Migrant Achievement Penalties in Western Europe: Do Educational Systems Matter?

Published version:

DOI:10.1093/esr/jcu067

(Article begins on next page)

Camilla Borgna; Dalit Contini. Migrant Achievement Penalties in Western Europe: Do Educational Systems Matter?. EUROPEAN SOCIOLOGICAL REVIEW. 30 (5) pp: 670-683.
DOI: 10.1093/esr/jcu067

Migrant Achievement Penalties in Western Europe:

Do Educational Systems Matter?

Camilla Borgna

Berlin Social Science Center

Dalit Contini

University of Torino

Abstract: This article presents a comparative examination of the educational underachievement of second-generation immigrants in Western Europe near the end of compulsory schooling, based on the 2006–2009 waves of the Programme for International Student Assessment survey. We propose a new measure of migrant educational penalty—revealing the relative position of immigrant students within the achievement distribution of natives with the same socio-economic background—and show that, in most countries, children of immigrants are substantially disadvantaged. We find that the severity of such penalties varies across countries in a way that can neither be reduced to compositional issues, nor equated to educational inequalities driven by socio-economic status. Based on a simple theoretical model of individual student achievement, we detect features of educational systems that might be specifically relevant for the relative disadvantage of immigrant students. By means of recursive partitioning methods, we explore the extent to which these features can explain the cross-country variability in migrant penalties. Our findings suggest that an early inclusion in the educational system may be beneficial for children of immigrants, as countries with high preschool attendance rates or early start of compulsory schooling display mild penalties. Finally, we find that another important institutional aspect is the degree to which second-generation immigrants are marginalized in low-quality schools, in stratified as well as comprehensive educational systems.

1. Introduction

International assessments on students' competencies show that in most European countries migrant educational underachievement is a serious issue. In the last years, public debates have called attention to the need of identifying schooling systems' characteristics able to tackle migrant educational inequalities. Indeed, endowing children of migrants with equal chances to succeed in school compared to their native peers is a major step toward their economic and social integration.

Sociological research on ethnic educational inequality has extensively explored its micro-level determinants (Heath and Brinbaum 2007). Home resources, parental class and qualifications explain the educational disadvantage of migrant children to a significant extent (Kristen and Granato 2007; Van De Werfhorst and Van Tubergen 2007). However, even after accounting for socio-economic background, a residual disadvantage persists (Rothon, 2007), sometimes labeled as ethnic "penalty" (Heath *et al.*, 2008). Comparative works show that educational penalties associated with migrant status differ across European countries (Schnepf 2007), even when same-origin migrants are contrasted (Crul, *et al.*, 2012; Dustmann *et al.*, 2012).

Less clear from previous research is the extent to which educational systems can be called in to explain such cross-country differences. A well established literature in labor economics has identified some institutional features explaining why countries differ in the way socio-economic status affects educational achievement (Hanushek and Woessmann 2011). For instance, the degree of stratification – and in particular age at first tracking – has been consistently found to increase educational inequalities driven by socio-economic background (Hanushek and Woessmann 2011; Van de Werfhorst and Mijs 2010). Like socio-economically disadvantaged families, immigrant families are likely to suffer from a lack of cultural resources relevant to make informed school choices. Hence, early tracking systems might be specifically detrimental to children of migrants. However, empirical evidence suggests that, given prior achievement, students with an immigrant background tend to make more ambitious educational choices with respect to their native peers (Kristen, Reimer, and Kogan, 2008; Cebolla Boado, 2011; Jackson *et al.,* 2012). This positive “secondary effect” (Boudon, 1974) might have different origins, from wishes of upward mobility to anticipation of discrimination on the labor market (Teney, Devleeshouwer, and Hanquinet 2013). Therefore, the role of stratification *per se* in explaining cross-country differences in migrant achievement penalties is not obvious.

More generally, do conventional institutional accounts of socio-economic differentials in school achievement help understanding why children of migrants suffer from more or less severe penalties in different receiving societies? Is migrant-specific disadvantage just another facet of socio-economic disadvantage, or rather are they distinct dimensions of educational inequality?

In this paper, we present a comparative analysis of the relative educational disadvantage of second-generation immigrants in Western Europe based on PISA 2006-2009 surveys on 15-year-old students. Our first aim is to provide new descriptive evidence on migrant-specific penalties in educational achievement across countries. Our second aim is to identify features of school systems that are theoretically relevant for immigrant-background students. Finally, by improving the comparability of receiving countries compared to previous works using international achievement data, we aim at providing new insights on the role of these features in explaining the observed cross-country variability of migrant penalties.

2. Educational institutions and migrant learning disadvantage

Most studies on ethnic educational inequalities focus on micro-level determinants. They consistently show that children of immigrants experience a double disadvantage: first, they underperform natives because they generally have access to fewer socio-economic resources; second, even after accounting for this lack of resources, they suffer from a negative penalty associated with migratory status (Rothon 2007; Heath *et al.*, 2008). This penalty is usually higher for first-generation than second-generation immigrants (Kalter *et al.* 2007), and increases with age at immigration (Ohinata and van Ours 2012).

The role of meso-level determinants of natives' and migrants' achievement – and in particular of school/classroom composition – has been addressed by a number of studies (Cebolla Boado 2007; Brunello and Rocco 2011; Cebolla Boado and Garrido Medina 2011; Agirdag *et al.*, 2012; Contini 2013). Despite the negative correlation between immigrant concentration and achievement, once accounting for socio-economic background and the non-randomness of sorting into schools, the effects become rather small or non-significant. By combining the literature on school factors with that on educational systems, Dronkers *et al*. (2012a, 2012b) use multilevel models to disentangle the direct system-level effects from those mediated by schools.

Cross-country comparisons of the educational performance of children of immigrants are mostly based on international achievement assessments. Fossati (2011), Schneeweis (2011), and Dustmann *et al.*(2012) find no significant differences between immigrant and native students in Britain and non-European countries with a long experience of immigration, while in Nordic and Continental Europe achievement gaps are significant and substantial. Drawing on a unique source of self-collected data, Crul *et al.* (2012) show that migrant educational disadvantage varies across countries even when examining the same minority group, i.e. descendants of Turkish immigrants.

The macro-level determinants of cross-country differences in migrant educational disadvantage have been investigated by some studies, based on two-step cross-country regressions (Schneeweis 2011), or multilevel regression models with individuals nested into countries (Cobb-Clark *et al.* 2012), or individuals nested into schools nested into countries (Fossati 2011; Dronkers *et al.* 2012b). While most of these contributions investigate the capability of educational systems to mitigate the disadvantage of immigrants relatively to natives, others focus on the absolute performance of immigrant students (Dronkers *et al.,* 2012a, 2012b).

These studies' role in raising the attention on the institutional determinants of migrant learning disadvantage is praiseworthy. However, they have important limitations. Firstly, the destination countries analyzed are heterogeneous with respect to their geographical position, developmental level, and societal structure. Moreover, they have different immigration histories, resulting in different immigrant populations. Failing to account for these compositional issues poses a threat to the identification of institutional effects[1]. Secondly, they lack a comprehensive theoretical framework and merely chose institutional variables among those used in the literature on educational inequalities driven by socio-economic background[2]. Thirdly, their regression models are based on restrictive assumptions. For instance, Cobb-Clark *et al.* (2012) force the social-background effect to be constant across countries, irrespective of the empirical evidence that this is stronger in late-tracking countries (Hanushek and Woessmann 2011). Most importantly, these authors do not consider possible interaction effects between the features of educational systems.

From these studies, no clear-cut evidence emerges on the role of educational systems. In particular, the effects of school stratification are unclear. Fossati (2011) and Cobb-Clark *et al.* (2012) find no significant effect of age at tracking, while according to Dronkers *et al.* (2012a) immigrant from favorable socio-economic background do benefit from comprehensive systems, and this effect is partly mediated by school composition. Among the other institutional variables examined, Schneeweis (2011) finds that extended instruction hours, high rates of preprimary enrollment and migrant segregation mitigate migrant underachievement, while social segregation worsens it. Cobb-Clark *et al.* (2012) find larger achievement gaps with higher educational expenditures and teachers' salaries, and reduced gaps if external students' evaluations are applied. On the contrary, they do not find significant effects of the starting age of compulsory schooling.

We contribute to the emerging literature on the macro-level determinants of migrant learning disadvantage in several respects. First, our study is carefully designed in order to address the shortcomings of previous works. We focus on the relative achievement disadvantage of second-generation immigrants, defined as individuals born in the country from both parents born abroad. Those are a rather homogeneous category unlike first-generation immigrants, who differ considerably according to age at immigration. Moreover, they have been fully exposed to the educational system of the receiving country, just like natives: this point is crucial, given our interest on the differential effects of educational systems on immigrants and natives. In order to attain greater comparability of receiving societies, we limit our sample to Western European countries. Beyond the societal and institutional similarities, these countries share a history of post-war labor immigration, as opposed to traditional settlement countries. Even so, immigrant populations across Western Europe are diverse in terms of origin[3]. Hence, with additional analyses, we check the robustness of our findings by contrasting children of Turkish origin only. Yet, this is possible only for destination countries providing information on the birthplace. Hence, in order to account for the remaining composition effects on the whole set of countries, we introduce an aggregate indicator of linguistic distance between origin-country's and destination-country's official languages.

As a second contribution, we define a simple theoretical model of student achievement and derive country-level implications that help formulating theoretically relevant research hypotheses on the features of educational systems likely to affect immigrant-native gaps.

Finally, we empirically evaluate the role of these system-level characteristics in explaining cross-country variability in migrant achievement penalties. Our analytical strategy, based on a two-step approach (Hanushek, and Woessmann 2006, Schneeweis 2011) allows the greatest parameters' flexibility. In a first step, we estimate country-specific individual-level regressions and introduce a new measure of migrant-specific penalty, revealing the *relative* position of immigrant students within the achievement distribution of natives sharing the same socio-economic background. In a second step, we analyze the cross-country variability in migrant penalties. Country-level analyses are based on recursive binary partitioning methods, which we employ to investigate the explanatory role of combinations of theoretically relevant institutions. Our perspective is unambiguously explorative and descriptive. We do not generalize results outside the set of countries under investigation, which is interesting *per se* and should not be thought as a sample drawn from a larger population of comparable units.

3. Institutional effects and research hypotheses

We now reflect on the micro-foundations of educational inequalities for children of immigrants, with the aim to identify features of educational systems potentially relevant for migrant-specific penalties. Firstly, we develop a theoretical model of the mechanisms affecting achievement in primary and secondary school. We then formalize the model and derive the country-level implications for migrant-native differentials. Finally, we identify the institutional aspects specifically relevant for the relative disadvantage of immigrant students, formulate and motivate our research hypotheses, and relate them to the existing literature.

Theoretical individual achievement model and country-level implications

Consider the simple model of individual achievement depicted in Figure 1.

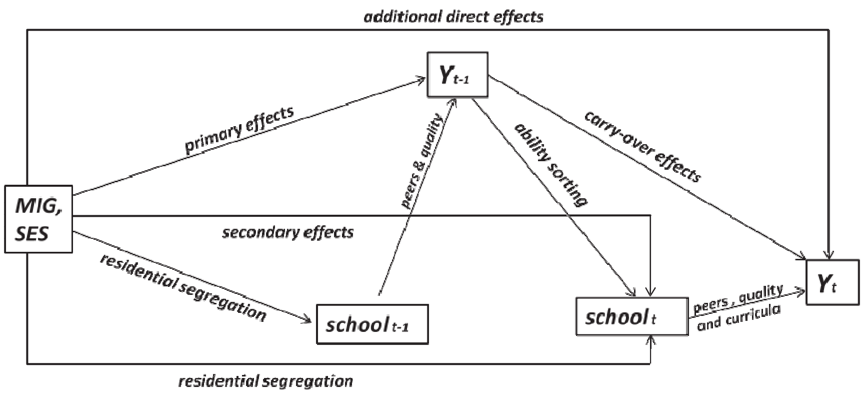


Figure 1. Theoretical individual-level model (own elaboration from Contini and Grand (2013) and Esser (2014))

Family background (migratory status and *SES*) directly influence primary school achievement (Y_{t-1}). Where residential segregation is pronounced and peer effects operate, family background also has an indirect effect through the school composition at time *t*-1. If primary schooling is not nationally standardized, school-quality effects may add on. Secondary school achievement (Y_t) depends on family background mainly via previous achievement and via the current school characteristics at time *t*. Schools are highly differentiated in stratified systems, as curricula and instruction levels differ between tracks, and there is often explicit ability sorting. Yet, there may be substantial variability also in comprehensive systems, due to residential segregation and/or disparities in resources allocation. In differentiated systems, school sorting also directly depends on family background (secondary effects).

In this framework, a stylized theoretical individual-level model for achievement at time *t* is:

$$y_{it} = \mu + \beta_1 SES_i + \beta_2 MIG_i + \gamma y_{i,t-1} + \delta s_{it} + \varepsilon_{it} \qquad (1)$$

where s_t are all the relevant secondary-school characteristics. β_1 and β_2 represent the additional direct effects of *SES* and migratory background between *t*-1 and *t*, net of previous ability and school factors. γ are carry-over effects of previous performance.

Note that in the "ideal" case with no residential segregation, no tracking, and full standardization, school factors are only marginally relevant, as they vary only due to random allocation of students and teachers.

Let us focus now on the country-level implications of individual model (1). The migrant/native gap at time *t* for a given *SES* in a specific country depends on the average previous achievement gap and on the average difference in school factors between migrants and natives:

$$\begin{aligned}E(y_t|MIG = 1, SES) - E(y_t|MIG = 0, SES) \\ = \beta_2 + \beta_1 SES + \gamma[E(y_{t-1}|MIG = 1, SES) - E(y_{t-1}|MIG = 0, SES)] \\ +\delta[E(s_t|MIG = 1, SES) - E(s_t|MIG = 0, SES)]\end{aligned}$$

This expression helps detecting system-level features relevant for migrant-specific inequalities for secondary school students. The component related to previous achievement represents the migrant-specific disadvantage up to time *t*-1. This gap should be influenced by the characteristics of primary and pre-primary schooling. Given the little institutional differentiation of primary schooling in Europe, we identify entry age in the (pre)schooling system as the main potentially relevant institutional factor affecting early achievement differentials. The component due to secondary schooling – school composition, curricula, resources, instruction quality – varies according to the degree of residential segregation, non-standardization and formal or informal tracking, and to the different allocation of migrants and natives in schools with different characteristics[4].

For further details, refer to the Appendix 1 in the Supplementary materials.

Research hypotheses

Our first research hypothesis is that *early entry in the (pre)school system helps contrasting migrant underachievement*, thereby reducing migrant penalties. As argued above, pre-primary and primary schooling can affect migrant-specific penalties in later age. Entry age in (pre)school can be crucial, as children's lives, previously fully spent within families and communities, become exposed to the surrounding society. While this is relevant transition for all children, for immigrant children in particular it may represent the first occasion for systematic interactions with natives. Entry age is affected by the start of compulsory education and by preschool participation.

Preschool attendance has been found to have a positive effect on cognitive development, especially for disadvantaged children (Carneiro and Heckman 2003; Magnuson *et al.*, 2006; Felfe and Hsin 2012). We expect it to be even more beneficial for children of immigrants because it provides a context to improve their linguistic skills in the destination-country language (Christensen and Stanat 2007). Moreover, early socialization with natives could reduce cultural distance and the lack of information experienced by their families (Schofield 2006). Empirical evidence, though limited to some studies on the US and Germany, suggests that preschool attendance boosts educational opportunities for first- and second-generation immigrants (Spiess *et al.*, 2003; Crosnoe 2007; Biedinger *et al.*, 2008).

Similar mechanisms could also apply to an early start of compulsory schooling; a formal educational context provides additional learning opportunities for children lacking the cultural capital specific to the destination country, as it is often the case for children of migrants. The age of compulsory education is relevant also because it determines the time when *all* children – regardless of their characteristics – are in school.

Secondly, we consider the extent to which children of immigrants are (or are not) marginalized in low-quality sectors of secondary school systems. Our second research hypothesis is that *high levels of marginalization widen migrant-native gaps*. Peer effects are a potential driving mechanism. The child's own performance is likely to be influenced by the performance of peers (Manski 1993) as teachers may adjust performance targets and lower the instruction level. A student's achievement could also be directly influenced by others' achievement: while good students may contribute establishing a positive competition climate, weak students may lose motivation and negatively affect peers' attitudes towards learning. An additional mechanism underlying the effect of marginalization is related to teaching quality. In systems where no specific incentives are given to foster the provision of high-quality teachers and additional resources to schools with low-performing children, the achievement of immigrant students may be further harmed. Highly qualified teachers have incentives and means to leave troublesome schools (Wyckoff and Boyd 2005). Evidence of lower quality teachers provided to schools with disadvantaged children is available for some countries (Barbieri *et al.*, 2010 on Italy; Bonesrønning *et al.*, 2005 on Norway; Schindler Rangvid, 2007 on Denmark). However, the latter also reports that immigrant students in Denmark tend to experience more favorable class sizes and teacher-student ratios.

But how is marginalization produced? We argue that major determinants are the degree of differentiation of educational programs and school quality on the one side, and the sorting process of children into schools and programs on the other. Such processes are related to the standardization and stratification of the school system and to school segregation. Standardization is the degree to which the quality of education meets the same targets nationwide: curricula, school-leaving examinations, teachers' training and financial resources may differ between areas of the country or from school to school (Allmendinger 1989). Stratification is the structural differentiation of the school system within given educational levels (*ivi*). Tracking into academic or vocational education – in Europe occurring between age 10 and 16 – is the most relevant form of stratification. Both low standardization and stratification yield to differentiated schooling systems.

In differentiated schooling systems, the sorting of children into schools and educational programs is (at least to some extent) driven by ability. Immigrant and socio-economically disadvantaged children generally display lower prior achievement. Hence, due to "primary effects" (Boudon 1974), immigrant children are likely to end up in schools with larger shares of low-performing children. However, family background can also have direct or "secondary" effects (*ivi*) on track placement, since educational choices depend on strategic information and cultural capital. Empirical evidence indicates that, while low *SES* drives negative secondary effects (Jackson 2013), immigrant status drives positive ones. Indeed, given prior achievement, immigrant-background students usually make more ambitious educational choices (Kristen *et al.*, 2008; Cebolla Boado 2011; Jackson *et al.*, 2012).

Another possible path to marginalization derives from social and/or immigrant school segregation, which might exist even in undifferentiated systems, since it can be triggered by residential segregation in poor districts or by discriminatory school-enrollment policies. Given the poorer average achievement of disadvantaged students, school segregation is likely to lead to a disproportionate concentration of immigrants in schools with low-performing peers.

Ultimately, we argue that standardization, stratification and school segregation become an issue for second-generation immigrants if they relegate them into *marginal* sectors of the schooling system, with low-quality teaching, low-performance targets and low-performing peers.

The institutional features discussed above could be particularly detrimental if the mother tongue of most immigrants is very different from that spoken in the destination country. Since the linguistic composition of immigrants varies across Western European countries, linguistic distance is an important contextual element to account for. Our final research hypothesis is that *high linguistic distance worsens migrant penalties*, particularly in late-entry systems.

4. Measuring migrant achievement penalties

Migrant underachievement is often measured as the average gap of migrants with respect to natives. We propose an alternative measure which reveals the average position of immigrant children (*M*) into the distribution of their native peers (*N*), expressed in terms of standard deviations. We define the "raw *z*-score" as:

$$\bar{z}_M = \frac{1}{n}\sum_i z_{i,M} = \frac{1}{n}\sum_i \frac{y_{i,M} - \bar{y}_N}{\hat{\sigma}_N} = \frac{\bar{y}_M - \bar{y}_N}{\hat{\sigma}_N}$$

Its interpretation is straightforward. A score of –0.5 implies that if the average migrant was placed into the natives' distribution, she would score 0.5 st.dev. below the mean. Assuming normality, this corresponds to the 31st percentile. The *z*-score metrics – as opposed to the average-gap metrics – also considers the existing variability in the receiving societies. For a given migrant-native gap in PISA scores, lack of immigrant integration is more severe in a society with less heterogeneity among native children, as the average immigrant child performs not only worse than the average native, but also worse than low-performing ones.

To isolate the migrant-specific disadvantage, we account for compositional effects due to socio-economic endowments *X* and use a modified version of the *z*-score, revealing the average position of second-generation migrants in the distribution of natives sharing their socio-economic status. This index – emphasizing the relative rather than absolute distance between scores of natives and migrants – is our measure of *migrant achievement penalty*. The "controlled z-score" is defined as:

$$\bar{z}_M^x = \frac{1}{n}\sum_i \frac{y_{i,M,x} - \bar{y}_{N|x}}{\hat{\sigma}_{N|x}} = \sum_x \bar{z}_{M|x}\, p_{M|x}$$

where n is the number of migrants, $\bar{z}_{M|x}$ the z-score for given x, and $p_{M|x}$ the proportion of migrants with X=x. Instead of evaluating $\bar{z}_M^x$ completely non parametrically, we refer to a simple model of performance Y, where we allow for differential returns to socio-economic resources for immigrants and natives:

$$Y_i = \alpha_N + \beta_N X_i + (\alpha_M - \alpha_N)MIG_i + (\beta_M - \beta_N)X_i\, MIG_i + \varepsilon_i \qquad (2)$$

MIG is a dummy indexing migrant background, α_M and α_N the intercepts for migrants and natives respectively, β_M and β_N the effects of socio-economic status. Thus:

$$\bar{z}_M^x = \frac{\sum_x[(\hat{\alpha}_M+\hat{\beta}_M x)-(\hat{\alpha}_N+\hat{\beta}_N x)]p_{M|x}}{\hat{\sigma}_\varepsilon} = \frac{\sum_x(\hat{\alpha}_M-\hat{\alpha}_N)+(\hat{\beta}_M-\hat{\beta}_N)\bar{x}_M}{\hat{\sigma}_\varepsilon} \qquad (3)$$

For further details, refer to Appendix 2 in the Supplementary Material.

5. Data and variable construction

Data

Analyses are based on representative data from the *Programme for International Student Assessment* (PISA) collected in the years 2006-2009[5]. PISA assesses 15-year-old students' competences in reading, mathematics, and science. Test scores are standardized on a common scale allowing direct cross-country comparisons. Individual, family and school background information is collected through questionnaires administered to students and school officials. PISA samples are derived from a two-stage stratified sampling procedure with schools selected in the first stage and students in the second one[6].

Our sample units are 15-year-old students in 17 Western European countries. Since mathematics literacy is less influenced by lack of linguistic skills than reading and science, we use the former as the educational outcome of main interest. This choice has the advantage of limiting compositional effects due to the origin country. Nonetheless, to gain leverage, we replicate analyses on reading and science.

All descriptive statistics are reported in Appendix 3 (Supplementary Materials).

Individual-level variables

Migrant categories are defined according to the information on birthplace provided by PISA: second-generation immigrants are native-born students with both foreign-born parents. Natives are students with at least one native-born parent. First-generation immigrants (foreign-born students with foreign-born parents) are excluded from the analyses. To operationalize the various background dimensions potentially affecting educational achievement, we used a synthetic measure provided by PISA: the index of economic, social and cultural status (*ESCS*), derived from the highest occupational status of parents, the highest parental education and home possessions (family wealth, cultural possessions, educational resources, number of books at home). For further details on this choice, refer to Appendix 4b in the Supplementary materials.

Country-level variables

Entry in the (pre)school system. We combine information on the starting age of compulsory schooling and the share of four-year-old children attending preschool in 1994-95 (Eurydice 1997, 34) and in 1996-97 (Eurydice 2000, 43)[7]. We standardize these measures and add them together; thus, the distance between countries is simply rescaled, and equal importance is given to the two[8].

Marginalization in schools with low-performing children. Evaluated as the relative risk for second-generation migrants (*vs.* natives) of attending “bad” schools – schools in the lowest-performing group, situated in the 10th percentile of the achievement distribution according to PISA average scores over all literacy domains[9].

Linguistic distance between origin- and destination-countries’ official languages: computed as follows: (1) assessment of linguistic distance between each origin-country language and the destination-country language; (2) aggregation of linguistic distance into a country index by shares of immigrant groups. For more details, refer to the Appendix 4.d in the Supplementary Materials.

6. Migrant achievement penalties in Western Europe

In order to compute the controlled *z*-score as a measure of migrant educational penalty, we run country-specific individual-level regressions on mathematics score[10] over migratory status, gender, *ESCS* and – where significant– an interaction between *ESCS* and migratory status[11]. These estimates – used to compute the z-scores – are reported in Table A5 in the Supplementary Materials[12].

As shown in Figure 2, raw *z*-scores provide a clear-cut picture of the issue of migrant underachievement in Western Europe: in most countries, second-generation immigrants lie below the 30th percentile of the distribution of natives, despite being born in the receiving society and having been fully exposed to its educational system. In Belgium-Flanders and Denmark the situation is critical: the average second-generation migrant lies around the 20th percentile of the natives’ distribution.

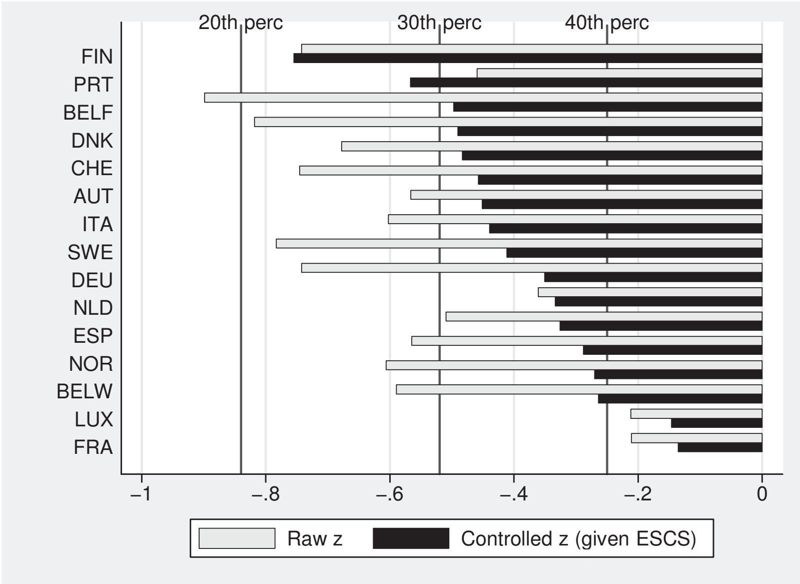


Figure 2. Overall underachievement and migrant achievement penalty for second-generation migrants. Source: PISA 2006– 2009, mathematical literacy (estimates with plausible values).

Sharp cross-country differences exist, not only in the levels of general underachievement, but also in the extent to which they are explained by socio-economic resources differentials. In the Netherlands, Luxembourg, and France, underachievement is more than halved when *SES* differences are accounted for, while in Finland, Portugal, Italy, and Spain more than 75% remains unexplained. Our measure of migrant educational penalty – the controlled z – reveals that in ten countries the average second-generation migrant child lies below the 35th percentile of the distribution of natives with the same socio-economic resources.

Are the countries with the greatest migrant-specific penalties also unequal with respect to socio-economic background? Figure 3-left displays how Western European countries perform with respect to these two dimensions of educational inequality[13].

Not only migrant-specific penalties do not coincide with socio-economic penalties, but in our sample of countries they are negatively correlated. This finding could be interpreted as signaling a policy tradeoff. This explanation was invoked by Fossati (2011), to account for her finding that where income dispersion is low – especially in Scandinavian countries –

immigrant children perform poorly. However, the correlation between migrant-specific and socio-economic penalties might also be spurious. In both cases, this evidence suggests that it is essential to reflect on the features of educational systems that may be specifically beneficial or detrimental to immigrant-background students.

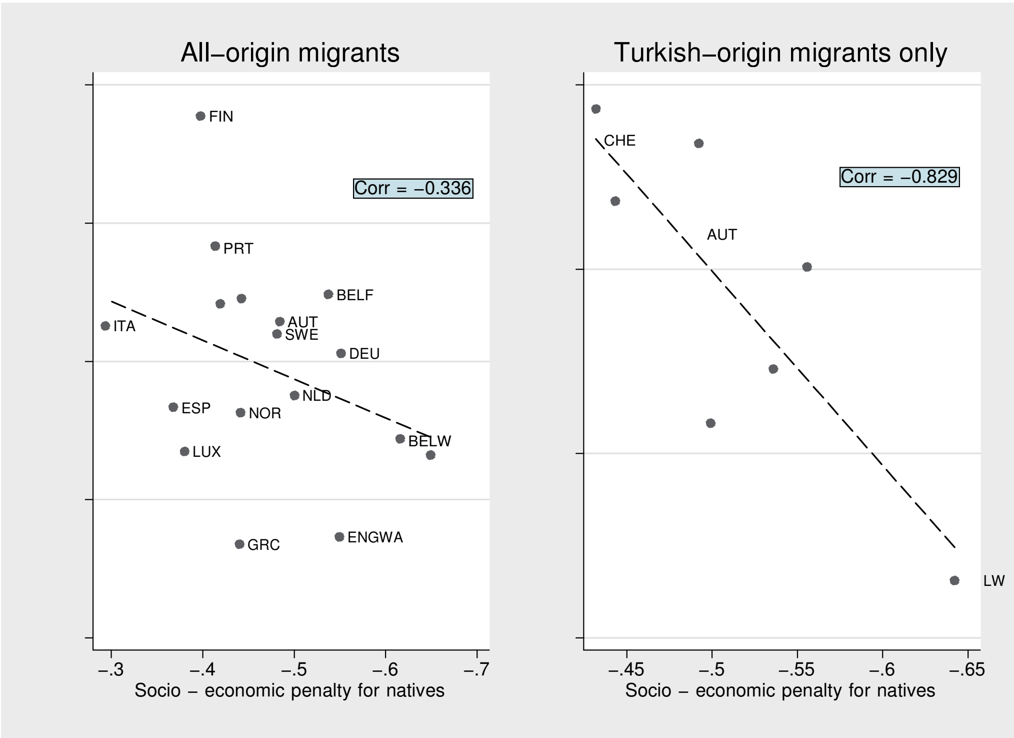


Figure 3. Overall underachievement and migrant achievement penalty for second-generation migrants. Source: PISA 2006– 2009, mathematical literacy (estimates with plausible values)

Cross-country variability of migrant penalties could be driven by compositional issues. This is why – following Crul *et al.* (2012) – we contrast students with Turkish immigrant mothers across countries providing information on parental birthplace and where samples are large enough. Figure 3-right shows that the ranking of most countries is unchanged. The inverse relationship with socio-economic penalties holds true, and appears even stronger.

To summarize, cross-country differences in migrant achievement penalties exist and cannot be reduced to compositional issues. Therefore, there is room for characteristics of educational systems to explain such variability. Moreover, migrant penalties and socio-economic penalties emerge as two distinct dimensions of educational inequalities. This provides empirical support to our idea that features of educational systems affecting migrant-specific educational disadvantage cannot be merely derived from those affecting class-driven educational inequalities.

7. Evidence on institutional effects

Can the dimensions of educational systems that we identified as theoretically relevant account for the cross-country differences in migrant achievement penalties? To address this question, we use regression–tree analysis. This is a multivariate data technique that recursively partitions the data space into smaller regions, according to the one binary question which minimizes the sum S of squared deviations from the subgroup means in the response variable. Each parent node is further divided into child nodes, and the procedure is repeated until the largest decrease in S falls below a given complexity threshold. Regression trees are particularly useful to detect complex interaction patterns, for which we have no *a priori* assumptions.

Figure 4 depicts results of the regression-tree analysis. As a guidance to the interpretation of the tree, note that variables with the best predictive power are those generating splits at the higher-level nodes and emerging again in subsequent divides, while those appearing for the first time in lower-level nodes are usually less important.

With the exception of Portugal in a fourth-level split, all partitions are consistent with our theoretical predictions: earlier entry is associated with milder migrant penalties, while high marginalization is generally associated with more severe penalties.

In the first step, countries are split according to the *(pre)school-system entry.* In Finland – the country with the largest migrant penalties – children enter the system particularly late. The remaining countries are differentiated according to the extent they *marginalize* second-generation immigrants in “bad” schools. Low-marginalizing countries generally display milder penalties than highly-marginalizing ones. Both groups are further split according to system entry and – again – marginalization, while linguistic distance does not emerge as a discriminating factor[14]. Aside from Finland, the most unequal systems for immigrant students combine late entry and high marginalization (Austria, Switzerland, Sweden, Denmark, and Portugal). On the contrary, systems with mild penalties have very little marginalization (Greece), very early entry (England and Wales) or display a combination of these two elements (Spain, France, Luxembourg).

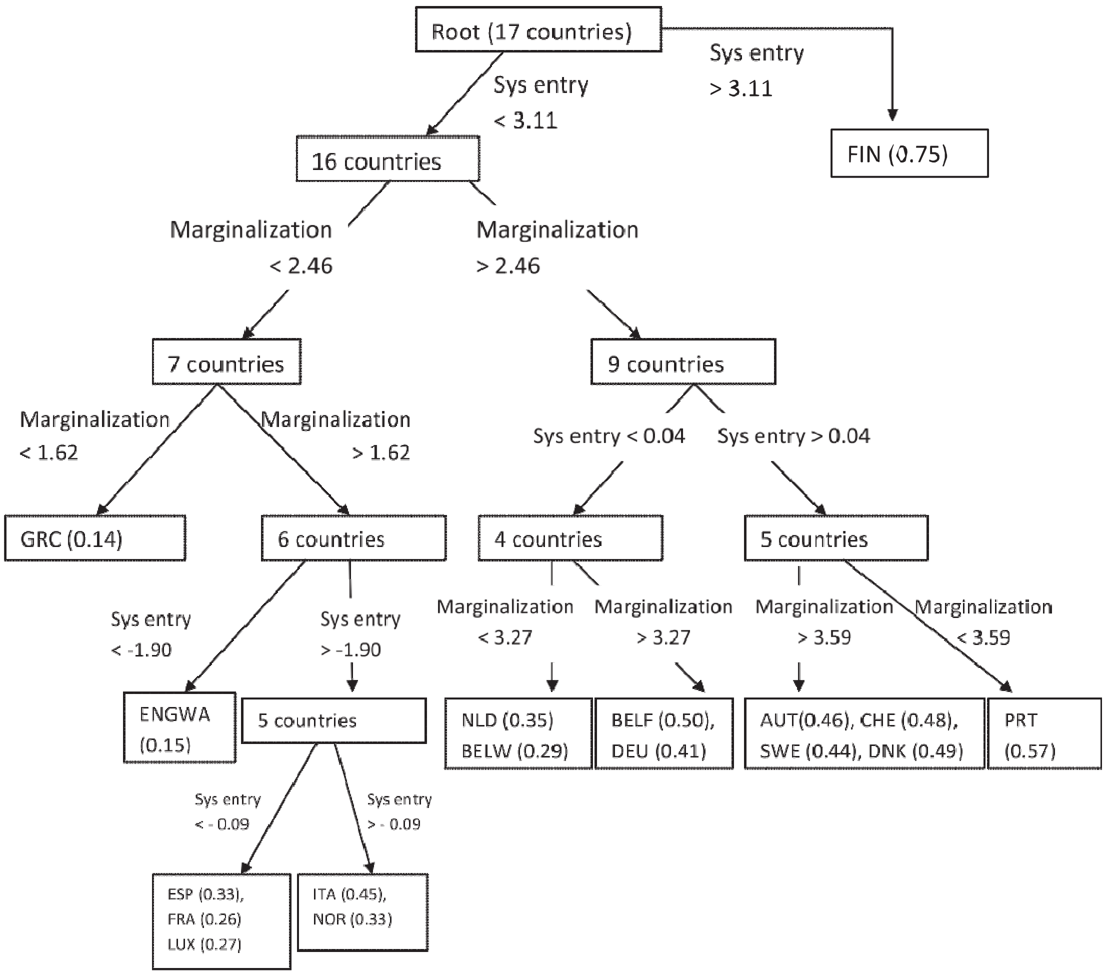


Figure 4 Results of regression tree analysis. Response variable: migrant-specific penalties in math achievement (absolute values reported in parentheses). Explanatory variables: marginalization, (pre)school system entry, linguistic distance. Analyses performed with package R 'rpart'. Method: 'ANOVA', complexity parameter 0.01.

8. Discussion and conclusions

In this paper, we explored migrant educational disadvantage in Western Europe. By using the 2006-2009 waves of the PISA survey on 15-year-old students, we provided new descriptive evidence on the relative educational disadvantage of second-generation immigrants. Our proposed measure clearly showed that migrant-specific penalties are severe in most Western European countries: in ten countries, the average second-generation immigrant lies below the 35^{th} percentile of the mathematics achievement distribution of natives with the same socio-economic background. Similar results are found for reading and science. Moreover, countries vary in the extent to which migratory status affects students' achievement. Cross-country differences cannot be reduced to the different origin-composition of immigrant populations, as we show with additional analyses on Turkish second-generation immigrants. Hence, educational systems may play a role in explaining why the relative performance of second-generation immigrants varies across Western European countries.

From our empirical analyses, migrant-specific and socio-economic penalties manifestly came forth as two distinct dimensions of educational inequalities. With an in-depth theoretical reflection on the institutional features specifically relevant for second-generation immigrant students, we contributed to the comparative literature on migrant educational disadvantage, so far mainly focused on the institutional features related to class-driven educational inequalities. In particular, we consider the moment when children enter (pre)school and the degree to which second-generation immigrants are marginalized in "bad" schools. With recursive partitioning methods, we then investigated how such theoretically-relevant features of educational systems combine with each other and with linguistic distance in producing more or less severe migrant penalties.

Our exploratory analyses indicate that the degree to which second-generation immigrants are relegated in marginal sectors of the school system is crucial to explain cross-country differences in migrant achievement penalties. Such marginalization can be produced by school stratification, but also by segregation, or by a lack of national standardization. In this sense, it is not surprising that previous studies failed to find significant effects of age at tracking on migrant learning disadvantage (Cobb-Clark *et al.*, 2012; Dronkers *et al.*, 2012a; Fossati, 2012). Our findings show that early tracking is associated with severe penalties only when coupled with high marginalization (Austria, Belgium-Flanders, Switzerland, and Germany). Moreover, penalties can be severe also where marginalization occurs despite late tracking (Sweden, Denmark, and Portugal). Finally, we found that high preschool attendance rates matter, as suggested by Schneeweis (2011), but are by no mean sufficient to avoid severe penalties. Formal-schooling start is an equally important element, as exemplified by the cases of England and Wales and the Netherlands, where schooling is already compulsory at age five. Most importantly, a delayed entry into (pre)school generally becomes an issue when coupled with other problematic aspects of the educational system. With the exception of Finland – where children enter school at age seven and preschool is limited – countries displaying severe penalties are also highly marginalizing. At the same time, countries where migrant penalties are mild generally cumulate several beneficial institutional factors. Contrary to our expectations, linguistic distance does not emerge as a relevant explaining factor of cross-country variability in migrant penalties for mathematics. Nevertheless, countries with high linguistic distance display larger penalties in reading and science.

We conclude with a note on the limitations of our work. Cross-country investigations of policy effects most often have an explorative and descriptive character and do not allow to make causal inference: the *ceteris paribus* assumption usually does not hold. However, given the limited institutional variability existing within national boundaries, international data are an essential tool to explore the role of educational systems in reducing inequalities. When examining migrant disadvantage, the heterogeneity of the immigrant populations is an additional obstacle. Despite our efforts to take this compositional issue seriously, immigrant populations might still not be completely comparable across countries. Additional research with more focused research designs (e.g. exploiting national reforms) and richer data (longitudinal, and with detailed information on the immigrant population) is needed to deepen the understanding of the causal impact of specific aspects of educational systems on the relative disadvantage of immigrant students in different receiving societies.

List of references

Agirdag, Orhan, Mieke Van Houtte, and Piet Van Avermaet (2012). Why Does the Ethnic and Socio-Economic Composition of Schools Influence Math Achievement? The Role of Sense of Futility and Futility Culture. *European Sociological Review* **28 (3)**, 366–378.

Allmendinger, Jutta (1989). Educational Systems and Labor Market Outcomes. *European Sociological Review* **5 (3)**, 231–50.

Barbieri, Gianna, Claudia Rossetti, and Paolo Sestito (2010). The determinants of teacher mobility. Evidence from a panel of Italian teachers." *Bank of Italy Working Paper* **761**.

Biedinger, Nicole, Birgit Becker, and Inge Rohling (2008). Early Ethnic Educational Inequality: The Influence of Duration of Preschool Attendance and Social Composition. *European Sociological Review* **24 (2)**, 243–256.

Bonesrønning, Hans, Torberg Falch, and J. Strom (2005). Teacher Sorting, Teacher Quality, and Student Composition. *European Economic Review* **49**.

Boudon, Raymond (1974). *Education, Opportunity and Social Inequality*. New York: Wiley.

Brunello, Giorgio, and Lorenzo Rocco (2011). The Effect of Immigration on the School Performance of Natives: Cross Country Evidence Using PISA Test Scores. *IZA Discussion Paper* **5479**.

Carneiro, Pedro, and James J. Heckman (2003). Human Capital Policy. *IZA Discussion Paper* **821**.

Cebolla Boado, Hector (2007). Immigrant Concentration in Schools: Peer Pressures in Place? *European Sociological Review* **23 (3)**, 341–356.

——— (2011). Primary and Secondary Effects in the Explanation of Disadvantage in Education: The Children of Immigrant Families in France. *British Journal of Sociology of Education* **32 (3)**, 407–430.

Cebolla Boado, Hector, and Luis Garrido Medina (2011). The Impact of Immigrant Concentration in Spanish Schools: School, Class, and Composition Effects. *European Sociological Review* **27 (5)**, 606– 623.

Christensen, Gayle, and Petra Stanat (2007). *Language Policies and Practices for Helping Immigrants and Second-Generation Students Succeed*. Migration Policy Institute and Bertelsmann Stiftung.

——— (2012). Migrant Youths' Educational Achievement The Role of Institutions. *The ANNALS of the American Academy of Political and Social Science* **643 (1)**, 18–45.

Contini, Dalit (2013). Immigrant Background Peer Effects in Italian Schools. *Social Science Research* **42 (4),** 1122–1142.

Contini, Dalit, and Elisa Grand (2013). On estimating achievement dynamic models from repeated cross-sections. *Department of Economics and Statistics "Cognetti de Martiis" Working Paper* **43/13**.

Crosnoe, Robert (2007). Early Child Care and the School Readiness of Children from Mexican Immigrant Families. *International Migration Review* **41 (1)**, 152–181.

Crul, Maurice, Jens Schneider, and Frans Lelie (Eds.). 2012. *The European Second Generation Compared. Does the Integration Context Matter?* Amsterdam: Amsterdam University Press.

Dronkers, Jaap, Manon de Heus, and Mark Levels (2012a). Immigrant Pupils' Scientific Performance: The Influence of Educational System Features of Origin and Destination Countries. *CReAM Discussion Paper* **1212**.

Dronkers, Jaap, Rolf van der Velden, and Allison Dunne (2012b). Why Are Migrant Students Better Off in Certain Types of Educational Systems or Schools than in Others? *European Educational Research Journal* **11 (1)**, 11–44.

Dustmann, Christian, Tommaso Frattini, and Gianandrea Lanzara (2012). Educational Achievement of Second-Generation Immigrants: An International Comparison. *Economic Policy* **27 (69)**, 143–185.

Esser, Hartmut (2014). *Sorting and School Effects. The Standard-Approach and the DVD-Puzzle on Effects of Education Systems on Educational Inequality*. AMCIS Conference presentation, February 13-14, Amsterdam.

Eurydice (1997). *Key Data on Education in Europe*. European Commission.

——— (2000). *Key Data on Education in Europe*. European Commission.

Fekjær, Silje Noack, and Gunn Elisabeth Birkelund (2007). Does the Ethnic Composition of Upper Secondary Schools Influence Educational Achievement and Attainment? A Multilevel Analysis of the Norwegian Case. *European Sociological Review* **23 (3)**, 309–323.

Felfe, Christina, and Amy Hsin (2012). Maternal Work Conditions and Child Development. *Economics of Education Review* **31 (6)**, 1037–1057.

Fossati, Flavia (2011). The Effect of Integration and Social Democratic Welfare States on Immigrants' Educational Attainment: A Multilevel Estimate. *Journal of European Social Policy* **21 (5)**, 391–412.

Hanushek, Eric A., and Ludger Woessmann (2011). The Economics of International Differences in Educational Achievement. In *Handbook of the Economics of Education*, by Eric Hanushek, Stephen Machin, and Ludger Woessmann (Eds.), 89–200. Oxford: Elsevier.

Hanushek, Eric A., and Ludger Woessmann (2006). Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries. *The Economic Journal* **116(150)**, C63-C76

Heath, Anthony, and Yaël Brinbaum (2007). Guest Editorial Explaining Ethnic Inequalities in Educational Attainment. *Ethnicities* **7 (3)**, 291–304.

Heath, Anthony, Catherine Rothon, and Elina Kilpi (2008). The Second Generation in Western Europe: Education, Unemployment, and Occupational Attainment. *Annual Review of Sociology* **34 (1)**, 211–235.

Jackson, Michelle, Jan O. Jonsson, and Frida Rudolphi (2012). Ethnic Inequality in Choice-Driven Education Systems A Longitudinal Study of Performance and Choice in England and Sweden. *Sociology of Education* **85 (2)**, 158–178.

Jackson, Michelle (Ed.) (2013). *Determined to Succeed? Performance versus Choice in Educational Attainment*. Stanford: Stanford University Press.

Kalter, Frank, Nadia Granato, Anthony Heath, and Sin Yi Cheung (2007). Educational Hurdles on the Way to Structural Assimilation in Germany. In *Unequal Chances. Ethnic Minorities in Western Labour Markets*, by Anthony F. Heath and Sin Yi Cheung (Eds.), 271–319. Oxford: Oxford University Press.

Kristen, Cornelia, and Nadia Granato (2007) The Educational Attainment of the Second Generation in Germany Social Origins and Ethnic Inequality. *Ethnicities* **7 (3)**, 343–366.

Kristen, Cornelia, David Reimer, and Irena Kogan (2008). Higher Education Entry of Turkish Immigrant Youth in Germany. *International Journal of Comparative Sociology* **49 (2-3)**: 127–151.

Magnuson, Katherine, Claudia Lahaie, and Jane Waldfogel (2006). Preschool and School Readiness of Children of Immigrants. *Social Science Quarterly* **87 (5)**, 1241–1262.

Manski, Charles F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies* **60 (3)**: 531–542.

Ohinata, Asako, and Jan C. van Ours (2012). Young Immigrant Children and Their Educational Attainment." *Economics Letters* **116 (3),** 288–290.

Rothon, Catherine (2007). Can Achievement Differentials Be Explained by Social Class Alone? An Examination of Minority Ethnic Educational Performance in England and Wales at the End of Compulsory Schooling. *Ethnicities* **7** (**3**), 306–322.

Schindler Rangvid, Beatrice (2007). Sources of Immigrants' Underachievement: Results from PISA—Copenhagen. *Education Economics* **15 (3),** 293–326.

Schneeweis, Nicole (2011). Educational Institutions and the Integration of Migrants. *Journal of Population Economics* **24 (4)**, 1281–1308.

Schnepf, Sylke. (2007). Immigrants' Educational Disadvantage: An Examination across Ten Countries and Three Surveys. *Journal of Population Economics* **20 (3)**, 527–545.

Schofield, Janet W. (2006). Migration Background, Minority-Group Membership and Academic Achievement Research Evidence from Social, Educational, and Developmental Psychology. *AKI-WZB Research Review* **5**.

Spiess, C.Katharina, Felix Büchel, and Gert G Wagner (2003). Children's School Placement in Germany: Does Kindergarten Attendance Matter? *Early Childhood Research Quarterly* **18 (2)**, 255–270.

Teney, Celine, Perrine Devleeshouwer, and Laurie Hanquinet (2013). Educational Aspirations among Ethnic Minority Youth in Brussels: Does the Perception of Ethnic Discrimination in the Labour Market Matter? A Mixed-Method Approach. *Ethnicities* **13(5)** 584–606.

Van De Werfhorst, Herman G, and Frank Van Tubergen (2007). Ethnicity, Schooling, and Merit in the Netherlands. *Ethnicities* **7 (3)**, 416–444.

Van de Werfhorst, Herman G., and Jonathan J.B. Mijs (2010). Achievement Inequality and the Institutional Structure of Educational Systems: A Comparative Perspective. *Annual Review of Sociology* **36 (1)**, 407–428.

Wyckoff, James, and Donald Boyd. (2005). Explaining the Short Careers of High-Achieving Teachers in Schools with Low-Performing Students. *American Economic Review* **95 (2)**, 166–171.

Endnotes

[1] Dronkers and colleagues include origin-country characteristics. Yet, since this information is unavailable for many countries, they must analyze a reduced but still heterogeneous set of countries. Schneeweis (2011) includes controls for origin macro-region, but this is possible only for few country-years.

[2] Typical institutions used to explain *SES*-driven inequalities are public vs. private provision, tracking age, central examinations, school autonomy, teachers' salary, expenditures, pupil/teacher ratio, preschool enrollment.

[3] Immigrant populations in Europe also differ for motives to migrate. Most immigrants have economic or family reasons, but in Sweden, Norway, and Switzerland several are asylum seekers, especially since the 1990s. However, this compositional problem could be less relevant for second-generation immigrants in PISA 2006-2009, as their parents most probably immigrated in the 1970-80s.

[4] Our analytical strategy is fully described in the following sections. However, in the light of the above argument, it is useful to acknowledge here an important data limitation. Given the cross-sectional nature of PISA, we observe achievement at age 15, while we do not observe individual achievement growth. Hence, we cannot disentangle inequalities developed in primary school from those developed in secondary school. Longitudinal data would allow to empirically relating the former with features of the educational system at the primary level, and the latter to the secondary school system. Our strategy, instead, consists in analyzing the cross-country variability of overall inequalities developed *up to* age 15 to system-level features of primary *and* secondary school (see also Appendix 1.2 in the Supplementary Material).

[5] We first performed separate analyses on waves 2006 and 2009. Since results proved consistent, we rerun the analyses on the pooled waves to ensure greater sample sizes for immigrant students.

[6] To account for this complex sampling structure, we used the final sampling weights and the 80 replicate sampling weights. To obtain unbiased estimates of the standard errors, we also used the five plausible values for students' proficiency, as recommended by PISA (OECD 2009: 129).

[7] Due to comparability issues of preschool systems at early ages, the rate at age 3 is unavailable for some countries. Children in our sample were age 4 in 1995 and 1998.

[8] Shares by immigrant status are not available. However, according to the information on preschool attendance provided by PISA, participation rates of natives and second-generation immigrants are similar in all countries.

[9] If children are randomly allocated to schools, the chances of ending up in bad schools are the same for all children and the index is nil. The choice of the 10th percentile to define "bad" schools is arbitrary; however, the use of the 20th percentile produces very similar results (available upon request).

[10] Results of analyses on reading and science literacy are reported in Table A6 in the Supplementary materials and are generally consistent with those on mathematics.

[11] Unlike previous works on migrant educational disadvantage based on PISA data, we do not control for language spoken at home, because it is endogenous to our dependent variable. Refer to the Appendix 4c in the Supplementary materials for further details.

[12] Consistently with previous analyses (Fekjær and Birkelund 2007; Kristen and Granato 2007; Schneeweis 2011), whenever the interaction term is significant, it is generally negative, meaning that migrants benefit less

than natives from socio-economic resources. However, interaction terms are rather small with respect to the migrant-dummy coefficients, indicating that underachievement of second-generation immigrants is not so much driven by differential returns to parental socio-economic resources, but rather by the lack of other resources (e.g. linguistic, cultural, relational).

[13] The "socio-economic penalty" measures by how many standard deviations a native individual with *ESCS*=*x*-1 lags behind the native individual with *ESCS*=*x*. This measure is computed only on natives to avoid compositional issues due to the lower socio-economic status of immigrants.

[14] However, linguistic distance is relevant to differentiate countries according to migrant penalties in the other literacy domains, as appears from additional analyses performed on reading and science (see Figures A6a-b in the Supplementary materials). Results are also consistent with our theoretical predictions on system entry and marginalization.

Migrant achievement penalties in Western Europe. Do educational systems matter?

Camilla Borgna and Dalit Contini

Supplementary Material

Appendix 1: Additional information on the analytical strategy

1.1 Theoretical individual-level model

Refer to Figure A.1 (and to Figure 1 and Section 3 in the main document). A stylized individual-level model for secondary school achievement Y_t in a given country is:

$$y_t = \mu + \beta x + \gamma y_{t-1} + \delta s_t + \varepsilon_t \tag{1}$$

where x are individual-level factors related to family background (SES and migrant status) and s_t are all the relevant current school factors. y_{t-1} is primary school achievement and ε_t is the usual error term, independent of all included explanatory variables.

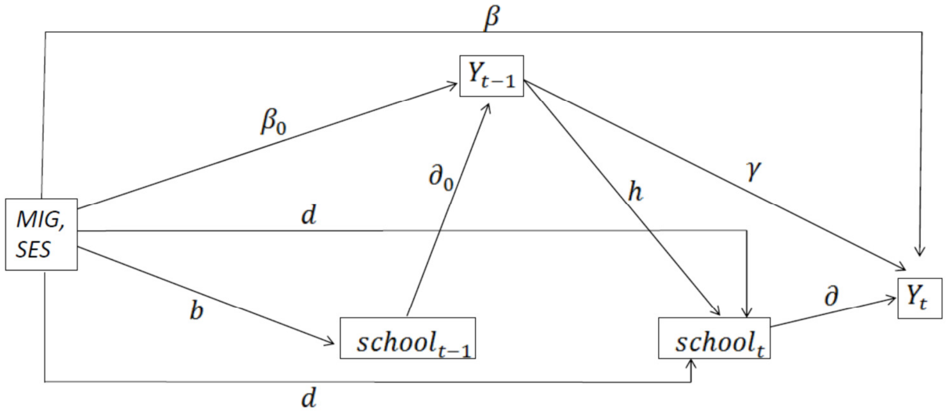
β represents the effect of x between *t*-1 and *t*, net of previous ability and school factors, γ carry-over effects of previous ability and δ the effect of current school factors, related to student body composition, teacher quality, management, resources, curricula.

Similarly, $y_{t-1} = \mu_0 + \beta_0 x + \delta_0 s_{t-1} + \varepsilon_{t-1}$, where the error ε_{t-1} also captures innate ability, while secondary school characteristics may depend on x and previous achievement:

$$s_{t-1} = a + bx + u_{t-1}$$
$$s_t = c + dx + hy_{t-1} + u_t$$

The extent to which school factors vary across family backgrounds is related to institutional features. For example, in "ideal" systems with no residential segregation, no stratification and perfect standardization, both primary and secondary school characteristics may vary only due to the random allocation of children and teachers. In systems with residential segregation, but no school differentiation, schools differ in family background composition, but should not differ in curricula and quality. If there is institutional differentiation, and tracking in particular, students' sorting is also driven by their previous ability, and schools differ also in curricula and potentially in teacher quality, management, resources.

Figure A.1 Theoretical individual-level model



Own elaboration from Contini and Grand (2013) and Esser (2014). See also Figure 1 in the text.

1.2 Estimated individual-level model

Given the lack of data on prior achievement in PISA we have to exclude y_{t-1} from the individual-level estimation. Since school choice is a crucial channel through which family background affects achievement, we also exclude school factors s_t. Thus, we obtain:

$$y_t = \mu + \beta x + \gamma y_{t-1} + \delta s_t + \varepsilon_t$$
$$= \mu + \beta x + \gamma(\mu_0 + \beta_0 x + \delta_0 s_{t-1} + \varepsilon_{t-1}) + \delta(c + dx + hy_{t-1} + u_t) + \varepsilon_t$$

which, with further substitutions becomes:

$$y_t = \mu' + (\beta + \gamma\beta_0 + \gamma\delta_0 b + \delta d + \beta_0 h\delta)x + (\delta u_t + \gamma\delta_0 u_{t-1} + \gamma\varepsilon_{t-1} + \varepsilon_t)$$

Hence, the coefficient of x is an *unbiased* estimate of the *total effect* of x. This is a quantity of interest, as it captures all direct and indirect effects of family background via previous achievement and school characteristics[1].

1.3 Country-level implications of the theoretical model

As we have argued above, the total effect of family background (SES and migrant status) on achievement is theoretically related to countries' institutional features. What are the country-level implications on family background differentials?

Elaborating on model (1), the average family background achievement differential at time *t* in country *c* is:

$$E(y_{ct}|x=1) - E(y_{ct}|x=0)$$
$$= \beta_c + \gamma_c[E(y_{ct-1}|x=1) - E(y_{ct-1}|x=0)] + \delta_c[E(s_{ct}|x=1) - E(s_{ct}|x=0)]$$

[1] In principle, we could also be interested in disentangling the total effect of family background into its different components related to previous achievement and school factors. Since we do not observe individual achievement growth, we cannot disentangle inequalities developed in primary school from those developed in secondary school. As for school factors, although a number of secondary school characteristics are available in PISA, they are likely to be affected by measurement error (school composition is measured from a random sample of 30 children per school, and school resources are subjectively reported by principals). Other issues like the omission of past school characteristic, of relevant school-level variables and the non-random allocation of children in schools are a threat to consistent estimation. As a general point, consider that estimating total effects is a much easier task than identifying the (causal) effects of the different factors at play.

comprising a component related to previous achievement and a component related to secondary school factors. Note that all the quantities involved are potentially country-specific. As we now explain, we expect varying cross-country differences on prior gaps and school allocation, depending on institutional arrangements.

In systems with no residential segregation, no stratification and perfect standardization, we can think of school characteristics as being randomly assigned to children. In this benchmark case, s_t varies little across schools, as variability is only due to random allocation of children and teachers.

In systems with residential segregation, but no stratification and perfect standardization, schools may differ only in terms of family background composition. We should observe schools with more and less disadvantaged children (according to the school location), but no other substantial differences in terms of teacher quality, resources, curricula). In this case, δ_0 and δ in model (1) represent school composition effects. We expect larger family-background achievement differentials in primary school and in current school factors with respect to the benchmark case mentioned above, and, in turn, larger differentials in y_t.

In systems with school differentiation – in particular (but not only) in tracked systems – secondary schools vary in terms of family background composition and ability, due to the so-called primary and secondary effects (effects via previous ability and effects net of previous ability; Boudon, 1974). Yet, schools also vary by curricula and level of the instruction, and possibly by other factors, such as teacher quality and resources. Here δ captures school composition effects *and* school quality effects. Larger differentials in school factors should bring about larger achievement differentials among low and high SES children, and between migrants and natives.

1.4 Rationale of the two-step strategy

In this perspective, it is sensible to apply a two-step analysis used, among others, by Hanushek and Woessmann (2006) and Schneeweis (2011). In the first step, we evaluate total background differentials within countries. In the second step, we relate these estimates to institutional features, with *country-level* analyses (in a purely exploratory/descriptive manner, as repeatedly pointed out in the main document). We conduct the analysis in two steps because this allows a greater flexibility of parameters (Achen 2005) and at the same time makes the explorative and descriptive character of the analysis more explicit.

Appendix 2: Additional information on our measure of migrant achievement penalties

Migrant underachievement is typically measured by the average gap of migrants with respect to natives. We propose an alternative measure that reveals the average position of immigrant children (*M*) into the distribution of their native peers (*N*), expressed in terms of standard deviations. We define the "raw *z*-score" as:

$$\bar{z}_M = \frac{1}{n}\sum_i z_{i,M} = \frac{1}{n}\sum_i \frac{y_{i,M} - \bar{y}_N}{\hat{\sigma}_N} = \frac{\bar{y}_M - \bar{y}_N}{\hat{\sigma}_N}$$

where *n* is the number of migrants. The *z*-score metrics – as opposed to the average-gap metrics – also considers the existing variability in the receiving societies. For a given migrant-native gap in PISA scores, lack of immigrant integration is more severe in a society with less heterogeneity among native children.

To isolate the migrant-specific disadvantage, we account for compositional effects due to socio-economic endowments *X* and use a modified version of the *z*-score, revealing the average position of second-generation migrants in the distribution of natives with the same socio-economic status. This index – emphasizing the relative rather than absolute distance between scores of natives and migrants – is our measure of *migrant achievement penalty*.

The "controlled *z*-score" is defined as:

$$\bar{z}_M^x = \frac{1}{n}\sum_i \frac{y_{i,M,x} - \bar{y}_{N|x}}{\hat{\sigma}_{N|x}} = \sum_x \bar{z}_{M|x}\, p_{M|x}$$

where $\bar{z}_{M|x}$ the *z*-score for given *x*, and $p_{M|x}$ the proportion of migrants with *X*=*x*. To evaluate $\bar{z}_M^x$ we refer to a simple model of performance *Y*, where we allow for differential returns to socio-economic resources for immigrants and natives:

$$Y_i = \alpha_N + \beta_N X_i + (\alpha_M - \alpha_N) MIG_i + (\beta_M - \beta_N) X_i\, MIG_i + \varepsilon_i \qquad (1)$$

MIG is a dummy indexing migrant background, α_M and α_N the intercepts for migrants and natives respectively, β_M and β_N the corresponding effects of socio-economic status. Thus:

$$\bar{z}_M^x = \frac{\sum_x[(\hat{\alpha}_M + \hat{\beta}_M x) - (\hat{\alpha}_N + \hat{\beta}_N x)]p_{M|x}}{\hat{\sigma}_\varepsilon} = \frac{\sum_x(\hat{\alpha}_M - \hat{\alpha}_N) + (\hat{\beta}_M - \hat{\beta}_N)\bar{X}_M}{\hat{\sigma}_\varepsilon} \qquad (2)$$

Incidentally, the numerator of (2) coincides with the unexplained component of the Blinder-Oaxaca decomposition of the absolute migrant-native achievement differential:

$$\bar{y}_M - \bar{y}_N = (\hat{\alpha}_M - \hat{\alpha}_N) + (\hat{\beta}_M - \hat{\beta}_N)\bar{X}_M + \hat{\beta}_N(\bar{X}_M - \bar{X}_N)$$

The last term is the explained component of the gap, i.e. the part that we can ascribe to compositional effects. The first two terms, instead, remain unexplained: the intercept difference is the average migrant-native gap at *X*=0, while the second term accounts for different returns to parental socio-economic status between migrants and natives.

Schneeweis (2011, pg. 1283) uses the unexplained component of the Blinder-Oaxaca decomposition as dependent variable. In interpreting it, she emphasizes differential returns. On the contrary, our empirical results show that the interaction effect is statistically significant only in a minority of countries, and that the unexplained component is mainly accounted by the difference in the intercepts (see Tables A5 and A6).

Appendix 3: Descriptive statistics

Table A3a. Sample sizes and descriptive statistics for math and *ESCS*, by country and migratory status

Country	**Natives**			**G2 migrants (all origins)**			**Turkish G2 migrants**		
	N	**Math**	***ESCS***	**N**	**Math**	***ESCS***	**N**	**Math**	***ESCS***
Austria	9838	511.1 (89.8)	0.226 (0.8)	869	444.2 (88.9)	-0.477 (0.8)	330	406.5 (77.9)	-0.850 (0.8)
Bel. Flanders	8859	547.7 (92.0)	0.256 (0.9)	380	465.1 (88.9)	-0.637 (1.0)	104	451.7 (85.8)	-0.978 (0.9)
Bel. Wallonia	5277	507.3 (96.2)	0.291 (0.9)	683	453.0 (95.6)	-0.285 (1.0)	123	430.7 (84.2)	-0.773 (0.9)
Switzerland	18333	548.3 (86.3)	0.208 (0.8)	3066	489.8 (91.0)	-0.340 (0.9)	368	454.0 (88.9)	-0.688 (0.9)
Germany	7692	521.8 (91.2)	0.367 (0.9)	763	450.3 (92.4)	-0.476 (1.0)	370	425.6 (80.4)	-0.703 (0.9)
Denmark	8633	514.6 (78.6)	0.370 (0.8)	1109	450.3 (79.9)	-0.446 (1.0)	375	423.9 (73.3)	-0.885 (0.9)
England+Wales	13117	497.4 (82.1)	0.221 (0.8)	556	480.0 (82.5)	0.056 (0.9)			
Spain	41778	487.5 (83.7)	-0.278 (1.1)	396	457.3 (81.2)	-0.366 (1.1)			
Finland	10228	546.3 (75.4)	0.322 (0.8)	69	490.3 (79.9)	0.246 (0.9)			
France	7722	505.2 (90.5)	-0.032 (0.8)	831	451.8 (93.1)	-0.627 (1.0)			
Greece	9004	467.4 (83.8)	-0.041 (1.0)	194	449.7 (86.5)	-0.258 (0.9)			
Italy	18729	506.3 (82.9)	0.053 (1.0)	306	459.3 (95.1)	-0.500 (1.1)			
Luxembourg	5658	509.9 (83.1)	0.489 (0.9)	1910	459.5 (84.5)	-0.487 (1.1)			
Netherlands	8414	536.3 (83.6)	0.361 (0.8)	764	474.3 (77.7)	-0.515 (1.0)	124	462.9 (69.9)	-0.656 (0.9)
Norway	8606	498.3 (81.8)	0.482 (0.7)	307	456.7 (90.6)	0.036 (0.9)			
Portugal	10701	480.1 (86.0)	-0.459 (1.2)	247	440.6 (97.4)	-0.321 (1.3)			
Sweden	7900	506.5 (83.6)	0.334 (0.8)	609	456.2 (83.2)	-0.043 (0.8)			-0.850 (0.8)

Source: PISA 2006-2009. Sample sizes: not weighted. Descriptive statistics: weighted, means and sd.dev. As motivated in the Appendix 4a, data from Italy exclude South and data from Germany exclude ethnic Germans.

Table A3b Shares of immigrants according to linguistic distance

Country	*No distance*	*Mild distance*	*High distance*	*V. high distance*	**Zero to mild**	**High to very high**
Austria	2.0	0.0	46.2	51.9	2.0	98.0
Bel. Flanders	1.8	1.8	33.4	63.0	3.6	96.4
Bel. Wallonia	4.5	58.3	8.6	28.5	62.8	37.2
Switzerland	6.0	0.2	62.8	31.0	6.2	93.8
Germany	2.5	2.7	41.6	53.2	5.2	94.8
Denmark	0.0	13.3	8.9	77.8	13.3	86.7
England+Wales	36.8	30.2	7.5	25.5	67.0	33.0
Spain	32.1	16.1	42.9	8.9	48.2	51.8
Finland	0.0	15.5	3.1	81.4	15.5	84.5
France	2.5	65.0	15.2	17.3	67.5	32.5
Greece	12.3	0.0	28.3	59.4	12.3	87.7
Italy	0.7	19.1	43.8	36.4	19.8	80.2
Luxembourg	2.9	8.5	77.5	11.1	11.3	88.7
Netherlands	22.2	1.2	14.9	61.7	23.5	76.5
Norway	2.2	5.5	0.0	92.3	7.7	92.3
Portugal	80.1	9.6	0.0	10.3	89.7	10.3
Sweden	0.0	16.0	39.4	44.6	16.0	84.0

Own calculation. Source: refer to Appendix 3c.

Table A3c Explanatory variables at a country level

Country	(Pre)school system entry	Marginalization	Linguistic distance
Austria	0.139	5.542	1
Bel. Flanders	-1.115	3.811	1
Bel. Wallonia	-1.115	3.094	0
Switzerland	1.834	3.137	1
Germany	-0.059	3.445	1
Denmark	1.664	7.074	1
England+Wales	-2.884	2.436	0
Spain	-1.093	1.509	0
Finland	3.733	1.209	1
France	-1.115	2.098	0
Greece	0.866	1.225	1
Italy	-0.873	2.019	1
Luxembourg	-0.939	1.908	1
Netherlands	-2.882	3.051	1
Norway	0.315	2.198	1
Portugal	0.844	2.485	0
Sweden	2.478	3.057	1

Own calculation. Source: refer to main manuscript, section 5.

Appendix 4: Additional information on data and variable construction

a. Details on selected countries

Belgium: split into (1) Belgium-Flanders and (2) Belgium-Wallonia (different educational systems).

Great Britain: split into (1) England and Wales, (2) Scotland and (3) Northern Ireland (different educational systems). Scotland and Northern Ireland dropped (insufficient sample sizes for immigrant students).

Italy: split into (1) Northern and Central regions, and (2) Southern regions. An accurate measure of relative disadvantage must contrast second-generation migrants with their own peers. However, given the poor general performance levels and the limited presence of second-generation migrants in the South, one should contrast second-generation migrants in Northern-Central regions to natives in Northern-Central regions, and second-generation migrants in Southern regions to natives in Southern regions. Southern regions dropped (insufficient sample sizes for immigrant students).

Germany: students whose mother was born in former USSR excluded from the sample (most probably ethnic-German return migrants, given their extremely high test scores and German as language spoken at home).

b. Individual-level regressions: details on ESCS as a control variable

The *ESCS* – provided by PISA – is a synthetic measure that allows parsimony while capturing several dimensions of social, economic, and cultural resources. Operationalizing these resources as distinct dimensions does not affect our estimates (sensitivity checks are available upon request).

Parental education (one of the components of *ESCS*) can be affected by measurement error. However, Engzell and Jonsson (2013) show that this source of error has no serious consequences on the estimates of the immigrant-origin effect given social origin. As robustness checks, we performed additional analyses controlling for parental occupation only. Results (available upon request) are consistent with those reported in the paper.

c. Individual-level regressions: details on decision not to control for language spoken at home

PISA provides information on whether students speak the destination-country language at home. Despite its predictive power documented by previous studies using PISA data, we do not control for this variable for two reasons:

(i) Endogeneity

Language spoken at home is endogenous to the achievement of second-generation immigrant children. In fact, whether or not children speak the destination-country language (L2) at home depends on a variety of factors, including children's and parents' linguistic skills in L2. The level of linguistic skills of immigrant *children* is highly endogenous to school achievement in all the literacy domains assessed by PISA. Clearly, the educational system affects the degree to which immigrant students master L2. *Parental* skills are also endogenous, because – through their children's schooling – parents can improve their linguistic competences. For example, if immigrant children generally attend preschool from early age, they will have more opportunities to improve their linguistic skills and therefore, be more inclined to speak L2 also at home, with siblings and parents. Therefore, controlling for the individual-level variable of language spoken at home would hinder the identification of the systemic effect of preschool, by capturing part of the desired effect.

(ii) Language spoken at home vs. linguistic skills

While language spoken at home is indeed related to linguistic skills, its precise meaning is unclear, especially from a theoretical perspective. While Esser (2006) argues that factors that promote L1 retention hinder immigrants' skills in L2, many educational scientists support the concurring hypothesis of "linguistic interdependence" (Cummins 1979). As lamented by Kristen *et. al* (2011), no empirical rigorous study has tested the validity of one over the other. What is more, the PISA variable is also a very rough indicator of parental linguistic skills in L2. It might well be the case that immigrant parents, while mastering L2, prefer to speak to their children in their mother tongue in order to preserve their cultural ties with the origin country. On the contrary, it is possible that immigrant parents, despite having a very poor knowledge of the host-country language, make an effort to use it to communicate with their children. In addition, pupils may speak the host language with their siblings and switch to the origin-country language with their parents. Unfortunately, the phrasing of the question in the PISA questionnaire is quite vague ("What language do you speak at home most of the time?"), so that is not possible to disentangle the two.

d. <u>Country-level regression-tree analysis: details on linguistic distance variable</u>

In the first step, we assessed the distance between each origin-country language and the destination-country language. Distance was assessed according to official languages' families and subfamilies (Lewis 2009) as follows: coinciding language: zero distance (e.g. American and British English); same linguistic sub-family: mild distance (e.g. Spanish and French as Romance languages); same family: high distance (e.g. Polish and French as Indo-European languages); different families: very high distance (e.g. Turkish as Altaic and German as Indo-European). However, when the destination-country language is widely spoken in the origin country, distance was also considered mild (e.g. French as widely spoken in Algeria).

In the second step, we aggregated linguistic distance into a country index by shares of immigrant groups. For countries where information on the country of birth of the mother was available in the national questionnaires, PISA 2006-09 data were used to compute shares of second-generation immigrants (Austria, Belgium-Flanders, Belgium-Wallonia, Switzerland, Germany, Denmark, Luxembourg, the Netherlands, Norway, Portugal). As a second best, we used UN-DESA data on immigration flows in the period 1975-1993 to proxy origins of parents of second-generation immigrants born in 1994 (England and Wales, Sweden). Where no international information on country of origin was available, we relied on national statistics: Spain (Observatorio Permanente de la Inmigración – Ministerio del Interior: data on foreign residents aged 16-64 in 2003); France (INED: data on foreign residents aged 25-54 in 2009); Italy (ISTAT: data on foreign residents in 2003).

Finally, since two clear-cut clusters of destination-countries emerge (see Table A3b), we recoded the indicator of linguistic distance as a dummy, with 0 indicating zero or mild linguistic distance, and 1 indicating high or very high linguistic distance.

Appendix 5: Estimates of individual-level regressions

Table A5. Estimates of individual-level regressions of mathematics scores

Country	$(\alpha_M - \alpha_N)$	β_N	$(\beta_M - \beta_N)$
Austria	-38.11** (5.44)	40.37** (2.13)	
Bel. Flanders	-56.43** (9.73)	44.63** (1.59)	-23.79** (6.07)
Bel. Wallonia	-30.21** (6.61)	51.80** (2.28)	-21.00** (5.19)
Switzerland	-39.51** (3.12)	34.25** (1.36)	
Germany	-41.15** (4.84)	45.33** (1.80)	-15.35** (4.00)
Denmark	-40.26** (4.82)	32.52** (1.32)	-9.39* (3.90)
England+Wales	-10.86** (4.00)	40.8** (1.58)	
Spain	-25.80** (6.59)	28.45** (0.97)	
Finland	-54.27** (12.04)	28.58** (1.17)	
France	-33.31** (6.73)	52.17** (1.81)	-19.28** (4.14)
Greece	-10.43 (6.96)	33.88** (1.56)	
Italy	-35.77** (8.38)	23.24** (1.41)	
Luxembourg	-21.07** (2.66)	29.66** (1.01)	
Netherlands	-35.71** (6.00)	38.47** (1.64)	-16.94** (3.78)
Norway	-25.41** (7.08)	34.43** (1.53)	
Portugal	-40.02** (8.34)	31.57** (1.31)	10.10* (5.14)
Sweden	-34.32** (4.60)	37.56** (1.60)	

Source: PISA 2006-2009. Country-specific regressions of mathematics scores estimated using replicate weights and plausible values. Model: refer to Equation (2) in the main manuscript. Additional controls: gender. Where the interaction term was not significant at 5% level, models were rerun without it.

** Sig. at 1% level * Sig. at 5% level. Standard errors in parentheses.

$(\alpha_M - \alpha_N)$ is the mean difference in the scores of migrants and natives at *ESCS*=0 (the OECD average), while β_N is the effect of one additional point in the *ESCS* scale for natives; where the interaction coefficient $(\beta_M - \beta_N)$ is non-significant, $(\alpha_M - \alpha_N)$ is the mean difference at all values of *ESCS*.

Appendix 6: Additional analyses on reading and science

Table A6. Estimates of additional individual-level regressions of reading and science scores

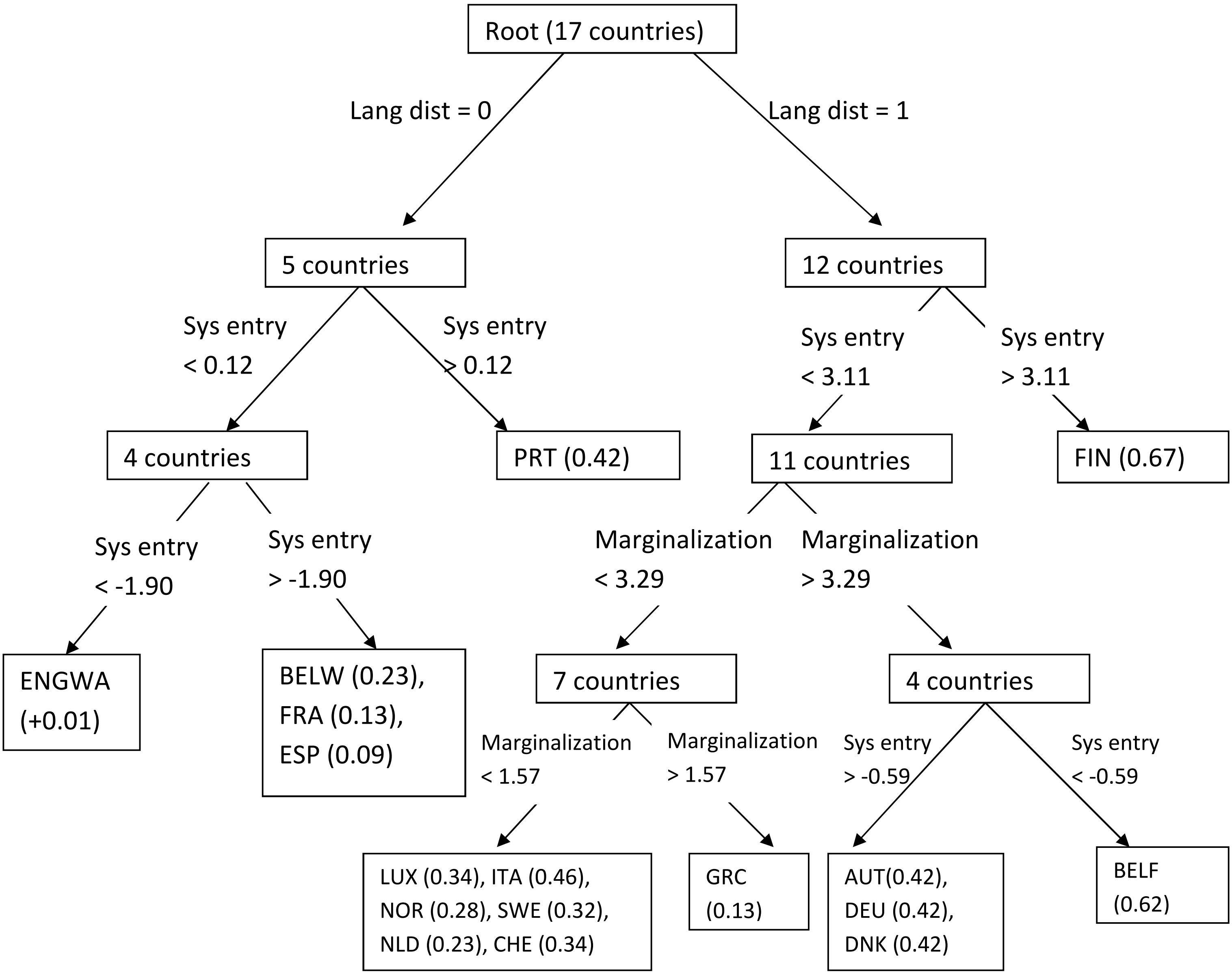
Country	$(\alpha_M - \alpha_N)$	β_N	$(\beta_M - \beta_N)$
		Reading scores	
Austria	-37.13** (7.00)	43.81** (1.95)	
Bel. Flanders	-61.31** (8.63)	42.43** (1.76)	-15.61** (5.77)
Bel. Wallonia	-25.71** (7.51)	51.84** (1.96)	-20.39** (5.91)
Switzerland	-26.06** (2.64)	34.28** (1.25)	
Germany	-40.08** (5.25)	44.14** (1.96)	-8.92** (3.92)
Denmark	-31.40** (5.14)	32.20** (1.29)	
England+Wales	0.74 (5.14)	43.83** (1.46)	
Spain	-6.75 (8.27)	27.18** (1.00)	
Finland	-47.66** (11.76)	27.65** (1.07)	
France	-25.27** (6.41)	51.66** (2.00)	-21.14** (4.79)
Greece	-10.89 (8.25)	35.13** (1.58)	
Italy	-36.84** (9.39)	27.78** (1.44)	
Luxembourg	-27.95** (3.00)	32.63** (1.09)	
Netherlands	-25.00** (6.44)	38.42** (1.54)	-12.13* (5.18)
Norway	-23.56** (7.06)	36.48** (1.80)	
Portugal	-32.25** (7.48)	31.77** (1.27)	
Sweden	-26.20** (4. 69)	36.39*** (1.73)	
		Science scores	
Austria	-53.67** (5.73)	42.24** (1.82)	
Bel. Flanders	-66.25** (7.90)	43.39** (1.54)	-19.92** (6.15)
Bel. Wallonia	-29.58** (7.48)	51.44** (2.02)	-23.43** (5.57)
Switzerland	-45.42** (2.87)	35.58** (1.28)	
Germany	-57.03** (4.87)	44.25** (1.70)	-11.62** (3.45)
Denmark	-48.85** (5.28)	36.00** (1.40)	
England+Wales	-10.32* (4.73)	47.61** (1.64)	
Spain	-13.68 (7.13)	28.54** (0.97)	
Finland	-59.81** (13.47)	28.85** (1.16)	
France	-34.92** (6.79)	54.03** (1.74)	-21.21** (4.26)
Greece	-16.83** (6.05)	33.79** (1.55)	
Italy	-46.63** (8.68)	25.39** (1.29)	
Luxembourg	-33.78** (3.01)	32.15** (0.96)	
Netherlands	-42.77** (7.93)	42.02** (1.49)	-16.73** (4.45)
Norway	-41.90** (7.27)	35.17** (1.59)	
Portugal	-33.69** (7.65)	29.61** (1.13)	
Sweden	-41.22** (5.08)	36.98** (1.46)	

Source: PISA 2006-2009. Country-specific regressions of reading and science scores estimated using replicate weights and plausible values. Model: refer to Equation (2) in the main manuscript. Additional controls: gender. Where the interaction term was not significant at 5% level, models were rerun without it.

** Sig. at 1% level * Sig. at 5% level. Standard errors in parentheses.

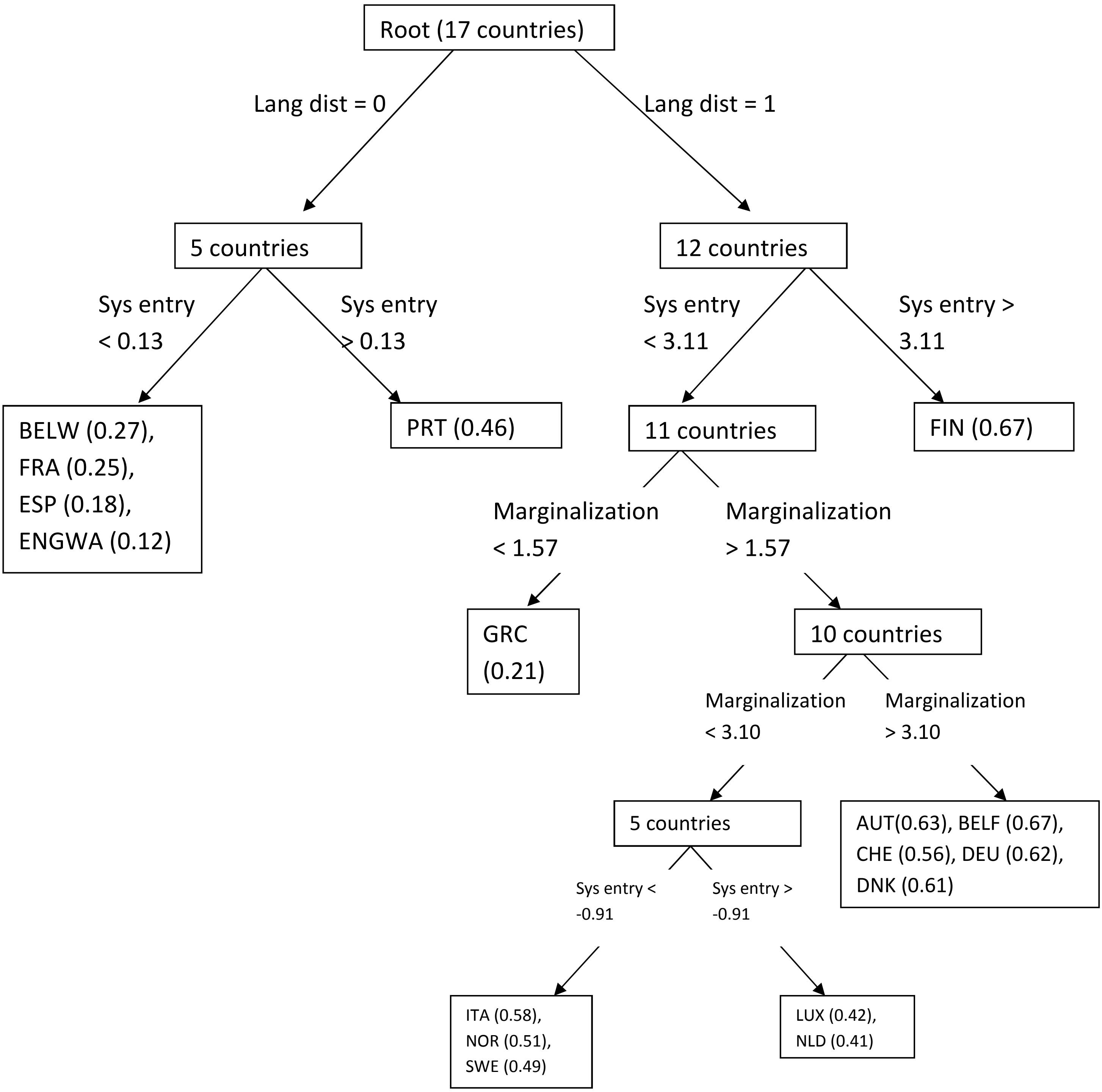
$(\alpha_M - \alpha_N)$ is the mean difference in the scores of migrants and natives at *ESCS*=0 (the OECD average), while β_N is the effect of one additional point in the *ESCS* scale for natives; where the interaction coefficient $(\beta_M - \beta_N)$ is non-significant, $(\alpha_M - \alpha_N)$ is the mean difference at all values of *ESCS*.

Figure A6a. Regression-tree analysis (reading)



Response variable: migrant-specific penalties in reading achievement (absolute values are reported in parentheses). Explanatory variables: marginalization, (pre)school-system entry, linguistic distance. Analyses performed with package R "rpart". Method: "anova", Complexity parameter: 0.04.

Figure A6b. Regression-tree analysis (science)



Response variable: migrant-specific penalties in science achievement (absolute values are reported in parentheses). Explanatory variables: marginalization, (pre)school-system entry, linguistic distance. Analyses performed with package R "rpart". Method: "anova", Complexity parameter: 0.02.

References for the Supplementary Material

Achen, Christopher H. (2005). Two-Step Hierarchical Estimation: Beyond Regression Analysis. Political Analysis **13(4)**, 447-456.

Contini, Dalit, and Elisa Grand (2013). On estimating achievement dynamic models from repeated cross-sections. *Department of Economics and Statistics "Cognetti de Martiis" Working Paper* **43/13**.

Engzell, Per, and Jan O. Jonsson. (2013). The Big Trade-off in School Surveys? Consequences for Stratification Research of Child Misreports and Parent Nonresponse in Socioeconomic Variables. *mimeo*.

Esser, Hartmut (2014). *Sorting and School Effects. The Standard-Approach and the DVD-Puzzle on Effects of Education Systems on Educational Inequality*. AMCIS Conference presentation, February 13-14, Amsterdam.

Hanushek, Eric A., and Ludger Woessmann (2006). Does Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries. *The Economic Journal* **116(150)**, C63-C76

Lewis, M.P. (Ed). (2009). *Ethnologue: Languages of the World, 16th Edition*. 16th ed. Dallas: SIL International.

Schneeweis, Nicole (2011). Educational Institutions and the Integration of Migrants. *Journal of Population Economics* **24 (4)**, 1281–1308.