OECD Education Working Papers No. 75

Immigrant Children's Age at Arrival and Assessment Results

Anthony Heath, Elina Kilpi-Jakonen

OECD

DIRECTORATE FOR EDUCATION

Immigrant children's age at arrival and assessment results

OECD Education Working Paper number 75

Anthony Heath of University of Oxford and Elina Kilpi-Jakonen of Otto-Friedrich-University, Bamberg.

JT03320879

OECD EDUCATION WORKING PAPERS SERIES

This series is designed to make available to a wider readership selected studies drawing on the work of the OECD Directorate for Education. Authorship is usually collective, but principal writers are named. The papers are generally available only in their original language (English or French) with a short summary available in the other.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

The opinions expressed in these papers are the sole responsibility of the author(s) and do not necessarily reflect those of the OECD or of the governments of its member countries.

Applications for permission to reproduce or translate all, or part of, this material should be sent to rights@oecd.org.

Comment on the series is welcome, and should be sent to edu.contact@oecd.org.

----
www.oecd.org/edu/workingpapers

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

ABSTRACT

While a number of single-country studies have been done to explore whether or not there is a "critical age" at which the arrival in a new country becomes a steep disadvantage to the immigrant student, this study aims to determine whether the steepness of the age-at-arrival/test score profile varies across origin or destination countries. As expected, the later the arrival, the greater the penalty. However results vary according to several factors, including language differences and whether the country of origin had higher or lower educational standards. Evidence shows the importance of helping young migrants with language difficulties, as well as with the subsequent adverse effects of these difficulties.

RÉSUMÉ

Tandis qu'un certain nombre d'études nationales ont été réalisées afin de déterminer s'il existe ou non un « âge critique » auquel l'arrivée dans un nouveau pays constitue un désavantage important pour les élèves immigrés, la présente étude tente d'analyser si le profil âge d'arrivée/résultats à l'évaluation varie en fonction du pays d'origine et du pays d'accueil. Comme escompté, plus l'arrivée est tardive, plus le désavantage est important. Les résultats varient cependant en fonction de plusieurs facteurs, notamment des différences linguistiques et du niveau (supérieur ou inférieur) des normes éducatives du pays d'origine. Les données recueillies montrent l'importance d'aider les jeunes migrants à faire face aux difficultés linguistiques auxquelles ils sont confrontés ainsi qu'aux effets négatifs qui peuvent en résulter.

TABLE OF CONTENTS

Tables

IMMIGRANT CHILDREN'S AGE AT ARRIVAL AND ASSESSMENT RESULTS

1.   The aim of this paper is to examine the effect of age at arrival on the performance of immigrant students in the PISA tests of literacy. We aim to investigate whether the effects of age-at-arrival vary for students coming from different origin countries and also for those arriving at different destination countries (the countries of current residence where the tests were held). Our general expectation following earlier research (*e.g.* OECD 2010) is that the later children arrive in a destination country, the further they will lag behind children from the majority group in that country. This might be due to language difficulties or to more general difficulties of adaptation to a different culture and school system, or to cross-national differences in educational standards at different ages.

2.   One important issue that has been debated in the previous literature is whether there is some 'critical age' for arrival after which performance falls off more sharply. Especially for young people who have to learn a new language in the country of destination (which will not of course be the case for all migrants), arriving after the prime ages for language learning may pose important obstacles additional to the other handicaps arising from adjusting to a new country and a new educational system.

3.   In this paper we focus particularly on whether the effect of age-at-arrival varies according to the country from which the young person comes and according to the country of destination in which they are now resident when taking the PISA test. Our primary focus, we should emphasize, is not on the question of whether migrants from particular countries do better in absolute terms than those coming from some other countries (for a detailed discussion of which see OECD 2010), but whether the steepness of the age-at-arrival/test score profile varies across origin or destination countries. That is to say, we are not so much interested in levels of performance but in the rate of change of performance with respect to age at arrival. Putting it more colloquially, our question is whether young people coming from or arriving in particular countries are especially vulnerable to what might be termed a 'late-arrival penalty'. These penalties may in turn have long-lasting implications for the young people's opportunities for making further progress in education, for their integration into the labour market, and for their broader life chances in the country of destination. Substantial penalties may leave long-term scars with important implications for both the migrants affected and for the wider society.

4.   Research on what might be termed the 'absolute' levels of test performance of students suggests that young people from certain origins (for example China) tend to score particularly well whichever country they migrate to. However, in the case of the 'late-arrival penalty', we suspect that it is the **combination** of particular origins with particular destinations that is likely to be of paramount importance. There are two main theoretical arguments that suggest this may be the case.

5.   First of all, it is likely that children arriving from, say, a developed western nation will already have achieved quite highly at school **before** migrating while those coming from a less-developed country with less-developed educational institutions and lower general levels of performance might lag further behind. For example, we know from PISA that test scores in Turkey tend to be below the OECD average, and so children arriving from Turkey in, say, Germany are likely to start at a considerable disadvantage. In contrast children migrating from the Netherlands to Germany are likely to be achieving already at a similar (or even at a higher) level to their peers in Germany. This suggests that the effect of age at arrival will be relatively flat in the case of immigrants from the Netherlands to Germany but could be steeper in the case of young people arriving from Turkey since the later they arrive in Germany, the more they are likely to have fallen behind.

6.   Second, it is likely that, even if children have already achieved quite highly in their origin country, they will suffer a penalty on arrival if they do not speak the destination-country language. For example, children migrating from, say, the UK to continental Europe will typically have to learn a new language, whereas those migrating from the UK to Australia or New Zealand will not. We therefore expect to find that the effect of age at arrival will be more marked where children have to learn a new language.

7.   To be sure, we do not exclude the possibility that there could also be some more general processes at work that affect all young people who migrate to a particular destination. These might for example relate to the kinds of help that countries or educational systems give to the integration of newcomers, but our general expectation is that the major processes are those that involve particular origin and destination combinations.

8.   While this kind of research cannot on its own tell us what policies to adopt, it can usefully indicate the sorts of students and countries where there are especially large 'late-arrival penalties', and where some kind of policy response may be warranted. Our findings may also indicate whether there are any countries which have been particularly successful at avoiding these penalties and whose institutional arrangements may be worthy of more in-depth study in order to assist policy transfer.

Previous research

9.   A small number of single-country studies have explored the effects of age at arrival, while the 2010 PISA report compares the effects of early and late arrival on immigrant students in different countries.

10.   A major focus of the single-country studies has been the question of whether there is a 'critical' age at arrival after which there is a strong negative impact on performance. Using Swedish register data, Böhlmark (2008) found that the critical age at arrival was about nine, with the slopes of the age-at-immigration/performance profiles being similar for boys and girls and for children from different family backgrounds, but varying widely by region of origin. For Asian children the profile was substantially steeper than for Western children, Böhlmark's interpretation being that the steep profile among the Asian children probably reflected "large differences in Sweden-specific skills between children of relatively high and those of relatively low age at migration" whereas the flatter profile for Western children probably reflected the fact that "there is not much to catch up on, *i.e.*, that the human capital they have acquired in the source country does not differ much from that acquired in Sweden." (Böhlmark 2008, p 1382). The finding that immigration at age 9 or later has a negative impact on GPA he attributes to the fact that these students have passed their prime age for language-learning and also because their acquisition of other subject skills is less efficient while the students are still struggling to learn the new language (Böhlmark 2009).

11.   In an American study Myers and his colleagues also investigate whether particular age(s) at arrival is critical. Using the 2000 census and focussing on experiences of Mexican immigrants, their results indicated that the effect of early arrival was much greater for English proficiency than for other socioeconomic outcomes. They found "little evidence at any age of a sharp discontinuity demarcating a 1.5 generation from older immigrants and, in fact, a series of classifications or a continuous measurement of age at arrival may be preferred in some cases" (Myers *et al* 2009, p205). More generally Myers and his colleagues show that linear and curvilinear models perform better than categorical ones, with curvilinear models being slightly superior to linear ones.

12.   In another study conducted in Israel Cahan and his colleagues, like Böhlmark in Sweden, found no difference between boys and girls in the effects of age-at-arrival on educational outcomes, but they found clear differences between ethnic background groups. However, in their case the decrease in

attainment with later arrival was greater for the Western groups (coming from Europe and North America) than among the Eastern group (coming from Asia and North Africa). (Cahan *et al* 2001, p 591). They also tested the 'vulnerable age' hypothesis which had been advanced by earlier scholars – that is to say a U-shaped relationship between age at arrival and educational performance with higher performance being found among early and late arrivals and the lowest performance among those arriving during the intermediate 'vulnerable' years. However, they found absolutely no support for this hypothesis.

13.   In a Dutch study, van Ours and Veenman compared age-at-arrival effects for Turks and Moroccans ('Mediterranean' groups) with those for Surinamese and Antilleans ('Caribbean' groups). Their main finding was that migration at an older age appears to be more of a disadvantage for the educational achievements (measured by level of education achieved not by test scores) of the Turks and Moroccans than for Surinamese and Antilleans. "… our parameter estimates imply the following. Caribbean females and Mediterranean males that came to the Netherlands in their teens will have on average 1 to 1.5 years less of schooling than those that came at an age between 0 and 3. Compared to those that came at a very young age Mediterranean females that entered between age 4 and 11 have about 1.5 years less of schooling while those that entered in their teens have about 2.5 years less of schooling."

14.   Van Ours and Veenman's interpretation of these differences in the effects of age at arrival on the two groups focussed on the congruence between the educational systems of the different sending countries and those of the Netherlands. "Since there are no indications that these differences are related to the receiving country (*e.g.* the characteristics of the [Dutch] educational system), the explanation must be searched for in the sending countries or in the groups' culture. In this respect it seems important that the educational system in Surinam and the Antilles, being (former) Dutch colonies, has a lot in common with the Dutch educational system. Since this is not the case with the Turkish and Moroccan educational system, this might explain the greater difficulties for the Turks and Moroccans. Migration to the Netherlands is for them a larger step than for the Surinamese and the Antilleans" (Van Ours and Veenman, 2006, pp. 314-6).

15.   They also checked whether their results were sensitive to the inclusion of statistical controls for parental education or interacted with them. Their conclusion was that the results barely changed after concluding parental controls or splitting the sample between higher and lower-educated parents.

16.   Turning next to the cross-national comparisons based on the PISA data, the 2010 OECD report confirms this picture of a late-arrival penalty. Table II.4.8 of the report shows that first-generation students who arrived in the host country at a younger age outperform those who arrived when they were older with a difference of 42 points (roughly equivalent to one school year or grade) between those who arrived when they were 5-years-old and those who arrived after they were 12-years-old. "This suggests that where the education system of the host country had a longer opportunity to shape the learning outcomes of immigrant students, it was able to improve student performance" (OECD 2010, p. 75).

17.   The size of these gaps, however, varied considerably across countries, with the largest late-arrival penalties being found in Italy, Belgium, Sweden and Ireland. There were also a small number of non-Western countries, such as Qatar and Macao-China, where there were late arrival 'premia' rather than penalties. That is to say in these countries the late arrivals performed better in the reading test than those who arrived earlier. One possible explanation for these puzzling cross-country variations that the 2010 report was unable to address is that the kinds of students, for example their national origins, who arrive early and late may vary between destination countries. We shall need to consider these 'compositional' differences carefully.
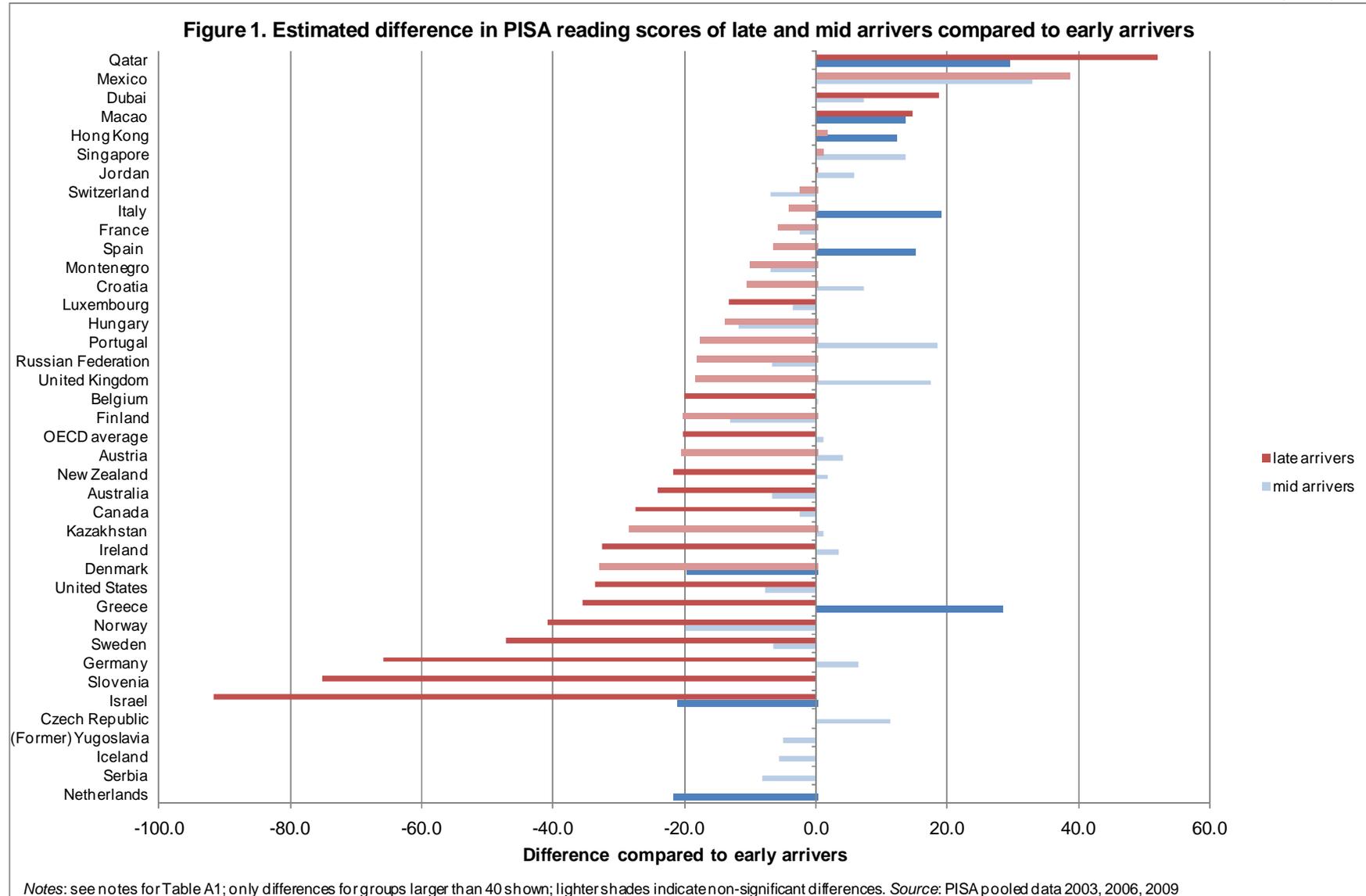
Country differences in late-arrival penalties

18.   We now turn to our own analysis of the PISA data. We focus on the test of reading literacy, which is particularly important for educational success and integration more generally, and for which age-at-arrival effects appear to be stronger than for mathematics or science. For this analysis we pool together all relevant years in which PISA collected the necessary data (2003, 2006 and 2009), all countries and all origin groups. We also pool the results for both boys and girls. We checked for differences by gender and while, as expected, girls generally perform better than boys, we find that there is no significant gender difference with respect to late-arrival penalties. In other words, the female advantage over males is broadly similar at all ages of arrival.

19.   We begin with some descriptive statistics showing for each participating country the differences in reading test scores between early and medium, and between early and late, arrivers. Following OECD (2010) we define the early arrivers as those who arrived at or before age 5 (which will generally cover children who arrived before the start of compulsory schooling). And we define the late arrivers as those who arrived after age 12. This latter group will thus have had a maximum of three years schooling in the destination country at the time of the test. Figure 1 shows the differences in the reading scores of the late- and mid-arrivers in comparison with those of the early-arrivers in each country (which are set to zero). (The full details are provided in Table A1 in the Appendix.) We show the results only for cases where there are reasonable numbers of immigrant students in the sample for each country. For a few countries (shown at the bottom of Figure 1) there are sufficient cases to show the differences between mid- and early-arrivers but not sufficient for investigating late-arrivers.

20.   It is important to remember that the patterns shown in Figure 1, and throughout this paper, are derived from cross-sectional not from panel data. That is to say, they show the differences in reading scores at the end of lower secondary education between fifteen-year olds who arrived in the destination country at different ages, and have thus spent a longer or shorter time in the country of current residence. As noted earlier, since they are not derived from a panel study, the figures do not show how test scores for the same individuals change over time, although it will often be reasonable to suppose that a panel study might show similar patterns to those documented here.

21.   We must emphasize, too, that these introductory descriptive statistics should be regarded as provisional and need to be checked with more sophisticated analysis in due course, in particular taking into account any 'compositional' differences in the origin countries from which the migrant students come. As we shall see in the next section, migrants from countries which also speak the test language tend to have flatter age-at-arrival profiles than do migrants who come from different linguistic backgrounds. A large late-arrival penalty may therefore simply reflect a large inflow of migrants who need to learn a new language in the destination country. Nevertheless, the overall comparisons do provide us with an overview of the main patterns in the data, and also provide us with a yardstick against which to check for important deviations. (We should also note that Figure 1 presents the 'gross' differences before controlling for parents' socio-economic background. Table A2 in the appendix shows the 'net' results after controls. As with van Ours and Veenman's study, the controls generally have little effect on the size of the late-arrival penalties.)

Figure 1. Estimated difference in PISA reading scores of late and mid arrivers compared to early arrivers

Qatar
Mexico
Dubai
Macao
Hong Kong
Singapore
Jordan
Switzerland
Italy
France
Spain
Montenegro
Croatia
Luxembourg
Hungary
Portugal
Russian Federation
United Kingdom
Belgium
Finland
OECD average
Austria
New Zealand
Australia
Canada
Kazakhstan
Ireland
Denmark
United States
Greece
Norway
Sweden
Germany
Slovenia
Israel
Czech Republic
(Former) Yugoslavia
Iceland
Serbia
Netherlands

■ late arrivers
■ mid arrivers

Difference compared to early arrivers

Notes: see notes for Table A1; only differences for groups larger than 40 shown; lighter shades indicate non-significant differences. *Source*: PISA pooled data 2003, 2006, 2009
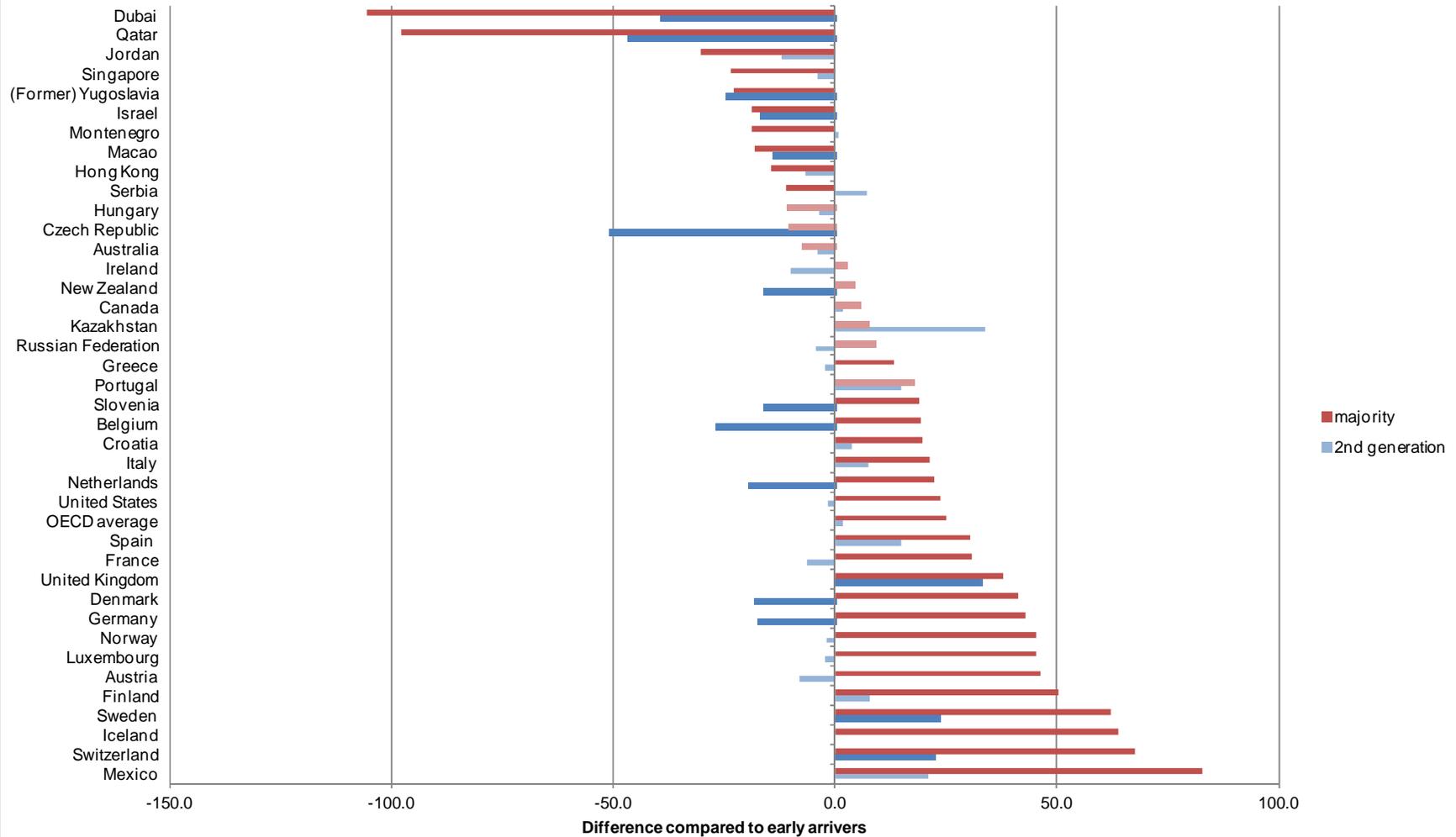
22.   As noted above, in Figure 1 we set the performance of the early arrivers at zero, and then compare the performance of the mid- and late-arrivers with that of the early arrivers. In general we see that there is a late arrival penalty, albeit of varying size, in almost all of the developed Western countries. The OECD average is a late-arrival penalty of 20 points, roughly equivalent to about half a year of schooling. Furthermore, in most countries we find that the mid-arrivers are rather closer to the early arrivers than to the late-arrivers. This suggests that the effect of age-at-arrival on test scores is not a linear one but becomes progressively larger the later the age at which the student migrates. This suggests that a curvilinear characterization of the age-at-arrival/performance profile is more appropriate than a linear one. More detailed analysis (available on request) indicates that there is no specific 'critical' age for arrival but that later arrivals are increasingly vulnerable.

23.   In contrast to this dominant pattern of a late-arrival penalty, in a small number of non-Western or less-developed countries such as Qatar, Mexico, Dubai and Macao, we see that late arrivers actually perform better than the early arrivers (although the difference is not statistically significant in Mexico). This may well reflect particular features of the migrant flows to these countries. As we shall show later, a late-arrival 'premium' rather than a penalty can occur when children migrate at an older age, having spent a larger part of their educational career in a country with higher educational standards.

24.   A further point to bear in mind is that even the early arrivers may perform less well than the majority group in the country of residence. We should not therefore jump too readily to the conclusion that the absence of a late-arrival penalty means that all is well for the immigrant children. This will become clearer if we compare the results of Figure 1 with those of Figure 2, which compares the test scores of the early arrivers with those of the majority group. Figure 2 also compares the early-arrivers' test scores with the scores of the native-born (second-generation) children from immigrant backgrounds.

Figure 2. Estimated difference in PISA reading scores of majority and 2nd generation compared to early arrivers

Notes: see notes for Table A1; only differences for groups larger than 40 shown; lighter shades indicate non-significant differences. Source: PISA pooled data 2003, 2006, 2009

25.   As in Figure 1, we set the performance of the early arrivers at zero, and we then compare the performance of the majority group, and of the second generation, with this baseline. Again, we can see that there is considerable cross-national variation, but with the majority group in many Western developed countries outperforming both the early arrivers and the second generation, sometimes by considerable margins. (As Table A2 in the appendix shows, the net differences between the majority and the immigrant groups after controls tend to be somewhat smaller than the gross ones shown in Figure 2, but in most cases the differences remain quite substantial and statistically significant.) Figure 2 also suggests that the scores of the second generation tend to be rather close to those of the early-arrivers; in some countries the second generation do rather better, while in others they score somewhat worse, but the overall average for OECD countries is virtually identical for early-arrivers and the second generation. This makes intuitive sense and suggests that any disadvantage that the early arrivers have in comparison with the majority group in the country of residence is due more to ethnic factors than to any problems associated with migration *per se*. We can interpret this as an 'ethnic penalty' as distinct from the 'late-arrival penalty' shown in Figure 1.

Figure 3. Estimated difference in PISA reading scores of late arrivers and majority compared to early arrivers

Late arrivers compared to early arrivers

Majority compared to early arrivers

line of best fit
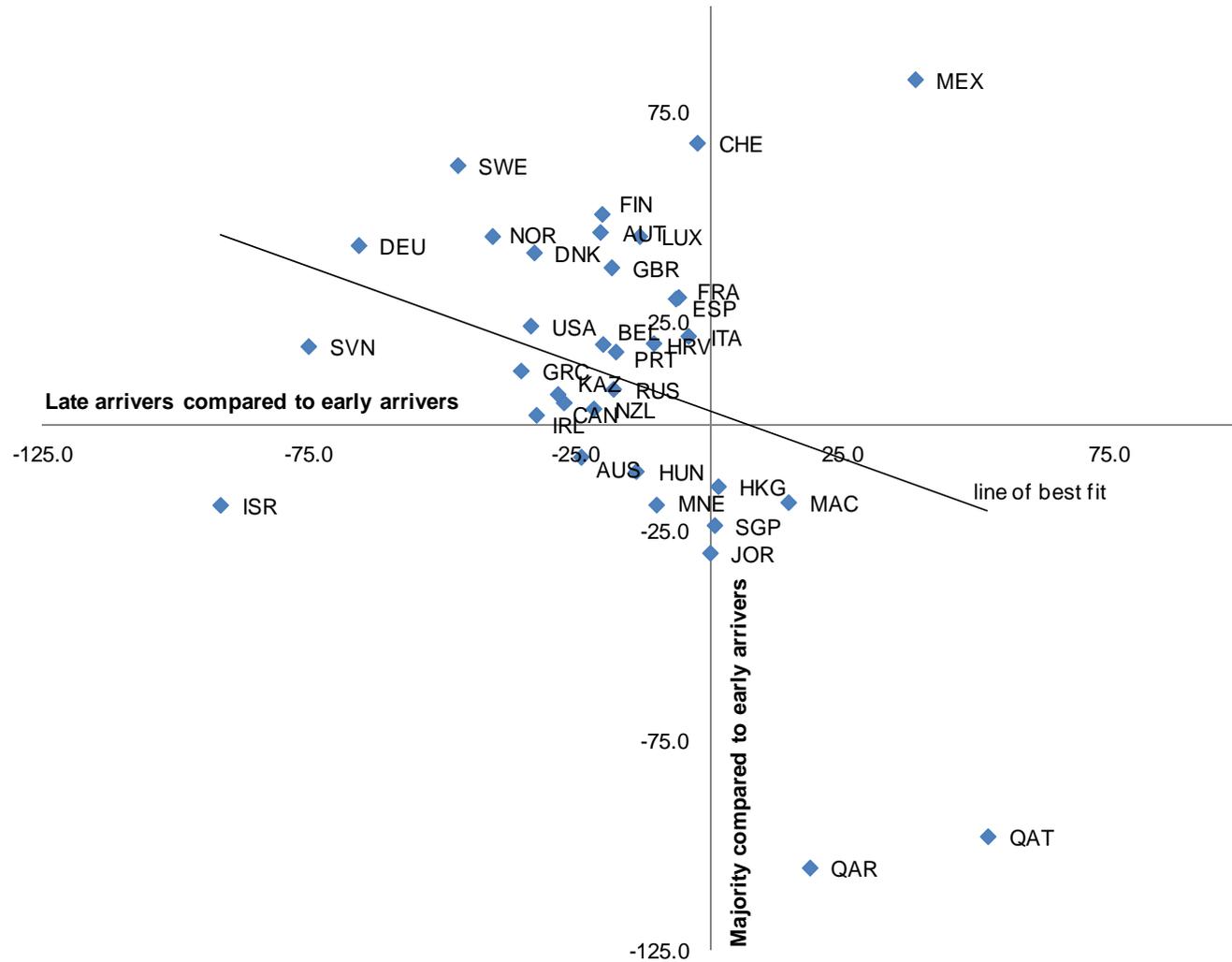
Notes: see notes for Table A1; only differences for groups larger than 40 shown. *Source*: PISA pooled data 2003, 2006, 2009

26.   In Figure 3 we plot the relationship between the early/late arriver difference (taken from Figure 1) and the early arriver/majority group difference (taken from Figure 2). As we can see, there is a general pattern for the size of the late-arrival penalty to be correlated with the size of the ethnic penalty, although Mexico is a striking outlier. The majority of OECD countries fall into the top left-hand quadrant of the diagram; in effect these are countries where there is both a late-arrival penalty and an ethnic penalty. Thus in countries such as Germany, Sweden and Norway we see that there is **both** a substantial late-arrival penalty (compared with the early arrivers) **and** a substantial early-arrival penalty (compared with the majority group). There thus appears to be a pattern of cumulative disadvantage in these countries. However, in many other countries the pattern is not cumulative. In some such as Australia and New Zealand (those below the line of best fit), late-arrival penalties are much larger than the ethnic penalties, while in others such as Switzerland the late-arrival penalties are rather small compared with the ethnic penalties. This suggests that perhaps the policy priorities in these countries need to be rather different.

27.   In contrast, in the lower right-hand quadrant of the diagram we see a number of non-Western countries, notably Qatar and Dubai, where the late arrivers score more highly than either the early-arrivers or the majority group, again probably reflecting distinctive features of the kinds of young migrants who arrive late in their school careers in these particular countries. This reinforces our suspicion that in these cases the late arrivers are a rather unusual set of migrants, atypical of the familiar pattern (in Western countries) of migrants coming from less-developed countries, and from less advantaged socio-economic backgrounds, to more developed countries.

28.   More generally, it is important to recognize that many of the patterns observed in Figure 3, including those in the top left-hand quadrant, may be due to the composition of the migrant groups. Thus there may be little in the way either of late-arrival or ethnic penalties in countries where most migrants come from countries at similar levels of development or speaking the same language, whereas larger penalties may be expected in countries where migrants typically come from less-developed origins and have to learn a new language. We explore this in more detail in the following sections of this paper.

Language and the late-arrival penalty

29.   Previous research has often found differences between origin groups in the extent to which later age-at-arrival depresses test performance, although the results vary across countries of residence. For example, in Israel Cahan *et al* (2001) found that children arriving from Western Europe had a much steeper age-at-arrival/performance profile than children arriving from North Africa and Western Asia. In contrast, in Sweden Böhlmark (2008) found that children arriving from Western Europe suffered little in the way of a late age-at-arrival penalty whereas there was a steep profile for those arriving from Asia. One reason for these differences is almost certainly, as Böhlmark suggests, the possession of destination-country specific skills, most notably language (although other factors such as the educational levels typically achieved in the country of origin may also be relevant as we shall see in the next section).

30.   The implication of this is that age-at-arrival profiles may have quite different characters and explanations from overall performance levels. Thus children from East Asia and India generally perform very well in performance tests, and appear to do so whatever the country they have migrated to. (Overall in the PISA datasets we find that first-generation Chinese outperformed the native-born by 4 points and first-generation Indians by 23 points in OECD countries.) In contrast, their age-at-arrival profiles are likely to vary according to the particular country which they have migrated to. For example, mainland Chinese migrating to Hong Kong, where Chinese is the language of instruction in schools, might be expected to have a flatter profile than those migrating to Australia, where they will be taught in English and might hence experience a steeper learning curve. In other words, it is the congruence or lack of congruence between the language of the home and that of the school that is likely to affect the age-at-arrival/test score profile.

31.   In practice it is not straightforward to investigate the differences in profiles of specific migrant groups in different countries of destination, as there are relatively few origin groups found in several different destination countries. There is also a selection problem – migrants often move to countries which are culturally and linguistically more similar, for example Britons tending to migrate to Australia or New Zealand, all English-speaking countries and hence ones where we might expect the age-at-arrival profiles of Britons to be fairly flat.

32.   However, we can compare Chinese who are found in reasonable numbers in our samples in Australia, Hong Kong, Macao and New Zealand. Our expectation is that Chinese migrating to Hong Kong and Macao will have relatively flat age-at-arrival profiles, since they have moved to countries with similar cultures and language of instruction in schools, whereas Chinese migrating to English-speaking Australia and New Zealand will have much steeper age-at-arrival profiles. It is important to recognize that this does not mean that Chinese or Asians in general have steeper profiles than Britons and other Western Europeans: our expectation is that the steepness of the profile will vary from one destination country to another. Some Western European groups migrating to countries where there is a different language of instruction in school (*e.g.* Germans migrating to Belgium) might therefore also be expected to have quite steep age-at-arrival profiles.

33.   In Table 1 we show age-at-arrival profiles for selected origin groups. Technically, for each country we regressed test scores on the natural logarithm of years since arrival (that is, the number of years spent in the destination country by the time of the test), fitting both main effects for years since arrival and interaction effects, allowing the coefficients to vary across origin groups. After experimenting with alternative functional forms, we decided that the most parsimonious and tractable form is to take the natural logarithm of the number of years spent in the destination country at the time of the test. The logarithmic transformation takes account of the fact that age-at-arrival differences are relatively small during the earlier years and that the differences gradually increase thereafter. It also provides a significantly better fit to the data than a simple linear model.[1] In Table 1 we report the main effect added to the interaction effect, thus showing the profile for that origin group in the particular country of destination. The larger the coefficient, the steeper the profile; that is the greater the late-arrival penalty. Note that negative coefficients indicate that a later age at arrival is associated with better, not worse, test scores which, as we shall see in the next section is a theoretically-important possibility. (The models also include controls for PISA year, gender and students' school year/grade. We fitted and report interactions only for those cases where at least 40 respondents from a given origin country were sampled in the country of destination.)

---

1 .   A quadratic function provides an even better fit to the data since it better captures the lower test scores of those who arrived in the first year of life. However, a quadratic function does not lend itself easily to modelling interaction terms, which are a key feature of our analysis, and we therefore prefer the more parsimonious logarithmic transformation.

Table 1.   Age at arrival profiles of young people from different countries of origin

| Country of residence	Country of origin									
	Africa	Britain	China	France	Germany	Portugal	Turkey	Former USSR	Former Yugoslavia	All immigrants
Australia	25.9	-5.7	48.5							12.9
Austria					*-22.1*		18.0		22.0	7.5
Belgium	26.2			9.7	24.0		*-2.5*			9.9
Croatia									6.2	4.2
Denmark									*-9.1*	23.7
Estonia								*8.4*		*9.0*
Finland								*52.6*		22.2
Germany							*48.9*	32.5	*36.5*	36.0
Greece								6.4	14.7	9.9
Hong Kong			-1.3							-1.5
Ireland		3.7								21.4
Israel	*70.9*							41.1		53.4
Kazakhstan								22.6		23.7
Luxembourg	*54.2*			4.3		18.4			60.2	6.9
Macao			-7.9							-7.4
Montenegro								4.4		6.0
New Zealand	41.5	15.1	35.0							16.5
Portugal	*36.2*									16.4
Russia								9.1		9.6
Serbia									17.9	17.4
Switzerland				*19.9*	5.8	27.9	*-7.4*		34.7	1.9
Yugoslavia									-4.3	-1.7

Source: PISA pooled data, 2003, 2006, 2009

34.   While standard errors for the coefficients reported in Table 1 are generally rather large (because of the small sample sizes), the overall pattern is nonetheless quite striking and in line with our expectations. Thus we do indeed find that the age-at-arrival profile is much steeper for Chinese in Australia and New Zealand than it is in Hong Kong or Macao. The profile for Germans migrating to Belgium is much steeper than for Germans moving to Austria or Switzerland (where we assume many will have moved to German-speaking parts of Switzerland). The profile for young people from the former USSR is much steeper in Israel, Finland and Germany than it is in Russia. And the profile for young people from the former Yugoslavia is much steeper in Luxembourg, Germany and Switzerland than it is in Serbo-Croat speaking Croatia or Serbia. In contrast the profiles for Britons or French, who typically migrate to other Anglophone or Francophone countries respectively, are uniformly fairly flat.

35.   Table 1 also shows in the final column the overall coefficient for young migrants in each country of destination, and it is informative to compare this with the coefficients for the individual origin groups within a given country. Consider Germany for example, which Figure 1 indicated was one of the countries with the largest late-arrival penalty. Overall, Germany does indeed exhibit a large age-at-arrival coefficient of 36.0. However, the main immigrant groups to Germany with sufficient numbers in the sample for detailed analysis are all ones from linguistically-dissimilar countries, namely from Turkey, the former-USSR and the former Yugoslavia. Conversely Australia, which overall displays a much smaller age-at-arrival coefficient of 12.9, also exhibits a large coefficient of 48.5 (as large as the biggest coefficient in Germany) for the linguistically-dissimilar migrants from China. The small overall Australian coefficient can almost certainly be explained by its large number of migrants from linguistically-similar Britain with

their small coefficient of -5.7. In other words, the overall differences in the magnitude of the late-arrival penalties observed in Figure 1 almost certainly owe a great deal to the composition of the immigrant flows, specifically whether the immigrants come from linguistically-similar or dissimilar origin countries.

Figure 4. PISA reading attainment by country of origin and age at arrival in certain countries of destination
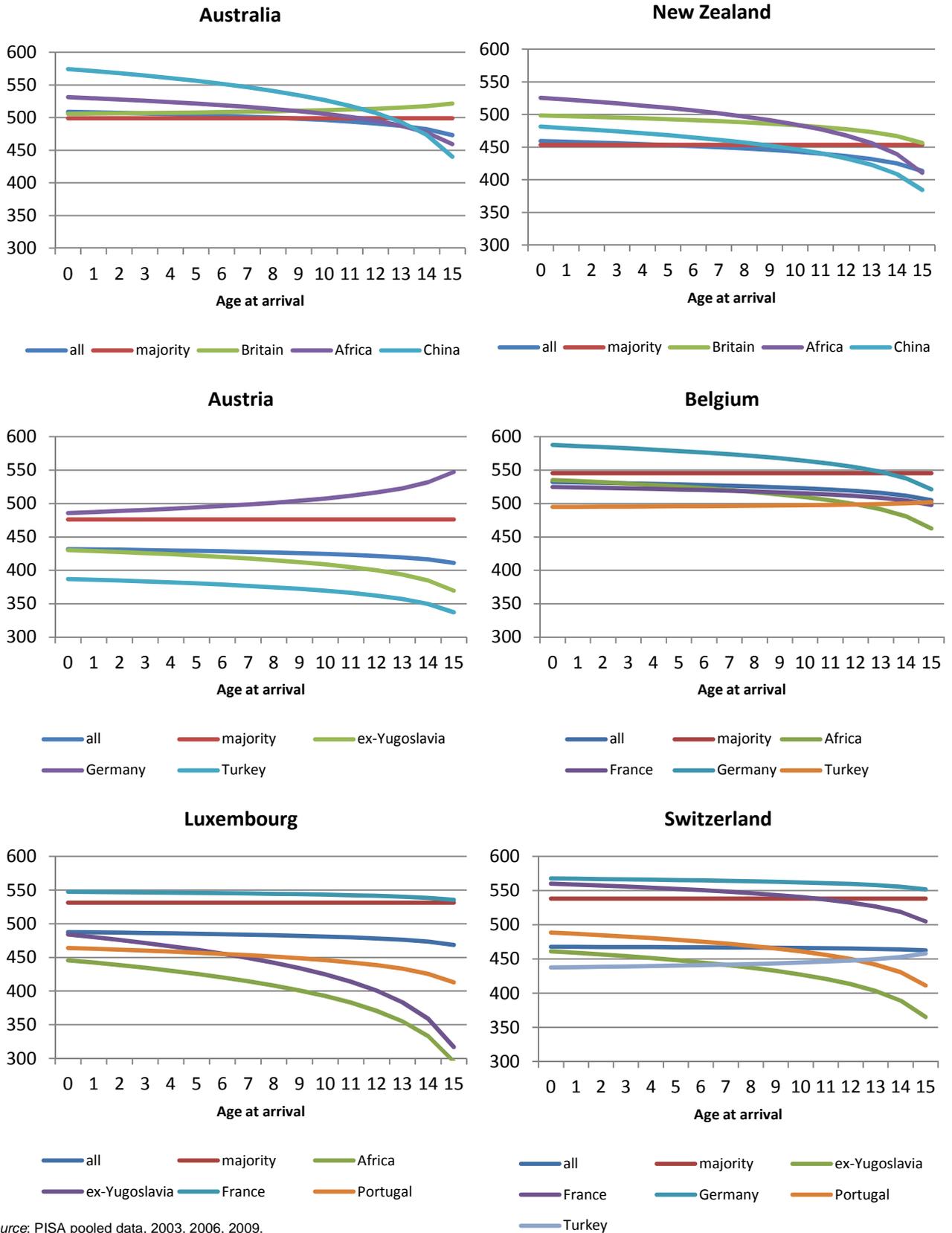
Source: PISA pooled data, 2003, 2006, 2009.
Note: all estimates control for PISA year (2009 as reference), gender (male as r

36.   To illustrate the patterns, we plot the curves for a selection of countries. Thus in the top panel of Figure 4 we show the age-at-arrival curves for Australia. Here we can see that the majority group is high-performing, that there is a flat curve for arrivals from Britain, and a steep curve for arrivals from Africa and even more so from China. Especially notable is the fact that the early arrivals from China clearly outperform the majority group, whereas the late arrivals score substantially lower.

37.   New Zealand, in the next panel, tells a rather similar story with steeper curves for the migrants from Africa and China, and with early arrivals outperforming the majority group and late arrivals faring worse.

38.   In Austria we see the later-arrival premium for the migrants from Germany, and the late-arrival penalties for the migrants from the former Yugoslavia and from Turkey. However, this figure also brings out the important point that the rather flat curves for the latter two groups do not tell a positive story about these migrants. Instead, what we see is that even the early arrivals from the former Yugoslavia and from Turkey have much lower scores than the majority group at age fifteen. This suggests that there are large ethnic penalties in Austria (in line with the position of Austria in Figure 3). In contrast, in Belgium the migrants from Turkey do not appear to suffer either an ethnic or a late-arrival penalty, performing at much the same level as the majority group throughout (although we should be careful about drawing any strong conclusions from this finding as it is based on a sample of only 61 migrants from Turkey).

39.   It is also striking that we see both late-arrival penalties and ethnic penalties for minorities coming from linguistically-dissimilar countries to Luxembourg and Switzerland. In contrast the curves are much flatter for migrants from France to Luxembourg, and for migrants from Germany to Switzerland, and there are no ethnic penalties for these latter groups of migrants either. Again, it is important to note that Switzerland, despite the absence of a late-arrival penalty overall, displays major late-arrival penalties for the late arrivers from Portugal and the former Yugoslavia.

40.   To be sure, there are some anomalies in Table 1 (and Figure 4), such as the profile for migrants from Turkey in Belgium, some of which may be due to sampling error given the small sample sizes involved. We therefore conducted a more formal test of the hypothesis that late-arrival penalties are larger when the young people come from linguistically-dissimilar countries. The PISA datasets record whether the young people were tested in the same language as they speak at home. We can use this measure to test statistically our informal observations based on the coefficients displayed in Table 1. The result is to confirm a significant difference in the age-at-arrival coefficients for those who do and do not speak the test language at home.[2]

41.   However, we need to recognize that the measure of language spoken at home is far from ideal. The measure gives the **current** language spoken at home, which may not always be the same as the language spoken on arrival. Thus many early arrivers (and their families) may have shifted from speaking their origin-country language in their early years in the new country to speaking the destination-country language by the time of the test. This is likely to bias the results of the statistical test although possibly the bias might lead us to underestimate the difference: thus we can divide the early arrivers into three groups: those who speak the test language at home throughout their lives, those who speak a non-test language when they migrate but switch to the test language by age 15 and those who speak a non-test language throughout their lives. If we assume that the test scores of these three groups go from highest to lowest in the order that we have just listed them in, then when we look at the age-of-arrival effect of current test language speakers, this includes a slightly worse-performing group of those who did not speak the test

---

2.   The co-efficient for those who speak the test language at home is 5.1 with a standard error of 2.0 and the interaction for those who do not speak the test language at home is 9.6 with a standard error of 2.5 (p=0.0002).

language on arrival, whereas the non-test group is missing the better-performing students (in comparison). If those who switched language do not differ in progress from those who have always spoken the test language and these two groups perform throughout at a higher level than those who speak a non-test language, then the age-at-arrival effect is correct for the test speakers but is dampened for the non-test speakers. But this does depend on what we assume about the test scores and progress of these three groups (which unfortunately we cannot check with the currently available data).

42.   While the need to learn a new language is clearly important in explaining the late-arrival penalty, we suspect that it is not the whole story. Even among those who speak the test language or who come from linguistically-similar origin countries, we still find a modest-sized late-arrival penalty. In the following section we therefore consider an additional explanation for the penalties, namely the disparities in educational standards between the origin and destination countries.

Origin/destination disparities in educational standards

43.   One possibility is that migrants moving from a country with lower educational standards to one with higher standards will have more ground to make up and thus will be particularly penalized if they arrive late (having spent a larger proportion of their school career in the country with lower standards and thus being even further behind). Just as in the case of language, the crucial factor is the extent of **difference** between the origin and receiving countries in educational standards, not simply on the level of performance in the receiving country. For example, migrants going from one high-achieving West European country to another will have little ground to make up in their general literacy skills, even if they speak a different specific language, and we hence will not expect much catching-up to be necessary. In contrast, where we have a big gap between average performance in the origin and destination countries, there is much more ground to make up and we expect a steeper profile. There is a further possibility too: some young people might be moving from a high-achieving to a low-achieving country. In this case we might expect that late age-at-arrival might be beneficial since the young people will have spent more time in the educational system of the country of origin. In other words, the usual pattern of a penalty for late arrival might be reversed with a premium for late arrivals. And indeed there were some hints of this in Table 1, where migrants from higher-achieving Germany (mean score in PISA 2009 of 497) had a late-arrival premium in lower-achieving Austria (mean score of 470) as shown by the negative coefficient.

44.   In Table 2 we repeat the same kind of analysis that we reported in Table 1 although now, in order to sidestep the issue of language differences, we show the age-at-arrival coefficients only for those students coming from linguistically-similar countries. (We have also carried out analogous analyses of students who reported speaking the test language at home. This analysis however is more problematic because of the possible biases, discussed earlier, arising when young people have shifted to speaking the test language at home after arrival in the destination country.)   As before we focus on countries of destination and show the coefficients (main effects plus the interaction effects with country of origin) after controls for PISA year, gender and school year/grade. Because of small sample sizes, now shrunken further because of the restriction to students coming from linguistically-similar countries, we have relatively few coefficients to inspect.

Table 2.   Age-at-arrival profiles of young people from different countries of origin in linguistically similar countries of destination

Country of residence	Country of origin													
	China (Shanghai 557)	New Zealand (526)	Germany (511)	France (505)	Britain (499)	Former USSR (Russia 464)	Former Yugoslavia (Serbia 442)	Brazil (416)	Jordan* (HDI 0.681)	Egypt* (HDI 0.620)	South Africa* (HDI 0.597)	Yemen* (HDI 0.439)	Palestine*	All immigrants
Hong Kong (535)	-1.3													-1.5
New Zealand (526)				15.1							*36.9*			16.5
Belgium (519)				9.7										9.9
Australia (515)		15.6			-5.7						11.3			12.9
Switzerland (513)			5.8	*19.9*										1.9
Estonia (505)						*8.4*								*9.0*
Ireland (502)				3.7										21.4
Luxembourg (495)				4.3										6.9
Portugal (492)								9.7						16.4
Greece (489)							14.7							9.9
Austria (482)			*-22.1*											7.5
Macao (482)	-7.9													-7.4
Qatar (331)									4.2	-14.0		-2.0	*-16.4*	-27.8
Notes: *non-test language speakers specifically excluded														
Minimum sample size: 40, estimates based on samples less than 100 italicized; all estimates control for PISA year (2009 as reference), gender (male as reference) and student's grade (10th as reference)														
Source: PISA pooled data 2003, 2006, 2009, OECD (2010) for country scores, and UNDP's Human Development Index (HDI) for 2009														

45.   We have arranged the countries in Table 2 according to the average PISA reading scores in 2009 of young people belonging to the majority group, with lower-scoring sending countries in the right-hand columns and lower-scoring receiving countries in the bottom rows of the table. Some of the (less-developed) origin countries did not participate in PISA, and we have arranged these according to their scores on the UNDP's Human Development Index (HDI) for 2009, since there is quite a strong overall relationship between HDI scores and PISA test scores. This is not at all surprising given that educational enrolment is one of the three indicators used to build the HDI (the others being per capita GDP, which is truncated in order not to give too high a weight to higher levels of GDP, and life expectancy at birth). The HDI score is not available for the Occupied Palestinian Territories (although on the two components which are available, the OPT comes close to Jordan). It is also worth noting that all four have lower HDI scores than Brazil, which comes next to them in the table.

46.   Our general expectation is that we will see the largest age-at-arrival penalties in the top right-hand corner of the table, where students have moved from a country with lower educational standards to one with higher standards. And we expect to see age-at-arrival premia in the bottom left-hand corner of the table, where students will have moved from a country with higher educational standards to one with lower standards.

47.   The picture is not especially clear as the table is very sparsely populated, and many of the coefficients are based on small sample sizes. However, we do see that the largest late-arrival penalty occurs as expected in the top-right corner of the table where we have migrants from less-developed South Africa to high-achieving New Zealand. (Note that we have explicitly excluded students from South Africa who did not speak the test language – in this case English.) We also see that the largest late-arrival premium occurs towards the bottom-left corner of the table where we have migrants from high-achieving Germany to lower-achieving Austria. We also see late-arrival premia among the students moving to Qatar, which is by far the lowest-achieving country in the table.

48.   However, the other coefficients in the table do not show a clear pattern, and the results overall are not sufficiently strong to give us great confidence in our hypothesis. They are however suggestive and indicate that the hypothesis should be taken seriously and warrants more detailed investigation as more data become available.

49.   One further complication which one needs to recognize is that migrants are often a highly-selected group from their country of origin. In other words, migrants from a less-developed country will often not be typical of the non-migrant population but may have greater social, cultural and human capital than do those families who do not attempt to migrate. Given that migration is a hazardous enterprise, which often requires resources well beyond the means of the poorer members of a developing country, we should assume that migrants, especially those from developing countries, may be rather atypical. To be sure, our theory of differences between the educational standards of sending and receiving countries is not one about individual differences in achievement but about differences in systemic educational standards. Nevertheless, we might expect migrants to come from more urbanized settings where educational opportunities and facilities are more developed. For a thorough test of the hypothesis we would therefore need to have more information about the specifics of the educational institutions from which the young migrants had come.

Combined analyses

50.   In order to assess the relative importance of these different processes, and to determine what country differences remain after taking account of these processes, we undertook a multivariate analysis of the dataset. Our dependent variable is, as before, the reading test scores of the young migrants. We include

only young migrants in the analysis (excluding the second generation and members of the majority group) in order to focus on the effects of late arrival. (The total sample size for these analyses is therefore 11,299 young migrants.) And we restrict our analyses to destination countries which are relatively developed, and which generally have high overall scores on the reading test and on the HDI, in order to sidestep the differing patterns to be found in the less-developed destinations (and which would have involved more complex interaction terms). We also exclude migrants with an unreported country of origin as they cannot be classified into the categories used in this analysis (described below). All the included countries are given equal weight in the analyses, the results of which are shown in Table 3.

Table 3. Age-at-arrival effects for different types of immigrants in western countries

	Model 1			Model 2			Model 3			Model 4		
	no interactions			interaction with immigrant category			interaction with country			both interactions		
	estimate	s.e.		estimate	s.e.		estimate	s.e.		estimate	s.e.	
Age at arrival (log transformed)	18.5	5.3	***	7.6	4.5	*	20.9	4.3	***	11.8	7.1	
Sending country (linguistically-similar western country as reference)												
Linguistically-dissimilar western country	-30.4	9.9	**	-41.4	22.7	*	-29.3	11.7	**	-41.5	62.0	
Linguistically-similar non-western country	-23.7	17.7		-15.7	18.5		-23.1	15.7		-10.7	24.0	
Linguistically-dissimilar non-western country	-51.4	12.2	***	-92.5	12.0	***	-50.5	12.9	***	-89.6	110.2	
Age at arrival interaction with immigrant category												
*different, western				7.5	13.0					8.0	14.1	
*similar, non-western				-4.2	10.1					-6.9	11.6	
*different, non-western				21.7	8.3	**				20.8	9.9	**
Country of destination (Australia as reference)												
Austria	-40.8	18.1	**	-44.5	19.3	**	-39.9	28.6		-25.1	25.9	
Belgium	-31.4	10.5	**	-33.3	10.6	**	-29.4	15.0	*	-36.7	15.9	**
Czech Republic	-31.2	23.5		-34.2	22.6		-51.1	61.3		-54.8	66.6	
Denmark	-13.6	21.0		-18.1	22.4		-44.5	46.7		-21.8	46.2	
Estonia	-9.3	22.0		-15.7	23.3		16.7	78.6		39.5	79.4	
Finland	67.0	26.6	**	64.5	27.2	**	4.3	49.1		23.8	46.5	
Germany	-1.7	21.8		-5.0	23.1		-42.4	30.5		-25.7	29.1	
Greece	-45.8	14.5	**	-49.0	15.7	**	-32.1	33.8		-10.9	30.3	
Ireland	11.8	13.3		12.6	12.7		40.1	27.0		31.2	28.5	
Israel	-24.7	12.3	**	-28.2	13.4	**	-113.0	37.1	**	-94.5	35.9	**
Italy	-51.2	13.3	***	-54.0	12.5	***	-26.9	27.0		-31.0	27.7	
Luxembourg	-48.9	15.8	**	-50.6	16.4	**	-36.6	18.1	**	-38.5	17.8	**
Netherlands	-9.8	16.7		-11.8	17.2		6.5	49.1		14.0	49.9	
New Zealand	-16.8	7.2	**	-16.4	7.3	**	-11.1	12.5		-13.1	12.5	
Portugal	-24.6	19.0		-26.0	20.0		-6.9	22.0		-22.0	23.0	
Switzerland	-25.2	17.4		-29.2	18.5		13.0	16.5		15.0	16.0	
United Kingdom	-29.1	13.7	**	-27.9	13.8	**	-59.5	32.0	*	-58.3	32.9	*
Age at arrival interaction with country of destination												
*Austria							-1.3	10.0		-9.6	10.2	
*Belgium							-1.5	6.8		1.9	6.7	
*Czech Republic							10.0	26.2		10.7	30.0	
*Denmark							12.2	20.5		0.5	21.4	

	Model 1 no interactions			Model 2 interaction with immigrant category			Model 3 interaction with country			Model 4 both interactions		
	estimate	s.e.		estimate	s.e.		estimate	s.e.		estimate	s.e.	
*Estonia							-11.4	32.7		-23.1	33.6	
*Finland							28.8	20.2		18.4	20.3	
*Germany							17.0	11.7		8.1	13.4	
*Greece							-6.9	12.2		-18.0	11.1	
*Ireland							-15.6	11.6		-10.1	12.1	
*Israel							37.5	15.2	**	27.6	15.2	*
*Italy							-14.0	12.7		-12.8	14.5	
*Luxembourg							-6.5	7.0		-6.4	10.0	
*Netherlands							-8.1	21.0		-12.3	22.5	
*New Zealand							-3.2	7.7		-1.6	7.9	
*Portugal							-11.0	9.8		-2.2	14.2	
*Switzerland							-17.8	7.9	**	-20.4	9.6	**
*United Kingdom							16.8	13.8		17.0	14.3	
Notes: *p<0.10, **p<0.05, ***p<0.001												
All countries have equal weight in these analyses; all estimates control for PISA year (2009 as reference), gender (male as reference) and student's grade (10th as reference)												
Source: PISA pooled data 2003, 2006, 2009												

51.   The first column of results in Table 3 shows the coefficients estimated from a model which includes age-at-arrival, destination country, and type of sending country as the predictors. We distinguish four types of sending country, namely

- Linguistically-similar Western countries (the reference category),

- Linguistically-dissimilar Western countries,

- Linguistically-similar non-Western countries, and

- Linguistically-dissimilar non-Western countries.

52.   As we can see, in this first model there are significant coefficients for age at arrival, for migration from linguistically-dissimilar countries (both Western and non-Western) and for nine countries of residence (relative to the reference country of Australia). We should note that, since age-at-arrival is coded as the natural logarithm of years spent in the destination country, the estimates for countries of origin and destination effectively tell us about the reading scores of young people who arrived in the destination country at ages 15 and 16 (which are combined in the coding of age at arrival). That is, late arrivals from a linguistically-dissimilar non-Western country are estimated to score on average 51 points less than late-arrivals from a linguistically-similar Western country; and late arrivals in Finland are estimated to score on average score 67 points more than late arrivals in Australia, while late arrivals in Italy on average score 51 points worse than those in Australia.

53.   In the second model we then introduce interactions between age-at-arrival and type of origin country. This in effect allows the late-arrival penalty to differ according to the type of country from which the migrants came. And the estimated coefficients indicate that late-arrival penalties are significantly greater for migrants coming from linguistically-dissimilar non-Western countries than they are for migrants coming from linguistically-similar Western countries.

54.   Perhaps the most helpful way to interpret these interaction terms is to add them to the 'main effect' of age-at-arrival, thus showing us what the effect of age-at-arrival is for migrants from a particular type of origin. Thus, for migrants moving from one Western country to another, linguistically-similar Western country, the age-at-arrival coefficient is a statistically-significant 7.6. Thus even these migrants experience a late-arrival penalty, possibly because of the difficulties of adjustment to a new educational system (or to unmeasured heterogeneity within our Western category, for example with respect to educational standards).

55.   For migrants arriving from linguistically-similar non-Western countries the age-at-arrival coefficient is effectively the same at 3.4 (7.6 – 4.2), while for those arriving from linguistically-dissimilar Western countries it is rather larger at 15.1 (7.6 + 7.5). Finally, for those arriving from linguistically-dissimilar non-Western countries it is significantly larger at 29.3 (7.6 + 21.7). It is this latter group of migrants, then, who have both to learn a new language and to make the transition from a non-Western to a Western educational environment, who experience much the largest late-arrival penalties.

56.   In the third model, we then introduce interactions between age-at-arrival and destination country. These interactions tell us whether the relationship between age-at-arrival and test scores is steeper in some countries than in the reference country of Australia. In effect, then, these tell us whether late-arrival penalties are greater in some countries than in others. (Note that this is importantly different from the meaning of the main effects of country, which simply tell us how the test scores of late arrivals in a particular country differ from those of late arrivals in Australia.) What we see is that only two of the interactions are statistically significant, those for Israel and for Switzerland – two countries which we saw in Figure 1 to be quite distinctive, Israel having the largest late-arrival penalty and Switzerland having the smallest.

57.   In the fourth model, we include both sets of interactions, but the story remains essentially the same: migrants from linguistically-dissimilar non-Western countries experience much larger late-arrival penalties than do those from other origins. Once we control for these differences in the composition of the migrant flow and the distinctive penalties which this particular group of migrants from linguistically-dissimilar non-Western countries experience, the remaining differences between Western destination countries in their late-arrival penalties prove to be statistically non-significant (apart from Israel and Switzerland).

58.   Given the absence of statistically-significant interaction terms, we should be careful about drawing any conclusions as to whether particular countries provide more effective institutional arrangements for integrating late arrivals in the school system. Even in the cases of Israel and Switzerland we need to be circumspect since their distinctive age-at-arrival/test score profiles may be due to unmeasured heterogeneity in their migrant flows; this is particularly likely to be the case in Israel where the migrant flows will be rather different from those going to any other destination, but we also saw in Figure 4 that late arrivals in Switzerland from the former Yugoslavia (many of whom were from Kossovo) experience large late-arrival penalties.

59.   Nevertheless, it is of some interest to report some of the data on language teaching that is available in PISA. Table 4 reports, for the same countries that we analysed in Table 3, a variety of indicators. In the first column we show the number of minutes that immigrant children reported that they spent each week on (test) language lessons each week and in the second column we show the difference from the time spent on language lessons reported by the majority group in each country. The following six columns then show what schools reported that they provided for fifteen-year-old students whose home language was not the test language: columns 3-5 show whether they received additional instruction designed to enhance skills in the test language, while columns 6-8 show whether such students attended a preparatory programme aimed at developing test language skills before transferring to regular classes.

Table 4. Minutes spent each week in test language lessons by immigrant and majority students and proportion of immigrant students in schools that provide support for students whose first language is not the test language (reported by schools)

			Does your school offer any of the following options to students in <national modal grade for 15-year-olds> whose <first language> is not <the test language>?					
	Minutes spent each week in test language lessons, reported by students		These students attend regular classes and receive additional periods of instruction aimed at developing <test language> skills			Before transferring to regular classes, these students attend a preparatory programme aimed at developing <test language> skills		
	Immigrants	Difference from majority	Arrival at or before age 5	Arrival ages 6-11	Arrival ages 12 and above	Arrival at or before age 5	Arrival ages 6-11	Arrival ages 12 and above
Australia	233	-5	68	68	69	41	35	41
Austria	167	28	56	80	66	5	4	7
Belgium	222	11	36	40	29	6	3	10
Czech Republic	193	11	26	x	x	6	x	x
Denmark	322	9	80	75	c	43	42	c
Finland	144	-7	96	x	x	59	x	x
Germany	200	18	36	42	c	18	16	c
Greece	189	-6	36	32	c	4	3	c
Ireland	184	4	c	94	97	c	23	20
Israel	199	-18	83	88	c	28	35	c
Italy	314	31	88	86	87	37	35	31
Luxembourg	227	45	53	53	43	24	37	73
Netherlands	163	2	63	52	c	16	36	c
New Zealand	246	4	95	95	95	72	74	75
Portugal	233	9	c	85	100	c	25	24
Switzerland	211	12	75	77	87	48	57	59
United Kingdom	215	-4	86	87	98	35	49	56
OECD average	229	14	67	75	76	35	43	49
Notes: Missing, not applicable, etc. excluded; too few observations in Estonia								
c: fewer than 40 observations, x - data included in early arrivers category								

Source: PISA 2009

60.	Table 4 certainly appears to show that arrangements differ from one country to another. For example some countries such as Austria, Greece, Ireland and Portugal make much greater use of extra tuition than of preparatory classes, while others such as Luxembourg, the Netherlands, and Switzerland make relatively greater use of preparatory classes. Belgium appears to make little use of either arrangement while New Zealand makes a great deal of use of both. These arrangements do not however show any clear relationship with the amount of time (as reported by the students) spent on instruction in the test language. Even more strikingly, there seems to be little relationship between these arrangements and the magnitude of the late-arrival penalty in a given country. Thus Israel and Switzerland were at opposite ends of the spectrum with regard to late-arrival penalties and yet appear to be very similar in their institutional arrangements. Conversely, countries such as Austria and Luxembourg which are very different in their arrangements turn out to be rather similar in the magnitude of their late-arrival penalties.

61.	Cross-sectional data such as this cannot on its own prove or disprove that certain institutional arrangements are more or less successful. There are two major problems with drawing any strong policy conclusions: first there are issues of causal direction: for example it could be that large late-arrival penalties may lead governments to provide greater resources for tackling the problem, and if the resources have only recently been established they may well not have had time to be effective. At the very least we need 'before and after' panel designs to get a handle on these issues of causal direction.

62.	Secondly, the effectiveness of particular arrangements may depend on the particular context in which they are being used. For example, the kinds of arrangements that need to be made for late (or early arrivals) may depend upon the diversity of the school; the requirements may be very different in a highly diverse school, where a majority of students lack fluency in the test language, from a less diverse school where there are just a few students, from a single origin country, who need to learn the test language.

Conclusions

63.	To summarize, our analyses have shown (more tentatively in some cases than in others):

- Test scores are typically lower for young people who arrive later in their school careers. In other words there are typically 'late-arrival penalties' for the first-generation students, although we also found some examples of 'late arrival premia'. The size of the penalty is much larger for late-arrivers than for mid-arrivers (relative to early arrivers).

- There is no difference in the profile of late-arrival penalties between boys and girls.

- Early arrivers (who arrive in the destination country before compulsory schooling begins) perform at roughly similar levels to second-generation students, although in many cases they remain disadvantaged compared with students from the majority group. In other words, there appear to be 'ethnic penalties' over and above 'late-arrival penalties'.

- The late-arrival penalty is greater when children migrate to a country where the test language differs from their home language. There is no general tendency for, say, West Europeans to have smaller late-arrival penalties and Chinese to have larger ones – it depends primarily on whether they have migrated to a linguistically-similar country or not.

- But language may not be the only factor involved. The difference between educational standards in the origin and destination countries may also be relevant. Thus if there is a big difference in

the educational standards between the origin and destination countries, the more the late arrivers will have fallen behind their contemporaries in the destination country.

- But in some cases children are moving from high-achieving to low-achieving countries, and in these cases there may be a late-arrival 'premium' rather than penalty.

- For students moving between countries with similar educational standards and with the same language, there is still a late-arrival penalty but it tends to be rather modest in size.

- Much of the apparent difference between destination countries in the magnitude of the late-arrival penalties appears to be due to the particular composition of the migrant flows (especially whether they are from linguistically-dissimilar and less-developed origin countries).

64.   These results clearly suggest that an especially vulnerable group consists of students arriving in the later years of lower secondary education from less-developed countries where the home language is different from the test language in the destination country. These students both have to gain rapid knowledge of the test language and have to catch up with the higher levels of attainment achieved by their peers in the country of destination, as well as coping with all the problems of adjusting to a new educational and social environment. This will present important challenges to educationalists, teachers and policy-makers.

65.   Cross-sectional analysis of the kind undertaken here cannot on its own provide a firm basis for policy recommendations. Essentially what our analyses can do is to identify which particular groups are most vulnerable, and it is very clear that it is the combination of needing to learn a new language together with coming from a less-developed to a more developed country that is especially problematic.

66.   Policies designed to counteract the problems can take at least two different forms. One, of the sort covered in Table 4 above, is to concentrate on providing extra language instruction in order to enable the late arrivals to overcome linguistic difficulties. This is clearly important but there is no clear evidence that any particular arrangements, or indeed any particular countries, have as yet really solved this problem. It may indeed be a rather intractable one, given the short period available for the late-arrivals to master a new language before taking the test.

67.   To be sure, test performance at age 15 is not in itself a critical issue providing that poor performance at this stage does not have major implications for the young people's future school careers. However, in most countries there are major decisions to be taken at age 16 about which track to follow in upper secondary schooling, or whether to continue with schooling at all. Relatively poor test performance at this stage may therefore have important implications for subsequent school careers. This suggests that, in addition to efforts to overcome language difficulties before these key decision points, arrangements should be made for subsequent help in order to mitigate the adverse consequences of the late-arrival penalties. These might take the form, for example, of further post-16 language instruction and flexible arrangements for these late arrivals to make delayed transitions to more attractive tracks in post-16 education. Flexible arrangements of this kind might be particularly important if, as we suspect though cannot rigorously establish, linguistic problems are not the only kinds of problem facing young migrants. In other words, while language instruction is clearly of the greatest importance, it may not be the only important angle. As the analyses of Table 3 showed, it is the **combination** of coming from a linguistically-dissimilar and less-developed country that is particularly disadvantageous for late arrivals. Coming from a linguistically-dissimilar but highly-developed country does not appear to pose the same risks. Now there could be many possible explanations for this, which we do not have the space or data to investigate in details, but the key point is that it is not solely language that is the source of vulnerability.

68.   In conclusion, then, we believe that our evidence shows the importance both of helping young migrants with their language difficulties while still in lower secondary school and also the importance of attempting to mitigate the adverse effects of these difficulties subsequently. Different social and educational contexts will probably require different solutions to these problems, but the problems appear to be of a rather general and wide-spread kind.

Acknowledgements

References

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

Böhlmark, Anders (2009) Integration of childhood immigrants in the short and long run – Swedish evidence, *International Migration Review* 43(2): 387-409.

Cahan, Sorel, Daniel Davis, Rachel Staub (2001) Age at immigration and scholastic achievement in school-age children: Is there a vulnerable age? *International Migration Review*, 35 (2): 587-595

Myers, Dowell, Xin Gao, and Amon Emeka (2009) The gradient of immigrant age-at-arrival effects on socioeconomic outcomes in the U.S., *International Migration Review* 43 (1): 205-229.

OECD (2010) *PISA 2009 Results: Overcoming Social Background – Equity in Learning Opportunities and Outcomes* (Vol II). Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264091504-en

van Ours, Jan C and Justus Veenman (2006) Age at immigration and educational attainment of young immigrants. *Economics Letters* 90 (3): 310–316

THE OECD EDUCATION WORKING PAPERS SERIES ON LINE

The OECD Education Working Papers Series may be found at:

- The OECD Directorate for Education website: www.oecd.org/edu/workingpapers

- The OECD's online library, SourceOECD: www.sourceoecd.org

- The Research Papers in Economics (RePEc) website: www.repec.org

If you wish to be informed about the release of new OECD Education working papers, please:

- Go to www.oecd.org

- Click on "My OECD"

- Sign up and create an account with "My OECD"

- Select "Education" as one of your favourite themes

- Choose "OECD Education Working Papers" as one of the newsletters you would like to receive

For further information on the OECD Education Working Papers Series, please write to: edu.contact@oecd.org.