Best of both worlds? Early cognitive and non-cognitive development of bilingual children.

Joanna Clifton-Sprigg*
School of Economics, the University of Edinburgh
31 Buccleuch Place, Edinburgh EH8 9JT
UK
email: j.m.clifton.sprigg@gmail.com, phone: +44 7772908836

January 10, 2015

Abstract

This paper looks at the effect of speaking two rather than one language at home on early childhood cognitive and non-cognitive performance. Using the Scottish Government data for under-six-year-olds in Scotland, I find that children have comparable cognitive and non-cognitive skills, irrespective of the number of languages used at home. Bilingual children score on average comparably to monolingual children also in the English Vocabulary Naming Exercise. However, bilingual families are a diverse group; this heterogeneity plays an additional role in children's outcomes. Specifically, bilingual children who have one foreign and one native parent perform worse in this task than native monolingual children at the age of 3 but catch up with them by the age of 5. Bilingual children of two foreign-born parents, on the other hand, may perform substantially worse than native monolingual children in the English Vocabulary Naming Exercise, although results are statistically insignificant.

JEL-Classification: I25, J13, J15, F22
Keywords: Childhood, Bilingualism, Skills, Immigration

*Advice given by Prof. Maia Güell, Prof. Mike Elsby and Dr Steven Dieterle has been a great help in the execution of this project. I would also like to thank Lesley Kelly and Paul Bradshaw at the Growing Up in Scotland project for data access as well as valuable advice on how to use it. My research work has also been supported by the ESRC scholarship (no: ES/J500136/1).

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; 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. 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 potential 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.

[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.

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,

[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.

randomly selected sample of children in Scotland: their family situation, socio-economic 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 rather descriptive in nature and relies on raw comparisons and simple regressions. The causal inference one can make from this analysis is threatened by unobserved heterogeneity and selection bias. 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.

Despite the shortcomings, this paper confirms many findings from the linguistic literature to date and further contributes to it on various fronts. 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 2. I discuss data and provide unconditional comparisons in Section 3. Section 4 contains regression analysis and its results. In Section 5 I discuss limitations of my approach and conclude.

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,[3] may be the main 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.

[3]Simultaneous bilingualism is a form of bilingualism that takes place when a child becomes bilingual by learning two languages simultaneously from birth.

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, 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.[4] 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. In this paper I demonstrate that bilingualism and family composition are strongly interlinked and key for a child's performance.

3 Data and descriptives

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.[5] 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 (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

[4] Note that the list proposed here is exemplary and by no means exhaustive.

[5] Detailed information about the project can be found on the website: http://growingupinscotland.org.uk

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 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.[6] 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.

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.

[6]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

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 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.[7] 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 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.

[7] There were only 70 children with both foreign parents in the combined sample in the final wave.

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 A.5.

Table 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

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 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 2.

Table 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.

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.

3.4.1 Household composition and socio-economic situation

I start by comparing monolingual and bilingual families (see Table 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.

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.

Table 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

3.4.2 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 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 reflec-

tion 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.

3.4.3 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).

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 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.[8] This observation holds for the total scores, as well as their composites (see Tables 4 and 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

[8]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.

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 27 in Appendix B.1.

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: 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 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 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 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 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.

4.1 Empirical specification

4.1.1 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:

$$Y_{it} = \alpha + \beta_1 bilingual_i + \beta_2 mixed_i + \beta_3 foreign_i + \beta_4 female_i + \beta_5 age_{it} + \beta_6 X_{it} + \gamma_t + \epsilon_{it} \quad (1)$$

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:

$$\begin{aligned} Y_{it} = \alpha + \beta_1 bilingual_i + \beta_2 mixed_i + \beta_3 foreign_i + \theta_1 mixed_i \times bilingual_i \\ + \theta_2 foreign_i \times bilingual_i + \beta_4 female_i + \beta_5 age_{it} + \beta_6 X_{it} + \gamma_t + \epsilon_{it} \end{aligned} \tag{2}$$

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.

4.1.2 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.

$$\begin{aligned} Y_{it} = \alpha &+ \beta_1 bilingual_i + \beta_2 mixed_i + \beta_3 foreign_i + \theta_1 mixed_i \times bilingual_i \\ &+ \theta_2 foreign_i \times bilingual_i + \beta_4 female_i + \lambda_1 bilingual_i \times female_i \\ &+ \lambda_2 mixed_i \times female_i + \lambda_3 foreign_i \times female_i + \beta_5 age_{it} + \beta_6 X_{it} + \gamma_t + \epsilon_{it} \end{aligned} \tag{3}$$

$$\begin{aligned} Y_{it} = \alpha &+ \beta_1 bilingual_i + \beta_2 mixed_i + \beta_3 foreign_i + \theta_1 mixed_i \times bilingual_i \\ &+ \theta_2 foreign_i \times bilingual_i + \beta_4 age_{it} + \lambda_1 bilingual_i \times age_{it} + \lambda_2 mixed_i \\ &\times age_{it} + \lambda_3 foreign_i \times age_{it} + \beta_5 female_i + \beta_6 X_{it} + \gamma_t + \epsilon_{it} \end{aligned} \tag{4}$$

4.2 Results

4.2.1 Cognitive outcomes

In Tables 8 and 9 I present results of the baseline regressions for BAS scores. The coefficients in columns (1) to (6) of Table 8 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 8, 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 9) 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 10. 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.[9] 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 C.

[9]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 8: Regression outcomes for BAS Vocabulary Naming Score

Dependent variable: BAS Vocabulary Naming Score							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	OLS	OLS	OLS	OLS	OLS	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 9: 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 10: 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

4.2.2 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 11 - 14).

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 13 *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 12 and 14.

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 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 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 14 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 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.[10] 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 14 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

[10] According to Wetherby and Prizant (2001) criterion levels for concern are set at more than 1.25 standard deviation below the mean.

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.

Table 11: 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 12: 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 13: Regression results for CSBS total score

Dependent variable: CSBS total score						
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	OLS	OLS	OLS	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 14: 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

4.2.3 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 15.

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 15). 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,[11] but found no significant effects.

[11]Note that I cannot investigate whether CSBS score differ by age as I only have a cross-sectional measure at hand.

Table 15: 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

4.2.4 Foreign parents can help their children catch up

In Section 4.2.3 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 16.

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 16: 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

5 Discussion and limitations

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.

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.

Various forms of sample selection pose another 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 bias 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.

If main attributes related to these two kinds of selection can be captured by the observable characteristics I have information about, the selection is explicitly controlled for in the regressions. Unfortunately, I cannot control for unobserved factors related to the choices to emigrate and to marry a foreigner, such as motivation or drive. Earlier I argued that selection into marriage is possibly not very large, as the profile of respondents in *mixed* families is similar to that of a migrant to Scotland in general.

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 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 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.

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.

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.

References

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.

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.

A Further elaboration on data and variables

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.

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 17: 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

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,[12] 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.

[12] 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.

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.

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.

A.5 Representativeness of families with at least one foreign parent

The portrait of a *mixed and fully foreign family* in Section 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.[13] Therefore, any comparisons are very rough.

[13] Only 28 children interviewed in wave 1 were born outside of the UK and only 108 were born in other countries in the UK.

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[14] - 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 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,[15] 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.

[14] 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.

[15] 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

A.6 Variables of interest

Table 18: 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	
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	

Continued on next page

Table 18 – continued from previous page

Variable	Description
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
BAS Vocabulary Naming	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.
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.
Can child be understood by respondent?	It is measured at ages 2,3,4 and 5.
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, except the pro-social score, and the higher score indicates greater evidence of difficulties.

Table 19: 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?

Continued on next page

Table 19 – continued from previous page

	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
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

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.

B Data summary and statistics

Table 20: 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 21: 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
once every 3 or 4 months	2.1	4.41	3.02	5.37	0	0

Continued on next page

Table 21 – continued from previous page

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
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

Continued on next page

Table 21 – continued from previous page

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
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

Continued on next page

Table 21 – continued from previous page

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 22: 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 23: 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 24: 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 25: 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

B.1 Children's outcomes

Table 26: 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 27: 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

Continued on next page

Table 27 – continued from previous page

			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		
	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

C Further regression results

Table 28: 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 29: 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 30: 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 31: 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 32: 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 33: 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