
			(0.0845)			(0.0844)			(0.0837)
Region of origin (Southern Africa)			0.452***			0.573***			0.389***
			(0.106)			(0.106)			(0.105)
Region of origin (Southern Europe)			-0.149***			-0.194***			-0.0813***
			(0.0278)			(0.0278)			(0.0276)
Region of origin (Western Africa)			-1.360			-1.296			-1.218
			(0.833)			(0.832)			(0.825)
Observations	515,087	515,087	515,087	515,087	515,087	515,087	515,087	515,087	515,087
R-squared	0.098	0.267	0.268	0.138	0.273	0.274	0.143	0.288	0.289

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

Table 4: The regression results for the nativity gap.

4.2. The Institutions

Typically, the educational outcome of the immigrant population is based on many different factors. Apart from the individual, geographical and family characteristics, the institutional arrangement of education system in the host countries influences the performance of immigrant peers and results in a variation between natives and immigrant peers.

Immigrant students' relative achievement is naturally related to host country's school starting age. First of all we are interested in the main effect of school starting age on both natives and immigrants. The results are summarized in Tables 5, 6, and 7. School starting age is controlled with a categorical variable where age 6 is the omitted category. If a student, whether she is an immigrant or native, goes to primary school at an early age (school starting age<=5), the achievement is decreasing in all three domains of interest. First column of the tables show that starting to primary school in age of 4 or 5 relative to age of 6 results in 0.164 standard deviation lower score in the math, 0.116 standard deviation lower score in reading. Nonetheless, we are not the first one investigating the effect of school starting age on immigrant's educational achievement. Cobb-Clark et al. (2012) also investigates this issue but our study has two distinctions. Firstly, our findings are consistently and intuitively more robust since we control for the school characteristics and the region of origin. Secondly, we can estimate both the main effects and the specific effects for the immigrant of institutions. The interaction between the school starting age and the immigrant dummies are generated to carry out an analysis investigating the specific effect of school starting age on the immigrant students. At first, the main effects of early school starting age appears a negative factor on the achievement of the overall students, however the early school starting age influences the achievement of both the 1st generation and 2nd generation immigrants positively. For example, the 1st generation immigrant students that start primary school at the age of 5 have higher score in the math domain by an amount of 0.419 standard deviation relative to starting age of 6 (see the column 2 in the Table 7). Moreover, the school starting age of 7 leads to

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Reading Score	Reading Score	Reading Score	Reading Score	Reading Score	Reading Score
Second Generation	-0.388***	-0.456	-0.447	-0.262	-0.563**	-0.482*
	(0.0222)	(0.280)	(0.280)	(0.280)	(0.265)	(0.266)
First Generation	-0.351***	-0.342***	-0.364***	-0.352***	-0.300***	-0.303***
	(0.0300)	(0.112)	(0.112)	(0.112)	(0.106)	(0.106)
Length of track	-0.00127	-0.000841	-0.00112	-0.00230*	-0.00111	-0.00103
	(0.00118)	(0.00118)	(0.00119)	(0.00119)	(0.00114)	(0.00114)
Time spent in education system	0.0549***	0.0503***	0.0710***	0.0691***	0.0990***	0.0994***
	(0.00334)	(0.00415)	(0.00946)	(0.00945)	(0.00897)	(0.00897)
Preprimary education	0.0160***	0.0161***	0.0164***	0.0223***	0.0427***	0.0423***
	(0.00224)	(0.00225)	(0.00225)	(0.00226)	(0.00222)	(0.00222)
School starting age≤5	-0.116***	-0.112***	-0.135***	-0.143***	-0.246***	-0.246***
	(0.00712)	(0.00766)	(0.0122)	(0.0122)	(0.0117)	(0.0117)
School starting age=7	0.310***	0.317***	0.337***	0.319***	0.383***	0.385***
	(0.00868)	(0.00914)	(0.0123)	(0.0123)	(0.0118)	(0.0118)
(School starting age≤5)*(1st generation)		0.335***	0.348***	0.327***	0.379***	0.363***
		(0.0887)	(0.0889)	(0.0889)	(0.0841)	(0.0842)
(School starting age=7)*(1st generation)		-0.383***	-0.388***	-0.337***	-0.395***	-0.398***
		(0.0642)	(0.0642)	(0.0642)	(0.0608)	(0.0608)
(School starting age≤5)*(2nd generation)		0.0253	0.0273	0.0382	0.0752	0.0663
		(0.0748)	(0.0748)	(0.0747)	(0.0707)	(0.0707)
(School starting age=7)*(2nd generation)		-0.168***	-0.167***	-0.134**	-0.139***	-0.145***
		(0.0548)	(0.0548)	(0.0548)	(0.0519)	(0.0519)
(Preprimary education)*(1st generation)		0.00879	0.00747	0.00753	0.0290	0.0110
		(0.0414)	(0.0414)	(0.0414)	(0.0392)	(0.0398)
(Preprimary education)*(2nd generation)		-0.0426	-0.0428	-0.0294	-0.0404	-0.0639**
		(0.0282)	(0.0282)	(0.0281)	(0.0267)	(0.0271)
(Time spent in education system)*(1st gen.)		0.0227***	0.0266***	0.0285***	0.0214***	0.0190***
		(0.00722)	(0.00752)	(0.00751)	(0.00711)	(0.00715)
(Time spent in education system)*(2nd gen.)		0.0482*	0.0473*	0.0302	0.0606**	0.0496*
		(0.0278)	(0.0278)	(0.0278)	(0.0263)	(0.0264)
(Length of track)*(2nd gen.)		-0.0519***	-0.0519***	-0.0440***	-0.0405***	-0.0382***
		(0.00976)	(0.00976)	(0.00975)	(0.00924)	(0.00925)
(Length of track)*(1st gen.)		-0.0391***	-0.0392***	-0.0251**	-0.0410***	-0.0365***
		(0.0127)	(0.0127)	(0.0127)	(0.0120)	(0.0121)
Age of arrival (0,4]			-0.0293*	-0.0120	-0.0349**	-0.0361**
			(0.0165)	(0.0165)	(0.0156)	(0.0156)
Age of arrival (4,11)			0.0298	0.0627***	0.121***	0.122***
			(0.0235)	(0.0235)	(0.0223)	(0.0223)
Age of arrival [11,13)			0.137**	0.195***	0.363***	0.364***
			(0.0642)	(0.0642)	(0.0609)	(0.0609)
Age of arrival [13,16)			0.196**	0.241***	0.548***	0.542***
			(0.0813)	(0.0812)	(0.0771)	(0.0771)
Test language at home				0.220***	0.142***	0.126***
				(0.00735)	(0.00696)	(0.00747)
Constant	-2.826***	-2.862***	-2.743***	-2.901***	-1.257***	-1.244***
	(0.0772)	(0.0785)	(0.0921)	(0.0922)	(0.0891)	(0.0891)
Observations	515,087	515,087	515,087	515,087	515,087	515,087
R-squared	0.271	0.271	0.271	0.272	0.349	0.349

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

NOTE: All regressions include controls for gender, parent's highest employment status, age of the student, extensive controls for parental socioeconomic status, and region of origins, year, and country level characteristics and school characteristics of the students.

Table 5: The regression results of institutions in the domain of reading

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Science Score	Science Score	Science Score	Science Score	Science Score	Science Score
Second Generation	-0.462***	-0.441	-0.432	-0.286	-0.591**	-0.497*
	(0.0221)	(0.279)	(0.279)	(0.279)	(0.264)	(0.265)
First Generation	-0.391***	-0.182	-0.214*	-0.204*	-0.186*	-0.190*
	(0.0298)	(0.111)	(0.111)	(0.111)	(0.106)	(0.106)
Length of track	-0.00760***	-0.00702***	-0.00748***	-0.00841***	-0.00635***	-0.00625***
	(0.00118)	(0.00118)	(0.00118)	(0.00118)	(0.00114)	(0.00114)
Time spent in education system	0.0442***	0.0386***	0.0729***	0.0714***	0.0968***	0.0973***
	(0.00332)	(0.00412)	(0.00940)	(0.00940)	(0.00894)	(0.00894)
Preprimary education	0.132***	0.133***	0.133***	0.138***	0.153***	0.152***
	(0.00222)	(0.00223)	(0.00224)	(0.00224)	(0.00221)	(0.00221)
School starting age≤5	0.0182**	0.0246***	-0.0135	-0.0204*	-0.110***	-0.110***
	(0.00708)	(0.00762)	(0.0122)	(0.0122)	(0.0117)	(0.0117)
School starting age=7	0.509***	0.515***	0.548***	0.534***	0.570***	0.572***
	(0.00862)	(0.00908)	(0.0122)	(0.0122)	(0.0118)	(0.0118)
(School starting age≤5)*(1st generation)		0.289***	0.309***	0.293***	0.344***	0.324***
		(0.0882)	(0.0884)	(0.0883)	(0.0839)	(0.0839)
(School starting age=7)*(1st generation)		-0.332***	-0.342***	-0.301***	-0.342***	-0.346***
		(0.0638)	(0.0638)	(0.0638)	(0.0606)	(0.0606)
(School starting age≤5)*(2nd generation)		-0.0790	-0.0759	-0.0673	-0.0247	-0.0353
		(0.0743)	(0.0743)	(0.0743)	(0.0705)	(0.0705)
(School starting age=7)*(2nd generation)		-0.262***	-0.260***	-0.234***	-0.233***	-0.240***
		(0.0545)	(0.0545)	(0.0545)	(0.0517)	(0.0517)
(Preprimary education)*(1st generation)		-0.0199	-0.0227	-0.0226	0.00171	-0.0230
		(0.0411)	(0.0412)	(0.0411)	(0.0391)	(0.0397)
(Preprimary education)*(2nd generation)		-0.151***	-0.151***	-0.141***	-0.150***	-0.178***
		(0.0280)	(0.0280)	(0.0280)	(0.0266)	(0.0270)
(Time spent in education system)*(1st gen.)		0.0245***	0.0302***	0.0317***	0.0257***	0.0226***
		(0.00718)	(0.00747)	(0.00747)	(0.00709)	(0.00712)
(Time spent in education system)*(2nd gen.)		0.0857***	0.0847***	0.0712***	0.100***	0.0874***
		(0.0276)	(0.0276)	(0.0276)	(0.0262)	(0.0263)
(Length of track)*(2nd gen.)		-0.0851***	-0.0849***	-0.0787***	-0.0726***	-0.0700***
		(0.00970)	(0.00970)	(0.00970)	(0.00921)	(0.00923)
(Length of track)*(1st gen.)		-0.0698***	-0.0700***	-0.0589***	-0.0726***	-0.0666***
		(0.0126)	(0.0126)	(0.0126)	(0.0120)	(0.0121)
Age of arrival (0,4]			-0.0412**	-0.0276*	-0.0510***	-0.0524***
			(0.0164)	(0.0164)	(0.0155)	(0.0155)
Age of arrival (4,11)			0.0492**	0.0752***	0.127***	0.129***
			(0.0234)	(0.0234)	(0.0222)	(0.0222)
Age of arrival [11,13)			0.200***	0.246***	0.392***	0.394***
			(0.0638)	(0.0638)	(0.0607)	(0.0607)
Age of arrival [13,16)			0.338***	0.373***	0.649***	0.640***
			(0.0808)	(0.0808)	(0.0768)	(0.0768)
Test language at home				0.174***	0.0977***	0.0772***
				(0.00730)	(0.00694)	(0.00745)
Constant	-3.432***	-3.484***	-3.289***	-3.413***	-1.698***	-1.682***
	(0.0767)	(0.0780)	(0.0916)	(0.0917)	(0.0888)	(0.0888)
Observations	515,087	515,087	515,087	515,087	515,087	515,087
R-squared	0.284	0.284	0.284	0.285	0.357	0.357

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

NOTE: All regressions include controls for gender, parent's highest employment status, age of the student, extensive controls for parental socioeconomic status, and region of origins, year, and country level characteristics and school characteristics of the students.

Table 6: The regression results of institutions in the domain of science

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Math Score	Math Score	Math Score	Math Score	Math Score	Math Score
Second Generation	-0.392***	-0.322	-0.296	-0.198	-0.520**	-0.440*
	(0.0218)	(0.276)	(0.276)	(0.276)	(0.259)	(0.260)
First Generation	-0.363***	-0.187*	-0.234**	-0.228**	-0.193*	-0.196*
	(0.0295)	(0.110)	(0.110)	(0.110)	(0.104)	(0.104)
Length of track	0.00610***	0.00652***	0.00566***	0.00503***	0.00717***	0.00725***
	(0.00116)	(0.00116)	(0.00117)	(0.00117)	(0.00112)	(0.00112)
Time spent in education system	0.0427***	0.0444***	0.106***	0.105***	0.131***	0.131***
	(0.00328)	(0.00408)	(0.00930)	(0.00930)	(0.00878)	(0.00878)
Preprimary education	0.161***	0.162***	0.163***	0.166***	0.181***	0.181***
	(0.00220)	(0.00221)	(0.00221)	(0.00222)	(0.00217)	(0.00217)
School starting age≤5	-0.164***	-0.168***	-0.236***	-0.241***	-0.337***	-0.336***
	(0.00700)	(0.00754)	(0.0120)	(0.0120)	(0.0115)	(0.0115)
School starting age=7	0.491***	0.500***	0.559***	0.549***	0.592***	0.593***
	(0.00853)	(0.00899)	(0.0121)	(0.0121)	(0.0115)	(0.0115)
(School starting age≤5)*(1st generation)		0.419***	0.457***	0.446***	0.485***	0.469***
		(0.0872)	(0.0874)	(0.0874)	(0.0823)	(0.0823)
(School starting age=7)*(1st generation)		-0.264***	-0.279***	-0.252***	-0.315***	-0.319***
		(0.0631)	(0.0631)	(0.0631)	(0.0595)	(0.0595)
(School starting age≤5)*(2nd generation)		-0.00617	0.000402	0.00623	0.0492	0.0405
		(0.0735)	(0.0735)	(0.0735)	(0.0692)	(0.0692)
(School starting age=7)*(2nd generation)		-0.190***	-0.189***	-0.171***	-0.177***	-0.183***
		(0.0539)	(0.0539)	(0.0539)	(0.0507)	(0.0507)
(Preprimary education)*(1st generation)		-0.00634	-0.00769	-0.00766	0.0169	0.000396
		(0.0407)	(0.0407)	(0.0407)	(0.0383)	(0.0389)
(Preprimary education)*(2nd generation)		-0.173***	-0.173***	-0.166***	-0.172***	-0.195***
		(0.0277)	(0.0277)	(0.0277)	(0.0261)	(0.0265)
(Time spent in education system)*(1st gen.)		0.00506	0.0104	0.0114	0.00529	0.00307
		(0.00710)	(0.00739)	(0.00739)	(0.00696)	(0.00699)
(Time spent in education system)*(2nd gen.)		0.0673**	0.0645**	0.0553**	0.0858***	0.0750***
		(0.0273)	(0.0273)	(0.0273)	(0.0257)	(0.0258)
(Length of track)*(2nd gen.)		-0.0510***	-0.0507***	-0.0465***	-0.0402***	-0.0379***
		(0.00959)	(0.00959)	(0.00959)	(0.00904)	(0.00905)
(Length of track)*(1st gen.)		-0.0413***	-0.0423***	-0.0348***	-0.0503***	-0.0461***
		(0.0124)	(0.0125)	(0.0125)	(0.0117)	(0.0118)
Age of arrival (0,4]			-0.0155	-0.00636	-0.0328**	-0.0339**
			(0.0162)	(0.0162)	(0.0152)	(0.0152)
Age of arrival (4,11)			0.0903***	0.108***	0.162***	0.163***
			(0.0231)	(0.0231)	(0.0218)	(0.0218)
Age of arrival [11,13)			0.415***	0.446***	0.599***	0.600***
			(0.0632)	(0.0632)	(0.0595)	(0.0595)
Age of arrival [13,16)			0.574***	0.597***	0.886***	0.880***
			(0.0799)	(0.0799)	(0.0754)	(0.0754)
Test language at home				0.117***	0.0385***	0.0224***
				(0.00723)	(0.00681)	(0.00731)
Constant	-3.814***	-3.810***	-3.463***	-3.547***	-1.736***	-1.724***
	(0.0759)	(0.0772)	(0.0906)	(0.0907)	(0.0871)	(0.0871)
Observations	515,087	515,087	515,087	515,087	514,379	514,379
R-squared	0.301	0.301	0.301	0.302	0.382	0.382

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

NOTE: All regressions include controls for gender, parent's highest employment status, age of the student, extensive controls for parental socioeconomic status, and region of origins, year, and country level characteristics and school characteristics of the students.

Table 7: The regression results of institutions in the domain of math

a decrease of a 0.264 standard deviation in math score with reference to the column 2 in the Table 7 relative to age of 6. In the same way, this patterns hold in both science and reading scores even though we control for the factors of language, age of arrival to the host country and school characteristics. Moreover, the findings are the same in the case of 2nd

generation immigrants. For instance, with reference to the column 2 in the Table 5 if a second generation migrant student starts primary school at the age of 7 the outcome of this student in reading decreases by 0.168 standard deviation relative to native peers starting age of 7. Apart from reading scores, the same tendency holds in two other domains of interest in spite of controlling for language, age of arrival to host country and school characteristics. The intuition behind this relies on the fact that early school starting age mitigates the effect of disadvantaged parental backgrounds for several reasons. In Early entry may also help immigrants increase cognitive capacities and expose the children to the educational system of the host country and this finding is consistent with the previous literature Currie and Thomas (1999), Cobb-Clark et al. (2012) and Heckman (2006).

In addition to the school starting age, we are interested in the effect of preprimary education on the student achievements. One extra year preprimary education increases math scores of an average student by a 0.161 standard deviation (see the column 1 in the Table 7), 0.132 standard deviation in the case of science scores (see the column 1 in the Table 6) and 0.0160 standard deviation in the case of reading scores (see the column 1 in the Table 5). The main effect of the preprimary education for both natives and the immigrants are still in the same direction and statistically significant in all three domains of interest when we control for language, age of arrival to host country and school characteristics. On the whole, these findings for the effect of preprimary education are consistent with the previous literature Schneeweis (2010), Datar (2006), Lubotsky (2009), Curie (2001) and Deming et al. (2008). However, we also generate the interaction of the preprimary education variable with the immigration status. We find no additional effect for the 1st generation immigrant students. For the scores of 2nd generation immigrant students, however, interaction term indicates a negative influence in both the science and math sections. Moreover, there is no effect on the reading scores of the 2nd generation immigrants since the coefficient is statistically significant. The total effect which is the sum of main and interaction effect are positive or slightly negative for the immigrants. For instance, the 1st generation immigrants in reading scores benefit from preprimary education when we consider the Tables 5, 6 and 7 with reference to the column 2. However, the preprimary education is affecting the 2nd generation immigrants negatively (see the column 2 in the Table 5, 6 and 7). The intuition behind this fact may be that these immigrants do not have access to preprimary education or that they drop out for several reasons.

It is popularly believed that the preprimary education increases the educational outcome of the immigrant students for several reasons such as increasing language skills and cognitive skills especially for the 2nd generation immigrants. Nonetheless, our data for preprimary education is at the country level and therefore we do not exactly know whether the immigrants attend to preprimary education for all three waves of data used in the analysis. Luckily, the PISA 2003 and 2009 have the data on preprimary attendance of students and thus we can use this variable instead of using a country level variable. Although the results are same for the 1st generation immigrant, for the 2nd generation immigrants the negative interaction terms positive insignificant for math and science but significant for reading as shown in Tables 8, 9 and 10 with reference to the column 2 for all tables.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Science Score	Science Score	Science Score	Science Score	Science Score
Second Generation	-0.139	-1.840***	-1.106*	-0.935	-1.213**
	(0.0880)	(0.584)	(0.593)	(0.593)	(0.567)
First Generation	-0.133	-0.444***	-0.0238	-0.0572	0.0423
	(0.0866)	(0.104)	(0.131)	(0.131)	(0.126)
Length of track	-0.000472	-5.97e-05	-0.000451	-0.00119	0.00314**
	(0.00152)	(0.00152)	(0.00153)	(0.00153)	(0.00147)
School starting age≤5	0.113***	0.125***	0.105***	0.103***	0.0516***
	(0.00883)	(0.00943)	(0.0156)	(0.0156)	(0.0151)
School starting age=7	0.498***	0.502***	0.520***	0.506***	0.558***
	(0.0112)	(0.0118)	(0.0160)	(0.0161)	(0.0155)
Preprimary education attendance in years	0.105***	0.104***	0.105***	0.107***	0.0910***
	(0.00305)	(0.00308)	(0.00308)	(0.00308)	(0.00301)
Time spent in education system	0.0340***	0.0267***	0.0431***	0.0411***	0.0614***
	(0.00428)	(0.00521)	(0.0125)	(0.0124)	(0.0119)
Time spent in education system)*(1st gen.)		0.0222**	0.0247**	0.0267***	0.0264***
		(0.00939)	(0.00985)	(0.00985)	(0.00940)
(Time spent in education system)*(2nd gen.)		0.179***	0.135**	0.116*	0.157***
		(0.0615)	(0.0620)	(0.0620)	(0.0593)
(Preprimary education)*(2nd generation)		0.000130	0.0252	0.0229	0.0200
		(0.0351)	(0.0353)	(0.0353)	(0.0337)
(Preprimary education)*(1st generation)		0.127***	0.0735**	0.0617*	0.0916***
		(0.0336)	(0.0344)	(0.0344)	(0.0328)
(School starting age≤5)*(1st generation)		0.369***	-0.0836	-0.0720	-0.0185
		(0.0801)	(0.127)	(0.126)	(0.121)
(School starting age=7)*(1st generation)		-0.595***	-0.405***	-0.376***	-0.443***
		(0.0735)	(0.0750)	(0.0750)	(0.0716)
(School starting age≤5)*(2nd generation)		0.134	-0.217	-0.180	-0.213
		(0.127)	(0.160)	(0.160)	(0.153)
(School starting age=7)*(2nd generation)		-0.177**	-0.0914	-0.0890	-0.119
		(0.0819)	(0.0840)	(0.0840)	(0.0803)
Length of track)*(2nd gen.)		-0.0792***	-0.0740***	-0.0692***	-0.0768***
		(0.0153)	(0.0176)	(0.0176)	(0.0168)
(Length of track)*(1st gen.)		-0.0430***	-0.0802***	-0.0718***	-0.0834***
		(0.0131)	(0.0139)	(0.0139)	(0.0133)
Age of arrival (0,4]			-0.00962	0.00138	-0.0490**
			(0.0210)	(0.0210)	(0.0201)
Age of arrival (4,11)			0.0697**	0.0897***	0.127***
			(0.0303)	(0.0303)	(0.0289)
Age of arrival [11,13)			0.0895	0.125	0.242***
			(0.0831)	(0.0831)	(0.0794)
Age of arrival [13,16)			0.131	0.163	0.391***
			(0.105)	(0.105)	(0.100)
Test language at home				0.144***	0.0627***
				(0.00950)	(0.00909)
Constant	-3.064***	-3.111***	-3.031***	-3.136***	-1.560***
	(0.0962)	(0.0980)	(0.117)	(0.117)	(0.114)
Observations	328,401	328,401	328,401	328,401	328,401
R-squared	0.273	0.273	0.274	0.274	0.340

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

NOTE: All regressions include controls for gender, parent's highest employment status, age of the student, extensive controls for parental socioeconomic status, and region of origins, year, and country level characteristics and school characteristics of the students.

Table 8: The regression results of new preprimary variable in the domain of science

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Math Score	Math Score	Math Score	Math Score	Math Score
Second Generation	-0.0943	-1.059*	-0.506	-0.416	-0.702
	(0.0872)	(0.578)	(0.587)	(0.587)	(0.558)
First Generation	-0.125	-0.315***	0.0365	0.0189	0.146
	(0.0858)	(0.103)	(0.130)	(0.130)	(0.123)
Length of track	0.0165***	0.0168***	0.0162***	0.0158***	0.0194***
	(0.00151)	(0.00151)	(0.00152)	(0.00152)	(0.00145)
School starting age≤5	-0.0440***	-0.0407***	-0.0779***	-0.0794***	-0.135***
	(0.00875)	(0.00934)	(0.0155)	(0.0155)	(0.0148)
School starting age=7	0.455***	0.462***	0.495***	0.488***	0.554***
	(0.0111)	(0.0117)	(0.0159)	(0.0159)	(0.0152)
Preprimary education attendance in years	0.140***	0.139***	0.140***	0.141***	0.126***
	(0.00302)	(0.00305)	(0.00305)	(0.00305)	(0.00296)
Time spent in education system	0.0363***	0.0348***	0.0673***	0.0663***	0.0876***
	(0.00424)	(0.00516)	(0.0123)	(0.0123)	(0.0117)
Time spent in education system)*(1st gen.)		0.00652	0.00444	0.00550	0.00493
		(0.00930)	(0.00976)	(0.00976)	(0.00925)
(Time spent in education system)*(2nd gen.)		0.0910	0.0613	0.0511	0.0950
		(0.0609)	(0.0614)	(0.0614)	(0.0583)
(Preprimary education)*(2nd generation)		-0.00756	0.0186	0.0173	0.00736
		(0.0348)	(0.0350)	(0.0350)	(0.0332)
(Preprimary education)*(1st generation)		0.0509	0.00444	-0.00176	0.0269
		(0.0333)	(0.0341)	(0.0341)	(0.0323)
(School starting age≤5)*(1st generation)		0.624***	0.188	0.194	0.242**
		(0.0793)	(0.125)	(0.125)	(0.119)
(School starting age=7)*(1st generation)		-0.566***	-0.401***	-0.386***	-0.476***
		(0.0728)	(0.0743)	(0.0743)	(0.0705)
(School starting age≤5)*(2nd generation)		0.490***	0.0925	0.112	0.0777
		(0.126)	(0.158)	(0.158)	(0.150)
(School starting age=7)*(2nd generation)		-0.184**	-0.115	-0.114	-0.158**
		(0.0811)	(0.0832)	(0.0832)	(0.0789)
Length of track)*(2nd gen.)		-0.0315**	-0.0342*	-0.0316*	-0.0402**
		(0.0152)	(0.0174)	(0.0174)	(0.0165)
(Length of track)*(1st gen.)		-0.00139	-0.0379***	-0.0334**	-0.0476***
		(0.0130)	(0.0137)	(0.0137)	(0.0131)
Age of arrival (0,4]			0.0201	0.0259	-0.0294
			(0.0208)	(0.0208)	(0.0197)
Age of arrival (4,11)			0.103***	0.114***	0.153***
			(0.0300)	(0.0300)	(0.0285)
Age of arrival [11,13)			0.226***	0.245***	0.367***
			(0.0823)	(0.0823)	(0.0781)
Age of arrival [13,16)			0.254**	0.271***	0.508***
			(0.104)	(0.104)	(0.0985)
Test language at home				0.0760***	-0.00714
				(0.00941)	(0.00894)
Constant	-3.412***	-3.415***	-3.242***	-3.298***	-1.670***
	(0.0953)	(0.0971)	(0.116)	(0.116)	(0.112)
Observations	328,401	328,401	328,401	328,401	328,401
R-squared	0.283	0.282	0.283	0.283	0.358

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

NOTE: All regressions include controls for gender, parent's highest employment status, age of the student, extensive controls for parental socioeconomic status, and region of origins, year, and country level characteristics and school characteristics of the students.

Table 9: The regression results of new preprimary variable in the domain of math

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Reading Score	Reading Score	Reading Score	Reading Score	Reading Score
Second Generation	-0.108	-1.075*	-0.493	-0.245	-0.568
	(0.0883)	(0.586)	(0.595)	(0.594)	(0.567)
First Generation	-0.166*	-0.515***	-0.180	-0.228*	-0.0823
	(0.0868)	(0.104)	(0.132)	(0.132)	(0.125)
Length of track	-0.00288*	-0.00243	-0.00280*	-0.00388**	0.000394
	(0.00153)	(0.00153)	(0.00154)	(0.00153)	(0.00147)
School starting age≤5	-0.102***	-0.0923***	-0.112***	-0.116***	-0.177***
	(0.00886)	(0.00945)	(0.0157)	(0.0157)	(0.0150)
School starting age=7	0.303***	0.306***	0.323***	0.304***	0.383***
	(0.0113)	(0.0118)	(0.0161)	(0.0161)	(0.0154)
Preprimary education attendance in years	0.0390***	0.0371***	0.0376***	0.0412***	0.0347***
	(0.00306)	(0.00309)	(0.00309)	(0.00309)	(0.00300)
Time spent in education system	0.0511***	0.0440***	0.0605***	0.0577***	0.0816***
	(0.00429)	(0.00523)	(0.0125)	(0.0125)	(0.0119)
Time spent in education system)*(1st gen.)		0.0281***	0.0301***	0.0330***	0.0312***
		(0.00942)	(0.00988)	(0.00988)	(0.00939)
(Time spent in education system)*(2nd gen.)		0.0832	0.0465	0.0185	0.0683
		(0.0617)	(0.0622)	(0.0621)	(0.0592)
(Preprimary education)*(2nd generation)		0.0431	0.0699**	0.0665*	0.0522
		(0.0352)	(0.0354)	(0.0354)	(0.0337)
(Preprimary education)*(1st generation)		0.167***	0.119***	0.102***	0.123***
		(0.0337)	(0.0345)	(0.0345)	(0.0328)
(School starting age≤5)*(1st generation)		0.538***	0.161	0.178	0.237**
		(0.0803)	(0.127)	(0.127)	(0.121)
(School starting age=7)*(1st generation)		-0.592***	-0.410***	-0.367***	-0.461***
		(0.0737)	(0.0753)	(0.0752)	(0.0716)
(School starting age≤5)*(2nd generation)		0.467***	0.179	0.234	0.186
		(0.128)	(0.160)	(0.160)	(0.152)
(School starting age=7)*(2nd generation)		-0.186**	-0.0998	-0.0964	-0.138*
		(0.0821)	(0.0843)	(0.0842)	(0.0802)
Length of track)*(2nd gen.)		-0.0297*	-0.0246	-0.0175	-0.0298*
		(0.0154)	(0.0177)	(0.0176)	(0.0168)
(Length of track)*(1st gen.)		-0.0371***	-0.0728***	-0.0605***	-0.0749***
		(0.0132)	(0.0139)	(0.0139)	(0.0133)
Age of arrival (0,4]			-0.00379	0.0122	-0.0374*
			(0.0210)	(0.0210)	(0.0200)
Age of arrival (4,11)			0.0583*	0.0873***	0.129***
			(0.0304)	(0.0304)	(0.0289)
Age of arrival [11,13)			0.105	0.156*	0.288***
			(0.0833)	(0.0833)	(0.0793)
Age of arrival [13,16)			0.132	0.179*	0.430***
			(0.105)	(0.105)	(0.100)
Test language at home				0.209***	0.122***
				(0.00953)	(0.00908)
Constant	-2.658***	-2.701***	-2.617***	-2.769***	-1.254***
	(0.0965)	(0.0983)	(0.118)	(0.118)	(0.114)
Observations	328,401	328,401	328,401	328,401	328,401
R-squared	0.261	0.260	0.261	0.262	0.334

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

NOTE: All regressions include controls for gender, parent's highest employment status, age of the student, extensive controls for parental socioeconomic status, and region of origins, year, and country level characteristics and school characteristics of the students.

Table 10: The regression results of new preprimary variable in the domain of reading

Moreover, these results are robust when we control for family background, language and school characteristics. Taking everything into account, it is likely that the preprimary education in the host country is a crucial factor on the achievement of immigrant students, a conclusion which differentiates our study from the previous literature that do not study this institutional arrangement such as Cobb-Clark et al. (2012), Dronkers and Velden (2012), Horn

(2012), Ludemann and Schwerdt (2010), Entorf and Miniou (2005), Schnepf (2007) and Schneeweis (2010).

The existence of tracking in the system as an institutional structure may also affect the achievement of students. With the intention of investigating the effect of tracking system we generate a track variable containing the time that a student spends in a track until the age of 15. One advantage of considering the total time spent in a track is that of observing the continuing permanent effect of tracking due to initial selection in time. Initially, the previous literature such as Waldinger (2006), Schuetz et al. (2008), Cobb-Clark et al. (2012), Horn (2012), and Hanushek and Woessmann (2006) considers the age of selection or early age of selection that ignores the time effect of tracking. Therefore, we carry out our analysis by using the time spent in a track. To put it another way, our expectation for the track variable is that the longer time the student spends in a track the larger will be the effects of track. Considering the regression results in Table 5, 6 and 7 for the column 1, the main effect of tracking is positive not considering the immigration status of a student. If a student spends one extra year in a track, the score in the math section is relatively a 0.00610 standard deviation higher (see the column1 in Table 7). In addition, taking into account family background, language and school characteristics, and the positive effect of spending one extra year in a track is still valid. In the science performance, the tracking has a negative effect even after we control for other variables. Interestingly, tracking has no significant effect on reading performance of a student with reference to the column1 in the Table 5. However when we control for the language variable it has a negative effect. While it is true to say that tracking has no consistent affect in all three domains of interest, in fact it results in a penalty for the immigrants in all domains. When we generate the interaction between the track variable and the status of immigration, we reach a conclusion that spending an extra one year in any track for both the 1^{st} generation and the 2^{nd} generation immigrants affects negatively their performance relative to the native ones. This effect of tracking on the immigrant population is consistent and statistically significant even after controlling for family background, language and school characteristics. According to Table 7, being a 2^{nd} generation immigrant results in a 0.00510 lower standard deviation performance in math and the penalty for being a 1^{st} generation immigrant is a 0.0413 lower standard deviation performance in the same domain. The trend in math performance for the immigrant is the same in both the reading and the science performance. That is to say, with reference to Table 5, 6 and 7 if a student is a 1^{st} generation immigrant then this student is exposed to a 0.0365 standard deviation penalty in

reading and a 0.0666 lower standard deviation score in science after controlling for family background, language and school characteristics. Moreover, the 2nd generation immigrant performs 0.0382 standard deviation lower in the reading domain and 0.07 standard deviation lower in science by controlling for family background, language and school characteristics (see Table 5 and 6). Unlike the previous studies of Cobb-Clark et al. (2012) claiming that the limited tracking[10] on ability is advantageous for the immigrant students. Our results for tracking reveals that the inequality between the immigrant and the natives increases as time spent in track increases and this result is robust to controlling school, family and language variables. The reason why we get a different result from Cobb-Clark et al. (2012) is that they use a different tracking measure focusing on the ability grouping[11] in the current school. However, we are interested in the effect of previous tracking mechanism on the current school success of a student. In addition, Cobb-Clark et al. (2012) investigates the effect of age of first selection which is a different construction of our tracking measure. Our results are consistent with their findings and more robust since after controlling for several variables such as school characteristics it is still significant. Our results reveal that tracking increases the inequality between immigrant and native students against the immigrant in education. The intuition behind this relies on the fact that immigrants are more likely to fall in lower tracks for several reasons. For instance, the lower family background conditions of immigrants relative to natives and lack of language ability in the host country results in lower tracks for immigrant students. Moreover, the family background effect for the immigrant students can increase the negative effects of tracking on them in the time of tracking as discussed in the previous literature Ludemann and Schwerdt (2010), Horn (2012), and Brunello and Checchi(2007)). The previous selection -tracking in our study- makes the students in lower tracks deprived of sufficient science or math education for several reasons. For instance, in Turkey the students that drop into vocational track suffer from insufficient science and math education. As a result, the time spent in the tracking may increase the gap between the natives and immigrants because of the reason as in the case of Turkey. Therefore, the lower track schools may not be as efficient as the higher track schools to educate the immigrant students. As a result, the immigrant students may have lower scores. In our data set, the 1st generation immigrant students have different age of arrival to host country. As a consequence, the familiarity of

[10] Ability grouping in current school for some subjects.

[11] We also control for ability grouping and the effect of ability grouping for both all subjects and some subjects is negative unlike the positive effect for some subjects in Cobb-Clark et al. (2012). This difference may be based on the fact that we control for the school resources and other institutional arrangements related to previous selection of students.

these students and their families to the host country's educational structure may vary by age of arrival. The familiarity to the educational structure may force the parents to be active in their educational investment for their siblings. Particularly, a parent may send her sibling to some special courses for her child to decrease the negative effect of tracking or to settle her child into a high quality school. In addition, we expect that the more years a student spends in the education system the more the student performs since the more years increase the adaptation of the student to the host country education system. With the purpose of understanding the effect of familiarity we carry out an analysis including a variable that indicates how many years a student spends in the educational system of a host country. The results are summarized in Table 5, 6 and 7. First of all, the main effect of staying one extra year in the education system is positive in all domains of interest (see the column 1 for all tables). For instance, the increase due to staying one extra year in the education system is a 0.0427 standard deviation in math, a 0.0549 standard deviation in reading and 0.0442 in science. The findings are robust when we control for family background, language and school characteristics. Secondly, the specific effect of staying one extra in the education of host country for the immigrants is positive. In other words the immigrant peers benefit more than the native born peers in general. For the 1st generation immigrant students, spending one extra year in the host country's education system increases their performance by 0.0227 standard deviation in reading and 0.0245 standard deviation in science however there is not significant effect in math score (see the column 2 for all tables). Regardless of the fact that there is no effect in reading for the 1st generation immigrant students, the total effect of staying in a host country's education system leads to a 0.149 standard deviation increase in reading and 0.119 in science respectively. In brief, 1st generation immigrants benefit more from spending more time in the host country's education system relative to the native born peers even when we control for family background, language and school characteristics (see Table 5, 6 and 7). Furthermore, the effect of staying one extra year in a host country's education system leads to additional increase in all domains for 2nd generation immigrant students. Indeed, 2nd generation immigrant students relative to native-born counterparts perform better in all three sections as shown in Table 5, 6 and 7. For example, the 2nd generation immigrant students perform a 0.0673 standard deviation more in math scores relative to native-born counterparts (see the column 2 in the Table 7). Moreover this finding is still robust and significant while controlling the school characteristic and language variable. The same pattern is valid in science and reading scores as summarized in Table 5 and 6. As a result, staying more in the education system of a host country increases the performance of immigrant students and they

benefit more than the native born ones. Thus the inequality between the immigrant and the natives decreases. To sum up, our findings suggest that the gap between the immigrants and native students as well as for better integration of immigrants to host countries may be decreased to minimum level if the host countries consider more preprimary education, early school starting age and more comprehensive education system in their education system. As a result, the importance of our study is that how a host country increases the education outcomes of immigrant students by revising its education system.

4.3 Age of arrival, Language and Country Specific Characteristics

Although our main interest in this study is to understand the role of educational institutions on the educational achievement of immigrant population, we also conduct our analysis to measure the effect of school characteristics, language spoken at home, age of arrival to host country and some country development indices. First of all, the age of arrival to a host country is one of the major indicators of educational achievements for immigrant students, with the best results observed at age 11, which is intuitive from the psychology literature. Initially, one may expect that the scores should be higher as the age of arrival decreases. However, the age of arrival both includes familiarity to the educational system argument and language ability thus the language ability component may be the dominant over the institutional and environmental familiarity. In fact, the psychology literature talks about "critical period" for second language ability. Snow et al. (1978) argue that the period of age of 12-15 is the most efficient time to learn a second language. By controlling for school start age our findings support this argument since our results show that the age of arrival of 11 and higher increases the scores. However, the effect of age of arrival turns to reverse direction – negative effect on scores- if we do not control for the school starting age. One potential explanation may be that the language ability is hidden in the age of arrival as in the case of Snow et al. (1978). The second explanation may be the fact that immigrants do not have sufficient information about the educational system in a host country. Controlling for school start age may represent knowledge of educational system and therefore the language hypothesis of Snow et al. (1978) occurs in our results.

Likewise, speaking the host country's language at home increases the performance of the migrants relative to ones who do not speak as shown in Table 5, 6 and 7. Another crucial finding is that all students, whether migrant or not, perform higher in the schools with higher average family background (summarized in Table 11).

	(1)	(2)	(3)
VARIABLES	Math Score	Science Score	Reading Score
Vocational School	-0.263***	-0.261***	-0.254***
	(0.00359)	(0.00366)	(0.00367)
School FB mean	0.596***	0.556***	0.582***
	(0.00275)	(0.00280)	(0.00281)
School standard deviation FB	0.00723	-0.0277***	0.0140*
	(0.00755)	(0.00770)	(0.00772)
Student-teacher ratio	-0.00163***	-0.00173***	-0.00141***
	(0.000159)	(0.000163)	(0.000163)
Teacher shortage	-0.0245***	-0.0158***	-0.0187***
	(0.00127)	(0.00129)	(0.00130)
Responsibility for curriculum & assessment index	-0.0175***	-0.0277***	-0.0235***
	(0.00124)	(0.00127)	(0.00127)
Shortage of instruction material (very little)	-0.0130***	-0.0167***	0.00523*
	(0.00282)	(0.00287)	(0.00288)
Shortage of instruction material (to some extent)	-0.0232***	-0.0450***	-0.0232***
	(0.00333)	(0.00340)	(0.00341)
Shortage of instruction material (A lot)	-0.00242	-0.0136**	-0.000178
	(0.00539)	(0.00550)	(0.00551)
School type (Private-Government)	0.116***	0.0921***	0.125***
	(0.00476)	(0.00486)	(0.00487)
School type (Public)	0.0797***	0.0839***	0.114***
	(0.00384)	(0.00392)	(0.00393)
Ability Grouping (For Some Subject)	-0.0736***	-0.0643***	-0.0601***
	(0.00274)	(0.00279)	(0.00280)
Ability Grouping (For All Subject)	-0.0670***	-0.0769***	-0.0533***
	(0.00334)	(0.00341)	(0.00342)
Human Development Index	0.0591***	0.0517***	0.0432***
	(0.00133)	(0.00136)	(0.00136)
Gini	-0.127***	-0.0507***	-0.163***
	(0.00301)	(0.00307)	(0.00308)
Teacher pupil ratio	0.0139***	0.00913***	0.00515***
	(0.000748)	(0.000763)	(0.000765)
Public expenditure on education percent of GDP	-0.0184***	-0.119***	0.0392***
	(0.00361)	(0.00368)	(0.00369)
Teacher salary	1.02e-05***	5.26e-06***	1.08e-05***
	(2.08e-07)	(2.12e-07)	(2.13e-07)
Constant	-1.736***	-1.698***	-1.257***
	(0.0871)	(0.0888)	(0.0891)
Observations	515,087	515,087	515,087
R-squared	0.382	0.357	0.349

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

NOTE: All regressions include controls for gender, parent's highest employment status, age of the student, extensive controls for parental socioeconomic status, and region of origins, year, and institutional arrangements and status of immigration, and age of arrival and language.

Table 11: The regression results for school characteristics and country level variables

Students in vocational schools have lower scores than students attending general schools which may be intuitive due to the reasons of low educational resources. As the student teacher ratio increases the performance of a student decreases because the instructional efficiency probably decreases. Likewise, this argument is valid for the shortage of instructional materials in a school as this hurts the performance of the students. What is more, the ability grouping in a school relative to no ability grouping leads to lower scores as the intensity of ability grouping increases. With reference to (Cobb-Clark et al. (2012) and Korthals (2012)), this finding is consistent with the previous studies.

Finally, students perform better in the countries where HDI, teacher-pupil ratio, teacher salary and public expenditure on education percent of GDP is high. However, they perform worse in the countries where the Gini coefficient is higher. Therefore, the educational achievement of both immigrant and native students is related to the wider social context as in the case of Gini coefficient findings as suggested in Schneeweis(2010).

5. Conclusion

We observe substantial heterogeneity in the educational achievement of immigrant students in OECD countries. Our goal in this study is to explain the reasons for this variation therefore we conduct an empirical analysis by using three waves of PISA tests. Our main consideration is that how much the role of institutional arrangement in the host countries is responsible of the variation in the educational achievement of immigrant population. While most of the literature on educational achievement of immigrant students is concerned with the causal effects of the institutional arrangements in the education system of host countries, they ignore the difference between the immigrants across countries due to their region of origin. We contribute the literature in the following two ways: firstly, we control the region of origin for the immigrant students and explain the cross-country difference between the immigrant groups by using the region of origin fixed effects. For instance, the tracking has no effect on 1st generation immigrants if we do not control for region of origin. However, if we control for region of origin the tracking has negative significant effect on immigrants. Secondly, we use in this study more comprehensive measure of institutional arrangements. In particular we generate new measures taking into account familiarity to host country's education system and time spent in an institutional arrangement. Moreover, we also control for school characteristics and how children are allocated to schools.

We find that achievement gaps are wider for the immigrant students in the host countries where the school starting age is late. The early starting school age is beneficial for both the 1st and 2nd generation immigrant students after controlling for school characteristics, language and region of origins. This finding is intuitively consistent owing to the fact that the early starting school age is increasing the cognitive and language abilities of the immigrant students. Moreover, the inequality between the immigrants and native-born counterparts decreases if the expected duration of preprimary education increases. The investigation of the effects of preprimary education on immigrant students is one of the contributions of this study. The tracking system results in the increase of inequality between the immigrant and native students. In other words the more the longer is the duration of track the more the immigrant student underperforms relative to native ones. We also find that the time spent in the educational system of a host country is a crucial determinant on the achievement of immigrant students and it affects their performance positively. This result suggests that performance of immigrants increase as familiarity with educational system increases. The gap between immigrants and natives is increasing if the age of arrival to host country increases.

Thus, the emigration in early ages for immigrant students is beneficial for their educational performance. In addition, the language spoken at home has also notable effect on the immigrant performance as well as country specific effects such as HDI, education spending, pupil-teacher ratio and teacher salary rates. However, the income inequality harms the performance of immigrant students.

Based on the findings of our study policymakers should take into account the institutional arrangements in education for the integration of immigrants to the host country. In particular, the school starting age, tracking and preprimary education can be rearranged to mitigate the negative effects of institutional arrangements on immigrant students. All in all, the immigrant students may benefit from the education system in host countries if the education system of these host countries has the structure of more preprimary education, early school starting age and more comprehensive educational arrangements rather than including tracking.

Appendix A: Distribution of the institutional arrangements and country level variables

COUNTRY	Duration of expected preprimary education	Time spent in a track	Primary School Starting Age	Gini	HDI	Public expenditure for health (% GDP)	Employment ratio	Education index	Pupil-teacher ratio	Education share in government expenditure	Public expenditure for education (% GDP)	Teacher salary
Australia	1.64434	0.0343839	5	5.5667	62.3	61.4333	0.9803	0.978	23.7	14.176	4.7488	47445
Austria	2.32851	5.82726	6	7.8	58.6333	57.3	0.8507	0.8217	18.9	11.0514	5.4113	40818
Belgium	2.98341	3.854408	6	7.2	53.4667	52.7667	0.8863	0.8757	11.6294	12.2675	6.0852	44076
Canada	1.566	0.0552045	6	6.8667	62.9333	63.3667	0.9087	0.9107	17.416	12.3	4.8737	54978
Czech Rep.	2.40113	4.864321	6	5.8333	60.1333	60.2333	0.9203	0.921	17.1403	9.8397	4.1161	19949
Denmark	2.74448	0.0231036	7	8.1	61.6667	63.2	0.9177	0.913	10.0276	15.3751	7.9404	50253
Estonia	2.669	0.8102749	7	3.8333	58.7667	61.6333	0.9177	0.92	12.7995	14.2312	5.0695	12576
Finland	2.22858	0.016841	7	6.0333	58.7	58.9	0.8783	0.8787	15.6614	12.5294	6.1215	37455
Germany	2.77993	5.82286	6	8	56.2667	54.3667	0.937	0.9237	14.199	10.1235	4.4956	55771
Greece	1.40482	0.7072798	6	5.9	52.5667	52.6	0.853	0.8503	11.7194	9.2236	4.0911	32387
Hungary	2.61203	4.724611	7	5.6667	49.9333	51.1	0.8887	0.8837	10.3056	10.5389	5.3218	13228
Iceland	2.87785	0.0206788	6	7.6667	72.6333	75.2667	0.9077	0.8937	17.8	17.7877	7.5141	27930
Ireland	1.69395	0.7037104	4	5.8	58.5667	62.3667	0.9617	0.9467	17.9257	13.7069	4.9991	53677
Italy	2.92354	1.734035	6	6.6667	47.9667	48.2667	0.8473	0.8307	10.6333	9.313	4.4852	32658
Japan	2.69328	0.7869273	6	6.5667	59.8333	60.4	0.883	0.8707	19.356	9.4518	3.4626	44788
Korea	2.38844	1.742143	6	3.2333	64.9667	65.6667	0.9343	0.911	28.7487	15.2696	4.3493	46338
Luxembourg	2.57936	2.822874	6	6.1667	58.7667	57.5667	0.776	0.7713	11.49	9.8215	3.7452	95043
Mexico	2.45474	3.719487	6	2.6	63.9	64.1333	0.7123	0.68	27.6088	21.5928	4.8693	18621
Netherlands	1.98824	3.723033	6	6.8667	63.0333	61.8333	0.9283	0.9043	23.9	11.9427	5.4268	50621
New Zealand	2.82496	0.0339614	5	6.9667	66.6333	66.9	0.998	0.9913	16.7761	17.2742	6.0263	41009
Norway	2.835	0.0398324	6	7.2333	66.4	65.1667	0.9873	0.9947	19.3	16.3565	6.63	35991
Poland	1.41875	0.0215615	7	4.3	55.4333	52.2	0.816	0.8103	11.1918	11.9931	5.1788	15186
Portugal	2.36845	0.7876182	6	6.7333	60.0667	61.2	0.7357	0.7073	11.561	11.2863	5.0696	37542
Slovakia	2.21206	4.756228	6	5.1	58.3333	57.0667	0.8677	0.8537	17.6812	10.3347	3.721	12688

Slovenia	2.325	1.723563	6	5.8	59	59.6	0.9343	0.9263	14.6285	12.3234	5.4345	32436
Spain	2.95631	0.060231	6	5.9	52	54.3667	0.8637	0.8367	13.6943	11.111	4.363	42846
Sweden	2.73377	0.0246262	7	7.2667	62.5667	62.6333	0.9087	0.9077	10.5972	12.7754	6.7362	33374
Switzerland	1.42232	3.7983	7	6.4667	66.0333	65.0667	0.867	0.8557	20.2	16.3161	5.2363	70784
Turkey	0.7117	4.873262	6	3.9333	46.3	45.8667	0.6	0.573	25.6	9.4	2.8625	24761
United Kingdom	2.78477	0.02045	5	6.8333	59.3	59.8333	0.822	0.808	17.7436	11.6228	5.4387	44145
USA	1.6717	0.0488934	6	7	62.2	64.5333	0.9887	0.974	14.414	14.0951	5.4635	45226

Bibliography

Aydemir, A., Chen, W., Corak, M. (2008). Intergenerational education mobility among the children of Canadian immigrants. *IZA Discussion Paper No. 3759.*

Bauer, P., Riphahn, R.,T. (2007).Heterogeneity in the intergenerational transmission of educational attainment: evidence from Switzerland on natives and second generation immigrants. *Journal of Population Economics 20 (1), 121-148*

Brunello, G. & Checchi, D. (2007). Does school tracking affect equality of opportunity? New international evidence. *Economic Policy* Vol. 22, no. 52, pp. 781-861.

Cobb-Clark, D., A, Sinning M. & Stilmann. (2012). Migrant youths' educational achievement: the role of institutions. *Annals of the American Academy of Political and Social Sciences, 643 (1), 18 - 45*

Currie, J., Thomas, D. (1999). Does head start help Hispanic children? *Journal of Public Economics 74 (2), 235-262*

Currie, J. (2001). Early childhood education programs. *The Journal of Economic Perspectives 15 (2), 213-238*

Datar, A. (2006). Does delaying kindergarten entrance give children a head start? *Economics of Education Review 25, (1), 43-62*

Deming, D., Dynarski, S. (2008). The lengthening of childhood. *The Journal of Economic Perspectives 22 (3), 71-92*

Dronkers, J., Velden, R. F. & Dunne, A. (2012). Why are migrant students better off in certain types of educational systems or schools than in others? *European Educational Research Journal,* Volume 11, No: 1, pp. 11-44

Dustmann, C., Frattini, T. and Lanzar, G. (2011). Educational achievement of second generation immigrants: an international comparison. *Center for Research and Analysis of Migration (CReAM) Discussion Paper 25:* London.

Elder, T., E., Lubotsky, D., H. (2009). Kindergarten entrance age and children's achievement, impacts of state policies, family background, and peers. *The Journal of Human Resources 44 (3), 641-683*

Entorf, H. and Minoiu, N. (2005). What a difference immigration policy makes: A comparison of PISA scores in Europe and traditional countries of immigration. *German Economic Review* vol. 6, no. 3, pp. 355-376.

Gang, I., N., Zimmermann, K., F. (2008). Is child like parent? Educational attainment and ethnic origin. *The Journal of Human Resources 35 (3), 550-569*

Hanushek, E. and Wossman, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across Countries. *The Economic Journal*, 116(150), pp. C63-C76.

Heckman, J., J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science 312, (5782), 1900-1902*

Horn, D. (2012). Age of selection counts: a cross-country comparison of educational institutions. *Educational Research and Evaluation*, Vol. 15, No. 4, pp. 343–367

Kahn, L. M. (2004). Immigration, skills and the labor market: International evidence.

Journal of Population Economics vol.17, no. 3, pp. 501-534.

Korthals, R. (2012). Selection and tracking in secondary education: A cross country analysis of student performance and educational opportunities. *Working Paper*

Ludemann, E. & Schwerdt, G. (2010). Migration background and educational tracking: Is there a double disadvantage for second-generation immigrants? *CESifo Working Paper Series No. 3256*

OECD (2012). *PISA 2009 technical report.* PISA, OECD Publishing.

OECD. (2005). *Education at a glance 2005.* Paris: OECD Publishing.

Snow, C., E. and Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: evidence from second language learning. *Child Development,* vol. 49, no.4, pp. 1114-1128

Schneeweis, N. (2011). Educational institutions and the integration of migrants. *Journal of Population Economics,* Volume 24, Issue 4, pp. 1281-1308

Schnepf, S.V. (2007). Immigrants' educational disadvantage: An examination across ten countries and three surveys. *Journal of Population Economics* vol. 20, no. 3, pp. 527-545.

Schuetz, G., Ursprung H. W. and Woessmann, L. (2008). Education policy and equality of opportunity." *Kyklos*, vol .61, no. 2, pp. 279-308

Stanat, P. and Christensen, G. (2006). *Where immigrant students succeed- a comparative review of performance and engagement in PISA 2003.* OECD Report.

Waldinger, F. (2006). Does tracking affect the importance of family background on student's test scores? *Working Paper*