The Effect of the Age of Migration on Learning Dutch as a Second Language

Or How Linguistic Distance, Gender, and Country of Origin matter for Attaining a Second Language as a Child

Abstract

This thesis investigates how the age of migration influences the integration possibilities of immigrant children in the Netherland. For this purpose, language test results are used as a predictor for integration possibilities. The research uses the PRIMA database, which consists of six cohorts of children in the grades 2, 4, 6 and 8 in primary school, throughout the years 1994-2004. Using an OLS-method, and controlling for important determinants of language test scores, such as grade, social-economic status and school-fixed effects, a continuous, significant and negative effect of arriving at a later age is found. No clear sign of a cut-off point – also known as the Critical Period – was found in the data. These effects are roughly similar per subgroup of the four main countries of origin – Suriname, the Antilles, Morocco, and Turkey. But the results document a significant difference in the performance of girls per country of origin. Girls from Turkey and Morocco perform relatively similar to their male counterparts, whereas girls from Suriname and the Antilles perform significantly better. Additionally, the results show a small but significant effect for linguistic distance, indicating that language of origin matters for second-language learning.

Gerben Muskee

434936

Supervisor: Dr. Anne C. Gielen

February 2017

ERASMUS UNIVERSITEIT ROTTERDAM

Acknowledgments

I would like to express my deepest gratitude to my supervisor, dr. Gielen. Her guidance, constructive feedback and positive attitude were crucial in aiding me to write this thesis.

I am also very thankful to Vishand Lachmansingh, Peter van den Heuvel, Marijn Verspaandonk and Koen van Ruijven for their comments and valuable discussions. Likewise, I would like to thank Walter Kuks for sharing not too few tea-breaks in the university library. I am also grateful for all the help I received from Esmée Zwiers, aiding me in getting a grip on the dataset and providing valuable tips for additional information.

Finally, I am grateful to Jans Muskee for being someone I had to explain every part of the thesis to in a straightforward manner, which in the end made it easier for me to write the thesis in a clear way.

All remaining errors are my own.

Table of Contents

1. Introduction

In recent years, we have seen a significant increase in the amount of asylum applicants in the Netherlands and other western countries. As the volume of immigrants increases, it is important that immigration and integration policy does not limit the future for immigrants and their children. Currently, the school performance gap between native and immigrant children in the Netherlands is decreasing, but it is still large (CBS, 2015). School performances around the age of 12, near the end of the primary school, are an important factor in qualification for a higher future high school degree. Consequently, the gap in early child educational performance is still a significant factor as it predicts a gap in final educational attainment, and thus on later labour-market performances and other long-term life outcomes (Bleakley & Chin, 2010; Böhlmark, 2008). Better knowledge about the factors influencing the integration of the children from immigrants in their new society, will help to better target policies at those groups who need the help the most. In doing so, it will limit inequality in a country.

Integration of immigrant children can be measured by how good children perform in school (Frick & Wagner, 2001), and especially at what level they master the language of their new country (Isphording, 2014). Isphording (2014) finds that inadequate linguistic skills in the language of the destination country represent a significant hurdle for the integration and assimilation into the labour market of destination countries. This would make the use of language test scores a decent proxy for integration at a younger age. Being fluent in a new, second language makes it easier for the children to fit in into the class and perform better at school. It is also shown that immigrant children have an arrear at an early age in nation specific subjects like national history and geography (Van den Berg et al, 2014). This translates into worse general performance later as the children of immigrants spent fewer years in the same school as non-immigrant children, and thus have less time to develop these nation specific subject skills (Böhlmark, 2008). The general performance therefore suffers, and this ultimately might lead to a lower standard of living later in life (Mincer, 1974).

A major factor influencing the acquisition of a new language is age. The older one is, the more difficult it is for a person to attain a new language. A straightforward explanation for this so-called age of migration effect is that older migrant children have less time in the educational system in the country of destination relative to the younger, and thus perform worse. But even considering the years spend in the education system, there still seems to be a difference between children below a certain age and above a certain point of age, which is also known as the Critical Period (Johnson & Newport, 1989).

The Critical Period Hypothesis (CPH), first proposed by Penfield and Roberts (1959), states that there is a period during which learners can acquire a second language easier and achieve a more native-speaker-like level. When the Critical Period ends, this acquirement process become more difficult (Ellis, 2015). Originating from psychology and biology research, the Critical Period is demonstrated by an increased sensitivity to environmental stimuli (Basu, 2010). The Critical Period is crucial for immigrant children as a smooth integration requires being able to speak the second-language proficiently.

This leads to the main research question of this paper: does the age of immigrant children matter for assimilation in the Netherlands? In which assimilation is measured by performance in language in primary school, relative to the “home-grown” children. This research will provide some insight into the topic of the nature of human beings’ language learning capability, and how this changes as they grow older. It also offers the opportunity to policymakers to better suit educational policies to children of a specific age. For example, around what age it is useful to give children additional language classes to catch up. If immigrants bring along children with ages just above a critical period, then such children could be at a higher risk of adverse future outcomes, and one may reallocate funding towards preventive educational policies for such children. Similarly, the analysis is relevant for adoption policies, since the children adopted above the age of a critical period will also be at a higher risk of integration problems, and therefore they may need special educational treatment after adoption.

From a more economic point of view, a better integrated immigrant leads to more benefits to society in general. As better integration leads to better possibilities of a well-paying job in the future, Gonzalez (2003) finds that the higher wages for migrants also benefit society. This is thanks to the increase in taxes and decrease in the public services used and received by immigrants. Weighing the costs and benefits, one could argue that educating the children at a younger age basically pays for itself. He also states that students with less than a high school education (which is likelier the older they are when migrating) have no economic gain.

In addition to the main question, it might also be interesting to see if there are significant differences in the performances of immigrants, depending on country of origin. Various studies on immigrant integration processes pointed towards differing difficulties for immigrants who are trying to acquire the main language in their new country. The distance between the language in the native and the origin country of a migrant might be a significant factor in this development (Chiswick & Miller, 2005).

This thesis has two important aspects in which it is different from other research in the field of age of migration and second-language learning. First, many linguistic studies use a language test scores as a proxy, but make use of a small database. In contrast, research from a sociological and economic perspective on this subject is typically done with larger sample sizes. However, it usually makes use of self-reported quality of attainment of the second language. Both approaches have disadvantages: small-scale studies cannot deal with differing countries of origin and equally differing languages of origin as explanatory factors for variances in language attainment. Furthermore, the legitimacy of the use of self-reported measures for language attainment in large-scale databases is not entirely clear (Van der Slik, 2010). For example, De Bot, Evers, De Quay, and Van der Slik (2005) found that lowly-educated students overrate their foreign language skills. This decreases the validity of research using questionnaires.

This thesis is not subject to these limitations. This is due to the usage of a large database of about 15,000 children who were not born in the Netherlands. The database provides information on language test scores for children of several grades in primary school and thus can be used to estimate the effect of age of migration on language test performance. An additional benefit of the usage of this database is that the second-language attainment is measured directly by means of language test scores instead of self-reported questionnaires. Besides these reasons, this study uniquely combines two databases from different research backgrounds, to research the possible influence of linguistic distance on the language test scores. A further explanation of this will follow in the next section of Relevant Literature

The results show that there is a significant difference for children arriving in their new country at a later age. However, no clear sign of a Critical Period can be detected. Country of Origin is not a deciding factor in the language test results, although gender seems to play a different role depending on the country of origin. Traditional gender roles might play a role deciding in how much girls are being motivated by their parents and penultimate their results in school. Additionally, linguistic distance has a small but significant contribution in predicting the language score results. The higher the linguistic distance relative to Dutch, the worse the language results will be. From the four main migrant groups in the Netherlands, the Turkish and Moroccan girls perform more equally to boys, contradicting with general educational research.

The remainder of this paper is organized as follows: the next section presents the relevant literature, which will give an overview of the findings of related research and the research method used therein. After that, a description of the data used to investigate the questions mentioned will be offered, followed by an explanation and discussion on the research method used by this paper. Afterwards, the outcomes of the research are presented, in combination with a discussion of the limitations and suggestions for further research. To wrap it all up, the last section will describe what these findings imply and what the impact could be for immigration and education policies.

2. Relevant Literature

This section is divided in three segments. The first is describing the relevant literature around the main question of the paper if the age of migration matters for primary school results and subsequently integration. This is accompanied by discussing the research methods of the papers. The following segments discuss the relevant literature for the two side questions about expected differences in results based on country of origin and gender.

2.1 The Effect of Age of Migration on Integration

There is no clear consensus in the literature on when the critical period for language learning ends or even if there is one to begin with. Thomas and Johnson (2008) and Birdsong (1999; 2006) present an outline of the biological literature on second-language attainment. Researchers have been divided on this topic into two camps, where one side argues the age of 5 or 6 to be a critical turning point. While others consider this point to be around to the age of 12 till 15 (Singleton, 2005).

Recent studies in economic research emphasize the importance of the relationship between the age of immigration and integration of migrants. For example, Schaafsma and Sweetman (2001) find evidence for varying results in the educational system when the age of migration changes. They also find that the salaries of migrants differ, with migrant children arriving before the age of 12 to earn a similar income to natives. For the United States, comparable results were found by Gonzalez (2003), and Bleakley and Chin (2010). Migrants who arrive at a younger age in the US integrate quicker than migrants who arrive in their teens. Subsequently, childhood migrants tend to perform more similar to their native counterpart than migrants who arrived during their teens.

The studies performed on this subject so far can be categorised into three different research methods: An Ordinary Least Squares (OLS), a Fixed-Effects (FE) method comparing siblings, and an Instrumental-Variables (IV) method. All relevant studies are divided by research method and discussed including the deficiencies of the chosen research method. First, the OLS studies are liable to omitted-variable bias. It is very hard to account for all factors which could influence the difference between groups. It is in fact always unsure if every possible explanatory variable is considered with such a study, as characteristics of the parents might influence both the decision to migrate and the ability/performance of their children. There is also the issue of selection bias, which occurs when the research method does not control for the factor that certain kind of immigrants arrive at an earlier age than others. One example of this is the distinction between refugees who migrate involuntarily and economic migrants, who choose when they migrate. The following studies perform such an analysis Isphording (2014), Gonzalez (2003), Van Ours and Veenman (2006), and Corak (2011).

Corak (2011) focuses on the degree to which high school graduation for immigrant children may change after a particular age at arrival in Canada. He finds that there is no apparent relation between graduating high-school and age of migration, but only for those migrants who came from English- or French-speaking nations. For countries of origin which are not English- or French-speaking, the levels in high-school graduation differ significantly. However, the shape of the effect remains the same: no difference between migrant children arriving at an early age and native children, followed by a distinct change when comparing natives with migrant children who arrived at a later age. This moment of distinct change in high-school graduation due to later age of arrival differs between countries of origin, but generally matches to the first years of primary school. For some countries this moment of distinct change takes place around the age of five, for other countries this is around the age of 9. Douglas Willms categorizes this moment of distinct change as a period in which children change from “learning to read” to “reading to learn”. In regards of the Netherlands, Van Ours and Veenman (2006) perform a similar study in which they find that Turks and Moroccans have a stronger disadvantage than the Suriname and Antilleans in regard of educational achievement. Furthermore, they find evidence that men suffer less from the disadvantages of migrating than women. However, the research entails migrants from the age of 15 to 29, so has nothing to say about any variance in the moment of distinct change for differing countries of origin. More about the possible reasons can be found in the following sections.

The second group of research designs is the fixed-effects method. Unlike the OLS method which is heavily susceptible to omitted-variable bias, the fixed-effects method deals with this bias by using data with a time or cohort dimension to control for unobserved but fixed characteristics, such as genetics and unobservable household characteristics like parenting style. Böhlmark (2008) uses such a research method and finds that a later age of migration results in lower educational performance, and that the critical age is around the age of 9. This research is done by comparing siblings in a family-fixed-effects framework. Although it controls for a certain type of omitted variable, fixed-effects estimates are prone to attenuation bias thanks to measurement error. This bias means that the estimate of the effect will be biased towards zero, due to large variances in the errors of the independent variable. A variation of the measurement-error problem comes up from the fact that the variation for which is controlled with a fixed-effects method, removes both good and bad variance. In other words, using a fixed-effects method can kill some of the omitted-variables-bias bathwater, but it also removes useful information about the baby (Angrist & Pischke, 2008). An example of such a case is shown by Bound and Solon (1999). They found that first-born children weigh more and have higher IQ-scores than the second-born children in twins. This means results could be biased from a within-twins research. Also, using a fixed-effects method to analyse the effect of immigration age on cognitive skills and education might be subject to spill-over effects. One sibling could help the other at school, later in life at work, integrating into the new environment (making friends), all this could put a downward pressure on the measured effect. Although the research method might not be perfect, Van den Berg et al. (2014) find a decline in cognitive test scores which sets in after the age of 7/8. Likewise, when performing their sensitivity analysis, they make use of a regression discontinuity design to check for discontinuities at potential critical ages. While checking they still use their FE model, but they do find significant discontinuities around the age of 5 and 9.

The third and final research design is the Instrumental-Variables design. This design is also called a Two Stage Least Squares, as it uses a regression upon a regression. This design recreates a random test by a variable called the 'instrument' which is assumed to influence the independent variable and no direct link to the dependent variable, while also being as good as randomly assigned. If these assumptions hold, there is no selection bias due to the random assignment of the instrument and it being independent of the potential outcome. Where the fixed-effects method still has some difficulties with omitted variable bias, the IV design, contingent on the power of the instrument, is equivalent to a random design (with non-compliance). Bleakley and Chin (2010) consider the question of what the effects of integration are later in life. It is worthwhile to note that they find a significant decrease in language-learning ability around 7 years old. This regression is done with an OLS (as a First-Stage part of an Instrumental Variables research design), so the exact values might be biased, but the direction of these studies point into the same direction as Van den Berg et al (2014).

Due to the openness of the Dutch school system, which provides relatively more "long routes" and "second chances" than for example their German neighbours, Crul and Schneider (2009) find that Turkish migrants are significantly more present in higher education. This means that in the Dutch education system it is more difficult to base any long-term expectations like income, unemployment and highest level of finished education on primary school results. Obviously, many other factors, like school choice and possible changes in social environments during high school and later in life, influence educational and career performance. By looking only at the immigrants in the Netherlands, mainly from Turkey and Morocco, it is difficult to tell if the results of this paper can be generalized to other countries and to other immigrant groups. This is amplified by the liberal immigration and integration policies in place in the Netherlands around the beginning of this century, where for example immigrants were encouraged to preserve their own cultural identity (Euwals et al., 2007).

2.2 Heterogeneity by Country of Origin

One of the reasons why age-of-migration estimates might vary by country is due to differing education traditions. In a sibling study in Sweden, Böhlmark (2008) finds several significant differences between immigrant groups, among which it is notable that Western children perform like Swedish children. They only experience a drop in their performance at school when arriving in Sweden after the age of 14. Performance in this study is however measured by GPA scores, which measures all cognitive school-related performance instead of only language skills. Thanks to this, the similarities between Swedes and Western immigrants, and the differences between immigrants from other parts of the world can be attributed to the quality of the school systems in each respective country. Böhlmark also singles out the Asian migrants, who are estimated to have a culture where education is valued higher. Results in the test however, do not show a significant increase in performance when comparing Asian migrants with other non-Western migrants, due to the higher linguistic distance from the English language for Asian languages. Gonzalez (2003) concurs with Böhlmark, as he finds that success in American schools depends on the degree of transferability of the education a migrant had in the country of origin. Immigrants from vastly different education systems relative to the education system in the country of arrival, will thus face more difficulty relative to children from countries with similar education systems. Therefore, it is possible that immigrants with a low age of migration complete more years of education in that nation-specific education style (and quality) of the country of arrival, and thus have more of their education time in the specific education than migrant children who arrived at a later age.

A second explaining factor is the linguistic distance between the first-language of a child and the language spoken in the country of arrival. An increase in linguistic distance decreases the potential language transfer, which is the application of knowledge in the mother tongue in the acquisition of the language of the destination country (Isphording, 2014). Providing an economic interpretation, the linguistic distance displays to what degree the human capital of the country of origin can be transferred to the country of arrival (Friedberg, 2000). Linguistic distance could also possibly vary across the distribution of age at arrival (Corak, 2011). This would be the case in the Netherlands if those arriving at younger ages are disproportionately from Dutch-speaking or from countries with a language close to Dutch (Germany for example). In such a case, the effect of age of migration on the language learning ability would differ between varying countries of origin. This hypothesis is confirmed by Schepens et al. (2013), who found robust first-language distance effects for immigrants trying to learn Dutch.

With respect to Dutch immigration, prior exposure to one of the official languages could vary, even when the source country is not a Dutch-speaking country. Age at arrival is not automatically equal to the age of first exposure to a second language. This depends on the country of origin, schooling, parental investments, exposure to relatives, or visits to Dutch-speaking countries (Corak, 2011). When those who arrive at a younger age are more likely to have been exposed to Dutch before arriving to the Netherlands, their outcomes may tend to be more favourable than otherwise and bias the findings. Parents who are aware of the difficulties of learning a second language could prepare the children beforehand such that these children have a slight benefit compare to children who are not trained in advance, for example refugees.

Van Ours and Veenman (2006) find that migration at a young age appears to be more of a disadvantage for the educational achievements of Turks and Moroccans than for Surinamese and Antilleans (the four main countries of origin for the Netherlands). As Suriname and the Antilles are former colonies of the Netherlands, their educational systems are more in line with the Dutch educational system. As such, it is easier to become familiar with the Dutch system for the Surinamese and Antillean migrant than for the Moroccan and Turkish migrants. Also, the aspect of language comes into play here, as the former colonies still practice the Dutch language – at least till some degree. In Suriname, Dutch is the main language, while for the Antilleans this is Papiamento. Although Dutch is still the second language being taught in middle and high school. As it is more likely children encounter the Dutch language before they migrate in the former colonies, they might also find it easier to integrate once arrived in the Netherlands, compared to the Turkish and Moroccan immigrants.

All in all, from the reviewed literature it can be deduced that assimilation success (measured by second-language acquisition) depends on differences in educational traditions, the linguistic distance, and the age of first exposure.

2.3 Heterogeneity by Gender

According to Basu (2011), the critical age for acquiring English as a second-language is 8 for boys, and 10 for girls. Genetically, gender mainly starts to impact the age of migration effects during puberty, also identified as the stage in which “identity” formation takes place. It is thus expected to only see significant differences after or during puberty, which girls experience about two years earlier than boys. Alternatively, spill-over effects could vary for girls, in the sense that girls are generally seen as more socially active which could be helpful for learning a new language. Another reason could be that parents prefer their sons, which could make them care less about their daughters’ age when they are considering the timing of their migration. In such a case, girls are likely to arrive at older ages than boys. Similarly, parents of girls in societies which can be characterised as ‘traditional’, may be more inclined to think their daughters should invest less time in education and more in household chores (Basu, 2011).

Dahl and Moretti (2008) find a variety of effects concentrated among first-born daughters in the United States. If the first-born child is a girl, she is less likely to live with their father than if the first-born child is a boy. Similarly, the number of children in a family is significantly higher if the first-born child is a girl. Generally speaking, an absentee father does not bode well for a child, and economically speaking it results on average in reduced family income and a higher chance of family poverty. Thanks to variances in educational performances between boys and girls which are likely stronger than any psychological effects arising from an absentee father, comparing boys and girls is a difficult task. Instead, Dahl and Moretti (2008) make a comparison between second-born children. They make a division between second-born children with a first-born daughter, and second-born children with a second-born son. Notably, children with first-born sibling as a girl have lower educational achievement, which points towards a general preference of (American) families (or fathers) towards sons. The explanation offered by Dahl and Moretti is that parents who prefer boys over girls, can pay less attention to girls or might be less incentivising towards girls and their educational performances.

Gender differences might also come into play at an earlier stage of life than puberty, as was argued by Basu (2011). Two international educational comparison studies, the PIRLS and TIMSS, showed a remarkable consistent picture where girls perform better in reading than boys in almost every country (Meelissen et al., 2012). In a meta-analysis, Voyer and Voyer (2014) find a stable advantage in school marks for girls across the years in the data retrieved (from 1914 to 2011), starting from an early age. Noteworthy is that the female advantage was largest for language courses and smallest for math courses. Possible explanations for the gender differences refer to the nature-nurture discussion. Some argue that biological differences are the root cause (Sax, 2005; Singleton et al., 1999), others point to differences in expectations and stereotypes regarding boys and girls (Chapman et al., 2007; McGeown et al., 2012).

In respect to the Dutch situation, Van Ours and Veenman (2006) note that male immigrants from the big four countries of origin (Turkey, Morocco, Suriname, and the Antilles) are generally less disadvantaged than women. As stated by them, this difference comes from the more traditional culture compared to the Netherlands, wherein women equality tends to be higher in terms of societal participation. In a more traditional culture, women are ‘protected’ by their family, which can be found in less contact with people of the country of arrival or living in a less stimulating environment at home. Things which in general would be helpful for adapting a new language and feeling integrated in a society.

For the Turkish immigrants, a lot of research has been done about the parental attitudes towards the schooling and careers of girls (Coenen, 2001; Crul, 2000; Lindo, 2000). Coenen (2000) has termed such attitudes “cultural carryovers”, which is the situation in which parents grew up on the countryside where it was normal for girls to be part of the household or family and education for girls was seen as unnecessary. These parents have a stronger tendency to ‘carryover’ these values unto their children. Daughters of these parents seem to be driven out of school, specifically higher education, by obstructing parents. Also, Turkish girls are more often present in less theoretical types of high school education such as lower vocational training. The parents consider the education the girls receive as preparation for marriage instead of training for a future career (Crul & Schneider, 2009).

Crul and Schneider (2009) argue that Turkish immigrants carry over their mainly traditional definition of the norms and values unto their children, resulting in a lower value given to the education of girls compared to the boys. The Turkish migrants are a major immigrant group in the Netherlands, as The Netherlands recruited substantial numbers of so-called ‘guest-workers’ from Turkey in the 1960s and early 1970s followed by family reunification immigration afterwards. This might make it interesting to see if there is any significant difference in the primary school grades between boys and girls and if this differs by country of origin.

3. Data Description

The discussing and describing of the data is divided into several sections. First, the sample is described in general. E.g., how was the sample collected, what are the countries from origin, what is the distribution of age of migration. After the general description, we move on towards specific parts of the data. In paragraph 3.2, an explanation is given of how the fact that the grade of children affects the language test scores is dealt with. The third paragraph describes the issue of school-fixed effects versus clustered standard errors, and the final section is about the factor linguistic distance and how different databases are used to create the variable for this sample.

3.1 Report of the Sample

The analysis made use of the research program "Cohortonderzoek Primair Onderwijs" – abbreviated as ‘PRIMA’ – is a longitudinal data collection for primary education. It describes and explains the development of cognitive and social skills, as well as social behaviour of pupils in primary education. Since the 1994/1995 school year, bi-annual research has been performed with respect to pupils in grade 2, 4, 6, and 8 in some 600 schools for (special) primary education. The grades correspond respectively with the age of 6, 8, 10, and 12. The data-set includes the Dutch CITO-test scores, a standardized test almost every child must take at the end of the primary school to select the most-fitting type of high school. The dataset also contains information about the age the immigrant children arrived in the Netherlands.

The PRIMA cohort studies have been collected by ITS Nijmegen together with SCO-Kohnstamm Institute in Amsterdam. About 30 children of each grade were randomly selected at each school. Some 60,000 pupils were involved in each cohort in this national project. The data from all six the PRIMA surveys were used for this article. PRIMA drew a cluster sample of schools, divided into a representative and a supplementary section. The former is representative of all primary schools in the Netherlands. The supplementary section comprises schools with an over-representation of students from socio-ethnically disadvantaged situations. This sub-sample was included to permit reliable appraisal of socio-ethnically disadvantaged minority groups. The PRIMA data originates from four sources: 1) teacher assessments of children’s non-cognitive skills; 2) parental survey responses on family background and children’s early childhood care and education; 3) school’s registry information, e.g. on general parental and child characteristics; and 4) language and cognitive test scores from the school administrations for the testing period in the middle of the second school grades (Driessen et al, 2000).

Table 1: Number of Students by Age & Year of Cohort

Age	Year 1994	1996	1998	2000	2002	2004	Total
3	0	2	31	11	1	0	45
4	23	21	249	130	100	2	525
5	9,257	10,325	10,910	10,184	10,233	10,057	60,966
6	4,427	4,898	5,406	5,171	4,847	4,612	29,361
7	8,820	8,605	9,862	8,729	8,335	8,533	52,884
8	5,689	5,515	6,508	5,739	5,740	5,290	34,481
9	8,012	8,147	8,753	8,491	8,066	7,741	49,210
10	5,127	5,338	5,931	5,968	5,615	5,480	33,459
11	7,668	7,859	8,462	8,061	8,142	7,839	48,031
12	4,910	4,908	5,567	5,614	5,662	5,333	31,994
13	686	545	663	533	513	510	3,450
14	29	16	21	21	5	9	101
Total	54,648	56,179	62,363	58,652	57,259	55,406	344,507

Source: created by the author based on the data of the PRIMA cohort studies.
Note: this does not include the dropped observations of children born in a 13th month or before the year of 1970.

In the first cohort of 1994, children born far before 1970 were included in the sample. As they would have been 22 years old at that time, it seems highly unlikely that these data points were filled in correctly. And if true, these observations would be likely to be outliers in other departments than age only. Subsequently, these observations were dropped from the database. In the same manner, children registered as being born in the 13th month of the year were removed from the database as there is no way these observations could be verified and placed in the correct month. The variable 'age' was constructed by deducting the date of birth (determined in month and year) from the year of the cohort. The language test was taken in several months throughout the country, of which the timing differed per cohort. For example, for the cohort of 1994 the language test was taken in the months September till December, and for the cohort of 1998 the test was taken from January till April. Seeing as it was impossible to tell which month a student took the test in, it was decided to only use the year of the cohort.

The indicators used to measure the integration of the migrated children are twofold. The first segment consists of the tests taken in all years of primary school. It contains a language and an arithmetic part. The second segment of indicators comes from the nationally taken Cito test in grade 8, which consists of four parts: language, arithmetic, information-processing and the test called World Orientation, which combines geography, history and biology. Information-processing consists of a general understanding of graphs, tables, maps, and how to search the internet. The Cito test is only taken once, at the end of primary school. Thanks to the process of grade retention and skipping, children might differ by age but should be more or less similarly prepared for taking such a test.

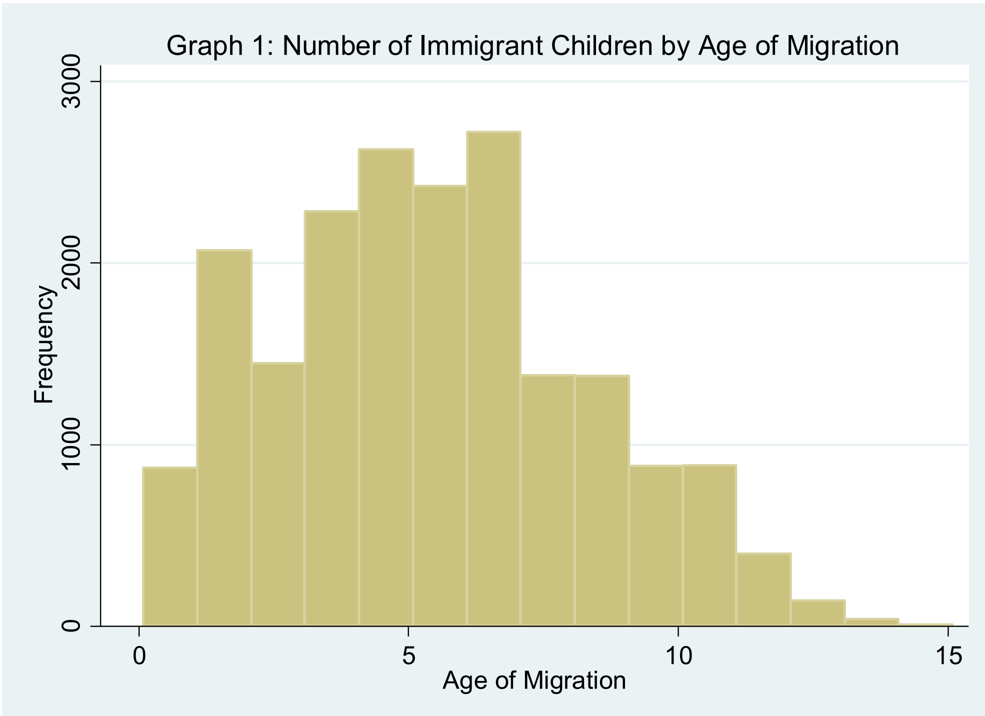
As only the language abilities of migrant children are of value for this thesis, all the native children observations were dropped. Due to a lack of uniformity of the observations used and the variable definition throughout the cohorts, several strict assumptions were made to create the required variables. The two main assumptions were in regard of the country of birth of the children and the age of migration. The former was identified in only one of the six cohorts and thus had to be constructed for the other cohorts, which was done by using the countries of birth of the parents. The somewhat limiting assumption was made that for a child to be born in a certain country, both his parents had to be from the same country. To also include one-parent families, foreign single-parents were identified and their children were given the same nationality if they were born outside the Netherlands. Parent-couples with differing nationalities are thus left out, as the country of birth of the children cannot be obtained with much certainty. As can be seen in the table below, still a significant portion of the children were identified as having Dutch parents. These are likely to be either adopted children or expats (people who stay in a foreign country for a limited amount of time, like diplomats).

Table 2: Amount and Percentage of Migrant Children by Country of Origin, and Gender

Country of Birth	Gender: Boy	Girl	Total
Netherlands	1,031 (6.60%)	1,130 (7.24%)	2,161 (13.84%)
Suriname	462 (2.96%)	493 (3.16%)	955 (6.12%)
Antilles	298 (1.91%)	330 (2.11%)	628 (4.02%)
Moluccas	10 (0.06%)	1 (0.01%)	11 (0.07%)
Turkey	1,262 (8.08%)	1,294 (8.29%)	2,556 (16.37%)
Morocco	1,312 (8.40%)	1,311 (8.40%)	2,623 (16,80%)
Greece	4 (0.03%)	5 (0.03%)	9 (0.06%)
Spain	3 (0.02%)	18 (0.12%)	21 (0.13%)
Italy	8 (0.05%)	7 (0.04%)	15 (0.10%)
Portugal	28 (0.18%)	33 (0.21%)	61 (0.39%)
Yugoslavia	455 (2.91%)	404 (2.59%)	859 (5.50%)
China	86 (0.55%)	98 (0.63%)	184 (1.18%)
Vietnam	46 (0.29%)	34 (0.22%)	80 (0.51%)
Other	2,794 (17.89%)	2,657 (17.02%)	5,451 (34.91%)
Total	7,799 (49.95%)	7,815 (50.05%)	15,614 (100%)

Source: created by the author based on the data of the PRIMA cohort studies.
Note: The first number in each cell contains the number of children of a certain gender and country of origin. The number between brackets represents the percentage of a cell relative to the total number of migrants. The children defined as from the Netherlands are children with Dutch parents but born in a foreign country. They are likely to be either adopted or children from expats.

The latter assumption constrained age of migration, a variable which was present in all cohorts mainly as a categorical variable. This means that the variable was divided into categories of a certain amount of years which had passed since migration. The categories ranged from ‘less than one year ago’, ‘between one and three years ago’, ‘between three and five years ago’ to ‘more than five years ago’. By using the average of each category, these categories were split into respectively 0.5, 2, 4, and 6 years, and were subsequently used to calculate the age of arrival for the children of migrants. Obviously, this results in some noise in regards to the age of migration, which makes it more difficult to precisely pinpoint the age after which the critical period ends. However, the effect of age of migration as a continuous variable roughly remains the same, meaning that the overall negative effect on language test scores is still present and not influenced by age of migration formerly being a categorical variable.



Source: created by the author based on the data of the PRIMA cohort studies.

The total sample size of immigrant children consisted of 7,799 boys and 7,815 girls. On average, there are about 650 observations for boys and 650 for girls at each age of arrival. As can be seen in the histogram in Graph 1 above, it was the case that the largest fraction of children arrived in the country between birth and the age of 7. This might point towards parents being aware of the extra difficulty their children will have integrating or learning a new language at a later age. If the awareness of this phenomenon is mainly present under parents with certain characteristics, and these characteristics directly influence the ability of children, the results are likely to be biased. Examples of these characteristics might be higher education, higher income, certain education styles. Comparing groups of children who arrived before or after a certain age is difficult in such a situation. As children of these parents will perform better in school compared to their peers, regardless of the age of migration effect, thanks to the advantages their parents offer. It should be noted though, that parents do not have full-decision making ability or authority when it comes to the age of migration as immigration applications can be delayed or might be rejected at first try (Corak, 2011). Seeing as there is a certain distinction after the age of 7 visible in the histogram, there is a danger of omitted-variable bias. This suspicion cannot easily be invalidated due to the structure of the PRIMA database, in which additional household characteristics like detailed information about income of the parents and family size and composition are only available for all children of all grades for the first cohort in the year 1994. In the following cohorts, these household characteristics were either dropped – partially or entirely – or only present for grades 2 and 8.

Table 3: Descriptive Statistics for Children of Different Ages of Migration

	0-7	8+	Difference
Age	8.47	11.38	2.912***
	(2.23)	(1.62)	(0.0330)
Gender	1.50	1.50	0.00377
	(0.50)	(0.50)	(0.00811)
Education Mother	2.08	2.15	0.0694***
	(1.19)	(1.30)	(0.0200)
Education Father	2.44	2.59	0.147***
	(1.38)	(1.46)	(0.0236)
Age of Migration	4.06	9.11	5.055***
	(1.85)	(1.53)	(0.0280)
Language Test Score	1023.54	1063.54	40.00***
	(61.25)	(46.46)	(0.973)
Math Test Score	304.59	469.50	164.9***
	(409.47)	(495.30)	(7.459)
Total Cito Score	530.39	527.89	-2.496***
	(10.01)	(11.34)	(0.390)
Cito Language Score	53.63	45.17	-8.457***
	(19.53)	(18.96)	(0.726)
Cito Math Score	39.48	37.29	-2.194***
	(11.38)	(12.38)	(0.448)
Cito Information Score	31.00	32.18	1.180**
	(10.00)	(11.72)	(0.412)
Cito Geography Score	41.25	39.10	-2.163***
	(12.90)	(12.98)	(0.553)

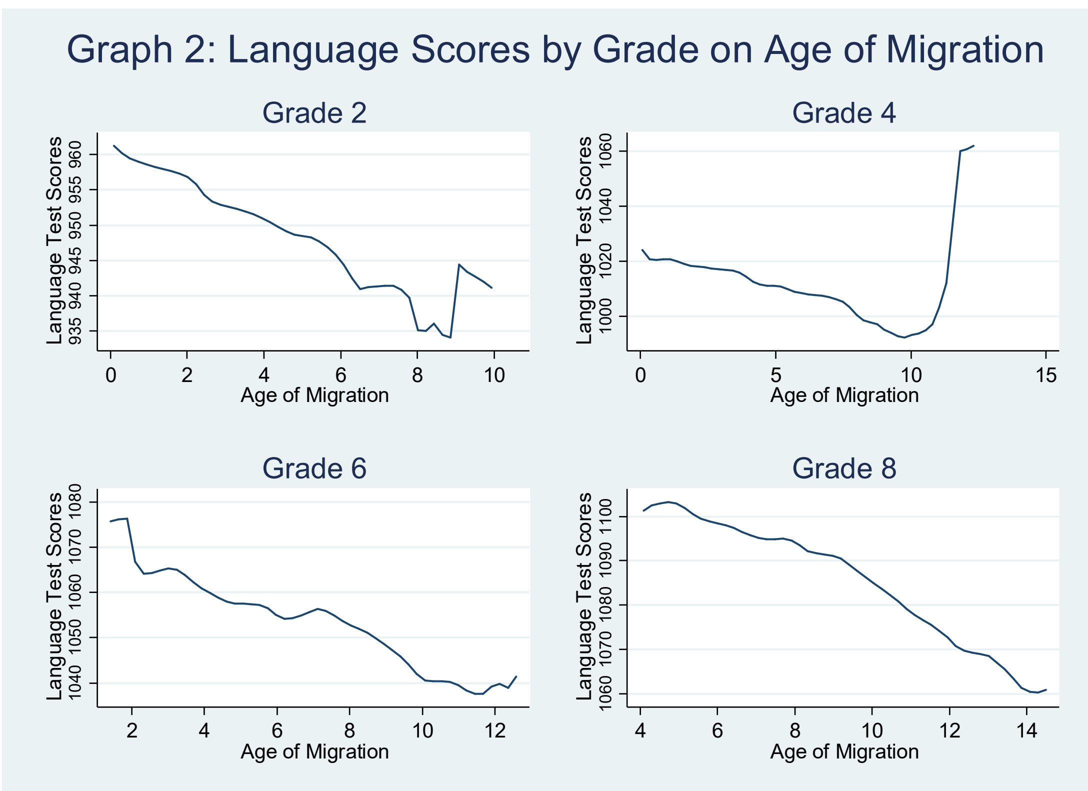
Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard errors in parentheses.
* p<0.05, ** p<0.01, *** p<0.001

It is noteworthy that the children arrived after the age of 8 seem to have a small, but significant difference in background characteristics. The education of the parents is slightly better and they are older at the moment of observation, which makes sense as the group of children arriving at a later age uses an average of older children. The 8+ group of migrants score significantly better on the language and math tests, very likely due to the grade effect. The difference in the Cito scores is more what one might expect from the relevant literature: the children arriving at a later score worse than children arriving before the age 8. Except for the Cito Information test section, which might be thanks to the absence of a necessity of nation-specific education for this school subject. The Cito scores differs from the former test scores in the Cito test having a uniform unit of measurement for children of all ages. Additionally, the test is only taken in the eighth grade. The breakdown of gender between native and immigrant children seems to be similar.

Moreover, striking is that the standard deviations for Math Test Scores for both immigrant and native children is higher than the mean test score. This indicates that there is an enormous dispersion in the results of the math test and this might make inference from this estimator more difficult. Although the math test is not the main variable of interest, it might be interesting to see how it compares with the language test. For example, to check if language is indeed more influenced by a later age of migration. The math tests are made on basis of the PRIMA-math test in the first two years. In further years, the math tests were based on the test of Calculating & Math from the Cito student tracking system. These scores of the latter four years can be transformed to the same level as the former years to make comparison possible, but any trend on basis of these test results should be handled with care (Kamphuis et al., 1998). Seeing as the data from this paper is no longer available for public use, this transformation method cannot be used. To compensate for this in a regression model, a dummy for year will be added. This should control for all the fluctuations due to a change in tracking system throughout the cohorts. However, to show the age of migration effect for math scores in a graph might prove to be more difficult, as it would have to be standardized by both grade and year.

3.2 The Influence of Grade

In Graph 2 below the age of migration effect is plotted against the language test results, split by grade. One can see that the older a child is when it migrates, the worse it seems to perform on the language test. But when all the grades are lumped together, there is an upward going line, indicating that performances increase when a child arrives at a later age. Not only does is this contradictory to general common sense, it also does not compute when focusing on the effect of age of migration divided by grade. The large sudden increase in grade 4 after the age of 10 does not fit the general idea about age of migration effect. To find out why, a scatter plot of the fourth grade has been added in which one can see that the curve of grade 4 is quite sensitive to the few outliers of children who arrived at a relative late age to be in the fourth grade as the general age for children in the fourth grade is 8. Age might play a significant role for these few and these observations are thus regarded as outliers.



Source: created by the author based on the data of the PRIMA cohort studies.
Note: The large uptick in language test scores in grade 4 is created by a few observations who arrived after the age of migration of 10. These observations have higher language test results and are significantly older than the “normal” grade 4 age; they are thus regarded as outliers.

When taking a closer look upon the y-axis of the graphs in Graph 2 or Table 4 below, one can see that the scores on the language test differ for immigrant children in different grades. There is a clear positive correlation between grade and language test scores: there is a bump in test scores when a child goes to the next grade, which is shown by the differing minimal and maximum scores in Table 4. This makes it difficult to compare children between grades. In a regression model, these discrepancies can be considered to by including a dummy variable for grade. However, showing the predicted output from the model in a graph would not result in a downward-sloping graph, as the effect of grade would be stronger than the age of migration – as can be seen in Graph 2. As the age of migration increases, the likelihood of entering the education system at a higher grade also goes up.

Table 4: Language Test Scores of Immigrant Children by Grade

Grade	N	mean	min	max
2	3266	953.4239	810.094	1133.3
4	4250	1012.248	841.8	1143.3
6	4154	1055.4	923.3	1197.4
8	3990	1091.729	904.2	1261.2
Total	15660	1031.677	810.094	1261.2

Source: created by the author based on the data of the PRIMA cohort studies.
Note: N is the number of children in a certain grade. Mean is the average score of all the immigrant children in that grade on the language test. Min and max are the minimum and maximum scores on the language test.

Table 5: Standardized Language Test Scores of Immigrant Children

Grade	N	mean	min	max
2	3266	-.5920971	-4.535909	4.357309
4	4250	-.6512312	-5.0812	2.754837
6	4154	-.5417457	-4.358998	3.56157
8	3990	-.5980414	-5.640549	3.958916
Total	15660	-.5963038	-5.640549	4.357309

Source: created by the author based on the data of the PRIMA cohort studies.
Note: N is the number of children in a certain grade. Mean is the average score of all the immigrant children in that grade on the language test. Min and max are the minimum and maximum scores on the language test.

However, to be able to show somewhat of the age of migration effect, the test results could be standardized by grade. This should result in test scores which are comparable between children in different grades. Seeing how the mean, minimum and maximum in Table 5 are not eerily similar, it was chosen to use both the standardized test results as the test results by grade. They are not ideal measures for the total picture, but to present the most complete overview of the age of migration effect, both measurements will be used.

3.3 Linguistic Distance

Linguistic distance is a measure of the difficulty of learning a new language. In this case, the closer to zero the linguistic distance, the easier it would be for an immigrant to learn Dutch. Suriname and the Antilles thus have a better score on linguistic distance, most likely thanks to the colonial links to the Netherlands through which Dutch as a language still plays a role in the educational system. Turkey and Morocco score relatively worse due to Turkish and Arabic not being languages which are closely-related to Dutch. See the table below for all included Countries of Origin and their linguistic distances. The countries which had 1% or more of the total immigrant children population were included (see also Table 2).

Table 6: Linguistic Distance Measures by Country of Birth

Birthcountry	mean	N
Suriname	498.93	1009
Antilles	518.03	636
Turkey	547.61	2672
Morocco	549.1	2746
Yugoslavia	533.3	877
China	536.39	187

Source: created by the author based on the data of Schepens (2015), Schepens et al. (2013) and Hessels (2005).
Note: the original scores of Schepens (2015) & Schepens et al. (2013) have been inversed to correspond better with the term 'linguistic distance'. The closer the mean is to zero, the smaller the linguistic distance and thus the easier it would be for an immigrant of said country to learn Dutch as a second language. For Suriname, a combination of Hindi and Dutch was used, and for Yugoslavia a combination of Albanian, Croatian, Serbian and Slovenian. N is the number of immigrant children from that country.

Each country of origin in the sample which was 1% or more of the migrant sample, was classified with a certain linguistic distance measure. This was done by using the main language spoken in the country of origin and how it related to Dutch, as it is spoken in the Netherlands. To do this, the measure of linguistic distance developed by Schepens, Van der Slik and Van der Hout (2013) in combination with Schepens (2015) was used. The dominant languages in the country of origin were used, except for Suriname and Yugoslavia. In Suriname, an average estimation of Hindi and Dutch was used due to the lack of available data for Surinamese, and Dutch not being a language which can be used as a starter language for learning Dutch. A similar combination of languages was sought for the linguistic distance of Yugoslavia due to the many differing languages and ethnicities from that region (Hessels, 2005).

4. Research Method

To prove the Critical Period Hypothesis in regards of second language development, it does not suffice that there is a relationship between age and language attainment. To qualify as a Critical Period, there needs be a marked difference in the line regressing second language attainment on age of migration on either side of the of the critical age point. Ordinary least squares (OLS) is used to address the main question: does the age of migration matter for assimilation in the Netherlands, where literacy is used as a proxy for assimilation.

The predicted values below capture the relationship between age of migration and the results on language tests, which is used as a proxy for literacy. In all models, the result of the language test is based on the year and month of arrival and the gender of the child. In this specific model, a continuous variable for age of migration is used to check how the effect looks. Pooling migrants from all cohorts, the empirical specification is:

$$Y = \beta_0 + \beta_1 AOM + \beta_2 Sex + X'\gamma \qquad (1)$$

The language test score 'Y' is estimated as a function of the age of migration 'AOM', gender 'Sex', and control variable 'X' which consists of the categorical variables social-economic background, school-fixed effects, and the grade the child was in (divided into grade 2, 4, 6, and 8). Due to these last three variables being categorical variables, they are used as dummies in the regression. This ensures that the regression no longer treats the categories from the categorical variables as additive or multiplicative.

Next, the second model specifies at which specific age the Critical Period ends. This is done by replacing the continuous variable for age of migration with a dummy indicating if children arrived before or after a certain age. This means that 'AOM' is replaced by 'AgeOfEntry', which is a dummy for arrival before a certain age. Several ages are tested, both above as below the by the literature expected age of 7. The independent variable 'Sex' and control variable 'X' stay the same.

$$Y = \beta_0 + \beta_1 AgeOfEntry + \beta_2 Sex + X'\gamma \qquad (2)$$

The second research question is if the estimated effect differs by country of origin. To address this, a specification to the two models above is used, but specifically for each of the four most important origin countries (i.e., Suriname, the Antilles, Morocco, and Turkey). These will be tested to see if the overall continuous effect of age of migration remains the same and if the Critical Period differs for Country of Origin.

To go into more detail about the country of origin, the model was specified towards the language most inhabitants spoke in their origin country. For this, the linguistic distance as measured by Schepens et al. (2013) and Schepens (2015) was combined with the original PRIMA database. As the linguistic distance measurement is based on language test of migrants who arrived in The Netherlands, this design also allows for controlling for origin-fixed effects capturing potentially omitted country characteristics, such as quality of education in country of origin and exposure to the second language. In the model, the result of the language test is additionally based on the linguistic distance parameter.

$$Y = \beta_0 + \beta_1 AOM + \beta_2 Sex + \beta_3 LD + \beta_4 LD \text{ x } AOM + X'\gamma \qquad (3)$$

$$Y = \beta_0 + \beta_1 AgeOfEntry + \beta_2 Sex + \beta_3 LD + \beta_5 LD \text{ x } AgeOfEntry + X'\gamma \qquad (4)$$

The interaction terms LD x AOM and LD x AgeOfEntry stand for an increase in the effect of the linguistic origin by age at arrival, as indicated in the psychobiological literature and referred to as the Critical Period hypothesis. A positive significant interaction effect would mean that an increase in the linguistic distance also leads to an increase in the age of migration effect.

Before using a model on basis of theoretic significant variables, it is wise to check for possible statistical biases. One of these biases is multicollinearity, which is the issue of having explaining variables being correlated with one another. This leads to higher standard errors which could lead to non-significant variables in the model while they might be significant. Early regression analysis showed that there might be a chance for multicollinearity present in the models. There are two main reasons for this: first is that the social-economic status of an immigrant child is highly correlated to the education of the father and the education of the mother (see table 9). The subsequent analysis showed that removing the variables for the parents' education would lead to more observations in the model. This is because of adding the education of both parents is restricting the sample: children with one parent are excluded due to non-existing information on the second parent. Taking this in mind, the social-economic status seems a good-enough proxy for the education of the parents and is maintained in the model instead of the parents' education variables.

Table 7: Correlation matrix for the variables SES, Education of the Father and Mother

	SES	Education Father	Education Mother
SES	1.0000		
Education Father	0.6533	1.0000	
Education Mother	0.7337	0.5192	1.0000

Source: created by the author based on the data of the PRIMA cohort studies.

The second issue of multicollinearity revolved around age. Namely, this is a factor which was used for calculating the age of migration. It also is heavily correlated with the grade of a child (see table 10), which on itself is an important determinant for the language test scores. Another way of measuring multicollinearity is the variance inflation factor (VIF), which assesses how much the variance of a variable increases if they are correlated. If no factors are correlated, the VIFs will all be 1. A VIF of more than 10 is usually regarded as too much (see table 11). It was thus decided to drop the variable of age.

Table 8: Correlation matrix for the variables Grade, Age, and Age of Migration

	Grade	Age	Age of Migration
Grade	1.0000		
Age	0.9386	1.0000	
Age of Migration	0.6437	0.7352	1.0000

Source: created by the author based on the data of the PRIMA cohort studies.

Table 9: Multicollinearity Diagnostics for the variables Grade, Age, and Age of Migration

	VIF	SQRT VIF	Tolerance
Grade	1.0000	2.96	0.1144
Age	0.9386	3.34	0.0898
Age of Migration	0.6437	1.50	0.4415

Source: created by the author based on the data of the PRIMA cohort studies.

In addition, the variables grade and age of migration are strongly linked. Inherently so, due to the age of migration logically being an important determinant of the grade a child is placed in. As both variables are of importance in the model, there is no way of dropping either. One thus should take in mind that the estimated effects for the regular language test scores might be slightly more significant as the variances of grade and age of migration could be inflated. A way to circumvent this problem, is looking at the Cito test scores as the Cito is only being conducted in the last grade. Another solution would be to make separate analyses for children of each grade. Both these options will be used in the further analysis.

5. Results

The results are divided into several sections. First, the results on how the age of migration reflects in the test scores of language and math are shown. Second, the differences between the four main countries of origin in the language tests are shown. The third section checks how the model holds if the Cito test is used instead of the language test for children for all grades. The fourth and final section adds the factor linguistic distance to both the regressions on the language test as on the Cito test. Only the main results are presented in this thesis. The complete results for the regressions can be found in the Appendix.

5.1 Baseline Effects

The estimated linear effects of age of migration on various outcomes – language and math test scores for children of every age – are presented in Table 12 below. Controls for social-economic status, grade of the child, year of the cohort, and school-fixed effects are included.

Table 10: Age of Migration Effects on Test Scores

	(1) Language	(2) Math
Age of Migration	-2.847***	-0.594*
	(0.144)	(0.232)
Girl	2.571***	-3.292***
	(0.518)	(0.834)
SES - LBO Other Im~s	6.567***	0.453
	(0.771)	(1.241)
LBO Native	16.87***	8.196*
	(2.416)	(3.827)
MBO	11.21***	3.325*
	(0.817)	(1.314)
HBO/WO	16.81***	4.444**
	(0.938)	(1.505)
Unknown	6.355***	4.569
	(1.665)	(2.682)
Grade=4	62.97***	57.57***
	(0.769)	(1.221)
Grade=6	110.3***	99.01***
	(0.885)	(1.420)
Grade=8	151.2***	130.0***
	(1.049)	(1.681)
Year=1996	8.644***	25.88***
	(1.309)	(2.086)
Year=1998	13.39***	-951.8***
	(1.351)	(2.167)
Year=2000	13.96***	-954.2***
	(1.408)	(2.251)
Year=2002	16.57***	-952.0***
	(1.447)	(2.316)
Year=2004	18.37***	-956.1***
	(1.530)	(2.445)
Constant	941.1***	960.9***
	(1.464)	(2.337)
Observations	14578	14277
R-squared	0.769	0.989
Adjusted R-squared	0.751	0.988

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: created by the author based on the data of the PRIMA cohort studies.

Note: controls for social-economic status, grade and year were included. The school-fixed effects are included in the model, but are not shown in the table due to the number of clusters (about 1020). The t-values of these school-fixed effects are 0.000 for both models (see also Appendix I for the regression outputs). The first category of SES is compressed. The full names of this variable label are: SES – LBO Other Immigrants. The different educational categories are LBO – lower vocational education; MBO – secondary vocational education; HBO – higher professional education; WO – University Education.

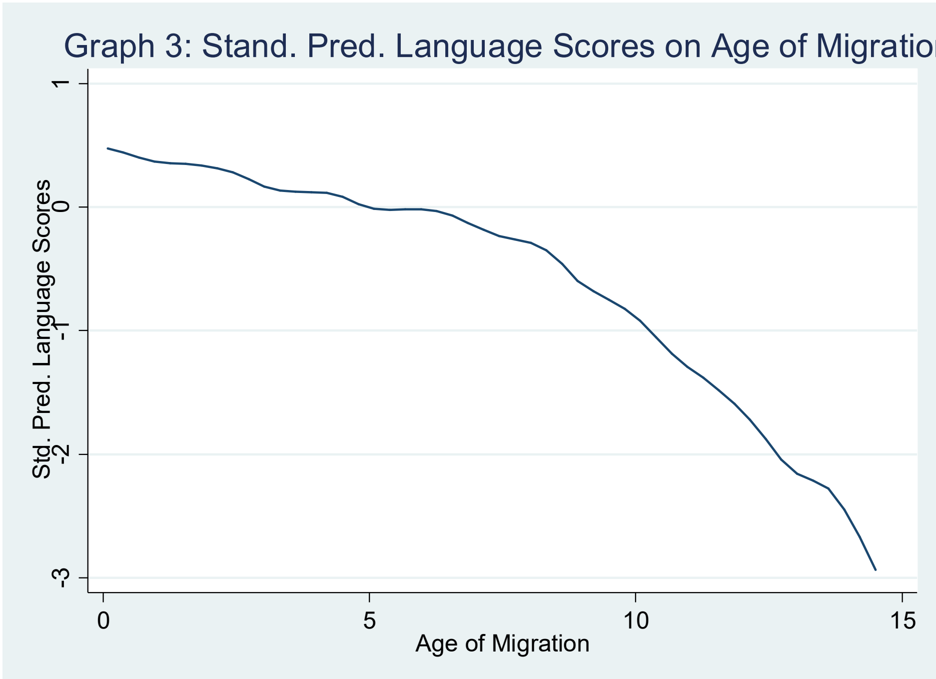
The first thing that stands out is the significant, negative continuous age of migration effect on both language and math test scores. As can be expected, the estimated age of migration effect on language test scores is stronger and more robust than the effect on math scores. Also, per expectations is the difference in gender performances between the two tests. Girls are estimated to perform better in language tests and boys in math tests. Social-economic status predicts a significant increase in the language test scores, with a noteworthy exception of MBO relative to its predecessor LBO native. This could be thanks to the division of migrant and native speakers in the previous two categories, while in MBO both are combined. Although the ability to learn might increase, the effect of being a non-native speaker might still predominate. For the math test scores, social-economic status is far less significant, which could indicate that math skills are less dependent on the social environment outside school and acquired more in the studying process at school. Next, there are the obvious significant increases in predicted test scores as children go to higher grades in primary school. The year coefficients show a small positive effect as time passes which might indicate an increase in the performance of the educational system. The negative coefficients for years in the math test are due to the change in testing systems, as explained before. The school-fixed effects are too many in number to show in one simple table. However, the regression output can be found in Appendix I where the school-fixed effects are shown to be of significant value (P-value is 0.000).

Table 11: Age of Migration Effects by Grade on Language Test Scores

	(1) Grade 2	(2) Grade 4	(3) Grade 6	(4) Grade 8
Age of Migration	-2.098***	-2.223***	-2.712***	-3.692***
	(0.464)	(0.306)	(0.244)	(0.303)
Girl	2.840*	3.898***	3.995***	0.0212
	(1.183)	(1.106)	(0.950)	(1.140)
Observations	3120	3987	3863	3608
R-squared	0.350	0.368	0.357	0.368
Adjusted R-squared	0.175	0.214	0.197	0.215

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001
Source: created by the author based on the data of the PRIMA cohort studies.
Note: controls for social-economic status, grade, year of cohort and school-fixed effects were included.

As can be seen in Table 13 above, splitting the results by grade makes a big difference. Interestingly is that the results in the eighth grade are similar for both sexes, while girl perform better than boys in all the other grades. Possibly this is due to gender-stereotype confirmation in which boys are motivated more by their teachers to perform well, while girls are taught that it’s acceptable to fail and less pressure or stimulation is put on them in order to succeed (Fox et al., 1981). The estimates for the age of migration effect remain similar to the previous estimates, although the effect seems to increase as children move to a higher grade. This makes sense as more children from a later age of migration are added to the sample at higher grades. A child arriving in the Netherlands at the age of 12 has a very low change of being put in grade 2 and so on.



Source: created by the author based on the data of the PRIMA cohort studies.
Note: controls for social-economic status, grade, year of cohort and school-fixed effects were included.

To check which specific ages are closest to a possible end of the Critical Period, and so to specify the regression, graph 3 shows the standardized predicted language test scores. The standardization of the test scores by grade creates a test score which measures the standard deviation from the average test score of all the children in a certain grade. Around age 7 or 8, the decline seems to increase which could point toward a Critical Period until the age of 7/8 so the specifications of the second model will be focused around that age period. The graph of the standardized math scores can be found in Appendix I, and shows a significant rise in the test scores as far as the age of immigration goes up. This is due to the aforementioned change in math tests throughout the cohorts. Standardizing test scores twice – once by grade, and once by year – would not result in believable results due to the standardizing process not being as accurate as the original test scores.

Table 12: Age of the End of the Expected Critical Period

	(1) Five	(2) Six	(3) Seven	(4) Eight	(5) Nine
Five	-6.988*** (0.720)				
Six		-8.804*** (0.667)			
Seven			-9.214*** (0.707)		
Eight				-11.24*** (0.774)	
Nine					-13.49*** (0.913)

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: created by the author based on the data of the PRIMA cohort studies.
Note: controls for social-economic status, grade, year of cohort and school-fixed effects were included.

Next, the regression was changed to model 2 from the Research Method. This means that the continuous variable for age of migration was replaced by a dummy, indicating if children arrived before or after a certain age. So, a dummy for the age of 7 would show what the difference is in language test scores between children who arrived before the age of 7 and children who arrived after the age of 7. In table 14, no distinct end of a Critical Period can be detected. For every age, the difference in test results is more negative, as can be expected, but there does not seem to be a clear difference which could point towards a change in the slope of the 'age of migration'-curve. In the Appendix two more tables are shown which present the results of two additional estimates with the same model but for the ages of 1 to 3, and 10 to 12. A clearer picture does not emerge from this. The data seems to point towards an almost inherent arrear from the moment an immigrant child arrives in the Netherlands

5.2 Heterogeneity by country of origin

For a closer look at each of the countries separately, the same regression models as in the previous graphs are run, but with a fixed country of origin. These countries are the four most-represented migrant groups in the Netherlands: Suriname, the Antilles, Turkey, and Morocco. For more precise estimates of the age of migration effects, we turn to the regression models as seen in Table 15. Instead of looking at the two main test sections – language and math – we focus at the language test scores and do a few regressions where the country of birth is kept fixed for one specific country of birth. The controls shown in the previous section were still used in the current regression but not shown for brevity purposes.

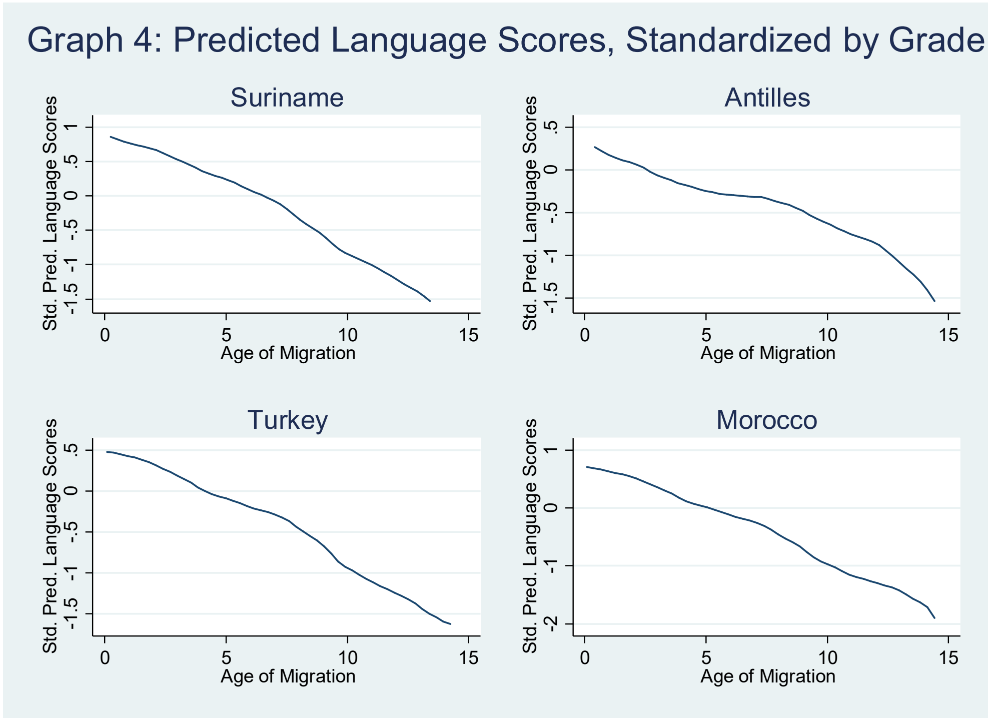
Table 13: Effect of Age of Migration per Country of Origin on language test scores

	(1) Suriname	(2) Antilles	(3) Turkey	(4) Morocco
Age of Migration	-2.499***	-2.933**	-1.711***	-2.337***
	(0.733)	(0.941)	(0.396)	(0.413)
Girl	6.674*	8.665*	2.392	2.434
	(2.585)	(3.483)	(1.253)	(1.253)
Observations	837	519	2310	2380
R-squared	0.786	0.822	0.821	0.795
Adjusted R-squared	0.708	0.718	0.789	0.764

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: created by the author based on the data of the PRIMA cohort studies.
Note: controls for social-economic status, grade, year of cohort and school-fixed effects are included.

Although the number of observations is relatively low, especially for the Antilles, some deductions about the effect of age of migration can be made. For all countries of origin, the age of effects is significant. Gender has a relatively large and significant effect in the case of Suriname and the Antilles. This contrasts with Turkey and Morocco, which have small and insignificant effects, implying that there is no difference between the performances of boys and girls in regards of language acquisition. Such a small difference might be explained by the mainly traditional culture of Turkish migrants (Crul & Schneider, 2009). Additionally, parental aspirations for both Turkish and Moroccan immigrants might play a role here as they are generally estimated to be lower for daughters than for sons. Notwithstanding that the girls themselves are not less aspiring than the boys (Phalet & Schönpflug, 2001). This could partially explain the insignificant differences for the gender variable for Turkey and Morocco: girls might be less stimulated to perform well at school, as their main function traditionally is meant to function as a husband or in the household.



Source: created by the author based on the data of the PRIMA cohort studies.
Note: controls for social-economic status, grade, year of cohort and school-fixed effects are included.

In Graph 4 above, the standardized predicted language scores are plotted against age of migration. The same model as in the previous section with the continuous age of migration variable was used. The predicted test scores were standardized to control for the differences in grade. A first look will not reveal any discernible Critical Period in either of the four countries. There are some differences to be spotted however. Looking at the y-axis, Suriname and Morocco seem to have a higher starting point (+1) compared to the Antilles and Turkey (+0.5), although Morocco is the one country which has the lowest ending point (-2) compared to the others (-1.5). Regarding the Critical Period for these four countries of origin, the tables can be found in Appendix I. From these estimates, no clear end of a Critical Period can be deducted, which confirms graph 4 where no Critical Period could be detected either. It is also similar to the previous section discussing the general case of the age of migration effect.

5.3 Cito Scores

In this section, the baseline model of section 5.1 will be used, but instead of the regular language test which is taken in all grades of primary school, the Cito test is the dependent variable. The Cito test exists of four subsets: language, math, information processing, and geography and biology. As the Cito test is only taken in the eighth grade, the variance in the model should not be dependent on grade anymore and show a more precise estimate of the age of migration effect. The estimated linear effects of age of migration on various Cito tests – total score, language, math, information processing, and geography test scores for children of grade 8 – are presented in Table 16 below. Controls for age, social-economic status, the year of the cohort, and school-fixed effects are included. The control variable for grade is removed in this model, as the Cito test is only taken in grade 8, which makes controlling for grade redundant.

Table 14: Age of Migration Effects on Cito Scores

	(1) Total	(2) Language	(3) Math	(4) Information	(5) Geography
Age of Migration	-0.983***	-1.155***	-0.683***	-0.817***	-0.691***
	(0.120)	(0.159)	(0.143)	(0.109)	(0.141)
Girl	-0.627	2.666***	-4.259***	0.315	-3.695***
	(0.435)	(0.565)	(0.509)	(0.387)	(0.498)
Observations	2446	2310	2303	2288	1847
R-squared	0.358	0.701	0.344	0.519	0.589
Adjusted R-squared	0.193	0.625	0.175	0.394	0.475

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001
Source: created by the author based on the data of the PRIMA cohort studies.
Note: controls for social-economic status, year of cohort, and school-fixed effects are included.

Age of migration has a significant, negative effect on all four different parts of the Cito tests, and also on the total score of the Cito test. The lower effects compared to the model with the regular language test exists because the scale of grading for the Cito tests is vastly smaller (920-1100 versus 30-60 points can be scored on respectively the regular and the Cito language test). The number of observations are significantly lower than in previous regressions due to only children in grade 8 having to take the Cito test. Age of migration has the largest effect on the language test scores, as expected.

Gender also plays a significant role in determining the level of the Cito scores. It is particualrly interesting to note that for the the total score at the Cito test there is no such difference, but there are marked differences for particular subsets of the Cito test. The language test is made significantly better by girls, whereas boys perform relatively better at the math and geography tests. The differences in math and language tests between gender correspond with the previous findings that boys seems to be better at math and girls better in languages. Apparently such a difference is also present in geography. This might be explained by girls being more socially interactive from a younger age which helps them especially in social-skills like learning a new language. While on the other hand, if they are slightly stimulated by their parents to perform at school, they would not necessarily perform as well as boys in other subjects at school.

For comparison of the intensity of the age of migration effects, a graph combining the four subparts of the Cito test can be found in Appendix II. Here it becomes more clear that the language test scores are most effected by the age of migration effect. Where in the language department of the Cito, scores drop by half as the age of migration increases, the other sections show a decrease of about a quarter.

5.4 Effect of linguistic distance

In this section, the third model is used. Here the variable of linguistic distance is added to the regression and we test if the factor has any significant value against several tests: the regular language test for children of every grade, the Cito test for children of the eighth grade, and the language subset of the Cito test. The estimated effects are presented in table 17 below. Controls for social-economic status, year of cohort, grade of the child, and school-fixed effects are included.

Table 15: Linguistic Distance Effect on Language Scores

	Language Test	Language Test	Cito Test Overall Score	Cito Test Language Score
Age of Migration	-2.489***	-7.994	-1.102***	-1.343***
	(0.216)	(4.245)	(0.181)	(0.249)
Linguistic Distance	-0.255***	-0.309***	0.0242	-0.00323
	(0.0363)	(0.0553)	(0.0298)	(0.0406)
Girl	2.940***	2.935***	-1.150	1.940*
	(0.736)	(0.736)	(0.596)	(0.798)
Interaction Effect		0.0103		
		(0.00791)		
Observations	6994	6994	1262	1180
R-squared	0.785	0.785	0.409	0.703
Adjusted R-squared	0.762	0.762	0.215	0.608

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for social-economic status, year of cohort, and school-fixed effects are included. Controls for grade are included only for model 1 and 2. The interaction effect constitutes the combined effect of age of migration with linguistic distance. The first model consists of the regular language test which is taken by children from all grades.
* p<0.05, ** p<0.01, *** p<0.001

When linguistic distance is added to the regular language test, it shows to have a significant, negative effect on the test scores. This means that the further away languages are from one another, the more difficult it is to acquire Dutch as a second language, indicating that it matters which language a child speaks or is familiar with in the country of origin.

In the second regression, an interaction effect is added to test if there is a combined effect of linguistic distance and the age of migration. In theory, it could be increasingly more difficult to learn Dutch for children with a high linguistic distance arriving at a later age in the Netherlands. Or children with a certain linguistic distance might arrive sooner or later than average. However, the interaction effect is insignificant and of a low value. This implies there is no or little contact between the two variables and the effect on the language test scores is low.

Adding the factor 'linguistic distance' to the models of the Cito test scores results in insignificant variables. This difference of the effect of linguistic distance test might be explained due to the Cito test not being completely focused on language, so that immigrant children can make up any possible negative differences in the language section with good performances in the math section. The insignificant effect of linguistic distance for the language section in the Cito test might possibly be caused by the set-up of the Cito test, and the grade or age of the child when taking the Cito test. Through the process of grade retention and skipping, only children of a certain skill-level can partake in the Cito test. So, where the regular language test allows for all children to participate, there is a certain minimum level of skills or ability children need to have before doing the Cito test.

6. Sensitivity Analysis

In the sensitivity analysis, the main model will be tested for robustness with several background characteristics. Afterwards, the model will be run without school-fixed effects. Lastly, the functional form of the model will be checked. The former is tested with all the available and relevant variables. Relevant variables herein mean that there will be no age of migration with age of children interaction effect due to the resulting multicollinearity. This also means the model cannot be tested against variables such as the region of country the school resides in or other possible household and school characteristics due to the limited availability of uniform variables in the dataset. The second test is an assessment of regression performance using clustered standard errors versus school-fixed effects. The last check is for functional form, which will be tested with Ramsey's RESET test and Linktest. If the results from these tests are insignificant, there is limited risk to the model being incorrectly specified.

6.1 Robustness Checks

In this section, the baseline model will be tested with several interaction effects to see if the previously found results are subject to change due to alterations in other variables. For example, the significant interaction effects in table 16 below confirm the former hypothesis that gender and age of migration have different effects on children in different grades. The table also indicates that both effects do not change over time; the age of migration and gender effects do not change significantly from 1994 to 2004. Alternatively, one can find similar tests for social-economic status, being a child in a school with a relative high number of high-performers, education of the father, and education of the mother in Appendix I. Except for education of the father and gender, these tests do not show any significant relationship between the variables. A more detailed version of how the effects per grade look, one can find in table 11, page 21 of this thesis.

Table 16: Interaction Effects with Year of Cohort and Grade

	Age of Migration – Year	Gender – Year	Age of Migration - Grade	Gender – Grade
Age of Migration	-43.82	-2.847***	-1.464***	-2.851***
	(67.25)	(0.144)	(0.367)	(0.144)
Girl	2.570***	56.28	2.545***	5.043***
	(0.518)	(334.9)	(0.518)	(1.309)
Interaction Effect	0.0205	-0.0269	-0.250***	-0.486*
	(0.0336)	(0.167)	(0.0609)	(0.236)
Observations	14578	14578	14578	14578
R-squared	0.769	0.769	0.769	0.769
Adjusted R-squared	0.751	0.751	0.752	0.752

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for social-economic status, year of cohort, grade, and school-fixed effects are included. The interaction effect constitutes of the two variables mentioned at the top of the table.
* p<0.05, ** p<0.01, *** p<0.001

To further specify the relationship between age of migration and gender and year of cohort, the results can be seen by year in table 17. One might expect beforehand that integration and/or integration procedures improve over time and thus influence the age of migration negatively, and possibly result in more gender-equality in migrant households as migrants turn more towards a more "Western"-like cultural value of the household and gender equality. However, the data in table 17 shows that this is not the case. The age of migration seems worse in the early years, but does increase slightly towards 2004. The gender effects on language test scores seems to be rather weak and are difficult to explain within the current dataset. Although girls seem to perform significantly better in 2000 and 2002, the cohort of 2004 results in an insignificant difference in gender. In Appendix I, one can also find the effects by year for each of the four main countries of origin (Suriname, the Antilles, Turkey, and Morocco). These do not show any significant effect of gender on language test results if split by year (except for Suriname and the Antilles in 2002). This means that these data do not support any conclusion towards an increasing move towards “Western”-values in regards of gender equality as there does not seem to be any trend developing towards better performances of girls in language test scores.

Table 17: Language Test Effects by Year of Cohort

	(1) 1994	(2) 1996	(3) 1998	(4) 2000	(5) 2002	(6) 2004
Age of Migration	-3.292***	-3.839***	-2.637***	-2.689***	-2.495***	-2.851***
	(0.494)	(0.358)	(0.334)	(0.342)	(0.371)	(0.475)
Girl	3.229	2.283	1.089	3.435**	3.860***	1.358
	(1.725)	(1.343)	(1.292)	(1.156)	(1.133)	(1.351)
Observations	1299	2351	2633	2839	3197	2259
R-squared	0.828	0.820	0.776	0.779	0.784	0.780
Adjusted R-squ~d	0.765	0.777	0.730	0.746	0.754	0.742

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001
Source: created by the author based on the data of the PRIMA cohort studies.
Note: controls for social-economic status, year of cohort, and school-fixed effects are included.

Another effect influencing the language test scores could be the month of birth. Namely, the Netherlands use an education system in which children until the month of October are being allowed into first grade, while children born after said month must wait for the next grade to start in the next year. This means that children born early in the school-year (October, November), are almost an entire year older than children born at the end of the year (August, September). The latter are more often sent to lower education schools thanks to lower performances at school (Luyten et al., 2013; Bedard & Dhuey, 2006). This does not seem to have any effect on the age of migration effect, as can be seen in table 18 below in the insignificant interaction effects. However, there is a significant co-relation between gender and month of birth. This is not unfamiliar to previous studies; Plug (2001) argued that girls mature at an earlier age and thus have a bigger month of the birth effect.

For differences of country of birth, we have seen in section 5.2 how the estimates differ between four main countries of origin. In the table below, one can see that the age of migration and gender effects statistically differ depending on the country of origin. Reasons as to why certain countries face more difficulties in regard of the age of migration effect can range from cultural values to educational quality or linguistic origin. Without a more precise and richer dataset, the exact reasons are difficult to pinpoint. In theory, the same goes for the differing effect of gender, except that more research has been done in this area. From the related literature, one can make an educated guess that a more conservative or traditional culture might form some blockade for girls to perform consistent with their actual ability.

Table 18: Interaction Effects with Migrant Schools and Month of Birth

	Age of Migration – Migrant School	Gender – Migrant School	Age of Migration – Month of Birth	Gender – Month of Birth	Age of Migration – Country of Birth	Gender – Country of Birth
Age of Migration	-2.577***	-2.849***	-2.902***	-3.133***	-3.013***	-2.585***
	(0.225)	(0.144)	(0.236)	(0.149)	(0.264)	(0.160)
Girl	2.579***	3.352***	2.485***	5.310***	2.157***	4.737***
	(0.519)	(0.916)	(0.517)	(1.104)	(0.563)	(1.237)
Interaction Effect	-0.357	-1.123	-0.406	-5.279**	0.0476*	-0.294*
	(0.227)	(1.112)	(0.325)	(1.827)	(0.0234)	(0.125)
Observations	14526	14526	14578	14578	11706	11706
R-squared	0.769	0.769	0.770	0.770	0.783	0.783
Adjusted R-squared	0.751	0.751	0.752	0.752	0.764	0.764

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for social-economic status, year of cohort, grade, and school-fixed effects are included. The third and fourth regressions also control for month of birth. The fifth and sixth regressions additionally to the baseline controls, also include dummies for country of birth. The interaction effect constitutes the combined effect of the two variables mentioned at the top of the table. A migrant school is defined as such if more than 25% of the children of the school sample are originally non-native.
* p<0.05, ** p<0.01, *** p<0.001

6.2 Clustered Standard Errors versus School-Fixed Effects

In this section the model will be tested with clustered standard errors instead of controlling for school-fixed effects. This model will be compared with the baseline model. This is done because the differences between language test results can partially be caused by differences in school quality. This hinges on a variety of aspects like teacher quality, intelligence of other pupils (peer-effects), price and quality of lunches, etc. All these factors are not uniformly present in the database, but there is data available that shows which children were together in school. This presents two possibilities: the first is to cluster the standard errors by school. The second possibility is to use the school-variable as a fixed-effect so that all unobservable school-effects are controlled for.

Table 19: Clustered OLS vs. FE

	(1) OLS	(2) FE
Age of Migration	-2.825***	-2.847***
	(0.185)	(0.144)
Girl	2.464***	2.571***
	(0.616)	(0.518)
Observations	14578	14578
R-squared	0.733	0.769
Adjusted R-squared	0.733	0.751

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001
Source: created by the author based on the data of the PRIMA cohort studies.
Controls for social-economic status, year of cohort, grade, and school-fixed effects are included.

Although the results of an OLS and FE present similar estimates for the effect of age of migration, the OLS model with clustered standard errors might be vulnerable to omitted variable bias. This is because the age of migration could be correlated with many individual- and school-level factors affecting a student's language test scores. It might for example be linked to the student's ability, which is influenced by parents' income and/or education. More importantly, the age of migration might be correlated with unobserved aspects of school quality. Using fixed effects at the school level to disentangle school from individual effects is a possible solution to this problem: they solve for the problem of school-level omitted variables and unobservables.

Table 20: FE with only Dummies for School Clusters

	(1) Language
Constant	1031.7***
	(0.456)
Observations	15660
R-squared	0.140
Adjusted R-squared	0.078

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: created by the author based on the data of the PRIMA cohort studies.

The adjusted R-squared of this model regressing language scores on school dummies is only 0.078. The adjusted R-squared shows the proportion of the variance in the dependent variable explained by the school indicators (adjusted for the degrees of freedom), this means that 7.8% of the variance in language test scores is between schools. Symmetrically, this means that 92.2% of the variance is within schools. All other variables are measured on an individual level. Consequently, running regular OLS would not be appropriate: without the robustness correction, it would show incorrect standard errors and statistical tests. Even with robust standard errors, it wouldn't present efficient estimates of these parameters.

6.3 Model Specification Tests

Another way to test the model, is with specification tests. One of these is the Linktest. This checks whether more variables are needed in the model by running a new regression with the observed language test scores against the predicted language test scores and the predicted language test scores-squared as independent variables. If the model is accurately specified, the predictor squared shouldn't have much predictive power apart from chance. The null hypothesis is that there is no specification error. If the p-value of said variable is not significant, then the null is rejected and it can be concluded that this model is correctly specified. As the p-value is insignificant (higher than 0.05), predictor squared does not have explanatory power, indicating that the model is specified correctly.

Table 21: Linktest

	Linktest
Predictor	0.712 ***
	(0.225)
Predictor Squared	0.0001
	(0.0001)
Constant	147.141
	(114.822)
R-Squared	0.7692
Adjusted R-Squared	0.7517
Observations	14,578

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses.
Controls for social-economic status, grade, year of cohort, and school-fixed effects are included.
* p<0.05, ** p<0.01, *** p<0.001

Ramsey's RESET test performs another test of regression model specification. The idea behind it is close to the Linktest. It generates new variables grounded on the predictors and refits the model with those new variables, checking if they are significant. The null hypothesis is that the model is correctly specified, but the p-value is 0.0003 which is below the usual threshold of 0.05 (95% significance), so that the null hypothesis is rejected. It can be concluded that the functional form of the model is incorrect. The rejection of the null hypothesis of the RESET test can be easily explained by the non-linear form of the model as can be seen in the graphs above. As the age of migration increases in strength over age, the curve cannot be linear.

7. Discussion

There are of course some caveats which should be mentioned in the interpretation of the main results of this paper. These caveats are mostly inherent to the identification strategy and the available data. Many researchers studying this subject encountered problems with self-selection of parents. This is the idea that parents with certain qualifications or characteristics which influence the performances of their children (such as higher education or income), might choose to migrate at an earlier age of their child. In other words, if highly-educated parents with children persistently choose to migrate earlier than lowly-educated parents, the effect measured will be biased upwards. Also, parents who have the choice of making such considerations (e.g., labour-market immigrants) are more likely to migrate with young children compared to parents who do not have this choice (e.g., refugees) (Böhlmark, 2008). An OLS approach fails to take this into account.

One of the possible ways to remove the selection bias would be to use a Fixed-Effects method. This could mean using within-sibling analysis to estimate how age of migration would influence the ability to learn Dutch as a second language. Obviously, twins should be excluded as one would expect them to arrive at the same age in The Netherlands. As the parents are the same for siblings, the selection bias would be identical for all the siblings of one family. Comparing the immigration-age effects between the siblings would then cancel out the selection bias. With the current PRIMA-dataset this is not possible: there are no uniform identifiers for siblings throughout the cohorts. This is a similar problem as with the other household characteristics which are only present in the first cohort for all grades. With the existing variables, it is impossible to create a decent proxy identifying siblings in the database.

The same logic goes for identifying twins. Webbink et al. (2006) showed that while using the same PRIMA-database, matching without family name proves to be incomplete and incorrect. By using all available background variables, they constructed a matching procedure similar to the correct method of using family name, date of birth, school and year of survey. The procedure does not function well for two reasons: it does not identify all twins, and it does the opposite, which is identifying twins who are not twins. Adding more variables to identify twins with results in a higher share of not identified twins, likely due to coding errors and missing values. As the dataset used in this thesis family names are absent, there is no sure way of identifying twins or siblings correctly.

Another solution would be to do research in a country or region which has an age restriction in their immigrant selection procedure. In such a case, a regression-discontinuity design could be used to get a more precise estimate of the age of migration effect. As such policies are generally considered unethical, finding such a country or region to research might prove to be difficult. Instead, with the recent implementation of granting a 'pardon' to the children of immigrants, there might lie some possibilities in the cut-off age with which children are legally considered eligible for such an arrangement. Another immigrant selecting system is the point-system used by Canada and Australia, among others. If such a system would be adopted, an extra point could be awarded for children under the age of 7. Nevertheless, it does not seem very plausible that the cut-off age will soon be the age of 7. A more plausible research method would be to set-up an experiment where a random number of potential immigrants could be sampled. To avoid the self-selection procedure, an experiment could be set up such that random people from certain countries asked to migrate to the Netherlands. This should be done in such a way because random sampling would already be non-random when the sample would be selected from the already-arrived immigrants. This is thanks to the decision of when a migrant arrives would already have been made by the migrant. Of course, a different self-selection process takes place between people not willing to migrate and those who are willing. However, this is not a relevant self-selection as the people who do not wish to migrate are unlikely to be willing to learn Dutch as a second language anyway – also seeing how Dutch is not a widely-used language in the world.

Another caveat for the results focusing on the performances of migrant girls, might be the increasing emancipation of women over time. Although this study does not find any indication of increasing equality in gender over time, this does not mean such an effect does not exist at all or did not take place in a significant manner after the year of 2004. Also, when focusing on the four main countries of origin. As stated by the Dutch Central Agency of Statistics, significant progress has been made since the time the sample collection ended (Fransman, 2016). Not only among the first-wave immigrants from Turkey and Morocco the views on gender roles has changed in a progressive way, but even more so for the second-wave immigrants (also known as the children of the first-arrivals). In the same report of the Agency, one can still see the differences in emancipation between the former colonies on one side, and Turkey and Morocco on the other side. This indicates that the division in the data seems valid, but that the database itself might be slightly outdated. For further research, it is advised to use a more up-to-date database to ensure current views on gender roles are included.

The third caveat comprises the set-up of the language test. Ideally, a language test would be taken in all grades with a uniform grading system throughout the grades, so that there are no changes in test scores over age which could possibly conflate the results. The current set-up of the language test has a non-standard grading system in which children from higher grades perform better. Although this might make sense on first sight, it makes comparing children from different grades difficult. A uniform grading system throughout the whole school system would solve this issue.

Seeing as a selection bias seems to be present in the model, the results likely point towards the right direction, but the accuracy might not be as high as hoped for. If parents with positive characteristics migrate before the child reaches the age of 7, then this would result in a database which is biased towards better-off children. Most likely is that the overall estimates are biased upwards due to the selection bias, as a completely random sample would feature more higher-ability children in the late-arrival subsample. It does however not explain the non-occurrence of the Critical Period in the data, as children who arrived before the age of 7 should have even higher language test scores than the children who arrived after. Nevertheless, more research (and data) into the different factors influencing the decision to migrate at a certain age of their children could be helpful. If these variables (or proxies of these variables) could be used in a database, an examination of the children can be done controlling for a large portion of the independent variables.

8. Conclusion

The main conclusions of this paper come at the backdrop of an increase of interest in the political debate about the so-called “migrant crisis” and the integration of immigrants in the Netherlands. In the upcoming elections in the Netherlands, one of the hot topics is the timing of the start of integration processes and language lessons of newly-arrived migrants. The key part of this integration process is learning Dutch, as acquiring Dutch is the minimum requirement for being officially integrated as stated by the official website of the Dutch government. (Rijksoverheid, 2016). Certainly, policies and policy changes can only render the desired outcomes if one precisely understands how integration processes are affected. The results of this paper show that the longer a child must wait with lessons to learn Dutch as a second language, the worse the child will perform. Put in a more positive way: if one starts learning Dutch as soon as possible, a child will be able to learn the language far easier and will perform more on par with his native classmates.

This thesis examined if the age of migration of children matters for assimilation in the Netherlands. Evidence is found that the age of migration effect is significantly negative on language test scores, and that it increases as the age of migration rises. No clear evidence is found for a Critical Period Hypothesis around the age of 7 or 8. Therefore it remains unclear if there is a need for additional educational help for children who specifically arrived after the hypothesized Critical Period. The underlying factors and the consequences of the results are unknown and more research is needed to see whether policy measures should be implemented.

The analysis on country of origin shows that it significantly matters what the migrant’s country of origin is. There are significant differences in second language learning, which point towards an easier integration of immigrants from countries which have strong links to the host country. Both country of origin and linguistic distance play a role here. This could be an important result for policymakers, as this might guide extra educational help to immigrant children from countries with a lower linguistic similarity. With more advanced research, children from certain countries of origin can be identified as ‘high-risk’-countries. Children from these countries could be offered additional help learning Dutch from an earlier age, as they can be found to be more prone to the age of migration effect.

The research finds that gender plays a significant role regarding general language test scores: being a girl correlates significantly with higher language test scores. There is a difference herein when it comes to country of origin: girls from Suriname and the Antilles perform better than boys of those countries. Whereas there is no significant difference in gender for immigrants from Turkey and Morocco. Possibly this is due to a relatively more traditional role for women in their societies. Conceivable policy improvements should involve the parents to create a mutual understanding of how to best stimulate children of both genders. Similar to the ‘Overstap’ policy where parents receive lessons on effective reading learning strategies. However, if traditional culture is the determining factor which prevents girls from performing as they potentially could, attitudes might be difficult to change. A potential solution might be maximizing the educational environment and time at school to compensate for the household morale. This could be done in extra after-class lessons or additional weekend classes.

Also, noteworthy related to gender is the decrease in significance of gender in the age of migration effect as the grade increases. This could either point towards gender-stereotype confirming bias where boys’ performances are valued more (societal influence), or towards a natural catch-up of boys in the acquirement of a second-language (genetic approach). Either way, further research might try to track children over a longer period to see if these gender differences in language-acquirement remain over a longer period. Moreover, further research could better identify whether the causes for similar performances at a later age lie in worse performances of girls or in a better performance of boys.

In conclusion, this thesis shows that policy makers who are focusing on educational and integration policies, should pay attention to a certain set of variables (among others). These are the age of migration, the linguistic distance, and gender in combination with the country of birth. The linguistic distance effect means that the language the child speaks as a first language should be taken into account. This is due to the increasing difficulties arising from learning a new language, if the first-known language is more unlike the second language. The effect of gender in combination with country of birth indicates that certain cultural blockades might prevent girls from performing consistent with their ability. To circumvent this problem, additional educational policies could be directed towards additional hours at school or towards directly confronting households with these gender inequality issues.

References

Angrist, J. D., & Pischke, J. S. (2008). Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Basu, S. (2010). Age-of-Entry Effects on Education of Immigrant Children: A Sibling Study. *Available at SSRN 1720573*.

Bedard, K., & Dhuey, E. (2006). The persistence of early childhood maturity: International evidence of long-run age effects. *The Quarterly Journal of Economics*, *121*(4), 1437-1472.

Birdsong, D. 1999. ―Introduction: Whys and why nots of the critical period hypothesis for second language acquisition. *Second Language Acquisition and the Critical Period Hypothesis*. D. Birdsong (ed.). Mahwah (N.J.). Lawrence Erlbaum Associates. p. 1–22.

Birdsong, D. 2006. ―Age and second language acquisition and processing: A selective overview, *Language Learning*. Vol. 56. p. 9–49.

Bleakley, H., & Chin, A. (2010). Age at arrival, English proficiency, and social assimilation among US immigrants. *American economic journal. Applied economics*, *2*(1), 165.

Böhlmark, A. (2008). Age at immigration and school performance: A siblings analysis using Swedish register data. *Labour Economics*, *15*(6), 1366-1387.

Bound, J. and Solon, G. (1999). Double Trouble: On the Value of Twins-based Estimation of the Returns of Schooling. *Economics of Education Review*, 18, 169.182.

CBS. (2015). *Jaarraport 2015 Landelijke Jeugdmonitor.* The Hague: CBS.

Chapman, M., Filipenko, M., McTavish, M., & Shapiro, J. (2007). First graders' preferences for narrative and/or information books and perceptions of other boys' and girl's book preferences. *Canadian Journal of Education*, 30(2), 531-553.

Coenen, L. (2001). *Word niet zoals wij: de veranderende betekenis van onderwijs bij Turkse gezinnen in Nederland* [Do not become like us: The changing meaning of education in Turkish families in the Netherlands]. Amsterdam: Het Spinhuis.

Corak, M. (2011). Age at Immigration and the Education Outcomes of Children. *Analytical Studies Branch Research Paper Series 2011336e*, Statistics Canada, Analytical Studies Branch.

Crul, M. (2000). *De sleutel tot succes. Over hulp, keuzes en kansen in de schoolloopbanen van Turkse en Marokkaanse jongeren van de tweede generatie* [The key to success: On support, choices, and chances in the school careers of Turkish and Moroccan youth of the second generation]. Amsterdam: Het Spinhuis.

De Bot, K., Evers, R., De Quay, P., & Van der Slik, F. W. P. (2005). Engelse taalvaardigheid in de onderbouw: Nederland vergeleken met zes andere Europese landen [English language proficiency in junior high school: The Netherlands compared to six other European countries]. *Levende Talen Tijdschrift*, *6*, 3 – 17.

Dahl, G. and Moretti, E. (2008). The Demand for Sons. The Review of Economic Studies 75: 1085-1120.

Driessen, G., van Langen, A., & Vierke, H. (2000). Basisonderwijs: veldwerkverslag, leerlinggegevens en oudervragenlijsten: basisraportage PRIMA-cohortonderzoek: derde meting 1998/99. ITS.

Ellis, R. (2015). *OAL:* Understanding Second Language Acquisition 2nd Edition*: Oxford Applied Linguistics*. Oxford University Press.

Euwals, Rob and Dagevos, Jaco and Gijsberts, Mérove and Roodenburg, Hans, Immigration, Integration and the Labour Market: Turkish Immigrants in Germany and the Netherlands (March 2007). IZA Discussion Paper No. 2677. Available at SSRN: https://ssrn.com/abstract=978762 or http://dx.doi.org/10.2139/ssrn.978762

Fox, L., Brody, L. and Tobin, D. (Eds) 1981 *Women and the Mathematical Mystique*, John Hopkins University Press, London.

Fransman, R. (2016, 12 22). *Het gaat ook heel goed met de integratie in Nederland.* Opgehaald van CBS.nl: https://www.cbs.nl/-/media/_pdf/2016/51/column-robin-fransman.pdf
Source of the table mentioned in the thesis:
http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=80270ned&D1=0-2,24-30,37-38&D2=0-3&D3=0-2,6-7&D4=a&HDR=T&STB=G1,G2,G3&VW=T

Frick, J. R., & Wagner, G. G. (2001). Economic and social perspectives of immigrant children in Germany.

Friedberg, R.M., 2000. You can't take it with you? Immigrant assimilation and the portability of human capital. J. Labor Economics. 18, 221–251.

Gijsberts, M. (2004). De onderwijspositie van vrouwen uit etnische minderheden [The educational position of women from ethnic minorities]. In M. Gijsberts & A. Merens (Eds.), *Emancipatie in estafette. De positie van vrouwen uit etnische minderheden* (pp. 31–58). Rijswijk, Netherlands: Social and Cultural Planning Office.

Gonzalez, A. (2003). The education and wages of immigrant children: the impact of age at arrival. *Economics of Education Review*, *22*(2), 203-212.

Hessels, T. (2005). Voormalig Joegoslaven in Nederland. *Bevolkingstrends*, *53*(1), 98-103.

Isphording, I. E. (2014). Disadvantages of linguistic origin—Evidence from immigrant literacy scores. *Economics Letters*, *123*(2), 236-239.

Kamphuis, F., Mulder, L., Vierke, H., Overmaat, M., & Koopman, P. (1998). De relatie tussen PRIMAtoetsen en toetsen uit het CITO-leerlingvolgsysteem. Arnhem/Nijmegen/Amsterdam: CITO, ITS, SCO-Kohnstamm Instituut

Luyten, H., Staman, L., & Visscher, A. J. (2013). Leerachterstanden van vertraagde leerlingen op normaalvorderende leeftijdgenoten. *Pedagogische studiën*, *90*(5), 45-57.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive psychology*,*21*(1), 60-99.

Lindo, F. (2000). The silent success: The social advancement of Southern European labour migrants in the Netherlands. In H. Vermeulen & R. Penninx (Eds.), *Immigrant integration: The Dutch case* (pp. 132–152). Amsterdam: Het Spinhuis.

McGeown, S., Goodwin, H., Henderson, N., & Wright, P. (2012). Gender differences in reading motivation: Does sex or gender identity provide a better account? *Journal of Research in Reading, 35(3)*, 328-336.

Meelissen, M., Netten, A., Drent, M., Punter, R., Droop, M., & Verhoeven, L. (2012). *PIRLS- en TIMMS-2011: Trends in leerprestaties in lezen, rekenen en natuuronderwijs*. Nijmegen: Radboud Universiteit.

Mincer, J. (1974). Schooling, Experience, and Earnings. Human Behavior & Social Institutions No. 2.

Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton University Press.

Plug, E. J. (2001). Season of birth, schooling and earnings. *Journal of Economic Psychology*, *22*(5), 641-660.

Phalet, K., & Schönpflug, U. (2001). Intergenerational transmission of collectivism and achievement values in two acculturation contexts: The case of Turkish families in Germany and Turkish and Moroccan families in the Netherlands. *Journal of cross-cultural psychology*, *32*(2), 186-201.

Rijksoverheid. (2016, 08/23). *Inburgering en integratie van nieuwkomers*. Opgehaald van rijksoverheid.nl: https://www.rijksoverheid.nl/onderwerpen/nieuw-in-nederland/inhoud/inburgering-en-integratie-van-nieuwkomers

Sax, L. (2005). *Why gender matters: What parents and teachers need to know about the emerging science of sex differences*. New York: Broadway Books.

Schaafsma, J., & Sweetman, A. (2001). Immigrant earnings: age at immigration matters. *Canadian journal of Economics*, 1066-1099.

Schepens, J. J. (2015). *Bridging linguistic gaps: The effects of linguistic distance on adult learnability of Dutch as an additional language*. Utrecht: LOT.

Schepens, Job, Frans Van der Slik, and Roeland Van Hout. 2013. The effect of linguistic distance across Indo-European mother tongues on learning Dutch as a second language. In Lars Borin and Anju Saxena (eds.), *Comparing Approaches to Measuring Linguistic Differences,* 199–230. Berlin: Mouton de Gruyter.

Singleton, C., Horne, J., & Thomas, K. (1999). Computerized baseline assessment of literacy. *Journal of Research in Reading, 22(1)*, 67-80.

Singleton, D. (2005). The Critical Period Hypothesis: A coat of many colours. *International Review of Applied Linguistics in Language Teaching*, *43*(4), 269-285.

Thomas, M. S., & Johnson, M. H. (2008). New advances in understanding sensitive periods in brain development. *Current directions in psychological science*, *17*(1), 1-5.

Van Ours, J. C., & Veenman, J. (2006). Age at immigration and educational attainment of young immigrants. *Economics Letters*, *90*(3), 310-316.

Voyer, D., & Voyer, S. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin, 140(4)*, 1174-1204.

Webbink, D., Roeleveld, J., & Visscher, P. M. (2006). Identification of twin pairs from large population based samples. *Twin Research and Human Genetics*, 9, 496−500.

Appendix I – Tables

Table A: Regression Output for the continuous Age of Migration Effect on Language Scores, all controls included

	Baseline
Age of Migration	-2.847***
	(0.144)
Girl	2.571***
	(0.518)
Social Economic Status	
LBO – Other Immigrants	6.567***
	(0.771)
LBO Natives	16.875***
	(2.416)
MBO	11.208***
	(0.817)
HBO/WO	16.812***
	(0.938)
Unknown	6.355***
	(1.670)
Grade	
4	62.967***
	(0.769)
6	110.259***
	(0.885)
8	151.169***
	(1.049)
Year	
1996	8.645***
	(1.309)
1998	13.389***
	(1.351)
2000	13.960***
	(1.408)
2002	16.566***
	(1.447)
2004	18.373***
	(1.530)
Constant	941.109***
	(1.464)

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for school-fixed effects are included. Number of observations is 14,578.
* p<0.05, ** p<0.01, *** p<0.001

Table B: Regression Output for the continuous Age of Migration Effect on Math Scores, all controls included

	Baseline
Age of Migration	-0.594*
	(0.232)
Girl	-3.291***
	(0.834)
Social Economic Status	
LBO – Other Immigrants	0.453
	(1.241)
LBO Natives	8.196*
	(3.827)
MBO	3.325*
	(1.314)
HBO/WO	4.444**
	(1.505)
Unknown	4.569
	(2.682)
Grade	
4	57.568***
	(1.221)
6	99.010***
	(1.420)
8	130.025***
	(1.681)
Year	
1996	25.881***
	(2.086)
1998	-951.843***
	(2.167)
2000	-954.160***
	(2.251)
2002	-951.979***
	(2.316)
2004	-956.128***
	(2.445)
Constant	960.880***
	(2.337)

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for school-fixed effects are included. Number of observations is 14,277.
* p<0.05, ** p<0.01, *** p<0.001

Table C – The Age of the end of the Critical Period (1-4)

	(1) One	(2) Two	(3) Three	(4) Four
One	-8.447***			
	(1.773)			
Two		-3.859***		
		(0.932)		
Three			-4.362***	
			(0.855)	
Four				-6.647***
				(0.723)

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for social-economic status, grade, year of cohort and school-fixed effects were included. Number of observations is 14,578.

Table D – The Age of the End of the Critical Period (10-12)

	(1) Ten	(2) Eleven	(3) Twelve
Ten	-15.11***		
	(1.110)		
Eleven		-16.69***	
		(1.610)	
Twelve			-19.66***
			(2.538)

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for social-economic status, grade, year of cohort and school-fixed effects were included. Number of observations is 14,578.

Table E: Critical Period for Suriname

	(1) Five	(2) Six	(3) Seven	(4) Eight	(5) Nine
Five	-3.890				
	(3.863)				
Six		-9.764**			
		(3.305)			
Seven			-9.349**		
			(3.613)		
Eight				-12.08**	
				(3.867)	
Nine					-13.73**
					(4.335)

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for social-economic status, grade, year of cohort and school-fixed effects were included. Number of observations is 837.

Table F: Critical Period for the Antilles

	(1) Five	(2) Six	(3) Seven	(4) Eight	(5) Nine
Five	-12.42**				
	(4.582)				
Six		-13.20**			
		(4.632)			
Seven			-8.124		
			(4.964)		
Eight				-11.33*	
				(5.091)	
Nine					-4.273
					(5.625)

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for social-economic status, grade, year of cohort and school-fixed effects were included. Number of observations is 519.

Table G: Critical Period for Turkey

	(1) Five	(2) Six	(3) Seven	(4) Eight	(5) Nine
Five	-4.319*				
	(1.828)				
Six		-7.287***			
		(1.725)			
Seven			-6.416***		
			(1.872)		
Eight				-8.328***	
				(2.132)	
Nine					-7.276**
					(2.580)

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for social-economic status, grade, year of cohort and school-fixed effects were included. Number of observations is 2310.

Table H: Critical Period for Morocco

	(1) Five	(2) Six	(3) Seven	(4) Eight	(5) Nine
Five	-4.624*				
	(1.936)				
Six		-4.452*			
		(1.783)			
Seven			-7.288***		
			(1.977)		
Eight				-11.11***	
				(2.302)	
Nine					-11.63***
					(2.907)

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for social-economic status, grade, year of cohort and school-fixed effects were included. Number of observations is 2380.

Table I: Interaction Effects for High Performing Cito Schools and Social-Economic Status

	Age of Migration – High Cito Scoring Schools	Gender – High Cito Scoring Schools	Age of Migration – Social Economic Status	Gender – Social Economic Status
Age of Migration	-2.776***	-2.846***	-2.948***	-2.847***
	(0.151)	(0.144)	(0.219)	(0.144)
Gender	2.561***	2.444***	2.570***	-2.597*
	(0.518)	(0.560)	(0.518)	(1.027)
Interaction Effect	-0.460	0.883	0.0374	-0.00980
	(0.294)	(1.477)	(0.0608)	(0.330)
Observations	14578	14578	14578	14578
R-squared	0.769	0.769	0.769	0.769
Adjusted R-squared	0.751	0.751	0.751	0.751

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for social-economic status, year of cohort, and school-fixed effects are included. A dummy for schools performing well on Cito tests are included for model 1 and 2. The interaction effect constitutes the combined effect of the two variables mentioned at the top of the table. A High Cito Scoring School is defined as such if the average Cito test score was above 535.
* p<0.05, ** p<0.01, *** p<0.001

Table J: Interaction Effects for Education of Father and Education of Mother

	Age of Migration – Education of the Father	Gender – Education of the Father	Age of Migration – Education of the Mother	Gender – Education of the Mother
Age of Migration	-2.563***	-2.768***	-2.743***	-2.764***
	(0.226)	(0.150)	(0.238)	(0.150)
Gender	2.513***	0.372	2.512***	0.925
	(0.539)	(1.048)	(0.539)	(1.024)
Interaction Effect	-0.0832	0.899*	-0.0256	0.808
	(0.0686)	(0.378)	(0.0811)	(0.444)
Observations	13399	13399	13399	13399
R-squared	0.773	0.773	0.773	0.773
Adjusted R-squared	0.755	0.755	0.755	0.755

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for education of the mother and father, year of cohort, and school-fixed effects are included. The interaction effect constitutes the combined effect of the two variables mentioned at the top of the table.
* p<0.05, ** p<0.01, *** p<0.001

Table K: Effects on the Language Test by year for Suriname

	(1)	(2)	(3)	(4)	(5)	(6)
	1994	1996	1998	2000	2002	2004
Age of Migration	-4.478	-4.014	-2.173	-4.626	-2.669	-2.150
	(3.286)	(2.626)	(1.832)	(2.559)	(1.487)	(2.831)
Girl	1.710	10.01	10.63	1.766	10.66**	6.428
	(11.43)	(11.72)	(8.973)	(9.823)	(3.916)	(8.809)
Observations	98	115	119	104	287	114
R-squared	0.870	0.910	0.879	0.786	0.811	0.848
Adjusted R-squared	0.618	0.689	0.735	0.585	0.757	0.727

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for social-economic status, grade, and school-fixed effects were included.
Number of observations is 837.

Table L: Effects on the Language Test by year for the Antilles

	(1)	(2)	(3)	(4)	(5)	(6)
	1994	1996	1998	2000	2002	2004
Age of Migration	11.84	-6.519	-3.027	-3.359	-2.112	-6.307
	(7.227)	(4.226)	(4.632)	(2.884)	(1.690)	(3.244)
Girl	2.072	-10.02	36.44	18.81	21.37**	-4.858
	(25.71)	(13.88)	(26.99)	(13.70)	(6.361)	(10.87)
Observations	46	64	53	70	181	105
R-squared	0.953	0.947	0.950	0.934	0.794	0.843
Adjusted R-squared	0.736	0.776	0.678	0.774	0.641	0.591

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for social-economic status, grade, and school-fixed effects were included.
Number of observations is 519.

Table M: Effects on the Language Test by year for Turkey

	(1)	(2)	(3)	(4)	(5)	(6)
	1994	1996	1998	2000	2002	2004
Age of Migration	0.137	-1.953	-1.407	-1.085	-2.377	-3.695**
	(0.983)	(1.036)	(1.145)	(0.879)	(1.226)	(1.424)
Girl	4.231	-1.859	-2.281	3.750	0.798	5.866
	(3.167)	(3.302)	(4.215)	(2.746)	(2.799)	(3.159)
Observations	273	292	307	460	560	418
R-squared	0.905	0.923	0.862	0.821	0.821	0.815
Adjusted R-squared	0.841	0.862	0.758	0.773	0.770	0.769

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for social-economic status, grade, and school-fixed effects were included.
Number of observations is 2310.

Table N: Effects on the Language Test by year for Morocco

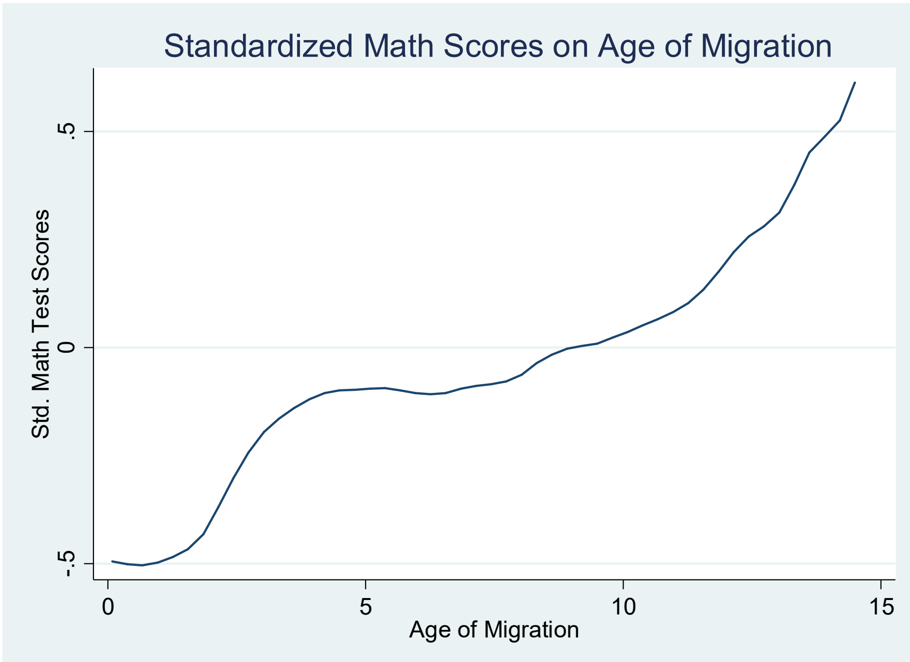
	(1) 1994	(2) 1996	(3) 1998	(4) 2000	(5) 2002	(6) 2004
Age of Migration	-2.283	-3.601**	-2.610*	-2.098	-3.176**	-3.137*
	(1.307)	(1.346)	(1.190)	(1.252)	(1.035)	(1.425)
Girl	1.025	3.723	6.009	2.554	2.561	0.891
	(4.510)	(4.108)	(4.354)	(2.875)	(2.432)	(2.730)
Observations	211	298	280	464	619	508
R-squared	0.857	0.850	0.850	0.810	0.811	0.784
Adjusted R-squared	0.756	0.757	0.753	0.764	0.771	0.740

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Source: created by the author based on the data of the PRIMA cohort studies.
Note: Standard Errors in parentheses. Controls for social-economic status, grade, and school-fixed effects were included. Number of observations is 2380.

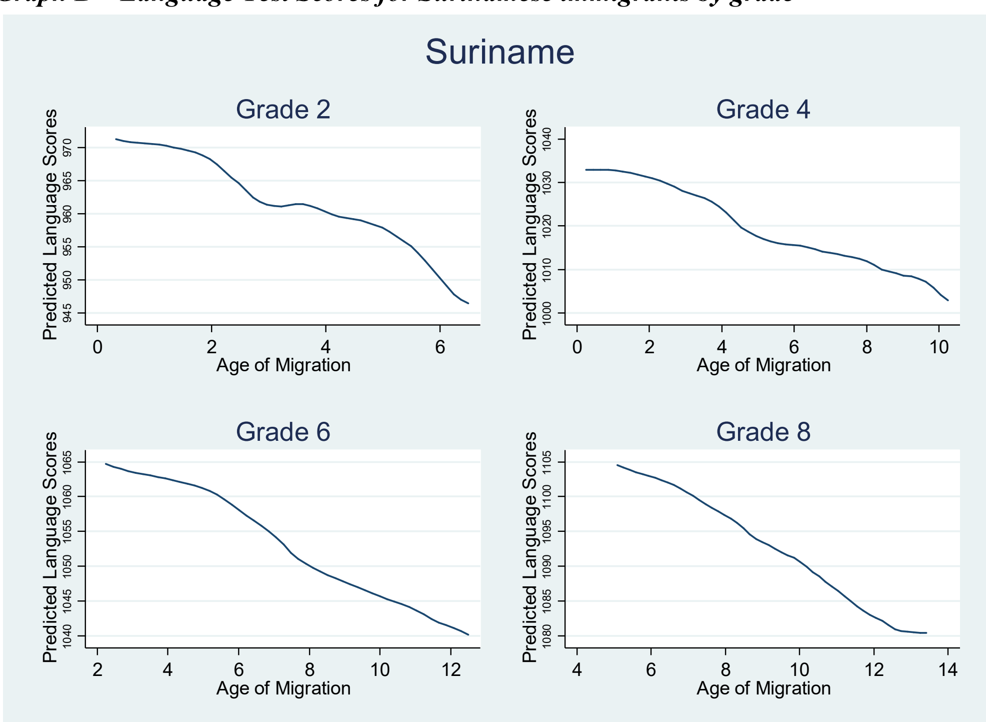
Appendix II – Graphs

Graph A – Standardized Math Test Scores on Age of Migration



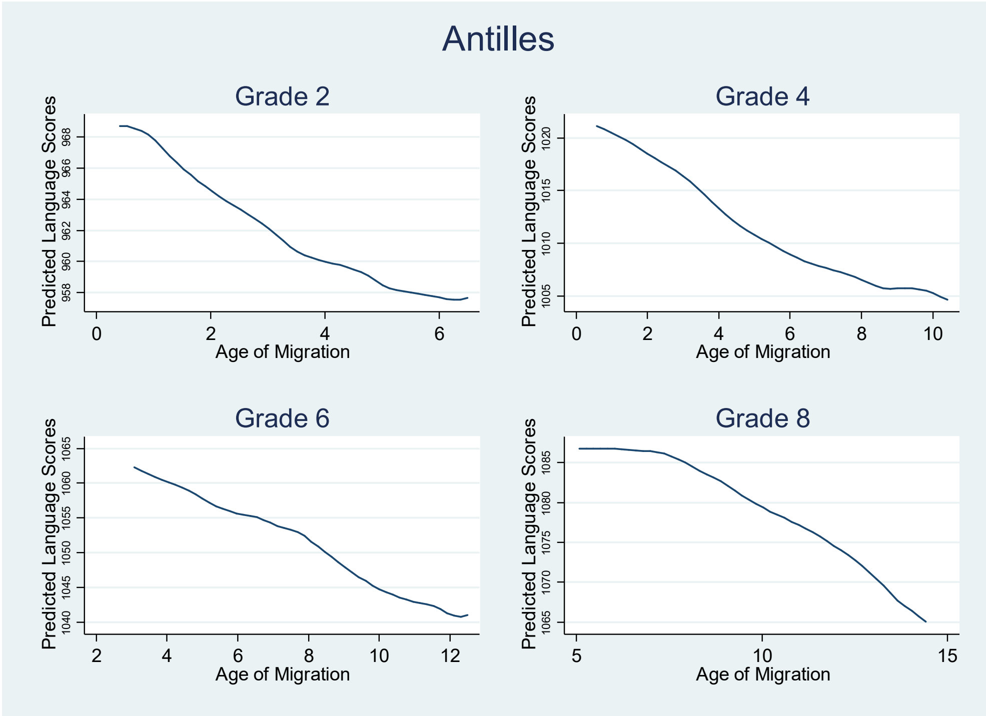
Source: created by the author based on the data of the PRIMA cohort studies.
Note: Number of observations is 14,277.

Graph B – Language Test Scores for Surinamese immigrants by grade



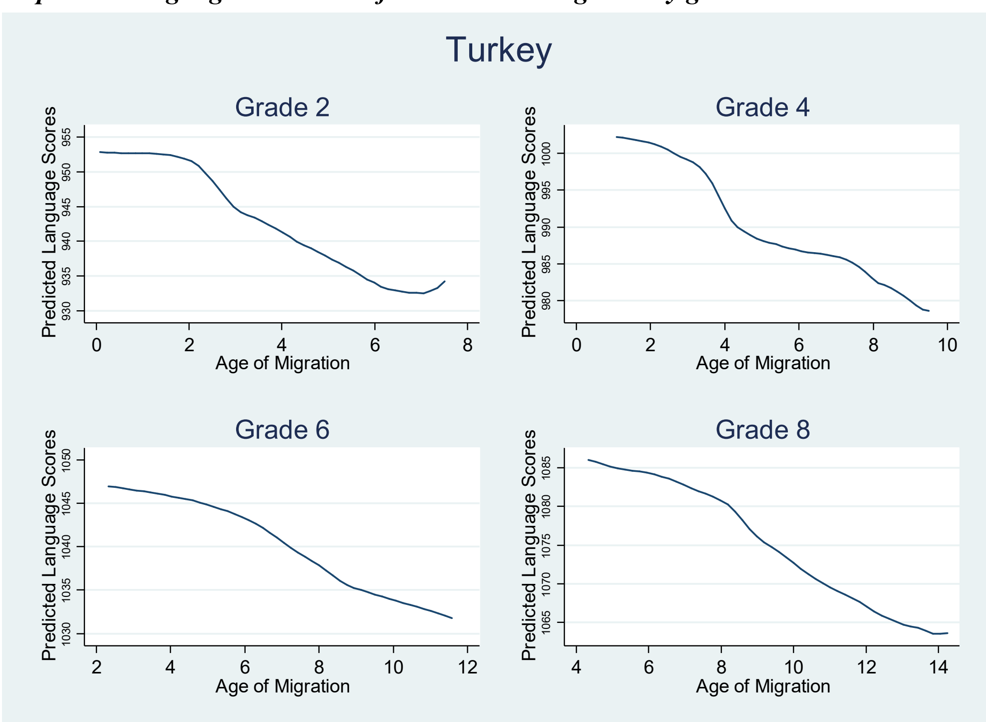
Source: created by the author based on the data of the PRIMA cohort studies.
Note: Number of observations is 837. Controls for social-economic status, grade, year of cohort and school-fixed effects were included for obtaining the predicted results.

Graph C – Language Test Scores for Antilles' immigrants by grade



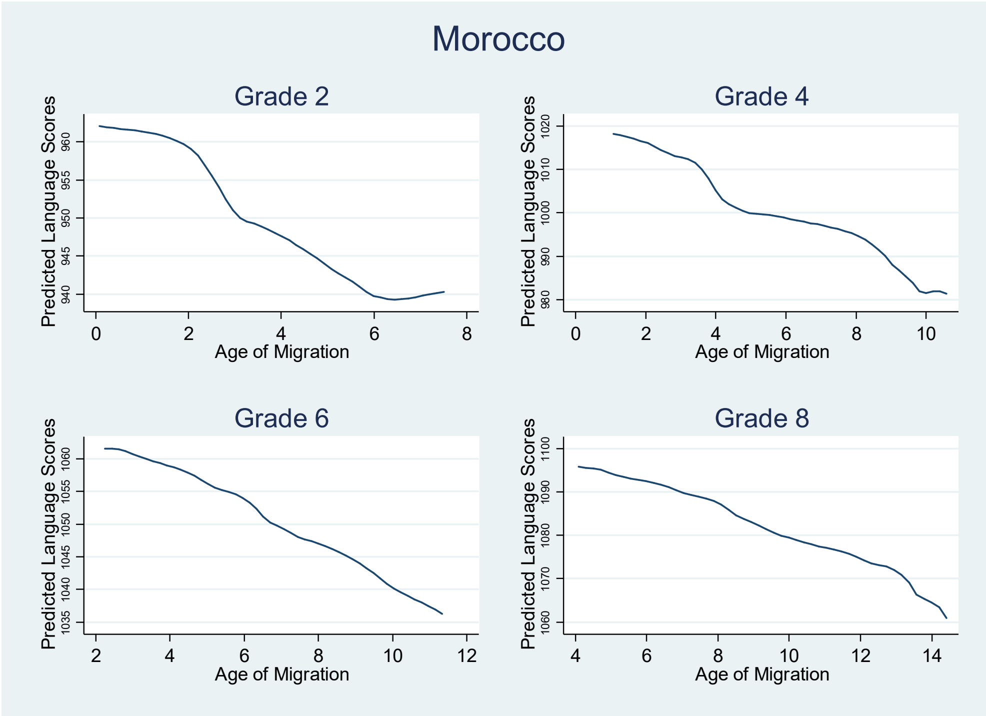
Source: created by the author based on the data of the PRIMA cohort studies.
Note: Number of observations is 519. Controls for social-economic status, grade, year of cohort and school-fixed effects were included for obtaining the predicted results.

Graph D – Language Test Scores for Turkish immigrants by grade



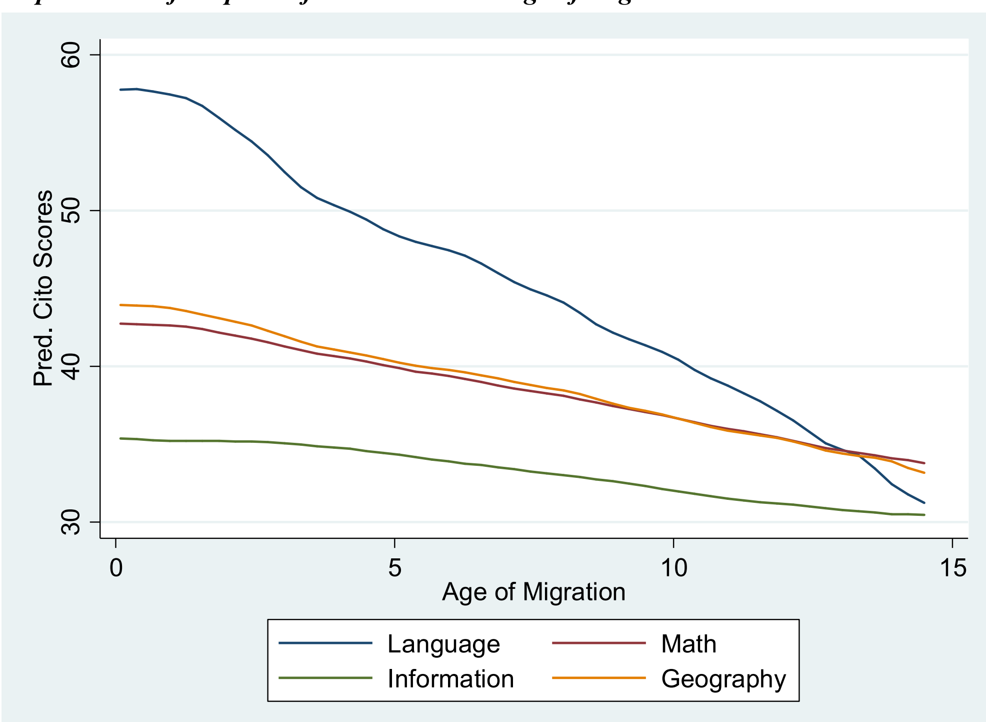
Source: created by the author based on the data of the PRIMA cohort studies.
Note: Number of observations is 2310. Controls for social-economic status, grade, year of cohort and school-fixed effects were included for obtaining the predicted results.

Graph E – Language Test Scores for Moroccan immigrants by grade



Source: created by the author based on the data of the PRIMA cohort studies.
Note: Number of observations is 2380. Controls for social-economic status, grade, year of cohort and school-fixed effects were included for obtaining the predicted results.

Graph F – The four parts of the Cito Test on Age of Migration



Source: created by the author based on the data of the PRIMA cohort studies. Controls for social-economic status, year of cohort and school-fixed effects were included for obtaining the predicted results.