To leave or not? Parental emigration decisions and human capital development of children.

Joanna Małgorzata Clifton-Sprigg

Doctor of Philosophy
The University of Edinburgh

2015

Declaration

I declare that this thesis was composed by myself and that the work contained is my own. No work by any other author has been used without due acknowledgement. This work has not been submitted for any other degree or professional qualification.

Abstract

This thesis explores empirically the links between parental migration and educational performance of children, as well as bilingualism and family status and cognitive development of children.

The first three chapters focus on a scenario in which households send one member, usually a parent, for temporary employment abroad. I firstly examine the implications of such a family structure on educational performance of teenagers. I then investigate whether this impact may spill over through peer interactions at school.

I have designed and collected a survey for the purpose of the analysis. I elaborate on the process in the first chapter. The gathered data contain information about over 2800 16-year-old pupils, including their socio-economic background, performance at school and migration experiences within a family over a period of three years. Parental migration is common in the studied population and is mostly characterised by relatively short, repeated spells of legal employment of fathers in other European countries. The nature of the migration experience sets it apart from cases considered so far.

In the second chapter I utilise the data to investigate the relationship between parental absence due to emigration and the child's performance at school in that period. I find that, on average, children's grades improve when they have a parent abroad. A negative impact may, however, emerge in cases of prolonged separation. Meanwhile, sibling emigration exerts a strong positive effect on educational attainment which accumulates over time. The results are plausible if parental emigration significantly increases household income, whilst not disproportionately burdening children by means of increased responsibility.

The third chapter extends the analysis by looking at the influence children with parents working abroad may exert on their classmates. I find that pupils in classes with a high proportion of children of migrant parents perform better. The impact is greater for those who experienced family migration themselves. I consider various possible explanations of the result and conclude that the positive individual effect found in the second chapter may spill over through the peer interactions. Increased teachers' involvement in classes with many migrant children may play an additional role.

Parental education is key to the positive effects found in both chapters. The children of parents, who have themselves graduated from high school, benefit most from their parents' emigration experience. They also are the influential group among their peers.

In the final chapter I consider a different scenario where families use two languages at home to investigate whether it affects the development of cognitive and non-cognitive skills of their children. Importantly, I notice that bilingualism may be an insufficient element to explain any differences, as bilingual families are a heterogeneous group.

Therefore, I also differentiate between families with two native, one native, one foreign and two foreign parents.

Using the data for families with children under the age of 6 in Scotland, I find that overall children's cognitive and non-cognitive skills are similar. The performance in the English Vocabulary Naming exercise is an exception. On average, bilingual children do not perform worse than monolinguals in the task. There is heterogeneity within the group, however. Bilingual mixed-nationality children lag behind the monolingual native children at the age of 3 but they catch up by the age of 5. However, there is some evidence that bilingual children who have two foreign born parents may perform worse than the monolingual native children and not improve with age.

Acknowledgements

First and foremost, I would like to thank my supervisors, Prof. Maia Güell and Prof. Mike Elsby, who have always lent me a helping hand, discussed ideas with me, asked the right questions and offered reassurance and invaluable guidance when needed.

I am also grateful for the discussions I had with Dr Steven Dieterle, who has been very enthusiastic and encouraging. He has shown interest and offered help well beyond that which was expected. I am equally grateful to my external examiner, Prof. Jackline Wahba, for her time, valuable feedback on my work and the effort put into ensuring my viva took place without any delays.

I am also obliged to Dr Tatiana Kornienko, Prof. Ed Hopkins and Prof. Michèle Belot, who continuously supported and encouraged me, particularly towards the end of this process.

I would also like to acknowledge the funding I have received from the Economic and Social Research Council to pursue my research and the Innovation Initiative Grant from the University of Edinburgh Development and Alumni, which financed the data collection project.

I am indebted to my PhD companions, with whom I shared my happiness and sorrows, and who often provided me with great day-to-day advice, keeping me sane. They let me share this otherwise lonely journey with them.

Last but not least, I would like to thank my family, who have always stood by me; my parents and my sister for their unending love and faith in me, Stef for tirelessly correcting my English and my husband for being my rock. PhD is definitely a testing ground for marriage and I am glad that we both have survived it. This adventure would not have started without the help of my aunts, Ela and Krysia.

Contents

Introduction

In this thesis I study empirically the role parental migration decisions play in development of the human capital of children. I consider various scenarios; I firstly look at the immediate impact of having a parent abroad on the educational outcomes of teenagers, whose parents work abroad. I then consider the impact those children may have on their classmates. Lastly, I study the role of growing up in a bilingual family on the development of skills in early childhood.

Human capital plays a crucial role in an individual's success in adult life. In aggregate it also contributes to countries' economic growth. For this reason economists have become increasingly interested in uncovering its determinants and evolution at different stages of life. Recent literature has also established that various skills and personality traits are developed very early in life and build upon each other throughout our growing up (Cunha and Heckman, 2007).

Crucially, the early childhood human capital is shaped by nature, as well as nurture. In particular, parental decisions regarding family life impact the child's skills through the effects they have on the upbringing. School and neighbourhood environment are additional factors determining skill development (Conti et al., 2011).

Migratory decisions within families have the potential to affect a child's human capital via two opposing channels; they often substantially alter the household composition and the growing up environment of children (separation effect) and result in an increase of household income (income effect). Parental emigration may result in lower parental inputs into upbringing, increased number of family responsibilities being put on children and the emotional burden of being away from a parent. This changed family situation may influence how children perform at school and in long term affect their labour market success (Antman, 2013).

The concern that the negative impacts may outweigh the financial gains of emigration has sparked a debate in many European countries about the consequences of families being split by migration (Tynelski, 2010). This is a particularly relevant topic for the new EU member states which observed a significant temporary migration to other EU countries following their EU accession (Eurostat, 2013).

However, the free movement of workers within the EU allows for temporary and circular migration over relatively short distances, which may relieve some of the burden of family separation. This different nature of migration may have other implications for the performance of children than what the literature has established so far.

The theory is ambiguous on the overall impact. It is also silent on the potential heterogeneous impacts of migration. Therefore, this is an empirical question. Various studies have analysed empirically the relationship between having a parent abroad and

educational attainment of children (Antman, 2013). They have focused on migratory movements which led to prolonged family separations with little or no contact with the migrant parent and where remittances could be limited or uncertain as the employment abroad was often illegal.

Moreover, the studies provide mixed results and do not always resolve key identification challenges involved. This is mostly due to the lack of suitable data. For example, in many studies the information about the exact timing of migration, as well as precise measures of educational attainment are unavailable. Parental absence during pre-school years may have different implications for a child than during teenage years. Moreover, measuring performance by school enrolment may not be adequate when analysing the phenomenon in countries where school attendance is exemplary. Thus, these elements are key for identification of the overall impact.

The lack of appropriate data is therefore a problem. This is especially true in the European context, where large scale migration is a relatively recent and difficult to record phenomenon. To investigate the relationship between the migratory behaviour of families and the school performance of children, I collected my own data; I surveyed one birth cohort of pupils in Poland, gathering information about their experiences when they were aged 14 to 16. I have designed and managed the project, secured finance, established contacts, put together, trained and led the data collection team. The data contain detailed information about the timing of migration, family background and school progress of pupils. The process and the resulting data set are described in detail in Chapter 1 of the thesis.

18% of the 2800 surveyed pupils have had a parent abroad over the three year period I capture in the data. This closely reflects the statistics on temporary migration from the 2011 Census in Poland. The migratory movement is father-dominated with a parent spending on average under 4 out of 6 observed semesters abroad, frequently returning home. Germany is the main destination country. Migration is predominantly low-skilled with over 60% of fathers having vocational qualifications.

This data is used for the analysis in the two following chapters. In Chapter 2, I consider the impact of parental foreign employment on their child's school grades. The main methodological challenge is the fact that migrants are a selected group and that migration decision may be driven by certain characteristics which also influence children's educational attainment. Unless these traits are explicitly captured in the analysis, the estimates will be confounded. Thanks to the panel nature of the data, I can isolate such characteristics from the effect, provided they do not vary over time.

Perhaps unexpectedly, I find that the immediate effect of having a parent abroad is positive. This implies that the effect of increased income is greater than the negative impact of family separation.

Importantly, this is true only for pupils whose migrant parents have a certain education level. For example, those whose parents have lower than secondary education, which amount to 67% of the entire group, neither gain nor lose from parental mi-

gration. On the other hand, pupils whose migrant parents are high school graduates benefit most. This is most likely for two reasons. Better educated migrant parents are more likely to succeed in the labour market. At the same time they may value their children's education more than parents with lower educational attainment, thus investing a higher proportion of income and their time in the child's schooling.

I also establish that prolonged migratory movements have a potential to significantly lower a pupil's performance, as the negative impact emerges after about a year from the parent's departure abroad. This suggests that even though initially the extent of the separation effect may be limited, it may become burdensome with time.

A significant proportion of respondents have also experienced sibling emigration. I undertake a similar analysis to that described above and find that having a sibling working abroad can only benefit pupils, irrespective of the length of the migratory experience. This finding highlights the different roles siblings and parents play in children's upbringing. Whilst sibling migration also increases the household budget, their absence from a household is unlikely to have as negative an effect as when the parent leaves. Moreover, siblings play a different, perhaps more influential role in children's career plans. As they are more likely to succeed in foreign labour markets, they may encourage better school performance. Despite the commonality of international migration of a young workforce, this seems to be the only analysis of this kind in literature; once again, the lack of data may be the reason why.

The findings in Chapter 2 add to the existing knowledge of the impacts of migration and highlight the heterogeneous effects. The overall influence depends on the socio-economic situation of the family, as well as the timing and duration of migration. Whom the household decides to send is also crucial.

Human capital acquisition depends also on the more broadly defined environment we grow up in. In the educational context, the classroom peers also play a role. Even if a pupil does not experience emigration within his family, his performance may be affected by the fact that his classmates' parents are working abroad. If parental emigration changes their children's attitudes towards school, behaviour and their performance, then they may in turn, through interactions, affect their classmates. Additionally, given the changes in children's behaviour, schools or teachers may change their approach, which will also have an impact on all pupils.

The existence of peer effects in schooling is well-established in economic literature (Black et al., 2013; Sacerdote, 2000). However, migration studies considered only the impact of immigrant children on their classmates, which is rather different from the scenario here.

The peer effect here depends on the individual impact parental emigration has on a child and on the interactions between children at school. Therefore, the spillover effect is ambiguous theoretically and perhaps was not studied until now due to the lack of appropriate data. In Chapter 3 I investigate the relationship between the number of classmates who have a parent abroad and one's average grade.

There are various challenges related to the estimation of such an impact. Of particular concern is potential non-random class composition where pupils are grouped on the basis of characteristics which affect their performance and, at the same time, determine the number of pupils with parents abroad in the class. I control for any differences between classes which do not vary over time by including class dummies into the regressions. My approach, however, does not address the issue of potential time-varying changes in the class environment which are correlated with pupils' performance and class composition.

I find that pupils benefit from the presence of classmates with parents abroad. Further, those whose migrant parents are high school graduates are most influential in the group. Moreover, pupils who themselves experienced family emigration benefit more than others.

Taking into account the findings in Chapter 2, I conclude that children who benefit from short-term parental migration positively influence their classmates by interacting with them. There is also an indication that the within-classroom interactions are stronger between pupils who have parents abroad and hence the effect on this group is greater.

Spillover may not be the only explanation for the findings. I consider alternative scenarios and eliminate resource re-allocations and school adjustments, as well as the grade inflation as possible options. Nonetheless, one other interpretation may be that teachers increase their efforts in response to the class composition, which also leads to grade improvements. I cannot exclude this possibility.

To summarise, the impacts of temporary migration are not limited to the migrant families, but may have further implications. Once again, the results are heterogeneous and depend on the type and intensity of interactions between pupils. The data at hand does not allow me to investigate these relationships explicitly, but this would be a logical extension to the analysis.

In the final chapter of this thesis I analyse how the family environment and use of two languages at home affect cognitive and non-cognitive skills of small children. Raising a bilingual child may be seen as an investment in the child's skill. This is because language is among the first skills an individual acquires and is essential for further acquisition of human capital. Early life language acquisition is often linked with a better performance at later stages of life (Kamhöfer, 2014).

The effect of bilingualism on children's outcomes is an empirical question as there are many mechanisms at play, often acting in opposite directions. Moreover, the empirical challenges behind the question are significant; I hope to address them thanks to the use of a very rich data set produced by the Scottish Government - Growing Up in Scotland data.

Given the importance of language for skill-development, raising bilingual children may be seen as an investment parents make in their human capital. By exposing children to two different languages early on, parents may increase their productive skills

and enable them to learn more efficiently in the future. Hence, bilingual children may have an educational advantage over their peers. However, bilingual children usually have at least one foreign-born parent and this may be to their disadvantage. Moreover, raising bilingual children is a high effort task and not everyone may succeed in doing so.

Bilingual families are a heterogeneous group. Families with two foreign-born and one native and one foreign parent are often bilingual. Yet, they may differ substantially from each other socio-economically as well as in terms of lifestyle.

Parents in mixed-nationality families are frequently positively selected (Lanzieri, 2012) and have the potential of creating a favourable environment for their children, encouraging academic success. Families with two foreign-born parents, on the other hand, may be at a disadvantage in terms of their linguistic skills, local cultural and institutional knowledge crucial for enhancing children's development.

These differences may define the ways in which they bring up their children and whether they succeed in teaching children two languages. Therefore, any potential gaps in children's performance may be driven by bilingualism as well as other family factors, which are more difficult to capture.

There is a consensus in current linguistic literature that bilingualism benefits children, even though it may be advantageous in some areas of cognition and harmful in others (?). The linguistic studies are usually based on experimental data and use very precise measures of development. However, arguably the subject groups of the studies are often selected since participation is voluntary.

I use the rich, representative Growing Up in Scotland data set compiled by the Scottish Government, which contains comprehensive information about family background of respondents, as well as objective measures of performance. These features of the data allow me to overcome the concerns in linguistic literature of participants' selection into the study.

I use both the fact of using two languages at home and the family status (as a proxy for the unobservable home environment) to establish whether they influence children's skills. I find that families with two native, one native and one foreign and two foreign parents differ socio-economically. As expected, mixed-nationality families are positively selected relative to the native families and the families with two foreign parents are worse off socio-economically. They also differ in lifestyle, defined by activities they engage in, and views related to children's upbringing.

All children perform comparably in most fields, with an exception of the English Vocabulary Naming exercise. On average, bilingual children do not perform worse than monolinguals in the task. There is heterogeneity within the group, however. Bilingual mixed-nationality children lag behind the monolingual native children at the age of 3 but they catch up by the time they are 5 years old. Further, monolingual mixed-nationality children perform better than monolingual native children. However, there is some evidence that bilingual children who have two foreign born parents may perform

worse than the monolingual native children and not improve with age; the effect is sizeable but insignificant, which may be due to the small number of observations in the data.

I analyse the contribution of various activities to the English Vocabulary Naming score. I find that some, but not all factors, including practising letters and visits to the zoo, have a higher payoff for bilingual and mixed-nationality families. This is expected if worse performance is related to linguistic skills as these activities facilitate language acquisition.

This is, to the best of my knowledge, the first analysis considering the impact of both language and family composition on the outcomes of very young children. I provide an initial insight into how bilingual and bicultural upbringing may shape children's skills. Importantly, I look at very young children at the language-learning stage, when any potential differences are likely to be visible. Moreover, the results highlight the importance of distinguishing between children who have one and two foreign-born parents.

Bibliography

ANTMAN, F. (2013): "The Impact of Migration on Family Left Behind," in *International Handbook on the Economics of Migration*, ed. by A. F. Constant, and K. F. Zimmermann, pp. 293–309. Edward Elgar, Northampton, MA.

BLACK, S., ET AL. (2013): "Under Pressure? The Effect of Peers on Outcomes of Young Adults," *Journal of Labor Economics*, 31, 119–153.

CONTI, G., ET AL. (2011): "Early Endowments, Education and Health," Human Capital and Economic Opportunity: A Global Working Group, Working Paper No. 2011-001.

CUNHA, F., AND J. HECKMAN (2007): "The Technology of Skill Formation," *The American Economic Review*, 97(2), 31–47.

EUROSTAT (2013): "International Migration and Asylum, Immigration," 10.09.2013, http://epp.eurostat.ec.europa.eu.

KAMHÖFER, D. (2014): "The Effect of Early Childhood Learning Programs on the Contemporary Formation of Grammar Skills," mimeo.

LANZIERI, G. (2012): "Mixed marriages in Europe, 1990-2010," in *Cross-Border Marriage: Global Trends and Diversity*, ed. by D.-S. Kim, pp. 81–121. Korea Institute for Health and Social Affairs (KIHASA).

SACERDOTE, B. (2000): "Peer effects with random assignment: results for Dartmouth roommates," Discussion paper, NBER Working Paper 7469.

TYNELSKI, A. (2010): "Sytuacja dziecka w rodzinie migrującej w Polsce," Discussion paper, Ministerstwo Edukacji Narodowej.

Chapter 1

Migration and Education of Children in Poland 2012 (MECP2012)

Acknowledgements

The project discussed in this chapter would not have happened without the help and support of Maciej Gruszecki at the Careers Office of the Opole University of Technology, who helped me with recruitment, as well as Agnieszka Kaczor and Karolina Janik, who were involved in establishing contact with local schools and coordinated the work of data collectors. I would also like to thank Anna Szatka, Ksenia Kornacka, Sabina Waluś, Malwina Wyrwich, Anna Maria Lisowska, Anna Kisiel, Henryk Kretek, Sabina Gordzielik, Natalia Lechowicz, Paulina Herba, Iwona Janczyszyn, Martyna Grzyb, Anna Hury and Damian Kolarczyk for their involvement in surveying and data imputation. I will also always owe my sister, Barbara Jankowska, for last minute trouble-shooting in times of despair. I would like to gratefully acknowledge the receipt of the Innovation Initiative Grant from the Development and Alumni of the University of Edinburgh which financed the project.

Brief motivation

Analysis of individual and spillover impacts of parental employment abroad, which I undertake in the following two chapters of this thesis, requires data containing information about emigration experience within households as well as school enrollment and school performance of children. To the best of my knowledge, no such data exist for any European countries and, perhaps for that reason, most economic migration research focused on traditional sending countries, like Mexico. The ability to make any predictions from previous research about the impact of parental temporary migration within Europe on children's educational performance is very limited, given the unique nature of the recent intra-European migratory movements. Therefore, Europe-specific analyses could help clarify whether the effects are negative, as commonly perceived.

To investigate the relationship between parental emigration and a child's performance, as well as its influence on classmates, I have designed a survey and collected new data during field work in Poland. It contains information on the individuals' class and school enrollment, academic performance, as well as some

family background information, including migration history. The data has a panel structure and covers 6 semesters of schooling.

This chapter provides an insight into the data collection process and sets the background for the subsequent studies. I begin by describing the education system in Poland and drawing a profile of the region where data come from. Their characteristics motivated the choices made at the outset of the project. I then explain how the information was collected and briefly describe the resulting data set.

1.1 Why Poland?

Poland is the largest of the EU member states which joined the organisation since 2004. It has also become the largest (in absolute terms) sending area. It is estimated that over 1.2 million Poles (3.1% of the population) left the country for temporary employment abroad between 2002 and 2011 Census. As much again decided to live abroad permanently, leaving Poland between 2003 and 2012 (The Central Statistical Office of Poland, 2012). Although not the largest outflow relative to population size, the pace of the migration movement is overwhelming and is changing the socio-economic reality in Poland. Temporary emigration, which I am interested in, has resulted in a phenomenon of leaving families behind by many Poles.

Choosing the largest migrant-sender among the EU member states which joined since 2004 increased the likelihood of capturing sufficiently large number of children with parents working abroad. At the same time, however, Poland is comparable in many respects to most of the remaining Member States which joined the European Union since 2004 - they share similar political history and have transitioned into democratic market economies in a short period of time.

1.2 Choice of region and age group

1.2.1 Education system in Poland

The education system in Poland is divided into three mandatory stages: *szkoła podstawowa* (children aged 7-12), *gimnazjum* (age 13-15) and *szkoła średnia* (age 16-18/19). During the first two stages pupils follow a common national curriculum and write a competence test at the end of each stage. Tracking begins at the age of 16 when pupils apply to institutions with different educational goals. One is obliged by law to remain in full-time education until the age of 18.

The data was collected among pupils aged 15 or 16, in their final year of *gimnazjum*, and records retrospectively their performance over a 3-year-period, i.e. the duration of lower secondary education. Hence, one can follow each pupil

throughout the 6 semesters he spent at *gimnazjum*.

I decided to target this particular age group as the comparisons of performance are still reliable at this stage of education, thanks to the common national curriculum. The choice also allowed for a collection of the biggest amount of information on educational attainment of pupils (6 five month long semesters worth of average grades, behavioural assessment and attendance). Moreover, the middle stage of education might be the most crucial in terms of impact of migration in the family on educational performance. *Gimnazjum* pupils have been pointed out by the recent Polish policy-makers' report as the most affected by the migratory outflows from the country (Tynelski, 2010). At this age, teenagers still rely on their parents, particularly when making career choices. Therefore, the consequences of family separation may be most visible in this age group. At the same time they are independent and sufficiently informed to successfully participate in the study. Thus, they seemed a suitable study population, given the topic I aimed to explore.

During *gimnazjum* most pupils are enrolled with their local school[1] and have limited opportunity to influence their class allocation. Some schools allow pupils to name their preferred classmates but the request is not always granted and there is no scope for a coordinated action of parents to create a favoured class. Nonetheless, allocation is not random.

Once created, the group does not change throughout the three years.[2] All classes are carried out in the same unit and pupils mostly interact at the class level. A degree of mixing takes place within the school but it has a more social character. Once allocated a group, the subject teacher also does not change, except in cases of retirement, maternity leave or illness, to ensure consistent assessment and education of pupils.

At the end of their education in *gimnazjum* pupils sit a national competence test in major subjects and are accepted to further education on the basis of their results in the national tests and the grades awarded by their schools.

1.2.2 Study area - Opolskie, Poland

Geography and economy

Opolskie voivodship is the smallest of 16 Polish voivodships and is located in southern Poland, along the border with Czech Republic, as well as in close proximity to Germany, with a population reaching just over 1 million inhabitants.

[1]Schools are obliged to admit all pupils from the catchment area and are allowed to consider applications from outside the catchment area if they have spare capacity. As a result they are often highly selective towards applicants from outside the local areas.

[2]Exceptions are cases when a pupil needs to repeat a year, moves away from the area, etc. Later I discuss the frequency of such cases and threats they pose to validity of the results.

According to the National Statistical Office of Poland, the registered unemployment rate in the area in 2012 was 14.4% (compared with 13.4% for Poland as a whole) and the region contributed 2.1% to the Polish GDP with a GDP per capita in Opolskie equal to 80.1% of the Polish GDP per capita (The Central Statistical Office of Poland, 2013b).

Opolskie's complicated past

Opolskie has been historically the highest out-migration region of Poland. The reasons behind the significant outflow of population from Opolskie are numerous and include among others historical, ethnic, cultural, political and economic motives (Jończy, 2007). The migration tradition in the region dates back to the 19th century.

Before the Second World War, the current territory of the Opolskie voivodeship constituted part of Germany; but being originally Polish, it was significantly inhabited by Polish citizens. As a result, the regional language and culture have been heavily influenced by the mixing of German and Polish nationals.

In 1945, however, the border of Poland was renegotiated and Opolskie voivodeship, along with other regions, was included again to Polish territory. It was then that the government of newly independent Poland extradited most of the population of German origin. The displacement of German nationals from the region coincided with an inflow of Polish citizens from other regions of Poland, particularly eastern, where the borders of the country had also shifted. In Opolskie voivodeship remained a significant group[3] of autochthonous or indigenous population, who were rendered Polish, despite their connections to Germany (Madajczyk, 1996).

This movement contributed to the first significant wave of migratory flows within the region. As a result of these historical influences, the autochthonous population of the Upper Silesia and Śląsk Opolski, which is a part of the voivodship, has developed a new identity and showed visible signs of ethnic and cultural separation. In fact, nowadays, the group advocates for minority status in Poland.[4] Many autochthons declare the feeling of belonging to Poland, but also having German origins.

At the same time, the Poles who newly arrived into the region after the Second World War have found it difficult to fully accept the distinct culture of autochthonous population. The socialist government in Poland also played its role by exercising a clearly anti-German politics and attempting to "polonise" the autochthonous population.

[3]Equal to 436900 people in 1950 (Jończy, 2007, p.84).

[4]According to the results of Census 2011, as many as 809 000 Poles defined themselves as belonging to the Silesian minority (The Central Statistical Office of Poland, 2013c).

This unique identity felt by many inhabitants of the voivodeship constituted an easing factor in migration decision-making and has been further encouraged by the German authorities, which allowed Polish autochthonous citizens of the region to apply for the German citizenship, provided they fulfilled certain criteria.[5]

Until 1990s, when the scale of immigration became too large, the German government was relatively welcoming to the Polish migrants, offering various means of support upon migration, particularly at times when the country needed an inflow into the labour force.

Hence, possession of a dual nationality (Polish-German) became a great opportunity for many inhabitants of the region, as it meant they could work and live freely in Germany and other EU countries before Poland accessed the European Union in 2004. Legal employment opportunities abroad were much more limited to other Poles, who often required a visa to live outside Poland and, therefore, saw less benefits to migration.

The scale of the phenomenon was non-negligible. Jończy (2007) discusses the changes in the composition of population of Opolskie over time. He points out a decline in the share of autochthonous population registered for permanent residence in the region from 436900 in 1950 to 350000 in 2000, which is largely due to migration. He also provides detailed information about the distribution of the autochthonous population in the region, with the visibly highest concentration in the eastern part of the voivodship. Between 50 and 90 percent of the population is native in these areas. Moreover, these areas are also defined by lower birth rate and a different demographic structure from the rest of the region, which result from significant migratory movements.

Jończy's arguments coincide to an extent with the statements of nationality by the inhabitants of the Opolskie region as reported during Census 2002. Unfortunately, no disaggregation of nationality to other than *not-Polish* is provided in the Census data, disabling a full verification. However, it can be deduced that to a large extent this will be German or Silesian nationality (minority) that has been declared by locals.

Among the inhabitants of Opolskie, the autochthonous population is most likely to qualify for the dual citizenship, as they would have lived in the area for generations and hence may be able to fulfill the residence criteria for German citizeship. Establishing how many of the autochthonous individuals have

[5]The criteria for German citizenship: If one's parents do not possess German nationality, he has to demonstrate German origins of his ancestors, as well as the fact that in years 1913-1945, i.e. before the restoration of Polish borders to their current position, the ancestor lived in the territory of the Third Reich or the Free City of Gdańsk. Fulfilment of the latter condition is highly likely for the native population of the region, given that it constituted a territory of the Third Reich before 1945. Should the family not had lived in these territories, one needs to demonstrate that his father's or grandfather's name was added to the German Volkslist during the Second World War. More information can be found on the website of the German Embassy in Warsaw: http://www.warschau.diplo.de

exercised the right to dual nationality is difficult, although it is expected to be approx. 200 000 citizens (Jończy, 2007), which would constitute more than a half of the autochthonous population and just over a fifth of overall population of the voivodeship. These estimates are likely to be inaccurate and lower than the actual number, because the acquisition of German nationality by Polish citizens has been for years discouraged by the government.

In his analysis Jończy (2007) argues that as much as 16.68% of surveyed by him autochthonous population worked only and permanently in other EU countries and only 42.5 % exclusively in Poland.

The major driver of the emigration decision, given the favourable conditions for migration from the region to Germany, were the differences between the living standards in source and destination country, as well as connections to Germany. The already established migrant networks may have also acted as an encouraging factor. Many households decided to send only certain family members to work abroad, choosing Poland as their primary place of residence. This may be due to their feeling of belonging in the region, lower living costs, etc.

New migration

Following Poland's accession to the European Union in 2004, immediate opening of labour markets by many EU member states, followed in 2011 by Germany, dual nationality became less pivotal in migration decision-making. Nonetheless, the strong connections with Germany meant that many temporary migrants chose the country as their destination.

According to the Polish Census there were 107 985 residents of Opolskie residing *temporarily abroad for at least 3 consecutive months* in 2011. Of them 94.5% emigrated to other EU countries, almost 62% to Germany (The Central Statistical Office of Poland, 2013a). It is clear from Figure 1.1 that Opolskie had the highest proportion of temporary emigrants per 1000 inhabitants in the entire country in 2011; this is also true of the region in the past.[6] Resultantly, 17.8% of all households in the region had *at least one emigrant* at the time of the 2011 Census.

The number of permanent emigrants from Opolskie is also significant and larger in relative terms than for the remainder of the country (See Figure 1.2). It is, however, less relevant for the discussion here, as permanent emigrants are most likely to uproot the entire family. I am predominantly interested in families split by migration, which usually have a temporary emigrant abroad.

[6]Note, however, that the gap in migration outflow between Opolskie and other regions of Poland has been closing following the entry of Poland to the European Union. Specifically, the migration levels remained relatively constant in Opolskie but other parts of Poland have experienced a migration shock following accession (The Central Statistical Office of Poland, 2013a).

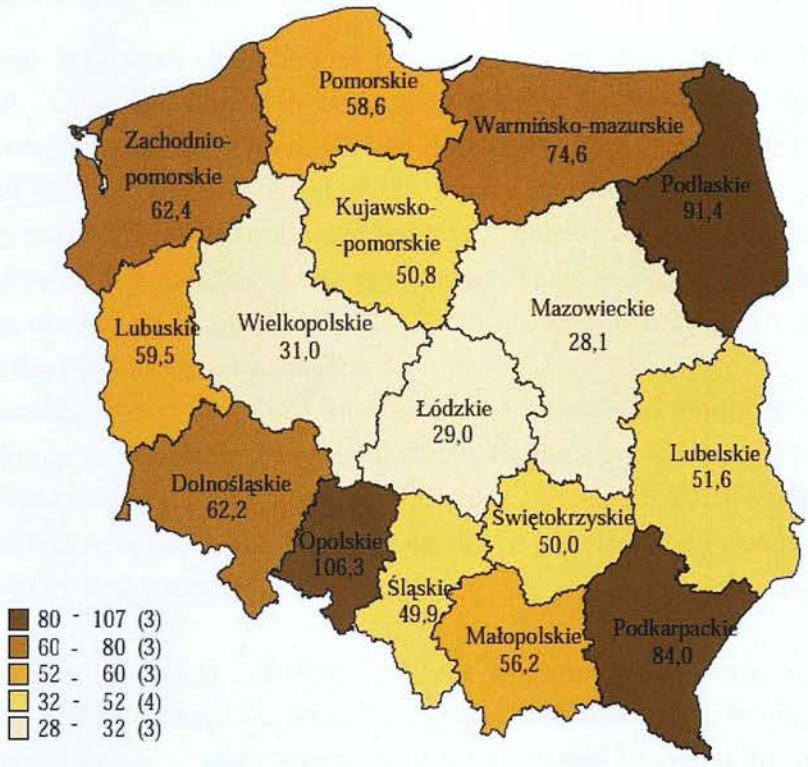


Figure 1.1: Number of emigrants from Polish regions per 1000 inhabitants in 2011, Source: The Central Statistical Office of Poland (2013a)

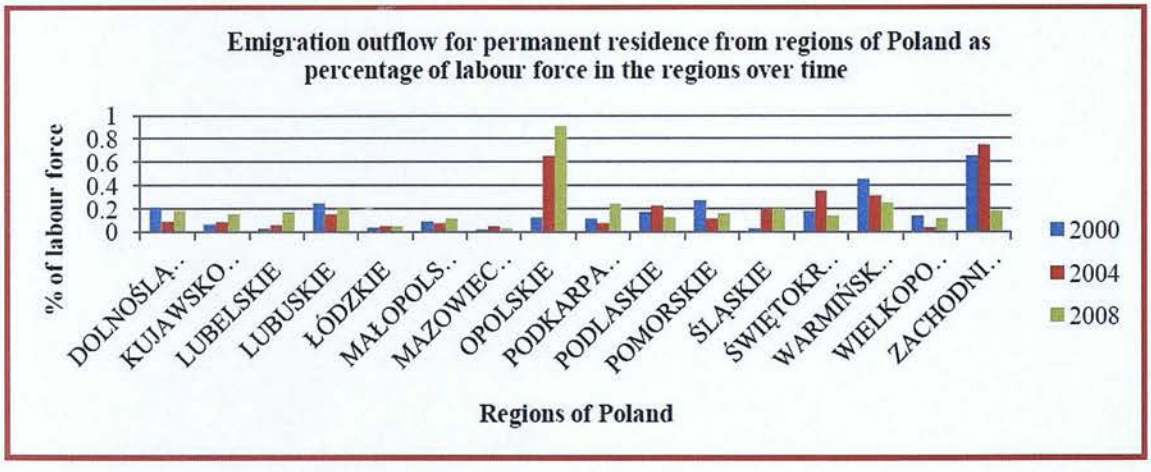


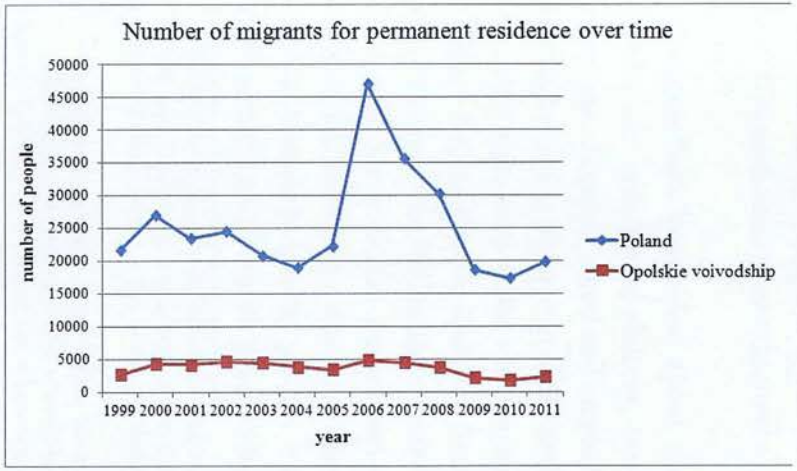
Figure 1.2: Outflow for permanent residence over time, Source: Central Statistical Office of Poland

Migration motives and impact on the region

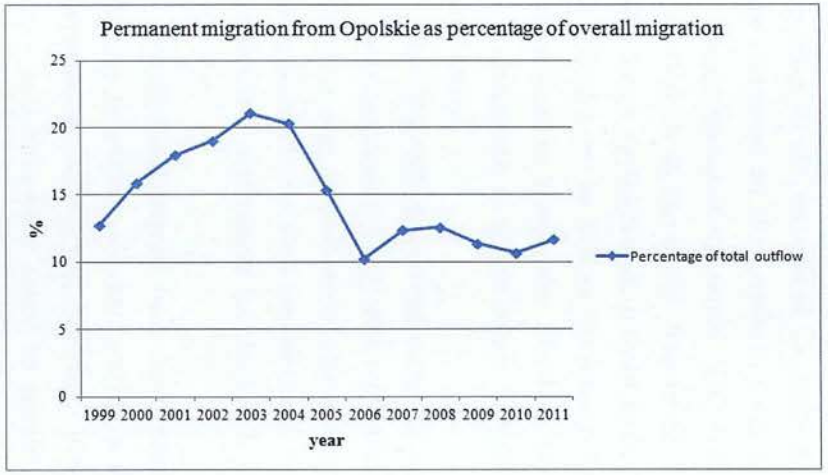
The 2011 Census estimates that 73% of temporary migrants have left Poland to work abroad. Of those, almost a third were seeking better wages and 31% could not find employment in Poland prior to departure (The Central Statistical Office of Poland, 2013a). In their recent study Jończy and Rokita-Poskart (2013) estimate that in 2010 12% of the total population of Opolskie were working abroad and on average spent 3.9 months of the year away. They earned approximately PLN 5.9 billion abroad and remitted PLN 4.2 billion, of which PLN 3.7 billion was spent and the rest allocated in banks.

Recently there has been a revived interest in the number of families divided by migration living in Opolskie (Tynelski, 2010; Regionalny Ośrodek Polityki Społecznej w Opolu, 2011). The latest publications (Tynelski, 2010) name the region as one of the areas with the highest number of pupils having one or both parents temporarily living abroad.

Focus on the area increased the likelihood of the migrant group in the sample being sufficiently large to obtain statistically significant results in regressions. Arguably, such a decision may have compromised the potential to generalise results of any analyses using the data. I consider it in detail in Section 1.5 of this chapter.



(a) Emigration for permanent residence from Poland and Opolskie over time, Source: Central Statistical Office of Poland



(b) Emigration from Opolskie as percentage of the overall outflow, Source: Central Statistical Office of Poland

Figure 1.3: Migratory outflow from Opolskie in absolute terms and as percentage of the overall population outflow from Poland

1.3 How the data was collected

There are 140 junior high schools (*gimnazjum*) for pupils aged 13-16 in Opolskie,[7] according to the records of the local Education Board. In the 2010/2011 school year they educated 30 605 pupils in total and approximately 9 500 16-year-olds, who are the target group of the study. The 114 largest schools[8] were contacted with a request to participate in the study. Among the contacted schools, 55 agreed to participate and 59 declined participation, mostly indicating timing of the project (close to the end of the school year) as well as the sensitivity of the issue to be investigated as a reason for their refusal to cooperate. Further, three schools initially committed to the project, have withdrawn at a late stage. Overall, 52 schools, educating 3423 final year pupils, participated.

The data collection required an introduction of a short questionnaire to students and school management along with collection of a time series of data on school performance, behaviour and attendance of the respondents. It can be found in Appendix 1.A. Data was collected in the first two weeks of June 2012, to ensure access to information about respondents' attainment for the last 6 semesters of schooling (September 2009 to June 2012).[9]

School management have been informed of the aim of the data collection when agreeing to participate and setting a suitable date for the survey to be conducted. Respondents themselves, however, were unaware of the project until the day of the survey and have been asked to answer the questions on the spot, which lowered the likelihood of them opting out of the process by not coming to school on the day of the survey. Research aims were explained to the respondents on the day to ensure informed consent.[10]

Students have been asked about their age, gender, nationality, as well as family situation, i.e. number of siblings, birth order, age of siblings, who they live with, parents' age, education level and employment status. They have also been asked about participation in any extra-curricular activities, plans to attend university and emigrate. Lastly, they have been asked whether any member of their close family (mother, father or sibling) has emigrated. Children from emigrant families were then asked additional questions about the destination country, period of absence of the parent, frequency of contact with the emigrant parent and whether they have experienced an increase in household responsibilities due to emigration. The respondents have not been asked about the household income as they may be unaware of the exact financial situation in their families and because it would have caused a controversy, potentially leading to

[7] After exclusion of schools for adults and for children with disabilities.

[8] It was not feasible to reach out to the smallest schools in the region due to financial and time constraints of the project. A list of the contacted schools and their responses can be found in Appendix 1.A.

[9] The school year finished on 29th June 2012 and the final grades of pupils for period from September 2011 to June 2012 were approved by schools by 27th June 2012.

[10] The project has also passed the ethics approval.

less schools participating in the project. Thus the only indication of the family social status can be drawn from the information about parental employment and education level.

Schools provided information regarding an average grade, behavioural grade and attendance of respondents. The average grade is an average over all courses taken in a given semester and ranges from 1 to 6, with 6 being a top mark awarded to a pupil for extracurricular achievements in the subject area. Pupils who mastered 100% of the curriculum in a given semester are usually awarded 5; 1 is a fail mark. Every semester schools also issue a grade for the overall behaviour of each pupil. The grade ranges from 1 to 6, with 6 being for extraordinarily good behaviour, including involvement in charity work, etc. Attendance is also recorded with the number of missed hours in a semester and the number of hours missed, but excused due to illness or any other justified reasons.

Some educational institutions also released data on respondents' performance in the national tests in Polish language, maths, history, sciences and foreign languages. Overall I have such data for 12.83% of the sample. These tests were taken by the students in their last semester in *gimnazjum*, organised nation-wide by one Exam Board and graded in percentage terms.

Apart from releasing information about respondents, schools also shared their impressions of the migration problem within families and its impact on pupils. The management of schools indicating existence of large migration in their community, have declared observable problems with behaviour, motivation and school attendance of pupils whose parents emigrated.

Overall data for 2822 students was collected. The questionnaire responses have been matched with the information provided by the school regarding respondents' average grades and school attendance each semester over the observed period. Occasionally, schools were unable to provide a full set of data for all 6 semesters, in which case the information about respondent's educational progress is limited. The full list of variables can be found in Appendix 1.A.

1.4 Data overview

There are 2822 respondents in the data, observed over a period of six semesters, between September 2009 and June 2012.[11] All of them provided information about migration experience within their family but only 2669 gave a detailed account of its timing over the 3 year period and were included in the analysis.

[11]This is true only in cases where complete information was provided in the survey and the school released a full history of academic performance. In some cases less than six semesters of data are available.

1.4.1 How common is migration?

Pupils are identified as children with parents working abroad (PWA children) if they had at least one parent abroad during the period between September 2009 and June 2012. The migration status can be identified from two variables in the questionnaire: about family having experienced migration in the 3-year-period and the exact timing of migration (see Appendix 1.A). Pupils were asked whether parents have been abroad at any time during the observed period and, based on this information, PWA children constitute almost 18% of the sample (see Table 4.3). This number closely reflects the 2011 Census statistic of 17.8% of the households in Opolskie having at least one emigrant. Respondents have also provided details regarding the timing of parental emigration, which was subsequently used to build a time-varying migration variable and to create a fraction of PWA pupils in the class at a given time t. The number of PWA children in the sample at given time t is lower than the overall measure over the entire period.

The migratory movement is father dominated and in only 64 cases a respondent indicated having both parents abroad. Moreover, only 40 respondents stated that both their parents were away at the same time. Additionally, over 40% of children from migrant families have experienced sibling migration.

Table 1.1: Emigration situation

Panel A: Summary (irrespective of the exact timing)			
	Absolute value	Percent	Percent
total sample (n)	2669	100	
migrants (incl. sibling)	685	25.67	
migrant parents - total	479	17.95	100
Who emigrated:			
only father	315	11.80	65.76
only mother	100	3.75	20.88
mother and father emigrated	64	2.40	13.36
sibling	333		
Panel B: Average duration of parental emigration (time spent abroad during the observed 6-semester-period)			
Father's emigration			4.40 semesters
Mother's emigration			1.29 semesters
Panel C: Frequency of meetings with the emigrant parent (Note: not all respondents provided this information)			
		N	
		mother	father
every month		119	254
every 6 months		28	63
every year		6	6
more rarely		2	27
total		155	350

Source: MECP2012

Table 1.2: Destinations of migrant parents

country	mothers		fathers	
	N	%	N	%
Germany	100	64.52	222	64.35
the Netherlands	38	24.52	63	18.26
the UK	7	4.52	16	4.64
Austria	1	0	8	2.30
Ireland	2	1.29	7	2.03
other destinations	7	4.52	29	8.41
total	155		345	

Source: MECP2012

Table 1.3: Patterns of emigration in the sample

	fathers	mothers	either parent
away entire time	123	16	139
left	57	14	71
returned	142	56	198
cyclical migration	137	67	204
total	459	153	612

Source: MECP2012

The main receiving country in the sample is Germany,[12] followed by the Netherlands and the UK, which points to the fact that emigration occurs over relatively short distances with the possibility of frequent returns.

Not only do families tend to send one member at a time for emigration, but also common patterns of the movement emerge within the sample. Parental migratory movements can be grouped into four main patterns (see Table 1.3).[13] There are parents who have been absent for at least 6 semesters, those who returned from or left for emigration during the period for which I have data. Lastly, there is a significant group of migrants who experience short, repetitive spells of emigration.

Overall, migration observed in the sample is characterised by rather short-term, circular movements, with respondents having frequent contact with the migrant parent. These features distinguish the new European migration spells from those most commonly analysed in research of cross-border families[14] and I expect them to have a bearing on the findings in my research. The patterns observed in the data are in line with the 2011 Census output and the literature on Polish emigration (The Central Statistical Office of Poland, 2013c; Kaczmarczyk and Okólski, 2008).

[12] Almost 65% of migrant mothers and 64% of migrant fathers left for Germany; these statistics closely reflect reports from the 2011 Census that 62% of temporary emigrants from Opolskie lived in Germany in 2011 (The Central Statistical Office of Poland, 2013c); see Table 1.2 for a summary of destinations of migrants.

[13] Note that the total number of observations differs between Tables 4.3 and 1.3. This is because I am unable to define patterns of all migrants reported in the sample.

[14] Studies of migration from traditional sending countries like Mexico or the Philippines highlight the fact that children are often left with distant family members for prolonged periods of time with little contact with the migrant parents (McKenzie and Rapoport, 2011; Cortes, forthcoming). This is not the case in my data.

1.4.2 Who are the migrant families?

Migrant and non-migrant families differ in terms of socio-demographic characteristics. Children from migrant families have on average more siblings and tend to be the younger ones in the family (birth order of 2.3 versus 1.8). A lower percentage of mothers in emigrant families work compared to those in non-migrant families. Importantly, a higher proportion of fathers and mothers in emigrant families have only vocational or secondary education and a lower proportion have completed tertiary education relative to non-migrant families–an outcome indicative of low-skilled emigration. Performance of children also differs across the two groups. Children from migrant families obtain on average .16 lower average grade than children from non-migrant families. They have also lower behavioural grades and miss on average more school hours than their peers from non-migrant families (see Table 1.4). These observations coincide with the feedback from schools indicating problems with behaviour, motivation and school attendance of pupils who have parents abroad.

Figure 3.1 portrays the distribution of respondents' grades focusing on the migrant-non-migrant divide. The average grade distribution of children with an emigrant parent is shifted to the left relative to the grade distribution for non-migrant children. This indicates an overall worse performance of children from migrant families.

Table 1.4: Characteristics of children and households

	Migrant				Non-migrant			
	mean	st.dev.	min	max	mean	st.dev.	min	max
number of siblings	1.74	1.17	0	4	1.62	1.12	0	4
mother's age	40.40	5.29	29.50	59	41.36	5.57	28.50	62
father's age	43.19	5.81	30.50	62	43.90	5.74	28.50	69
child participates in after-school activities	0.55	0.50	0	1	0.55	0.50	0	1
child's average grade	3.49	0.83	1.40	5.50	3.65	0.85	1	5.88
child's behavioural grade	4.31	1.32			4.56	1.20		
child's no of hours missed	15.90	32.35			10.37	27.41		
Mother's education	N		% of group		N		% of group	
primary	16		5.71		229		10.02	
vocational	122		43.57		793		34.70	
secondary	101		36.07		769		33.65	
tertiary	41		14.64		494		21.62	
Mother works	189		69.23		1,542		72.36	
Father's education	N		% of group		N		% of group	
primary	11		4.10		213		9.70	
vocational	168		62.69		1031		46.27	
secondary	78		29.10		644		29.31	
tertiary	11		4.10		309		14.06	
Father works	241		91.98		1855		90.53	
% respondents female	57.82				50.90			

Source: MECP2012

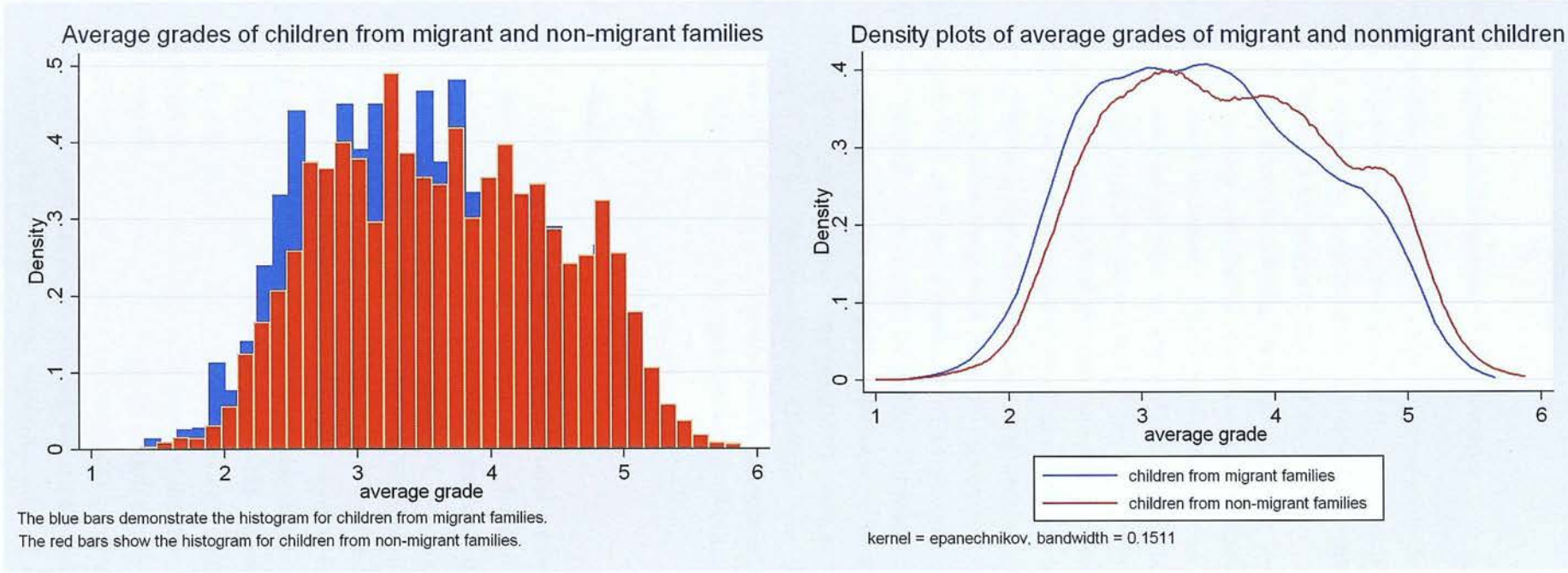


Figure 1.4: Average grades of respondents from migrant and non-migrant families
Source: MECP2012

1.5 Representativeness of the sample

The population of interest are children in lower secondary education. Given the choices made during the data collection process (i.e. the high migration region in Poland, the main migrant-sending country among the new EU member states) and the financial and time constraints of the project, one may be concerned about the internal and external validity of any analysis utilising the data.

Firstly, one may ask whether the data truly characterise the situation of the PWA children in lower secondary schools in Opolskie, even though the initial descriptive sample statistics reflect what we know about migrant families in the region. Arguably, schools and participants can opt out of the study, which may compromise the representativeness of the sample if the non-participation is not random. In Section 1.5.1 I argue that school and participant selection can be thought of as almost random.

Another worry is that Opolskie may not be representative of the situation in Europe, as it has been experiencing high levels of population outflow, both historically and in recent years. The scale and persistence of the phenomenon may have led to a different response of families to temporary migration. For example, there may be policies in place to support migrant-sending families. Moreover, if having a parent working abroad is perceived as a norm, children may differently react to it than if migration was a new phenomenon. Thus, the situation in Opolskie may differ from that in the rest of Poland and other European migrant-sending countries. I discuss the extent to which these concerns may limit any generalisations in Section 1.5.2.

1.5.1 Internal validity

School selection

The selectivity of schools in the process may raise concerns, particularly if those which opted out from cooperation, are believed to be differently affected by the phenomenon studied; for instance, I may find that PWA children's grades do not differ from other pupils' and conclude no impact of emigration on school performance. However, due to self-selection of schools, there might be a number of differences between the PWA children in participating schools and those, who were excluded from the study. Perhaps the participating schools agreed to cooperate because they do not perceive emigration as problematic and the children included in the sample were not affected, whilst those excluded might have been. In such a case the target population will not be well-represented and results may not have a causal interpretation.

The feedback given by schools, however, undermines the argument of schools' self-selection into the study when emigration within pupils' families does not cause problems. Participant schools perceive emigration as problematic.

Although the problem of schools' self-selection should not be neglected, participation decisions might not have been driven by migration situation in the school. The

engagement in the project required additional effort from the schools' administration in form of grade provision and their pupils' time. This in itself became a discouraging factor. The negative attitude might explain why as many as 35 of the institutions, who refused cooperation provided no sound reason for the refusal; 13 schools expressed concerns about the timing of the project, which coincided with audits, lay-offs of teachers and school trips. Only 8 schools stated clearly that the problem lay in the request to access information on performance of children and their family situation; this data was perceived as sensitive.

As can be seen from Figure 1.5, the participating schools are equally spread across the entire region. The highest percentages of respondents in the whole sample come from *opolski*, *oleski* and *strzelecki* counties; these areas are also among the top five emigration areas in the region. Only *krapkowicki* and *kędzierzyńsko-kozielski* counties with the highest number of migrants in 2002 could be of concern, given that local schools were reluctant to cooperate there.

In most cases, again with exception of *krapkowicki* county, the refusal of schools to cooperate coincided with low population density in the area (see Figure 1.5), indicating that the most populous areas have been well captured in the study.

The following counties have been particularly well-covered: *oleski, namysłowski* and *strzelecki.* As mentioned before, *oleski* and *strzelecki* counties are characterised by one of the highest emigration rates. A response much below the voivodship average has occurred in *brzeski, glubczycki* and *kędzierzyńsko-kozielski* counties. The last one might be of concern, given a relatively high temporary out-migration from the region. However, a different light may be shed on the earlier concern about the underrepresentation in *krapkowicki* county; the participation rate in the study in this county is still lower than the voivodship average, but it is not the lowest across the areas covered.

Further, the counties with a large number of temporary migrants staying abroad according to the 2002 Census are relatively well-represented in the study (see Table 1.5).[15] At the voivodship level, almost half of the contacted schools participated in the study, providing a capture of over a third of all students (see Figure 1.6).

Even if the areas are unequally represented in the data set, the counties do not differ strikingly in terms of their local economy. From Table 1.5 it is clear that the average gross salary and wages in 2011 mostly varied between 2730.02 PLN (in *prudnicki* county) to 2872.04 PLN (in *opolski* county) (The Central Statistical Office of Poland, 2012). The only exceptions are *krapkowicki* and *kędzierzyńsko-kozielski* counties, where the average gross salaries reach 3798.54 PLN and 3518.97 PLN respectively. These two outcomes are closer to the national average which was 3315.38 PLN in 2011 (The Central Statistical Office of Poland, 2012). The difference is driven by the existence of an industry in both counties, in contrast with the rest of the predominantly rural

[15] I am predominantly interested in the areas with high number of temporary emigrants as they are more likely to leave families behind in Poland. Deregistration from an address in Poland to emigrate is usually equivalent to an uprooting of the whole family from Poland.

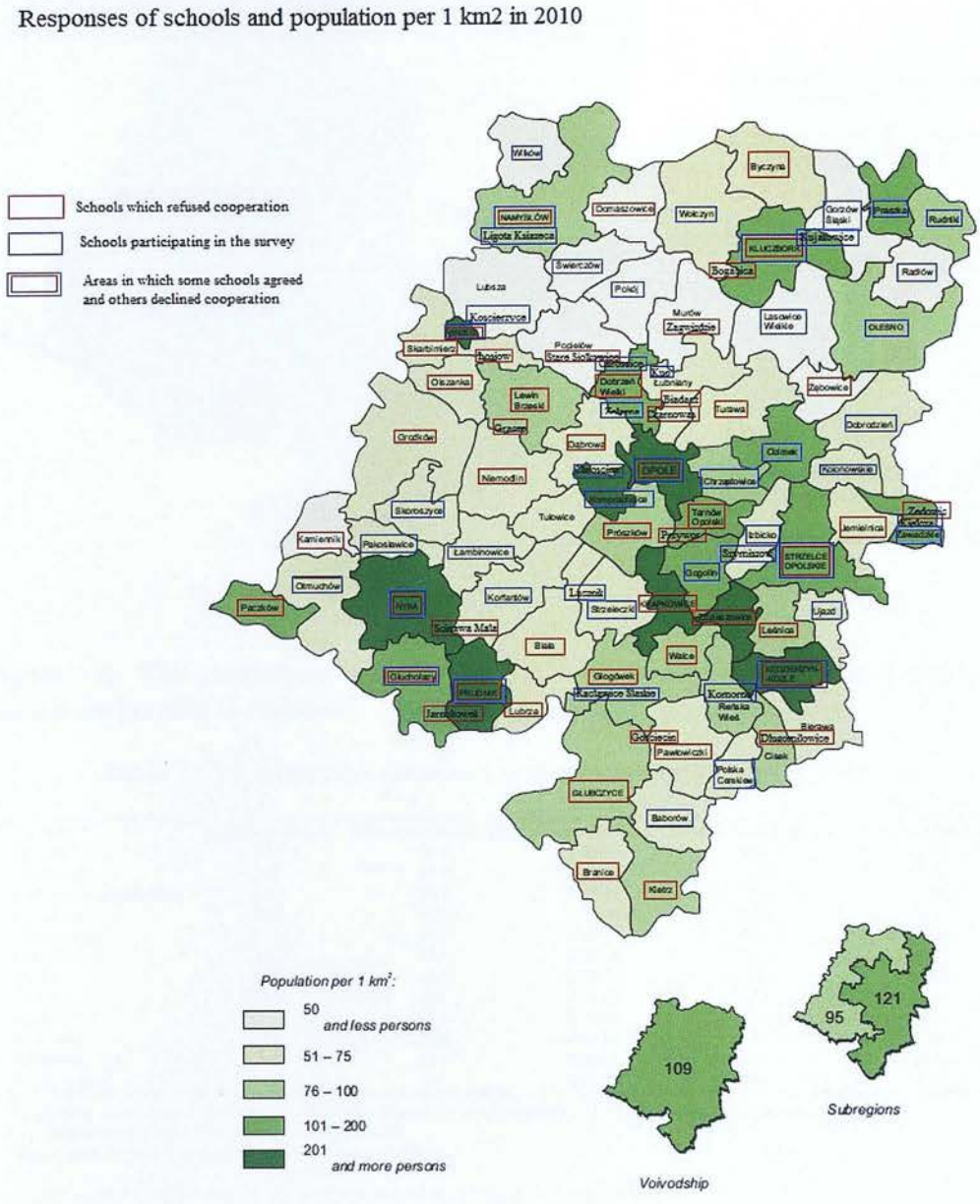


Figure 1.5: Map of school responses, Source: Central Statistical Office of Poland and own calculations

voivodship.

Failure to fully represent areas of higher average income might impact the analysis. Given relatively higher incomes of families and the relationship between house-

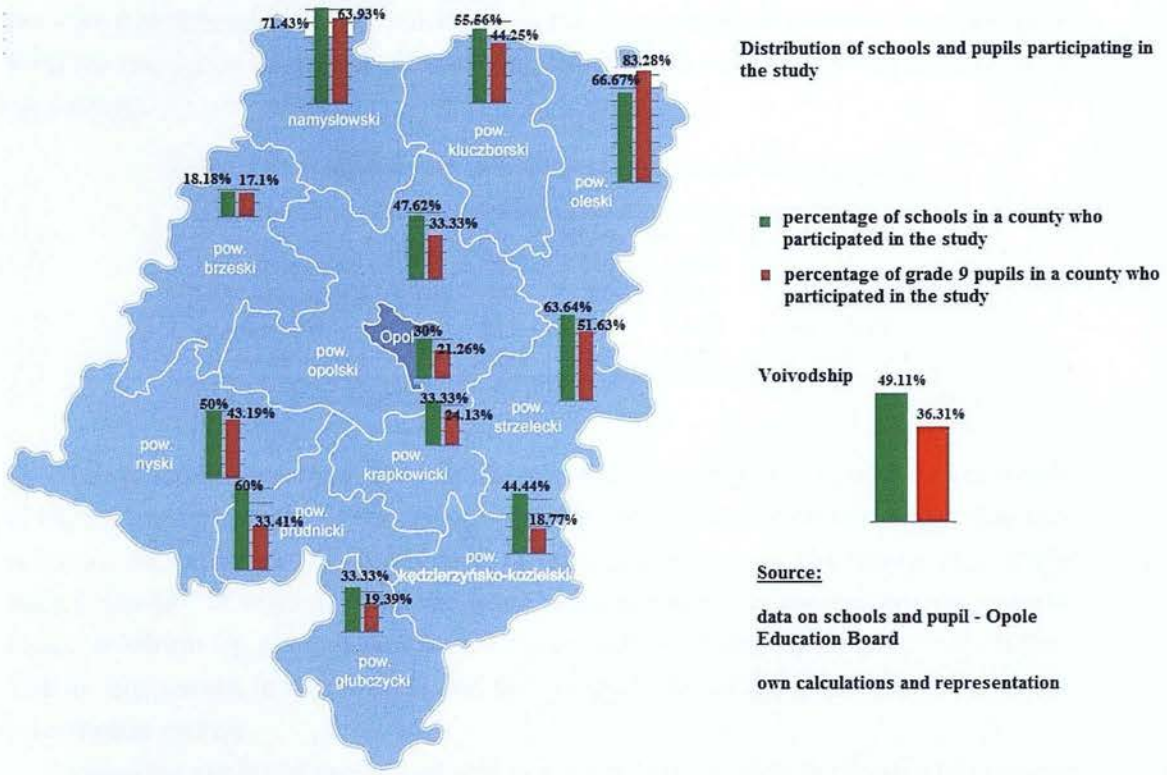


Figure 1.6: The participant schools and pupils as a percentage of total number of schools and pupils in counties

Table 1.5: Emigration rates and economic situation in the counties

County	Emigration (%)	Unemployment (%)	Wages (PLN)	% of 3rd year pupils	% of respondents
Brzeski	3.11	20.5	2795.69	10.09	4.75
Glubczycki	5.57	17.9	2878.02	4.89	2.08
Kedzierzynsko-kozielski	12.65	13.1	3518.97	9.62	5.61
Kluczborski	8.08	15.5	2848.38	7.35	10.43
Krapkowicki	16.60	10.9	3798.54	6.45	4.54
Namyslowski	4.35	18.6	2833.22	4.62	6.30
Nyski	4.63	19.4	2733.31	14.71	10.25
Oleski	12.07	8.9	2731.82	6.74	15.80
Opolski	17.98	13.1	2872.04	12.64	16.38
Prudnicki	9.95	18.6	2730.02	5.76	6.58
Strzelecki	17.27	11.7	2929.69	7.53	11.40
miasto Opole	4.99	6.4	3541.80	9.60	5.89

Emigration: number of people staying temporarily abroad for over 2 months as % of the population in Census 2002
Unemployment: registered unemployment rate in 2011
Wages: average gross salaries and wages in 2011, in PLN
Source:
emigration data: the 2002 Census, Central Statistical Office of Poland, own calculations
unemployment and wages data: Central Statistical Office of Poland

hold budget and educational attainment of the offspring, children in *krapkowicki* and *kędzierzyńsko-kozielski* counties might be on average better off before, as well as after, parental emigration relative to children in other areas. Moreover, considering the high out-migration from the two areas, the increased average income might signal a significant remittance flow, not just the existence of local industry. All of these factors may lead to better school performance of children from the area.

Looking at the statistics presented in Table 1.6, however, it becomes clear that

the school performance of respondents from the two counties in question does not differ from the average in the sample; if anything, the children seem to perform slightly worse on average.

Table 1.6: Respondents' school performance (average grade)

	n	mean	std. dev.	min	max
overall sample	2822	3.610	0.850	1	5.88
kędzierzyńsko-kozielski	340	3.392	0.852	1.4	5.79
krapkowicki	150	3.575	0.845	1.66	5.77
all other schools	2332	3.621	0.849	1	5.88

Source: MECP2012

The variance in unemployment in the voivodship is much higher, with a clear divide of higher unemployment in the western part of the region, where the emigration rate is lower. As expected, the lowest level of unemployment is in the capital city of the region, Opole. It is likely that the lower unemployment in the eastern part of the region is driven by a significant and regular outflow of the working-age population. The unemployment in *krapkowicki* and *kędzierzyńsko-kozielski* counties are close to the voivodeship average.

I assess the quality of participant and non-participant schools in the area by comparing the average outcomes of their pupils in the final exam in 2012.[16] Any differences in performance between the two groups may suggest that indeed schools have selected into the study in a non-random way. The results are presented in Table 1.7. Pupils in non-participant schools performed worse on average in the final exam, but the differences are insignificant and support the conclusion that the respondent group is representative of the entire population.

Table 1.7: Average test scores in 2012 in schools in Opolskie

	Participant schools				Non-participant schools				
	Mean	St.dev.	Min	Max	Mean	St.dev.	Min	Max	T-stat
Humanities test score	62.251	6.152	51.75	88.65	61.572	6.197	38.95	79.5	.629
Science test score	48.554	7.249	39.1	82.1	48.142	6.088	29.15	66.4	.344
N	52				88				

Source: MECP2012

Pupils' participation decision

Another estimation challenge arises if respondents select into the study in a non-random manner. A request to disclose personal information is more likely to prompt a refusal to answer the questionnaire. One particular worry is that, given the sensitive nature of migration in Poland, individuals in the treatment group may refuse to cooperate or may answer the questions partially. It may also be argued that even when students

[16]The exam was taken by the final year pupils, which are the respondent group in this study.

do not self-select actively, their non-attendance to school on the day is a form of self-selection. This should be of concern if we believe that students who are more likely to miss school on the day differ significantly from their peers, especially if they also are PWA children. Then the results do not reflect the situation fully.

Table 1.8: Survey response rate

	n		
total of pupils in surveyed schools	3423		
pupils present during the survey	2863		
total number of responses	2822		
	average	min	max
response rate of total pupils of the school	82.47	58.54	98.39

Source: MECP2012

Table 1.9: Average outcomes for respondents and non-respondents

	Respondents				Non-respondents				
	Mean	St.dev.	Min	Max	Mean	St.dev.	Min	Max	T-stat
average grade	3.61	0.850	1	5.88	3.412	0.851	1	5.72	4.987
behavioural grade	4.489	1.240	1	6	4.259	1.291	1	6	3.840
number of hours missed not excused	12.131	30.802	0	542	19.904	46.281	0	540	3.932
N	2822				548				

Source: MECP2012

There have been no signs of self-selection within the chosen schools, however. As can be seen in Table 1.8, the majority of pupils present at school on the day of the study filled in the questionnaire. The response rate among the pupils present varied from 89.66 to 100% across participating schools. The number of respondents constituted on average 82.47% of the overall school population.

In Table 1.9 I present a summary of outcomes for students who took part in the survey and those who did not respond or were absent at school on the day. The non-respondents have on average lower average grade, worse behavioural grade and miss more school without an excuse. The differences between students who participated in the study and those who didn't are statistically significant. This will have implications for validity of the results I present in later chapters if the non-participation was non-random. In particular, one may be concerned that PWA pupils are overrepresented among the non-respondents and that their parents' migration decision is related to their worse school performance. I discuss this issue in more detail when introducing the empirical framework for analyses in subsequent chapters.

1.5.2 External validity

Poland and the region of study were not chosen randomly. They were targeted in order to capture a sufficiently large sample of children in lower secondary education who have parents temporarily employed abroad.

To be able to generalise the results in the following two chapters, relating to the impact of parental emigration on school performance of children and, in turn, the classroom spillover effect, I need to ensure that 1) Opolskie does not differ from the rest of Poland and other regions in Europe in terms of the scale of the phenomenon studied and 2) that the effect is likely to be similar in other areas.

The relationship between parental migration and students' educational performance may depend on the broader context in which migration takes place. The scale and persistence of the phenomenon may play a role. As mentioned before, schooling in a region which has newly experienced migration may be differently affected by it than in a region in which migration is a norm.

I will highlight similarities between Opolskie and other regions of Poland, as well as other European countries to argue that the results can be generalised to an extent. Unfortunately, the discussion is constrained by the scarce data on temporary migration and PWA children in other regions.

Migration levels in Opolskie and other regions of Poland

Opolskie has experienced prolonged high levels of emigration, both temporary and permanent. The region had the highest in Poland number of temporary migrants per 1000 residents in Census 2011, which was twice the country average.

However, I am interested in a very specific subgroup of all temporary migrants - those who have teenage children still living in Poland. Migration levels for this group may differ from the overall outcome. The 2011 Census does not provide information about temporary migrants by family status. Therefore it is difficult to establish whether migration rates for this particular group are also disproportionately high when compared with the rest of Poland.

I make comparisons across age groups, relying on the fact that the average age for a migrant mother in the sample is 40 years old and for a migrant father 43 years old. In Table 1.10 I present information about the rates of temporary migration in all regions of Poland for individuals aged 30-39 and 40-49 as well as the percentage of households in a region with at least one temporary migrant. Also here the migration rates are the highest in Opolskie.

The age-specific group comparisons are not ideal, however, as being of certain age is not a perfect indicator of having a child in lower secondary education. There should, nonetheless, be a correlation between the two.

PWA children and their parents in other regions of Poland

Ideally, I would like to compare the statistics on a number of PWA pupils in other regions of Poland. According to the sample statistics almost 18% of pupils in lower secondary schools in Opolskie had a parent abroad at some point in the observed 3 year period. On average, however, about 7% had a parent abroad in a given semester t, since migratory movements in question are temporary. In majority of cases fathers

Table 1.10: Further migration information from 2011 Census

Region of Poland	% of temporary migrants by age group		% of households with a temporary migrant
	Aged 30-39	Aged 40-49	
Dolnośląskie	11.4	6.9	11.0
Kujawsko-Pomorskie	9.3	5.2	9.6
Lubelskie	10.4	5.4	9.9
Łódzkie	5.6	2.9	11.2
Małopolskie	10.4	6.5	5.6
Mazowieckie	4.7	3.3	5.4
Opolskie	19.4	15.9	17.8
Podkarpackie	16.0	8.3	15.9
Podlaskie	17	10.7	15.2
Pomorskie	10.2	6.8	10.8
Śląskie	8.6	6.2	8.9
Świętokrzyskie	9.9	5.1	10.0
Warmińsko-mazurskie	13.5	7.5	13.3
Wielkopolskie	5.8	3.1	6.6
Zachodnio-pomorskie	11.7	6.6	11.2
Poland	9.5	5.9	9.9

Source: 2011 Census, the Central Statistical Office of Poland

engaged in emigration and spent relatively short periods of time abroad.

Comprehensive comparisons with other regions of Poland are impossible as no detailed data on PWA children in Poland exist. However, a few localised studies, results of which are outlined in Table 1.11, have been undertaken in recent years.

It is important to acknowledge the limitations of this exercise. The studies mentioned here are piecemeal, descriptive and sometimes based on a small sample of respondents. They also asked respondents different questions, so the statistics are not always directly comparable. All of them, however, have selected participants in the given region in a randomised manner.

Between 7 and 11% of pupils of lower secondary schools in the studied regions had a parent abroad. This is a lower percentage than in my data. However, the respondents in these cases were usually asked whether their parent was abroad at the time of the survey; the studies did not collect retrospective information. I find that on average only 7% of respondents in MECP2012 data had a parent abroad at a given point in time, which is closer to what the studies in question report. Moreover, official statistics based on information provided by schools may underestimate the scale of the problem, as parents often do not inform schools of their emigration.

The report by Walczak (2006) on the situation in Mazowieckie region of the country, capturing the capital city of Warsaw, is by far the most comprehensive and reliable. It provides further information about the situation of PWA children in the region. In particular, the author points out that in majority of cases only one parent emigrates, usually the father. Parental migration is short term, with an average length of stay of 5.7 months. These characteristics closely reflect those of the PWA families in MECP2012

Table 1.11: PWA pupils in other regions of Poland

Region	Author	Info on PWA pupils
Zachodnio-pomorskie	Zajączkowska (2008)	7% of pupils in lower secondary schools almost always one parent abroad
Mazowieckie	Walczak (2006)	11.3% of pupils in lower secondary schools 9.1% had a father abroad 3.7% had a mother abroad average stay abroad: 5.7 months average age of migrant mother: 38 average age of migrant father: 43
Podlaskie	Regionalny Ośrodek Polityki Społecznej w Białymstoku (2011)	6% of pupils had a father abroad 2% of pupils had a mother abroad
Poland	DG Employment, Social Affairs and Inclusion (2012)	15% of pupils aged 9-18

data. The two groups are also comparable with respect to PWA parents' age and education levels.

Importantly, a report by the European Commission (DG Employment, Social Affairs and Inclusion, 2012) suggests that 15% of Polish children are PWA children. This is closer to the statistics obtained from MECP2012 data. It is stressed, however, that the estimates bear significant uncertainty and imprecision.

Although rather scant, the evidence is suggestive of similarities in terms of the scale of the phenomenon across regions of Poland.

Different response to migration depending on the scale and circumstances

There is some evidence that the percentage of PWA children in Opolskie is not as high relative to the rest of the country as the general migration statistics are suggesting. The greater migration from Opolskie will be problematic for generalisation of results if it shapes differently the relationship between having a parent abroad and school outcomes or the spillover effect of that relationship.

This is likely if, due to the prevalence of the phenomenon, the region authorities have introduced policies to target PWA children. To the best of my knowledge no such policies exist. However, it is still possible that individuals respond differently to having a parent abroad if parental migration is a common occurrence in their environment. In particular, the effect of separation from a parent in a high migration region may be less pronounced as having a parent abroad is perceived as normal. However, as will be stipulated in the following chapters, in case of Poland the burden of separation is thought to be relatively small due to the nature of parental migration.

If the large scale migration had a different impact in the specific context of schooling, one may expect to see different performance of pupils across the regions of Poland depending on their experience of migration. I do not possess the individual level schooling data for Poland, but in Table 1.12 I present the average results of the national exams for the relevant cohort, by region. Students in Opolskie performed on or close to the average in the exam. There are no strong indications of their differential performance.

Table 1.12: Average Test Scores (%) in the National Exam in 2012

Region of Poland	Polish	History	Maths	Science
Dolnośląskie	63	60	46	49
Kujawsko-Pomorskie	64	59	46	49
Lubelskie	67	61	47	50
Łódzkie	65	60	46	49
Małopolskie	69	63	50	51
Mazowieckie	67	63	50	52
Opolskie	63	60	47	49
Podkarpackie	67	62	49	51
Podlaskie	64	61	49	51
Pomorskie	62	59	47	49
Śląskie	66	61	47	50
Świętokrzyskie	65	60	46	49
Warmińsko-mazurskie	62	60	46	49
Wielkopolskie	63	60	46	49
Zachodnio-pomorskie	62	59	45	49
Poland	65	61	47	50

Source: Centralna Komisja Egzaminacyjna

PWA children in other countries

Can the results be generalised beyond Poland? Poland and most countries which joined the European Union in and after 2004 share common experiences related to the economic and political changes over the past 20 years. Many of them have moved from socialist to market based economies and introduced democracy over a short period of time. These changes put them on a similar footing in terms of economic performance nowadays, although they all face country-specific difficulties.

In particular, they all strived to join the European Union, which committed them to meeting certain economic and political conditions to allow free movement of goods, services and individuals. Following the EU accession they have all experienced significant migratory flows, which were in some instances restricted over the initial membership period. According to DG Employment, Social Affairs and Inclusion (2012) the stock of the EU-10 nationals residing in the old Member States tripled over the period between 2003 and 2009.

It is difficult, however, to precisely estimate the resultant migratory flows (particularly when they are temporary) due to the nature of the movements within the EU. In particular, statistics provided by Eurostat are based on immigration data supplied by the member states using administrative records, sample surveys or estimates; these sources focus mostly on long-term migrants and are unlikely to capture temporary migration well.

Moreover, the generalised migration statistics are unlikely to provide accurate information about the group of interest. The situation of PWA children in the EU and Europe more broadly has not been extensively studied. However, in Table 1.13 I provide information on the few available analyses which may help shed light on the scale of

the problem. Once again, this is a collection of results from various sources, based on different measures and often targeting different groups of children. Importantly, there is considerable uncertainty attached to the estimates.

Most analyses indicate that between 15 and 22% of children in various new EU member states and other Eastern European countries have a parent working abroad. This number reflects my findings in the MECP2012 data, which appears reassuring for representativeness of the data for Europe.

At the same time, some sources provide much lower estimates. In particular, according to the official statistics for Romania 2% or 7% of children have a migrant parent. This may be due to severe underreporting of temporary migration to the authorities. The estimates on the higher end of the spectrum probably reflect the actual situation more closely.

Table 1.13: PWA pupils in other European countries

Region	Source	Info on PWA pupils	Comments
EU as a whole	DG Employment, Social Affairs and Inclusion (2012)	500 000 PWA children	the estimate is uncertain
EU countries			
Poland	DG Employment, Social Affairs and Inclusion (2012)	15% of pupils have a parent abroad	info for all children aged 9-18
Romania	Botezat and Pfeiffer (2014)	2% of children aged 0-18 have a parent abroad	official statistics data for year 2008
		20% of students enrolled in high school have a parent abroad	an independent survey in 2007
		21% of children in grades 5 to 8 have a parent abroad	data from Gallup International for year 2007 parent abroad for at least 12 months
	DG Employment, Social Affairs and Inclusion (2012)	85 000 children have at least one parent abroad	official government report in 2011
	DG Employment, Social Affairs and Inclusion (2012)	350 000 children have at least one parent abroad i.e. 7% of all children aged 0-18	UNICEF report in 2007
Lithuania	DG Employment, Social Affairs and Inclusion (2012)	9500 PWA children	
Bulgaria	DG Employment, Social Affairs and Inclusion (2012)	problem recognised but no data	
Latvia	DG Employment, Social Affairs and Inclusion (2012)	problem recognised but no data	
Estonia	DG Employment, Social Affairs and Inclusion (2012)	problem recognised but no data	
Other European countries			
Moldova	Botezat and Pfeiffer (2014)	17% had at least one parent abroad	info for all children aged 0-18
Albania	Botezat and Pfeiffer (2014)	22% live in a migrant household	info for all children aged 0-18

Bibliography

Botezat, A., and F. Pfeiffer (2014): "The Impact of Parents' Migration on the Well-being of Children Left Behind: Initial Evidence from Romania," IZA Discussion Paper 8225.

Cortes, P. (forthcoming): "Feminisation of international migration and its effects on the children left behind. Evidence from the Philippines," *World Development*.

DG Employment, Social Affairs and Inclusion (2012): "Social Impact of Emigration and Rural-Urban Migration in Central and Eastern Europe," Synthesis Report.

Jończy, R. (2007): "Migracje zarobkowe z województwa opolskiego. Przyczyny i charakterystyka pracy za granicą oraz warunki powrotu," Discussion paper, Fundacja Rozwoju Śląska oraz Wspierania Inicjatyw Lokalnych and Regionalny Ośrodek Europejskiego Funduszu Społecznego and Akademia Ekonomiczna we Wrocławiu, Katedra Ekonomii i Gospodarowania Środowiskiem, Opole.

Jończy, R., and D. Rokita-Poskart (2013): "Fiscal effects of migration as a factor increasing disparities in local and regional development - an example of the Opole Voivodship in Poland," *Economic and Environmental Studies*, 13, 9–21.

Kaczmarczyk, P., and M. Okólski (2008): "Demographic and labour market impacts of migration on Poland," *Oxford Review of Economic Policy*, 24(3), 599–624.

Madajczyk, P. (1996): *Przyłączenie Śląska Opolskiego do Polski 1945-1948*. Instytut Studiów Politycznych PAN, Warszawa.

McKenzie, D., and H. Rapoport (2011): "Can migration reduce educational attainment? Evidence from Mexico," *Journal of Population Economics*, 24, 1331–1358.

Regionalny Ośrodek Polityki Społecznej w Białymstoku (2011): "Eurosieroctwo - skala zjawiska w województwie podlaskim," .

Regionalny Ośrodek Polityki Społecznej w Opolu (2011): "Zjawisko eurosieroctwa w województwie opolskim," Raport z badań przeprowadzonych przez Regionalny Ośrodek Polityki Społecznej w Opolu.

The Central Statistical Office of Poland (2012): "Poland in figures, 2012," .

——— (2013a): "Migracje zagraniczne ludności. Narodowy Spis Powszechny Ludności i Mieszkań 2011," .

——— (2013b): "Vademecum Samorządowca 2013. Województwo Opolskie.," Discussion paper, Warsaw.

——— (2013c): "Wybrane tablice dotyczące emigracji na pobyt czasowy powyżej 3 miesięcy - wyniki Spisu Ludności i Mieszkań 2011," .

Tynelski, A. (2010): "Sytuacja dziecka w rodzinie migrującej w Polsce," Discussion paper, Ministerstwo Edukacji Narodowej.

Walczak, B. (2006): "Migracje poakcesyjne z perspektywy ucznia," Pedagogium.

Zajączkowska, K. (2008): "Wyniki diagnozy problemowej zjawiska zwanego eurosieroctwem, wynikającym z rosnącej skali emigracji zarobkowej do krajów wspólnoty europejskiej rodziców uczniów szkół podstawowych i gimnazjów województwa zachodniopomorskiego," Kuratorium Oświaty w Szczecinie.

Appendix

1.A Project execution and variables

Table 1.A.1: Participant schools

Location	Number of 16 year olds	Type of area
Baborów	68	4
Brzeg	121	6
Chróścice	26	3
Chróścina	41	2
Chrząstowice (Dębska Kuźnia)	48	2
Dobrodzień	97	3
Głuchołazy	98	6
Gogolin	102	5
Gorzów Śląski	80	4
Izbicko	45	2
Kędzierzyn-Koźle	70	4
Kędzierzyn-Koźle	39	4
Kielcza	26	2
Kluczbork	104	6
Kluczbork	47	6
Kolonowskie	59	3
Komorno	32	2
Komprachcice	79	3
Kościerzyce	55	1
Kujakowice Górne	36	1
Kup	40	1
Łącznik	32	1
Łambinowice	95	3
Lasowice Wielkie	38	1
Ligota Książęca	33	1
Namysłów	107	6
Nysa	40	6
Olesno	19	5
Olesno	56	5
Opole	103	7
Opole	67	7
Opole	47	7
Otmuchów	96	4
Ozimek	48	5
Ozimek	121	5
Pakosławice	34	1
Pokój	53	1
Polska Cerekiew	39	2
Praszka	135	4
Prudnik	174	6
Racławice Śląskie	28	2
Radłów	34	3
Rudniki	89	1
Strzelce Opolskie	75	6
Strzeleczki	75	2
Świerczów	28	1
Szymiszów	72	2
Ujazd	61	2
Wilków	61	1
Wołczyn	144	4
Zawadzkie	76	4
Zelazna	28	1
Total students:	**3353**	

Table 1.A.2: Schools which refused participation

Location	Number of 16 year olds	Type of area	Reason for denial
Biadacz	70	1	x
Biała	69	3	lack of time
Bogacica	50	2	x
Branice	76	3	lack of time
Brzeg	136	6	x
Brzeg	103	6	x
Byczyna	100	3	x
Czarnowąsy	42	3	x
Dąbrowa	31	1	lack of time
Długomiłowice	60	2	could not find a suitable date
Dobrzeń Wielki	67	4	x
Domaszowice	40	2	sensitive data
Głogówek	100	4	sensitive data
Głubczyce	152	5	lack of time
Głuchołazy	85	5	sensitive data
Gościęcin	29	1	x
Gracze	45	2	lack of time
Grodków	100	4	x
Grodków	78	4	x
Grodków	115	4	sensitive data
Jarnołtówek	70	1	x
Jemielnica	60	3	lack of time
Kamiennik	34	1	x
Kędzierzyn-Koźle	104	6	lack of time
Kędzierzyn-Koźle	140	6	sensitive data
Kędzierzyn-Koźle	140	6	x
Kietrz	129	4	x
Kluczbork	135	6	lack of time
Krapkowice	70	5	lack of time
Krapkowice	80	5	lack of time
Leśnica	86	3	x
Lewin Brzeski	90	4	x
Łosiów	27	2	no problem of migration
Lubrza	34	1	x
Namysłów	95	6	x
Niemodlin	77	4	x
Nysa	163	6	x
Nysa	72	6	x
Nysa	200	6	lack of time
Olszanka	54	1	x
Opole	166	7	lack of time
Opole	162	7	x
Opole	143	7	x
Paczków	118	4	x
Pawłowiczki	50	2	x
Prószków	30	3	sensitive data
Prudnik	120	6	sensitive data
Przywory	37	2	x
Ścinawa Mala	27	1	x
Skarbimierz	70	1	x
Stare Siołkowice	80	1	lack of time
Strzelce Opolskie	88	6	lack of time
Tarnów Opolski	80	3	x
Turawa	74	1	sensitive data
Walce	65	2	x
Zagwiździe	30	1	x
Zdzieszowice	164	5	x
Zębowice	40	2	x
Zędowice	22	3	
Total students:	**4974**		

Table 1.A.3: Other schools in the area

Location	number of 16 year olds	Type of area	Reason
Bierawa	48	2	no contact made
Bogdanowice	10	1	no contact made
Brzeg	45	6	special needs
Chocianowice	30	1	did not answer the phone
Cisek	50	2	no contact made
Głogówek	10	4	no contact made
Głubczyce	45	5	no contact made
Jędrzejów	30	1	did not answer the phone
Kędzierzyn-Koźle	10	6	did not answer the phone
Kędzierzyn-Koźle	15	6	did not answer the phone
Kędzierzyn-Koźle	20	6	did not answer the phone
Kędzierzyn-Koźle	30	6	did not answer the phone
Kluczbork	15	6	no contact made
Krapkowice	25	5	no contact made
Lisięcice	63	1	no contact made
Opole	49	7	no contact made
Pietrowice	10	2	no contact made
Prószków	30	3	did not answer the phone
Skorogoszcz	30	2	did not answer the phone
Smogorzów	25	1	did not answer the phone
Solarnia	20	1	no contact made
Strzelce Opolskie	70	6	no contact made
Tułowice	45	3	did not answer the phone
Zimnice Wielkie	25	1	did not answer the phone
Nysa	0	6	did not have a class
Opole	0	7	did not have a class
Korfantów	74	2	withdrawn at a later stage
Skoroszyce	70	2	withdrawn at a later stage
Nysa	121	6	withdrawn at a later stage
Total students	**1015**		

Table 1.A.4: Classification of settlements

population	classification	code
up to 1000	small village	1
1000 - 2250	village	2
2250 - 4500	intermediate settlement	3
4500 - 10000	small town	4
10000-18000	medium town	5
18 000 - 75000	large town	6
above 75 000	city	7

THE UNIVERSITY
of EDINBURGH

Thank you for agreeing to participate in the survey. Your answers will be treated **anonymously** and will only be used in my PhD project. Answer the questions honestly. There are no incorrect answers. If you feel that none of the given choices reflects your situation, try to choose the best alternative answer. **Circle your chosen answer or tick the box if required**. If you choose 'other' as your answer, please elaborate.

1. Class .....
2. Number in class register .....
3. Year of birth .....

4. Gender: Female / Male

5. What nationality are you?
   Polish / Polish and German / Other:.....

6. What is your parents' nationality?

	Polish	Polish & German	Other (what?)
mother			
father			

7. How many siblings do you have?

0	1	2	3	more

8. What is your birth order in the family (e.g. if you are the oldest, circle 1)?

1	2	3	4	Higher order

9. What age group are your siblings? More than one answer is possible.

0-5	6-10	11-15	16-18	Older than 18

10. Do you live with both parents?
    Yes / No
    a. If not, why? .....

11. If you do not live with both parents, who do you live with?
    mother / father / grandparents / another family member / someone else: .....

12. Does your mother have a job? Yes / No
13. What is your mother's age? .....

14. What type of school has your mother graduated from?

podstawową	zawodową	liceum / technikum	studia

15. What is your mother's profession? (e.g. teacher, cook, hairdresser)
    .....

16. Does your father have a job? Yes / No
17. What is your father's age? .....

18. What type of school has your father graduated from?

podstawową	zawodową	liceum / technikum	studia

19. What is your father's profession? (e.g. mechanic, doctor, driver)
    .....

20. Do you attend any after-school activities/ classes?
    Yes / No
    a. If yes, of what kind. Tick all correct answers

sports	
languages	
art	
course related	
volunteering	
other (what kind?)	

21. Are these after-school activities organised by your school?
Yes / No / Some of them are

22. Are you intending to apply to attend university after secondary school?
Yes / No

23. Are you planning to leave for work abroad after graduating from secondary school?
Yes / No

24. Do your parents have family or friends abroad?
Yes / No

25. Has your mother, father or sibling recently moved abroad?
Yes / No

If you answered positively to Question 25, please answer the remaining questions. Otherwise this is the end of the survey.

26. Who in your family has recently emigrated?

mother	father	siblings

27. Where has your family member left to? Please state country and city/region.

	mother	father	siblings
country			

28. Has any of family or friends living abroad helped your family member to emigrate? (e.g. by finding job or accommodation)

Yes / No / I do not know

29. **Tick the box only if** your mum/ your dad / your sibling **was staying abroad** in a given time period:

	mother	father	siblings
just before September 2009			
I semester of 1st year (2009/2010)			
II semester of 1st year (2009/2010)			
summer of 2010			
I semester of 2nd year (2010/2011)			
II semester of 2nd year (2010/2011)			
summer of 2011			
I semester of 3rd year (2011/2012)			
II semester of 3rd year (2011/2012)			
He/ she has emigrated permanently			

Only answer questions 30 – 33 if one or both of your parents emigrated.

30. If your parent is/ was abroad, how often do/did you see her / him?

	once a month or more often	every 6 months	once a year	more rarely than once a year
mother				
father				

31. If your parent is/ was abroad, have you ever been to visit him/her?
Yes / No

32. Do you feel you have more home responsibilities (e.g. housework, caring for siblings), because of the departure abroad of your relative?
Yes / No

33. Do you think your grades have been affected by the emigration in your family?
Yes / No

a. If yes, in what way? Positive / Negative

Table 1.A.5: Variables available from the data set

Variable	Description
ID	student's ID number
day	day of birth
month	month of birth
year	year of birth
sex	sex of the respondent: = 1 woman, = 0 man
Nationality	
Polish	respondent's nationality: =1 if only Polish
dual	respondent's nationality: =1 if Polish and German
other	respondent's nationality: =1 if not 'Polish' or 'dual'
mumpolish	mother's nationality: =1 if only Polish
mumdual	mother's nationality: =1 if Polish and German
mumother	mother's nationality: =1 if other than Polish or Polish and German
dadpolish	father's nationality: =1 if only Polish
daddual	father's nationality: =1 if Polish and German
dadother	father's nationality: =1 if not 'Polish' or 'dual'
Family demographics	
sibling	number of siblings: = 0 if 0, =1 if 1, =2 if 2, =3 if 3, =4 if more
BirthOrder	birth order among siblings: = 1 if the oldest, =2 if 2nd, =3 if 3rd, = 4 if 4th, =5 if further
Sibling0-5	siblings aged 0-5, 1 = yes
Sibling6-10	siblings aged 6-10, 1 = yes
Sibling11-15	siblings aged 11-15, 1 = yes
Sibling16-18	siblings aged 16-18, 1 = yes
OlderSibling	siblings older than 18, 1 = yes
BothParents	Does the respondent live with both parents? =1 if yes
WhyDivorce	: =1 if parents are divorced
WhyDeath	: = 1 if one or both parents dead
WhyMigration	: = 1 if one or both parents emigrated
WhyOther	: = 1 if other reasons for not living with both parents
WhichDad	: = 1 lives with dad
WhichMum	: = 1 lives with mum
WhichGrand	: = 1 lives with grandparents

Continued on next page

Table 1.A.5 – continued from previous page

Variable	Description
WhichOther	: = 1 lives with another member of the family
WhichInstit	: = 1 is put into care
Parental background	
MumWork	Does the respondent's mother work? = 1 if yes
MumAge	Mother's age
mum ed-pods	Mother's education - elementary: =1 if yes
mum ed-zas	Mother's highest education level - vocational: =1 if yes
mum ed - secon	Mother's highest education level - secondary: = 1 if yes
mum ed- tert	Mother's highest education level - tertiary: = 1 if yes
MumJob	Mother's profession
DadWork	Does the respondent's father work? = 1 if yes
DadAge	Father's age
dad ed-pods	Father's education - elementary: =1 if yes
dad ed-zas	Father's highest education level - vocational: =1 if yes
dad ed - secon	Father's highest education level - secondary: = 1 if yes
dad ed- tert	Father's highest education level - tertiary: = 1 if yes
DadJob	Father's profession
After school activities:	
AfterSchool	Does the respondent participate in any after school activities? = 1 if yes
Sport	sports = 1 if yes
Lan	languages := 1 if yes
Art	arts := 1 if yes
Course	course-work related =1 if yes
charity	charity := 1 if yes
otherAS	other := 1 if yes
otherAStype	if other, what kind?
ASSchool	the activities are organised only by the school := 1 if yes
ASBoth	the activities are organised by the school and private bodies := 1 if yes
Ambitions	
Uni	planning to attend university :=1 if yes
Emigration	
EmigrPlan	planning emigration :=1 if yes

Continued on next page

Table 1.A.5 – continued from previous page

Variable	Description
Friends	have family or friends abroad :=1 if yes
emigr	someone in the closest family emigrated :=1 if yes
emigrMum	mother emigrated :=1 if yes
emigrDad	father emigrated :=1 if yes
emigrSib	sibling emigrated :=1 if yes
emigrwhereM	destination of mother's emigration
emigrwhereD	destination of father's emigration
emigrwhereS	destination of sibling's emigration
Network	help of family already abroad in emigration :=1 if yes
Mother's emigration	
temigrm	emigration during term time: =1 if abroad at time t (t=6)
semigrm	emigration in summer: =1 if abroad during summer s (s=2010,2011)
pemigrm	permanent emigration: =1 if yes
Father's emigration	
temigrd	emigration during term time: =1 if abroad at time t (t=6)
semigrd	emigration in summer: =1 if abroad during summer s (s=2010,2011)
pemigrd	permanent emigration: =1 if yes
Sibling's emigration	
temigrs	emigration during term time: =1 if abroad at time t (t=6)
semigrs	emigration in summer: =1 if abroad during summer s (s=2010,2011)
pemigrs	permanent emigration: =1 if yes
Frequency of visits:	
seemum	frequency of seeing mother: =1 once a month or more often, =2 once every 6 months, =3 once a year, =4 more rarely than once a year
seedad	frequency of seeing father: =1 once a month or more often, =2 once every 6 months, =3 once a year, =4 more rarely than once a year
abroad	visited a parent abroad := 1 if yes
Subjective perception:	
respon	more responsibilities due to migration : =1 if yes
school	impact of emigration on grades =1 if yes
schoolHow	what impact on grades : = 1 if negative
Performance:	
av	average grade of a respondent in period t (t=6)
behav	behavioural grade of a respondent in period t (t=6)

Continued on next page

Table 1.A.5 – continued from previous page

Variable	Description
School attendance:	
attend	number of teaching hours missed not excused by parent or doctor in period t (t=6)
attendex	number of teaching hours missed excused by a parent or doctor in period t (t=6)
Control information:	
fullanswer	1 if pupil filled the questionnaire fully
School information:	
avhum	school's average result of the final humanities exam in 2012 (in %)
avmath	school's average result of the final maths and sciences exam in 2012 (in %)
avlan	school's average result of the final languages exam in 2012 (in %)
Derived variables	
Class ratio	proportion of PWA children in the class at time t
Class sex	gender composition of the class
Class sibling	average number of siblings pupils in the class have
Class employment	fraction of mums or dads in employment at time t
Class parental education	average parental education level (coded as separate variables)
Class parental age	average age of parents (coded as two separate variables)
County ratio	proportion of PWA children in classes in the county at time t

Chapter 2

Out of sight, out of mind? Educational outcomes of children with parents working abroad

2.1 Introduction

The recent enlargements of the European Union resulted in new migration trends. An increasing number of households decide to send a member abroad, leading to family separation. The Polish Ministry of Education reports that 20% of Polish educational institutions surveyed in 2010 had pupils for whom one or both parents emigrated abroad.[1] In this chapter I analyse the impact of parental emigration on educational attainment of Polish children whose parents work abroad (henceforth PWA children).[2]

Large scale parental emigration raises questions about the impact family separation may have on children. There are concerns for children's immediate welfare[3] as well as long term socio-economic implications. In light of the theoretical literature to date, however, it is ambiguous whether the impacts of parental employment abroad are negative or not.[4]

These considerations are crucial because human capital acquisition early in life depends largely on parental decisions and is vital for short and long-term outcomes of individuals. It also plays an important role in economic development.[5] One's skills are shaped by both nature and nurture. They depend on the initial level of human capital as well as investments made, and these two elements complement each other. For most of childhood, parents decide which investments to make in children. For example, they may spend quality time with children or invest money in their education.

In that sense also, the emigration decision may have a bearing on a child's development. Emigration leads to family separation and less quality time with the migrant parent. Children may also be given greater household responsibilities if a parent emigrates. Moreover, family member emigration may change the perception of returns to

[1] See Tynelski (2010).

[2] Children whose parents left for employment abroad have been called in the literature *the left behind children*. I refrain from using the phrase in this study as I do not perceive children of temporary migrants, who stay with the other parent in the home country and see the migrant parent on the regular basis as left behind.

[3] For example, in a policy report Tynelski (2010) expresses the worry that children whose parents work abroad feel abandoned and lonely. He emphasises that they may struggle with their identity, definition of priorities and with their educational responsibilities, which may influence school attendance and overall performance.

[4] See Chen et al. (2009); Dustmann and Glitz (2011); Amuedo-Dorantes et al. (2008).

[5] See Apps et al. (2012); Aizer and Cunha (2012); Behrman et al. (2006); Feinstein (2003); Barro (2001).

education, depending on the demand for labour in the destination countries.[6] At the same time, migration usually results in an increase of household income[7], which may benefit children.

Since there are forces acting in opposite directions, the question of the relationship between parental migration and children's schooling is an empirical one. Moreover, the theory is silent on potential heterogeneous impacts, depending on the family background and nature of migration. The literature to date has provided mixed results and has not always dealt with the key identification issues. This is mostly due to data limitations; it is difficult to obtain data matching educational performance of children and the emigration situation in the family. Therefore, in 2012 I created and collected a data set for this purpose (See Migration and Education of Children in Poland 2012 data in Chapter 1). In particular, I obtained detailed information about migration experiences in the families and their timing, family background and school progress of pupils measured by grades, rather than drop-out rates. As becomes clear in the analysis, the timing of emigration and precise measures of educational attainment are key for establishing the relationship; so far they have been rarely used in the literature due to lack of such data.

I analyse how three outcomes - the average grade, the behavioural grade and number of school hours missed by a respondent - are impacted by parental employment abroad at a given point in time.

The ordinary least squares regression results indicate a negative, significant in size, relationship between parental emigration and a pupil's grade. They reflect the fact that PWA children perform on average worse at school, irrespective of the emigration decision of their parents.

This initial approach does not account, however, for the fact that migration decisions may be endogenous. There may be unobserved characteristics of migrant-sending families which simultaneously influence the decision to emigrate and the child's school performance. They may confound the estimates of the true effect. To resolve the problem, I employ individual fixed effects approach. It accounts for any time-invariant unobserved differences between respondents.

I find a positive, statistically significant, but very small immediate impact of parental emigration on the educational attainment of children. It suggests that, when a parent is abroad, the grade of a child increases by up to 5% of a standard deviation. There is no effect on the behavioural grade and school attendance. One potential explanation for such a result is that the gain from the increased household income following migration dominates any negative effects of family separation, which are lower than in previously studied cases because of the nature of migration in question.

Parental education appears pivotal. PWA children of high school graduates gain most, relative to their non-PWA peers whose parents have equivalent educational at-

[6]See de Brauw and Giles (2006); Kandel and Kao (2001); Chand and Clemens (2008).
[7]See Antman (2012).

tainment. Moreover, PWA children whose parents have lower than secondary education (67% of the overall group) do not perform better, on average, than their peers from a similar background. I suggest that more educated parents are more likely to succeed abroad by securing better employment and assimilating to a greater extent. As a result, they may have higher incomes and life satisfaction. This, in turn, is likely to influence the family life and may be reflected in a child's school performance. Moreover, a child's education may be of greater importance to more educated parents. Then they may allocate a higher share of the household income to children's schooling and be more personally involved in their child's education.

I allow for a delayed response to emigration by including lags of the parental migration status in the regression. I find that the grade is negatively affected after 2 semesters from parental departure. Since emigration in the sample is temporary in nature, I consider various migration patterns. If the emigration episode is short-lived, the size of the impact is not striking. A departure for 2 or more semesters, however, has the potential of significantly lowering a child's grades. It is possible that family separation becomes more burdensome and the income flow falls the longer a parent is abroad. The negative impact gradually disappears following a parent's return.

I extend the analysis to sibling emigration and find large, positive, significant and persistent effects on the attainment of pupils. This is only true for those siblings, whose first migratory experience occurred within the observed 3-year-period. The same cannot be said about siblings who have migrated recently but have also been living abroad prior to September 2009, when the survey began. The positive impact may be related to income effects as well as a change in perceived returns to education. Siblings with longer migratory experiences are likely to be older and have their own families. Hence, they may remit less and exert lesser influence on younger relatives.

The analysis is not without limitations. The approach does not cater for situations in which time-varying changes, affecting both the school outcomes of pupils and the migration decisions of parents, take place. I discuss various such potential limitations later. One possible scenario arises if a teacher's evaluation of pupils is inconsistent over time and the resultant changes in pupils' grades coincide with the migratory spells in the data. An economic shock to the region, which affects the staffing of schools (and hence potentially grades) as well as migratory movements in the area is another example.

My findings may be limited in scope as I cannot provide detailed insight into the mechanisms behind the effects I find. Nonetheless, this chapter contributes to empirical migration research in several ways. The results of empirical studies to date are mixed and often difficult to reconcile (See Antman (2013) for an overview). My work highlights the importance of using adequately defined variables to accurately measure the effects of parental emigration.

Firstly, the source countries which emerged as a result of the EU enlargements bear little resemblance to the traditional emigration states like Mexico or Philippines.

The main differences lie in economic performance, culture, tradition and their history. The focus on a new EU member state provides an opportunity to reevaluate claims in the literature and to investigate whether the divergence between source countries leads to differential outcomes. Although some analyses for Eastern European countries are available,[8] they often focus on other indicators of children's well-being and are less flexible in terms of the analysed migration patterns.

The migratory movements captured in the sample differ largely from those studied before; they are usually temporary, legal, circular and characterised by one family member working abroad, whilst others stay in the home country. For that reason PWA children are less burdened by parental departure and may still gain from the increased income.

Parental emigration in middle-income countries is unlikely to lead to school drop-outs,[9] which is how educational attainment is usually captured in empirical studies. Rather, it impacts school grades, school attendance and children's behaviour. Therefore, I look at the influence of emigration on these outcomes, which may be more informative about the exact mechanisms behind changes in children's performance.

Most importantly, the analysis reveals the complexity of the effect migration can have on children. The impacts I find depend on the socio-economic background of the family, as well as on the timing and duration of migratory movements. In particular the analysis of timing of migration provides new insights into the changes which occur as a result of parental migration.

The choice of family member to emigrate is also crucial; parental emigration may not benefit children, but foreign experiences of siblings may be favourable.

In Section 2.2 I provide a brief overview of the literature. Section 2.3 describes the data. Section 2.4 outlines the methodology employed and related concerns one may have. Results are presented in Section 2.5 and an extension to sibling emigration in Section 2.6. Section 2.7 concludes.

2.2 Literature

The research presented in this chapter is motivated by and draws upon different strands of economic literature, applying general theory to a migration context.

Firstly, the economic literature of education has been recently dominated with the discussion of the role of human capital in one's individual outcomes (Cunha and Heckman, 2007; Postlewaite and Silverman, 2006; Agan, 2011) as well as in the economic development of countries (Barro, 2001).

It is recognised that both cognitive and non-cognitive abilities of individuals determine their success in the labour market. It is also believed that intra-family environ-

[8] See Botezat and Pfeiffer (2014); Gassmann et al. (2013); Giannelli and Mangiavacchi (2010).

[9] This certainly is a more likely scenario in Poland, where 95.1% of 16-18 year-olds have been reported as attending educational institutions in 2011 (The Central Statistical Office of Poland, 2012).

ment plays a key role in this context and that both nature and nurture contribute to human capital development.

Therefore, apart from investigating at which stage of life and how these skills are acquired, economists became interested in knowing what would happen if the process of their acquisition was disrupted. Do changes in family environment, such as divorce or emigration of a parent, have a bearing on one's skills?

The ideas have been tested in the context of single-parenthood (divorce or death of a parent) and family separations due to occupational commitments (Lyle, 2006). It has been argued that children from one parent families are likely to perform worse academically, but the negative effect weakens with the age of the child at the time of marital disruption. Steele et al. (2009) claim that the type of family disruption is irrelevant for the overall negative outcome. The effects are contingent on gender of a child (Zaslow, 1988), its age (Reneflot, 2007), race and ethnicity (Sun and Li, 2007). Resource deprivation and limited, often diminishing, contact with the parent are the reasons behind the negative impact of family disruptions. The literature is relevant to migration context but migration, although leading to family disruption, has a different, less drastic effect on a child's well-being, predominantly because it is expected to result in an increase in household income.

Until now most migration research has focused on the traditional sending states, like Mexico, or countries with a large scale internal migration, like China. The results of the studies are mixed and inconclusive, but the perception of negative impacts of emigration dominates. For example McKenzie and Rapoport (2011) find in their paper that migration lowered schooling for 16-18 year old boys in rural Mexico and argue that it may impart a disincentive effect on children due to increased burden of housework. Positive or no impacts of emigration are found in Hanson and Woodruff (2003) and Chen et al. (2009). Antman (2013) provides a comprehensive overview of the literature to date.

Recent years have seen emergence of studies for Eastern European countries, following the growing incidence of leaving children in the home country by migrant parents. The evidence is mixed and scarce, however, partly due to lack of data, different research methods and economic heterogeneity among the countries.

Botezat and Pfeiffer (2014) discuss the scale of family separations due to migration in Romania and, using instrumental variable regressions, provide evidence that parental emigration has a positive impact on school grades of children; children with migrant parents seem to spend more time studying than their peers. Migration has, however, negative implications for their health and emotional well-being.

The nature of migration in the study is similar to that observed in the Polish data. However, Botezat and Pfeiffer (2014) consider only cases where migrants have been abroad for at least 12 consecutive months. Migration spells are often shorter, which is why I allow for more flexibility in my approach.

Gassmann et al. (2013), using data for Moldova, find that migration is not associated

with negative outcomes on children's well-being. They focus, however, on an overall index of well-being, rather than educational attainment.

Lastly, Giannelli and Mangiavacchi (2010) consider schooling of children in Albania and argue that father's emigration increases probability of dropping out of school or delaying school progression. They analyse school attendance, rather than grades, and rely on a duration model to do so.

Despite looking at different states and lacking comprehensive data, all studies approximate that 19-22% of children in Albania, Moldova and Romania have at least one parent abroad.

2.3 Data and descriptive statistics

2.3.1 Overview of the sample

The data set used for the analysis is described in detail in Chapter 1 of the thesis. It contains information about 2822 Polish pupils from one cohort, followed over a period of six semesters between September 2009 and June 2012, when they were aged 14 to 16. It includes individual and household characteristics, some socio-economic indicators, school performance and migration experience within the family.

Pupils are identified as children with parents working abroad (PWA) if they had at least one parent abroad at any point during the observed period. All respondents provided information about migration experience within the household but only 2669 gave a detailed account of its timing and were included in the analysis.[10]

As observed in Chapter 1, 18% of respondents indicated having a parent abroad at some point during the 3 year period. Almost 26% of respondents indicated that their sibling has emigrated abroad. The migratory spells in the data are relatively short-term, repeated and occur over rather short distances. Households usually send just one parent abroad, in most cases the father.

Migrants from households in Opolskie are low skilled with 44% of mothers and 63% of fathers having finished vocational schooling and 36% of mothers and 29% of fathers high school.

Variable definitions

Children with parents working abroad (PWA)

I define **a PWA child** as a child who has had at least one parent abroad in a given semester and stayed in the home country during parental emigration experience. Given such definition, one may have one or both parents abroad at the same time; moreover, a migrant parent may be absent in one semester and return to Poland in another and

[10]The exclusion of some observations from the data set may influence the outcomes of analysis. In particular, it may change the educational profile of the class and result in underestimation of the class size, as well as number of PWA pupils in the class. I run regressions including all observations and find that results do not differ.

this change will be reflected in a change in the PWA child status.

School performance variables
The main dependent variable is **the grade of a pupil**. The grade is taken as an average over all courses taken in a given semester and ranges from 1 to 6, with 6 being a top mark awarded to a pupil for extracurricular achievement in the subject area. Pupils who mastered 100% of the curriculum in a given semester are usually awarded 5; 1 is a fail mark. The grade is awarded internally but based on the requirements of the national curriculum for a given year. The average grade in the sample has a mean of 3.61 and a standard deviation of .851. As can be seen in Table 2.1, PWA children obtain lower grades on average.

Test scores in the national exam respondents took in the final semester of *gimnazjum* are another measure of academic performance. Prior to completion of *gimnazjum* and progression to the next education stage, pupils are tested in the following areas: Polish language and literature, history, maths, science and foreign languages. The exams are organised nation-wide by one Exam Board and blind-graded in percentage terms. Unfortunately, I only possess information about the exam results for under 13% of the sample, which is insufficient to use for the analysis. I will, however, rely on it for some background checks.

Schools also issue a grade for the overall **behaviour of each pupil** every semester. The grade ranges from 1 to 6, with 6 being awarded for extraordinarily good behaviour, including involvement in charity work, etc. The behavioural grade has a mean of 4.489 and a standard deviation of 1.241, which is significantly larger than the standard deviation of the average grade. A PWA child has on average .256 lower behavioural grade.

Attendance is also recorded with the number of hours at school a pupil missed in a given semester, as well as the number of hours missed but excused due to illness or any other justified reasons. In the analysis I will focus on the number of hours missed by a pupil at school but not excused, as they most likely are indicative of problems with a pupil. Pupils with a parent abroad on average miss 5.5 more hours of school in a semester.

Table 2.1: Some summary statistics

Panel A: Pupils' performance at school

	PWA children				non-PWA children			
	mean	st.dev.	min	max	mean	st.dev.	min	max
child's average grade	3.489	0.829	1.4	5.5	3.649	0.853	1	5.88
child's behavioural grade	4.308	1.319	1	6	4.564	1.197	1	6
number of hours missed at school	15.897	32.346			10.370	27.408		

Panel B: National exam results

(Note: this information refers only to a subsample of pupils)

	mean	st.dev.	min	max	N
literature test score	62.232	18.250	19	100	334
history test score	59.023	18.103	15	100	334
math test score	45.469	23.449	7	100	334
science test score	48.264	15.442	15	100	334
language test score	54.279	24.965	11.5	100	334

Panel C: Parental education levels

(Note: not all respondents provided this information)

	migrant families		non-migrant families	
Mother's education	N	% of group	N	% of group
primary	16	5.71	229	10.02
vocational	122	43.57	793	34.7
secondary	101	36.07	769	33.65
tertiary	41	14.64	494	21.62
Mother works	189	69.23	1542	72.36
Father's education	N	% of group	N	% of group
primary	11	4.1	213	9.7
vocational	168	62.69	1031	46.27
secondary	78	29.1	644	29.31
tertiary	11	4.1	309	14.06
Father works	241	91.98	1855	90.53

Source: MECP2012

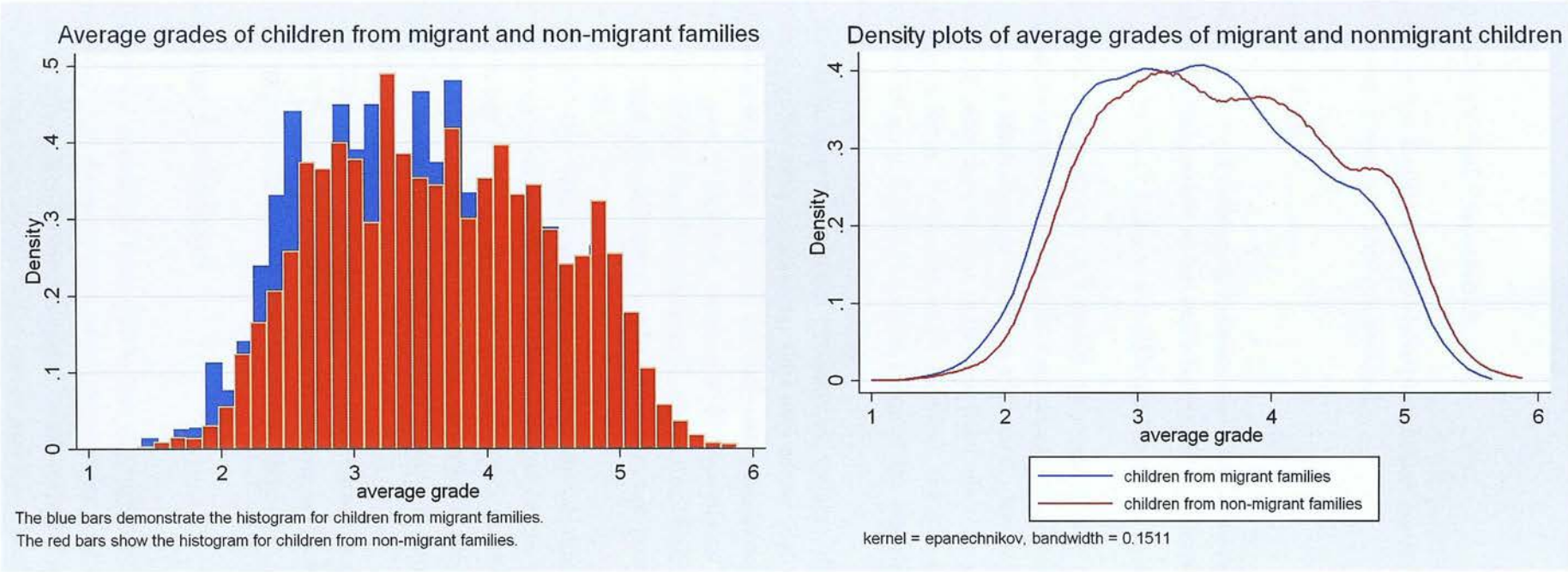


Figure 2.1: Average grades of respondents from migrant and non-migrant families
Source: MECP2012

2.4 Empirical framework

In this section I outline the preferred estimation equation, discuss problems related to the approach and how they are tackled.

Specification

I investigate the relationship between one's individual school performance and the experience of emigration within one's family. The preferred specification is the following:

$$Y_{it} = \alpha + \beta EmigrParent_{it} + \gamma_i + \theta_t + \epsilon_{it} \tag{2.1}$$

where Y_{it} is a performance variable investigated (the average grade, the behavioural grade or the number of hours missed at school) for individual i in semester t, $EmigrParent_{it}$ is a dummy variable equal to 1 if an individual i has at least one parent abroad at time t and zero otherwise, γ_i is an individual fixed effect, θ_t are semester fixed effects and ϵ_{it} is the error term. Unless otherwise specified, the standard errors in the regression are clustered at an individual level, as I expect the individual outcomes to be correlated over time.

The definition of the main explanatory variable, $EmigrParent_{it}$, is to an extent dictated by the data constraints. One may argue that it would be optimal to use two emigration dummies, allowing for differentiation between having one or two parents abroad at time t. However, since only 40 respondents had two parents abroad at the same time during the observed period, separating those with one or two parents abroad leads to imprecise estimates in the regression, not providing any further insights into the analysis.

Equally, one could separate the emigration variable to account for the role of gender of the emigrant parent in the overall impact on the child's performance (See Cortes (forthcoming)). Also in this case the coefficients on maternal emigration become statistically insignificant, since not many mothers in the sample engage in employment abroad. Given the data at hand, there is a trade off between exploring the relationship in more detail and the estimation precision.

The parameter of interest is β. It explains how the school performance of pupil i in semester t changes when the emigration status in the family changes, i.e. at least one parent emigrates or returns from abroad at time t.

Individual fixed effects

I include the individual fixed effects into the regression to control for any unobserved individual level characteristics *which do not vary over time.* This will isolate any confounding effect these factors may have on the estimate of interest, if they are correlated with the emigration status in the family and the school performance of children.

Many characteristics which influence children's performance at school are also correlated with migration decisions of parents. Parental education or socio-economic characteristics of the household are an example; from the summary statistics it is clear that low-skilled parents engage in temporary employment abroad more often than parents with higher qualifications. At the same time, one may argue that children's school performance is likely to be correlated with educational attainment of their parents. Hence, children of low-skilled parents are likely to perform worse at school and to have a parent abroad. If no individual fixed effects are included in the regression, the regression coefficient of interest, β, will capture the impact of the individual characteristics, as well as of emigration experience.

The fixed effects approach will also eliminate the risk of reverse causality in β. Arguably, the educational attainment of children may cause the migration event, rather than the other way round. This is, however, unlikely in the Polish situation. Based on the results of the survey, the general perception in respondent schools is that parents often do not appreciate the potential impacts emigration may have on their children and that their decision is primarily driven by income considerations. I check for reverse causality by including leads of the emigration variable into regression and find no evidence of the problem. Nonetheless, exploring the panel dimension of the data and allowing for identification to be made upon a change in the emigration status, resolves the potential problem.

Semester fixed effects

The material studied at school changes and becomes more difficult with time. Since the pupils' performance is tracked over a 3 year period, one may notice a change in pupils' grades which is attributable to the advancement in their studies and not to other circumstances.

There is, in fact, a clear pattern to the average grade over time in the sample (see Appendix 2.B). Each year there is a systematic improvement in pupils' grades in the second semester, when compared with the first semester of that year. Further, the gap in grades between first and second semester in each year widens further into *gimnazjum*. It may confound the estimate of β. Hence, I include semester fixed effects to isolate the changes in grades over time which are common to all.

Further concerns

Are the average grade and behavioural grade good measures of performance?
Grades are awarded internally. The assessment of pupils against the national curriculum is at the teachers' discretion. Hence, grades may be subjective; pupils may be awarded different grades for comparable performance by different teachers and be assessed relative to their classmates.

However, I expect that teachers are consistent in the way they assess pupils over time. If so, then any differences in average grades would be teacher-specific and time-invariant, and will be isolated by use of class or individual fixed effects.

Moreover, the identification of the estimate is based on a change in one's individual grade, rather than its absolute value.

One can still argue that some teachers are inconsistent in their assessment over time. One such case may be when inexperienced teachers learn over time and adjust their assessment of pupils accordingly. It is difficult to predict whether such a behaviour would result in an improvement or deterioration of pupils' grades.

This would be problematic if the changes in grades driven by the teacher's learning process coincided with parental migration and occurred on a significant scale. I find the scenario unlikely for the following reason: if teachers were adjusting grades as they learn, the changes should be gradual and occurring in the same direction (i.e. improvement or worsening of grades) until they reach a point at which the assessment of pupils is deemed adequate. On the other hand, migratory movements in the data are circular, short term and vary in timing. It is therefore unlikely that the two patterns consistently coincide to explain the results presented in this chapter.

Alternatively, teachers may become lenient towards a PWA child or provide more support for the child upon learning that his parent has emigrated. Then the improvement in the pupil's grades may indeed coincide with parental departure. It is difficult to rule out such a scenario here.

Nonetheless, I run a check to ensure that the average grade is a satisfactory measure of school performance in this case. I am in possession of the individual scores from the national tests respondents took at the end of *gimnazjum* for almost 13% of the sample. Even though the number of observations is insufficient to obtain robust results and I only observe a test score at one point in time, which means I am unable to run a regression with individual fixed effects, I rerun a regression of an average test score on the emigration experience within a pupil's family. The results, although statistically insignificant, are in line with the output presented in the next section. Details can be found in Appendix 2.B.2.

The behavioural grade is even more subjective in nature and based on more general guidance related to behaviour expected of pupils in *gimnazjum*. Unfortunately, there is no other measure which could be used to check whether the relationship I find would hold if pupil's behaviour was assessed externally.

Other time-varying changes

I already provided a couple of examples of time-varying changes which may be correlated with the school performance of children and migration decisions of parents (and hence bias the estimates) when discussing the assessment by teachers. Here I consider further cases.

For example, an economic shock to the region may influence both migration or

return decisions. At the same time it could trigger a change in the availability and composition of teachers in the region, which may have a direct impact on grades. However, there is no indication that the region was either severely or positively economically affected in the observed period of 2009-2012. Moreover, changes in economic conditions of destination countries did not discourage emigration from Opolskie in this period.[11]

Alternatively, a change in a child's performance following parental emigration may prompt parental decision to return. To check whether this appears to be a viable threat to estimation results in this context, I run regressions including leads of the emigration variable. The results suggest that the future emigration situation in the family does not predict current school performance (See Table 2.B.1 in Appendix 2.B).

Endogeneity of emigration

Households select into emigration (Gibson et al., 2010). More specifically, the following elements of the emigration decision may be endogenous:

1. Households decide to engage in emigration.
2. Households also decide whether one family member or all should emigrate.
3. Some emigrants decide to return from emigration, whilst others stay abroad permanently.
4. Emigrants also decide on duration of their emigration experience.

I will now briefly elaborate on these aspects of emigration decision and discuss the extent to which they may constitute a threat to validity of the results.

Firstly, the decision to send a family member abroad may be correlated with certain characteristics, such as the socio-economics discussed above, which also influence a child's performance at school. Provided these traits do not vary over time, they will be isolated by the fixed effects approach. I expect the bias due to the fixed differences between the migrants and non-migrants to be negative. This is because families with a lower socio-economic background (which is associated with worse school performance) are more likely to engage in migration.

Of greater concern is selection into migration which is correlated with time-varying characteristics key for school performance. It will not be isolated by the fixed effects approach. To provide an example, assume that parental emigration has no effect on a child's performance. Now take two identical families without migrants and consider a random shock, such that the father in one of the families loses a job. Suppose that the job loss has a short-term immediate negative impact on the child's grade and triggers migration decision of the parent, but only after an unsuccessful job search at home. If the child's grades recover after one period and this improvement coincides in time with

[11] In Appendix 2.B.1 I provide some statistics on the local economy in the observed period.

(but is not caused by) the departure of the parent, then the regression estimates will be positively biased.

Alternatively, one may observe a fall in a child's school performance which coincides with a parental departure and is driven by a third factor. Once again, imagine that there are two identical families with no migrants and migration as such has no impact on school performance. Then suppose that one of the families experiences a negative, unexpected shock which has a negative effect on a child's schooling and drives a parent abroad. One example may be an unexpected illness in the family or a sudden increase in a number of dependants in the household (e.g. an elderly family member moving in). The situation may have a negative effect on a child, e.g. due to sudden crowding in the house or greater household responsibilities. It is likely to also impose additional financial burden on the family, resulting in migration of one of the parents. In this case the regression estimates would be negatively biased if the fall in the child's grades occurred at the same time as parental migration, making it look like parental departure triggered the fall in performance.

I will later consider alternative estimators that impose different assumptions on the dynamic responses.

A further source of selectivity is the decision of a household whether all or only some family members should emigrate. Naturally, when entire families emigrate, they are not captured in the data and the approach essentially compares households which never had emigrants with those who sent only some family members abroad.

If the households who emigrate with children differ from those who leave children behind, e.g. are wealthier, and these differences affect educational attainment, then the estimates will be biased, as the comparison will be made only between a selected group of migrants and non-migrant families (Steinmayr, 2013; Gibson et al., 2010; McKenzie and Rapoport, 2007). Although I cannot control for this type of selection in the sample, I argue that it is unlikely or the scale of the problem is rather small in my sample.

Looking at Table 2.A.2 in Appendix 2.A, it is clear that, although many students disappear at some point from the class register, this is due to failing to pass the year, a change of class or change of school. Only 67 students disappear from the register for unknown reasons; even then it is unlikely all of them leave the country.

To further infer what percentage of these pupils might have left for abroad, I refer to regional deregistration statistics. When an entire household leaves the region, they should deregister from the address at which they were residing in Poland.[12] According to the register only 187 pupils born in 1996 (age cohort surveyed) left Opolskie to live abroad between January 2002[13] and December 2011. Given the cohort size of 9 500, these flows are very small.

[12]The records are very accurate for internal migration. There is a degree of uncertainty about its precision in case of international migration. The numbers captured in these statistics are likely underestimating the scale of the phenomenon (The Central Statistical Office of Poland, 2009).

[13]when the children were not yet of school age

Thirdly, selectivity may also be driven by the decision of some migrants to return. Gibson et al. (2010) argue that this form of selectivity is only challenging if the return migrants are wrongly classified in the survey as *never migrants*. This is because they would differ from never migrants but would be grouped with them in comparisons of migrants and non-migrants. This should not pose a problem as I allow for returns from migration and ask about migration experiences pre-2009.

Moreover, I do not find much evidence of returns of children from emigration in the data. Only 106 new pupils joined the cohort in participant schools throughout the three observed years; as before, it is unlikely that all of them arrived from abroad, but I cannot specify precisely what their past experiences were. I also do not know whether children, who I observe in the data throughout the entire 3 year period, have emigrated with the family and then returned prior to joining the school. However, looking at the data on registration at a local address, I find that 28 children of this age group arrived in Opolskie from abroad between January 2002 and December 2011. I cannot distinguish between Polish and foreign children. Nonetheless, this is a very small group relative to the sample and cohort size.

Gibson et al. (2010) also point out that another form of selectivity is visible in the return migrant's decision regarding the duration of the stay abroad. I can control for that (at least in the observed period) thanks to the precise information on the timing of migration and duration of migratory spells.

All these elements of migration decision may be problematic, despite the use of fixed effects, only if they depend on time-varying characteristics that are also key for school performance of children.

Non-respondents and dropouts

One further complication I cannot control for is posed by the fact that some pupils have dropped out of the class at some point over the observed period and before the survey took place. Others, on the other hand, have not responded to the survey. As a result they were not included in the sample.

If those who have a parent abroad were more likely to drop out, then the fact that they are omitted from the sample may introduce bias into the estimation.

If many PWA students do not progress to the next level at school and it is due to their parents' emigration, my analysis may underestimate potential negative impacts of emigration by not considering class failure in the regression and focusing on grades, conditional on having progressed to the final year of school. To consider how problematic this concern is, I would like to know the proportion of students born in 1996 (respondents' cohort) who failed at least one year (hence are not represented in the sample) and come from migrant families.

I do not possess the information; however, I observe students born in 1995 and earlier who have repeated at least one year at school. On the basis of this group, I

make inferences about the potential situation among pupils born in 1996. As can be seen in Table 2.A.1 in Appendix 2.A, I find that there are 94 pupils in the sample (3%) who have repeated a year at school and, among those, only 17% declared having had a migrant parent in the family. Hence, PWA students do not appear to dominate the group of under-performers.

A similar problem arises if the non-response to the survey is not random. In particular, one may be concerned that due to the sensitive nature of the survey PWA pupils were more likely to refuse participation in the study. If so, they may be overrepresented in the group of non-respondents. Unfortunately, I have no way of checking if this was the case.

As mentioned in the introductory data chapter, non-respondents performed worse on average relative to the respondents. If the worse performance of this group is due to the increased presence of PWA children and was driven by parental migration rather than other factors, then the results I present may be upward biased as they do not account for the outcomes of this group.

I cannot, however, establish if parental migration is the driver of the lower grades among the non-respondents. Given that migrant parents are negatively selected, any worse performance of potential PWA non-respondents could be due to selection as well as other factors.

Alternative specifications

One alternative regression specification may involve using various individual level controls instead of individual fixed effects. However, given the limitations of the data, many of the characteristics crucial for educational attainment of children and emigration within the family are not available.

Therefore, the estimates of β in such a case would be most likely biased. In particular, I expect the coefficient to be biased downwards; since low-skilled parents are more likely to emigrate and children's outcomes are correlated with those of their parents, it is likely that poor school performance is correlated with, but not caused by, emigration experience of a parent.

Nonetheless, for illustration purposes I include results of regressions without individual level fixed effects and with some individual level controls, such as parental education level, employment and number of siblings, in the results table. To isolate as many characteristics which may confound the estimate of interest as possible, I also include school or class fixed effects to control for the role the learning environment may play in performance, and the semester fixed effects, for reasons explained before.

Given the concern that some of the omitted variables key for the analysed relationship may vary over time I also considered a lagged dependent variable regression setup. This is because school performance of children is likely to be driven by its historic values as students' performance is correlated over time. The past performance

variable may be capturing some environmental and individual characteristics crucial for the future outcomes of pupils as well as migration decision of parents. A specification, including lagged school performance, provides therefore an alternative to my estimation approach. It is particularly useful when considering threats to validity of estimation results because LDV and FE models have a bracketing property which may be informative of the true relationship being analysed (Angrist and Pischke, 2009).

The detailed results of the analysis can be found in Appendix 2.B.3. The bounding exercise based on comparisons of the estimates in the FE and LDV regressions leads to two conclusions. Firstly, one can rule out existence of large negative effects of parental emigration in the contemporaneous case. Secondly, there is no large positive long run effect of having a parent abroad.

2.5 Influence of parental emigration

2.5.1 Immediate impacts

I look at the instantaneous relationship between average grades, behavioural grades and school attendance of respondents and emigration by one or both their parents, as described by Equation 2.1. The regression results are presented in Table 2.1.

For the average grade, the regressions without individual level fixed effects (columns (1)-(3)) produce negative, statistically significant coefficients on emigration, varying between -.119 and -.091. The results suggest that having a parent abroad can lower pupil's current average grade by up to 14% of a standard deviation, which reflects the findings in the summary statistics of worse average performance of children with working parents abroad. However, the estimates of β may be biased due to unobserved time-invariant differences between individuals, which impact the average grade and are correlated with the family's migration decision. In particular, the PWA children are expected to perform worse on average, irrespective of the emigration decision of their parents, given the socio-economic characteristics of the families they come from.

The individual fixed effects regression estimates in columns (4)-(5) of Table 2.1 imply a small, positive impact, however; parental absence in semester t increases the average grade by .024-.045, which is equivalent with 2.8-5% of its standard deviation. The coefficient is statistically insignificant when semester fixed effects are included in the regression, though. This may be due to clustering of emigration over time and a degree of confounding of time and migration effects.[14]

[14]Inclusion of time fixed effects or difference-in-difference approaches may not be the best methods to analyse the data at hand as they do not allow great flexibility. The observed migration behaviour in the surveyed group is very general and not subject to strict restrictions. Emigration can occur at any point in the observed period and I allow for returns, departures, as well as circular migratory patterns. There are also no restrictions on duration.
Looking at the summary statistics, I conclude the following: (i) one should consider years of education separately but allow for correlation across semesters, as semester 2 grade in a given year takes into account pupil's performance in previous semester, (ii) there is a systematic improvement in semester 2 grades in every year and mixed performance across years for all students, (iii) looking at differences

The OLS results for behaviour (columns (6)-(7)) are comparable to those for average grades, indicating a negative impact equivalent to 12-14% of standard deviation of behavioural grade. However, the estimates with individual level fixed effects suggest that behaviour remains unaffected.

I also consider the numbers of hours missed (and not excused) by a pupil as another indicator of the overall school performance. Here, all regression coefficients are of comparable scale and identical sign, although the estimates from regressions with individual fixed effects are not statistically significant. They indicate that a child misses on average 2 to 4 more hours of schooling a semester following parental departure, which is equivalent to about half a day at school.

The positive or almost no influence of parental emigration on the average grade and behaviour may be surprising, especially given the general perception that emigration imposes a burden on young people. It is often suggested that separation from parents badly influences behaviour and academic performance of children (Tynelski, 2010). Nonetheless, I argue below that such an outcome is plausible.

The literature on *left behind children* and their school performance outlines the following likely mechanisms at play when a child is left in the home country during parental employment abroad:[15]

1. A positive income effect - emigration is usually motivated by income considerations and, upon parental migration, the financial situation of the household improves. If the budget constraint is relaxed, part of the increased income may be directly or indirectly invested in a child's education.
   For Polish migrants emigration can potentially lead to a three- or fourfold increase in earnings, depending on the employment abroad (See Table 2.3). As discussed in Chapter 1, there is evidence suggesting that temporary migrants from Opolskie remit an overwhelming proportion of their earnings.

2. A negative impact of family separation - parental departure imposes a psychological burden on a child and may also change the expectations of a family towards the child. Children whose parents are employed abroad frequently have to take over many household responsibilities at the cost of time spent in education. Additionally, parental inputs in the child's upbringing are likely to fall.
   This mechanism is unlikely to play such a detrimental role here. Given the circumstances of migration in the sample, if only one parent emigrates and the other

between semesters within each year, in Year 1 I observe a greater improvement in grades of PWA children but worsening of behaviour and school attendance relative to non-PWA children, (iv) in Years 2 and 3 the improvement in grades is smaller for PWA children. Hence, it seems that the impact depends to a large extent on when the parent was abroad and for how long.
Plotting the mean of migration at time t over time, I observe signs of clustering in migration. A lot of identification comes from the beginning of the observed period; there is a high ratio of returns in the first observed semester and a visible increase in departures from semester 4 onwards. Hence it is impossible to separate time effects fully.

[15]For a detailed discussion see Antman (2011b).

stays at home with the family, children may not be faced with any additional responsibilities as a result of migration. In fact, only 27% of PWA children indicated that their responsibilities increased as a result of parental departure. This is a likely case here as a lower percentage of mothers in migrant-sending families are employed, compared to non-migrant families (See Table 1.4). Further, they also maintain frequent contact with the migrant parent who engages in employment abroad only temporarily.

Additionally, one may argue that, if emigration is driven by the lack of employment in the home country, the family separation may not have such detrimental impacts, depending on the situation prior to a parent's departure. For instance, imagine a family where both or one parent is unemployed prior to emigration and the unemployment not only negatively affects the family finances, but also introduces tension into the household. Then one parent's employment abroad may be a better alternative, even if it leads to separation.

3. Depending on the destination country of migrants and their experiences of foreign cultures and labour markets, the perception of returns to education may change (Kandel and Kao, 2001).

The overall impact is difficult to predict theoretically as the interplay between these effects depends significantly on the context of migration and the target population studied. In this case I expect the income effect to be larger than the potential burden of separation. This is because the Polish parents (usually fathers) emigrate over short distances and short time periods, often return and have frequent contact with the family. At the same time many mothers in migrant-sending households stay at home, caring for children. These characteristics of migration minimise the effects of separation. Meantime, the income effect is potentially very substantial (See Jończy and Rokita-Poskart (2013) and Table 2.3). The positive effect found here confirms that.

Table 2.1: Contemporaneous impacts of parental emigration on children's outcomes

Dependent variable	Average grade					Behavioural grade			Hours missed not excused		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$EmigrParent_{it}$	-.119***	-.107***	-.091*	.045*	.024	-.168***	-.195***	-.002	2.772*	4.416**	2.701
	(.026)	(.026)	(.053)	(.024)	(.037)	(.044)	(.044)	(.061)	(1.457)	(1.750)	(5.298)
Individual level controls	no	yes	yes	no	no	no	yes	no	no	yes	no
Semester FE	no	no	yes	no	yes	no	no	no	no	no	no
Individual level FE	no	no	no	yes	yes	no	no	yes	no	no	yes
No of observations	13860	10859	10859	13860	13860	12348	9630	12348	6738	5446	6738
No of students	2669	2071	2071	2669	2669	2470	1920	2470	1604	1280	1604

Source: MECP2012

Individual level controls include gender, number of siblings, parents' age, employment and education level, whether parents divorced or one parent died, and school fixed effects.

Standard errors clustered at individual level in parentheses.

Statistical significance: *** 1 % ** 5% * 10%

2.5.2 Parental education level matters

Even though the migration captured in this study is predominantly low-skilled, migrant parents constitute a mixed group in terms of their educational attainment; among fathers 61% have vocational qualifications (below A-level equivalent), 31% completed secondary education and the remaining 8% either have tertiary or lower secondary education.

Parental education level, as well as the family socio-economic situation in general, are crucial for a child's educational attainment as human and cultural capital are transmitted across generations (Black et al., 2005; Black and Devereux, 2011). The usual expectation is that the children's school performance improves with parental education. It is also the case here: the higher the parents' education level, the higher the average grade and behavioural grade of the child.[16] See Appendix, Section 2.D for the results.

I have controlled for parental education as another factor influencing school performance, but migration experience of parents may impact children differently, depending on parental education level. This may be due to different earning potential, but also investments (time and financial) made in their children.

Given these considerations, I interact parental educational attainment dummies with the emigration status in the family to see whether differential effects emerge. I choose father's education as the indicator of parental education as fathers have a higher propensity to emigrate. Since parental education levels in the sample are highly correlated, I do not expect this decision to be crucial for the results. Moreover, I combine together the parents with elementary and vocational education into one category.
The main regression equation now becomes:

$$\begin{aligned} Y_{it} = \alpha + \beta \text{EmigrParent}_{it} + \lambda_1 \text{EmigrParent}_{it} \times FatherLow_i + \lambda_2 \text{EmigrParent}_{it} \\ \times FatherSecon_i + \lambda_3 \text{EmigrParent}_{it} \times FatherTert_i + \gamma_i + \theta_t + \epsilon_{it} \end{aligned} \tag{2.2}$$

where Y_{it} is the average grade, behavioural grade or numbers of hours missed by pupil i in semester t, $FatherLow_i$, $FatherSecon_i$ and $FatherTert_i$ are dummy variables equal to one if a father's highest educational attainment is below A-level, A-level equivalent or degree education, respectively. As before, EmigrParent_{it} is the emigration dummy variable, γ_i is the individual fixed effect and θ_t are semester fixed effects.

The results are presented in Table 2.2. In the first 4 columns I present results for the average grade, followed by the behavioural grade (columns (5)-(8)) and the number of hours missed at school (columns (9)-(12)). I only report the results of regressions which include individual level controls and class and semester fixed effects or individual and semester fixed effects.

[16] The relationship between parental education and the number of hours missed at school is less clear cut. The relationship between father's education and attendance is not significant. Children of more educated mothers seem to miss more hours of school; the effect is only significant at 10% level and is not of a large scale.

For the average grade, all regression specifications produce similar output with a negative coefficient on the emigration experience, positive on education dummies and a positive interaction term between the two variables.

Using the outcomes from column (4), I conclude that, compared to non-PWA children whose parents have an equivalent education level, the average grade of PWA pupils whose parents have lower than secondary education is .008 higher on average. This is an impact equivalent with only .24% of the grade's standard deviation and it is statistically insignificant. The finding is important as these students dominate the overall group of pupils with working parents abroad. Other PWA children gain relative to their non-migrant peers whose parents have the same qualifications. In particular, PWA pupils whose parents have secondary education have on average .108 better grade than their non-PWA peers whose fathers have an equivalent education level.

It is difficult to draw any conclusions about the differential impact of migration on behaviour and school attendance, depending on the migrant parent's education level, because majority of the regression coefficients are insignificant. There is, however, some weak evidence to suggest that children of PWA parents who are high school graduates may miss less schooling hours on average following the parent's departure. The effect is not very large and only marginally statistically significant.

The differential impact of parental emigration on school performance, depending on parents' educational attainment, may be related to potentially different success in employment abroad and distinctive perception of importance of education for children's future well-being.

Firstly, the positive income effect of emigration may differ, depending on the education level of migrant parents. Better educated migrant parents may be employed in better paying jobs relative to parents with a lower educational attainment, if jobs require specific qualifications or knowledge of the language of the destination country. However, many temporary migrants are likely to be underemployed.

Further, better educated migrants assimilate quicker (Card, 2005; Amuedo-Dorantes and de la Rica, 2007), which may improve their foreign experience due to their exposure to different cultures, more diverse network of contacts and better access to the labour market.

If better educated parents earn higher wages abroad, they are more likely to remit more in absolute terms and more money can be invested in a child's well-being, including education.

It is impossible to evaluate these statements using this data set, as it does not include any income information. However, data from Eurostat regarding mean earnings in various EU countries and findings of Jończy and Rokita-Poskart (2013) on remittances to Opolskie provide some support for the scenario outlined above.

Table 2.2: Differential impact of parental emigration depending on educational attainment

Dependent variable	average grade				behavioural grade				no of hours missed			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$EmigrParent_{it}$	-.921***	-.762****	-.061	-.081	-.399	-.221	.146	.103	7.188*	14.772***	8.361	9.461
	(.200)	(0.201)	(0.052)	(0.062)	(.306)	(.308)	(.229)	(.235)	(4.150)	(4.204)	(5.688)	(5.820)
father's education secondary	0.090**	0.090**			.116**	.116**			-1.019	-.934		
	(0.039)	(0.039)			(.050)	(.050)			(1.421)	(1.421)		
father's education tertiary	0.209***	0.209***			.179***	.179***			-1.202	-1.143		
	(0.053)	(0.053)			(.068)	(.068)			(1.835)	(1.840)		
$EmigrParent_{it}$ * low	.882***	.721***	.078	.093	.250	.068	-.119	-.089	-3.774	-11.170***	-2.054	-1.746
	(0.180)	(0.181)	(0.058)	(0.066)	(.295)	(.296)	(.232)	(.239)	(2.618)	(2.727)	(6.273)	(6.522)
$EmigrParent_{it}$ * secondary	.780***	.618***	.188***	.189***	.163	-.019	-.263	.247	-6.621**	-13.876***	-12.817**	-13.021*
	(0.180)	(0.181)	(0.064)	(0.070)	(.276)	(.278)	(.235)	(.242)	(3.251)	(3.338)	(5.865)	(6.071)
$EmigrParent_{it}$ * tertiary	.125	-.045	.125	.084	-.489	-.674	.158	.133	30.301**	21.356*	6.076	-.244
	(0.331)	(0.330)	(0.095)	(0.105)	(.526)	(.524)	(.263)	(.258)	(12.337)	(12.243)	(6.848)	(6.091)
Individual controls	yes	yes	no	no	yes	yes	no	no	yes	yes	no	no
Fixed effects	class	class and semester	individual	individual and semester	class	class and semester	individual	individual and semester	class	class and semester	individual	individual and semester
no of respondents	1985	1985	2436	2436	1863	1863	2470	2470	1262	1262	1604	1604

Source: MECP2012

Note: Individual controls in all specifications include gender, number of siblings, both parents' education, age and employment.
Two educational groups (vocational and elementary) were combined into a low education category.

Looking at Table 2.3, one may conclude that irrespective of the education level, by seeking employment abroad, Poles have a chance to increase their income three- to fourfold. Moreover, the higher their education level, the greater the gain to be made (in absolute terms).

One may argue, however, that emigrants are unlikely to be employed in their own professions, especially if they are staying abroad temporarily. For instance, Barrett and McCarthy (2007) find that immigrants from the new EU Member States earn on average 31% less than the natives in Ireland. Nonetheless, even taking into account a large wage disadvantage, there are still significant financial gains to be made and they are likely to increase in absolute terms with the educational attainment of the migrant parent.

However, the results in this chapter suggest that the gains are not only absolute, but also relative to peers of a similar background. From Table 2.3 it is clear that a migrant parent is likely to earn more than a parent with the same education level staying in Poland, even if he works below his qualifications and faces a wage disadvantage. The gain is smaller, however, for lower levels of education. Thus, there may be a threshold at which the income gain is sufficiently big to exert positive impact on a pupil's performance at school.

Even if the income gains are not significant enough to result in differential impacts by parental education levels, there may be other factors crucial for the size of the overall effect. For instance, parents' priorities with regards to their children may differ, depending on their education level (Guryan et al., 2008). In particular, parents with higher educational attainment may see their children's education as very important and spend a higher proportion of income on schooling or take other steps to ensure their children perform well at school - work with them at home, etc.

Once again, it is difficult to establish whether this indeed is the case here. However, looking at Table 2.4, it is clear that, with exception of families where parents have tertiary education, in migrant-sending households, a lower proportion of parents staying in Poland are employed. One of the reasons for such a situation may be that parents consciously choose to remain at home to compensate for the absence of a family member and ensure well-being of their children.

Given that the income abroad may be significantly more than double what one earns in Poland, it may suffice to improve the household finances, despite one parent leaving a job. The parents' presence at home may increase the benefits of migration, if it results in a significant increase in quality time with children. This may be particularly the case when parents are better educated and invest their time with children in activities which foster better school performance (Guryan et al., 2008; Carneiro et al., 2013).

Further, if better educated migrant parents assimilate better in the destination countries and enjoy their experience, they may also transfer some of the gained cultural capital onto their children, which may be beneficial to school performance.

Therefore, on balance, gains to be made from parental emigration may increase

Table 2.3: Mean annual earnings in construction, industry and services in 2010 by education level (€)

overall			
	total	male	female
European Union (15 countries)	35268	39440	30459
Germany	38735	43377	32870
Netherlands	41149	45664	36358
Poland	10233	11089	9287
United Kingdom	34817	41119	28386
Pre-primary and primary education			
European Union (15 countries)	22152	24040	19206
Netherlands	28418	31426	24556
Poland	6977	7894	5750
United Kingdom	21460	23115	17775
Lower secondary education			
European Union (15 countries)	25056	27396	22094
Germany	22577	24410	20812
Netherlands	29819	32880	26150
Poland	6132	6271	5550
United Kingdom	26558	30846	22351
Upper secondary education			
European Union (15 countries)	33315	36935	28939
Germany	37308	40858	32591
the Netherlands	37209	41122	33327
Poland	8292	9008	7298
United Kingdom	29322	34274	24037
First stage of tertiary education			
European Union (15 countries)	47980	56711	39338
Germany	62873	71953	50344
the Netherlands	56356	63433	49116
Poland	14823	18466	12733
United Kingdom	42183	50295	34187
Second stage of tertiary education			
European Union (15 countries)	37395	43119	31874
the Netherlands	51985	61002	42419
Poland	13055	15319	10573
United Kingdom	37861	45959	29805

Source: Eurostat

with parental educational attainment thanks to greater income potential and different investments made in children. Moreover, children may benefit relative to their classmates whose parents have the same educational attainment, as the migrant parent is still likely to earn more and the other parent may be able to invest more time in interactions with children, e.g. by leaving employment.

Table 2.4: Proportion of parents employed by migration and education status

Panel A: If a migrant family, father emigrated				Panel B: If a migrant family, mother emigrated			
overall							
	migrant	non-migrant	difference		migrant	non-migrant	difference
mother works	0.641	0.725	0.084	mother works	0.753	0.711	-0.042
father works	0.953	0.901	-0.052	father works	0.851	0.913	0.062
elementary education							
	migrant	non-migrant	difference		migrant	non-migrant	difference
mother works	0.533	0.668	0.135	mother works	0.647	0.656	0.009
father works	0.897	0.832	-0.065	father works	0.722	0.847	0.125
vocational education							
	migrant	non-migrant	difference		migrant	non-migrant	difference
mother works	0.59	0.673	0.083	mother works	0.75	0.652	-0.098
father works	0.957	0.901	-0.056	father works	0.841	0.916	0.075
secondary education							
	migrant	non-migrant	difference		migrant	non-migrant	difference
mother works	0.701	0.764	0.063	mother works	0.827	0.749	-0.078
father works	0.959	0.925	-0.034	father works	0.884	0.934	0.05
tertiary education							
	migrant	non-migrant	difference		migrant	non-migrant	difference
mother works	0.893	0.892	-0.001	mother works	0.875	0.892	0.017
father works	0.923	0.947	0.024	father works	0.813	0.948	0.135

Source: MECP2012

If the hypothesis of the gains increasing in education, relative to children from similar backgrounds, was true, one may argue that an even bigger effect should be found for children whose migrant parents have tertiary education. It is difficult to put forward a reliable argument in this case as the number of migrant parents with tertiary education is very low and hence precise estimation of the impact is impossible. However, it is important to note that, among families where parents have tertiary education level, the employment levels for either parent are very high and comparable, irrespective of the migration experience in the family. This implies that, even if the family experiences migration, usually both parents remain employed; hence, it may be difficult for the parent remaining with children in the home country to compensate for the separation. Then the positive effect of increased household income may be offset by the impact parental departure may exert on the family.

2.5.3 Lagged impacts

The baseline analysis leaves an array of questions unanswered. I observe a multitude of patterns in the sample. A big proportion of respondents indicated that their parents have been away a few times for short periods of time, rather than leaving the country for a long-term employment abroad.

This observation prompts a question about the effect of parental returns and subsequent departures on a child's performance. It is also difficult to predict the impact of the circular movements; even though they ensure frequent contact with the parent and hence a stronger bond, they also introduce a source of further instability into the household.

Moreover, the realisation of the emigration effects may be delayed and the full scale of the impact may be uncovered only after a certain amount of time has elapsed since departure; particularly when the separation from a migrant parent is prolonged.

Hence, it is reasonable to suspect a degree of dynamics in the emigration impacts on school performance of a child. I include 4 lags of the emigration status into the regression to see if the relationship between the average grade, behavioural grade and school attendance and parental migration changes. Otherwise the specification is defined as before. The results are presented in Table 2.5.

The estimates of the average grade regressions without individual fixed effects (columns (1)-(2), Table 2.5) are mostly insignificant and sensitive to inclusion of controls.

Table 2.5: Lagged impacts of parental emigration on children's outcomes

Dependent variable	Average grade			Behavioural grade			Hours missed not excused		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EmigrParent_{it}	-.007	-.141**	.057	-.036	-.200*	.019	.245	-.307	-3.388
	(.070)	(.068)	(.042)	(.107)	(.108)	(.146)	(5.090)	(5.460)	(4.846)
$\text{EmigrParent}_{i(t-1)}$	.163**	.062	.057	.091	.050	.282*	8.564	5.118	9.883**
	(.070)	(.072)	(.058)	(.120)	(.116)	(.151)	(8.414)	(5.358)	(4.999)
$\text{EmigrParent}_{i(t-2)}$	-.077	.042	-.121**	-.122	-.006	.114	-4.361	-1.482	-.148
	(.077)	(.075)	(.062)	(.113)	(.120)	(.158)	(5.827)	(5.001)	(4.937)
$\text{EmigrParent}_{i(t-3)}$	-.104	.018	-.203***	-.032	.098	-.071	-7.006	-5.447	-5.250
	(.066)	(.075)	(.061)	(.118)	(.133)	(.146)	(6.248)	(6.122)	(4.905)
$\text{EmigrParent}_{i(t-4)}$	-.115	-.136*	-.079	-.135	-.241*	-.132	7.956	9.739*	3.179
	(.080)	(.078)	(.050)	(.131)	(.140)	(.120)	(5.390)	(5.165)	(4.764)
Individual controls	no	yes	no	no	yes	no	no	yes	no
School FE	no	yes	no	no	yes	no	no	yes	no
Individual FE	no	no	yes	no	no	yes	no	no	yes
No of observations	4716	3670	4716	4094	3181	4094	1854	1508	1854
No of students	2629	2051	2629	2314	1809	2314	1163	947	1163

Source: MECP2012

Individual controls include gender, number of siblings, parents' age, employment, educational level, whether parents divorced or one parent died. Note that no semester fixed effects are included in the specification as they are dropped when 4 lags of the main explanatory variable are included.

Standard errors clustered by individual level in parentheses.

Statistical significance: *** 1% ** 5% * 10%

Table 2.6: Impacts on the average grade of various migration patterns over time

Panel A: Parent abroad for 2 consecutive semesters					
	Abroad now	Just came back	Has been back for a semester	Has been back for a year	Has been back for 3 semesters
Overall effect on the average grade	0.114	-0.064	-0.324***	-.282***	-.079
Panel B: Parent abroad for 3 consecutive semesters					
	Abroad now	Just came back	Has been back for a semester	Has been back for a year	Has been back for 3 semesters
Overall effect on the average grade	-.007	-.267***	-.403***	-.282***	-.079
Panel C: Parent has been away every other semester in the last 3 years					
	Home now	Away now			
Overall effect on the average grade	-.146***	-.143			

Note: The calculations are based on the results for the average grade presented in Table 2.5. Stars next to the impacts indicate statistical significance.

The fixed effect results are clearer with the coefficients on current emigration dummy and its first lag positive, though statistically insignificant. In fact they are comparable with the results obtained in the fixed effects regression without lags, which reiterates the idea of no negative immediate impact of parental emigration on children's grades. From lag 2 onwards, however, the coefficients are relatively large, negative and statistically significant, ranging from 14 to 24% of the grade's standard deviation which suggests that the full effect of parental departure realises after about a year.

The lagged effect is more detrimental than the instantaneous impact, and persistent. One should be cautious when interpreting the size of these results as the majority of emigration observed in the sample is temporary and characterised by returns and subsequent departures. Less than a quarter of migrant parents have been away permanently. An average migrant parent spends 2 out of 6 semesters abroad and 40% of migrant fathers return and subsequently depart.

I consider various scenarios to shed light on the impact, given the migration patterns. My calculations are presented in Table 2.6. In Panel A I look at cases when a parent has been abroad for 2 consecutive semesters. The impact is only positive if the parent is abroad now. Importantly, the effect upon return becomes negative, but dies off with time.

In Panel B I present the expected effects if a parent has been abroad for 3 consecutive semesters and find that there are gains to be made upon return, but the negative effect sets in after a year since return and is large. Again, it dies off gradually.

However, almost a third of migrant parents engage in circular migration. This case is considered in Panel C, where I assume that a parent is away for one period, back the next semester and away again throughout the three years. Then the negative effect is much smaller than in the previous two cases.

The coefficients in the analysis of behavioural grade (columns (4)-(6) of Table 2.5) reflect the patterns seen for the average grades. This may be due to the correlation of 0.69 between the behavioural and average grade in the sample. The coefficients in regressions without individual fixed effects are mostly statistically insignificant, large and negative. Most of the fixed effects coefficients are also insignificant. Nonetheless, it is clear that they become negative only from lag 3 onwards, i.e. the adverse impact, if any, is channelled with an even longer delay than in case of the average grade. The only statistically significant coefficient is on the first lag of emigration and it is positive.

The negative effect emerging after a while may be explained by various factors. For instance, detachment from a parent may be easier accepted by a child, when it is temporary and recent. Children may realise the difficulty of being apart only *ex-post* and when the parent has been abroad for long enough. The full effect of additional income flows may also be realised with a delay if migrant parents need time to settle in the destination country before sending remittances.

Children may also wish to join their migrant parent abroad; such a desire is likely to influence attitudes towards schooling and educational attainment, depending on the

perception of returns to education abroad relative to the home country.

For short-term, one-off, migration episodes the observed outcome might be explained by the expected fall in income upon parental return. In cases of prolonged emigration it might be that either the remittances fall with time, as the migrant parent establishes himself abroad and develops a more comfortable lifestyle, or no change in remittances takes place but the effect of separation is experienced to a greater extent. These interconnections are further complicated if one considers a possibility that children's future plans change, conditional on parental experiences; children of emigrant parents may want to emigrate too (see Kandel and Kao, 2001) and lose motivation to excel at school as their perception of returns to education changes.

Both scenarios are reasonable. Unfortunately, I do not have means of testing the hypotheses with the data.

The results for school attendance (columns (7)-(9)) provide a different story. As in the analysis of instantaneous impacts, the majority of the regression coefficients are statistically insignificant. In the OLS regressions only the coefficient on the fourth lag of emigration is significant and positive. In the fixed effects regressions the coefficient on the first lag of the emigration status is significant and also positive. According to these estimates, a PWA child seems to initially miss more school following parents' emigration, but the behaviour is reversed and the situation worsens again roughly two years after departure.

This seems counter-intuitive, but taking into account that majority of emigration in the sample is short term and repetitive, it may be that after 3 semesters the migrant parent is back and can ensure school attendance of the child and then leaves again. Note that the number of hours missed by a student due to emigration is not striking, perhaps because most PWA children stay with the other parent, who supervises them. The supervision may not be, however, as diligent as when both parents were home.

Given the statistical insignificance of most coefficients on behaviour and attendance and a very limited scope for testing any explanations of the results, drawing firm conclusions is impossible. The results are merely suggestive. The imprecision of the estimates may be caused by a small number of observations in the sample. The addition of 4 lags may be too demanding for the data set I am utilising.

2.5.4 Inference on test scores

An important question to be asked is to what extent an experience of parental emigration during adolescence determines future prospects. Are the grades obtained at this stage crucial for a child's further education?

They might be pivotal. At the end of *gimnazjum* students apply to new schools which differ in level of difficulty and determine pupils' opportunity to apply to a university or college. During the recruitment process, 90% of a credit given to an applicant

is based on his grades and results of national tests.[17]

Hence, the question is how the impact observed in grades translates to the test performance. I am in possession of data on respondents' performance in the national tests for under 13% of students. Given the cross-sectional nature of the data, a fixed effects regression with the scores as a dependent variable is impossible. I can draw inferences about a potential impact of parental emigration on these results from the regressions presented so far.

Students sit five competence tests[18] to independently evaluate their knowledge gained in the preceding three years and allow progression to the next stage of education. I rerun the fixed effects regressions of the average grade on the emigration status in the family, using only the subsample of students, for whom I possess information about the exam performance. I then run regressions of the test results on the average grade and use the coefficients to make inferences from the fixed effects regressions about potential impacts on the respondent's performance in the national exams.[19]

The instantaneous impacts are negligible, ranging from .497 to .940 of a test score, where the maximum test score is 100. Looking at the lagged regressions, the impacts from the scenarios presented above translate into between -5.190 to -2.749 points when a parent stayed abroad once and only for one semester, and between -7.248 and -3.839 points in case of prolonged emigration. The latter impact is relatively high, given that the average test scores vary between 44.74 and 62.19, depending on the subject.

2.6 Does sibling emigration play a different role?

Sibling influence on educational performance of children has not been extensively investigated. The majority of lessons related to the role of siblings come from literature analysing interegenerational transmissions and correlations (Bingley and Capellari, 2012; Black and Devereux, 2011) and find correlation in labour market and education outcomes of siblings. In migration literature Kuhn (2006) finds that emigration of brothers had a positive effect on schooling of children in rural Bangladesh, whereas sisters' emigration did not affect children's attainment. He argues that the result may be driven by differential income capacity, and hence remittances, of migrating siblings. He does not, however, correct for selection biases and endogeneity in his work.

Income aside, siblings also play a crucial role in one's decision making. They often act as role models and can motivate or discourage younger children from studying, influencing human capital accumulation of the left behind children. Biavaschi et al. (forthcoming) argue that siblings, who stay at home during parental migration, play a particularly important role when the family experiences migration. They show that

[17] A simplified version of the process is following: schools award marks for national test scores (maximum 50 marks), grades obtained in the final year of *gimnazjum* (maximum 40 points) and extra-curricular achievements, e.g. competitions, charity work (10 points). Students are then accepted to a school on the basis of their classification among all the applicants until all places have been filled.

[18] Polish language and literature, math, sciences, history and foreign language

[19] Details can be found in Appendix 2.C.

sibling influence on schooling performance is stronger among left-behind children, most likely due to changes in family roles following migration.

To investigate whether sibling's own emigration experience may influence school performance of pupils, I add a sibling emigration dummy as a right hand side variable into the equation.

The equation then becomes:

$$\text{average grade}_{it} = \alpha + \beta \text{EmigrSibling}_{it} + \theta \text{EmigrParent}_{it} + \gamma_i + \theta_t + \epsilon_{it} \quad (2.3)$$

where EmigrSibling_{it} is a dummy variable equal to 1 if at least one sibling of pupil i was abroad in semester t, EmigrParent_{it} is defined as before, γ_i are individual fixed effects, θ_t semester fixed effects and ϵ_{it} an error term.

Following on, I also add 4 lags of sibling emigration to see if there is a scope for a delayed effect.

I possess information on whether one's sibling has engaged in migratory experience prior to the observed period, although without much further detail. This division is important as those who have been migrating prior to September 2009 must be at least 6 years older than respondents and their relationship may differ, relative to one with only slightly older siblings. For that reason I separately run two sets of regressions with different groups of interest: siblings who emigrated recently and before September 2009 and siblings who have only engaged in emigration post-September 2009. All results are presented in Table 2.1.

As can be seen in Table 2.1, impact of siblings who migrated recently is strong and significant, but only when I allow for delayed effects (columns (4)-(6)). The OLS results for this group indicate a negative instantaneous and no delayed impact of emigration, once again most likely capturing not only the effects of emigration, but also the socio-economic family background. Results of regressions with individual fixed effects, however, reveal existence of a large, positive and statistically significant impact in lagged regressions; sibling migration immediately increases educational attainment by .202 - an equivalent of almost 24% of standard deviation of an average grade. If a sibling left for abroad 3 years ago and hasn't returned, the impact reaches 76.9% of standard deviation of an average grade. The effect of a one off, one semester departure dies off after a year. However, I find no impacts of emigration of those siblings who have first left before 2009.

The rationale for the positive and large impact of sibling emigration, that is persistent and accumulates over time, could be that many migrant siblings remit, positively influencing the household budgets (although Kuhn (2006) argues that these remittances are much lower than the ones sent by migrant parents). At the same time their absence leaves parental time inputs into family life unaltered. Moreover, it may be that siblings play a crucial role in encouraging educational success, especially if their migratory experience indicates high returns to education (Chand and Clemens, 2008).

Table 2.1: Impact of sibling emigration on children's average grade

Dependent variable	Average grade											
	Siblings who have emigrated post-2009 only						Siblings who migrated pre- and post-2009					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$EmigrSibling_{it}$	-.041	-.106*	.005	.093	-.035	.202***	-.254**	-.129	-.085	.296*	-.026	.001
	(.103)	(.060)	(.049)	(.124)	(.100)	(.050)	(.107)	(.127)	(.098)	(.155)	(.233)	(.130)
$EmigrSibling_{i(t-1)}$				-.108	.097	.262***				.211	.177	-.148
				(.130)	(.138)	(.074)				(.154)	(.184)	(.128)
$EmigrSibling_{i(t-2)}$				-.051	-.140	.134				-.477***	-.125	-.263**
				(.151)	(.118)	(.083)				(.161)	(.154)	(.115)
$EmigrSibling_{i(t-3)}$				.137	-.023	.124				-.169	.151	-.160
				(.136)	(.096)	(.085)				(.139)	(.126)	(.107)
$EmigrSibling_{i(t-4)}$				-.107	-.153	.190*				-.142	-.308*	-.066
				(.198)	(.114)	(.099)				(.231)	(.170)	(.074)
Individual controls	no	yes	no	no	yes	no	no	yes	no	no	yes	no
school FE	no	yes	no	no	yes	no	no	yes	no	no	yes	no
individual FE	no	no	yes	no	no	yes	no	no	yes	no	no	yes
No of observations	13837	10077	13413	4685	3390	4562	13525	10133	13525	4665	5243	4562
No of students	2661	1999	2583	2612	1960	2531	2620	1932	1932	2616	1932	1932

Source: MECP2012

Individual controls include: gender, no of siblings, emigration history of parents, employment, age and education level of parents.

Standard errors clustered by individual level in parentheses.

Statistical significance: *** 1% ** 5% * 10%

This may be the case since young people's migratory experience and employment opportunities often differ substantially from those awaiting their parents - they often know the language of the destination country, are more entrepreneurial, flexible and mobile (Nowicka, 2002).

More puzzling may be the fact that this impact is driven by siblings who only embarked onto migration post-September 2009. I expect them to be closer in age to the respondents than siblings with prior migratory experience. They may, therefore, still have a very strong bond with the household, remit and visit often, maintaining close relationships with younger siblings and influencing their decisions. Assuming that they were of age at point of emigration, siblings who left the country prior to 2009 must be at least 6 years older than respondents. In this case the age gap may change the relationship between siblings into a more parental one. Additionally, these siblings are likely to already have their own families and hence neither remit nor come back to their parents' home as often. Therefore, the impact may diminish.

Unfortunately, I cannot test these hypotheses at this stage.

2.7 Conclusions

This chapter explores impacts of parental and sibling emigration on children's outcomes, with particular focus on educational attainment of 16-year-olds. I utilise a unique data set with student-level information about teenagers in a high migration region of Poland and estimate the results using regressions with individual and semester fixed effects, to minimise the problems inherent in estimation of migration impacts, such as the endogeneity of migration decision or reverse causality in the relationship.

I find that parental emigration has a small and positive immediate impact on educational attainment of children. The positive effect, although counter-intuitive, may be thanks to the short-term, circular nature of parental migration in the sample; it is likely to lower any potential burden on PWA children and to more effectively channel positive aspects of international experiences, such as increased income, exposure to other cultures and possibly changed perception of returns to education.

However, not everyone benefits. The greatest gains are made by children whose migrant parents are at least high school graduates. More importantly, PWA children whose parents have lower than secondary education, who constitute about 67% of the overall group, do not improve their performance as a result of their parents' emigration, relative to their peers of similar socio-economic background. This may be due to different allocation of household resources and parental involvement in child's education, depending on the parents' own education level. Sufficiently educated parents may reap greater benefits of migration, including higher income and cultural gains, which they can pass onto their children. They may also value their children's education more highly and ensure that their children perform well at school despite their departure for abroad.

A finding of positive migration impact on children contradicts a number of outcomes from other case studies. Perhaps the reason behind it is the difference between the economic, cultural and social situation in Poland and in other source countries. In the Polish case, parental emigration is less likely to force children to abandon schooling altogether in order to engage in paid employment. Rather, the impact is more subtle and limited to an increase in household responsibilities. Moreover, given that emigration is legal and travel to Poland relatively short and affordable, the respondents in my study are in frequent contact with the parent abroad and suffer less from the feeling of abandonment than, say, children in Mexico whose parents work illegally in the USA. Contact is also maintained thanks to wide-spread use of internet communication and mobile networks. Additionally, in most cases the emigrant parents leave children with a family member, who takes over parental duties. A similar argument was used by Chen et al. (2009) in their study of Chinese rural-urban migration; children were left with family members and were not burdened with additional workload, whilst the household as a whole received an increased income. Biavaschi et al. (forthcoming) also argue that adjustments within the family left behind may generate benefits or at least reduce hardships. This is not to say that the emotional burden of family detachment is negligible.

My findings are in line with studies of Chen et al. (2009), Antman (2011b) and Hanson and Woodruff (2003), all of which demonstrate that parental emigration can have none or positive impact on the education of children. However, the magnitude of the positive effect observed by me in the study seems less striking in comparison with conclusions drawn by Hanson and Woodruff (2003) for Mexico. It is likely to occur because of a smaller income effect; the differences between the economic situation, standard of living, incomes and purchasing power in Poland and the destination countries of Poles are smaller than between Mexico and the destination countries of Mexicans. Hence, the potential for increased income due to emigration is also smaller. Should the negative impact of family detachment due to emigration be comparable in the two cases, the overall positive influence of emigration will naturally be smaller in the Polish case. The difference stems from a changed balance between the effects at play relative to earlier studies.

After taking into account the delay in average grade's response to migratory movements, I see that emigration might exert a negative effect. It seems to be delayed by a year for the average grade and its scale depends on the migration patterns observed in individual families. I consider different scenarios and present a possible range of impacts on the child's average grade which also translate to a poorer performance in the national tests. The negative relationship observed in the regression suggests that the income effect is outweighed by the negative influences of migration. The results are in line with various publications in the field (McKenzie and Rapoport, 2011; Antman, 2011a; Cortes, forthcoming; Kandel and Kao, 2001). However, the justification for such outcomes differs due to divergence in migration contexts. It is unlikely that respondents

in my study perform worse at school due to greater responsibilities, which is usually the argument proposed in literature. In most cases they stay with the other parent during emigration and hence do not need to take on adult responsibilities in the household. I do not observe school dropouts either.

I also find a strong, positive impact of sibling emigration on average grades of 16-year-olds. This may be because of their potential influence on children's decision making, a positive change in perceived returns to education due to the migratory experience and increase in the household's income. Sibling migration is likely to foster openness and provide an international outlook on opportunities for young people, which may facilitate human capital accumulation. Unfortunately, due to data limitations, I am unable to test these suggestions.

Despite its various caveats, this analysis sheds new light on the role migration plays in human capital accumulation. It seems that migratory experiences, which are temporary, repeated and rather short-term in nature, and during which a sibling or only one parent engages in employment abroad, may, under certain circumstances, benefit some children staying in the home country. The story is very complicated, however, and depends heavily on family circumstances, as well as the type of migration a family member engages in.

Therefore, my work also highlights the importance of heterogeneity analysis in this context and of use of an array of approaches to create a comprehensive view of the analysed situation. Even if partial, this analysis is one of the first few to acknowledge the different nature of European migration from the labour flows studied to date and to investigate various elements of the complicated temporary migration patterns observed in Europe.

Given that such migratory movements are increasingly common among the new EU member states, these findings may be informative of the situation in Europe.

Bibliography

AGAN, A. (2011): "Non-cognitive skills and crime," IZA Conference Paper.

AIZER, A., AND F. CUNHA (2012): "The production of human capital: endowments, investments and productivity.," NBER Working Paper Series 18429.

AMUEDO-DORANTES, C., AND S. DE LA RICA (2007): "Labour Market Assimilation of Recent Immigrants in Spain," *British Journal of Industrial Relations*, 42, 257–284.

AMUEDO-DORANTES, C., ET AL. (2008): "Migration, remittances and children's schooling in Haiti," IZA Discussion Paper 3657.

ANGRIST, J. D., AND J.-S. PISCHKE (2009): *Mostly harmless econometrics*. Princeton University Press, Woodstock, Oxfordshire.

ANTMAN, F. (2011a): "The intergenerational effects of parental migration on schooling and work: what can we learn from children's time allocation?," *Journal of Development Economics*, 96(2), 200–208.

——— (2013): "The Impact of Migration on Family Left Behind," in *International Handbook on the Economics of Migration*, ed. by A. F. Constant, and K. F. Zimmermann, pp. 293–309. Edward Elgar, Northampton, MA.

ANTMAN, F. M. (2011b): "International migration and gender discrimination among children left behind," *American Economic Review*, 101(3), 645–49.

——— (2012): "Gender, Educational Attainment, and the Impact of Parental Emigration on Children Left Behind," *Journal of Population Economics*, 25, 1187–1214.

APPS, P., ET AL. (2012): "The impact of pre-school on adolescents' outcomes: evidence from a recent English cohort," IZA Discussion Paper 6971.

BARRETT, A., AND Y. MCCARTHY (2007): "Earnings of immigrants in Ireland: results from the 2005 EU Survey of Income and Living Conditions," IZA Discussion Paper No. 2990.

BARRO, R. J. (2001): "Human Capital and Growth," *American Economic Review*, 91(2), 12–17.

BEHRMAN, J. R., ET AL. (2006): "What Determines Adult Cognitive Skills? Impacts of Pre-Schooling, Schooling and Post-Schooling Experiences in Guatemala," *Latin American Economic Review*, 23(4).

BIAVASCHI, C., ET AL. (forthcoming): "Sibling Influence on the Human Capital of the Left-Behind," *Journal of Human Capital*.

BINGLEY, P., AND L. CAPELLARI (2012): "Alike in many ways: Intergenerational and sibling correlations of brothers' earnings," CESIfo Working Paper No. 3994.

BLACK, S., AND P. DEVEREUX (2011): "Recent developments in intergenerational mobility," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card. North Holland Press.

BLACK, S. E., ET AL. (2005): "The more the merrier? The effect of family size and birth order on children's education," *The Quarterly Journal of Economics*.

BOTEZAT, A., AND F. PFEIFFER (2014): "The Impact of Parents Migration on the Well-being of children left behind: Initial Evidence from Romania," IZA Discussion Paper No. 8225.

CARD, D. (2005): "Is the new immigration really so bad?," *The Economic Journal*, 115, F300–F323.

CARNEIRO, P., ET AL. (2013): "Maternal education, home environments, and the development of children and adolescents.," *Journal of the European Economic Association*, 11, 123–160.

CHAND, S., AND M. CLEMENS (2008): "Skilled Emigration and Skill Creation: A quasi-experiment," Working Paper no. 152.

CHEN, X., ET AL. (2009): "Effect of Migration on Children's Educational Performance in Rural China," *Comparative Economic Studies*, 51, 323–343.

CORTES, P. (forthcoming): "The feminisation of international migration and its effects on the children left behind. Evidence from the Philippines.," *World Development*.

CUNHA, F., AND J. HECKMAN (2007): "The Technology of Skill Formation," *The American Economic Review*, 97(2), 31–47.

DE BRAUW, A., AND J. GILES (2006): "Migrant Opportunity and the Educational Attainment of Youth in Rural China," IZA Discussion Paper 2326.

DUSTMANN, C., AND A. GLITZ (2011): "Migration and Education," Centre for Research and Economic Analysis of Migration, No 05/11.

FEINSTEIN, L. (2003): "Inequality in the Early Cognitive Development of British Children in the 1970 Cohort," *Economica*, 70, 73–97.

GASSMANN, F., ET AL. (2013): "The impact of migration on children left behind in Moldova," Discussion paper.

GIANNELLI, G., AND L. MANGIAVACCHI (2010): "Children's Schooling and Parental Migration: empirical evidence on the left-behind generation in Albania," *Labour*, 24, 76–92.

GIBSON, J., ET AL. (2010): "Accounting for selectivity and duration-dependent heterogeneity when estimating the impact of emigration on incomes and poverty in sending areas.," Policy Research Working Paper Series 5268.

GURYAN, J., ET AL. (2008): "Parental education and parental time with children," NBER Working Paper 13993.

HANSON, G. H., AND C. WOODRUFF (2003): "Emigration and educational attainment in Mexico," Working Paper.

JOŃCZY, R., AND D. ROKITA-POSKART (2013): "Fiscal effects of migration as a factor increasing disparities in local and regional development - an example of the Opole Voivodship in Poland," *Economic and Environmental Studies*, 13, 9–21.

KANDEL, W., AND G. KAO (2001): "The Impact of Temporary Labour Migration on Mexican Children's Educational Aspirations and Performance," *International Migration Review*, 35(4), 1205–1231.

KUHN, R. (2006): "The Effects of Fathers' and Siblings' Migration on Children's Pace of Schooling in Rural Bangladesh," *Asian Population Studies*, 2(1), 69–92.

LYLE, D. S. (2006): "Using Military Deployments and Job Assignments to Estimate the Effect of Parental Absences and Household Relocations on Children's Academic Achievement," *Journal of Labor Economics*, 24(2), 319–350.

MCKENZIE, D., AND H. RAPOPORT (2007): "Self-selection patterns in Mexico-US migration: the role of migration networks.," *The Review of Economics and Statistics*, 92, 811–821.

——— (2011): "Can migration reduce educational attainment? Evidence from Mexico.," *Journal of Population Economics*, 24, 1331–1358.

NOWICKA, M. (2002): "Portret polskiego emigranta. Struktura i wielkosc emigracji z Polski przed i po rozszerzeniu Unii Europejskiej na wschod," Masters' Thesis, Jagiellonian University.

POSTLEWAITE, A., AND D. SILVERMAN (2006): "Non-cognitive skills, social success and labor market outcomes," Working Paper.

RENEFLOT, A. (2007): "Family Dissolution and Children's Educational Attainment: the Norwegian Story," Discussion paper, Centre for Advanced Studies, Oslo.

STEELE, F., ET AL. (2009): "Consequences of family disruption on children's educational outcomes in Norway," *Demography*, 46(3), 553–574.

STEINMAYR, A. (2013): "When a random sample is not random. Bounds on the effect of migration on children left behind.," mimeo.

SUN, Y., AND Y. LI (2007): "Parental divorce, sibship size, family resources, and children's academic performance," *Social Science Research*, 38, 622–634.

THE CENTRAL STATISTICAL OFFICE OF POLAND (2009): "Metoda szacunku liczby osob (stalych mieszkancow Polski) przebywajacych czasowo za granica w latach 2002-2008," .

——— (2012): "Poland in figures, 2012," .

TYNELSKI, A. (2010): "Sytuacja dziecka w rodzinie migrującej w Polsce," Discussion paper, Ministerstwo Edukacji Narodowej.

ZASLOW, M. (1988): "Sex differences in children's response to parental divorce. Research methodology and postdivorce family forms.," *American Journal of Orthopsychiatry*, 58, 355–378.

Appendix

2.A PWA pupils in the schools

Table 2.A.1: Changes in class composition over the observed period

total number of registered students:			3423
number of surveyed students:			2822
dropped out:	229	joined the school	109
failed a year	229		
transferred to another school	1	transferred from another school	3
went abroad	1	came from abroad	0
died	1		
do not know why	67	do not know why	106
Transferred to another class in the same school:			10

Source: MECP2012

Table 2.A.2: Departures and arrivals of children born in 1996 from abroad into Opolskie

Year	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	total
Emigration	16	14	17	7	24	25	28	13	14	25	183
Immigration	4	4	3	5	2	10	3	7	2	4	4

Data come from the Central Statistical Office of Poland.
The flows are approximated on the basis of family deregistration from an address in Poland.

Table 2.A.3: Migration situation of pupils by birth year

	Went to school early	Started school on time	Failed at least one year				
born in	1997	1996	1995	1994	1993	no info	total
number of pupils	15	2413	81	12	1	300	2822
migration in general	3	692	39	7	0	69	809
parental migration	1	285	14	2	0	25	327
sibling migration	2	129	7	0	0	9	147
Summary for the group of older students:							
no of pupils		94					
parent abroad		16	17.02 %				
sibling abroad		7	7.45 %				
emigration in general		45	47.87 %				

Source: MECP2012

2.B Empirical framework - further checks

Table 2.B.1: Regression results with leads of emigration variable

Dependent variable	Average grade											
	OLS	OLS	FE	OLS	OLS	FE	OLS	OLS	FE	OLS	OLS	FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$emigration_t$	-.044	-.002	.053*	-.054	.009	.082**	-.077	-.019	.059	-.041	-.004	.046
	(.042)	(.044)	(.029)	(.052)	(.055)	(.035)	(.067)	(.074)	(.042)	(.085)	(.091)	(.043)
$emigration_{t+1}$	-.110***	-.108**	.001	-.056	-.050	.024	-.094*	-.057	.034	-.114	-.081	.020
	(.041)	(.042)	(.024)	(.040)	(.041)	(.030)	(.052)	(.053)	(.037)	(.071)	(.077)	(.062)
$emigration_{t+2}$				-.051	-.063	.003	.017	.015	.031	-.002	.030	.081
				(.048)	(.047)	(.028)	(.050)	(.049)	(.038)	(.073)	(.074)	(.055)
$emigration_{t+3}$							-.015	-.039	.058	.054	.061	.117**
							(.066)	(.066)	(.042)	(.074)	(.075)	(.054)
$emigration_{t+4}$										-.076	-.111	-.012
										(.066)	(.068)	(.044)
controls	no	yes	no	no	yes	no	no	yes	no	no	yes	no
no of observations	11593	9099	11593	9127	7177	9127	6588	5192	6588	4493	3541	4493
no of students	2653	2067	2653	2573	2011	2573	2541	1993	2541	2462	1936	2462

Source: MECP2012
Controls include: sex, sibling, education, employment and age of parents, and school fixed effects
Clustered standard errors in parentheses.
Statistical significance: *** 1% ** 5% * 10%

Table 2.B.2: Differences in outcomes between semester 2 and semester 1 within each year

	Year 1			Year 2			Year 3		
	Overall	Migrant	Non-migrant	Overall	Migrant	Non-migrant	Overall	Migrant	Non-migrant
Δ Average grade	0.059	.069	.056	.136	.134	.137	.262	.256	.264
	(.307)	(.313)	(.305)	(.285)	(.287)	(.285)	(.284)	(.259)	(.295)
N	2161	622	1513	2222	641	1552	2215	662	1528
Δ Behaviour	.001	-.024	.007	.159	.195	.144	.342	.361	.334
	(.774)	(.794)	(.766)	(.750)	(.732)	(.756)	(.732)	(.771)	(.717)
N	1892	544	1323	1999	586	1386	1898	568	1307
Δ School attendance	6.163	8.678	5.132	8.348	8.711	8.228	11.985	16.767	10.403
	(26.339)	(29.920)	(24.834)	(25.493)	(28.571)	(24.249)	(29.385)	(35.581)	(26.852)
N	1151	329	811	1171	346	813	726	189	528

Source: MECP2012

Note: standard deviation provided in parentheses

2.B.1 Local economy of Opolskie in years 2009-2012

Table 2.B.3: Economic indicators for Opolskie in period 2009-2012

Economic indicator	Unemployment rate (%)				Average wages (PLN)			
Year	2009	2010	2011	2012	2009	2010	2011	2012
Opolskie province	12.9	13.6	13.3	14.4	2987.87	3137.29	3249.58	3358.42
By county								
brzeski	18.7	20.5	20.3	21.0	2687.60	2795.69	2962.78	3067.56
glubczycki	16.3	17.9	18.0	19.8	2750.66	2878.02	3031.11	3111.54
kedzierzynsko-kozielski	12.5	13.1	12.9	14.5	3363.79	3518.97	3753.82	3793.84
kluczborski	15.1	15.5	15.2	15.2	2730.08	2848.38	2994.99	3200.22
krapkowicki	11.6	10.9	10.1	10.9	3602.56	3798.54	3597.89	3720.00
namyslowski	17.8	18.6	18.1	19.2	2671.86	2833.22	2974.33	3152.42
nyski	18.5	19.4	19.0	21.4	2612.02	2733.31	2846.85	3012.34
oleski	8.6	8.9	10.2	10.6	2622.70	2731.82	2868.91	3013.57
opolski	12.1	13.1	12.2	13.5	2681.01	2872.04	2785.98	2948.63
prudnicki	16.8	18.6	18.6	19.5	2594.65	2730.02	2958.74	3052.97
strzelecki	11.8	11.7	10.2	11.3	2839.95	2929.69	3079.75	3221.15
city of Opole	5.9	6.4	6.2	7.1	3352.46	3541.80	3714.16	3771.22

Source: Central Statistical Office of Poland, database can be accessed on
http : //stat.gov.pl/bdlen/app/strona.html? p_name = indeks

2.B.2 Regressions using national exam scores instead of average grades

The analysis in this chapter relies mostly on the average grade as a dependent variable. The average grade, however, is awarded internally and may not objectively reflect pupils' skills. To check whether the average grade is a satisfactory measure of school performance, I rerun the baseline regressions using the national exam results of almost 13% of respondents.

I have information about pupils' results in exams in the following subject areas: literature, history, math, science and foreign language. The average grade used in the analysis is an average over all courses taken by a pupil, which include the examined subject areas. Therefore, to make the two measures comparable in terms of the knowledge and skills they are assessing, I create a new variable, which is an average test score for an individual, based on all the exam results. It is aimed to capture a pupil's overall performance in all 5 exams.

The results are presented in Table 2.B.4. Although statistically insignificant (due to sample size), the results suggest existence of a positive relationship between parental emigration and a pupil's performance in the national test.

Table 2.B.4: Impact of parental emigration on test scores

Panel A: Average test score statistics					
	mean	st.dev.	min	max	n
average test score	53.853	16.998	20.4	96.2	334
Panel B: Regression results					
	(1)		(2)		
$emigration_{it}$	6.720**		5.052		
	(2.796)		(3.390)		
Individual controls	no		yes		
N	334		268		
R-squared	.014		.165		

Source: MECP2012
The regressions in this table are based on observations for a subsample of respondents for whom exam results data are available.
The dependent variable is the average exam result (an average of all exams pupils took). The main explanatory variable is $emigration_{it}$.
Robust standard errors in parentheses.
Statistical significance: *** 10%, ** 5%, *1%

2.B.3 Lagged Dependent Variable specification

In this section I explore in more detail an alternative specification to the ones presented in Section 2.4.

As mentioned, the fixed effects specifications isolate any time-invariant changes specific to a student which influence the school outcomes of a pupil and may be correlated with migration decision of the parent. One may argue, however, that some of the important omitted variables vary over time. In particular, past school performance is likely to explain a large proportion of the current performance and may be correlated with the migration decisions of parents.

A specification including a lag of the dependent variable as an explanatory variable may shed some light on the issue of which changes in particular drive the results. By including the lagged dependent variable into the regression I am hoping to capture any remaining unobserved characteristics (not captured by individual and semester fixed effects) which may be influencing current school performance. Then the decision of parents to emigrate needs to be exogenous only to changes in the school performance and not its overall level. I will only focus on the average grade and the baseline regressions for the purpose of this exercise.
I run the following regressions:

$$Y_{it} = \alpha + \beta EmigrParent_{it} + \lambda Y_{i(t-1)} + \gamma X_{ict} + \theta_t + \epsilon_{it} \tag{2.4}$$

$$\begin{aligned} Y_{it} = \alpha + \beta EmigrParent_{it} + \beta_1 EmigrParent_{i(t-1)} + \beta_2 EmigrParent_{i(t-2)} \\ + \beta_3 EmigrParent_{i(t-3)} + \beta_4 EmigrParent_{i(t-4)} + \lambda Y_{i(t-1)} + \gamma X_{ict} + \theta_t + \epsilon_{it} \end{aligned} \tag{2.5}$$

where Y_{it} is the average grade of individual i at the end of given semester t, $Y_{i(t-1)}$ is its first lag, $EmigrParent_{it}$ is a dummy variable equal to 1 if pupil i's parent was abroad at time t, and is the main variable of interest. X_{it} is a set of individual level controls, θ_t semester fixed effects. Standard errors are clustered at individual level.

In Table 2.B.5 I present the results of the regression outlined above. I also restate the results of the regressions with individual fixed effects presented in Tables 2.1 and 2.5 as I will be referring to them in my discussion.

Table 2.B.5: Individual FE vs. LDV Specification

Dependent variable	Average grade									
	FE	FE	LDV	LDV	LDV	FE	FE	LDV	LDV	LDV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$EmigrParent_{it}$	.045*	.034	-.006	-.013	-.021	.057	.018	-.003	.013	-.008
	(.024)	(.024)	(.012)	(.012)	(.013)	(.042)	(.032)	(.029)	(.030)	(.029)
$EmigrParent_{i(t-1)}$						.057	.007	.064*	.009	-.002
						(.059)	(.047)	(.038)	(.037)	(.037)
$EmigrParent_{i(t-2)}$						-.121*	-.061	-.023	-.006	-.010
						(.062)	(.045)	(.045)	(.043)	(.041)
$EmigrParent_{i(t-3)}$						-.204***	-.087*	-.050	-.025	-025
						(.061)	(.044)	(.036)	(.037)	(.038)
$EmigrParent_{i(t-4)}$						-.079	-.023	.015	.001	.030
						(.050)	(.034)	(.033)	(.037)	(.039)
$Y_{i(t-1)}$			.920*	.918*	.894***			.908***	.910***	.896***
			(.004)	(.005)	(.006)			(.006)	(.007)	(.008)
Controls										
Individual	no	no	no	yes	yes	no	no	no	yes	yes
Semester FE	no	yes	no	yes	yes	no	yes	no	yes	yes
Class FE	no	no	no	no	yes	no	no	no	no	yes
N	2657	2657	2475	1851	1851	2629	2629	2435	1820	1820

Source: MECP2012
The dependent variable is the individual average grade at time t.
The main explanatory variable is the dummy variable for having a parent abroad at time t.
Other individual controls include gender, number of siblings, mother and father's age and education.
Standard errors are clustered at the class level and reported in parentheses.
FE stands for individual fixed effects, LDV stands for the lagged dependent variable specification
Statistical significance levels *** - 1%, ** - 5%, * - 10%

As can be seen in Table 2.B.5, for contemporaneous regressions, the fixed effects results indicate a positive although almost insignificant impact, whereas the outcomes in LDV regressions point towards a negative (but insignificant) relationship between having a parent abroad and the average grade.

Differences in estimation coefficients emerge also in the regressions with lags of migration dummy. As discussed before, the fixed effects approach suggests that the contemporaneous and first lag impact is positive but becomes negative from the year after parental departure onwards. The delayed negative effect can be substantial and outweighs the immediate positive influence. The situation is less clear-cut in LDV regressions because the coefficients on variables of interest are small and insignificant.

It is difficult to establish which approach is best suited in this case as it depends on the belief about the behaviour of the omitted variables, i.e. whether they are time-invariant or not. The lagged dependent variable and individual fixed effects models

are not nested and the distinction may play a role here. Angrist and Pischke (2009) point out, however, that LDV and FE models have a bracketing property which may be informative of the true relationship being analysed. If LDV is the correct approach but fixed effects are used, then the estimates of the positive effect will tend to be too big. If the reverse is true, then the estimates of the positive effect will be too small.

The FE estimates for the contemporaneous impact presented in Table 2.B.5 are larger than the LDV estimates and could provide an upper bound on the effect if the lagged dependent variable approach was more appropriate. In this case, the LDV estimates indicate either the correct impact or its lower bound. Since the LDV estimates are negative, very small and insignificant, the bounding exercise seems to rule out large negative effects in this case.

In case of the analysis of any delayed effects of migration, following the same logic as before, one may conclude that the bounding exercise rules out existence of large positive effects in a long run.

2.C Test scores inferences

Table 2.C.1: Regressions of the average grade for the sub-sample of students with test scores

	(1)	(2)
$EmigrParent_{it}$	.061	.268***
	(.048)	(.053)
$EmigrParent_{i(t-1)}$	-	.301**
		(.125)
$EmigrParent_{i(t-2)}$	-	-.137
		(.158)
$EmigrParent_{i(t-3)}$	-	-.218**
		(.105)
$EmigrParent_{i(t-4)}$	-	-.076
		(.128)
Individual controls	no	no
Individual FE	yes	yes
Semester FE	yes	no
N	328	316

Source: MECP2012
The dependent variable is the average grade is semester t.
The main explanatory variables are the dummy equal to 1 if parent abroad at time t and its 4 lags.
Clustered standard errors are reported in parentheses.
Statistical significance: *** 1% ** 5% * 10%

Table 2.C.2: Regressions of the test scores on the average grade of respondent

The dependent variable here is a score in the national test taken in April 2012				
	(1)	(2)	(3)	(4)
	literature score	math score	history score	science score
average grade	10.258***	14.913***	9.865***	7.896***
	(.571)	(.745)	(.559)	(.559)
N	349	349	349	349

Source: MECP2012
Robust standard errors in parentheses.
Statistical significance: *** 1%, ** 5%, * 10%

Table 2.C.3: Inferred fixed effects impacts of emigration on the test scores

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	literature score	literature score	math score	math score	history score	history score	science score	science score
$EmigrParent_{it}$	0.646	2.749***	0.940	3.804***	0.621	2.644***	0.497	2.116***
$EmigrParent_{i(t-1)}$		3.088**		4.489**		2.969**		2.377**
$EmigrParent_{i(t-2)}$		-1.405		-2.043		-1.352		-1.082
$EmigrParent_{i(t-3)}$		-2.236**		-3.251**		-2.151**		-1.721
$EmigrParent_{i(t-4)}$		-.780		-1.133		-.750		-.600

Source: MECP2012
Robust standard errors in parentheses.
Statistical significance: *** 1%, ** 5%, * 10%

2.D Role of parental education in children's schooling outcomes

Table 2.D.1: Impact of parental education on a child's schooling outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Dependent variable	Average grade					Behavioural grade					Number of hours missed				
$FatherSecon_i$	.037	.073*	.073*	.093* .076**	.038	.072	.072	.089*	.085*	-.696	-2.509*	-2.447*	-3.085**	-2.107	
	(.037)	(.039)	(.039)	(.038)	(.037)	(.052)	(.053)	(.053)	(.053)	(.051)	(1.511)	(1.415)	(1.407)	(1.482)	(1.432)
$FatherTert_i$	.334***	.318***	.319***	.228***	.173***	.218***	.208***	.211***	.155**	.099	-2.066	-2.211	-1.714	-2.506	-1.756
	(.057)	(.056)	(.056)	(.057)	(.053)	(.080)	(.077)	(.077)	(.079)	(.073)	(2.105)	(2.085)	(2.084)	(2.149)	(2.111)
$MotherSecon_i$	.139***	.126***	.126***	.126***	.110***	.054	.057	.056	.069	.049	1.847	2.741	2.697*	2.939*	2.485**
	(.037)	(.039)	(.038)	(.039)	(.038)	(.052)	(.052)	(.052)	(.053)	(.050)	(1.464)	(1.509)	(1.502)	(1.409)	(1.260)
$MotherTert_i$	.376***	.384***	.384***	.344***	.279***	.205***	.221***	.221***	.220***	.160**	.690	.726	.948	2.016	3.079*
	(.049)	(.049)	(.049)	(.049)	(.046)	(.069)	(.070)	(.070)	(.070)	(.066)	(2.062)	(2.020)	(2.005)	(1.944)	(1.737)
Individual level controls	no	yes	yes	yes	yes	no	yes	yes	yes	yes	no	yes	yes	yes	yes
School FE	no	no	no	yes	no	no	no	no	yes	no	no	no	no	yes	no
Class FE	no	no	no	no	yes	no	no	no	no	yes	no	no	no	no	yes
Semester FE	no	no	yes	yes	yes	no	no	yes	yes	yes	no	no	yes	yes	yes
N	2379	1983	1983	1983	1983	2210	1838	1838	1838	1838	1454	1245	1245	1245	1245

Source: MECP2012

This table contains results of a regression of the child's school outcomes on parental education dummies. Additional controls are progressively added to the equation and include: migration dummy, number of siblings, parental age and whether parents work, school or class fixed effects and semester fixed effects. Only the coefficients on education dummies are reported to show the positive relationship between the education level of parents and school performance of their children.

I include the following education dummies: $FatherSecon_i$ if father has completed secondary education, $FatherTert_i$ if father has graduated from university, $MotherSecon_i$ if mother has completed secondary education, $MotherTert_i$ if mother has graduated from university. $FatherLow_i$ and $MotherLow_i$ are categories for father or mother having education level below secondary, respectively, and are omitted.

Robust standard errors in parentheses.

Statistical significance: *** 1%, ** 5%, * 10%

Chapter 3

Educational spillovers and parental migration

3.1 Introduction

In 2006 the flow of temporary immigrants to the OECD countries reached 2.5 million and was three times higher than the flow of permanent immigrants.[1] The Polish Ministry of Education reports that 20% of Polish educational institutions surveyed in 2010 had pupils for whom one or both parents have emigrated.[2] This raises concerns about the immediate impact parental emigration has on pupils as well as on their classroom peers. In Chapter 2 I consider the former. Here, using data for Poland, I analyse whether children whose parents work abroad (henceforth PWA children) influence the school performance of their classmates.[3] The question is very relevant and potentially important given the scale of temporary migration, the role temporary migrants play in the labour markets and the fact that many of them have left their families behind.

Large scale parental emigration raises questions about children's immediate welfare as well as long term socio-economic implications, although it is theoretically ambiguous whether the impacts of parental employment abroad are negative or not.[4]

These considerations are crucial because early life human capital acquisition depends on nature as well as nurture and is vital for outcomes of adult individuals.[5]

Parental decisions to emigrate may impact the educational attainment of a child, as emigration leads to family separation, less quality time with the migrant parent and possibly greater household responsibilities for children.[6] At the same time, it usually results in an increase of household income, which can be directly and indirectly invested in a child's education. In some cases parental emigration may also change the perception of returns to education.

[1]Note that the estimate is based on statistics for 20 countries and relies predominantly on the count of permits issued and hence does not adequately capture the migration within the EU Member States where the labour movement is unrestricted. Therefore, this is likely an underestimate of the phenomenon. Moreover, it is estimated that between 20 and 50% of migrants return home within the first five years of arrival in the destination country. Family reunions are among the reasons for return (OECD, 2008). In 2011 almost 10% of the Polish households had at least one member residing temporarily abroad (The Central Statistical Office of Poland, 2013a).

[2]See Tynelski (2010).

[3]Children whose parents left for employment abroad have been called in the literature *the left behind children*. I refrain from using this phrase in this study as I do not perceive children of temporary migrants, who stay with the other parent in the home country and see the migrant parent on the regular basis as left behind.

[4]See Chen et al. (2009); Dustmann and Glitz (2011); Amuedo-Dorantes et al. (2008).

[5]See Apps et al. (2012); Aizer and Cunha (2012); Behrman et al. (2006); Feinstein (2003); Barro (2001).

[6]Note, however, that the effect of separation in this case is most likely limited due to the nature of migration in the sample.

Human capital is also shaped by one's surroundings, particularly educational environment. Thus, even if a child does not experience parental emigration personally, it may be impacted by emigration of a classmate's parents.

The presence of PWA children in the class may have an effect on the performance of their peers; if parental emigration affects a child's behaviour or performance at school, it will also influence the learning environment of other children in the class and their performance.[7] Suppose the parental emigration improves a PWA child's performance and positively changes its attitude towards education as the income gain dominates any potential negative effects of family separation. The child may then directly motivate peers by providing good example. The better performance will also increase the average academic quality of the class. The presence of PWA children in a class may also change the teachers' or schools' approach, which will affect all pupils.

Such a spillover effect cannot be determined by theory. Its sign and magnitude depend both on the impact parental emigration has on their children and on the interactions between pupils in the class. Therefore, whether PWA children influence their classroom peers is an empirical question.

Migration literature has considered the impacts parental emigration may have on children and what the contributing factors are.[8] To the best of my knowledge, however, no studies have addressed the question of classroom spillover. Lack of adequate data combining educational outcomes of children, their class and school allocation with the migration history in the family may be one of the reasons why.

I created and collected a data set for the purpose of this analysis. In particular, I obtained detailed information about migration experiences in the families and their timing, family background, school allocation, classroom composition and academic progress of pupils (See Chapter 1). The information about the timing of migration is key for identification which exploits the within-class variation in the proportion of PWA pupils over time.

I introduce class fixed effects into the regression to control for time-invariant unobserved characteristics of classes which may be influencing the performance of pupils and be related to class composition.

I then find that pupils benefit from the presence of the PWA peers in the class and the effect is non-negligible. A one standard deviation increase in the fraction of PWA pupils in the class is associated with about 3% of the standard deviation increase in the average grade. The result may seem counterintuitive given the aforementioned concerns about the negative effect of parental emigration. However, in Chapter 2, exploring the same data, I find that parental emigration experience exerts a positive impact on children's average grades. It is plausible that the improved school performance of PWA pupils is channelled onto their peers, through their motivation and the interaction

[7]The peer effects literature has already established that children are likely to be influenced by their school friends (Sacerdote, 2000; Black et al., 2013; Carrell et al., 2008).

[8]See Antman (2013) for the literature review.

between them.

One particular group of influencers emerges - PWA pupils, whose parents have completed secondary education. Increasing their proportion in the class yields a greater increment in the average grade than when the overall proportion of PWA pupils changes. It is difficult to establish the exact mechanism behind such an outcome. One may think that these PWA children gain more from their parents' experience abroad, as better educated parents have the potential for earning higher income abroad and hence remitting more. At the same time, they may invest a higher proportion of earned income or their own time in their children's schooling. In that case, this particular group of children has the potential of sharing the positive influence with their classmates. They may also be more effective in influencing the educational outcomes of others by sharing information on the value of education.

The presence of PWA peers in the class is more beneficial for those who themselves were exposed to emigration in the family. It may be because of the types of interactions within the class; perhaps children who have a parent abroad interact more with like children, as they have more in common and, therefore, are more influenced by them.

The analysis is not without limitations. The approach does not cater for situations in which time-varying changes, affecting both the average grade of pupils and the class composition, take place. I discuss such potential limitations.

I consider alternative explanations for the effect and eliminate cases in which schools reallocate resources to support PWA pupils or teachers inflate grades in classes with a higher concentration of PWA children. It is possible, however, that teachers put more effort in teaching classes with PWA pupils to overcompensate for having parents abroad.

Despite its various caveats, this analysis sheds new light on the role migration plays in human capital accumulation. It highlights the fact that impacts of migration are not limited to the affected families, but may spill over onto those surrounding them. This study also reveals heterogeneity within the group of migrants. Not all PWA children influence their peers. The impact depends on the socio-economic background of the migrant family; only children whose parents are sufficiently educated benefit from migration and positively influence their peers.

The migration in question differs from the migratory movements studied before. It is temporary, repeated and rather short-term in nature. Usually only one parent engages in employment abroad and remains in frequent contact with the family. Thus this analysis is not only the first to look at peer effects in this context, but considers new migratory movements, which are increasingly common in Europe. As such its findings may be informative for current policy setting.

The remainder of the chapter is structured as follows. In Section 3.2 I briefly review the literature. Section 3.3 discusses the data and Section 3.4 the empirical framework. In Section 3.5 I present results and in Section 3.6 consider explanations other than peer effect for the findings. I then consider threats to validity of the results and conclude.

3.2 Literature

This study draws on various concepts in economic literature. The idea to look at the school performance of teenagers relates to the fact that economists, among other researchers, see development of cognitive and non-cognitive skills as crucial for both short-term and long-term outcomes of individuals (Apps et al., 2012; Aizer and Cunha, 2012; Behrman et al., 2006; Feinstein, 2003). Such skills also play an important role in economic development (Hanushek and Woessmann, 2009). Therefore, there has been increased interest in what happens to children's performance at various stages of education.

There is a consensus that human capital is moulded by nature, as well as nurture (Cunha et al., 2010; Cunha and Heckman, 2007); thus, many studies consider the role family plays in one's human capital development. What is more relevant to the analysis here is the fact that environmental factors also play a role.

In particular, friends or peers at school may impart a great influence on one's performance; they can influence one's capacity to acquire new skills by being a component of their learning environment and, potentially, directly influencing their behaviour and attitudes towards learning. The education economics literature has investigated particular schooling settings to establish whether peer effects arise and what are the key factors behind them. The research began with the analysis of *pure academic effects* (Sacerdote, 2000; Black et al., 2013; Burke and Sass, 2006; Carrell et al., 2008; Evans et al., 1992; Hanushek et al., 2001; Zimmerman, 2003), followed by various studies of *the white-black score gap* (Guryan, 2004; Card and Rothstein, 2007; Angrist and Lang, 2004; Hoxby, 2000).

To the best of my knowledge, the peer effects in relation to PWA children have not been studied before. As suggested previously, this may be due to the amount and type of information required to identify the spillover and isolate it from other possible effects. Therefore the results presented in this chapter cannot be directly compared with any other research.

However, some similarities can be sought in the literature on the immigrant peer effects and on the influence interruptions to family life have on children and their peers.

The context of analyses of immigrant peer effects differs substantially from the one I am considering as PWA children are native to the area they live in. Therefore, unlike immigrants, they do not face linguistic barriers and do not need to assimilate. However, the studies rely on similar methodology and analyse how an increase in the proportion of foreign-born pupils in the class affects the peers. Findings vary, depending on the country of study, age group, type of immigrants and measures of academic performance used.[9]

[9]For example, Brunello and Rocco (2013), using PISA data for 19 countries, find that increasing the share of immigrant pupils in secondary schools negatively affects the test scores of natives. Ohinata and van Ours (2013), on the other hand, argue that, even though the immigrant children affect the learning environment in the school, they exert no negative spillover on Dutch students.

Since parental migration changes the home environment, which may influence children's behaviour, it can be related to the literature on disruptive or difficult children. Also here the peer effects vary, depending on the type of disruption considered.[10]

Peers may also have a positive impact on classmates, as found by Bobonis and Finan (2009) and Lalive and Cattaneo (2009) in their studies of the spillover effects of PROGRESA programme participation in Mexico. The authors highlight the social multiplier as one of the channels of the effect.

I also discover that parental education level may play an additional role. The positive influencers in my study are pupils whose parents have completed secondary education. I seek explanation for this finding in the fact that migrant parents differently influence their children and that this differentiated effect is then transmitted onto friends. The reasoning is linked with concepts of intergenerational transmission of human and cultural capital (Black et al., 2005; Black and Devereux, 2011; Holmlund et al., 2011). I draw on the idea that children's attitudes towards school and aspirations are highly correlated with those of their parents (Heckman and Rubinstein, 2001). I speculate that better educated parents value education differently and invest more in their children's schooling. This then may be further reinforced by their children's positive approach to performance at school.

3.3 Data and descriptive statistics

3.3.1 Overview of the sample

The data set used for the analysis is described in detail in Chapter 1. It contains information about 2822 16-year-old Polish pupils over a period of six semesters between September 2009 and June 2012, including individual and household characteristics, some socio-economic indicators, school performance and migration experience within the family.

Pupils are identified as children with parents working abroad (PWA) if they had at least one parent abroad at any point during the observed period. All respondents provided information about migration experience within the household but only 2669 gave a detailed account of its timing and were included in the analysis.

As observed in Chapter 1, 18% of respondents indicated having a parent abroad at some point during the 3 year period. Almost 26% of respondents indicated that their sibling has emigrated abroad. The migratory spells in the data are relatively short-term, repeated and occur over rather short distances. Households usually send just one parent abroad, in most cases the father.

Migrants from households in Opolskie are low skilled with 44% of mothers and 63% of fathers having finished vocational schooling and 36% of mothers and 29% of fathers

[10]Carrell and Hoekstra (2010) find that children exposed to domestic violence negatively influence their peers and Kristoffersen et al. (2015) conclude more generally that adding potentially disruptive children to a class lowers the academic achievement of peers.

high school.

3.3.2 Variable definitions

Children with parents working abroad (PWA)

I define **a PWA child** as a child who has had at least one parent abroad in a given semester and stayed in the home country during parental emigration experience. Given such a definition, one may have one or both parents abroad at the same time; moreover, a migrant parent may be absent in one semester and return to Poland in another and this change will be reflected in a change in the PWA child status.

Proportion of PWA children in the class

The information about migration within the family and its timing was used to construct the main explanatory variable for the analysis - the fraction of PWA pupils in the class. Specifically,

$$Fraction_{-ict} = \frac{M_{-ict}}{C_{-ict}} \tag{3.1}$$

where M_{-ict} is the number of pupils with a parent abroad in class c (excluding person i) in semester t and C_{-ict} is the total number of pupils in class c in semester t. By construction $Fraction_{-ict}$ varies over time but alternative, time-invariant specifications are also feasible. I discuss the options in Appendix 3.A.3.

School performance

The main dependent variable is **the grade of a pupil**. The grade is taken as an average over all courses taken in a given semester and ranges from 1 to 6, with 6 being a top mark awarded to a pupil for extracurricular achievement in the subject area. Pupils who mastered 100% of the curriculum in a given semester are usually awarded 5; 1 is a fail mark. The grade is awarded internally but based on the requirements of the national curriculum for a given year. The average grade in the sample has a mean of 3.61 and a standard deviation of .851.

Test scores in the national exam respondents took in the final semester of *gimnazjum* are another measure of academic performance. Prior to completion of *gimnazjum* and progression to the next education stage, pupils are tested in the following areas: Polish language and literature, history, maths, science and foreign languages. The exams are organised nation-wide by one Exam Board and blind-graded in percentage terms. Unfortunately, I only possess information about the exam results for under 13% of the sample, which is insufficient to use for the analysis. I will, however, rely on it for some background checks.

3.3.3 PWA pupils in the class and the pupils' grades

There is variation in the proportion of PWA children in the classes. The variable has a much lower mean than the overall number of PWA children in the sample would imply since it is based on the parental absence *in a given period of time* and varies across semesters.

The individual average grade is lower in classes where the fraction of PWA pupils is above the median, which indicates lower performance in these classes on average and is to be expected given worse average performance of the PWA students in the sample. The standard deviation of the variable is also lower indicating smaller variation across PWA-dominated classes. The density plots of the average grade in classes with below and above median concentration of PWA pupils differ mostly in upper tails, suggesting that the difference in performance comes from having less high scoring pupils and more average performers (See Figure 3.1).

The correlation between the academic performance of children and the fraction of PWA pupils in the class is almost zero.

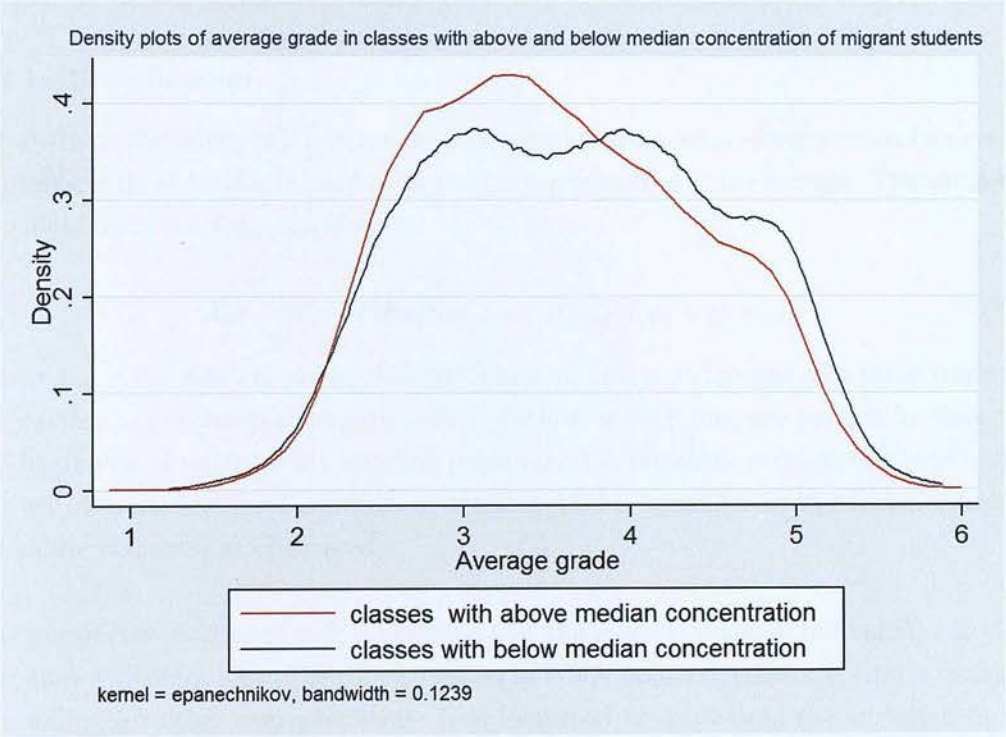


Figure 3.1: Distribution of average grades of respondents from classes with above median and below median proportion of PWA pupils
Source: MECP2012

Table 3.1: PWA children and school performance

Panel A: Fraction of PWA children at time t					
	mean	st.dev.	within variation	min	max
class level	.064	.075	.029	0	.454
Panel B: Average school performance					
	mean	st.dev.	within variation	min	max
individual average grade	3.610	.850	.280	1	5.88
Panel C : Classes with different proportions of PWA pupils					
	mean	st.dev.		min	max
average grade (below median)	3.677	.877		1.23	5.88
average grade (above median)	3.558	.824		1	5.87
Panel D : Correlations					
Corr(average grade, class fraction)					-0.043

Source: MECP2012

3.4 Empirical framework

I now present the empirical relationship explored in this chapter. I outline the preferred estimation equation, discuss problems related to the approach and how they are tackled.

3.4.1 Specification

I investigate the relationship between one's individual school performance and a number of pupils with at least one parent abroad as a proportion of one's class. The preferred specification is the following:

$$Y_{ict} = \alpha + \delta Fraction_{-ict} + \beta X_{ict} + \rho_t + \eta_c + \varepsilon_{ict} \tag{3.2}$$

where Y_{ict} is the average grade of individual i in class c at the end of a given semester t, $Fraction_{-ict}$ represents the proportion of students with migrant parents in class c at the beginning of semester t, excluding pupil i, and is the main variable of interest; X_{ict} is a set of individual level controls, η_c are class and ρ_t semester fixed effects. Standard errors are clustered at class level.

The parameter of interest is δ; it explains how the average grade of individual i in class c at time t changes when the concentration of PWA pupils in class c at time t changes, controlling for other characteristics. It is identified by exploiting the variation in the fraction of PWA pupils within the same class across different semesters, i.e. the change in the PWA fraction in each class over time.

I will now discuss concerns related to the estimation of spillover effects and how they are addressed in this specification.

Individual level controls

One may be concerned that certain characteristics of an individual may affect his performance at school and be correlated with the proportion of PWA pupils in the class. If so, failure to include them explicitly in the regression will result in coefficient δ reflecting not only the pure spillover effect but also the impact of those characteristics. Therefore, I include in the regression a series of individual level characteristics to control for pupils' observable personal or family traits which may influence their school performance.

I account for gender, as girls and boys are likely to perform differently at school and also be differently influenced by classmates.

I also include number of siblings as an explanatory variable as family size is deemed crucial for one's school attainment (Black and Devereux, 2011; Ginther and Pollak, 2004) and can also act as a proxy for one's socio-economic background. Moreover, based on the summary statistics of the data, families with migrants have on average more children. If classes are created in a non-random way and PWA children are grouped together, then the number of siblings may be correlated with the proportion of PWA classmates.

Given the lack of the household income variable, to proxy for the socio-economic background of students, the specification contains information about the parents' highest obtained education level and age. I expect children's performance to be correlated with parental education (Dickson et al., 2013). Further, the majority of migrant parents are low-skilled; if classes are created in a non-random way, parental education level may be correlated with the fraction of PWA pupils in the class. For example, sorting weaker pupils into one class may result in grouping many PWA children together.

Since the results in Chapter 2 suggest that parental migration influences child's school performance and because construction of $Fraction_{-ict}$ is based on the migration experiences of peers, I also incorporate the dummy variable indicating whether pupil i's parent was abroad in semester t. As argued in the previous paragraph, classes may be formed in a non-random manner, resulting in PWA children being grouped together. Therefore, the migration variable is likely to be correlated with both the pupil's average grade and the fraction of PWA peers in the class.

Role of the fixed effects

The material studied at school changes and becomes more difficult with time. Since the pupils' performance is tracked over a 3 year period, one may notice a change in pupils' grades which is attributable to the advancement in their studies and not to other circumstances. The semester fixed effects isolate the changes in grades over time which are common to all classes. There is, in fact, a clear pattern to the average grade over time in the sample (see Appendix 3.B). Each year there is a systematic improvement in pupils' grades in the second semester, when compared with the first semester of that

year. Further, the gap in grades between first and second semester in each year widens further into *gimnazjum*.

Class fixed effects are introduced to the specification to control for any *time-invariant* unobserved differences across classes. Such differences may be due to a range of factors, such as having different teachers, smarter or less able pupils in certain classes or different resources. If these differences persisted and were correlated with the proportion of PWA pupils in the class, failure to control for them would result in a biased estimate of δ; δ would capture the effect due to class composition as well as due to the class-specific features. Below I consider various reasons for which class fixed effects should be included in the regression.

Is the average grade a good measure?
Firstly, grades are awarded internally. The assessment of pupils against the national curriculum is at the teachers' discretion. Hence, pupils may be awarded different grades for comparable performance by different teachers. They may also be scored relative to their classmates. The situation is particularly problematic if teachers' assessment depends on the class composition, for example on the number of pupils with parents abroad. Therefore, the grades may be correlated within classes and across time.

However, I expect that teachers are consistent in the way they assess pupils over time; for example, a lenient teacher will remain lenient over the period of 3 years. If this is the case, the differences in average grades due to teachers' subjective assessment will be class-specific and time-invariant, and therefore captured by class fixed effects.

Nonetheless, I run a further check to ensure that the average grade is a satisfactory measure of school performance. I am in possession of the individual scores from the national tests respondents took at the end of *gimnazjum* for almost 13% of the sample. Even though the number of observations is insufficient to obtain robust results, I rerun the regression as specified above, replacing the average grade with test scores of pupils; the results, although statistically insignificant, imply a similar relationship between the concentration of PWA peers in the class and performance. Details can be found in Appendix 3.A.2.

Reflection problem and sorting
The causal interpretation of δ coefficient may also be challenged by the existence of the reflection problem. One may argue that, in peer effects analysis, the individual outcomes may reflect the behaviour of the peer group due to three different types of effects: endogenous, contextual and correlated (Manski, 1993). Importantly, it may be impossible to individually identify the endogenous (causal) effects in the reduced form linear analysis, as they are intertwined with the correlated effects.

The contextual effects are driven by the characteristics of the group one is a member of and, if they do not change over time, can be isolated by inclusion of control variables

or group fixed effects.

The correlated effects, linked to the fact that an individual and the group may behave similarly due to sorting or being in similar environments, pose a bigger estimation problem.

I have claimed already that the initial class allocation may be not random. Specifically, schools may group pupils on the basis of certain criteria, such as past performance or schools they came from. This is a reason for concern if pupils who perform comparably (e.g. either well or poorly) are grouped into classes together and if the proportion of PWA pupils in these classes is also high.[11] It is highly unlikely that migration status of the family is the main determinant of the initial class composition, though, as schools have limited knowledge of the household situation of applicants. Parents may also attempt to influence school's decisions, particularly if they would like their child to go to school with a certain group of peers. Nonetheless, if the class allocation is non-random at the point of enrollment, but does not change afterwards, it can be seen as a time-invariant characteristic of the class which will be controlled for by the class fixed effects. This is assuming that the impact the initial non-random allocation has on the class is constant over time.

Frequently the correlated effects problem effectively means that the dependent variable is pre-determined by the performance of the group and that the group's performance is also determined by the individual. The specification in this chapter limits the extent of the issue as I investigate the relationship between the number of PWA pupils in the class (not their performance as such) and the average grade of their classmates.

[11] Literature to date provides various solutions to the problem of non-random group composition, starting from randomisation of peer assignment and reliance on quasi-experiments (Guryan, 2004; Kugler et al., 2012; Sacerdote, 2000). If class assignment is random, there should not be any concerns about selection into groups. Such cases are, however, rare and arise as a result of an exogenous shock, e.g. reallocation of pupils due to a hurricane (Kugler et al., 2012), or of a specific experiment, although even then randomisation is often debated.

Another identification strategy is to rely on idiosyncratic variation in exposure of different cohorts to the influence within the same school (Hoxby, 2000; Hanushek et al., 2001; Gould et al., 2009). The method relies on the concept that, having controlled for the total number of migrant pupils in a school, their number in a given cohort is determined by random factors and hence conditioning on the variable removes a substantial portion of bias. Unfortunately, the exercise requires data for at least two cohorts within each school; I only observe one cohort across several schools over a period of 3 years.

A third solution is to minimise sorting bias by aggregation of the data to a higher geographical area, bringing the pupil allocation across the areas closer to random (Card and Rothstein, 2007; Evans et al., 1992; Cutler and Glaeser, 1997). The idea is that, even though students of differing abilities can sort to classes and schools within a city or county, they are less likely to do so across larger areas. An approach following Card and Rothstein (2007) is feasible with the data at hand but it would change the interpretation of the results and may not be a significant improvement on the fixed effects approach in this context. In particular, Card and Rothstein (2007) firstly aggregate the data to eliminate the sorting bias by averaging the outcomes of black and white students to the metropolitan level and then take across-race differences for each metropolitan area to further control for any across-city differences in average unobserved abilities of students which may be correlated with the control variables included in their regressions. Such a setup results in an analysis of a link between segregation and performance gap. I consider it unsuitable in this study as there is no strong evidence of segregation in Opolskie. An interpretation in terms of concentration of migrant pupils seems more intuitive. Given the argument that the majority of sorting takes place prior to enrolment, the fixed effects approach should deliver similar results to aggregation.

The proportion of pupils with migrant parents in the class is determined by individual parental decisions to engage in employment abroad, which are unlikely to be driven by or to directly affect other children's performance in the class. The channel of impact on the other children's performance in the class is via PWA children's behaviour.[12]

Overall, the use of class fixed effects allows to control for the pre-determined group characteristics, eliminating the time-invariant component of the reflection problem. A similar argument is put forward in academic peer effects studies of immigrant concentration and domestic violence (Schneeweis, 2013; Carrell et al., 2008).

Since I suspect that the grades are correlated within class and over time, I cluster the standard errors in the regression at the class level.

3.4.2 Some issues remain

The chosen specification does not cater for a scenario in which the unobserved characteristics, crucial for one's school performance, are time-varying and correlated with the explanatory variables explicitly included in the regression.

Therefore, the extent to which causality may be claimed is limited. I will discuss various such scenarios when providing alternative justifications for my findings in Section 3.6.

3.4.3 Alternative specifications

The list of individual controls included in the regression presented above is by no means exhaustive and many characteristics are not captured. Hence, an alternative specification may involve including individual level fixed effects, rather than a series of controls in the regression.

Individual level fixed effects would isolate all individual time-invariant characteristics which may impact school performance (such as intelligence, talent, etc.), whilst also controlling for class specific time-invariant characteristics. They are unable to control for personal circumstances which may change over time and influence pupils' performance. However, if they are uncorrelated with the proportion of PWA classmates, they should not influence the estimate of δ. It is impossible to include both class and individual level fixed effects as pupils do not change classes over time. I present results for both specifications to show that they yield similar results.

[12]For the issue of simultaneous determination of outcomes to arise in this context not only the proportion of PWA children in the class would have to influence a pupil's performance but also a pupil's performance would need to somehow affect the number of PWA peers. Although it is reasonable to think that one's child's performance may influence an individual decision to leave, it is unlikely that that child's peers' performance triggers migration within a family. Suppose, however, that such simultaneity arose and parental migration influenced and was in turn influenced by other pupils' grades. Then a positive δ would indicate that having good peers is correlated with a greater number of parents emigrating. It is difficult to imagine why good peers would encourage parental migration. One possible instance may be that parents are more likely to leave their children when they are not worried about the quality of teaching and their offspring's school performance; this is a more plausible scenario in classes with better performing pupils. Nonetheless, the case seems rather unusual and finds no support either in schools' perceptions of the migration phenomenon or the literature (Ryan and Sales, 2013).

The fact that some of the omitted variables key for the analysed relationship may vary over time is the biggest concern. In particular, the average grade at time t is likely to be largely driven by its historic values as students' performance is correlated over time. A lagged dependent variable specification, including lagged school performance, would capture this relationship and may help me control for some unobserved environmental factors which matter for performance, providing another alternative to the approaches discussed so far (Kristoffersen et al., 2015). Lagged dependent variable regressions produce similar results to those presented in this chapter, although many coefficients are statistically insignificant, most likely due to the sample size issues when the lagged average grade is included. The details can be found in the Appendix, Section 3.A.5.

3.5 Results

In this section I present outcomes of the empirical analysis just described. They are accompanied by results of differently specified regressions, to demonstrate the role played by the various elements I just examined. In particular, I include results of an ordinary least squares regression (OLS): 1) without any controls, 2) including individual and class level controls and semester fixed effects but no class fixed effects, 3) including individual rather than class fixed effects.

3.5.1 Baseline

In Table 3.1 I present the results for the relationship between the concentration of children with migrant parents in the class and the individual school performance.

The OLS coefficients in columns (1) and (2) are negative and become statistically insignificant once individual level controls and semester fixed effects are included in the regression. They suggest existence of a negative correlation between the concentration of PWA pupils in the class and respondents' academic performance, which reflects the findings in the summary statistics of worse average performance in classes with higher concentration of PWA children. However, as I mentioned, the estimates of δ may be biased due to unobserved time-invariant differences between individuals and classes, which impact the average grade and are correlated with the proportion of PWA pupils in the class.

I focus on results in columns (3) to (9) of Table 3.1, as they are obtained from regression specifications with either class or individual level fixed effects, gradually adding individual migration experience, other individual level controls and semester fixed effects. Irrespective of the exact controls included, these outcomes consistently suggest that an increased presence of PWA children in the class is associated with a higher average grade.

However, the effects derived from regressions with individual fixed effects, capturing individual as well as class time-invariant differences, are suggestive of a more modest

Table 3.1: Impact of concentration of PWA children in the class

	OLS			class FE			individual FE		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$Fraction_{-ict}$	-.720*	-.565	.410*	.325*	.393**	.282*	.328*	.322*	.174
	(.388)	(.344)	(.234)	(.187)	(.181)	(.156)	(.176)	(.175)	(.150)
Controls									
individual level migration	no	yes	no	yes	yes	yes	no	yes	yes
other individual controls	no	yes	no	no	yes	yes	no	no	no
semester FE	no	yes	no	no	no	yes	no	no	yes
class FE	no	no	yes	yes	yes	yes	no	no	no
individual FE	no	no	no	no	no	no	yes	yes	yes
No of observations	13842	10853	13842	13842	10853	10853	13842	13842	13842
No of respondents	2669	2070	2669	2669	2070	2070	2669	2669	2669
No of classes	159	159	159	159	159	159	159	159	159

Source: MECP2012
The dependent variable is the individual average grade at time t.
The main explanatory variable is the fraction of PWA pupils in a class at time t.
Other individual controls include gender, number of siblings, mother and father's age and education.
Standard errors are clustered at the class level and reported in parentheses.
Note: the difference in the number of observations in regressions which include individual controls is due to the fact that not all respondents provided the information. All individual level controls, except for the timing of migration experience, are time-invariant and drop out when individual FE are included.
Statistical significance levels *** - 1%, ** - 5%, * - 10%

impact, compared to those based on regressions with class fixed effects. Inclusion of semester fixed effects further lowers the estimates. Even though they range from .174 to .410, the coefficients from regressions with individual or class fixed effects are not statistically different from each other. Unfortunately, individually they are either marginally significant or insignificant.

According to the results in columns (3) to (9), a one standard deviation increase in the proportion of PWA children in the class (equal to .075) is correlated with a .013 to .031 increase in the average grade of a pupil, an equivalent of 1.53% to 3.65% of a standard deviation of the individual average grade. This implies that adding an extra PWA pupil to a class of 20 may be associated with a 1-2.8% increase in an average grade. This is similar to the effects reported by the class size literature (2-5% of a standard deviation (Angrist and Lavy, 1999)) but lower than the effect found by Carrell and Hoekstra (2010) in their study of the impact of disruptive children.[13]

As mentioned in the introduction to this analysis, a positive effect would suggest that PWA pupils are benefiting from parental migration experience and, through their improved school performance, influencing their peers. The positive impact of parental migration is likely due to the income gains from migration dominating any potential negative effect of family separation, which may be mitigated by the short-term and circular nature of parental departures.

The positive impact is non-trivial but smaller than the individual effect of em-

[13]Carrell and Hoekstra (2010) report a nearly 7% of st.dev. reduction in boys' test scores as a result of adding one troubled boy to a class of 20.

igration experience found in Chapter 2.[14] It may indicate a greater role of family experiences than peers in shaping academic outcomes.

3.5.2 Parental education matters

Given that a very high proportion of migrant parents (80% of mothers and 92% of fathers) only completed vocational or secondary education, it is plausible to expect that the effect is driven by a group of migrants with specific characteristics.

Parental education level, as well as the family socio-economic situation in general, are crucial for a child's educational attainment as human and cultural capital are transmitted across generations (Black et al., 2005; Black and Devereux, 2011). I have controlled for parental education as another factor influencing school performance, but migration experience of parents may impact children differently, depending on parental education level. If so, some groups of PWA pupils may become more influential than others.

For example, more educated migrant parents may be employed in better paying jobs relative to parents with a lower educational attainment if jobs require specific qualifications or knowledge of the language of the destination country, although many temporary migrants are likely to be underemployed. Further, better educated migrants assimilate quicker (Card and Rothstein, 2007), which may improve their foreign experience thanks to exposure to different cultures, a more diverse network of contacts and better access to the labour market. If better educated parents earn higher wages abroad, they are more likely to remit more in absolute terms and more money can be invested in child's well-being, including education.

Even if this is not the case, parents' priorities with regards to their children may differ, depending on their education level (Guryan, 2004). In particular, parents with higher educational attainment may see their children's education as very important and spend a higher proportion of income on schooling or take other steps to ensure their children perform well at school - work with them at home, etc.

If educated migrant parents assimilate better in the destination country and enjoy their experience, they may transfer some of the gained cultural capital onto their children, which may be beneficial to school performance. In Chapter 2 I provide some evidence demonstrating that these considerations are relevant in the context of Polish migration.

If better educated parents' migration experience is reflected to a greater extent in their children's improved school performance, then these children may be also more influential in interactions with peers. Equally, perhaps some parents' migration experience does not affect their children at all or does so in a negative way. Then their children's influence on peers may be negative or none.

[14] Given that in Chapter 2 I find significant delayed impacts of individual emigration experience on average grade, I specify alternative regressions with lags of individual migration as additional controls. It does not influence the coefficient of interest.

I run an alternative set of regressions, splitting the fraction of migrant students in the class according to education levels of their parents. The regressions mirror the approach discussed in Section 3.4 but now $Fraction_{-ict}$ is replaced with 4 different variables: $FractionElementary_{-ict}$, $FractionVocational_{-ict}$,$FractionSecondary_{-ict}$ and $FractionTertiary_{-ict}$ which are defined in the following way:

$$FractionX_{-ict} = \frac{MX_{-ict}}{CX_{-ict}} \tag{3.3}$$

where X = (Elementary, Vocational, Secondary, Tertiary), MX_{-ict} is the number of pupils with a parent abroad with the highest educational attainment X in class c (excluding person i) in semester t and CX_{-ict} is the total number of pupils whose parent has educational qualification X in class c in semester t. Alternative specifications, for example using mother's or father's education levels only, do not lead to different conclusions. This is partly due to high correlation in education levels of parents in the sample.

Table 3.2: Education-dependent concentration of PWA pupils in the class

	mean	st.dev.	min	max
$Fraction_{-ict}$	0.064	0.075	0	0.45
$FractionElementary_{-ict}$	0.051	0.192	0	1
$FractionVocational_{-ict}$	0.077	0.097	0	0.5
$FractionSecondary_{-ict}$	0.080	0.116	0	1
$FractionTertiary_{-ict}$	0.044	0.148	0	1

Source: MECP2012

The newly defined fractions reflect the structure of parental education found in initial statistics (Table 3.2). Fractions of PWA pupils whose parents have vocational or secondary education have higher means than the general fraction used initially.

Results of the analysis are presented in Table 3.3. The outputs are comparable across various regression specifications. However, the only estimates which are consistently statistically significant, irrespective of the regression specification, are the coefficients on $FractionSecondary_{-ict}$; This suggests that, among all PWA pupils, children of migrants, who graduated from high school, have positive impact on their peers. They constitute about 30% of the entire migrant group.

There is no strong evidence of influence by the most numerous group of PWA children, whose parents have vocational qualifications. Even though all specifications return negative coefficients, they are mostly statistically insignificant.

The coefficients on other fractions are statistically insignificant, which could be partly due to a much lower representation of parents in other educational groups.

Table 3.3: Impact of PWA pupils by parental education level

	OLS		class FE				individual FE		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$FractionElementary_{-ict}$	.176	.156	-.067**	-.065**	-.058**	-.008	-.049	-.050*	.010
	(.173)	(.125)	(.028)	(.028)	(.025)	(.038)	(.030)	(.030)	(.033)
$FractionVocational_{-ict}$	-.858***	-.528**	-.176	-.156	-.100	-.097	-.193	-.203	-.195*
	(.303)	(.282)	(.141)	(.141)	(.149)	(.124)	(.136)	(.137)	(.107)
$FractionSecondary_{-ict}$	.240	.040	.384***	.402***	.407***	.315**	.413***	.405***	.281**
	(.214)	(.197)	(.137)	(.141)	(.142)	(.121)	(.131)	(.130)	(.109)
$FractionTertiary_{-ict}$	.050	.218**	.128**	.132**	.118**	.030	.123**	.027	.022
	(.113)	(.098)	(.054)	(.055)	(.054)	(.045)	(.055)	(.025)	(.038)
Controls									
individual level migration	no	yes	no	yes	yes	yes	no	yes	yes
other individual controls	no	yes	no	no	yes	yes	no	no	no
semester fixed effects	no	yes	no	no	no	yes	no	no	yes
class fixed effects	no	no	yes	yes	yes	yes	no	no	no
individual fixed effects	no	no	no	no	no	no	yes	yes	yes
No of observations	13842	10853	13842	13842	10853	10853	13842	13842	13842
No of respondents	2669	2070	2669	2669	2070	2070	2669	2669	2669
No of classes	159	159	159	159	159	159	159	159	159

Source: MECP2012

The dependent variable is the individual average grade at time t.

The main explanatory variables are defined in this section - fractions of PWA pupils by their parents' education level.

Other individual controls include gender, number of siblings, mother and father's age and education.

Standard errors are clustered at the class level and reported in parentheses.

Note: the difference in the number of observations in regressions which include individual controls is due to the fact that not all respondents provided the information. All individual level controls, except for the timing of migration experience, are time-invariant and drop out when individual fixed effects are included.

Statistical significance levels *** - 1%, ** - 5%, * - 10%

Using the estimates from regressions with individual or class fixed effects, the coefficient on $FractionSecondary_{-ict}$ varies between .281 and .413; this implies that a one standard deviation increase in the proportion of PWA students, whose parents are high school graduates, results in a .033 to .048 increase in the individual average grade. This is equal to 3.88% to 5.65% of the average grade's standard deviation and significantly greater than the impact found in the previous section.

I make similar observations in Chapter 2. There the positive contemporaneous impact of parental emigration appears to be most pronounced in families, where parents have completed secondary education.

The fact that PWA children whose parents graduated from high school are the impact group suggests that the influence of migration differs, depending on the family situation. As mentioned earlier, better educated parents potentially have greater employment opportunities abroad, which may lead to larger remittances; they are also more likely to reap other benefits of migration to the full, thanks to faster assimilation and greater exposure to a different culture. They may also care about their children's education more than parents who have lower qualifications.

In such a case, arguably, I should have found an even bigger influence of PWA pupils, whose parents graduated from university. This group is, however, negligible in size and hence no significant effects emerged.

3.5.3 Role of migration background and gender

I also consider the roles one's family migration experiences and gender play in the class peer effect. PWA students may interact more with other PWA peers, in which case the spillover will be more pronounced within the group. Peer groups may also be formed around gender with boys interacting more frequently with boys and girls with girls.

I firstly look at the interaction between the family migration experience and then between gender and concentration of PWA pupils in the class. The results are presented in Table 3.4. In the first two columns one can find the outcomes for migration status and in the last two columns for gender. In both cases, I start from displaying results of a regression using $Fraction_{-ict}$ as the main explanatory variable (columns (1) and (3)). I then include output of a regression using $FractionSecondary_{-ict}$ as the explanatory variable, since I established that the PWA students with parents who graduated from high school are the influential group in this study (columns (2) and (4)).[15] I only report the results of regressions including class effects, individual level controls and semester fixed effects.

[15]Note that another possibility would be to include all 4 $FractionX_{-ict}$ as the main explanatory variables and interact them all with either individual migration experience dummy or gender. However, such a specification proves highly demanding on the data. See Appendix 3.A.4 for details.

Table 3.4: Differential impact by gender and migration status

	Impact by migration status		Impact by gender	
	(1)	(2)	(3)	(4)
Main explanatory variable:	$Fraction_{-ict}$	$FractionSecondary_{-ict}$	$Fraction_{-ict}$	$FractionSecondary_{-ict}$
fraction	.299*	.200*	.634*	.473**
	(.157)	(.116)	(.369)	(.223)
parent abroad at time t	-.180***	-.248***		
	(.062)	(.065)		
fraction*Parent abroad	.610	.640**		
	(.427)	(.277)		
female			.433***	.418***
			(.052)	(.046)
fraction*female			-.639	-.330
			(.553)	(.301)
No of observations	10853	10853	10853	10853
No of students	2070	2070	2070	2070
No of classes	159	159	159	159

Source: MECP2012

The dependent variable is the individual average grade at time t.

All specifications include individual level controls (gender, number of siblings, parental education and age, and parental emigration dummy), semester and class fixed effects.

The main explanatory variable in columns (1) and (3) is defined as a fraction of children in the class with at least one parent abroad at time t. In columns (2) and (4), the main explanatory variable is the fraction of PWA pupils in the class whose parents have graduated from high school.

Standard errors clustered at class level in parentheses.

Statistical significance: *** 1% ** 5% * 10%

From columns (1) and (2) of Table 3.4 I conclude that, even though all students benefit from the presence of PWA peers, PWA pupils benefit more in comparison with their non-PWA classmates. The effect is statistically significant only when I rely on the fraction of PWA students with parents who graduated from high school to explain the relationship, rather that the overall proportion of PWA students in the class.

In particular, the estimates using $Fraction_{-ict}$ as the main explanatory variable imply that a one standard deviation increase in $Fraction_{-ict}$ would increase an average grade by .299 (2.6% of a standard deviation) for PWA pupils and by .729 (6.43% of a standard deviation) for other pupils. Using $FractionSecondary_{-ict}$ as an explanatory variable, considering the same scenario as before, one would conclude that the impact equals .592 (or 8.08% of standard deviation of an average grade) for PWA pupils and .200 (or 2.73% of a standard deviation) for other pupils.

It is intuitive that PWA pupils gain more from having like peers in a class. They are likely to interact more with each other and to find parental emigration easier to brave if they can share the experience with others, who understand their situation very well.

I find no differential impacts of the concentration of PWA pupils in the class by gender.

3.6 Is this really a positive spillover?

The positive association between the proportion of PWA pupils in the class and the average grade may be due to peer effects. However, there may be other plausible explanations of such an outcome, such as a time-varying change which increases the average grades in the class and coincides with an increase in a proportion of PWA pupils in the class.

3.6.1 Resource and teacher allocations

One concern is that PWA students' performance triggers changes in the school environment.

For example, teachers may adjust their methods and expectations towards a class in response to its composition. It is a common preconception in Poland that parental emigration negatively influences children (Czeladko and Kopacz, 2008), even though there is no strong evidence that this is indeed the case. I already mentioned that the observed poor performance of PWA children may be due to selection rather than the impact of parental emigration. Nonetheless, if teachers think PWA children are disadvantaged and problematic, they may attempt to overcompensate the pupils. As a result, they may shift their attention solely to the PWA pupils or, more likely, become personally involved and dedicate more time and effort to classes with PWA pupils. I do not find strong evidence to suggest that any particular group of pupils explicitly loses out in academic terms. In most cases the effect I find is none or positive. Thus the

scenario of teachers focusing on PWA pupils is unlikely. The idea of teachers' greater involvement, which positively affects all pupils, is more persuasive. The scope for such a mechanism to play a role is limited as, to the best of my knowledge, currently there are no organised schemes to support teachers of classes with specific migration background. Hence, any efforts to help PWA pupils would be individual and independent. Nonetheless, I cannot eliminate such a scenario.

Alternatively, one may argue that, as a result of high concentration of PWA students in certain classes within school, the resources are being reallocated to support these particular groups. The increased investment of resources in PWA-dominated classes could improve pupils' performance but it is unlikely.

I do not possess specific information about the funding schools or classes receive, but the scenario is difficult to imagine, given the financing system of Polish schools. All local schools are financed by the county administration and receive funding per pupil, the value of which is established by the Ministry of Education. Any additional resources usually come from private sponsorship or the European Union funds. However, their acquisition is a lengthy process and therefore obtaining additional resources in response to the class composition in a cohort that is in the school for only 3 years is unlikely; especially as current school composition is not a perfect predictor of the migrant situation in future cohorts.

It is also questionable whether school management have the scope for unevenly spreading available financial resources across classes. If schools were to treat certain groups of pupils differently, they could achieve it in the following ways: 1) increase teaching hours for PWA students, 2) allocate better, more experienced teachers once they realise the class composition, 3) assign a support teacher, 4) split the groups they perceive as disadvantaged into smaller groups. All these measures are conspicuous and would provoke significant controversy among parents, whose children were not offered the additional facilities. They would also result in singling out of the PWA pupils, potentially introducing tensions in the school community. Should the situation arise, the Education Board overseeing schools is likely to be informed and object to differential treatment of pupils.

The feasibility of such changes is also debatable, given the short time pupils spend in *gimnazjum*. Firstly, schools do not know the situation within the classes *a priori*. In fact, they often do not realise the full extent of the migration situation in the class at all, as it is up to parents to inform the school about their employment abroad and many do not do so. Moreover, once the class is created, changes are difficult to introduce as they require a coordinated approach, affecting more than one group of students.

Given these considerations, I do not think reallocation of resources by schools, even though possible, could be occurring on a large enough scale to explain the results of my analysis.

3.6.2 Change in the local economy

Similarly, the class environment of a pupil could change if the region faced an economic downturn. This may affect both the parental migration decision and the availability and composition of teachers in the schools. Thus there may be a change in investments in children, and hence grades. However, there is no indication that the region was severely economically affected in the observed period of 2009-2012. Moreover, changes in economic conditions of destination countries did not discourage emigration from Opolskie and emigration flows remained steady over the period.

3.6.3 Average grade not a good measure?

Since the average grade may be a measure of relative performance, arguably if the performance of classmates changed in the observed period, it may also bias the estimated parameter of interest. In particular, if parental emigration worsened the school performance of PWA children and the teachers graded pupils relative to each other, then non-PWA children's performance will appear to have improved and would not be driven by the positive spillover. Such an interpretation assumes, however, that rather than lowering the PWA child's grades to reflect the worsening in performance, the teacher rewards the other pupils, whose performance did not change but is now better relative to their peers. If this was indeed the case, then I should have found a differential impact by PWA status in which the PWA children are negatively influenced by the like peers and non-PWA children benefit.

I have argued that I expect teachers to be consistent in their assessment over time so that any across class differences in grades due to the teachers' discretion are captured by the class fixed effects. However, it is possible that some teachers change their assessment over time, even with respect to the same group of students. For example, young teachers may become more or less lenient as they gain experience. If so, the grades in the classes they teach may change over time even if the pupils' ability has not changed. This is problematic only if such a change coincides with a change in the number of PWA pupils in the class. A significant proportion of teachers in classes with PWA pupils would have to fit into the category and would have to become more lenient over the period of three years for this scenario to drive the results I report. I have no way of checking the number of young teachers in the data and whether they are more likely to teach classes with PWA pupils but I suspect that the problem does not occur on a large scale.

3.6.4 Grade inflation

If schools do not adjust their behaviour, then another justification can be sought in grade inflation. The grade inflation literature is interested in establishing the drivers of different attainment of pupils with seemingly equal abilities. It is predominantly focused on tracking the changes in the value of grades over time (Jewell et al., 2013;

Jewell and McPherson, 2012; Oleinik, 2009) but within-cohort comparisons have also been made. Hinnerlich et al. (2011) investigate the grade difference between girls and boys in Sweden and find that blind grading substantially lowers the grades, suggesting that personal ties and non-blind grading may lead to grade inflation. Lavy (2008) claims that part of the gender difference in grades is due to statistical discrimination of male students in grading; according to his evidence, teachers' beliefs about girls performing better than boys affect non-blind grading.

The literature does not provide any intuition as to what would happen, if anything, in the context of migration. There are no publications analysing such relationships. Perhaps the positive impact of having children with migrant parents in the class is driven by the fact that teachers become more lenient towards classes where many pupils have parents abroad?[16]

I cannot investigate the case thoroughly due to data limitations but I run basic regressions, which may reveal existence of correlations in this field.

I have information about the individual exam performance in the national exams for just under 13% of the overall sample. In addition to that, for all schools in Opolskie (participating and not), I have the school average test scores from the national exams taken by the respondents, as these are publicly available.

Using the exam results and information about the average grades of pupils, I look at the claim that teachers in classes with a high concentration of children with migrant parents become more lenient.

In such a case I expect to see a smaller correlation between the grades awarded by teachers and the pupils' performance in the national exams. This is because grades are awarded internally, and hence subject to manipulation, whereas the national exam is taken by all pupils in the country and blind-graded. If grades are inflated, they will be reflecting the pupils' actual skills and knowledge to a lesser extent and the correlation with the exam results is likely to be lower.

The summary statistics for various controls measured at individual and school level can be found in Table 3.1.

[16] Another way of looking at this issue is to consider children's performance in the context of differences in assessment across schools. Dardanoni et al. (2009) look at grading standards across a sample of 16 countries and find that in all countries, except Ireland and the USA, there is conspicuous heterogeneity in standards across schools with evidence of grading on a curve, which means that grading standards increase with average competence of the school's students. Betts and Grogger (2003) find that higher standards raise test scores throughout the distribution of achievement but the increase is greatest toward the top of the test score distribution.
Perhaps schools with many PWA pupils lower their overall standards relative to the other schools as opposed to teachers selectively inflating grades of pupils they see as disadvantaged? This explanation is only plausible if the change in standards occurred within the observed period. Any time-invariant differences, including different standards across schools, would be captured by class fixed effects I include in the regression. Hence, the question to ask is how possible it is that many schools significantly change their standards over a period of 3 years. I consider it to be an unlikely scenario. Nonetheless, following the approach of Betts and Grogger (2003), I run regressions as in Equation 3.5, adding school fixed effects and comparing the coefficients on school fixed effects for signs of differential grading standards across schools. I find no evidence to support the claim that the differential grading standards in schools are driving the results.

Table 3.1: Summary statistics for students with test results

	mean	st.dev.	min	max
Individual level data*				
average grade	3.638	.853	1.4	5.8
literature test score	62.2	18.120	19	100
maths test score	44.743	23.352	7	100
fraction of PWA pupils in the class	.073	.080	0	.307
School level data				
average grade	3.599	.321	2.537	4.977
literature test score	63.923	6.116	53.7	88.6
maths test score	46.477	7.740	34.4	86.5
fraction of PWA pupils in the class	.067	.053	0	.308

Source: MECP2012
This table contains information about the average grade, test scores in literature and maths in the national exam, as well as the proportion of PWA pupils in the class for two groups: a 13% subsample of students for whom I have individual level results and the sample for which I have school average results of the national exams
* These statistics refer only to respondents for whom I have information about individual test results in the national exam at the end of *gimnazjum* they took in 2012.

Regressions using individual level data

The basic regression is:

$$AverageGrade_{icT} = \alpha + \beta Emigration_{iT} + \delta Fraction_{-icT} + \theta TestResult_{icT} + \psi Fraction_{-icT} * TestResult_{icT} + \lambda X_{iT} + \epsilon_{icT} \quad (3.4)$$

where $Emigration_{iT}$ indicates emigration experience within a family, X_{iT} contains individual level controls as used in all previous regressions in this analysis, $Fraction_{-icT}$ is defined as before and $TestResult_{icT}$ is a score a pupil obtained in the national exam.

Note that all these variables are measured at time T, the final semester of *gimnazjum*. This is because student performance was assessed externally only at one point in time, at the end of *gimnazjum*. I consider two different test scores: for the exam in literature and humanities and for the exam in maths and science. Pupils also wrote exams in languages, but results of these are less comparable as there is a choice of languages to be examined on. Therefore, I do not employ them in the analysis. I report the regression results in Panel A of Table 3.2.

I expect $\theta > 0$ in all cases, because better test results should always be positively correlated with a higher average grade. If teachers are lenient, however, $\psi > 0$ as, in case of grade inflation, out of two pupils with the same test score, the one from a class with more PWA children should have a higher grade.

Table 3.2: Grade inflation - regression results

Panel A: Individual level regressions

$$AverageGrade_{icT} = \alpha + \beta Emigration_{iT} + \gamma Fraction_{-icT} + \theta TestResult_{icT} + \psi Fraction_{-icT} * TestResult_{icT} + \lambda X_{iT} + \epsilon_{icT}$$

literature test score:	$\gamma = 3.183 * *$	$\theta = .026 * **$	$\psi = -.040*$
	(1.593)	(.003)	(.021)
maths test score:	$\gamma = 2.250*$	$\theta = .019 * **$	$\psi = -.030$
	(1.325)	(.002)	(.019)

Panel B: School level regressions

$$AverageGrade_{sT} = \alpha + \beta Fraction_{sT} + \theta TestResult_{sT} + \psi Fraction_{sT} * TestResult_{sT} + \gamma X_{sT} + \epsilon_{sT}$$

literature test score	$\beta = 10.127 * **$	$\theta = .044 * **$	$\psi = -.167 * **$
	(.326)	(.0003)	(.005)
maths test score	$\beta = -2.330 * **$	$\theta = .029 * **$	$\psi = .043 * **$
	(.316)	(.0002)	(.007)

Source: MECP2012
Standard errors clustered at a class or school level in parentheses.
Statistical significance: *** 1%, ** 5%, * 10%

This is not the case with $\theta \in [.019, .026]$ and $\psi \in [-.040, -.030]$, both statistically significant. The results suggest that the better one's test score, the better also his average (as expected) but less so in classes with higher concentration of children with parents employed abroad. In particular, a pupil who scored equally to someone else, but comes from a class with 1 standard deviation higher proportion of PWA children, will have .003 lower average grade. It is a negligible impact, equivalent to only .353% of a standard deviation in the average grade.

Regressions using school level data

The analysis above is based on a small number of observations. I now look at analogous regressions at a school level, since school average test results for the national test are readily accessible, allowing me to compare schools in the entire region.

$$\begin{aligned} AverageGrade_{sT} = \alpha + \beta Fraction_{sT} + \theta TestResult_{sT} \\ + \psi Fraction_{sT} * TestResult_{sT} + \gamma X_{sT} + \epsilon_{sT} \end{aligned} \tag{3.5}$$

where s stands for school and T is the final observed semester. Now $AverageGrade_{sT}$ is the average grade of all pupils in a given school s in the last semester, T, $Fraction_{sT}$ is the concentration of PWA pupils in school s at time T and $TestResult_{sT}$ is the average test score for school s at time T. I also include school level controls, which are based on the individual and family characteristics of the respondents within a given school, averaged across the entire school.

The results are presented in Panel B of Table 3.2 and seem in line with the regression outcomes at the individual level. All else equal, among schools which have the same average test scores in literature, those with a 1 standard deviation higher proportion of PWA pupils have .008 lower average grades (1% of a standard deviation of an average grade). However, when maths scores are used for the comparison, the schools have .002 higher average grades. In both cases the impact is negligible.

3.6.5 Non-respondents

One further worry is related to potential selection into the study. In Chapter 1 I argue that the sample is representative of the population of interest. One of the reasons is the high response rate among the targeted pupils. I find, however, that the school performance of pupils who did not respond to the survey is worse than that of the respondents. The difference is statistically significant.

This is problematic if the non-response to the survey is not random. In particular, one may be concerned that due to the sensitive nature of the survey PWA pupils were more likely to refuse participation in the study. If so, they may be overrepresented in the group of non-respondents. This would result in undercounting of the number of PWA children in the class and mismeasurement of $Fraction_{-ict}$. Unfortunately, I cannot establish how many non-respondents had a parent abroad.

The matter would be particularly troubling if the outcomes of the PWA pupils not captured in the data deteriorated due to parental emigration and, in turn, the students exerted negative influence on their peers which would be unaccounted for. Then the positive association I find would be upward biased. I cannot, however, establish if parental migration is the driver of the lower grades among the non-respondents. Given that migrant parents are negatively selected, any worse performance of potential PWA non-respondents could be due to selection as well as other factors.

Although I do not have any background information about the non-respondents as they have not participated in the survey, I know their grades over the three year period. Thus, even though I cannot correct $Fraction_{-ict}$, I can check whether there is a differential relationship between the performance of non-respondents and the presence of PWA children in the class. Inclusion of the observations for non-respondents into the sample does not change the results of the analysis.

3.6.6 Students' responses to class composition

Students may have actively responded to the migration situation once the class has been put together and, as a result, changed the class composition before the survey was conducted. In such a case there would be an unmeasured change triggered by the presence of the PWA pupils in the class.

Failure to accurately record class composition and pupils grades should not affect the results if changes in class register are rare and random. However, if they are linked to the studied phenomenon and occur rather systematically, then the estimated relationship between the concentration of the PWA pupils in the class and the average grade may be biased.

I do not know exactly what happens to pupils who disappear from the class register and I have no information about their family background, since they did not participate in the survey. However, for most of them, I am in possession of the school performance record up to the point of dropping out. This allows me to look closer at the problem, its scale and whether it seems to be related to the class composition. I start by outlining various circumstances which may explain a pupil disappearing from the register. Then I relate these to the data to assess how likely they are in this context.

The following scenarios may be behind pupil drop-outs:

1. Some good students may choose to change class or school. This may be for various personal reasons unrelated to the situation in the class (e.g. moving away) or, in the worst case for the study, because they do not want to be in the class with peers who have migrant parents. If this was the case, it would most likely indicate that they are concerned about the influence of PWA peers, given the general perception of them being made worse off by their parents' departure.
   If the decision is only driven by the perception of negative impact and not many

pupils change classes, it does not pose an estimation problem. If, however, the pupils changed the class because their performance was actually negatively affected, the estimated positive impact is greater than it ought to be, as I fail to account for the group who dropped out.
If, on the other hand, the PWA peers actually exerted positive influence on dropouts, I would underestimate the effect. This is an unlikely scenario, as I suspect that pupils would not change classes if their grades were improving.

2. Pupils, who disappear from register, simply fail the year and do not consciously choose to change the class. This case proves problematic for the estimation of the spillover effect if the failure to pass the year is linked to the class composition. It may be that the poor performing students who drop out would have passed if they were not negatively influenced by their PWA peers. In such a case, again, the estimates reported so far will not account for the negative impact pupils with migrant parents had on certain classmates, prior to them dropping out.
Given that I establish a positive impact for the remaining pupils in the class, such an argument would imply that students whose parents work abroad influence their classmates in various ways - some benefit from their presence and some, most likely already poor performers, are so badly affected they fail the year.

3. A further complication occurs if pupils, who disappear, were PWA children themselves. Then they may have been influencing their peers prior to dropping out and the impact is not captured.
If they were a good influence before they left, the initial improvement may be attributed to a smaller number of PWA peers than were actually present in the class and lead to an overestimate of the impact.
If the dropping out PWA students were a bad influence, which is likely if they failed a year due to their poor performance at school, then the negative effect would drive the reported estimate down.
If PWA students disappear from register due to their poor performance and are suspected of having exerted negative impact on their peers, this would imply that the group is diverse and some PWA pupils, potentially those already performing poorly at school, may be negatively influenced by their parents' migration, whilst others benefit from it.

I now consider these scenarios in turn and discuss the likelihood of them playing a role in this context. Overall there is a record of 229 students dropping out of the participating classes, 109 joining them and 10 transferring between them. For some of the pupils I was able to establish reasons for departure or arrival (see Table 3.3).

Scenario 1. Good students change school because of the PWA classmates. Looking at Panel A in Table 3.3 it is apparent that the majority of dropouts from the

register are due to poor performance of pupils, rather than their conscious decision to change the class. Only 67 pupils disappear for unknown reasons. Pupils who disappear from the register obtained a much lower average grade than the respondents of the survey and there is a significant correlation between being a drop-out and the average grade (Corr(drop-out, average grade)=.327). To further check whether the disappearance from register is related to poor school performance, I run probit regressions of the fact of dropping out on the average grade of the pupils and find a negative, statistically significant coefficient (see Panel B, Table 3.3).

Scenario 2. Dropouts occur due to poor performance caused by PWA peers
If poor performing students, who drop out of the class, were negatively affected by their PWA peers, I should find correlation between the fact of dropping out and the concentration of pupils who have migrant parents in the class.

Firstly, there is almost no correlation between dropping out and the proportion of PWA pupils in the class (Corr(drop-out, $Fraction_{-ict}$)=-.033).

To further check the relationship, I run a probit regression of the fact of dropping out from class register on the proportion of PWA pupils in the class, class and semester fixed effects and find no relationship; the coefficient is close to zero, negative and statistically insignificant, indicating that $Fraction_{-ict}$ did not influence the disappearance from register (See Panel C, Table 3.3).

I also find no evidence of a relationship between the average grade of pupils who drop out and the concentration of PWA pupils in their class (see Panel D, Table 3.3)

Scenario 3. Dropping out pupils have parents abroad
Unfortunately, since pupils who disappear from the register did not participate in the survey, I do not know their individual migration situation and hence I am unable to exactly estimate the number of such cases or provide any information on their school performance.

However I can look at respondents who are older than their peers, because they repeated a year, and establish how many among them have a parent abroad. This information may shed light on the potential number of pupils among the dropouts likely to be PWA children.

As can be seen in Table 3.4, there are 94 pupils in the sample (3%) who repeated the class at least once and, among those, 17% declared having a parent working abroad, which is below the average for the overall sample. Thus, it is unlikely that PWA pupils are overrepresented among the dropouts and that failure to include them in the study significantly affects the estimates.

It is possible, though, that PWA students dropping out of register leave the class for reasons unrelated to their class performance. In particular, they may be moving abroad to join the migrant parent. In this case, what I know about PWA pupils who repeated a year is not informative as their number does not shed light on how many

Table 3.3: Information about pupils who drop out

Panel A: Changes in class composition - scale of the problem

total number of registered students			3423
number of surveyed students			2822
Students who:			
dropped out			229
transferred to another class in the same school			10
joined the school			109
Reasons for dropping out or joining the school:			
Dropping out		Joining the school	
failed a year	159	transferred from another school	3
transferred to another school	1	do not know why	106
went abroad	1		
died	1		
do not know why	67		

Panel B: The fact of dropping out and school performance

$$Prob(Dropout_{ict}) = \alpha + \beta Average_{ict} + \theta_c + \delta_t + \epsilon_{ict}$$
$$\beta = -.067 * **$$
$$(.005)$$

Panel C: The fact of dropping out and concentration of PWA children in the class

$$Prob(Dropout_{ict}) = \alpha + \beta Fraction_{-ict} + \theta_c + \delta_t + \epsilon_{ict}$$
$$\beta = -.002$$
$$(.474)$$

Panel D: School performance of drop-outs and concentration of PWA children in the class

$$Average_{ict} = \alpha + \beta Fraction_{-ict} + \theta_c + \delta_t + \epsilon_{ict}$$
$$\beta = -.004$$
$$(.845)$$

Source: MECP2012
In the regressions in Panel B - Panel D, $Dropout_{ict}$ is a dummy variable equal to one if a student disappears from register at any point in the observed period, $Average_{ict}$ is the average grade of individual i in class c at time t, $Fraction_{-ict}$ is defined as throughout the analysis, θ_c is a set of class fixed effects and δ_t is a set of semester fixed effects.
Standard errors clustered at class level in parentheses.
Statistical significance: *** 1%, ** 5%, * 10%

may have left the country. Among the students who dropped out of the sample only one indicated moving abroad (see Panel A, Table 3.3). However, for 67 dropouts the cause of changing school is unknown.

Judging by the information for Opolskie about the overall number of pupils born in 1996 (respondents' age) who left the country or returned from abroad between years 2002 and 2011 (See Panel B, Table 3.4), the percentage of PWA pupils dropping out to move abroad is likely to be very small.

Table 3.4: Migrants among dropping out pupils

Panel A: Migration situation of respondents by birth year

	Went to school early	Started school on time	Failed at least one year			no info	total
born in	1997	1996	1995	1994	1993		
number of pupils	15	2413	81	12	1	300	2822
migration in general	3	692	39	7	0	69	809
parental migration	1	285	14	2	0	25	327
sibling migration	2	129	7	0	0	9	147
Summary for the group of older students:							
		no of pupils		94	100%		
		parent abroad		16	17.02%		
		sibling abroad		7	7.45%		
		emigration in general		45	47.87%		

Panel B: International migration of children born in 1996 from Opolskie

Year	2002-2005	2006	2007	2008	2009	2010	2011
emigration	54	24	25	28	13	14	25
immigration	16	2	10	3	7	2	4

Panel A data source: MECP2012

Notes to Panel B:

This data come from the Central Statistical Office of Poland and reflect the number of children born in 1996 registered as arriving from abroad or leaving for abroad. The flows are approximated on the basis of family registering at or deregistering from an address in Opolskie and stating that they are moving abroad.

Another check on the impact of the failure to account for the dropping out students may be to include all the information about their school performance into the data set and repeat the analysis. This way the information about the class size as well as students' performance across semesters is more accurate. It is not ideal, however, as the migration situation of the dropping out students is unknown and hence it is not reflected in $Fraction_{-ict}$. The regression estimates (see Table 3.5), although statistically insignificant, are in line with those reported in Section 3.5, suggesting a positive effect of the concentration of PWA pupils in the class, even once dropouts are included.

Table 3.5: Baseline regressions including drop-outs' outcomes

	(1)	(2)
$Fraction_{-ict}$	.218	.284
	(.164)	(.244)
Controls:		
semester FE	yes	yes
class FE	no	yes
individual FE	yes	no
N	15239	15239
clusters	159	159

Source: MECP2012
Note that these regressions include observations for respondents as well as the pupils who dropped out of register prior to the survey. Since the drop-outs were not surveyed, no individual level information is available for them. Thus, the regressions, although mirroring those relied upon in the baseline analysis, do not include individual level controls.
Standard errors clustered at class level in parentheses.
Statistical significance: *** 1%, ** 5%, * 10%

3.7 Conclusions

I analyse the relationship between the proportion of children with migrant parents in a class and the academic achievement of its pupils. To the best of my knowledge, this is a first attempt of this kind in the literature. Given the particular nature of migration in the data, which differentiates the migration experiences of families in Poland from those in traditional sending countries and which I believe is becoming common within the borderless European Union, its outcomes may be informative for policy-makers in Poland, as well as other new or candidate member states.

The unique structure of the data allows me to exploit the variation in the proportion of PWA peers within classes over time which eliminates various estimation concerns related to the peer effect analysis.

The outcomes indicate the existence of a positive relationship between the overall proportion of children with working parents abroad in the class on academic performance of their peers. Further analysis reveals that PWA children whose parents are high school graduates are the driving force behind the association. I also establish that, even though all pupils benefit, those, who themselves have a parent abroad, gain most from having like peers in the class. I find no evidence of differential impacts by gender.

The results are in line with the research on individual level impacts of parental emigration in Chapter 2 and suggest that, in general, individual experiences are more

deterministic for academic performance than influence of the peer group. The positive effect, although counter-intuitive, may be thanks to the short-term, circular nature of parental migration in the sample, which lowers the burden on PWA children and potentially more effectively channels positive aspects of international experiences, such as increased income, exposure to other cultures and possibly changed perception of returns to education.

It is also plausible for PWA children of high school graduates to benefit most from their parents' migration experience and become the influencers in the group. Sufficiently educated parents may reap greater benefits of migration, including higher income and cultural gains, which they can pass onto their children. They may also value their children's education more highly and ensure that their children perform well at school despite their departure for abroad.

Despite various possible explanations, it seems most plausible that the positive influence may be caused by a genuine beneficial impact of the interaction with PWA children in the class, who are more driven and motivate their peers. It is also tenable, however, that teachers change their ways of teaching and adjust to pupils' needs when they realise the proportion of PWA children in the class.

Given the choice of the study area, questions may arise regarding the degree of external validity of these findings. Compared to the rest of Poland, Opolskie does have a very unique history of steady migration for employment over relatively short distances. However, it does not differ significantly from the country average in terms of its economy (The Central Statistical Office of Poland, 2013b). More importantly, students from the area have been performing comparably to the country average in national tests, since they were introduced in 2002 (Centralna Komisja Edukacyjna, 2002-2012). Despite these similarities one could argue that the commonality of migration in the area may mean that children react differently to the migration experience, seeing it as a norm, and hence the group is not representative of a broader population.

Secondly, the outflow from Opolskie can be described as steady, unlike the sudden increase in emigration across other areas of the country following the EU enlargement and opening of the foreign labour markets to Polish workers in 2004. However, if the migration outflow from other areas of Poland is maintained in the future, they may match Opolskie in migration characteristics.

It therefore follows that the results I present may be valid for countries or areas characterised as developed or middle-income which experience steady migration outflow but where migrants engage in short-term, circular and legal employment abroad over relatively short distances. As a result children in such families experience the negative aspects of parental departure to a lesser extent and reap a greater share of its benefits. The setup clearly differs from the situations considered so far, e.g. Mexican migration to the US, but may be closer to the new European reality, particularly if the migration from the new member states stabilises at a certain level and is sustained, following the initial shock.

Bibliography

AIZER, A., AND F. CUNHA (2012): "The production of human capital: endowments, investments and productivity.," NBER Working Paper Series 18429.

AMUEDO-DORANTES, C., ET AL. (2008): "Migration, remittances and children's schooling in Haiti," IZA Discussion Paper 3657.

ANGRIST, J., AND V. LAVY (1999): "Using Maimonides' Rule To Estimate The Effect Of Class Size On Scholastic Achievement," *The Quarterly Journal of Economics*, 114, 533–575.

ANGRIST, J. D., AND K. LANG (2004): "Does school integration generate peer effects? Evidence from Boston's Metco program," *American Economic Review*, 94, 1613–1634.

ANGRIST, J. D., AND J.-S. PISCHKE (2009): *Mostly harmless econometrics*. Princeton University Press, Woodstock, Oxfordshire.

ANTMAN, F. (2013): "The Impact of Migration on Family Left Behind," in *International Handbook on the Economics of Migration*, ed. by A. F. Constant, and K. F. Zimmermann, pp. 293–309. Edward Elgar, Northampton, MA.

APPS, P., ET AL. (2012): "The impact of pre-school on adolescents' outcomes: evidence from a recent English cohort," IZA Discussion Paper 6971.

BARRO, R. J. (2001): "Human Capital and Growth," *American Economic Review*, 91(2), 12–17.

BEHRMAN, J. R., ET AL. (2006): "What Determines Adult Cognitive Skills? Impacts of Pre-Schooling, Schooling and Post-Schooling Experiences in Guatemala," *Latin American Economic Review*, 23(4).

BETTS, J. R., AND J. GROGGER (2003): "The impact of grading standards on student achievement, educational attainment, and entry-level earnings," *Economics of Education Review*, 22, 343–352.

BLACK, S., AND P. DEVEREUX (2011): "Recent Developments in Intergenerational Mobility," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 4. Elsevier.

BLACK, S., ET AL. (2005): "Why the apple doesn't fall far: understanding intergenerational transmission of human capital," *American Economic Review*, 95, 437–449.

——— (2013): "Under Pressure? The Effect of Peers on Outcomes of Young Adults," *Journal of Labor Economics*, 31, 119–153.

BOBONIS, G., AND F. FINAN (2009): "Neighborhood peer effects in secondary school enrollment decisions," *The Review of Economics and Statistics*, 91, 695–713.

BRUNELLO, G., AND L. ROCCO (2013): "The effect of immigration on the school performance of natives: cross-country evidence using PISA test scores," *Economics of Education Review*, 32, 234–246.

BURKE, M. A., AND T. R. SASS (2006): "Classroom Peer Effects and Student Achievement," Public Policy Discussion Papers No. 11-5.

CARD, D., AND J. ROTHSTEIN (2007): "Racial segregation and the black-white test score gap," *Journal of Public Economics*, 91, 2158–2184.

CARRELL, S., AND M. HOEKSTRA (2010): "Externalities in the classroom: How children exposed to domestic violence affect everyone's kids," *American Economic Journal: Applied Economics*, 2, 211–228.

CARRELL, S. E., ET AL. (2008): "Does your cohort matter? Measuring peer effects in college achievement," NBER Working Paper Series 14032.

CENTRALNA KOMISJA EDUKACYJNA (2002-2012): "Osiągniecia uczniów kończących gimnazjum," Discussion paper, Warsaw.

CHEN, X., ET AL. (2009): "Effect of Migration on Children's Educational Performance in Rural China," *Comparative Economic Studies*, 51, 323–343.

CUNHA, F., ET AL. (2010): "Estimating the technology of cognitive and non-cognitive skill formation," NBER Working Papers 15664.

CUNHA, F., AND J. HECKMAN (2007): "The Technology of Skill Formation," *The American Economic Review*, 97(2), 31–47.

CUTLER, D. M., AND E. L. GLAESER (1997): "Are ghettos good or bad?," *The Quarterly Journal of Economics*, 112(3), 827–872.

CZELADKO, R., AND K. KOPACZ (2008): "Pokolenie eurosierot," Rzeczpospolita, 22.01.2008, http://www.rp.pl/artykul/86021.html.

DARDANONI, V., ET AL. (2009): "Grading across schools," *The B.E. Journal of Economic Analysis and Policy*, 9, 1–14.

DICKSON, M., ET AL. (2013): "Early, Late or Never? When Does Parental Education Impact Child Outcomes?," IZA Discussion Paper no. 7123.

DUSTMANN, C., AND A. GLITZ (2011): "Migration and Education," in *Handbook of the Economics of Education*, ed. by Hanushek, Machin, and Woessman, vol. 4, chap. 4, pp. 327–441. Elsevier.

EVANS, W. N., ET AL. (1992): "Measuring peer group effects: a study of teenage behaviour," *The Journal of Political Economy*, 100, 966–991.

FEINSTEIN, L. (2003): "Inequality in the Early Cognitive Development of British Children in the 1970 Cohort," *Economica*, 70, 73–97.

GINTHER, D. K., AND R. A. POLLAK (2004): "Family structure and children's educational outcomes: blended families, stylized facts, and descriptive regressions," *Demography*, 41(4), 671–696.

GOULD, E. D., ET AL. (2009): "Does immigration affect the long-term educational outcomes of natives? Quasi-experimental evidence," *The Economic Journal*, 119, 1243–1269.

GURYAN, J. (2004): "Desegregation and black dropout rates," *American Economic Review*, 94, 919–943.

HANUSHEK, E. A., ET AL. (2001): "Does peer ability affect student achievement?," NBER Working Paper 8502.

HANUSHEK, E. A., AND L. WOESSMANN (2009): "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation," IZA Discussion Paper 4575.

HECKMAN, J., AND Y. RUBINSTEIN (2001): "The importance of noncognitive skills: lessons from the GED testing program," *The American Economic Review*, 91, 145–149.

HINNERLICH, B. T., ET AL. (2011): "Are boys discriminated in Swedish high schools?," *Economics of Education Review*, 30, 682–690.

HOLMLUND, H., ET AL. (2011): "The causal effect of parents' schooling on children's schooling: a comparison of estimation methods.," *Journal of Economic Literature*, 49, 615–651.

HOXBY, C. (2000): "Peer effects in the classroom: Learning from gender and race variation," NBER Working Paper no. 7867.

JEWELL, R. T., ET AL. (2013): "Whose fault is it? Assigning blame for grade inflation in higher education," *Applied economics*, 45:9, 1185–1200.

JEWELL, R. T., AND M. A. MCPHERSON (2012): "Instructor-specific grade inflation: incentives, gender and ethnicity," *Social Science Quarterly*, 93:1, 95–109.

KRISTOFFERSEN, J., ET AL. (2015): "Disruptive School Peers and Student Outcomes," *Economics of Education Review*, 45, 1–13.

KUGLER, A. D., ET AL. (2012): "Katrina's children: evidence on the structure of peer effects from hurricane evacuees," *American Economic Review*, 102(5), 2048–82.

LALIVE, R., AND A. CATTANEO (2009): "Social interactions and schooling decisions," *The Review of Economics and Statistics*, 91, 457–477.

LAVY, V. (2008): "Do gender stereotypes reduce girls' human capital outcomes? Evidence from a natural experiment," *Journal of Public Economics*, 62, 2083.

MANSKI, C. (1993): "Identification of Endógenous Social Effects: the Reflection Problem," *Review of Economic Studies*, 60, 531–542.

OECD (2008): "International Migration Outlook 2008," OECD Publishing.

OHINATA, A., AND J. VAN OURS (2013): "How immigrant children affect the academic achievement of native Dutch children," *The Economic Journal*, 123, F308–F331.

OLEINIK, A. (2009): "Does education corrupt? Theories of grade inflation," *Educational Research Review*, 4, 156–164.

RYAN, L., AND R. SALES (2013): "Family Migration: The Role of Children and Education in Family Decision-Making Strategies of Polish Migrants in London," *International Migration*, 51(2), 90–104.

SACERDOTE, B. (2000): "Peer effects with random assignment: results for Dartmouth roommates," Discussion paper, NBER Working Paper 7469.

SCHNEEWEIS, N. (2013): "Immigrant concentration in schools: Consequences for native and migrant students," IZA Discussion Paper 7230.

THE CENTRAL STATISTICAL OFFICE OF POLAND (2013a): "Migracje zagraniczne ludności. Narodowy Spis Powszechny Ludności i Mieszkań 2011," .

——— (2013b): "Vademecum Samorządowca 2013. Województwo Opolskie.," Discussion paper, Warsaw.

TYNELSKI, A. (2010): "Sytuacja dziecka w rodzinie migrującej w Polsce," Discussion paper, Ministerstwo Edukacji Narodowej.

ZIMMERMAN, D. (2003): "Peer effects in academic outcomes: evidence from a natural experiment," *The Review of Economics and Statistics*, 85, 9–23.

Appendix

3.A Regression analysis

3.A.1 Analysis including observations without full migration history

The original dataset contains 2822 observations but only 2669 are used in the analysis due to lack of detailed information about the migration experience of 153 (5% of the sample) families.

This can influence results by:

1. lowering the estimated class sizes
2. potentially leading to an underestimate of the number of emigrants in the class
3. changing the distribution of grades in the class

Despite the lack of detailed history of migration within these families, a lot is known about the excluded group. Therefore, I provide summary statistics and a set of counterfactual regressions to argue that the decision to exclude these observations does not affect the validity of the results. I possess information about the number of PWA pupils in the excluded sample, as they have stated whether the family has experienced emigration during the sampled period; they failed to provide timing for the migratory movements. As can be seen in Table 3.A.1 the percentage of PWA pupils in the group is significantly higher than in the rest of the sample. Moreover, the excluded pupils have lower average grade relative to the rest of the sample.

However, given that there are 159 classes in the sample, excluding 153 observations may not affect the results significantly, if the impact is spread evenly across all classes and not highly concentrated in particular groups. Having investigated the data, I find that these observations are spread across 92 different classes, i.e. just over 1.5 per class.

To check whether inclusion of these observations changes the regression results, I redefine the main explanatory variable: $Fraction_{-ict}$. I adjust the class size by including the additional observations bringing it to a correct level. At the same time I adjust the numerator in the following way: PWA pupils from the main sample are still included in the numerator if they had at least one parent abroad at time t (time-varying element) but since I do not possess the same information about the pupils who were excluded from analysis, I use information about any migration experience in the family and assume that if a child stated a parent was abroad in the observed period, he/she was absent for all 6 semesters. This way I am likely overestimating the number of PWA pupils in the class at any time t, especially given that for majority of pupils the parent

Table 3.A.1: Summary statistics for the excluded group

Panel A: Migration situation			
	Absolute value	Percent	Percent
total sample (n)	153	100	
migrants (incl. sibling)	125	81.70	100
migrant parents - total	99	64.71	79.2
Who emigrated:			
only father	65	42.48	52
only mother	17	11.11	13.6
mother and father emigrated	17	11.11	13.6
Panel B: Student performance			
		mean	st.dev.
average grade		3.401	.830
Panel C: Fraction of PWA students			
old fraction		.064	.075
new fraction		.067	.078

Source: MECP2012

is abroad only for part of the 3 years. As can be seen from Panel C of Table 3.A.1, there is not a drastic change in the mean and standard deviation of the new variable, relative to the original one used in the analysis.

I repeat the baseline regressions using the newly defined fraction. The results are presented in Table 3.A.2 and are in line with the baseline results in this chapter.

Table 3.A.2: Regression of the average grade on newly defined fraction

	(1)	(2)	(3)	(4)	(5)
$NewFraction_{-ict}$	-.688*	-.409	.321**	.366**	.214
	(.375)	(.318)	(.153)	(.171)	(.148)
Individual level controls	no	yes	yes	no	no
Semester FE	no	yes	yes	no	yes
Class FE	no	no	yes	no	no
Individual level FE	no	no	no	yes	yes
n	2810	2203	2203	2810	2810

Source: MECP2012
Note: the dependent variable is the individual average grade. The main explanatory variable has been redefined to include students who did not provide information about the detailed family migration history. Regressions run here are as specified in the empirical framework. Standard errors are clustered at class level.
Statistical significance of coefficients: *** 1%, ** 5% , * 10%

3.A.2 Regressions using national exam scores instead of average grades

The analysis relies on the average grade as a dependent variable. The average grade, however, is awarded internally and may not objectively reflect pupils' skills. To check whether the average grade is a satisfactory measure of school performance, I rerun the

regressions outlined in Section 3.4 using the national exam results of almost 13% of respondents.

I have information about pupils' results in exams in the following subject areas: literature, history, math, science and foreign languages. The average grade used in the analysis is an average over all courses taken by a pupil, which include the examined subject areas. Therefore, to make the two measures comparable in terms of the knowledge and skills they are assessing, I create a new variable, which is an average test score for an individual, based on all the exam results. It is aimed to capture a pupil's overall performance in all 5 exams.

I present the results of the regressions in Table 3.A.3. Although statistically insignificant (due to sample size), the results suggest existence of a positive relationship between the concentration of PWA pupils in the class and the average exam performance.

Table 3.A.3: Regressions using test scores

Panel A: Average test score statistics					
	mean	st.dev.	min	max	n
average test score	53.853	16.998	20.4	96.2	334
Panel B: Regression results					
		(1)	(2)	(3)	(4)
Fraction		10.36	4.639	92.62	66.06
		(12.41)	(10.74)	(99.27)	(77.75)
Controls					
individual level controls		no	yes	no	no
class FE		no	no	yes	yes
N		334	271	334	271

Source: MECP2012
The regressions in this table are based on observations for a subsample of respondents for whom exam results data were available. The dependent variable is the average exam result (an average of all 5 tests pupils took). The main explanatory variable is $Fraction_{-ict}$ as defined throughout the analysis.
Standard errors clustered at class level in parentheses.
Statistical significance: *** 10%, ** 5%, * 1%

3.A.3 Alternative definitions of proportion of PWA pupils in the class

Overall concentration of PWA pupils in the class

Throughout the analysis I rely on the concentration of PWA pupils in the class defined according to the following equation:

$$Fraction_{-ict} = \frac{M_{-ict}}{C_{-ict}} \tag{3.6}$$

where M_{-ict} is the number of pupils with a parent abroad in class c (excluding person i) in semester t and C_{-ict} is the total number of pupils in class c in semester t.
By construction $Fraction_{-ict}$ varies over time but alternative, time-invariant specifications are also feasible.

I considered the following alternatives during the analysis:

1. Using mother's and father's emigration experience separately
2. Using information about a family ever experiencing paternal emigration, rather than the exact timing of the emigration experience

I refrain from splitting the parental emigration experience by parent's gender due to the nature of migration in the data. Specifically, it is usually the fathers, who emigrate and it is their migration experience which is mostly reflected in the results of my research. There are not enough observations for emigration of mothers to use it as a separate indicator in a regression; it returns statistically insignificant coefficients. Pooling the two groups - migrant fathers and mothers - together improves the precision of coefficients.

Alternatively, one may argue for the use of information about a family having ever experienced emigration in the observed period instead of the exact timing of emigration in the family. It certainly increases the number of observed cases per semester, but neglects important information about the PWA pupils by not allowing for returns. Barely any of the pupils were exposed to parental emigration for the entire 3 years and the short-time nature of emigration is key. In fact, regressions relying on the concentration of PWA pupils defined this way produce similar output, but the coefficients are often insignificant, due to much bigger standard errors. Another pitfall is that, when defined in such a way, the proportion of pupils with parents abroad does not vary over time, and hence use of dummy variables is limited.

Combination of the two cases discussed above could also be considered, evoking the same concerns regarding statistical significance and usefulness of the variable in capturing the temporary nature of migration.

Concentration of PWA pupils in the class by parents' education level

As in the case of the overall concentration of PWA pupils in the class, the fractions defined by education levels of parents may have also been defined differently.

In the analysis I relied on the following definition of the variables $FractionElementary_{-ict}$, $FractionVocational_{-ict}$, $FractionSecondary_{-ict}$ and $FractionTertiary_{-ict}$:

$$FractionX_{-ict} = \frac{MX_{-ict}}{CX_{-ict}} \tag{3.7}$$

where X = (Elementary, Vocational, Secondary, Tertiary) MX_{-ict} is the number of pupils with a parent abroad with the highest educational attainment X in class c (excluding person i) in semester t and CX_{-ict} is the total number of pupils whose parent has educational qualification X in class c in semester t. When identifying the educational levels I allowed either a mother or a father to have obtained a given education level. PWA pupils were identified, as before, on the basis of having at least one parent abroad at a given time t, so that the resultant $FractionX_{-ict}$ could vary over time.

The following options were considered:

1. using either a father's or a mother's emigration as an indicator of educational attainment
2. using information about a family ever experiencing paternal emigration, rather than the exact timing of the emigration experience to identify PWA pupils
3. rather than expressing the proportion relative to the number of all parents with the given education level, comparing it to the entire class size

The decision whether to use either parent's education level or to focus on a particular parent does not appear important for the results. My conclusions about the impact of certain groups of pupils do not depend on it. This is because parental education levels are highly correlated; hence the various definitions do not alter $FractionX_{-ict}$ significantly.

Redefining the fraction of PWA pupils whose parents have a given education level in relation to the entire class size also does not change the outcomes. It only results in the coefficients being rescaled. However, I find its interpretation less intuitive.

As for why I used the timing of parental emigration to identify PWA pupils, the reasons remain as before - to capture more precisely nature of the migration experience.

3.A.4 Impacts by gender and migration - further regressions

Table 3.A.4: Differential impacts depending on individual migration experience

	OLS		class FE			individual FE	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$FractionElementary_{-ict}$	.150	.176	-.082*	-.075**	-.024	-.082***	-.019
	(.166)	(.118)	(.043)	(.031)	(.044)	(.031)	(.037)
$FractionVocational_{-ict}$	-.794**	-.657**	-.128	-.127	-.126	-.209	-.203*
	(.315)	(.290)	(.159)	(.167)	(.145)	(.143)	(.112)
$FractionSecondary_{-ict}$	.176	-.063	.350**	.379**	.282*	.447***	.306***
	(.239)	(.212)	(.158)	(.171)	(.147)	(.136)	(.115)
$FractionTertiary_{-ict}$	.011	.217**	.092	.103*	.015	.134**	.039
	(.123)	(.105)	(.061)	(.062)	(.051)	(.055)	(.041)
individual emigration	-.256***	-.363***	-.163*	-.233***	-.244***	.069	.045
	(.093)	(.095)	(.086)	(.088)	(.088)	(.046)	(.043)
$FractionElementary_{-ict}$*Individual	.193	-.056	.099	.118	.113	.221***	.194***
	(.183)	(.173)	(.223)	(.176)	(.176)	(.038)	(.049)
$FractionVocational_{-ict}$*Individual	.166	.895*	.036	.506	.540	-.048	.014
	(.553)	(.488)	(.528)	(.442)	(.443)	(.178)	(.167)
$FractionSecondary_{-ict}$*Individual	.557*	.597*	.365	.239	.262	-.266**	-.160
	(.307)	(.321)	(.261)	(.336)	(.329)	(.117)	(.110)
$FractionTertiary_{-ict}$*Individual	.276*	.025	.275**	.114	.118	-.075	-.104
	(.141)	(.131)	(.144)	(.128)	(.126)	(.091)	(.076)
Controls							
individual level controls	no	yes	no	yes	yes	no	no
semester FE	no	yes	no	no	yes	no	yes
class FE	no	no	yes	yes	yes	no	no
individual FE	no	no	no	no	no	yes	yes
no of observations	13842	10853	13842	10853	10853	13842	13842
no of individuals	2669	2070	2669	2070	2070	2669	2669
no of classes	159	159	159	159	159	159	159

Source: MECP2012
The dependent variable is the average grade of a pupil. $FractionX_{-ict}$ are defined as throughout analysis, individual migration experience is a dummy variable equal to 1 if one's parent was abroad at time t.
Standard errors clustered at class level in parentheses.
Statistical significance: *** 1%, ** 5%, * 10%

3.A.5 Lagged Dependent Variable Specification vs. Individual Fixed Effects regressions

As mentioned, the fixed effects specifications (whether at the class or individual level) isolate any time-invariant changes specific to a class or student which influence the average grade of pupil and may be correlated with the $Fraction_{-ict}$. One may argue, however, that some of the important omitted variables vary over time. In particular, past values of the average grade are likely to explain a large proportion of the current average grade of a pupil and may be correlated with the proportion of PWA pupils in the class, even if I argue that such a case is unlikely.

A lagged dependent variable specification, with a lag of average grade included as an explanatory variable, may shed some light on the issue of which changes in particular drive the results in the main paper. By including the lagged average grade into the regression I am hoping to capture any remaining unobserved characteristics (not captured by class fixed effects) which may be influencing current average grade. Then the change in $Fraction_{-ict}$ needs to be exogenous only to changes in the average grade and not its level.

Table 3.A.5: Differential impacts depending on pupil's gender

	OLS		class FE		individual FE	
	(1)	(2)	(3)	(4)	(5)	(6)
$FractionElementary_{-ict}$	.255	.247	-.049	.065	-.042	.017
	(.247)	(.187)	(.101)	(.094)	(.052)	(.057)
$FractionVocational_{-ict}$	-.457	-.313	.111	.120	-.069	-.054
	(.371)	(.359)	(.284)	(.288)	(.183)	(.157)
$FractionSecondary_{-ict}$	.233	.105	.432**	.397*	.406**	.254*
	(.227)	(.233)	(.208)	(.226)	(.159)	(.141)
$FractionTertiary_{-ict}$	-.092	.143	-.043	-.060	.014	-.093*
	(.165)	(.138)	(.109)	(.117)	(.074)	(.049)
female	.443***	.464***	.404***	.434***		
	(.047)	(.054)	(.049)	(.053)		
$FractionElementary_{-ict}$*female	-.140	-.211	-.027	-.153	-.010	-.011
	(.180)	(.178)	(.203)	(.186)	(.083)	(.080)
$FractionVocational_{-ict}$*female	-.760*	-.502	-.570	-.408	-.273	-.307*
	(.440)	(.511)	(.453)	(.505)	(.175)	(.176)
$FractionSecondary_{-ict}$*female	-.013	-.130	-.079	-.143	.028	.061
	(.286)	(.326)	(.268)	(.301)	(.149)	(.141)
$FractionTertiary_{-ict}$*female	.261	.147	.287*	.155	.189***	.198***
	(.181)	(.199)	(.155)	(.178)	(.068)	(.069)
Controls						
individual level controls	no	yes	no	yes	no	no
semester FE	no	yes	no	yes	no	yes
class FE	no	no	yes	yes	no	no
individual FE	no	no	no	no	yes	yes
no of observations	13842	10853	13842	10853	13842	13842
no of individuals	2669	2070	2669	2070	2669	2669
no of classes	159	159	159	159	159	159

Source: MECP2012
The dependent variable is the average grade of a pupil. $FractionX_{-ict}$ are defined as throughout analysis, female is a dummy variable equal to 1 if pupil is female.
Standard errors clustered at class level in parentheses.
Statistical significance: *** 1%, ** 5%, * 10%

I run the following regression:

$$Y_{ict} = \alpha + \delta Fraction_{-ict} + \beta X_{ict} + \rho_t + \eta_c + \theta Y_{ic(t-1)} + \varepsilon_{ict} \tag{3.8}$$

where Y_{ict} is the average grade of individual i in class c at the end of given semester t, $Y_{ic(t-1)}$ is its first lag, $Fraction_{-ict}$ represents the proportion of students with migrant parents in class c at the beginning of semester t, excluding pupil i, and is the main variable of interest. X_{ict} is a set of individual level controls, η_c are class and ρ_t semester fixed effects. Standard errors are clustered at class level.

In Table 3.A.6 I present the results of the regression outlined above. I also restate the results of the regressions with individual fixed effects presented in Section 3.5, as I will be referring to them in my discussion.

As can be seen in Table 3.A.6, all regressions produce positive coefficients, which are smaller than in the main baseline regressions with class fixed effects. Most of them are also statistically insignificant. Nonetheless, all three specifications seem to imply existence of a positive relationship between the proportion of PWA pupils in the class

Table 3.A.6: Individual FE vs. LDV Specification

	individual FE (as in the baseline)	individual FE (as in the baseline)	LDV	LDV
	(1)	(2)	(3)	(4)
$Fraction_{-ict}$	.328*	.174	.207	.163
	(.176)	(.150)	(.225)	(.174)
Controls				
$Y_{ic(t-1)}$	no	no	yes	yes
individual level migration	no	yes	no	yes
other individual controls	no	no	no	yes
semester fixed effects	no	yes	no	yes
class fixed effects	no	no	yes	yes
individual fixed effects	yes	yes	no	no

Source: MECP2012
The dependent variable is the individual average grade at time t.
The main explanatory variable is the fraction of PWA pupils in a class at time t.
Other individual controls include gender, number of siblings, mother and father's age and education.
Standard errors are clustered at the class level and reported in parentheses.
FE stands for fixed effects, LDV stands for the lagged dependent variable specification
Statistical significance levels *** - 1%, ** - 5%, * - 10%

and the average grade.

It is difficult to establish which approach is best suited in this case as it depends on the belief about the behaviour of the omitted variables, i.e. whether they are time-invariant or not. The lagged dependent variable and individual fixed effects models are not nested and the distinction may play a role here. Angrist and Pischke (2009) point out, however, that LDV and FE models have a bracketing property which may be informative of the true relationship being analysed. If LDV is the correct approach but fixed effects are used, then the estimates of the positive effect will tend to be too big. If the reverse is true, then the estimates of the positive effect will be too small.

The FE estimates presented in Table 3.A.6 are only slightly larger than the LDV estimates and could provide an upper bound on the effect if the lagged dependent variable approach was more appropriate. In this case, the LDV estimates indicate either the correct impact or its lower bound.

Even though the results presented in the main paper may not be capturing the causal relationship between $Fraction_{-ict}$ and the average grade perfectly, it is reasonable to conclude that there is a positive association of a similar magnitude between the two variables.

3.A.6 Lagged impacts of concentration of PWA pupils

In Chapter 2 I also present results including lagged regressions, to account for the fact that migration may have a delayed influence; it is equally likely at an individual level and in a case of spillover. Thus, below I include a table with results of an analogous regression for the concentration of PWA pupils in the class. It is important to note that the full specification, including individual controls and class fixed effects produces insignificant regression coefficients, hence no conclusions can be drawn. This is most likely due to the fact that a specification with so many lags of the main explanatory variable is too demanding on the data set.

Table 3.A.7: Lagged impacts of the concentration of PWA pupils in the class

	OLS		class FE	
	(1)	(2)	(3)	(4)
$Fraction_{-ict}$	0.468	0.693	0.511	0.250
	(0.878)	(0.670)	(0.590)	(3.188)
first lag of $Fraction_{-ict}$	0.103	-0.674	-0.516	2.454
	(0.837)	(0.684)	(0.632)	(1.831)
second lag of $Fraction_{-ict}$	-1.232	-0.411	-0.624	-1.522
	(1.044)	(0.835)	(0.801)	(2.105)
third lag of $Fraction_{-ict}$	1.590	1.701	1.887**	-1.210
	(1.213)	(1.083)	(0.924)	(2.239)
fourth lag of $Fraction_{-ict}$	1.250	1.289	1.694**	-1.665
	(0.987)	(.888)	(0.842)	(2.164)
fifth lag of $Fraction_{-ict}$	-3.121***	-2.931***	-3.197***	1.380
	(1.129)	(0.911)	(0.846)	(1.705)
Controls				
individual level controls	no	yes	yes	yes
class FE	no	no	no	yes
no of observations	2252	1778	1776	1776
no of classes	137	137	137	137

Source: MECP2012
Individual level controls included in the regressions are as specified in the main analysis. Now the regression includes current concentration of pupils in the class and its five lags.
Standard errors are clustered at class level.
Statistical significance: *** 1%, ** 5%, * 10%

3.B Pattern of changes in average grade over time

Table 3.B.1: Average grades - further analysis

Panel A: Average grade over time												
	Overall sample				PWA pupils				non-PWA pupils			
	mean	st.dev.	min	max	mean	st.dev.	min	max	mean	st.dev.	min	max
YEAR 1												
semester 1	3.647	0.782	1.31	5.57	3.499	0.702	1.84	4.93	3.659	0.788	1.31	5.57
semester 2	3.662	0.85	1.4	5.85	3.546	0.822	1.4	5.14	3.671	0.853	1.6	5.85
YEAR 2												
semester 1	3.515	0.827	1.4	5.7	3.354	0.828	1.8	5.13	3.527	0.826	1.4	5.7
semester 2	3.626	0.878	1.57	5.88	3.532	0.824	2	5.3	3.632	0.881	1.57	5.88
YEAR 3												
semester 1	3.51	0.845	1	5.64	3.415	0.775	1.7	5.09	3.517	0.85	1	5.64
semester 2	3.775	0.87	1	5.88	3.683	0.827	2.08	5.38	3.783	0.874	1	5.88

Panel B: Differences in outcomes between semester 2 and semester 1 within each year									
	Year 1			**Year 2**			**Year 3**		
	Overall	PWA	non-PWA	Overall	PWA	non-PWA	Overall	PWA	non-PWA
Δ Average grade	0.059	.069	.056	.136	.134	.137	.262	.256	.264
	(.307)	(.313)	(.305)	(.285)	(.287)	(.285)	(.284)	(.259)	(.295)
N	2161	622	1513	2222	641	1552	2215	662	1528
Δ Behaviour	.001	-.024	.007	.159	.195	.144	.342	.361	.334
	(.774)	(.794)	(.766)	(.750)	(.732)	(.756)	(.732)	(.771)	(.717)
N	1892	544	1323	1999	586	1386	1898	568	1307
Δ School attendance	6.163	8.678	5.132	8.348	8.711	8.228	11.985	16.767	10.403
	(26.339)	(29.920)	(24.834)	(25.493)	(28.571)	(24.249)	(29.385)	(35.581)	(26.852)
N	1151	329	811	1171	346	813	726	189	528

Note: standard deviation provided in parentheses
Source: MECP2012

Chapter 4

Best of both worlds? Early cognitive and non-cognitive development of bilingual children.

4.1 Introduction

Language is a key instrument for human capital acquisition. Children develop linguistic skills very early in life and rely on them for further learning. Moreover, the early life language acquisition is often linked to better performance at later stages of schooling (Kamhöfer, 2014). This may be because language constitutes a *learning skill*,[1] returns to which are said to be very high in adult life (Neal, 2014).

Home is the first environment in which we learn language from the moment we are born. In this paper I analyse empirically whether cognitive and non-cognitive performance of children differs depending on whether they speak one or two languages at home. As I explain below this is an empirical question as there are many mechanisms at play, often acting in opposite directions. Moreover, answering the question requires overcoming significant empirical challenges; some of those challenges can be overcome thanks to the use of a very rich data set produced by the Scottish Government - Growing Up in Scotland data. It contains very comprehensive information about socio-economics and life-style of a randomly selected sample of families of small children in Scotland along with a series of objective measures of children's performance.

Given the importance of language for skill-development, raising bilingual children may be seen as an investment parents make in their human capital. Early life investments in human capital result in creation of various skills, which are complementary and build upon each other over time (Carneiro et al., 2013). Cognitive and non-cognitive skills of an individual are crucial for adult life outcomes, such as the labour market success (Borghans et al., 2008). At an aggregate level they also contribute to a country's economic growth through the labour market channel.

By exposing children to two different languages early on, parents may increase their productive skills and enable them to learn more efficiently in the future. Hence, bilingual children may have an educational advantage over their peers.

However, bilingual children come from families where at least one parent is foreign and this may be a disadvantage. Raising bilingual children is a high effort task. It requires extensive involvement and skill from the parents. Some parents may be more successful than others in teaching their children two languages. Therefore, any poten-

[1]Neal (2014) differentiates between productive and learning skills. He defines learning skills as those which not only increase one's productivity but also facilitate further learning. Productive skills are those acquisition of which increases productivity.

tial difference in performance of children from bilingual and monolingual families may also depend on factors other than the language they speak at home, e.g. the family background, culture, parental views and attitudes, etc.

This is important because bilingual families constitute a heterogeneous group. They consist of families with two foreign-born parents (henceforth fully foreign families) or one foreign and one native parent (mixed-nationality families) which differ substantially from each other and from the native families. These differences may play an additional role in children's skill development as the families may create a different growing up environment for their offspring.

For example, parents in mixed-nationality families are usually positively selected in terms of education and socio-economic background (Lanzieri, 2012). Thus they are in a better position to create a favourable upbringing environment for children and consciously engage in activities which enhance children's cognitive and non-cognitive development.

Families with two foreign parents, on the other hand, may be less proficient in the native language of the country and may lack location-specific knowledge essential for a child's upbringing, e.g. institutional arrangements. They may also have smaller networks than native families. These factors may obstruct a child's development by limiting its exposure to and interaction with native children.

For the reasons outlined above, it is theoretically ambiguous what is the effect of bilingualism on children's cognitive and non-cognitive performance. Thus an empirical analysis is required. To the best of my knowledge there are not many analyses of this kind in economic literature.[2] There is a consensus in current linguistic literature that bilingualism can benefit children. It is argued that children who learn two languages may experience a delay in speech (Baker, 1999). This may temporarily affect their cognitive and non-cognitive skills but their overall abilities remain largely unaffected over the long term (Kaushanskaya and Marian, 2007). It is highlighted, however, that bilinguals may be advantaged or not relative to monolinguals, depending on the specific nature of the task they are facing (Sorace, 2011). The linguistic studies are usually based on experimental data and use very precise measures of development. However, arguably the subject groups of the studies are often selected since participation is voluntary.

I exploit a rich, representative Growing Up in Scotland (GUS) dataset managed by ScotCen and the Scottish Government which provides information about a large, randomly selected sample of children in Scotland: their family situation, socio-economic

[2]Duncan and Trejo (2011) compare the outcomes of children from mixed-nationality and foreign families. They do not focus on linguistic skills, however, and consider school performance of teenagers. Their research cannot therefore answer a question as to how early, if at all, potential performance gaps emerge and whether they depend on children's linguistic skills.

Studies looking at small children, on the other hand, have mainly focused on the US and its Latino communities, and considered those from disadvantaged backgrounds or second generation immigrants (Reardon and Galindo, 2009; Fuller et al., 2009). Their growing up environment is different from the one created by bilingual families and thus the outcomes are not informative in this context.

characteristics, activities they engage in, parenting methods in the family and a series of children's performance measures. Importantly, apart from the parent-reported measures of children's performance, it also provides information on how a child performed in cognitive tests taken on the day of the survey.

The advantage of the data is that it is nationally representative. It is one of a very few data sets which, even though it was collected via a survey, contain a breadth of background information on the families and, importantly, provide objective measures of children's skills. Thus it allows me to overcome concerns of the linguistic literature that participants select into the study.

Given the expectation of heterogeneity among bilingual families, I firstly compare the socio-economic characteristics of mixed-nationality, fully foreign and native families in Scotland who have children under the age of 6. I analyse their lifestyle by considering the activities they engage in and their children's performance in various tests of cognitive and non-cognitive skills and of physical development.

I use the language spoken at home, a mixed family composition, i.e. whether a child has one foreign-born and one native parent, and a fully foreign family composition, i.e. when both parents were born outside of the UK, as indicators of the bicultural and bilingual environment. Even though the main focus is on performance of bilingual children, I single out mixed-nationality and foreign families to highlight the differences between the two groups. However, the results for foreign families are frequently too imprecise to draw any firm conclusions; this is due to a small number of families with two foreign-born parents in the sample.

I find that mixed, fully foreign and native families differ socio-economically. Further, they differ in the way they spend time with their children and in the views they hold on their children's upbringing and future career.

Children from all backgrounds perform comparably in most fields, with an exception of the English Vocabulary Naming exercise. On average, bilingual children do not perform worse than monolinguals in the task. There is heterogeneity within the group, however. Bilingual mixed-nationality children lag behind the monolingual native children at the age of 3 but they catch up by the time they are 5 years old. Further, monolingual mixed-nationality children perform better than monolingual native children. However, there is some evidence that bilingual children who have two foreign born parents may perform worse than the monolingual native children and not improve with age; the effect is sizeable but insignificant, which may be due to the small number of observations in the data.

I analyse the contribution of various activities to the English Vocabulary Naming score. I find that some, but not all factors, including practising letters and visits to the zoo, have a higher payoff for bilingual and mixed-nationality families than for native children. This is expected if worse performance is related to linguistic skills as these activities facilitate language acquisition.

It is clear that mixed families are not equivalent with the families where both

parents are foreign. Fully foreign families constitute a very small proportion of the GUS sample (1%). Therefore a comprehensive and robust analysis of outcomes for such a small subset is not possible. However, I establish that children from fully foreign families underperform in the same aspects of cognitive and non-cognitive development as children from mixed families, but the gap between them and native children is much larger. Moreover, unlike the children whose both parents are foreign-born, children from mixed-nationality families improve their performance with age. This may be because children from fully foreign families not only face the challenge of mastering two languages, but their parents may be in a worse position to help them catch up due to the lack of institutional, cultural and linguistic knowledge.

The analysis is descriptive in nature and relies on raw comparisons and simple regressions. The causal inference one can make is threatened by unobserved heterogeneity and selection bias. Of particular concern is the fact that foreign-born parents select into migration and into intermarriage. Moreover, teaching a child two languages is a choice and as such is endogenous. I discuss the extent to which these factors are problematic and stress that the results are robust to the inclusion of a variety of controls which may matter for both language used at home and children's skills.

Given the simplicity of the approach and inability to control for sources of bias, one may question the value of any additional input from economists. I argue that this research is novel and important for the understanding of the role skills play in economic outcomes.

I have already indicated that this work improves on a significant proportion of linguistic research by basing the analysis on a representative sample of population and thus eliminating the bias stemming from potential selection of participants into the study.

I do not provide a method which fully controls for sources of bias in this study. Neither do I claim, however, that the results presented here are causal. I am aware of and consider the threats to validity of the results; providing solution to these issues (by e.g. using an instrumental variable approach) is left for future research on the topic. As already pointed out, this analysis constitutes an important starting point in an area of interest to economists.

It is not uncommon for economists to present descriptive analyses uncovering patterns in certain phenomena prior to undertaking robust econometric analysis and obtaining causal relationships, particularly in new research fields. Examples can be found in economics of education. For instance, analyses of gender and racial gaps or differences in performance between immigrant and native children often rely on raw across-group comparisons and then gradually add controls, uncovering the degree to which the potential gaps can be explained by other factors (Fryer and Levitt, 2004, 2006, 2010; Duncan and Trejo, 2011).[3]

[3]Fryer and Levitt (2004, 2006) start their analyses of the black-white gap trajectory by analysing the raw differences in children's performance at different stages of kindergarten and early schooling.

Despite the shortcomings, this paper confirms many findings from the linguistic literature to date and provides further contributions. Firstly, I show that the role bilingualism plays in children's cognition varies depending on the family situation. Families with two foreign-born parents seem to be particularly disadvantaged. Thus, for policy purposes it is important to understand how the upbringing process differs in such environments. Another important novelty of this paper is the analysis of performance of children under the age of 6.

Further, I have at my disposal various, often complementary, measures of development. Their use allows me to conclude that cognition, broadly speaking, is not affected by bilingualism. The only effects are related to English language skills and are affected only early in life. The difference disappears by the age of 5.

The paper is structured in the following way. I provide a brief overview of relevant literature in Section 4.2. I discuss data and provide unconditional comparisons in Section 4.3. Section 4.4 contains regression analysis and its results. In Section 4.5 I discuss limitations of my approach and conclude.

4.2 Literature

This paper reflects ideas from various strands of economic, sociological and linguistic literature. Economists have argued that development of cognitive and non-cognitive skills is vital for short-term (Apps et al., 2012) and long-term outcomes of individuals (Aizer and Cunha, 2012; Behrman et al., 2014; Feinstein, 2003) and plays an important role in economic development (Hanushek and Woessmann, 2009).

These skills develop very early in life (Heckman and Conti, 2012; Carneiro et al., 2007) and depend on the initial level of human capital as well as investments made, which are complementary. What kinds of investments are most effective has been subject to a debate (Keane and Fiorini, 2012).

It is also important which skills parents invest in. Recent research, apart from distinguishing cognitive and non-cognitive traits, categorises the skills into productive and learning. Investments in learning skills contribute to both future productive and learning skills and enable further progress in learning. Investments in productive skills, on the other hand, return only future productive skills. Well-educated adults, who forego their earnings early in the career to invest in further education, possess greater learning capacities already when growing up. This is because their families invested in their learning skills during their childhood. The higher early investment in learning skills, the higher the payoff in the future (Neal, 2014). Teaching children two languages early in life may be seen as a parental investment in a learning skill.

I argue that language, specifically simultaneous bilingualism,[4] may be the main

They then include controls to demonstrate that a significant proportion of the gap can be explained by observed differences between the groups. Interestingly, they find a diverse trajectory across groups and ages. Only then they consider mechanisms which could explain the remaining performance gap.

[4]Simultaneous bilingualism is a form of bilingualism that takes place when a child becomes bilingual

channel of difference in performance between children, which brings me to linguistics literature. Baker (1999) provides an extensive overview of the impacts of bilingualism on cognitive outcomes in children. Bilinguals seem to have an advantage in certain thinking dimensions, such as divergent thinking, creativity, early metalinguistic awareness and communicative sensitivity. At the same time, bilingual children may initially possess a smaller vocabulary in each of their languages (Oller and Eilers, 2002; Portocarrero et al., 2007; Bialystok, 2009). Nonetheless, so far research found no correlations between bilingualism and IQ (Kaushanskaya and Marian, 2007) and it is suggested that many cognitive skills remain unaffected by bilingualism (Baker, 1999; Sorace, 2007). Most recent research indicates that bilingualism can slow down cognitive ageing by exerting a positive effect on later-life cognition (Bak et al., 2014). However, many of those findings are based on experiments run on a relatively selected sample.

Bilingualism on its own is unlikely to fully explain differences in performance between children. Bilingual families differ from each other and, since parental roles in early childhood are crucial, one should also account for the family background. Human and cultural capital are transmitted across generations and can influence educational outcomes (Black et al., 2005; Black and Devereux, 2011; Holmlund et al., 2011; Bjorklund and Salvanes, 2010). Children's attitudes towards school, aspirations and non-cognitive skills are highly correlated with those of their parents (Heckman and Rubinstein, 2001; Borghans et al., 2008; Carneiro et al., 2007). Activities families engage in and lifestyle, which form cultural capital, are usually learnt from parents and have influence on cognitive and non-cognitive skills (Meier Jaeger, 2011). De Philippis (2014) argues that culture is so persistent, it can explain correlation in PISA test scores between second generation immigrants and natives in their home countries.

Research suggests that family characteristics such as income and education (Ermisch, 2008; Hartas, 2011) but also time spent reading, writing or practising rhymes (Melhuish et al., 2008) may all influence children's cognitive and non-cognitive performance. Keane and Fiorini (2012) find that time spent in educational activities is the most productive input into cognitive skill development.

Thus, I expect the performance of bilingual children from mixed-nationality and fully foreign families to differ from each other because of the different environments they are growing up in.

The roles family background and culture play in outcomes have been recognised in migration studies. Economic literature established the existence of a performance gap for first generation immigrants, relative to the native population. The extent of the difference and whether it disappears with time depend crucially on the age at arrival in the country (Boehlmark, 2008) as well as the length of stay before the gap is measured (Glick and Hohmann-Marriott, 2007; Glick et al., 2012). The divide is also visible for second generation immigrants but varies across countries (Dustmann et al., 2012). In fact, studying second generation immigrants from minority groups in Britain,

by learning two languages simultaneously from birth.

Dustmann et al. (2010b) find that, for some minorities, the pupils not only catch up but even outperform their native peers. For this group whether the gap closes depends, among other factors, on ethnicity and country of birth (Reardon and Galindo, 2009; Glick et al., 2012), parental education levels (Fuller et al., 2009) and language spoken at home (Dustmann et al., 2010b; Rosenthal et al., 1983). Activities parents engage in are also central to the discussion (Brooks-Gunn and Markman, 2005). For example, Becker (2010) finds that in terms of language development Turkish children benefit more from activities outside the household than their German peers.

Most studies focus on immigrants past the early childhood stage. Dustmann et al. (2010a) consider 5-16 year olds in the UK, whilst Dustmann et al. (2012), Dronkers and de Heus (2012) and Kornder and Dronkers (2012) look at 15 year olds in Europe and Nordin and Rooth (2007) look at labour market outcomes of grown up second generation immigrants. With the exception of a few studies, little is known about immigrant children's performance at earlier stages of life.[5] Reardon and Galindo (2009) look at development of cognitive and non-cognitive skills of pre-schoolers and Fuller et al. (2009) of toddlers, but they focus specifically on Latino communities in the US. Hence, my analysis adds to the work in this area.

Further, research has generally focused on second generation immigrants and the literature on performance of children from mixed marriages is rather limited. Duncan and Trejo (2011) study outcomes of 16-17 year olds from Mexican-American mixed families and find that they outperform other Mexican second generation immigrants. They do not compare the group with the native population though.

Very little is said about language as a channel for closing of the performance gap identified in the migration literature. I demonstrate that bilingualism and family composition are strongly interlinked and key for a child's performance.

4.3 Data and descriptives

4.3.1 Data

The data used in this analysis come from the Growing Up in Scotland (GUS) longitudinal study. It has been commissioned by the then Scottish Executive Education Department and is managed by ScotCen Social Research.[6] It gathers information about physical, cognitive and non-cognitive development of children born in Scotland, as well as demographic and socio-economic details of the households they live in. The main topics covered by the study include the household composition and family background (parental education, income, employment, etc.), parental relationships, support parents receive and their views on parenting, childcare, pre-school and subsequently school enrollment, the child's health and development, the activities the child is involved in

[5]Note that the list proposed here is exemplary and by no means exhaustive.

[6]Detailed information about the project can be found on the website: http://growingupinscotland.org.uk

(including outings, physical and intellectual activities at home), social networks and children's development assessments.

Most importantly, the set of children's performance measures is diverse. Non-cognitive skills and physical development are assessed on the basis of questionnaires, such as the Strength and Difficulties Questionnaire and the Communication and Symbolic Behaviour Score Questionnaire, filled by parents or guardians. The cognitive skills of a child, however, are tested using the British Ability Scales during the interview. These measures in particular are therefore objective and reliable.

The participating families were randomly selected using Child Benefit records for Scotland and data was further weighted to adjust for initial selection as well as attrition. I apply the longitudinal weights throughout the analysis. For more details of the selection and weighting procedures, see Appendix 4.A.

The study now captures three cohorts of children: Child Cohort (CC) of around 3000 born in 2002/2003, Birth cohort 1 (BC1) of circa 5000 children born in 2004/2005 and Birth Cohort 2 (BC2) of about 6000 children born in 2010/2011. Due to data availability I rely on the BC1 and CC data for the purpose of this analysis.[7] The data for CC comprise 4 annual waves following children from age 3 to 6; the data for BC1 has been collected for 6 annual waves from when the children were 10 months of age until 6 years old. I apply the relevant weights and then combine the data to focus on analysis by age, rather than cohort. At a final wave the achieved sample size for both cohorts is 5857. This group participated in all waves of the study but observations are also available for those who participated only in some waves.

4.3.2 Identification of bilingual children

I identify bilingual children on the basis of language spoken at home. Using the data at hand, I separate groups who speak only English at home, English and another language and another language only (this is a negligible group).

This measure is not perfect. It does not provide much information about the families concerned. In particular, British born native speakers of English, foreign born residents of Scotland who are native speakers of English (e.g. if they come from the USA or Australia, etc.) as well as those who may choose to speak English, rather than their first language at home, will all be identified as monolingual families. On the other hand, speakers of Gaelic or Scots who identified themselves as speaking "other language" at home even though they are native residents of Scotland will be identified as bilingual, along with families where one or both parents were born outside of the UK. Therefore, the linguistic groups will be heterogeneous in terms of their cultural background.

For this reason, I also group families into categories on the basis of parents' origins. In particular, I define a child as coming from a *mixed family* if one of its parents was

[7] So far only one wave of data for Birth Cohort 2 has been released and it is not as informative for the purpose of this analysis. More information about the study and resultant research can be found on the project website, growingupinscotland.org.uk

born outside of the UK and one in the UK and as *native* if both parents were born in the UK. I also identify children from the *fully foreign* families if both their parents were born abroad. This group constitutes a very small proportion of the sample.[8] I separate *mixed* and *fully foreign* children in analysis, but given the small number of observations for *fully foreign* families, the results for this group should be treated with caution.

Definition of migrant status on the basis of country of birth is standard in the literature (Ozden et al., 2011) but has its limitations, as I cannot distinguish certain groups from each other. For example, a parent born abroad to two British citizens who then moved back to the UK will be identified as foreign born in this study. Equally a parent who is a second generation immigrant himself will be identified as *native* as he was born in the UK. In the majority of such cases I expect, however, the definition to imply that a child is brought up by parents of different nationalities, cultures and potentially in two different languages. Migration status is also often determined on basis of one's nationality, but this too has its drawbacks and is impossible to apply in this case, as no nationality information was collected during GUS.

Although intertwined, language and migration status may have different implications for children's development. Admittedly, there may be heterogeneity among bilingual children depending on whether they come from a *mixed* or *fully foreign* family. At early stages of development, children who speak the native language of the country they are growing up in, may find it easier to assimilate and interact with society (Rosenthal et al., 1983). Those who speak two languages may require further support from their parents and the level of help they receive will depend on the family composition.

Language and migration status are closely related. As can be seen in Table 4.1, 49% of children in *mixed and foreign families* speak English and another language or another language only at home. The corresponding group among *natives* reaches only 1%. Despite the high correlation, I will be using both language and family composition to identify the channel of the effect, if differences between children emerge. In particular, I would like to answer the question whether the difference is purely driven by language or whether unobserved characteristics of the families also contribute to the outcome.

From here on *mixed* family composition is defined by the variable *mixed*, the *fully foreign* family composition by the variable *fully foreign* and *bilingualism* is identified by the variable *bilingual*. It will become clear that often they are equivalent in terms of the results I obtain.

I compare the families in this study with what is known about immigrants to Scotland and conclude that the group is representative of the foreign and mixed-nationality families in Scotland. Details can be found in Appendix 4.A.5.

[8]There were only 70 children with both foreign parents in the combined sample in the final wave.

Table 4.1: Sample size and language

Panel A: Sample size at final wave			
	mixed families	fully foreign families	native families
Birth cohort	318	45	3344
Child cohort	179	25	2021
Total	497	70	5365
Panel B: Language spoken at home			
	only English	English and other	other only
overall	94%	5%	1%
mixed and foreign family	51%	40%	9%
native family	99%	1%	0%
Panel C: Correlations			
	Corr(language, mixed)		0.369
	Corr(language, foreign)		0.534

Data source: Growing Up in Scotland, ScotCen and the Scottish Government

4.3.3 Outcome variables for children

I analyse various measures of cognitive and non-cognitive development which were collected for participating children. As a check and to argue that there are unlikely to be differences in other aspects of development, I also briefly look at measures of motor and physical development available in the data. Below I describe how the outcome variables were created as well as which cohort and age group they are available for.

The Strength and Difficulties Questionnaire (SDQ) is a behavioural screening questionnaire. It was undertaken for children in both cohorts at ages 4, 5 and 6 and *filled by the child's parent* on the day of the survey. It includes 25 questions used to measure five aspects of a child's development - emotional symptoms, conduct problems, hyperactivity or inattention, peer relationship problems and pro-social behaviour. A score is calculated for each aspect and the total score is a sum of the scores from all the scales except the pro-social. The main indicator, total SDQ score, is a variable on the scale of 0-34 with the higher score indicating worse performance.

The Communication and Symbolic Behaviour Score (CSBS) measures non-cognitive development of children and was only used with Birth Cohort children at the age of 2. *Respondents were asked to complete questions* which assessed their child's communication, emotional development, understanding and interaction with peers. The 24 questions were grouped into clusters of individual scores. Clusters can be added into three composite scores assessing social communication, expressive language and symbolic functioning. A total score is the sum of the three composites and ranges from 0 to 57, with the higher score indicating better performance.

The exact questions and groupings which contribute to each score in SDQ and CSBS can be found in Appendix 4.A.

The British Ability Scales measure cognitive development. Children participating in GUS were subjected to two tests, Naming Vocabulary and Picture Similarities Exercise, which were *conducted by the surveyor, not reported by the parent.* The vocabulary test involves the child naming *in English* coloured pictures from a booklet he is shown one at a time and is aimed at assessment of spoken vocabulary. The exercise captures expressive language ability as well as the recall skill and depends on the child's existing vocabulary. The Picture Similarities test consists of a booklet with four images on each page and a set of cards with a single image. The child is asked to match the card with a picture in the booklet on the basis of them sharing an element or a concept. I use the percentile normative scores in the analysis. The normative scores are derived from standard tables and defined with the reference to the standardisation sample used in developing of the assessment (see Bradshaw et al. (2009) for details).

Respondents were also asked to assess **the child's speech development** from age 2 onwards. This is *a subjective measure* which was based on whether: 1) the child can be understood by strangers, 2) the child can be understood by family and friends and 3) the child can be understood by the respondent. The answer was to be given on the scale from 1 to 3 where 1 indicated mostly, 2 sometimes and 3 not at all.

Children in both cohorts were also assessed in terms of their **physical and motor development**. The test for babies took place at the age of 1 and for toddlers at the age of 3. Hence, CC was tested only once (age 3) and BC was subject to a baby test at age 1 and to a toddler test at age 3.

Availability of the outcomes for both cohorts at any given age is presented in Table 4.2.

Table 4.2: Availability of outcome measures across cohorts and age

Age	1		2		3		4		5		6	
	BC1	CC	BC1	CC	BC1	CC	BC1	CC	BC1	CC	BC1	CC
SDQ score							×	×	×	×	×	×
CSBS score			×									
BAS score					×				×			
Child's speech			×		×	×	×	×	×			
Motor development	×				×	×						

Note: here × indicates that data are available for this age group and cohort.

4.3.4 What do we learn about Scottish families - unconditional analysis

To study the effect of bilingualism on outcomes, it is important to control for characteristics of the families. Therefore, in this section I investigate socio-economic differences

between families, the way they spend time, their views and attitudes and their children's performance. I use weighted data, but do not control for any other characteristics in the comparisons.

Household composition and socio-economic situation

I start by comparing monolingual and bilingual families (see Table 4.3). I find that a higher percentage of parents in monolingual families are lone parents. Both types of families are relatively equally represented in all NS-SEC categories. They seem to be alike in terms of education levels of parents, although there is a degree of polarisation within the bilingual category with relatively high percentage of very highly educated parents and parents with no qualifications.

The division by language hides a significant heterogeneity between the families in the sample. Whether families are mono- or bilingual depends largely on parents' origins. Whilst almost all native families are monolingual, the bilingual group combines together mostly families with one and two foreign parents. The environment they can create for children to grow up in, which also contributes to the children's linguistic proficiency, may be better captured in an analysis on the basis of where the parents were born.

The *mixed*, *fully foreign* and *native* families in the study differ from each other in socio-economic characteristics. The families are similar in size, but a higher percentage of respondents in *native* families are lone parents in comparison with *mixed families*. In particular, 17% of native parents were lone parents when their child was 6 in contrast to only 11% of parents from *mixed* families and 1.4% of parents from *fully foreign* families.

Pronounced differences emerge also in terms of education with 47% of mixed parents and 29% of native parents having completed a degree. The higher educational attainment in *mixed* families is only partly channelled into their equivalised household incomes which are comparable with those of natives, except for the bottom quintiles. *Mixed* households are more likely to be classified higher in the NS-SEC classification with 67% falling into managerial and professional classification, compared with 53% of the native households. A higher percentage of *mixed* families live in the 20% least deprived areas of Scotland.

Thus, so far I find no indication of children in *mixed* families being at any material disadvantage relative to native children. In fact, given higher educational attainment of their parents on average, one may be inclined to conclude the opposite.

On the other hand, *fully foreign* families seem disadvantaged relative to *native* families. Even though almost 40% of parents from *foreign* families have a degree qualification, only 42% are employed in managerial and professional occupations. Notably, a higher percentage of them are small business owners when compared to *mixed* and *native* families. These families are also overrepresented in the bottom quintile of the household income distribution and almost a third live in the 20% most deprived areas of Scotland.

Table 4.3: Summary statistics - family socio-economics

% of lone parents	**ALL foreign**	***mixed***	***fully foreign***	**native**	**monolingual**	**bilingual**
Child's age	(N=567)	*(N=497)*	*(N=70)*	(N=5365)	(N=5856)	(N=267)
3	4.61	5.2	1.6	16.89	16.33	6.42
4	5.97	7	0	15.71	15.24	7.30
5	7.98	9.17	1.27	17.31	16.94	7.87
6	9.76	11.1	1.43	17.05	16.88	7.38
Household NSSEC category (%) ***	**ALL foreign**	***mixed***	***fully foreign***	**native**	**monolingual**	**bilingual**
Managerial and professional occupations	63.37	*67.4*	*41.9*	52.62	53.48	56.77
Intermediate occupations	8.79	*8.8*	*8.5*	14.29	14.02	10.28
Small employers and own account workers	10.1	*8.8*	*17*	6.53	6.46	12.96
Lower supervisory and technical occupations	5.96	*6.1*	*5.3*	8.22	8.07	6.86
Semi-routine and routine occupations	11.15	*8.7*	*24.1*	16.67	16.39	11.45
Never worked	0.63	*0.2*	*3.2*	1.67	1.57	1.67
Respondent's highest educational attainment (%) ***	**ALL foreign**	***mixed***	***fully foreign***	**native**	**monolingual**	**bilingual**
Degree or equivalent	46.03	*47.3*	*39.7*	29.11	29.83	45.19
Vocational qualification below degree	27.94	*29.1*	*22.2*	39.08	38.86	24.95
Higher Grade or equivalent	7.87	*7.3*	*10.6*	7.44	7.42	8.85
Standard Grade or equivalent	7.54	*8.6*	*2.3*	16.51	16.14	8.40
Other	2.17	*1.3*	*6.4*	0.12	0.18	2.55
No Qualifications	8.45	*6.3*	*18.9*	7.74	7.58	10.06
Equivalised income (%)***	**ALL foreign**	***mixed***	***fully foreign***	**native**	**monolingual**	**bilingual**
Bottom Quintile (<11,875)	18.78	*15.5*	*37.8*	20.39	19.93	35.42
2nd Quintile ($\leq$ 11,875 < 19,444)	19.03	*18.95*	*19.5*	20.41	20.27	19.27
3rd Quintile ($\geq$ 19,444 < 25,625)	18.56	*19.3*	*14.1*	19.04	19.16	13.02
4th Quintile ($\geq$ 25,625 < 37,500)	22.34	*23.4*	*16.1*	20.78	21.01	18.23
Top Quintile ($\geq$ 37,500)	21.29	*22.8*	*12.4*	19.38	19.62	4.06
Area of living by deprivation (%)	**ALL foreign**	***mixed***	***fully foreign***	**native**	**monolingual**	**bilingual**
20% least deprived	27	*28.20*	*21.25*	20.66	21.14	22.68
20-40% least deprived	19.96	*22.19*	*9.11*	21.68	21.80	16.64
40-60% least deprived	19.67	*21.26*	*11.95*	20.68	20.64	21.80
60-80% least deprived	17.16	*15.53*	*25.24*	16.83	16.63	20.73
20% most deprives	16.2	*12.82*	*32.45*	20.15	19.79	18.15

Data source: Growing Up in Scotland, ScotCen and the Scottish Government

Note: ALL foreign families is a pooled sample of mixed and fully foreign families. Mixed families are those in which only one parent was born outside of the UK. In fully foreign families both parents were born outside of the UK. In native families both parents were born in the UK.

Stars here indicate whether the distributions differ statistically across the groups.

Lack of stars suggests that the differences are not statistically significant. Significance levels: *** p<.01, ** p<.05, *p<.1

The observations about socio-economic characteristics prompt the question of potential selection of the non-native families. The concern is justified by the evidence in literature that more educated immigrants have a higher propensity to intermarry with natives (Sandefur and McKinnell, 1986; Lichter and Qian, 2001; Meng and Gregory, 2005; Chiswick and Houseworth, 2011) and the theory of assortative mating (Greenwood et al., 2014), suggesting that the group may be positively selected. On the other hand, the families with two foreign-born parents may be negatively selected. Therefore, it will be important to control for the family characteristics in the analysis.

Since the analysis by language used at home masks some important differences between the families, I will compare parental investments in children, parents' views and the way they spend time with children considering the family composition.

Parental investments in children

Given the richness of the data, it is possible to shed light on the activities children living in Scotland engage in and investigate whether *native*, *mixed* and *fully foreign* families make different *investments* in children.

Investments in this context encompass any activities parents involve in with children - educational, physical, social. The idea is to see whether a child's general environment differs in terms of their exposure to various factors which may contribute to development in early years. It is a key element of *nurture*, which may be correlated with parents' culture, hence contributing to human capital accumulation of the child (Keane and Fiorini, 2012). These *investments* may be a result of conscious choices parents make to ensure a child's development or a reflection of their lifestyle.

I consider unconditional differences between children from mixed, foreign and native families in every day activities they are involved in. The differences are taken over percentages of respondents from these families stating that they engage in a given activity. Here I just highlight some tendencies. Details can be found in Appendix 4.B.

Overall, families participate in similar kinds of activities and with a comparable frequency, particularly with respect to outdoor play. *Foreign* respondents do, however, on average visit friends with children less frequently than parents from *native* or *mixed* families do. In particular, 12% less *foreign* respondents visit friends with children most days when the child is 2 years old, but by the age of 6 the difference is only 2%.

Children in *mixed* families are less frequently involved in educational activities such as reading books or practising rhymes and songs. However, the differences are small and disappear with age. The situation is different for *fully foreign* families, where lower percentage of parents read to the child or practise letters with the child every day. Further, children in *mixed and foreign* families watch less TV on average, although the differences die off as they grow older.

Differences also emerge in types of entertainment outside home that parents provide for their children. For example, a lower percentage of parents from native families state

they have taken a child to the library or museum *in the previous year*, relative to *mixed* families. On the other hand, a higher percentage of respondents from *native* families have been to the swimming pool or zoo, compared with the *mixed* families. In contrast, children from *fully foreign* families are much less likely to engage in any such activities; for example, 40% more respondents in *native* families have taken their children to swimming pool.

The observations suggest heterogeneity in families' lifestyles, which may be a reflection of parents' lifestyles in general, e.g. whether they are physically active or have passion for literature, irrespective of having a child and be correlated with their socio-economic characteristics. They may, equally, be a result of conscious decisions made by parents regarding their children's upbringing. In particular, parents in *mixed* families may spend more time with their children practising letters as they feel a need to do so, given that children in many cases are bilingual and are learning two languages simultaneously.

Parental views and ambitions

Parents have distinctive ambitions for their children and views regarding upbringing. A higher percentage of parents in *mixed* than in *native* families hope for their child to complete a postgraduate degree. The difference may not be so surprising, bearing in mind that it is unconditional and that a higher percentage of *mixed family* respondents have completed tertiary education. The disparity narrows, however, with the age of the child. On the other hand, respondents from *foreign* families are less likely than natives to wish that their child completed an undergraduate degree and this difference persists as the child grows up.

Greater differences between respondents from *mixed*, *foreign* and *native* families are visible in their attitudes towards parenting. Specifically, native respondents were more likely to say that they agree or strongly agree that nobody can teach them how to be a good parent, although the gap narrows with the age of the child and becomes insignificant by the age of 4. A difference emerges also in the view that it is better for children to have two parents than one where about 16% more *mixed family* and 30% *foreign family* parents than *native* respondents agree or strongly agree with the statement. At the same time, respondents from *mixed and foreign* families are less likely to have used disciplining techniques, such as naughty step or ignoring bad behaviour with the child. They are also less likely to say that they smack the child or use a raised voice.

These contrasts in opinions may be partly a reflection of the family situation, with a higher percentage of *native* households being lone parent families. They may also suggest that families differ in their approach to upbringing on difficult to measure dimensions. There is potential for this heterogeneity to translate into child's outcomes, particularly in sphere of non-cognitive skills and behaviour (Borghans et al., 2008;

Carneiro et al., 2007; Heckman and Rubinstein, 2001).

4.3.5 Child outcomes

I start by comparing the performance of bilingual and monolingual children without conditioning on any other variables, to identify whether their outcomes differ. Since the bilingual group is rather heterogenous, I also make these comparisons by family composition.

From Table 4.5 it is clear that bilingual children score much lower than monolingual children in the (English) Vocabulary Naming exercise at both age 3 and 5, but the gap between the average score of the two groups narrows with age.

The observation is also true when comparing children from both *mixed* and *fully foreign* families with native children. Importantly, however, the gap is much smaller for the group of children from *mixed* families and almost closes by the age of 5. Children from *foreign* families score 25% lower than native children at the age of 3, which is equal to 45% of the average score. Further, they still perform much worse than native peers at the age of 5.

The same cannot be said about the Picture Similarities scores, where there are no significant differences between the groups, irrespective of whether the comparison is made on the basis of language or family composition; if anything, *mixed and foreign* family children seem to overtake *native* children and *bilingual* children overtake the *monolingual* children.

There are also visible differences in the percentage of children who, according to the respondents, can be *mostly* understood by strangers. The gap between bilingual and monolingual children is significant at the age of 2 but the outcome equalises with age. Differentiating by family composition, once again the gap exists only for *foreign* children, closes with age and disappears entirely by the age of 5. A similar pattern emerges when parents are asked whether the child can be understood by family and friends, but not if the child can be understood by the respondent.

Children perform comparably in non-cognitive and behavioural assessments, such as CSBS and SDQ.[9] This observation holds for the total scores, as well as their composites (see Tables 4.4 and 4.6). However, the difference in CSBS total score (.761 for mixed and 1.623 for foreign children, equivalent to 14% and 30% of standard deviation, respectively), is statistically significant; it is due to the difference in performance of children in CSBS social and symbolic composite part of the test. Statistically significant differences also emerge between native and foreign children in the total SDQ score and its peer relationships component.

I consider measures of motor and physical development and find no differences across the groups. The results of this analysis can be found in Table 4.B.8 in Appendix 4.B.1.

[9]Note: CSBS is the Communication and Symbolic Behaviour Score measured at age 2, SDQ is the Strengths and Difficultires Questionnaire score measured at ages 4, 5 and 6.

To summarise, the observations so far suggest that both language and family composition may play a role in children's performance. Based on the unconditional comparisons, children from *mixed and foreign families* lag behind in cognitive outcomes that are most likely driven by language skills (i.e. speech-related) and not any other aspects of development. Moreover, in case of mixed-nationality children, this is true only early on in life; they catch up with native children by the age of 5. Similar conclusion is reached when comparing bilingual and monolingual children. This is most likely because the family composition is closely related to bilingualism; almost all native families are monolingual whilst mixed and fully foreign families are more likely to be bilingual.

Importantly, there are no differences in non-cognitive and behavioural outcomes. Judging by the average Picture Similarities score, the cognitive skills are also not affected (at least to the extent measured by the test). However, it becomes clear that children from *foreign* families perform visibly worse in the exercises and the initial gap between *mixed family* and *native* children is much smaller.

Table 4.4: Children's outcomes - CSBS average scores

	native	mixed	fully foreign	monolingual	bilingual	min	max
social composite	22.571	22.25	22	23	21.42	0	26
speech composite	11.273	11.238	10.985	11.5	10.45	0	14
symbolic composite	15.357	15.006	14.3	15	13.79	0	17
total score	49.312	48.551	47.689	50.75	45.58	0	57

	$\Delta_{(NM)}$	$\Delta_{(NF)}$	$\Delta_{(MB)}$
social composite	0.321	0.571	1.58
speech composite	0.035	0.288	1.05
symbolic composite	0.351	1.057	1.21
total score	0.761	1.623	5.17

Data: Growing Up in Scotland, ScotCen and the Scottish Government
CSBS stands for Communication and Symbolic Behaviour Score, measured at the age of 2
Note: $\Delta_{(NM)}$ is the difference in outcomes between children from native and mixed families
$\Delta_{(NF)}$ is the difference in outcomes between children from native and fully foreign families
$\Delta_{(MB)}$ is the difference in outcomes between monolingual and bilingual children

Table 4.5: Children's outcomes - British Ability Scales average scores

	native		mixed		fully foreign		monolingual		bilingual		min	max
age	3	5	3	5	3	5	3	5	3	5		
Picture Similarities	51.093	72.13	51.445	74.238	52.745	73.893	61.61	79.58	50.26	73.6	1	99
(English) Vocabulary Naming	56.99	74.386	51.718	72.808	31.259	51.021	53.15	69.16	23.82	46.88	1	99

	Age 3			Age 5		
	$\Delta_{(NM)}$	$\Delta_{(NF)}$	$\Delta_{(MB)}$	$\Delta_{(NM)}$	$\Delta_{(NF)}$	$\Delta_{(MB)}$
Picture Similarities	-0.352	-1.652	11.35	-2.108	-1.763	5.98
(English) Vocabulary Naming	5.272	25.731	29.33	1.578	23.365	22.28

Data: Growing Up in Scotland, ScotCen and the Scottish Government
Note: $\Delta_{(NM)}$ is the difference in outcomes between children from native and mixed families
$\Delta_{(NF)}$ is the difference in outcomes between children from native and fully foreign families
$\Delta_{(MB)}$ is the difference in outcomes between monolingual and bilingual children

Table 4.6: Children's outcomes - SDQ average scores

	native			mixed			fully foreign			monolingual			bilingual				
age	4	5	6	4	5	6	4	5	6	4	5	6	4	5	6	min	max
emotional score	1.178	1.245	1.275	1.207	1.346	1.366	1.367	1.507	1.223	.562	.666	1	1.44	1.65	1.55	0	5
conduct score	1.978	1.77	1.586	1.883	1.712	1.494	2.127	1.693	1.575	1.5	1.66	1.76	2.44	1.67	1.67	0	10
hyperactivity score	3.588	3.605	3.416	3.453	3.575	3.564	3.709	3.888	3.446	3	3.93	3.85	3.82	3.89	3.69	0	10
peer relationships	1.136	1.04	0.975	1.183	1.104	1.081	2	1.893	1.485	.875	1.73	1.30	2.26	1.8	1.692	0	10
pro-social	7.849	8.176	8.399	7.847	8.154	8.339	7.395	7.986	8.058	7.68	7.73	7.62	7.38	8.07	8.15	0	10
total score	7.846	7.647	7.239	7.66	7.722	7.493	9.176	9.028	7.769	5.93	8	7.92	9.94	9.10	8.72	0	35

Differences

	Age 4			Age 5			Age 6		
	$\Delta_(NM)$	$\Delta_(NF)$	$\Delta_(MB)$	$\Delta_(NM)$	$\Delta_(NF)$	$\Delta_(MB)$	$\Delta_(NM)$	$\Delta_(NF)$	$\Delta_(MB)$
emotional score	-0.029	-0.189	-0.878	-0.101	-0.262	-0.65	-0.091	0.052	-0.11
conduct score	0.095	-0.149	-0.94	0.058	0.077	-0.01	0.092	0.011	-0.009
hyperactivity score	0.135	-0.121	-0.082	0.03	-0.283	-0.04	-0.148	-0.03	0.16
peer relationships	-0.047	-0.864	-1.385	-0.064	-0.853	-0.07	-0.106	-0.51	-.0392
pro-social	0.002	0.454	0.3	0.022	0.19	-0.34	0.06	0.341	-0.53
total score	0.186	-1.33	-4.01	-0.075	-1.381	-1.10	-0.254	-0.53	-0.8

Data: Growing Up in Scotland, ScotCen and the Scottish Government
SDQ stands for the Strengths and Difficulties Questionnaire, measured at ages 4, 5 and 6
Note: $\Delta_(NM)$ is the difference in outcomes between children from native and mixed families
$\Delta_(NF)$ is the difference in outcomes between children from native and fully foreign families
$\Delta_(MB)$ is the difference in outcomes between monolingual and bilingual children

Table 4.7: Children's outcomes - Child can be understood by strangers... (%)

	native				mixed				fully foreign				monolingual				bilingual			
age	2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
mostly	36.76	69.06	82.63	89.46	34.03	65.88	81.76	90.61	26.03	44.34	64.04	86.27	36.70	69.00	82.65	89.57	31.22	60.83	77.43	87.80
sometimes	44.91	25.53	14.62	8.88	45.37	26.08	14.47	8.66	42.47	31.13	23.6	11.76	45.10	25.53	14.55	8.84	42.44	28.34	17.01	10.37
rarely	18.33	5.42	2.75	1.66	20.6	8.04	3.77	0.72	31.51	24.53	12.36	1.96	18.20	5.47	2.80	1.59	26.34	10.83	5.56	1.83

	Age 2			Age 3			Age 4			Age 5		
	$\Delta_(NM)$	$\Delta_(NF)$	$\Delta_(MB)$	$\Delta_(NM)$	$\Delta_(NF)$	$\Delta_(MB)$	$\Delta_(NM)$	$\Delta_(NF)$	$\Delta_(MB)$	$\Delta_(NM)$	$\Delta_(NF)$	$\Delta_(MB)$
mostly	2.73	10.73	5.48	3.18	24.72	8.17	0.87	18.59	5.22	-1.15	3.19	1.77
sometimes	-0.46	2.44	2.66	-0.55	-5.6	-2.81	0.15	-8.98	-2.46	0.22	-2.88	-1.53
rarely	-2.27	-13.18	-8.14	-2.62	-19.11	-5.36	-1.02	-9.61	-2.76	0.94	-0.3	-0.24

Data: Growing Up in Scotland, ScotCen and the Scottish Government
Note: $\Delta_(NM)$ is the difference in outcomes between children from native and mixed families
$\Delta_(NF)$ is the difference in outcomes between children from native and fully foreign families
$\Delta_(MB)$ is the difference in outcomes between monolingual and bilingual children

4.4 Conditional comparisons

The unconditional analysis suggested that, although there are some differences between families in terms of socio-economics and in parental views, parents often engage in similar activities with children and children perform comparably in majority of dimensions, except for those related to speech. I hypothesise that the difference in the latter is driven by the fact that high percentage of children in *mixed and foreign families* are bilingual or speak a language other than English.

Comparisons not taking into account any socio-economic circumstances of the families are likely to produce a misleading picture, particularly given the fact that children's outcomes are often correlated with parental education levels (Black et al., 2005), income and social status (Meier Jaeger, 2011) as well as environmental factors, such as the number of siblings, social interactions the child is exposed to, etc. (Heckman and Conti, 2012; Hartas, 2011). It is vital to control for these factors to isolate the effect related purely to the family composition or language. The raw comparisons do not control for heterogeneity within the groups, whereas distinctions may emerge given specific circumstances.

Conditioning on variables key for children's performance may also, at least partly, mitigate the effect of selection of migrants. Specifically, given the presumption that *mixed* families are positively selected on socio-economics, I would like to control for the selection. Further, positively selected *mixed-nationality* parents may realise the disadvantage their children are at and consciously attempt to compensate for it. An example of such compensation could be the higher frequency with which they practice letters with children or visit the library. Since such types of investments in children matter for their cognitive development (Keane and Fiorini, 2012), they need to be accounted for in the analysis.

4.4.1 Empirical specification

Baseline

I start the analysis from a simple regression of various outcomes on the language and family composition, controlling for socio-economic characteristics of the household, activities parents engage in at home, physical activity and parenting methods as proxied by attitude to discipline. The regression equation becomes:

$$Y_{it} = \alpha + \beta_1 bilingual_i + \beta_2 mixed_i + \beta_3 foreign_i + \beta_4 female_i + \beta_5 age_{it} + \beta_6 X_{it} + \gamma_t + \epsilon_{it} \tag{4.1}$$

where Y_{it} are various outcome measures for child i at time t, *bilingual* is a dummy equal to 1 if a child speaks English and another language at home, *mixed* is a dummy variable equal to 1 if one of child's parents was born outside of the UK, *foreign* is a dummy variable equal to 1 if both parents were born outside of the UK, *female* is a dummy variable equal to 1 if the child is female, *age* is a variable reflecting child's age

in years, X_{it} contains household characteristics such as number of siblings, whether a two-parent family, parental education level and NS-SEC classification, geographical location by the index of deprivation, and variables directly related to child's upbringing such as activities child engages in at home (rhymes, letter and reading practice, use of computer, watching TV), physical activity (play outside, running, jumping, etc.), outings (visits to library, museum, zoo, gallery, swimming pool, cinema) and discipline (use of naughty step, time out etc.). Where the outcome variable was measured at more than one point in time, I cluster standard errors at an individual level and include time fixed effect γ_t. Some elements of X_{it}, such as the activities families engage in, may be endogenous as they are likely simultaneously determined with the child's outcomes. However, as will become clear from the output tables, excluding them from regressions does not change the results. I include both the family composition dummies and a bilingual dummy as I have already argued that they may jointly determine children's outcomes.

The measures considered here are the BAS outcomes (Picture Similarities and Vocabulary Naming score), Strength and Difficulties Questionnaire and Communication and Symbolic Behaviour Scale. I use OLS to estimate the impact on these. I also briefly look at the respondent-assessed speech development of the children, which is measured using an ordinal variable (1-3). For this outcome I rely on OLS and ordered probit but do not report all results, as the relationship is insignificant once controls are included.

Since the family composition and the languages spoken at home are correlated with each other and children from fully foreign families seem to perform much worse than others, I introduce an interaction between language spoken at home and family composition to further explore the relationship between these two variables. The regression becomes:

$$\begin{aligned} Y_{it} = \alpha + \beta_1 bilingual_i + \beta_2 mixed_i + \beta_3 foreign_i + \theta_1 mixed_i \times bilingual_i \\ + \theta_2 foreign_i \times bilingual_i + \beta_4 female_i + \beta_5 age_{it} + \beta_6 X_{it} + \gamma_t + \epsilon_{it} \end{aligned} \tag{4.2}$$

All the controls remain unchanged. I exclude monolingual speakers of another language from the regression (n=68) as combining bilingual children with monolingual speakers of language other than English is problematic, as the children are likely to face different challenges. Bilingual children learn two languages simultaneously, but when they master them, they are fluent in English and hence their interaction with other members of the society is eased. Children who only speak another language are likely to face a new set of difficulties upon beginning school when they need to learn English. The group of monolingual speakers of another language is negligible in the data and their exclusion from the regression does not change the results.

Differential impacts

I analyse further the outcomes for which I find an effect of being in a mixed or foreign family or being bilingual. In particular, I am interested in gender and age-variation in

performance.

It is reasonable to think that girls may develop differently from boys, also in the context of bilingualism and multiculturalism.

Given the observation in linguistic literature (Baker, 1999) that, although bilingual children are at a disadvantage in certain areas of development in early years, they catch up with or even supersede their peers and the fact that in unconditional comparisons gaps seem to narrow with age, I look at changes in the difference with age. Hence, I introduce further interaction terms of *mixed*, *foreign* and *bilingual* with age and gender into the regressions.

$$\begin{aligned} Y_{it} = \alpha + \beta_1 bilingual_i + \beta_2 mixed_i + \beta_3 foreign_i + \theta_1 mixed_i \times bilingual_i \\ + \theta_2 foreign_i \times bilingual_i + \beta_4 female_i + \lambda_1 bilingual_i \times female_i \\ + \lambda_2 mixed_i \times female_i + \lambda_3 foreign_i \times female_i + \beta_5 age_{it} + \beta_6 X_{it} + \gamma_t + \epsilon_{it} \end{aligned} \tag{4.3}$$

$$\begin{aligned} Y_{it} = \alpha + \beta_1 bilingual_i + \beta_2 mixed_i + \beta_3 foreign_i + \theta_1 mixed_i \times bilingual_i \\ + \theta_2 foreign_i \times bilingual_i + \beta_4 age_{it} + \lambda_1 bilingual_i \times age_{it} + \lambda_2 mixed_i \\ \times age_{it} + \lambda_3 foreign_i \times age_{it} + \beta_5 female_i + \beta_6 X_{it} + \gamma_t + \epsilon_{it} \end{aligned} \tag{4.4}$$

4.4.2 Results

Cognitive outcomes

In Tables 4.1 and 4.2 I present results of the baseline regressions for BAS scores. The coefficients in columns (1) to (6) of Table 4.1 suggest that a bilingual child scores on average almost 10% lower in the English Vocabulary Naming Exercise than a monolingual child. Moreover, children from fully foreign families score over 11% lower than native children. These are large impacts, equal to roughly one fifth of the score's mean in the sample. Importantly, mixed-nationality children do not score differently from the native children.

However, looking at results in column (7) of Table 4.1, heterogeneity within the bilingual group emerges upon inclusion of the interaction term between family composition and language. The results suggest that, on average, children of mixed nationality speaking English only score better than native children. Further, bilingual children in mixed families perform comparably to native children, but children from *fully foreign* families who speak English and another language score almost 11% lower in the exercise relative to native monolingual children. It should be noted that many regression coefficients become insignificant, which can be expected given the high correlation between the variables and the small number of observations for children from *fully foreign* families.

The results for BAS Picture Similarities test (Table 4.2) are statistically insignificant and negligible in size, confirming the previous observation that the cognitive skills of children may not differ across the groups.

I also consider impacts on the ability of a child to be understood by strangers and report results in Table 4.3. There is no differential impact between mixed family and native children, as well as between bilingual and monolingual children. Further, any effect of being in a fully foreign family, although positive and large, becomes statistically insignificant upon inclusion of control variables. This does not necessarily imply that there are no differences in performance; most likely the results are imprecise due to the sample size.

I relate the less conclusive findings to the fact that the measure is subjective and depends on parental perception of what being understood means. Nonetheless, a positive coefficient would suggest that a child from a *fully foreign* family or who speaks a different language is less likely to be understood by strangers.[10] I repeat a similar analysis for the two remaining questions in the respondent-assessment measure of speech development but find no significant results. The output can be found in Appendix 4.C.

[10]Note that the speech-assessment variable was coded in the following way: 1 - often, 2 - sometimes, 3 - rarely or not at all. Therefore, a positive coefficient in ordinal probit regression indicates that a child is less likely to be understood by strangers.

Table 4.1: Regression outcomes for BAS Vocabulary Naming Score

Dependent variable: BAS Vocabulary Naming Score							
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS
bilingual	-11.215***	-10.529***	-9.994***	-9.695***	-9.775***	-9.772***	-3.120
	(2.190)	(2.034)	(2.024)	(1.978)	(1.974)	(1.970)	(3.216)
mixed family	1.146	-.716	-.736	-1.153	-1.172	-1.199	10.872**
	(1.698)	(1.673)	(1.661)	(1.634)	(1.633)	(1.629)	(5.070)
fully foreign family	-13.316***	-10.827**	-11.421***	-11.623***	-11.678***	-11.770***	-1.736
	(4.650)	(4.677)	(4.651)	(4.447)	(4.449)	(4.446)	(12.762)
mixed*bilingual							-8.347
							(6.055)
fully foreign* bilingual							-10.552**
							(4.157)
Controls							
household controls	no	yes	yes	yes	yes	yes	yes
activities at home	no	no	yes	yes	yes	yes	yes
outings	no	no	no	yes	yes	yes	yes
physical activity	no	no	no	no	yes	yes	yes
discipline	no	no	no	no	no	yes	yes
N	4054	4054	4018	3975	3974	3974	3974
R-squared	.116	.199	.212	.222	.223	.224	.225

Data: Growing Up in Scotland, ScotCen and the Scottish Government

Note: All regressions include gender, sweep and the cohort dummy as controls. Household controls include: number of siblings, parental education, NS-SEC household classification, equivalised income, geographical area and whether one is a single parent. Activities at home are frequency of painting/drawing, practising rhymes, practising letters, use of computer and TV. Outings include visits to many attractions, including museums, library, etc. Physical activity includes swimming, playing in the park, running, etc. And disciplining techniques include time out and naughty step. The independent variable mixed family is a dummy equal to 1 if one of child's parents was born outside of the UK and zero otherwise. The variable fully foreign family is a dummy equal to 1 if both child's parents were born outside of the UK and zero otherwise. Bilingual is a dummy variable equal to 1 if child speaks English and another language or just another language at home.

Errors are clustered at individual level.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 4.2: Regression outcomes for BAS Picture Similarities Score

Dependent variable: BAS Picture Similarities Score							
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) OLS
bilingual	-1.994	-2.376	-2.091	-2.054	- 2.363	-2.360	-1.101
	(1.732)	(1.680)	(1.690)	(1.671)	(1.665)	(1.663)	(2.642)
mixed family	2.340*	.638	.257	.043	.077	.065	1.293
	(1.398)	(1.401)	(1.396)	(1.391)	(1.383)	(1.380)	(4.288)
fully foreign family	4.258	4.923	4.493	4.510	4.496	4.447	11.454
	(3.284)	(3.389)	(3.413)	(3.326)	(3.321)	(3.306)	(8.959)
mixed*bilingual							-4.201
							(5.013)
fully foreign*bilingual							-1.220
							(3.559)
Controls							
household controls	no	yes	yes	yes	yes	yes	yes
activities at home	no	no	yes	yes	yes	yes	yes
outings	no	no	no	yes	yes	yes	yes
physical activity	no	no	no	no	yes	yes	yes
discipline	no	no	no	no	no	yes	yes
N	4189	4054	4018	3975	3974	3974	3974
R-squared	.140	.187	.191	.196	.199	.199	.190

Data: Growing Up in Scotland, ScotCen and the Scottish Government

Note: All regressions include gender, sweep and the cohort dummy as controls. Household controls include: number of siblings, parental education, NS-SEC household classification, equivalised income, geographical area and whether one is a single parent. Activities at home are frequency of painting/drawing, practising rhymes, practising letters, use of computer and TV. Outings include visits to many attractions, including museums, library, etc. Physical activity includes swimming, playing in the park, running, etc. And disciplining techniques include time out and naughty step.

The independent variable mixed family is a dummy equal to 1 if one of child's parents was born outside of the UK and zero otherwise. The variable fully foreign family is a dummy equal to 1 if both child's parents were born outside of the UK and zero otherwise. Bilingual is a dummy variable equal to 1 if child speaks English and another language or just another language at home.

Errors are clustered at individual level.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 4.3: Regression outcomes for self-reported speech assessment

Dependent variable: Can the child be understood by strangers?							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OP	OP	OP	OP	OP	OP	OP
bilingual	.029	.063	.084	.026	.022	.026	.004
	(.026)	(.054)	(.065)	(.102)	(.103)	(.103)	(.204)
mixed family	.011	.050	.051	.035	.030	.034	-.029
	(.021)	(.046)	(.060)	(.082)	(.082)	(.082)	(.279)
fully foreign family	.135**	.193	.247*	.278	.280	.286	.352
	(.061)	(.132)	(.146)	(.233)	(.233)	(.235)	(.623)
mixed*bilingual							-.023
							(.343)
fully foreign*bilingual							.051
							(.238)
Controls							
household controls	no	yes	yes	yes	yes	yes	yes
activities at home	no	no	yes	yes	yes	yes	yes
outings	no	no	no	yes	yes	yes	yes
physical activity	no	no	no	no	yes	yes	yes
discipline	no	no	no	no	no	yes	yes
N	8069	8069	6605	6975	3974	3974	3974
R-squared	.010	.138	.138	.091	.093	.094	.094

Data: Growing Up in Scotland, ScotCen and the Scottish Government
OP stands for ordered probit, the possible answers to the question were: 1 - mostly, 2 - sometimes, 3 - rarely
Note: All regressions include gender, sweep and the cohort dummy as controls. Household controls include: number of siblings, parental education, NS-SEC household classification, equivalised income, geographical area and whether one is a single parent. Activities at home are frequency of painting/drawing, practising rhymes, practising letters, use of computer and TV. Outings include visits to many attractions, including museums, library, etc. Physical activity includes swimming, playing in the park, running, etc. And disciplining techniques include time out and naughty step.
The independent variable mixed family is a dummy equal to 1 if one of child's parents was born outside of the UK and zero otherwise. The variable fully foreign is a dummy equal to 1 if both child's parents were born outside of the UK and zero otherwise. Bilingual is a dummy variable equal to 1 if child speaks English and another language or just another language at home.
Errors are clustered at individual level.
Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Non-cognitive and physical development

I analyse the non-cognitive development by first looking at the total scores for the Strength and Difficulties Questionnaire and the Communication and Symbolic Behaviour Score (see Tables 4.4 - 4.7).

The regression results for SDQ total score suggest a slightly worse performance of children in *mixed* and *bilingual* families and a better performance of children in *fully foreign* families, relative to native children. The coefficients are, however, statistically insignificant, confirming what was clear also in summary statistics, that the groups score comparably in SDQ.

The total CSBS score is affected by a child's family composition. According to

columns (1) to (5) of Table 4.6 *bilingual* children and children from *mixed* and *fully foreign* families score worse than native children in this assessment. However, the results are only statistically significant for mixed-nationality children.

A child from a *mixed* family scores on average .773 less than native children, which is equivalent to 1.8% of the mean score but as much as 17% of the score's standard deviation. The results are very close to the unconditional differences discussed before. A larger but insignificant impact is found for fully foreign children.

Where could this difference in impacts between SDQ and CSBS be coming from, given that they both measure non-cognitive aspects of child's development? Unlike the SDQ test, the CSBS test was only taken at one point in time (age 2) and on one cohort of children (BC). It is possible that differences were more visible at this stage, but it is difficult to assess robustness of this result given the cross-sectional nature of the outcome. I cannot investigate whether the performance changes with age either. The Strength and Difficulties Questionnaire was used to assess children from the age of 4 onwards and, although it measures similar aspects of non-cognitive development, questions respondents were asked differ from those involved in CSBS analysis making comparisons infeasible. It is possible, however, that by the age of 4 children improve their performance and hence SDQ tests do not reveal any differences.

This initial analysis does not provide firm conclusions regarding non-cognitive development of children. The nature of the tests and the arbitrary way in which total scores are obtained (by summing up the composite scores), may raise questions about validity of the findings and whether some existing differences become invisible due to aggregation. I replace the total scores with clusters of SDQ and composite scores of CSBS as dependent variables and run separate regressions for these elements of assessments only. The results of fully specified regressions, including all previously used controls, can be seen in Tables 4.5 and 4.7.

Among the subcomponents of the total SDQ score, the only ones affected by the family composition or language are the *peer problems* and *emotional symptoms scores.*

The regressions imply that children in mixed-nationality families perform comparably to native children but children from *fully foreign* families and bilingual children score higher in the peer relationships test, which suggests that they face greater difficulties in relations with peers.

It is difficult to comment on the degree of interaction between the foreign family composition and language, as the coefficients in the regressions with interactions are insignificant. As is outlined in Appendix 4.A.6, the peer problems score is calculated on the basis of questions regarding the child having friends, liking other children, being bullied by other children and getting on better with adults than children. The effect on the peer problems score is channelled through the child being picked on and getting on better with adults (see regressions in Appendix 4.C). It is possible that these elements of relationships with peers are influenced by language and ability to communicate.

Children from *fully foreign* families also score lower in the emotional symptoms

score, although the evidence is weaker in this case. There is also no indication that this effect is channelled through language.

It is clear from Table 4.7 that the factor driving the result on the total CSBS score is the symbolic composite. The component is aimed to capture children's understanding of words (reaction to own name, understanding of phrases) and object use (appropriate use of objects, ability to stack blocks, interest in playing with objects and pretend playing with toys).

The results suggest that a child from a mixed family scores .369 lower in the symbolic component which is equivalent with 2.5% of the mean score. A child from a fully foreign family scores .610 lower than a native child, which is equivalent to 4% of the score's mean. Both elements of the symbolic composite are negatively affected by the child's status; children from *mixed and foreign families* know a lower number of words and are less likely to use objects appropriately (see detailed analysis by question in Appendix 4.C). There is no strong evidence to suggest that language is a channel of the effect here.

The data set also provides an alternative measure for CSBS, which takes a form of a dummy variable equal to 1 if child falls into the "concern group" given the assessment's threshold points.[11] Using this variable as a benchmark indicator of performance may be more suitable as, although still arbitrary, it highlights a more important aspect of the assessment - whether the children are performing well below the average. The results of the regressions can also be found in Table 4.7 and suggest that bilingual children are more likely to fall into a concern group with respect to most of the elements of CSBS assessment. This is consistent with the findings for the overall score which, although insignificant, may be indicative of bilingual children's poorer performance in the test. This may be suggesting that children fall behind. Once again, it is likely due to the language acquisition process. One should remember, however, that this measure is self-reported and taken at the age of 2 only. Therefore, it is less reliable than the BAS scores.

[11] According to Wetherby and Prizant (2001) criterion levels for concern are set at more than 1.25 standard deviation below the mean.

Table 4.4: Regression results for SDQ total score

	Dependent variable: SDQ total score						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	OLS	OLS	OLS	OLS	OLS	OLS
bilingual	.316*	.102	.144	.199	1.91	.210	.380
	(.184)	(.189)	(.183)	(.221)	(.221)	(.219)	(.264)
mixed family	.134	.131	.111	.186	.195	.198	.413
	(.140)	(.141)	(.129)	(.164)	(.164)	(.163)	(.561)
fully foreign family	-.009	-.032	-.217	-.332	-.326	-.343	.370
	(.334)	(.349)	(.362)	(.424)	(.419)	(.415)	(.998)
mixed*bilingual							-.447
							(.579)
fully foreign*bilingual							-.201
							(.900)
Controls							
household controls	no	yes	yes	yes	yes	yes	yes
activities at home	no	no	yes	yes	yes	yes	yes
outings	no	no	no	yes	yes	yes	yes
physical activity	no	no	no	no	yes	yes	yes
discipline	no	no	no	no	no	yes	yes
N	7155	6971	6605	3975	3974	3974	3974
R-squared	.397	.419	.554	.659	.660	.662	.662

Data: Growing Up in Scotland, ScotCen and the Scottish Government
SDQ is the Strengths and Difficulties Questionnaire measured at the ages of 4, 5 and 6.
All regressions include gender, sweep and the cohort dummy as controls. Household controls include number of siblings, parental education, NS-SEC household classification, equivalised income, geographical area and whether one is a single parent. Activities at home include frequency of painting/drawing, practising rhymes, practising letters, use of computer and TV. Outings include visits to many attractions, including museums, library, etc. Physical activity includes swimming, playing in the park, running, etc. And disciplining techniques include time out and naughty step. The independent variable mixed family is a dummy equal to 1 if one of child's parents was born outside of the UK and zero otherwise. The variable fully foreign family is a dummy equal to 1 if both child's parents were born outside of the UK and zero otherwise. Bilingual is a dummy variable equal to 1 if child speaks English and another language or just another language at home.
Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 4.5: Regression results for SDQ clusters

Dependent variable	Emotional symptoms		Conduct problems		Hyperactivity		Peer problems		Pro-social	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
mixed family	.051	.227	.031	.049	.074	.002	.007	.140	-.116*	-.282
	(.058)	(.196)	(.054)	(.210)	(.083)	(.260)	(.046)	(.174)	(.069)	(.203)
fully foreign family	-.297***	-.387*	-.102	.053	-.077	.071	.239*	.725**	-.066	-.326
	(.113)	(.225)	(.136)	(.346)	(.233)	(.573)	(.131)	(.359)	(.164)	(.487)
bilingual	.129*	.203	.039	.062	-.074	-.093	.116*	.226**	-.038	-.142
	(.069)	(.124)	(.083)	(.094)	(.102)	(.131)	(.070)	(.115)	(.078)	(.121)
mixed*bilingual		.011		-.090		-.067		-.302		.183
		(.157)		(.188)		(.289)		(.211)		(.274)
fully foreign*bilingual		-.148		-.018		.058		-.126		.148
		(.162)		(.171)		(.208)		(.151)		(.163)
N	3974	3974	3974	3974	3974	3974	3974	3974	3974	3974
R-squared	.551	.551	.647	.647	.682	.682	.514	.545	.937	.937

Data: Growing Up in Scotland, ScotCen and the Scottish Government
SDQ is the Strengths and Difficulties Questionnaire measured at ages 4, 5 and 6.
In this setup the regression controls include all: household controls, activities at home, outings, physical activity and discipline. All regressions include gender, sweep and the cohort dummy as controls. Household controls include number of siblings, parental education, NS-SEC household classification, equivalised income, geographical area and whether one is a single parent. Activities at home include frequency of painting/drawing, practising rhymes, practising letters, use of computer and TV. Outings include visits to many attractions, including museums, library, etc. Physical activity includes swimming, playing in the park, running, etc. And disciplining techniques include time out and naughty step.
The independent variable mixed family is a dummy equal to 1 if one of child's parents was born outside of the UK and zero otherwise. The variable fully foreign family is a dummy equal to 1 if both child's parents were born outside of the UK and zero otherwise. Bilingual is a dummy variable equal to 1 if child speaks English and another language or just another language at home.
Errors are clustered at individual level.
Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 4.6: Regression results for CSBS total score

Dependent variable: CSBS total score						
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS
mixed family	-.312	-.485	-.620	-.734*	-.773**	-1.357
	(.342)	(.363)	(.381)	(.381)	(.381)	(1.271)
fully foreign family	-.188	.130	-.863	-.778	-.801	-1.939
	(.899)	(.875)	(.922)	(.907)	(.901)	(2.387)
bilingual	-1.221**	-1.201**	-.559	-.344	-.342	-.758
	(.500)	(.536)	(.470)	(.469)	(.468)	(.876)
mixed*bilingual						.770
						(1.294)
fully foreign*bilingual						.536
						(1.082)
Controls						
household controls	no	yes	yes	yes	yes	yes
activities at home	no	no	yes	yes	yes	yes
outings	no	no	no	yes	yes	yes
discipline	no	no	no	no	yes	yes
N	4198	3883	3106	3106	3106	3106
R-squared	.021	.044	.124	.132	.135	.135

Data: Growing Up in Scotland, ScotCen and the Scottish Government
CSBS is the Communication and Symbolic Behaviour Score tested at the age of 2. In this setup the regression controls include all: household controls, activities at home, outings, physical activity and discipline. All regressions include gender, sweep and the cohort dummy as controls. Household controls include number of siblings, parental education, NS-SEC household classification, equivalised income, geographical area and whether one is a single parent. Activities at home include frequency of painting/drawing, practising rhymes, practising letters, use of computer and TV. Outings include visits to many attractions, including museums, library, etc. Physical activity includes swimming, playing in the park, running, etc. And disciplining techniques include time out and naughty step.
The independent variable mixed family is a dummy equal to 1 if one of child's parents was born outside of the UK and zero otherwise. The variable fully foreign family is a dummy equal to 1 if both child's parents were born outside of the UK and zero otherwise. Bilingual is a dummy variable equal to 1 if child speaks English and another language or just another language at home.
Errors are clustered at individual level.
Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 4.7: Regression results for CSBS composites

Dependent variable	Social		Speech		Symbolic		Whether child falls in the concern group: Social	Speech	Symbolic	Total score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS	OLS	OLS	OLS	OLS	OLS	probit	probit	probit	probit
mixed family	-.169	-.580	-.257	-.474	-.369***	-.283	-.026	-.019	-.019	.003
	(.202)	(.704)	(.173)	(.519)	(.125)	(.395)	(.088)	(.085)	(.092)	(.088)
fully foreign family	-.285	.382	-.043	-1.656	-.610*	-.787	-.207	.029	-.050	-.179
	(.482)	(1.211)	(.452)	(1.345)	(.352)	(.684)	(.208)	(.194)	(.216)	(.210)
bilingual	-.045	-.176	-.035	-.299	-.197	-.169	.187***	.133	.244**	.293***
	(.252)	(.512)	(.194)	(.342)	(.149)	(.229)	(.106)	(.100)	(.108)	(.106)
mixed*bilingual		-.264		.937		.074				
		(.714)		(.657)		(.355)				
fully foreign*bilingual		.338		.226		-.070				
		(.608)		(.425)		(.321)				
N	3144	3144	3304	3304	3296	3296	5542	5542	5542	5542
R-squared	.027	.027	.132	.133	.168	.170	.299	.256	.350	.284

Data: Growing Up in Scotland, ScotCen and the Scottish Government

CSBS is the Communication and Symbolic Behaviour Score tested at the age of 2.

In this setup the regression controls include all: household controls, activities at home, outings, physical activity and discipline.

All regressions include gender, sweep and the cohort dummy as controls. Household controls include number of siblings, parental education, NS-SEC household classification, equivalised income, geographical area and whether one is a single parent. Activities at home include frequency of painting/drawing, practising rhymes, practising letters, use of computer and TV. Outings include visits to many attractions, including museums, library, etc. Physical activity includes swimming, playing in the park, running, etc. And disciplining techniques include time out and naughty step.

The independent variable mixed family is a dummy equal to 1 if one of child's parents was born outside of the UK and zero otherwise. The variable fully foreign family is a dummy equal to 1 if both child's parents were born outside of the UK and zero otherwise. Bilingual is a dummy variable equal to 1 if child speaks English and another language or just another language at home.

Errors are clustered at individual level.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Variation by age and gender

Given that I find impacts on BAS scores, I allow them to differ by age and gender to see whether there are differential impacts in these two dimensions. The results are presented in Table 4.8.

I find differential impacts on the outcomes of the English Vocabulary Naming exercise. In particular, when including an interaction term of family composition or language with gender, I find that girls outperform boys, scoring almost 6% better. The difference does not depend on language status or family composition.

Children's performance improves with age (see column (1), Table 4.8). On average all children improve their scores by almost 7% per each year of life. Mixed bilingual children perform worse than native monolingual children at the age of 3 but catch up by over 4% a year. This means that the gap closes by the time they are five years old. There is some evidence that bilingual children who have two foreign parents perform worse than native children in the English Vocabulary Naming Exercise and that their performance does not improve, although the coefficients are insignificant. The results for this group may be inconclusive due to a small number of observations in the sample.

I find no differential impacts on Picture Similarities score which further confirms the initial finding that children score comparably in the exercise, irrespective of their background. As expected, however, the score improves with the child's age and girls score higher than boys.

When analysing the respondent-assessed child's ability to be understood by strangers, I conclude that gender does not matter for performance of children from mixed families, although girls in general are reported to be easier understood. There is some evidence of a change in impact due to age; children from mixed or foreign families, as well as bilingual children are less likely to be understood by strangers but the situation may be improving with age, more than for native children. The coefficients in the regression are insignificant.

I considered differential impacts by gender for SDQ and CSBS score and by age for SDQ scores,[12] but found no significant effects.

[12]Note that I cannot investigate whether CSBS score differ by age as I only have a cross-sectional measure at hand.

Table 4.8: Outcome dynamics - differences by age and gender

	BAS English Vocabulary Naming Exercise		BAS Picture Similarities Score		Can be understood by strangers	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	OLS	OP	OP
bilingual	-3.121	-3.142	-1.101	-1.087	.004	.004
	(3.216)	(3.219)	(2.642)	(2.641)	(.204)	(.204)
mixed family	8.651	13.156**	.449	-.441	.043	-.056
	(5.343)	(5.245)	(4.533)	(4.430)	(.284)	(.282)
foreign family	-2.962	-.448	10.375	12.981	.536	.388
	(13.351)	(12.766)	(9.748)	(8.396)	(.660)	(.611)
bilingual*mixed	-10.512**	-10.476**	-1.204	-1.281	.049	.049
	(4.159)	(4.136)	(3.562)	(3.555)	(.238)	(.238)
bilingual*foreign	-8.284	-8.214	-4.139	-3.986	-.038	-.019
	(6.081)	(6.083)	(5.011)	(5.009)	(.350)	(.348)
age	6.875***		9.165***		-.602***	
	(.471)		(.469)		(.055)	
mixed*age	4.330**		1.646		-.193	
	(2.118)		(2.179)		(.121)	
foreign*age	2.390		2.112		-.448	
	(5.089)		(6.091)		(.311)	
female		6.066***		3.062***		-.288***
		(.834)		(.762)		(.044)
mixed*female		-5.278		4.027		.073
		(2.949)		(2.620)		(.141)
foreign*female		-3.554		-4.680		-.104
		(7.669)		(6.234)		(.446)
N	3974	3974	3974	3974	3974	3974
R-squared	.225	.225	.199	.199	.094	.094

Data: Growing Up in Scotland, ScotCen and the Scottish Government

OP stands for ordered probit; the possible answers to a question about other people's understanding of a child were 1 - mostly, 2 - sometimes, 3 - rarely

Note: All regressions include gender, sweep and the cohort dummy as controls. Household controls include: number of siblings, parental education, NS-SEC household classification, equivalised income, geographical area and whether one is a single parent. Activities at home are frequency of painting/drawing, practising rhymes, practising letters, use of computer and TV. Outings include visits to many attractions, including museums, library, etc. Physical activity includes swimming, playing in the park, running, etc. And disciplining techniques include time out and naughty step.

The independent variable mixed family is a dummy equal to 1 if one of child's parents was born outside of the UK and zero otherwise. The variable fully foreign family is a dummy equal to 1 if both child's parents were born outside of the UK and zero otherwise. Bilingual is a dummy variable equal to 1 if child speaks English and another language or just another language at home.

Errors are clustered at individual level.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Foreign parents can help their children catch up

In Section 4.4.2 I noted that children from *mixed* families catch up with *native* children in the Vocabulary Naming Exercise as they grow. What drives this convergence? Is it a natural process related to simultaneous acquisition of two languages or are there different returns to parental inputs into children's upbringing, depending on the family composition? Do foreign parents engage in different types of activities with their children because they have group-specific returns to cognitive and non-cognitive skills?

I find that a child's involvement in various activities can explain children's outcomes to an extent, confirming the role nurture plays in a child's development (Cunha et al., 2010). Is this the channel through which children catch up? To shed light on this matter, I investigate further interactions of socio-economic factors, as well as indicators of various investments in children, with child's family and language status in relation to the BAS (English) Vocabulary Naming score. I do not present such analyses for the Picture Similarities, respondent-assessed speech development and SDQ, as I found no robust impacts on these outcomes so far. I also do not analyse the CSBS score in much detail as it is a cross-sectional measure taken at the age of 2, so it is difficult to claim that any investments would have already paid off. I present the results in Table 4.9.

Firstly, most variables I consider matter for children's performance in the exercise. In particular, higher socio-economic classification of the family (NS-SEC categorisation and parental education) improves the score. Similar observations can be made about various activities children engage in. However, not all these factors have a differential impact on performance of *native* and *mixed, foreign* family or bilingual children.

There is some evidence that bilingual children and children from mixed and foreign families who frequently practice letters gain more than the equivalent native group. This is particularly the case for the bilingual and mixed-family children. The result for children from the fully foreign families, although large, is statistically insignificant, most likely due to a low number of observations in the data.

I also find that bilingual and mixed-family children benefit from outings, for example visits to the zoo. It may be because such activities provide them with an opportunity to interact with other children and grown ups, improving their language skills.

Table 4.9: Contribution of other factors to performance of children in English Vocabulary Naming Exercise

	Dependent variable: BAS Vocabulary Naming score															
interacted variable	degree education		no qualifications		NS-SEC category		drawing		singing		practicing letters			visits to the zoo		
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
bilingual	-9.209***	-7.872***	-8.703***	-8.624***	-8.575**	-10.315***	-8.619*	-9.465**	-7.488	-9.139**	-11.791***	-15.080***	-9.787***	-13.336***	-13.113***	-9.082***
	(2.683)	(2.318)	(2.094)	(2.057)	(3.959)	(3.244)	(4.583)	(3.848)	(6.088)	(4.442)	(4.429)	(3.647)	(2.335)	(3.457)	(2.730)	(1.981)
mixed family	.588	-.660	.104	-.289	.172	.211	3.422	2.305	-7.488	2.499	-2.758	2.321	-3.570	-5.068	-1.661	-7.101**
	(2.318)	(1.658)	(1.679)	(1.652)	(3.294)	(1.996)	(4.067)	(2.005)	(5.198)	(1.997)	(3.893)	(2.007)	(3.500)	(3.408)	(1.627)	(3.145)
fully foreign family	-4.733	-8.241*	-8.560*	-8.102*	-17.273**	-9.757**	-16.498*	-10.389***	-19.864**	-9.702**	-16.405**	-10.116**	-18.539***	-7.190	-12.496***	-12.053**
	(5.657)	(4.549)	(4.768)	(4.493)	(8.730)	(4.891)	(8.621)	(4.806)	(9.671)	(4.846)	(7.443)	(4.811)	(5.904)	(6.860)	(4.469)	(5.681)
interacted var	3.363	5.607*	-13.188**	-10.419**	-3.391**	-4.322***	-.207	-.417	1.499	1.173	-.868	-1.513***	-.421**	-4.428	-3.957	1.441
	(4.035)	(3.149)	(6.331)	(4.070)	(1.476)	(1.008)	(.904)	(.705)	(1.052)	(.761)	(.879)	(.674)	(.179)	(4.311)	(3.327)	(.911)
bilingual*interacted	.749	-1.662	7.153	4.094	-.481	.452	-.279	-.079	-.411	-.109	.441	1.143*		5.788	5.491*	
	(3.869)	(2.818)	(6.036)	(3.387)	(1.433)	(.915)	(.874)	(.640)	(1.023)	(.692)	(.852)	(.612)		(4.156)	(3.032)	
mixed * interacted	-2.445		-10.099		.031		-.236		-.505		1.076		1.247*	4.266		6.790**
	(3.293)		(8.534)		(1.497)		(.814)		(.882)		(.736)		(.635)	(3.791)		(3.397)
foreign * interacted	-7.626		.278		3.258		1.394		2.019		1.382		1.850	-8.546		-2.438
	(9.221)		(11.260)		(2.705)		(1.667)		(1.710)		(1.595)		(1.184)	(8.989)		(7.046)
All controls included	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
N	3975	3975	3975	3975	3975	3975	3975	3975	3975	3975	3975	3975	3975	3975	3975	3975
R-squared	.221	.221	.221	.222	.221	.221	.221	.221	.221	.222	.221	.221	.221	.189	.189	.189

Data: Growing Up in Scotland, ScotCen and the Scottish Government

Note: All regressions include gender, sweep and the cohort dummy as controls. Household controls include: number of siblings, parental education, NS-SEC household classification, equivalised income, geographical area and whether one is a single parent. Activities at home are frequency of painting/drawing, practising rhymes, practising letters, use of computer and TV. Outings include visits to many attractions, including museums, library, etc. Physical activity includes swimming, playing in the park, running, etc. And disciplining techniques include time out and naughty step.

The independent variable mixed family is a dummy equal to 1 if one of child's parents was born outside of the UK and zero otherwise. The variable fully foreign family is a dummy equal to 1 if both child's parents were born outside of the UK and zero otherwise. Bilingual is a dummy variable equal to 1 if child speaks English and another language or just another language at home.

Errors are clustered at individual level.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

4.5 Discussion and limitations

Summary

I present evidence for existence of early life performance gaps between children in foreign/bilingual and native families, using data for Scotland. I show that families differ at the outset in their socio-economic characteristics, views and attitudes as well as lifestyle, measured by types of activities they engage in. Children perform comparably on an array of measures, including cognitive (Picture Similarities), non-cognitive (Strength and Difficulties Questionnaire) and motor development.

Where the differences do emerge (Vocabulary Naming, CSBS and speech assessment), the outcomes are likely to be related to speech and linguistic skills. The bilingual families are, however, a heterogeneous group and I find that children with two foreign-born parents are at a particular disadvantage. Bilingual children from mixed-nationality families do not fall far behind the native children in the English Vocabulary Naming Exercise.

Analysing further the affected outcomes, I find convergence with age among the mixed-nationality children. Its rate is sufficient for the gap to close fully by the age of 5, the last point for which data are available.

Given the initial differences between the families, especially in socio-economic situation, the question arises whether the various investments parents make in children affect them differently and hence narrow or widen the gap. I find that, for mixed family and bilingual children, participation in educational activities, such as practicing letters, and visits to various places, e.g. zoo, contribute more to their English Vocabulary Naming Score than for native children. The evidence for fully foreign families is weaker; this may be due to a small number of observations in the sample.

The difference between *mixed* and *fully foreign* families comes clearly to light in the analysis. Even though children in both types of families are affected in the same way, those who grow up in households with two foreign-born parents are at a greater disadvantage at an early age. This may be because *fully foreign* families are less assimilated with the society in the receiving country and may be less able to provide the child with all necessary support to learn English.

Limitations

The effects may not be causal and the implications are more likely qualitative. I rely on the least squares regression as the main analytical approach and caution is needed when drawing conclusions as various forms of selection and unobserved heterogeneity may invalidate the results.

The OLS regression coefficients will be biased if there is an unobserved heterogeneity among children, which is crucial for their performance and correlated with the explanatory variables already included in the regression. Given what we know about the role of various factors in explaining children's early life performance, inclusion of a

rich set of controls capturing the socio-economic situation of the families, investments made in children, as well as views and attitudes of their parents, may significantly limit the extent of the problem. Naturally, factors such as ability or IQ remain excluded but they are proxied with the variables contained in the regression. The fact that the results are robust to the inclusion of further controls is also reassuring.

Selection

Nonetheless, various forms of selection still pose a problem. The authors of the data set provide an extensive evidence that the participants were randomly selected from the universe of families with children of relevant age living in Scotland (Bradshaw et al., 2009). They also propose weights to correct for non-response and attrition in the overall sample, which I apply to all elements of the analysis.

However, immigrants are a selected group from the populations in their countries of origin. This should not be of particular concern here as I compare the immigrants to natives of the receiving, rather than sending country.

Further, among those who emigrate, the more educated have a higher propensity to intermarry (Meng and Gregory, 2005; Lichter and Qian, 2001), forming the *mixed* families. The positive selection may lead to an upward bias in the estimates as children of more educated parents are likely to perform better, which would close the gap between the *mixed* and *native* children. This becomes partially visible when I split the group into children with one and two foreign-born parents and a gap emerges between the two groups. However, the effect on children in *fully foreign* families is most likely different not only due to differences in socio-economics between the two groups (selection), but also due to the role cultural knowledge and assimilation play in child's upbringing. Children who have one foreign parent still are usually exposed to two cultures and languages, which is why I find a consistent effect across the two groups for measures of development which are related to linguistic ability, but they are in a better position to adapt. Hence, I correctly identify the aspects of child's development affected by family composition or bilingualism, but the size of the impact for *mixed* families may be dampened by positive selection.

Lastly, parental decision to use two rather than one language when communicating with a child is endogenous. There may be various reasons for which parents decide to raise their child bilingually and if the decision hinges on particular characteristics of the family or the child which are correlated with the outcome of interest and the explanatory variable, the estimate of the relationship between the bilingualism and children's skills will be biased.

The scale of the issue and the sign of the arising bias is difficult to assess. I focus on three specific groups: native, mixed and fully foreign. Among them hardly any families with two UK-born children are bilingual. About 40% of the mixed and foreign families use English and another language at home. Thus, I am predominantly interested in the mixed families as almost no native families are bilingual and there are not many fully

foreign families in the sample. The question is what makes parents in mixed and foreign families speak two languages and whether they differ significantly from the parents who teach their children only English.

Hypothetically one could imagine that more educated parents may want to raise bilingual children as they realise the value learning two languages as a child may have. If so, then the group of bilingual children may be positively selected relative to peers and the gaps I find between mono- and bilingual children may be underestimated by OLS.

At the same time, parents may not want to use a language other than English at home for various reasons; they may worry that the child will not be accepted by peers or not see a point in teaching the child a rare (not useful) language. They may also no longer have links with the country of origin, which would provide motivation for the child to speak the language in case of travel. It is difficult to establish whether these considerations would translate to differences between the groups in characteristics which are key for language acquisition and skill development of children. Predicting the direction of potential bias in this case is also difficult.

For example, more assimilated parents may be considering arguments mentioned above and not teach their children their mother tongue. However, parents in the mixed families (who are perceived as more assimilated) are positively selected relative to the other two groups in the study. I have argued that more educated parents are more likely to teach their child two languages. Hence, the two scenarios clash.

Raw comparisons of socio-economic outcomes for bilingual and monolingual families in the three groups do not shed a light on the issue. It appears that bilingual mixed families do not differ from monolingual families in terms of family structure or parental education. However, a higher proportion of mothers in bilingual mixed families never worked.

If main attributes related to selection can be captured or proxied by the observable characteristics I have information about, the selection is explicitly controlled for in the regressions. In fact, inclusion of a rich set of controls does not change the results of the analysis.

Unfortunately, I cannot control for unobserved factors related to the choices to emigrate, to marry a foreigner or to teach a child another language, such as motivation or drive. Whilst I argued that selection into migration may not be overly relevant here and into marriage is possibly not very large[13], I cannot make such conclusions with respect to the parental choice of the language used at home.

Attrition and non-response

A higher attrition rate among the *mixed and foreign* families compared to natives is also a concern, if the characteristics related to attrition differ among the natives and

[13]Recall that the profile of respondents in *mixed* families is similar to that of a migrant to Scotland in general

foreigners. In such a case the weights proposed in the data will not correct sufficiently for the dropout rate among migrants. I discuss the problem in Appendix 4.A.4 and argue that, although for all groups attrition is related to lower socio-economic characteristics, the differences are smaller among the *mixed* families. If the weights applied do not correct for the fact that *mixed* families with lower socio-economic status drop out of the study more, then, as with positive selection, the gap between *mixed* and *native* families will be smaller than it in fact is. Attrition among *mixed families* is closer to random, though, suggesting that the bias should be small. Moreover, the results in the paper do not change in size or scale when I do not weight the data.

I also have no information about the initial non-response rate of the group when first contacted by the project organisers. Given that weights were created on the basis of modelling which, for the first wave, only took into account respondent's age, gender and number of children in the household, the situation will be remedied only if *native*, *mixed* and *fully foreign* families responded similarly to the project.

Comparison group

One may ask whether there exists a suitable comparison group for mixed and fully foreign families and, if so, are native families best to compare to. The choice is debatable, but I argue that this is potentially the best existing group. All children participating in the study were born in Scotland and share similar environmental factors (neighbourhood, schooling, policies) from birth onwards. The main difference between them is the origin of their parents. In fact, in the majority of cases they have one British-born parent and the difference really stems from the cultural and national background of the other parent. Hence, I find them more suited for comparisons than, for example, children from foreign parents' sending countries who were subject to other institutional and cultural factors, parental nationality aside. Duncan and Trejo (2011) compare second generation immigrants from *mixed* and *fully foreign* families to assess the role of assimilation in development of skills; the extent to which I am able to follow this approach is limited as there are barely any children with two foreign parents in the GUS data.

Problematic measures?

Some of the development measures used in the analysis may be seen as subjective and hence not representative. However, the main indicators I consider, BAS assessments, are an objective evaluation of child's cognitive performance and are used in the literature as standard. Further, when used in regressions, results based on them are robust to addition of controls.

Despite these concerns, this paper provides a valuable contribution to understanding whether bilingual children are at a disadvantage at an early age and if language skills can affect outcomes differently depending on the family situation, i.e. whether children

come from *mixed* and *foreign* families. It fills in the gap in our knowledge about the human capital development process by providing an insight into the evolution until the age of 5. Small bilingual children perform worse than natives, which would confirm that a performance gap does emerge early (Heckman and Conti, 2012), but only in some aspects of cognition. Moreover, parents seem to address the initial impediment and the children's skills gradually converge. Moreover, bilingual children who have two foreign-born parents are additionally disadvantaged.

The affected outcomes are linked to linguistic ability, reinstating the role of language in the gap (Dustmann et al., 2010b). Unlike in previous studies, however, children from *mixed* families in the sample are either English speakers or bilingual, which is one of the reasons why they may catch up with peers with time. The narrowing of the gap with age is consistent with the linguistic literature arguing that bilingual children are at a disadvantage, if at all, only early on in life (Baker, 1999).

Like me, Reardon and Galindo (2009) also argue that second generation minority immigrants in the US catch up with the native children as they grow up. The finding of a closing gap is in line with the little that we know about second generation immigrants in the UK. Dustmann et al. (2010b), looking at older children, find that the performance gap closes with age and varies across minorities.

My analysis also reinstates the importance of household income, education and investments parents make in their children (Keane and Fiorini, 2012; Ermisch, 2008; Hartas, 2011) and that they play a greater role for children from *mixed* families.

In their research, Duncan and Trejo (2011) suggest that, in the US, second generation immigrants from mixed marriages perform better relative to those from fully foreign families. I reach similar conclusions for children in Scotland and propose a further argument that children from *mixed* families do not lose out relative to native children. However, I focus on a broadly defined group which may be masking heterogeneity related to one's origins.

Even if qualitative in nature, this analysis constitutes a starting point on the way to defining when exactly the educational gap may be emerging, what drives the differences and which factors play a role in narrowing it. It is hoped to shed light on early years' gaps and whether exposure to two cultures and languages fosters or hinders child's development.

Scotland has experienced a new wave of migration since the data was collected and it would be ideal to undertake a similar analysis for children who participated in Birth Cohort 2, as their parents are more likely to be new immigrants. Their length of stay in the UK may be key for child's development since assimilation takes time and parents may lack Scotland-specific 'cultural knowledge'. The composition of the migrant group may have changed as well, mostly in terms of socio-economic characteristics.

Controlling for parental country of origin would be another extension adding an insight into the types of culture which matter for children's upbringing. Reardon and Galindo (2009) have argued that significant variation exists within Latino groups which

is key for the children's outcomes. Similar considerations should apply here. However, at the moment data does not allow for such distinctions.

Bibliography

AIZER, A., AND F. CUNHA (2012): "The production of human capital: endowments, investments and productivity.," NBER Working Paper Series 18429.

ALLEN, W. (2013): "Geographical Distribution and Characteristics of International Migration Flows to Scotland," Discussion paper, The Migration Observatory, Oxford.

APPS, P., ET AL. (2012): "The impact of pre-school on adolescents' outcomes: evidence from a recent English cohort," IZA Discussion Paper 6971.

BAK, T., ET AL. (2014): "Does bilingualism influence cognitive aging?," *Annals of Neurology*, 00, 1–5.

BAKER, C. (1999): *Foundations of Bilingual Education and Bilingualism*. Multilingual Matters, Bristol.

BECKER, B. (2010): "Who Profits Most from Early Parental Investments? The Effects of Activities Inside and Outside the Family on German and Turkish Children's Language Development," *Child indicators research*, 3, 29–46.

BEHRMAN, J. R., ET AL. (2014): "What determines adult cognitive skills? Influences of pre-school, school, and post-school experiences in Guatemala," *Latin American Economic Review*, 23(1).

BIALYSTOK, E. (2009): "Bilingualism: the good, the bad and the indifferent," *Bilingualism: language and cognition*, 12, 3–11.

BJORKLUND, A., AND K. SALVANES (2010): "Education and Family Background: Mechanisms and Policies," IZA Discussion Paper 5002.

BLACK, S., AND P. DEVEREUX (2011): "Recent Developments in Intergenerational Mobility," in *Handbook of Labor Economics*. Elsevier.

BLACK, S., ET AL. (2005): "Why the apple doesn't fall far: understanding intergenerational transmission of human capital," *American Economic Review*, 95, 437–449.

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

BORGHANS, L., ET AL. (2008): "The economics and psychology of personality traits," NBER Working Paper 13810.

BRADSHAW, P., ET AL. (2009): "Growing Up in Scotland, Sweep 5. User guide," Scottish Centre for Social Research.

BROOKS-GUNN, J., AND L. MARKMAN (2005): "The Contribution of Parenting to Ethnic and Racial Gaps in School Readiness," *The Future of Children*, 15, 139–168.

CARNEIRO, P., ET AL. (2007): "Impact of early cognitive and non-cognitive skills on later outcomes," Discussion paper.

——— (2013): "Maternal education, home environments, and the development of children and adolescents.," *Journal of the European Economic Association*, 11, 123–160.

CHISWICK, B., AND C. HOUSEWORTH (2011): "Ethnics intermarriage among immigrants: human capital and assortative mating," *Review of Economics of the Household*, 9(2), 149–180.

CUNHA, F., ET AL. (2010): "Estimating the technology of cognitive and noncognitive skill formation.," NBER Working Papers 15664.

DE PHILIPPIS, M. (2014): "Parents' Country of Origin and Intergenerational Mobility of School Performance.," unpublished.

DOCQUIER, F., AND A. MARFOUK (2006): "International migration by educational attainment (1990-2000) - Release 1.1," in *International Migration, Remittances and Development*, ed. by C. Ozden, and M. Schiff. Palgrave Macmillan, New York.

DRONKERS, J., AND M. DE HEUS (2012): "The Educational Performance of Children of Immigrants in Sixteen OECD Countries," Discussion paper.

DUNCAN, B., AND S. TREJO (2011): "Intermarriage and the intergenerational transmission of ethnic identity and human capital for Mexican Americans," *Journal of Labor Economics*, 29, 195–227.

DUSTMANN, C., ET AL. (2010a): "Ethnicity and educational achievement in compulsory schooling," *The Economic Journal*, 120, F272–F297.

——— (2010b): "Ethnicity and second generation immigrants in Britain," Discussion paper.

——— (2012): "Educational achievement of second-generation immigrants: an international comparison," *Economic Policy*, 27, 143–185.

EIRICH, F. (2011): "Characteristics and intentions of immigrants to and emigrants from Scotland - review of existing evidence," Discussion paper, Scottish Government Social Research, Edinburgh.

ERMISCH, J. (2008): "Origins of Social Immobility and Inequality: Parenting and Early Child Development," *Journal of the National Institute of Economic and Social Research*, 205, 62–71.

FEINSTEIN, L. (2003): "Inequality in the Early Cognitive Development of British Children in the 1970 Cohort," *Economica*, 70, 73–97.

FRYER, R. G., AND S. D. LEVITT (2004): "Understanding the black-white test score gap in the first two years of school," *The Review of Economics and Statistics*, 86, 447–464.

——— (2006): "The Black-White test score gap through third grade," *American Law and Economics Review*, 8(2), 249–281.

——— (2010): "An Empirical Analysis of the Gender Gap in Mathematics," *American Economic Journal: Applied Economics*, 2(2), 210–240.

FULLER, B., ET AL. (2009): "The Health and Cognitive Growth of Latino Toddlers: At Risk or Immigrant Paradox?," *Maternal and Child Health Journal*, 13, 755–768.

GLICK, J. E., ET AL. (2012): "Migration Timing and Parenting Practices: Contributions to Social Development in Preschoolers With Foreign-Born and Native-Born Mothers," *Child Development*, 83, 1527–1542.

GLICK, J. E., AND B. HOHMANN-MARRIOTT (2007): "Academic Performance of Young Children in Immigrant Families: The Significance of Race, Ethnicity, and National Origins," *International Migration Research*, 41, 371–402.

GREENWOOD, J., ET AL. (2014): "Marry Your Like: Assortative Mating and Income Inequality," Discussion paper.

HANUSHEK, E. A., AND L. WOESSMANN (2009): "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation," IZA Discussion Paper 4575.

HARTAS, D. (2011): "Families' social backgrounds matter: socio-economic factors, home learning and young children's language, literacy and social outcomes," *British Educational Research Journal*, 37, 893–914.

HECKMAN, J., AND Y. RUBINSTEIN (2001): "The importance of noncognitive skills: lessons from the GED testing program," *The American Economic Review*, 91, 145–149.

HECKMAN, J. J., AND G. CONTI (2012): "The economics of child well-being," NBER Working Paper 18466.

HM REVENUE AND CUSTOMS (2012): "Child Benefit, Child Tax Credit and Working Tax Credit Take Up Rates 2010-11," Discussion paper.

HOLMLUND, H., ET AL. (2011): "The causal effect of parents' schooling on children's schooling: a comparison of estimation methods.," *Journal of Economic Literature*, 49, 615–651.

KAMHÖFER, D. (2014): "The Effect of Early Childhood Learning Programs on the Contemporary Formation of Grammar Skills," mimeo.

KAUSHANSKAYA, M., AND V. MARIAN (2007): "Bilingual Language Processing and Interference in Bilinguals: Evidence From Eye Tracking and Picture Naming," *Language Learning*, 57, 119–163.

KEANE, M., AND M. FIORINI (2012): "How the Allocation of Children's Time Affects Cognitive and Non-Cognitive Development," Discussion paper, University of Oxford, Department of Economics.

KORNDER, N., AND J. DRONKERS (2012): "Do Migrant Girls Always Perform Better? Differences between the Reading and Math Scores of 15-Year-Old Daughters and Sons of Migrants in PISA 2009 and Variations by Region of Origin and Country of Destination," Discussion paper.

LANZIERI, G. (2012): "Mixed marriages in Europe, 1990-2010," in *Cross-Border Marriage: Global Trends and Diversity*, ed. by D.-S. Kim, pp. 81–121. Korea Institute for Health and Social Affairs (KIHASA).

LICHTER, D., AND Z. QIAN (2001): "Measuring marital assimilation: intermarriage among natives and immigrants.," *Social Science Research*, 30, 289–312.

MEIER JAEGER, M. (2011): "Does Cultural Capital Really Affect Academic Achievement? New Evidence from Combined Sibling and Panel Data," *Sociology of education*, 84, 281–298.

MELHUISH, E., ET AL. (2008): "Effects of the Home Learning Environment and preschool center experience upon literacy and numeracy development in early primary school," *Journal of Social Issues*, 64, 157–188.

MENG, X., AND R. GREGORY (2005): "Intermarriage and the economic assimilation of immigrants.," *Journal of Labor Economics*, 23, 135–176.

NATIONAL RECORDS OF SCOTLAND (2013a): "2011 Census: Detailed characteristics on Ethnicity, Identity, Language and Religion in Scotland - Release 3A," Discussion paper.

——— (2013b): "2011 Census: Key Results on Population, Ethnicity, Identity, Language, Religion, Health, Housing and Accommodation in Scotland - Release 2A," Discussion paper.

NEAL, D. (2014): "Lectures on the Economics of Education," unpublished draft.

NORDIN, M., AND D.-O. ROOTH (2007): "The Income Gap Between Natives and Second Generation Immigrants in Sweden: Is Skill the Explanation?," Discussion paper.

OLLER, K., AND R. EILERS (2002): *Language and Literacy in Bilingual Children*. Multilingual Matters.

OZDEN, C., ET AL. (2011): "Where on Earth is Every-body? The evolution of global bilateral migration 1960-2000," *World Bank Economic Review*, 25(1), 12–56.

PORTOCARRERO, J., ET AL. (2007): "Vocabulary and verbal fluency of bilingual and monolingual college students," *Archives of Clinical Neuropsychology*, 22, 415–422.

REARDON, S., AND C. GALINDO (2009): "The Hispanic-White Achievement Gap in Math and Reading in the elementary grades," *American Educational Research Journal*, 46(3), 853–891.

RIENZO, C. (2013): "Characteristics and Outcomes of Migrants in the UK Labour Market," Discussion paper, The Migration Observatory, Oxford.

ROSENTHAL, A., ET AL. (1983): "The effect of language background on achievement level and learning among elementary school students," *Sociology of Education*, 56, 157–169.

SANDEFUR, G., AND T. MCKINNELL (1986): "American Indian intermarriage," *Social Science Research*, 347, 347–348.

SORACE, A. (2007): "The more, the merrier: facts and beliefs about the bilingual mind," in *Tall Tales about the mind and the brain: separating fact from fiction*, ed. by S. D. Salla, pp. 193–203. Oxford Scholarship Online.

——— (2011): "Cognitive advantages of bilingualism: is there a bilingual paradox?," in *Multilingualism, language, power and knowledge*, ed. by Valore, pp. 335–358. Pisa:Edistudio.

VARGAS-SILVA, C. (2013a): "Migrants in Scotland: An Overview," Discussion paper, The Migration Observatory, Oxford.

——— (2013b): "Migrants in the Scottish Labour Market," Discussion paper, The Migration Observatory, Oxford.

WETHERBY, A., AND B. PRIZANT (2001): "Communication and Symbolic Behavior Scales Developmental Profile," Paul H. Brookes Publishing Co.

Appendix

4.A Further elaboration on data and variables

4.A.1 Selection for the project

In preparation for the first wave a named sample of approximately 10 700 children was selected from Child Benefit records to give an achieved sample of 8 000 overall. The sampling frame was based on the geographical Data Zones for Scotland used by the Scottish Executive for purposes of releasing the small area statistics. The areas are nested within the Local Authority areas in Scotland and contain between 500 and 1000 household residents each. The zones were aggregated, sorted by Local Authority and the Scottish Index of Multiple Deprivation Score. Of those, 130 areas were selected at random and data for all children fitting the birth date criteria and living within the areas was released by the Department of Work and Pensions (DWP). Within each sample point, all eligible babies and three-fifths of toddlers were selected. Exclusions were made for 'sensitive' cases and children that had been sampled for research by the DWP in the previous 3 years. If more than one child was eligible within a household, one was selected at random.

One concern with such a selection procedure is that potentially not all families residing in Scotland register for Child Benefit, which every child under the age of 16 (or under the age of 20 if in education) is entitled to. In such a case, the initial population which was subjected to selection for the purposes of the project will not be equivalent with the universe of children residing in Scotland. This should not pose problems if we believe that those not claiming the Child Benefit do not differ from the rest of the population. It is unlikely to be the case however. It is reasonable to think that people not claiming the entitlement are either sufficiently well-off not to see a need of doing so or they are under-informed and do not know they can claim the benefit. One could argue that in the case of foreign-born citizens, many may not have sufficient knowledge of the British welfare system and be under-represented in the data. This would be contradicting the observation that the proportion of foreign born respondents in the data is similar to the overall proportion of migrants in Scotland, as registered by 2011 Census. Moreover, according to HM Revenue and Customs (2012), the uptake of the Child Benefit is persistently high, oscillating between 97% in 2006 and 96% in 2010, suggesting that the scale of the problem may be negligible. Analyses for earlier years are not available but are likely to be in line with the information cited here. In this situation, I see this limitation of the sampling procedure as a minor issue, especially given the fact that data was additionally weighted to closely match the population.

4.A.2 Response rates

The response rates within the first sweep reached around 80% of all in-scope. By the final sweep the response rate among those who initially participated in the study falls to 77% for natives in the child cohort and 71% for natives in the birth cohort. The attrition rate for the combined group of *mixed* and *fully foreign* families is larger - 72% of those in the child cohort and 62% of those in the birth cohort initial sample have still participated at a final sweep.

Table 4.A.1: Sample response rates (%)

	Child cohort		Birth cohort	
	native	mixed and foreign	native	mixed and foreign
sweep 1	100 (N=2609)	100 (N=250)	100 (N=4715)	100 (N=502)
sweep 2	87.543	86.400	86.957	82.072
sweep 3	81.794	79.200	81.103	73.506
sweep 4	77.463	71.600	77.243	70.120
sweep 5			74.337	65.339
sweep 6			70.923	62.351

Data: Growing Up in Scotland, ScotCen and the Scottish Government

4.A.3 Weighting procedures

The data was weighted to correct for selection and attrition. Weights were created for the sample after each sweep and different weights are suggested for cross-sectional and longitudinal analyses. I discuss only the panel weights as they are relevant for this study. For more details, please consult Bradshaw et al. (2009).

At every sweep except 1,[14] the weights were based on a response behaviour modelled using a logistic regression. The predictor variables were a set of socio-demographic respondent and household characteristics collected from the previous sweeps. Non-response was associated with the following characteristics in all sweeps: renting the property, not working, being a younger mother (under the age of 20) and living in the 20% of the most deprived Data Zones. The predicted probability of response was then inversed to create the non-response weights. Hence, respondents who had a low predicted probability are allocated a larger weight, increasing their representation in the sample.

The final sweep weight is the product of the sweep's non-response weight and the previous sweep's interview weight. For each cohort the final weights were scaled to the responding sweep sample size to make the weighted sample size match the unweighted sample size.

[14] At sweep 1 there was no prior information about the respondents, so the modelling was based on information from the Child Benefit records, such as age of claimant, sex of claimant, number of children in the household and the method of benefit payment. The other variables were Scottish index of multiple deprivation (quintiles), population density measured by the number of persons in private households per hectare and ONS urban rural indicator.

4.A.4 Attrition among mixed and fully foreign families

The attrition rate among the *mixed and foreign* families is higher than the average attrition in the sample, particularly for the birth cohort. 71 *mixed and foreign* families drop out of the CC by wave 4 and 189 disappear from the BC1. This raises concerns for representativeness of the group, if the weighting applied in the study does not correct sufficiently for it.

Weights are created on a basis of logistic model detecting characteristics of respondents in the sample related to higher likelihood of attrition, which include lower household incomes, lone parent households, households with younger mothers and living in the more deprived areas. The weighting applied to the data will not work well for *mixed and foreign* families if their attrition is driven by different characteristics.

I compare the dropouts to the stayers in the combined sample of *mixed and fully foreign* families to identify differences between them which may be related to attrition, focusing in particular on characteristics identified as correlated with attrition in the overall sample. I find that a higher proportion of dropouts have low household incomes and a higher proportion of respondents who drop out are young mothers. Those dropping out are also more likely to live in more deprived areas of Scotland. The differences within this group, however, are significantly smaller than for the group of native respondents. I also find no difference in % of lone parents among stayers and dropouts, which is identified as a determinant of attrition for the overall sample. Thus, the characteristics related to attrition are more pronounced for natives. Although the patterns are maintained for *mixed and foreign* families, no very clear selection emerges.

Results of a logistic regression of non-response on the family composition (mixed or foreign vs. native) confirm that the *mixed or foreign* families are more likely to disappear from the study. To investigate whether different socio-economic characteristics trigger attrition among families, I replicate the analysis undertaken by the authors of the data, to identify the characteristics correlated with attrition. I then repeat the same analysis on two subsets of data - for *mixed or foreign* and *native* families. I find that different characteristics matter to both groups, although they are all related to lower socio-economic outcomes of families and there is a degree of overlap in factors which matter. The elements also vary in importance - some factors are more influential for attrition among natives than among foreign-born. Overall the associations are weaker for *mixed and foreign* families suggesting that attrition is closer to random than for *native* families. Hence, the weights proposed in the study may not be most suitable for the purpose of my analysis.

The question is whether the weights matter at all then. I repeat all regressions presented in the paper on the unweighted data and find that the results remain unchanged, which is reassuring. All results for this analysis can be provided upon request.

4.A.5 Representativeness of families with at least one foreign parent

The portrait of a *mixed and fully foreign family* in Section 4.3 reflects, or at least does not contradict, what we know about immigrants to Scotland. The information about immigrants to Scotland is rather limited, however. The majority of studies focus on the UK in general, without singling out specific countries (e.g. Rienzo (2013)). Scotland-specific studies mostly provide information about the distribution and flows of immigrants to Scotland (Allen, 2013) or the labour market outcomes of immigrants (Vargas-Silva, 2013a,b), although Eirich (2011) sheds light on characteristics of migrants to and from Scotland, drawing on various UK data sources. The most comprehensive source of information is the 2011 Census, results of which are being gradually released (National Records of Scotland, 2013b,a). Even then, however, very little can be inferred about migrant families as its main focus is to report the migrant stock in various areas of Scotland, migrants' education levels and labour market outcomes. It does encompass the entire legal migrant population resident in Scotland at the time of the Census, but does not (as yet) provide detailed information on migrants' family situation. According to Eirich (2011), 23% of foreign-born residents of Scotland were living in a family with a child. Hence, only about a quarter of the migrant Census respondents constitute a potentially comparable group to the GUS respondents. It must be noted, however, that in *mixed* families usually just one of the parents was born abroad and the *fully foreign* families are a small subgroup in the sample, which further complicates any comparisons. Importantly, Census data capture the situation in Scotland in 2011; the group participating in GUS must have been residing in Scotland already in 2005 when the project started and beforehand, especially since over 98% of children in the sample were born in Scotland.[15] Therefore, any comparisons are very rough.

Nonetheless, according to the 2011 Census, 7% of Scottish residents were born outside of the UK and 5% of children in GUS data have at least one parent born outside of the UK. Further, according to the Census, almost 6% of Scottish residents spoke a *foreign* language at home[16] - exactly the same proportion as in the data I rely on.

The migrant group in the data also seems to approximately match the Scottish migrant population in terms of their socio-economic characteristics. For example, looking at NS-SEC classification of migrants, both males and females are concentrated in the lowest paid (18.2%) and in the two highest paid occupational categories (32.5%) (Vargas-Silva, 2013b). In GUS, respondents and their partners are mostly represented in the professional category (42% and 51% respectively). Still, 22% of respondents and 17% of their partners work in semi-routine and routine occupations. One could

[15]Only 28 children interviewed in wave 1 were born outside of the UK and only 108 were born in other countries in the UK.

[16]In particular, 5.56% of Census respondents aged 3 and over spoke language other than English, Gaelic or Scottish at home. Bear in mind, however, that GUS data does not necessarily exclude Gaelic and Scottish from the "foreign language" category.

argue that the polarisation is less visible in my data, but this may be due to the fact that respondents in GUS are likely to be a specific group of migrants - middle aged, with children, potentially further into their career. Moreover, recent migration from A8 countries following the EU enlargements (2004 onwards) changed the composition of migrant stock in Scotland. The shift may have not been captured in GUS, but is becoming visible in the Census.

Similarities are also visible in terms of education with 50% of recent migrants and 33% of migrants in general in the Census having a degree qualification, compared with 46% of foreign-born respondents in GUS. Moreover, Docquier and Marfouk (2006) estimate that in 1990 40% of the migrants living in the UK had tertiary education. The number reached 49% in year 2000. Although the result is not Scotland-specific, it is in line with what I find in the data.

Despite the limitations,[17] there are some indications that the group of *mixed and fully foreign families* may be representative of the migrant population in Scotland. Their size and percentage speaking foreign language is as expected and they seem similar to migrants in Scotland overall in terms of their education. Larger discrepancies emerge in NS-SEC classification but this may be due to the age structure and professional experience of the group.

[17] 1) limited studies on Scotland, 2) statistics come from various data sources, 3) no focus on migrant or mixed families, 4) many outcomes not comparable and differently defined

4.A.6 Variables of interest

Table 4.A.2: Variables used in the analysis

Variable	Description
Child indicators	
Age	
Gender	
Language spoken at home	The data set contains information whether in the language spoken in the house is: English, English and another language, a language different from English only.
Family socio-economics	
NS-SEC category	National Statistics Socio-economic Classification classifies groups on basis of employment relations, career prospects, autonomy, mode of payment and period of notice. The GUS dataset includes a five category system.
Highest education level	
Equivalised household annual income	Income data collected during the interview are adjusted to reflect the household composition.
Scottish Index of Multiple Deprivation	SIMD identifies areas of concentration of multiple deprivation across Scotland, based on indicatiors such as Current Income, Employment, Health, Education Skills and Training, Geographic Access to Services, Housing and Crime. The zones are then ranked from most deprived to least deprived.
Lone parents	Whether a parent is a lone parent.
Study child's birth order	
Number of siblings	
Parental investment measures	
Activities at home	
how often takes a child to park	
how often takes a child to friends who have children	
how often read books	
painting or drawing last week	
nursery rhymes last week	
letters and shapes practised last week	
frequency played computer last week	
hours of TV watched	
Physical activity	
child rode a bicycle last week	
child kicked a ball last week	

Continued on next page

Table 4.A.2 – continued from previous page

Variable	Description
child danced last week	
child ran/jumped last week	
child swam last week	
Outings	
Has been to library since last year	
Has been to a concert/play since last year	
Has been to a swimming pool since last year	
Has been to a museum, gallery etc. since last year	
Has been to a zoo, aquarium etc. since last year	
Has been to a cinema since last year	
Has been to a sports event since last year	
Disciplining techniques	
rewards/stickers	Respondents were asked whether they used the disciplining techniques with the child (yes/no question).
naughty step	
time out	
Child outcomes	
BAS Picture Recognition	The British Ability Scales aims to measure cognitive abilities and educational achievements of children. The English Vocabulary Naming Exercise assessed a spoken vocabulary and may reflect expressive language ability, vocabulary knowledge of nouns, ability to attach labels to pictures, general knowledge and level of language stimulation. It may reflect non-verbal problem solving, visual perception and analysis, ability to attach meaning to pictures, use of verbal mediation and general knowledge. The scores are expressed as percentiles.
BAS Vocabulary Naming	
Subjective speech assessment:	The assessment is based on asking the respondent to score the level of understanding using the following scale: mostly, sometimes and rarely. It is measured at ages 2,3,4 and 5.
Can child be understood by respondent?	
Can child be understood by family?	
Can child be understood by strangers?	
Communication and Symbolic Behaviour Scales Score	This is an outcome derived only for the Birth Cohort at age 2. It is a measure of non-cognitive skills, assessing the child's communication, emotional development, understanding and interaction with peers. The composite score ranges from 0 to 57 and higher scores indicate better performance.
Strength and Difficulties Questionnaire Score	It is a score based on a behavioural screening questionnaire and measures: emotional symptoms, conduct problems, hyperactivity, inattention, peer relationship problems and pro-social behaviour. The overall score is a sum of all aspects,

Continued on next page

Table 4.A.2 – continued from previous page

Variable	Description
	except the pro-social score, and the higher score indicates greater evidence of difficulties.

Table 4.A.3: Questions asked to compile SDQ and CSBS scores

Panel A: SDQ score	
Total score = emotional symptoms + conduct problems + hyper-activity + peer problems	
Emotional symptoms score	
MSDQ03	X often complains of headaches, stomach-aches or sickness
MSDQ08	X has many worries, often seems worried
MSDQ13	X is often unhappy, down-hearted or tearful
MSDQ16	X is nervous or clingy in new situations, easily loses confidence
MSDQ24	X has many fears, is easily scared
Conduct problems score	
MSDQ05	X often has temper tantrums or hot tempers
MSDQ07	X is generally obedient, usually does what adults request
MSDQ12	X often fights with other children or bullies them
MSDQ18	X often lies or cheats
MSDQ22	X steals from home, school or elsewhere
Hyper-activity score	
MSDQ02	X is restless, overactive, cannot stay still for long
MSDQ10	X is constantly fidgeting or squirming
MSDQ15	X is easily distracted, concentration wanders
MSDQ21	X thinks things out before acting
MSDQ25	X sees tasks through to the end, good attention span
Peer problems score	
MSDQ06	X is rather solitary, tends to play alone
MSDQ11	X has at least one good friend
MSDQ14	X is generally liked by other children
MSDQ19	X is picked on or bullied by other children
MSDQ23	X gets on better with adults than with other children
Pro-social score	
MSDQ01	X is considerate of other people's feelings
MSDQ04	X shares readily with other children (treats, toys, pencils etc.)
MSDQ09	X is helpful if someone is hurt, upset or feeling ill
MSDQ17	X is kind to younger children
MSDQ20	X often volunteers to help others (parents, teachers, other children)
Note: all questions in SDQ had the following possible answers: 1) not true, 2) somewhat true, 3) certainly true	
Panel B: CSBS score	
Social composite = emotion and eye gaze + communication + gestures	
Cluster 1: Emotion and eye gaze	
MCSBS01	Do you know when X is happy and when X is upset?
MCSBS02	When X plays with toys, does he look at you to see if you are watching?
MCSBS03	Does X smile or laugh while looking at you?
MCSBS04	When you look at and point to a toy across the room, does X look at it?
Cluster 2: Communication	
MCSBS05	Does X let you know that he needs help or wants an object out of reach?
MCSBS06	When you are not paying attention to X, does he try to get your attention?
MCSBS07	Does X do things just to get you to laugh?
MCSBS08	Does X try to get you to notice interesting objects - just to get you to look at the objects, not to get you to do anything with them?
Cluster 3: Gestures	

Continued on next page

Table 4.A.3 – continued from previous page

MCSBS09	Does X pick up objects and give them to you?
MCSBS10	Does X show objects to you without giving you the object?
MCSBS11	Does X wave to greet people?
MCSBS12	Does X point to objects?
MCSBS13	Does X nod his head to indicate yes?
	Speech composite = sounds + words
	Cluster 4: Sounds
MCSBS14	Does X use sounds or words to get attention or help?
MCSBS15	Does X string sounds or words together such as uh oh, mama, gaga, bye
MCSBS16	About how many of these sounds does X use: ma, na, ba, da, ga, wa, la, ya, sa , sha?
	Cluster 5: Words
MCSBS17	About how many different words does X use so that you know what he means ?
MCSBS18	Does X put two words together (such as 'more biccies'; bye-bye)?
	Symbolic composite = understanding of words + object use
	Cluster 6: Understanding
MCSBS19	When you call X's name, does he respond by looking or turning toward you?
MCSBS20	About how many different words or phrases does X understand without showing or pointing?
	Cluster 7: Object use
MCSBS21	Does X show interest in playing with a variety of objects?
MCSBS22	About how many of the following objects does your child use appropriately: cup, bottle, bowl, spoon, comb or brush, toothbrush, washcloth, ball, toy vehicle, toy, telephone?
MCSBS23	About how many blocks (or rings) can X stack?
MCSBS24	Does your child pretend to play with toys (for example, feed a stuffed animal, put a doll to sleep, put an animal figure in a vehicle)?

Note: all questions in CSBS have the following possible answers: 1) not yet, 2) sometimes, 3)often

4.A.7 Explanatory variables

Some variables in the data set which I use in regression analysis are specific to the Scottish data. I briefly discuss how they are created and what they reflect. The explanations come from Bradshaw et al. (2009) who provide an overview of all the variables in the data set.

The National Statistics Socio-economic Classification (NS-SEC) is a social classification system that classifies groups on basis of employment relations, including career prospects, autonomy, mode of payment and period of notice. In GUS the classification contains 5 employment categories: managerial and professional, intermediate, small employers and own account workers, lower supervisory and technical, and semi-routine and routine occupations. The data set contains categorisations for the respondent, partner and the household as a whole. I use the household NS-SEC classification in the analysis. Further information on NS-SEC is available from the National Statistics website at:

http://www.statistics.gov.uk/methods_ quality/ns_sec/cat_subcat_class.asp

The Scottish Index of Multiple Deprivation (SIMD) identifies small area concentrations of multiple deprivation across Scotland based on seven individual domains of Current Income, Employment, Health, Education Skills and Training, Geographic Access to Services, Housing and new Crime. It is obtained at data zone level, ranking areas of median population size of 769, from the most deprived to the least deprived. In the dataset, the data zones are grouped into quintiles. Further details on SIMD can be found on the Scottish Government Website:
http://www.scotland.gov.uk/Topics/Statistics/SIMD/Overview

The Equivalised Household Annual Income variable is a household income variable adjusted for a household's size and composition. Official income statistics use the 'Modified OECD' equivalence scale, in which an adult couple with no dependent children is taken as the benchmark with an equivalence scale of one and the scale is adjusted accordingly for other configurations within the household. The distribution of income for the population of the United Kingdom as a whole is taken from the most recent available data from the Family Resources Survey. The data and methodology are the same as those used by the Government in its annual Households Below Average Income publication.

4.B Data summary and statistics

Table 4.B.1: Mixed families and language spoken in the sample

Panel A: Percent of children with a parent born outside of the UK			
	Birth cohort and child cohort		
	foreign mother	foreign father	either parent foreign
sweep1	6%	5%	9%
sweep 2	6%	5%	5%
sweep 3	6%	5%	5%
sweep 4	6%	5%	5%
	Birth cohort only		
sweep 5	7%	6%	6%
sweep 6	6%	5%	5%

Panel B: Language spoken at home			
	only English	English and other	other only
overall	94%	5%	1%
mixed or fully foreign family	51%	40%	9%
native family	99%	1%	0%

Data: Growing Up in Scotland, ScotCen and the Scottish Government

Table 4.B.2: How often do you/ does the child..... ? (%)

take child to a park

	native			mixed			fully foreign		
age	3	5	6	3	5	6	3	5	6
every day/most days	17.38	15.01	10.47	24.14	18.77	10.45	22.22	13.73	6.67
once or twice a week	60.13	55.37	52.72	54.19	50.9	53.36	57.78	45.1	53.33
once a fortnight	12.7	17.12	20.2	13.3	12.64	15.67	4.44	11.76	20
once every 1 or 2 months	5.72	6.76	8.26	3.94	6.14	9.33	8.89	23.53	8.89
once every 3 or 4 months	0.92	1.08	1.71	0.99	2.17	2.24	0	0	2.22
once every 6 months	0.27	0.49	0.6	0	1.81	1.12	0	0	2.22
once a year or less often	0.19	0.26	0.21	0	0	0	0	0	4.44
varies too much to say	1.92	2.54	3.83	1.97	3.61	4.85	4.44	1.96	0
never	0.77	1.37	2	1.48	3.97	2.99	2.22	3.92	2.22

visit friends with children

	native					mixed					fully foreign				
age	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
every day/most days	18.8	14.64	15.18	13.21	9.7	14.58	10.39	12.73	11.76	6.49	6.58	9.26	5.62	7.14	12
once or twice a week	47.41	48.44	49.29	47.14	45.35	47.92	46.67	49.06	47.06	44.16	44.74	44.44	40.45	28.57	48
once a fortnight	13.17	14.98	15.4	16.92	20.15	12.8	18.63	14.61	15.29	20.13	9.21	12.04	22.47	17.86	16
once every 1 or 2 months	8.22	9.5	9.2	9.79	11.49	11.01	9.8	11.48	15.29	15.58	19.74	14.81	12.36	17.86	20
once every 3 or 4 months	2.29	2.43	2.08	2.76	2.38	3.87	3.33	2.3	1.76	3.25	2.63	2.78	3.37	3.57	0
once every 6 months	1.2	1.18	1.23	1.41	1.73	1.79	0.59	1.67	4.12	1.95	0	3.7	1.12	3.57	4
once a year or less often	0.51	0.68	0.62	0.8	0.35	0	1.18	0.63	1.18	0	1.32	1.85	1.12	3.57	0
varies too much to say	2.34	1.74	2.46	2.58	2.97	2.68	2.75	2.71	0.59	3.9	3.95	3.7	3.37	7.14	0
never	6.05	6.41	4.54	5.39	5.89	5.36	6.67	4.8	2.94	4.55	11.84	7.41	10.11	10.71	0

take the child to the library

	native		mixed		fully foreign	
age	1	3	1	3	1	3
every day/most days	0.28	0.35	0.5	0	0.96	0
once or twice a week	3.84	5.87	5.04	9.76	6.73	6.67
once a fortnight	5.24	12	7.81	15.12	8.65	11.11
once every 1 or 2 months	8.59	14.8	8.82	16.1	7.69	17.78

Continued on next page

Table 4.B.2 – continued from previous page

once every 3 or 4 months					2.1	4.41	3.02	5.37	0	0					
once every 6 months					2.16	3.07	2.52	4.88	2.88	2.22					
once a year or less often					1.85	2.65	3.27	2.44	2.88	0					
varies too much to say					0.7	1.5	1.26	3.9	0	0					
never					75.24	55.37	67.76	42.44	72.12	62.22					
read books to child															
	native					mixed					fully foreign				
age	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
never	1.61	0.99	0.64	1.19	3.04	1.79	1.31	1.25	0.89	1.66	4	3.17	2.27	1.27	4.29
1 day a week	1.63	0.97	0.95	1.35	2.05	0.6	0.98	0.63	0.89	1.9	2.67	0	1.14	1.27	2.86
2 days a week	3.59	2.9	2.75	3.26	5.16	3.87	1.97	3.97	2.46	3.32	10.67	1.59	3.41	5.06	7.14
3 days a week	4.05	3.56	3.56	4.29	6.17	4.46	3.93	3.34	5.37	4.74	2.67	6.35	7.95	6.33	5.71
4 days a week	3.68	3.69	3.63	4.63	5.59	1.79	2.95	3.55	4.03	5.21	5.33	1.59	4.55	5.06	5.71
5 days a week	3.81	4.71	6.05	5.96	7.51	2.68	3.93	5.01	5.15	6.64	5.33	7.94	9.09	7.59	8.57
6 days a week	1.12	1.86	1.57	1.67	2.16	1.49	1.31	1.25	1.79	2.37	1.33	0	2.27	2.53	4.29
every day	80.5	81.33	80.85	77.64	68.31	83.33	83.61	81	79.42	74.17	68	79.37	69.32	70.89	61.43
play outdoor games															
	native					mixed					fully foreign				
age	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
never	8.72	6.54	5.27	4.64	6.94	9.52	5.25	5.85	6.47	5.92	19.18	19.05	13.48	14.29	15.71
1 day a week	4.62	3.9	3.23	2.67	5.29	3.27	2.95	3.34	3.53	5.45	5.48	7.94	7.87	7.14	7.14
2 days a week	8.63	8.61	7.6	6.61	10.09	5.95	5.9	6.47	8.82	8.77	9.59	11.11	8.99	7.14	18.57
3 days a week	7.53	8.27	7.99	6.33	7.1	5.36	11.15	7.1	5.88	8.06	8.22	12.7	15.73	14.29	10
4 days a week	6.62	7.51	7.2	6.85	5.8	5.65	7.21	8.98	5.29	7.11	6.85	6.35	7.87	14.29	4.29
5 days a week	7.55	10.59	10.84	10.04	8.37	9.82	8.85	8.98	5.29	9.48	4.11	11.11	10.11	3.57	5.71
6 days a week	3.15	3.98	3.68	3.38	4.31	5.06	5.9	4.38	5.29	3.55	8.22	1.59	2.25	3.57	4.29
every day	53.18	50.61	54.19	59.47	52.11	55.36	52.79	54.91	59.41	51.66	38.36	30.16	33.71	35.71	34.29
paint or draw															
	native					mixed					fully foreign				
age	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
never	9.43	3.32	2.58	5.5	7.81	12.5	5.57	1.88	2.24	7.11	13.7	9.52	3.41	5.06	8.57
1 day a week	7.92	4.58	3.21	4.49	6.97	8.33	4.59	3.55	5.15	7.82	15.07	3.17	5.68	3.8	4.29
2 days a week	17.88	13.78	8.83	11.99	14.84	13.69	11.15	8.56	13.87	13.98	13.7	11.11	12.5	10.13	15.71

Continued on next page

Table 4.B.2 – continued from previous page

3 days a week	16.17	17.73	13.49	11.3	14.3	14.58	16.72	12.73	11.63	13.98	13.7	9.52	9.09	15.19	18.57
4 days a week	10.8	14.65	12.34	10.15	10.1	15.18	15.41	11.27	8.05	12.09	8.22	12.7	9.09	16.46	11.43
5 days a week	8.28	11.56	21.02	13.71	9.45	8.63	12.46	21.92	14.54	11.14	12.33	15.87	26.14	10.13	11.43
6 days a week	1.81	3.35	3.17	2.68	2.63	2.68	4.92	3.13	3.58	2.13	4.11	0	1.14	1.27	1.43
every day	27.71	31.02	35.34	40.19	33.9	24.4	29.18	36.95	40.94	31.75	19.18	38.1	32.95	37.97	28.57
							practise rhymes								
			native					mixed					fully foreign		
age	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
never	12.86	3.79	2.03	3.75	9.1	12.5	2.95	3.35	4.25	10.43	16.22	6.45	4.6	7.59	11.43
1 day a week	2.37	1.7	1.27	2.41	4.62	3.27	2.95	0.84	3.13	4.74	6.76	1.61	0	5.06	7.14
2 days a week	6.86	5.73	4.82	6.57	9.19	5.06	6.23	4.81	8.05	10.9	4.05	9.68	3.45	18.99	12.86
3 days a week	6.95	6.46	5.71	6.64	7.68	7.44	5.57	5.86	8.05	10.43	1.35	4.84	13.79	7.59	7.14
4 days a week	5.76	5.49	6.3	6.34	6.27	5.65	7.54	6.9	7.38	5.92	8.11	3.23	4.6	12.66	10
5 days a week	5.1	6.62	11.78	9.09	6.75	6.25	8.2	12.97	10.74	6.4	5.41	11.29	16.09	1.27	4.29
6 days a week	2.03	2.56	2.3	2.54	2.01	2.68	4.59	3.14	1.34	2.37	1.35	1.61	1.15	2.53	1.43
every day	58.08	67.65	65.81	62.67	54.37	57.14	61.97	62.13	57.05	48.82	56.76	61.29	56.32	44.3	45.71
							learn letters and shapes								
			native					mixed				fully foreign			
age		2	3	4	5	2	3	4	5	2	3	4	5		
never		30.62	12.21	5.93	5.54	33.93	16.39	6.92	6.04	42.47	24.19	9.09	10.13		
1 day a week		4.45	5.06	3.76	3.3	4.17	5.57	2.52	2.24	4.11	6.45	2.27	1.27		
2 days a week		11.58	13.15	9.42	9.35	8.04	10.49	8.6	7.61	5.48	9.68	5.68	2.53		
3 days a week		9.94	13.02	12.02	8.48	10.71	13.44	10.27	8.95	5.48	6.45	12.5	11.39		
4 days a week		6.27	9.75	10.83	8.57	5.95	8.2	12.58	10.29	9.59	8.06	3.41	12.66		
5 days a week		5.78	10.58	16.37	14.29	7.14	9.51	14.26	14.32	6.85	9.68	27.27	17.72		
6 days a week		1.35	1.99	2.59	2.89	1.49	2.3	2.73	4.47	2.74	1.61	6.82	7.59		
every day		30.03	34.24	39.07	47.57	28.57	34.1	42.14	46.09	23.29	33.87	32.95	36.71		
							use computer								
			native					mixed					fully foreign		
age	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
never	84.73	64.18	32.15	22.39	17.51	82.14	67.21	35.56	26.4	20.19	77.03	68.25	29.07	29.11	27.14
1 day a week	4.4	10.54	11.33	12.61	11.9	4.46	9.84	12.97	15.88	12.11	6.76	6.35	10.47	10.13	10
2 days a week	4.25	9.16	15.57	16.65	17.38	4.76	7.21	15.48	13.2	20.19	6.76	4.76	13.95	15.19	10

Continued on next page

Table 4.B.2 – continued from previous page

3 days a week	1.71	5.57	11.3	11.63	11.84	2.38	3.93	10.46	11.41	14.01	1.35	3.17	12.79	8.86	7.14
4 days a week	1.15	2.59	6.61	7.56	9.36	1.49	3.93	6.28	5.37	5.46	0	4.76	5.81	7.59	7.14
5 days a week	0.51	1.75	7.11	6.5	5.87	1.19	1.31	5.23	6.49	5.46	0	0	6.98	3.8	7.14
6 days a week	0.15	0.31	0.58	1.31	1.19	0	0.33	0.84	1.57	1.19	0	0	0	2.53	2.86
every day	3.1	5.89	15.36	21.34	24.93	3.57	6.23	13.18	19.69	21.38	8.11	12.7	20.93	22.78	28.57
							watch TV (weekdays)								
			native					mixed					fully foreign		
age	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
None	1.23	0.6	3.57	2.46	5.56	1.46	1.36	3.88	4.76	8.89	3.39	0	1.15	3.9	2.94
Up to 30 minutes	28.64	11.22	6.73	3.45	2.47	24.45	10.88	7.33	3.85	1.92	32.2	5.17	5.75	6.49	0
30 minutes to 1 hour	25.15	16.94	9.93	11.68	10.31	27.74	19.05	10.78	16.55	12.98	13.56	22.41	11.49	11.69	16.18
1 to 2 hours	30.08	36.36	38.9	41.94	46.29	30.29	35.71	44.4	40.82	48.08	28.81	15.52	36.78	36.36	41.18
2 to 3 hours	11.12	23.61	26.38	28.02	25.24	12.04	22.79	21.77	26.08	21.63	11.86	34.48	21.84	22.08	25
3 to 4 hours	1.99	6.35	8.47	7.72	6.54	2.55	6.46	6.9	5.22	4.09	3.39	10.34	9.2	12.99	13.24
4 to 5 hours	0.86	2.61	3.36	2.77	2.24	1.09	2.72	3.23	2.04	1.2	3.39	8.62	8.05	2.6	1.47
5 or more hours	0.95	2.31	2.66	1.96	1.36	0.36	1.02	1.72	0.68	1.2	3.39	3.45	5.75	3.9	0
							watch TV (weekends)								
age			native					mixed					fully foreign		
None	2	3	4	5	6	2	3	4	5	6	2	3	4	5	6
Up to 30 minutes	7.09	4.12	5.88	1.96	5.41	8.39	4.08	5.18	3.61	7.89	5.08	1.72	8.05	0	4
30 minutes to 1 hour	21.61	7.44	4.82	3.87	0.85	17.15	6.12	5.18	4.82	1.97	25.42	8.62	6.9	11.54	8
1 to 2 hours	21.03	12.43	5.86	7.74	1.4	24.09	14.29	6.26	6.63	0	13.56	12.07	8.05	15.38	8
2 to 3 hours	25.33	26.82	21.33	23.3	14.77	23.72	27.89	27.21	21.08	15.13	30.51	20.69	25.29	15.38	8
3 to 4 hours	16.61	24.04	24.51	28.27	24.49	18.61	22.45	22.46	23.49	28.29	16.95	20.69	16.09	34.62	24
4 to 5 hours	3.32	7.03	8.99	9.84	10.57	3.65	7.48	9.07	13.25	11.84	1.69	10.34	5.75	15.38	8
5 or more hours	3.1	10.79	13.31	13.23	17.68	2.55	11.22	11.23	16.87	13.16	1.69	13.79	11.49	0	24
	1.9	7.33	15.3	11.8	24.84	1.82	6.46	13.39	10.24	21.71	5.08	12.07	18.39	7.69	16

Data: Growing Up in Scotland, ScotCen and the Scottish Government

Table 4.B.3: Children's participation in various events

Child has been to..... since last year. (%)									
	native			mixed			fully foreign		
age	2	4	6	2	4	6	2	4	6
concert,play	25.83	64.97	77.09	26.79	62.84	77.27	18.42	37.08	60
swimming pool	86.02	88.53	91.54	79.17	83.51	86.36	42.11	52.81	64
sport event	18.32	26.61	30.23	18.45	24.22	31.17	9.21	11.24	24
museum, gallery	30.54	47.71	58.68	49.11	56.16	68.83	34.21	43.82	68
zoo, aquarium	75.34	79.18	71.4	76.49	78.29	66.88	55.26	62.92	52
cinema	4.71	55.96	81.79	8.63	55.11	87.01	10.53	35.96	72
religious event	33.61	39.5	53.64	36.01	44.47	57.14	38.16	52.81	44

Data: Growing Up in Scotland, ScotCen and the Scottish Government

Table 4.B.4: Physical activity of children (%)

Child ..... last week(%)						
	native		mixed		fully foreign	
age	3	5	3	5	3	5
rode a bicycle	58.99	63.9	59.34	59.55	52.38	50.63
kicked a ball	92.93	87.17	94.75	86.77	90.48	82.05
danced	61.15	63.24	64.47	63.31	58.73	63.29
ran/jumped	98.35	98.54	98.69	98.43	100	97.47
swam	29	40.98	30.16	44.97	9.52	22.78
played in a play area	39.1	31.54	36.07	29.53	26.98	13.92
played in a park	68.08	65.62	68.52	63.98	44.44	56.96
did another active sport	19.06	17.41	23.61	23.27	17.46	11.39

Data: Growing Up in Scotland, ScotCen and the Scottish Government

Table 4.B.5: Use of disciplining techniques (%)

Used with child									
	native			mixed			fully foreign		
	2	4	6	2	4	6	2	4	6
time out	30.78	65.05	72.59	31.85	64.23	65.58	26.67	47.19	56
rewards	8.24	56.93	70.56	12.2	59.62	70.13	12	50.56	76
ignored bad behaviour	67.49	69.99	62.99	62.5	66.32	61.04	49.33	61.8	60
smacking	16	41.1	46.41	14.29	36.61	39.61	21.33	39.33	44
naughty step	35.8	69.43	69.12	29.76	65.27	62.99	26.67	46.07	72
raised voice	62.05	79.24	85.85	65.18	80.33	81.17	65.33	71.91	76
removing treats	30.56	76.74	86.84	22.32	72.59	84.42	25.33	60.67	84

Data: Growing Up in Scotland, ScotCen and the Scottish Government

Table 4.B.6: Parental attitudes and ambitions for the child (%)

	How far do you see your child's education?					
	native		mixed		fully foreign	
Age	5	6	5	6	5	6
Standard Grades	0.28	2	0	0.48	0	0
Higher Grades	2.62	7.4	1.82	4.07	9.09	2.94
one or two year college course	15.19	59.9	5.45	61.72	18.18	63.24
three or four year degree course	12.15	17.02	9.09	17.46	0	8.82
a Masters, PhD, medical	37.71	11.53	45.45	14.83	36.36	23.53
I do not really mind	32.04	2.15	38.18	1.44	36.36	1.47

	Nobody can teach you how to be a good parent											
	native				mixed				fully foreign			
Age	1	3	4	6	1	3	4	6	1	3	4	6
Strongly agree	21.83	20.71	12.16	12.63	13.35	19.02	10.44	13.64	13.46	6.67	9.09	12
Agree	41.23	39.16	41.6	40.42	37.78	32.68	34.01	35.06	32.69	46.67	47.27	40
Neither agree nor disagree	15.23	15	18.29	17.19	17.13	16.1	20.54	18.18	18.27	22.22	18.18	20
Disagree	20.11	22.86	26.47	28.08	28.21	28.29	33.33	31.17	33.65	17.78	25.45	24
Strongly Disagree	1.59	2.26	1.48	1.68	3.53	3.9	1.68	1.95	1.92	6.67	0	4

	It is more important to go with what the child wants than to stick to a routine											
	native				mixed				fully foreign			
Age	1	3	4	6	1	3	4	6	1	3	4	6
Strongly agree	8.15	4.52	3.1	3.12	7.83	3.9	3.7	1.95	13.59	4.44	3.64	4
Agree	26.16	19.02	18.98	15.66	30.3	20.98	22.22	24.68	41.75	44.44	43.64	32
Neither agree nor disagree	18.77	16.76	23.67	20.56	20.96	21.46	20.2	18.18	16.5	28.89	16.36	44
Disagree	38.3	47.35	46.47	51.14	31.57	40	47.47	49.35	25.24	20	34.55	20
Strongly Disagree	8.62	12.35	7.77	9.51	9.34	13.66	6.4	5.84	2.91	2.22	1.82	0

	It is better for children to have two parents than one											
	native				mixed				fully foreign			
Age	1	3	4	6	1	3	4	6	1	3	4	6
Strongly agree	21.46	19.12	14.11	15.99	33.42	34.15	26.26	24.84	51.43	53.33	43.64	36
Agree	36.58	36.64	41.71	40.07	40.51	37.56	45.45	46.41	36.19	31.11	45.45	52
Neither agree nor disagree	21.32	21.77	22.01	21.55	16.2	15.61	17.51	18.95	10.48	13.33	5.45	12
Disagree	17.83	19.2	19.5	19.96	8.86	10.73	10.1	9.15	1.9	2.22	5.45	0
Strongly Disagree	2.8	3.26	2.67	2.43	1.01	1.95	0.67	0.65	0	0	0	0

Data: Growing Up in Scotland, ScotCen and the Scottish Government

4.B.1 Children's outcomes

Table 4.B.7: Child can be understood by .... (%)

							the respondent					
	native				mixed				fully foreign			
age	2	3	4	5	2	3	4	5	2	3	4	5
mostly	83.99	96.19	97.97	98.34	83.04	95.29	97.91	99.28	81.58	89.81	94.38	100
sometimes	13.25	3.05	1.55	1.37	13.39	4.31	1.67	0.72	15.79	7.41	5.62	0
rarely	2.76	0.76	0.47	0.29	3.57	0.39	0.42	0	2.63	2.78	0	0
			Age 2		Age 3		Age 4		Age 5			
			$\Delta_(NM)$	$\Delta_(NF)$	$\Delta_(NM)$	$\Delta_(NF)$	$\Delta_(NM)$	$\Delta_(NF)$	$\Delta_(NM)$	$\Delta_(NF)$		
mostly			0.95	2.41	0.9	6.38	0.06	3.59	-0.94	-1.66		
sometimes			-0.14	-2.54	-1.26	-4.36	-0.12	-4.07	0.65	1.37		
rarely			-0.81	0.13	0.37	-2.02	0.05	0.47	0.29	0.29		
							family and friends					
	native				mixed				fully foreign			
age	2	3	4	5	2	3	4	5	2	3	4	5
mostly	59.82	83.71	90.36	93.83	55.36	82.35	90.19	96.03	53.95	69.44	78.65	94.12
sometimes	33.69	14.28	8.8	5.65	35.42	15.1	9.39	3.61	38.16	24.07	21.35	5.88
rarely	6.49	2.01	0.84	0.51	9.23	2.55	0.42	0.36	7.89	6.48	0	0
			Age 2		Age 3		Age 4		Age 5			
			$\Delta_(NM)$	$\Delta_(NF)$	$\Delta_(NM)$	$\Delta_(NF)$	$\Delta_(NM)$	$\Delta_(NF)$	$\Delta_(NM)$	$\Delta_(NF)$		
mostly			4.46	5.87	1.36	14.27	0.17	11.71	-2.2	-0.29		
sometimes			-1.73	-4.47	-0.82	-9.79	-0.59	-12.55	2.04	-0.23		
rarely			-2.74	-1.4	-0.54	-4.47	0.42	0.84	0.15	0.51		

Data: Growing Up in Scotland, ScotCen and the Scottish Government
Note: $\Delta_(NM)$ is the difference in outcomes between children from native and mixed families
$\Delta_(NF)$ is the difference in outcomes between children from native and fully foreign families

Table 4.B.8: Children's motor development measures

Baby development measures at the age of 1 (%)					
			smiled		
	native	mixed	fully foreign	$\Delta_(NM)$	$\Delta_(NF)$
often	99.77	99.75	100	0.02	-0.23
once	0.21	0.25	0	-0.04	0.21
not yet	0.02	0	0	0.02	0.02
			sat		
	native	mixed	fully foreign	$\Delta_(NM)$	$\Delta_(NF)$
often	97.75	99.49	98.08	-1.74	-0.33
once	1.06	0.25	0	0.81	1.06
not yet	1.19	0.25	1.92	0.94	-0.73
			stood up		
	native	mixed	fully foreign	$\Delta_(NM)$	$\Delta_(NF)$
often	80.03	81.57	86.54	-1.54	-6.51
once	8.05	6.31	3.85	1.74	4.2
not yet	11.93	12.12	9.62	-0.19	2.31
			put hands together		
	native	mixed	fully foreign	$\Delta_(NM)$	$\Delta_(NF)$
often	91.19	90.38	90.38	0.81	0.81
once	5.52	5.32	3.85	0.2	1.67
not yet	3.29	4.3	5.77	-1.01	-2.48
			grabbed objects		
	native	mixed	fully foreign	$\Delta_(NM)$	$\Delta_(NF)$
often	99.58	99.24	99.04	0.34	0.54
once	0.28	0.76	0.96	-0.48	-0.68
not yet	0.15	0	0	0.15	0.15
			picked up objects		
	native	mixed	fully foreign	$\Delta_(NM)$	$\Delta_(NF)$
often	93.65	92.11	91.26	1.54	2.39
once	3.75	5.09	5.83	-1.34	-2.08
not yet	2.6	2.8	2.91	-0.2	-0.31
			passed a toy		
	native	mixed	fully foreign	$\Delta_(NM)$	$\Delta_(NF)$
often	96.43	94.7	97.12	1.73	-0.69
once	2.49	3.79	2.88	-1.3	-0.39
not yet	1.08	1.52	0	-0.44	1.08
			walked a few steps		
	native	mixed	fully foreign	$\Delta_(NM)$	$\Delta_(NF)$
often	13.37	10.35	18.27	3.02	-4.9
once	9.89	9.09	9.62	0.8	0.27
not yet	76.74	80.56	72.12	-3.82	4.62
			reached out		
	native	mixed	fully foreign	$\Delta_(NM)$	$\Delta_(NF)$
often	77.32	71.21	74.04	6.11	3.28
once	16.73	19.44	17.31	-2.71	-0.58
not yet	5.94	9.34	8.65	-3.4	-2.71
			waved		

Continued on next page

Table 4.B.8 – continued from previous page

	native	mixed	fully foreign	$\Delta_(NM)$	$\Delta_(NF)$
often	56.62	45.45	40.78	11.17	15.84
once	26.16	29.55	26.21	-3.39	-0.05
not yet	17.22	25	33.01	-7.78	-15.79
			extended arms		
	native	mixed	fully foreign	$\Delta_(NM)$	$\Delta_(NF)$
often	85.9	82.58	93.27	3.32	-7.37
once	10.68	13.38	3.85	-2.7	6.83
not yet	3.42	4.04	2.88	-0.62	0.54
			nodded		
	native	mixed	fully foreign	$\Delta_(NM)$	$\Delta_(NF)$
often	9.73	7.36	16.35	2.37	-6.62
once	10.09	9.14	8.65	0.95	1.44
not yet	80.17	83.5	75	-3.33	5.17
			Toddler development measures at the age of 3		
The child can.... (%)					
	native	mixed	fully foreign	$\Delta_(NM)$	$\Delta_(NF)$
walk on level	99.31	99.51	100	-0.2	-0.69
balance on one foot	88.62	87.81	80.58	0.81	8.04
hop on one foot	77.1	77.82	78.43	-0.72	-1.33
throw a ball	99.73	100	99.07	-0.27	0.66
grasp small objects	99.69	99.61	98.15	0.08	1.54
undo big buttons	79.9	78.09	71.03	1.81	8.87
draw a circle	80.65	80.4	84.26	0.25	-3.61
hold a pencil	99.59	99.61	98.15	-0.02	1.44
copy a square	40.95	41.94	35.29	-0.99	5.66
drink from a cup	98.78	98.82	98.15	-0.04	0.63
brush his teeth	96.3	94.3	85.98	2	10.32
put a t-shirt on	74.43	70.89	71.3	3.54	3.13
get dressed	41.68	37.28	43.52	4.4	-1.84

Data: Growing Up in Scotland, ScotCen and the Scottish Government

Note: $\Delta_(NM)$ is the difference in outcomes between children from native and mixed families $\Delta_(NF)$ is the difference in outcomes between children from native and fully foreign families

4.C Further regression results

Table 4.C.1: Regression results for other speech indicators not influenced by language or family composition

Dependent variable: Can the child be understood by respondent?							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OP	OP	OP	OP	OP	OP	OP
mixed family	-.003	-.035	-.073	-.113	-.122	-.119	.263
	(.018)	(.060)	(.093)	(.187)	(.187)	(.186)	(.500)
fully foreign family	-.001	.048	-.007	-.326	-.328	-.325	-.928
	(.050)	(.151)	(.187)	(.378)	(.372)	(.376)	(1.150)
bilingual	.018	.151**	.199**	.136	.145	.162	.281
	(.023)	(.074)	(.092)	(.198)	(.197)	(.196)	(.340)
mixed*bilingual							-.296
							(.416)
foreign*bilingual							-.195
							(.553)
Controls:							
household controls	no	yes	yes	yes	yes	yes	yes
activities at home	no	no	yes	yes	yes	yes	yes
outings	no	no	no	yes	yes	yes	yes
physical activity	no	no	no	no	yes	yes	yes
discipline	no	no	no	no	no	yes	yes
N	7447	6790	6616	4061	4058	4058	4058
pseudo R-squared	.094	.120	.135	.094	.102	.103	.100

Data: Growing Up in Scotland, ScotCen and the Scottish Government
OP stands for the ordered probit
Note: All regressions include gender, sweep and the cohort dummy as controls. Household controls include: number of siblings, parental education, NS-SEC household classification, equivalised income, geographical area and whether one is a single parent. Activities at home are frequency of painting/drawing, practising rhymes, practising letters, use of computer and TV. Outings include visits to many attractions, including museums, library, etc. Physical activity includes swimming, playing in the park, running, etc. And disciplining techniques include time out and naughty step.
The independent variable mixed family is a dummy equal to 1 if one of child's parents was born outside of the UK and zero otherwise. The variable fully foreign family is a dummy equal to 1 if both child's parents were born outside of the UK and zero otherwise. Bilingual is a dummy variable equal to 1 if child speaks English and another language or just another language at home.
Errors are clustered at individual level.
Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 4.C.2: Regression results for other speech indicators not influenced by language or family composition

Dependent variable - Can child be understood by the family and friends							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OP	OP	OP	OP	OP	OP	OP
mixed family	.015	.057	.076	.0009	-.015	-.013	-.236
	(.021)	(.049)	(.065)	(.104)	(.104)	(.104)	(.334)
fully foreign family	.082	.164	.216	.198	.213	.223	.815
	(.056)	(.125)	(.147)	(.247)	(.245)	(.246)	(.679)
bilingual	.002	.047	.019	-.022	-.022	-.021	-.051
	(.025)	(.059)	(.074)	(.120)	(.122)	(.122)	(.221)
mixed*bilingual							.164
							(.271)
foreign*bilingual							-.295
							(.381)
Controls:							
household controls	no	yes	yes	yes	yes	yes	yes
activities at home	no	no	yes	yes	yes	yes	yes
outings	no	no	no	yes	yes	yes	yes
physical activity	no	no	no	yes	yes	yes	yes
discipline	no	no	no	no	yes	yes	yes
N	7447	6790	6616	4058	4058	4058	4058
R-squared	.087	.100	.114	.150	.113	.118	.096

Data: Growing Up in Scotland, ScotCen and the Scottish Government
OP stands for ordered probit
Note: All regressions include gender, sweep and the cohort dummy as controls. Household controls include: number of siblings, parental education, NS-SEC household classification, equivalised income, geographical area and whether one is a single parent. Activities at home are frequency of painting/drawing, practising rhymes, practising letters, use of computer and TV. Outings include visits to many attractions, including museums, library, etc. Physical activity includes swimming, playing in the park, running, etc. And disciplining techniques include time out and naughty step.
The independent variable mixed family is a dummy equal to 1 if one of child's parents was born outside of the UK and zero otherwise. The variable fully foreign family is a dummy equal to 1 if both child's parents were born outside of the UK and zero otherwise. Bilingual is a dummy variable equal to 1 if child speaks English and another language or just another language at home.
Errors are clustered at individual level.
Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 4.C.3: Regression results for other measures not influenced by language or family composition

	CSBS score - clusters						
	Emotion and eye gaze	Communication	Gestures	Sounds	Words	Understanding of words	Use of objects
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OP	OP	OP	OP	OP	OP	OP
mixed	.034	-.048	-.150*	-.168**	-.059	-.136	-.285***
	(.082)	(.083)	(.079)	(.078)	(.080)	(.093)	(.080)
fully foreign	.271	-.094	-.279	-.029	-.042	-.134	-.489***
	(.226)	(.189)	(.189)	(.175)	(.199)	(.194)	(.188)
bilingual	-.151	.170	-.070	.003	-.061	-.182*	-.006
	(.108)	(.105)	(.093)	(.086)	(.089)	(.099)	(.089)

	CSBS - individual elements of the affected clusters									
	Gestures cluster									
	picks objects		*shows objects*		*waves*		*points to objects*		*nods*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS	OP	OLS	OP	OLS	OP	OLS	OP	OLS	OP
mixed	-.007	-.074	-.012	-.044	-.035	-.125	-.037*	-.297**	-.017	-.048
	(.017)	(.122)	(.027)	(.083)	(.026)	(.108)	(.022)	(.119)	(.032)	(.088)
fully foreign	-.063	-.274	-.095	-.295	-.091	-.303	.014	.293	-.099	-.222
	(.066)	(.228)	(.078)	(.191)	(.083)	(.239)	(.067)	(.359)	(.101)	(.220)
bilingual	-.067**	-.224*	-.071	-.141	-.026	-.020	-.076**	-.267**	-.054	-.068
	(.031)	(.121)	(.039)	(.109)	(.035)	(.119)	(.033)	(.125)	(.045)	(.111)
all controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
	Understanding of words						**Use of objects**			
	responds to name		*no of words known*		*plays with variety of objects*		*appropriate use of objects*		*stacking blocks*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS	OP	OLS	OP	OLS	OP	OLS	OP	OLS	OP
mixed	-.0008	-.026	-.009	-.120	-.046*	-.267***	-.066	-.168*	-.053*	-.192**
	(.017)	(.131)	(.016)	(.163)	(.029)	(.101)	(.042)	(.086)	(.028)	(.093)
fully foreign	-.030	.032	-.084	-.524**	-.147	-.490**	-.373	-.578*	-.062	-.199
	(.068)	(.257)	(.070)	(.266)	(.111)	(.206)	(.244)	(.301)	(.081)	(.200)
bilingual	-.070**	-.268**	-.051*	-.114	-.112**	-.115	-.016	.059	-.089**	-.189*
	(.031)	(.128)	(.029)	(.148)	(.051)	(.108)	(.081)	(.130)	(.031)	(.106)
all controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

Data: Growing Up in Scotland, ScotCen and the Scottish Government

OP stands for ordered probit

Note: All regressions include gender, sweep and the cohort dummy as controls. Household controls include: number of siblings, parental education, NS-SEC household classification, equivalised income, geographical area and whether one is a single parent. Activities at home are frequency of painting/drawing, practising rhymes, practising letters, use of computer and TV. Outings include visits to many attractions, including museums, library, etc. Physical activity includes swimming, playing in the park, running, etc. And disciplining techniques include time out and naughty step.

The independent variable mixed family is a dummy equal to 1 if one of child's parents was born outside of the UK and zero otherwise. The variable fully foreign family is a dummy equal to 1 if both child's parents were born outside of the UK and zero otherwise. Language is a dummy variable equal to 1 if child speaks English and another language or just another language at home.

Errors are clustered at individual level.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

No other components of the overall SDQ score are influenced by language spoken at home or family composition.

Table 4.C.4: Regression results for other measures not influenced by language or family composition

	SDQ - individual elements of the affected clusters									
	Peer problems score									
	solitary		has a good friend		liked by other children		picked on		gets on better with adults	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS	probit	OLS	probit	OLS	probit	OLS	probit	OLS	probit
mixed	.002	.007	.015	.227	.003	.197	-.004	-.008	-.039	-.135
	(.026)	(.076)	(.010)	(.155)	(.004)	(.230)	(.016)	(.102)	(.024)	(.087)
fully foreign	-.089	-.244	.020	.211	.008**	.008**	.044	.166	.103	.243
	(.068)	(.199)	(.028)	(.311)	(.004)	(.004)	(.050)	(.213)	(.068)	(.190)
bilingual	.062	.174*	-.017	-.202	.003	.199	.030	.153	.091***	.280***
	(.032)	(.090)	(.014)	(.139)	(.003)	(.279)	(.021)	(.106)	(.031)	(.092)
all controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
	Other questions									
	Considerate of people's feelings		shares with other children		often has tantrums		fights with other children		has many fears	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS	probit	OLS	probit	OLS	probit	OLS	probit	OLS	probit
mixed	-.017	-.198*	-.019*	-.240**	-.026	-.071	.013	.087	.041	.127*
	(.012)	(.116)	(.012)	(.122)	(.026)	(.073)	(.016)	(.098)	(.026)	(.077)
fully foreign	-.038	-.272	-.003	-.099	-.086	-.249	-.006	-.014	.019	.063
	(.038)	(.233)	(.032)	(.287)	(.067)	(.181)	(.046)	(.250)	(.065)	(.191)
bilingual	-.025	-.177	-.019	-.174	-.029	-.089	.022	.117	-.003	-.011
	(.017)	(.120)	(.015)	(.130)	(.032)	(.087)	(.021)	(.114)	(.029)	(.090)
all controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

Data: Growing Up in Scotland, ScotCen and the Scottish Government

Note: All regressions include gender, sweep and the cohort dummy as controls. Household controls include: number of siblings, parental education, NS-SEC household classification, equivalised income, geographical area and whether one is a single parent. Activities at home are frequency of painting/drawing, practising rhymes, practising letters, use of computer and TV. Outings include visits to many attractions, including museums, library, etc. Physical activity includes swimming, playing in the park, running, etc. And disciplining techniques include time out and naughty step.

The independent variable mixed family is a dummy equal to 1 if one of child's parents was born outside of the UK and zero otherwise. The variable fully foreign family is a dummy equal to 1 if both child's parents were born outside of the UK and zero otherwise. Bilingual is a dummy variable equal to 1 if child speaks English and another language or just another language at home.

Errors are clustered at individual level.

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

No other components of the overall SDQ score are influenced by language spoken at home or family composition.

Table 4.C.5: Regression results for motor and physical development indicators

	Baby measures (birth cohort only, age 1)											
	child smiles	sits	stands up	puts hands together	grabs objects	picks up objects	passes toys	walks	reaches out	waves	extends arms	nods
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	OP	OP	OP	OP	OP	OP	OP	OP	OP	OP	OP	OP
mixed family	4.141***	.725*	-.014	.074	-.067	-.050	-.165	-.112	-.109	-.189***	-.117	-.162*
	(.732)	(.431)	(.090)	(.108)	(.275)	(.124)	(.136)	(.089)	(.083)	(.071)	(.088)	(.093)
fully foreign	4.631***	.218	.091	.376*	-.398	.179	.028	.123	-.055	-.212	.315	-.041
	(.753)	(.529)	(.205)	(.222)	(.585)	(.251)	(.330)	(.186)	(.193)	(.183)	(.237)	(.192)
bilingual	-.780**	-.223	.015	-.294***	-.014	-.336***	-.014	.004	-.091	-.167*	-.036	.178*
	(.363)	(.302)	(.107)	(.110)	(.342)	(.123)	(.168)	(.105)	(.100)	(.091)	(.116)	(.100)
all controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
N	4671	4671	4671	4669	4670	4649	4663	4671	4670	4669	4670	4666
pseudo-R-squared	.090	.058	.008	.013	.021	.016	.010	.008	.010	.025	.008	.014

Data: Growing Up in Scotland, ScotCen and the Scottish Government

OP stands for ordered probit

Note: The answers to baby measures are coded in the following way: 0 - no, 1-sometimes, 2-often. Hence, a positive coefficient indicates better performance

Controls in the regressions include: child's gender, number of siblings, mother's education, household equivalised income, area of deprivation and whether family a full family.

Robust standard errors in parentheses.

Statistical significance: *** p< 0.01, ** p < 0.05, * p < 0.1

Table 4.C.6: Regression results for motor and physical development indicators

	Toddler measures (child cohort, age 3)												
	walks up steps	balances on foot	hops on foot	throws a ball	grasps objects	undoes buttons	draws a circle	scribbles	copies a square	drinks from a cup	brushes teeth	puts a T-shirt on	dresses himself
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	probit	probit	probit	probit	probit	probit	probit	probit	probit	probit	probit	probit	probit
mixed family	-.228	-.085	-.044	.072	.050	-.009	-.045	.043	.164	.031	.006	-.024	.003
	(.373)	(.150)	(.089)	(.079)	(.070)	(.377)	(.080)	(.081)	(.347)	(.074)	(.212)	(.119)	(.076)
fully foreign	.320	.341	-.421**	-.084	-.125	-.985	-.242	.111	-.215	-.318*	-.309	-.256	-.019
	(.464)	(.494)	(.200)	(.182)	(.761)	(.637)	(.183)	(.193)	(.351)	(.171)	(.335)	(.242)	(.178)
bilingual	-.120	-.135	.012	-.038	-.409	.190	.024	-.062	-.494**	.041	-.006	-.403***	-.121
	(.166)	(.177)	(.105)	(.090)	(.402)	(.488)	(.095)	(.094)	(.216)	(.085)	(.209)	(.120)	(.088)
all controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
N	6443	6170	6216	6444	6444	6246	6371	6444	5967	6443	6438	6411	6430
pseudo-R-squared	.047	.003	.005	.009	.029	.051	.031	.030	.033	.009	.011	.015	.051

Data: Growing Up in Scotland, ScotCen and the Scottish Government
Note: Measures of toddler development were coded as dummy variables equal to 1 if a child can do a given thing.
Controls in the regressions include: child's gender, number of siblings, mother's education, household equivalised income, area of deprivation and whether family a full family.
Robust standard errors in parentheses.
Statistical significance: *** p< 0.01, ** p < 0.05, * p < 0.1