The progression of achievement gap between immigrant and native-born students from primary to secondary education

Aigul Alieva, Vincent A. Hildebrand, Philippe Van Kerm

PII: S0276-5624(24)00074-X

DOI: https://doi.org/10.1016/j.rssm.2024.100961

Reference: RSSM100961

To appear in: *Research in Social Stratification and Mobility*

Received date: 21 August 2023
Revised date: 16 July 2024
Accepted date: 23 July 2024

Please cite this article as: Aigul Alieva, Vincent A. Hildebrand and Philippe Van Kerm, The progression of achievement gap between immigrant and native-born students from primary to secondary education, *Research in Social Stratification and Mobility,* (2024) doi:https://doi.org/10.1016/j.rssm.2024.100961

The progression of achievement gap between immigrant and native-born students from primary to secondary education*

Aigul Alieva[†]
LISER, Luxembourg

Vincent A. Hildebrand[‡]
Glendon College, York University, Canada

Philippe Van Kerm[§]
University of Luxembourg and LISER, Luxembourg

July 16, 2024

Abstract

This paper depicts the evolution of gaps in academic performance between native and immigrant background students, as they progress from primary to secondary education. We study three cohorts of students in European and traditional English-speaking immigration countries using combinations of international assessment studies (PIRLS, TIMSS and PISA). To address the issue of comparability of test scores across surveys and over time, we exploit rank-based measures of relative performance, which only require ordinal comparability of the data. We do not find significant differences between the academic achievements of immigrant children and their native-born peers in English-speaking receiving countries. By contrast, immigrant-background children – both of first- and of second-generation – exhibit a large achievement gap in primary school in Europe, even when accounting for observable differences in socioeconomic characteristics. The gap tends to narrow down in secondary education in both reading and mathematics but is not fully absorbed in most countries. This finding is noteworthy among second-generation students in systems with early tracking. The performance of students with mixed parents is not markedly different from native students. Diverging educational progress between immigrant children in traditional immigration countries and our sample of European countries seems to reinforce the importance of the initial socioeconomic endowment in shaping the academic trajectories of immigrant children.

*We thank two anonymous referees and the editor for their insightful comments and helpful suggestions. . This research is part of the 'The Persistence of Social and Ethnic Disadvantages in Primary and Secondary Schools' (PERSIST) project funded by the 'Fonds National de la Recherche Luxembourg' (grant CORE 2012 C12/SC/3943127). Vincent Hildebrand thanks the staff of the Living Conditions Department for hosting part of his sabbatical leave at the Luxembourg Institute of Socio-Economic Research (LISER) where a large part of this research was carried on.

[†]LISER, Maison des Sciences Humaines, 11, Porte des Sciences, L-4366 Esch-sur-Alzette, Luxembourg, aigul.alieva@liser.lu.

[‡]Department of Economics, Glendon College, York University, Canada, vincenth@yorku.ca.

[§]Unversity of Luxembourg, Maison des Sciences Humaines, 11, Porte des Sciences, L-4366 Esch-sur-Alzette, Luxembourg, philippe.vankerm@uni.lu.

Keywords: Achievement gap, immigrant-background students, primary education, secondary education, test scores comparability, Europe, Traditional Immigration Countries, TIMSS, PIRLS, PISA

The progression of achievement gap between immigrant and native-born students from primary to secondary education

July 24, 2024

Abstract

This paper depicts the evolution of gaps in academic performance between native and immigrant background students, as they progress from primary to secondary education. We study three cohorts of students in European and traditional English-speaking immigration countries using combinations of international assessment studies (PIRLS, TIMSS and PISA). To address the issue of comparability of test scores across surveys and over time, we exploit rank-based measures of relative performance, which only require ordinal comparability of the data. We do not find significant differences between the academic achievements of immigrant children and their native-born peers in English-speaking receiving countries. By contrast, immigrant-background children – both of first- and of second-generation – exhibit a large achievement gap in primary school in Europe, even when accounting for observable differences in socioeconomic characteristics. The gap tends to narrow down in secondary education in both reading and mathematics but is not fully absorbed in most countries. This finding is noteworthy among second-generation students in systems with early tracking. The performance of students with mixed parents is not markedly different from native students. Diverging educational progress between immigrant children in traditional immigration countries and our sample of European countries seems to reinforce the importance of the initial socioeconomic endowment in shaping the academic trajectories of immigrant children.

Keywords: Achievement gap, immigrant-background students, primary education, secondary education, test scores comparability, Europe, Traditional Immigration Countries, TIMSS, PIRLS, PISA

1 Introduction

The continuous inflow of migrants is rapidly changing the classroom environment in many societies. In 2022, the share of 15 year-old immigrant origin students ranged between 15 and 35 per cent in Western European and the traditional English-speaking immigration countries, surpassing the OECD average at 13 per cent (OECD, 2023). Between 2012 and 2022, the fastest growth was observed in Germany, Switzerland and Austria, ranging from 10 to 12 percentage points, while Traditional immigration countries (TIC) experienced a slower growth around 5-6 percentage points (OECD, 2023). Recent newcomers to Europe include large numbers of asylum seekers, with estimated 2 million of children lodging a first-time asylum application in between 2012 and 2022 (Eurostat, 2022). This unprecedented influx of humanitarian migrants has been driven mainly by the 2015–2016 Middle Eastern crisis and the ongoing Russian War against Ukraine since 2022. In contrast, statistics from TIC, such as Australia and Canada, show that these countries received overwhelmingly skilled newcomers from Asia and Europe over that period, with only a small proportion of asylum seekers originating from these tumultuous regions (Statistics Canada, 2021, Australian Bureau of Statistics, 2024). In this context, evidence of persistent achievement gaps between immigrant and native-born children in EU countries is a cause for concerns and motivates the need to better document the integration trajectories of immigrant children into the host country education systems (Borgna, 2016, Dollmann, 2021, Klein and Neugebauer, 2023).

A substantial part of achievement gaps can be attributed to socioeconomic factors. A large proportion of non-native students come from less developed countries and are up to three times more at risk of living in a poor household than their native counterparts, even among those living with highly educated parents (OECD/EU, 2015a). Many immigrant children, in particular among newcomers, grow up in households that do not speak the host country's language at home – another critical factor of academic success (Alba et al., 2011, Isphording, 2014). Results from PISA data persistently point to substantial language penalties in countries like Luxembourg, France and Switzerland (OECD/EU, 2015b). Language barriers may also be higher for some immigrant groups.

For instance, research on the linguistic distance between languages in host and origin countries explains part of the language gap (Isphording and Otten, 2014) and the wage gap in the adult population (Isphording, 2014). Immigration policy to TIC, which requires language skills screening next to selection based on socioeconomic and occupational status, further contributes to the positive selection of families. Moreover, TIC countries are overwhelmingly English-speaking countries, and with the language being a global *lingua franca* in economy, media & technology, education, and culture (Melitz, 2018), learning opportunities and costs

have become increasingly accessible.

Receiving countries vary substantially with respect to their institutional settings (Algan et al., 2010, Kogan, 2016). Education policies, including the structure of the education system, not only shape the human capital of the younger generation but also mediate socioeconomic inequalities (Pfeffer, 2008). For instance, participation in early childhood education and care (ECEC) is shown to benefit disadvantaged immigrant students (Schneeweis, 2011), with an estimated advantage of 55 points on OECD PISA tests (Biroli et al., 2017, Klein and Becker, 2017, OECD/EU, 2018). However, only 70 per cent of immigrant children attend preschool programs across OECD countries, in part due to associated costs or shortage of places in ECEC (OECD/EU, 2015b).

Many European countries have committed significant resources to address the educational needs of vulnerable children, including immigrant background students (see, e.g., Anderson et al., 2015, Klein and Becker, 2017, Riederer and Verwiebe, 2015). Yet, lackluster educational performances of students with a migration background in EU countries is underscored by large-scale comparative assessments studies, such as the Progress in International Reading Literacy Study (PIRLS), the Programme for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS). These three surveys provide data on skills and knowledge of students in mathematics, science and reading in primary and secondary education over the span of two decades either at a specific grade (TIMSS in grades 4 and 8 and PIRLS in grade 4), or at a specific age (15 years old in PISA). Studies based on these surveys have documented gaps in test scores between immigrant background and native children in either primary or secondary schools (Ammermüller, 2013, Cobb-Clark et al., 2012, Schneeweis, 2011). Other contributions have explored the academic progress between primary and secondary schools, however these contributions did not specifically focus on the scholastic progress of immigrant background children (see, e.g., Hanushek and Woessmann, 2006, Schubert and Becker, 2010, Jakubowski and Pokropek, 2015).

The present study contributes to this literature in two distinctive ways. First, we document the educational achievement of migrant children in a comparative perspective. Although a number of studies have documented the immigrant-native gap at a particular point in their education, to the best of our knowledge, little is known about the academic *progress* of immigrant children from childhood to adolescence as measured by the change over time in the academic performance gap relative to their native-born peers (henceforth "the immigrant achievement-gap change") in many European countries.[1]

[1] Vonnahme (2021), which documents the immigrant achievement-gap change for Germany by leveraging recent longitudinal data from NEPS, is a notable exception. Earlier, Andon et al. (2014) provided evidence for OECD countries as a whole, and more recently, Hanushek et al. (2022) documented the immigrant achievement-gap

We provide new measures of the progress (or regression) of immigrant-background children's knowledge and academic skills as they move through different stages of the school system. While it may not be surprising to find that immigrant-background students face a disadvantage in the early stage of their educational trajectories – for reasons outlined above, e.g., because they need to adapt and acquire new linguistic skills – evidence of persisting achievement gaps across successive educational stages is a source of concern as these signal failure to catch up to natives and therefore a risk of consolidating persisting long-run disadvantages. Comparisons of the evolution of the achievement gap along educational trajectories from primary to secondary education provide clearer evidence of differentiated achievement patterns, which is ultimately relevant to helping shape targeted interventions for the most vulnerable groups of students.

We document the immigrant achievement-gap change for three cohorts of students from three generations of immigrant children (first (1G) and second generation (2G) as well second generation of "Mixed" origin, which we define in section 3.1) between the ages of approximately 10 and 14-15 using data from PIRLS, TIMSS and PISA for twelve Western European countries and four traditional immigration countries (TIC). Our study design allows us to document the academic progress of immigrant children *within* cohorts in two academic subjects: reading (Cohort 1 and 2) and mathematics (Cohort 3), as the fundamental skills and knowledge for successful integration. While *between* cohorts comparisons do provide another interesting temporal dimension of academic progress, our study does not formally discuss these results as compositional changes in the migrant population over time as well as changes in both educational and migration policies make their interpretation difficult.

Our second contribution is methodological. While both the International Association for the Evaluation of Educational Achievement (IEA) and the OECD collect valuable assessment data for policy and academic research, PIRLS, PISA and TIMSS surveys were not originally designed to be exploited jointly. A number of studies have used these surveys in a quasi-longitudinal setup to explore the change in achievement of specific student cohorts between primary and secondary school in absence of genuine longitudinal data (Hanushek and Woessmann, 2006, Schubert and Becker, 2010, Jakubowski and Pokropek, 2015, Dämmrich and Triventi, 2018). These studies typically rely on the implicit (but strong) assumption that standardized test scores of different groups of students from PIRLS, PISA and TIMSS are *cardinally* comparable within and across datasets.

As we discuss below, this is questionable and we address this issue by relying on a measure of academic progress over time, which compares the relative *ranking* of native and immigrant children in the distribution of tests scores using Gastwirth's PROB Index (Gastwirth, 1975).

change for the United States.

This ordinal measure has the benefit over more traditional cardinal measures of the achievement gap of not relying on the assumption that the scaling of scores as the same interpretation across different surveys, and/or time periods.

Our findings reveal a stark contrast in the level and progression of achievement gaps in traditional immigration countries and in European countries. We observe comparable academic trajectories between immigrant children and their native peers in traditional immigration countries, both in reading and in mathematics. By contrast, immigrant children in Europe generally exhibit substantially larger achievement gaps in primary school, which narrow down over time but are not fully absorbed in many countries. Although we do not attempt to pin down the reasons for such contrasted trajectories, we do find some evidence of a negative association between the progression of academic gaps over time (mainly for 2G students) and early tracking. Within this broad characterisation of differences between TIC and European countries, our results reveal a great deal of heterogeneity, not only between countries, but also across student cohorts within countries.

2 Background

Our study is anchored in an extensive literature that we group into individual- and institutional-level determinants of the achievement gap. We briefly review this literature and then state the specific research questions that we aim to address here.

2.1 *Socio-economic status and language skills*

A literature based on longitudinal data has explored racial and ethnic disparities in educational achievement in the US, originally motivated by the resurgence of growing disparities between Black and White children as they progressed from kindergarten to elementary school in the late 1980s, early 1990s. The compositional changes in environmental factors between Black and White children, the reappearance of segregation, and the lower school quality attended by Black students have been identified as important risk factors forging the racial achievement gap in the US (Fiel, 2013, Hanushek and Rivkin, 2006, 2009). By the age of five, chronic exposure to poverty together with poor home environment account for up to 80 percent of the cognitive gap between Black and White children (Brooks-Gunn et al., 1996).

Evidence of disparities in educational achievement in the US is not confined to Black students. Hispanic children exhibit lower test scores in reading and mathematics upon entering kindergarten than both Blacks and Whites (Fryer and Levitt, 2004, Reardon and Galindo, 2009). A large proportion of Hispanic children also face poor environmental factors from an early age

due to their weaker socioeconomic position. Hispanics arguably form a more heterogeneous group than Blacks, with a large proportion of children with a recent immigration background who often do not speak English at home (Reardon and Galindo, 2009). Interestingly, the academic trajectory of Hispanic children improves throughout kindergarten and grade school, whereas that of Blacks continues to deteriorate (Fryer and Levitt, 2004, 2006). The achievement gap between Hispanic and White children narrows significantly during the first two years of elementary school, but flattens thereafter (Fryer and Levitt, 2006, Reardon and Galindo, 2009). By contrast, the achievement gap of Black students continues to widen throughout grade school (Fryer and Levitt, 2006, Reardon and Galindo, 2009). These findings point to i) the presence of a large achievement gap before the start of primary school among disadvantaged students, ii) a potential further widening of the achievement gap as students progress through school, and iii) heterogeneous academic trajectories between different groups of children sharing vulnerable socioeconomic backgrounds.

By contrast to US studies, research documenting the educational achievement gap among immigrant children in Europe is largely cross-sectional. Some consistent findings have been reported, even if the use of different surveys, time periods, and the focus on different immigrant groups (e.g., by generational status or country/region of origin) make it difficult to clearly identify the evolution of the native-immigrant gap over time. The achievement gap, both at the primary and secondary levels in reading, mathematics, and science, is of a larger magnitude in European countries than in TIC (Cobb-Clark et al., 2012, Schnepf, 2007, Entorf and Minoiu, 2005). The marked differences between these two groups of countries are largely attributed to the selective migration policies of TIC (Sakellariou, 2018). A significant share of immigrant children in Europe come from families with low socioeconomic status and educational capital (SES) and, in turn, SES remains a key individual-level predictor of both educational achievement and attainment both over time and across countries, regardless of migration status (Breen and Jonsson, 2005, Breen et al., 2009, Bukodi and Goldthorpe, 2013, Bernardi and Ballarino, 2016, Erikson and Goldthorpe, 2002, Erola et al., 2016, van de Werfhorst, 2019).

SES-driven disparities in cognitive functioning, language, and vocabulary development are significant in the early years of a child's development (Noble et al., 2005, 2015, Klein et al., 2014). A recent longitudinal study exploring the role of socioeconomic inequalities in Germany, the Netherlands, and the United Kingdom reports that between 50 and 80 percent of the achievement gap in primary school precedes the start of formal education (Passaretta et al., 2022). Similar results were found in the USA and Canada (Brooks-Gunn et al., 1996, Merry, 2013). Differences in parenting styles, the quality of early childcare, both formal and informal (Ruzek et al., 2014), and participation in preschool education (Heckman, 2006, Janssen et al., 2023) are among the key mechanisms to offset the initial disadvantage. However, research

has consistently shown lower pre-school attendance among children from disadvantaged and immigrant backgrounds (Hogrebe and Strietholt, 2016, OECD/EU, 2018). Additionally, young immigrant children, particularly newcomers, suffer from insufficient exposure to the language of the host society, which could be addressed, at least in part, through enrollment in formal preschool care (Klein and Becker, 2017).

2.2 *The role of education systems*

At the start of primary education, disadvantaged and immigrant students need to acquire not only curriculum-specific knowledge but also boost their language skills. This is challenging in a context in which the schools attended by more vulnerable student populations tend to differ in their quality: less experienced teaching staff, fewer resources, and a higher concentration of disadvantaged students (Hanushek and Rivkin, 2006, 2009), resulting in a higher likelihood of grade retention (Hanushek and Rivkin, 2006, Pedraja-Chaparro et al., 2016, Park and Sandefur, 2010) and poor school grades (Alesina et al., 2018, Luedemann and Schwerdt, 2013), thus hindering catching up to native students.

The transition to secondary school is another important step in the educational trajectory of immigrant students, particularly in Europe, owing to more heterogeneity in the structure of secondary education systems (Bol and van de Werfhorst, 2013). Academic tracking (curriculum differentiation and ability grouping) is of particular interest because it is known to exacerbate educational inequalities among the most vulnerable learners (Alba et al., 2011, Bol et al., 2014, Dustmann et al., 2017, Gamoran and Mare, 1989, Hanushek and Woessmann, 2006). Tracking usually takes place when students transition from one educational stage to the next, e.g. from primary to secondary schools, or from lower to upper secondary. Immigrant students face unequal chances of entering academic tracks because of biases in placement decisions (Glock et al., 2013). Early tracking has been shown to be more detrimental to the achievement of the most vulnerable students than late (or absent) tracking (Mulhlenweg, 2008, Lavrijsen and Nicaise, 2016, Burger, 2016), particularly for first-generation students (Mulhlenweg, 2008).

Vocational training offered in a tracking system improves the odds of a successful transition to the labour market (Brunello and Checchi, 2007). German education and training system, which combines early tracking (after four years of primary school) with an extensive vocational education and training system that performs well, is a particularly fitting example (DiPrete et al., 2017). It is also important to acknowledge that some form of ability grouping exists in the majority of Western societies, such as *course-by-course* (within-school) tracking, with students in the same class being offered course material with different levels of difficulty (Chmielewski, 2014). Nevertheless, Chmielewski (2014), Engzell and Raabe (2023) suggest

that *academic/vocational* (between-school) tracking generates larger achievement gaps driven by socioeconomic background. These results align with the findings of varying explanatory power of SES across European education systems: it is found to be larger in Western countries applying early between-school tracking, and weaker in Northern and Southern countries, with late or within-school tracking (OECD, 2019).

Navigating through the education system requires the support and knowledge of the system, which also sets apart high- and low-SES parents (Forster and van de Werfhorst, 2019, Triventi et al., 2020). Parents with higher SES have a greater understanding of institutional differences to foster better educational outcomes for their offspring (ibid.). They are also more likely to make extensive use of shadow education to reach this goal (Zwier et al., 2020). Low-SES immigrant families might lack the knowledge to fully understand the host country's education system, leaving them unequipped to adequately support and guide their children throughout their academic journey (Antony-Newman, 2019). However, in line with the US literature and the *immigrant optimism* hypothesis (Kao and Tienda, 1998), studies in Europe have also reported high educational aspirations among immigrant parents and their offspring (Becker et al., 2023, Jonsson and Rudolphi, 2011). These ambitions help migrant students complete compulsory school and enter tertiary education; however, they remain at higher risk of dropping out (Crul, 2013, Klein and Neugebauer, 2023, Jackson et al., 2012).

Our study also includes students of mixed origin (one of the parents is native-born), as this generational group is steadily growing in size across many Western societies (Alba and Foner, 2015, Cerna et al., 2021, Kalmijn, 2015). According to research, while in some countries, such as Italy and the UK, education and occupation outcomes of the mixed-origin group are similar to those of the native population Azzolini and Barone (2013), Muttarak (2013), in other countries, including some TIC destinations, the results are mixed (Alba and Foner, 2015, Kalmijn, 2015, Sakellariou, 2018).

2.3 *The progression of achievement gaps*

The findings discussed above are in line with the *cumulative (dis-)advantage* theory (DiPrete and Eirich, 2006), in which initial advantages or disadvantages accumulate over time and lead to growing differences in outcomes between these groups throughout the life course. While immigrant positive self-selection with respect to socioeconomic status is important, research also shows that the *institutional context* and settings in the receiving countries matter for immigrant integration (Cobb-Clark et al., 2012, Hillmert, 2013, Schneeweis, 2011).

Whether the native immigrant gap shrinks or widens over the course of education remains a largely unanswered question in the absence of longitudinal data. In the past two decades, there

has been a push to set up nationally representative longitudinal studies such as the German National Educational Panel Study (NEPS). However, longitudinal surveys are not available in most countries. Skopek and Passaretta (2020) exploit NEPS data to document the achievement gap between high- and low-SES students in Germany from childhood to adolescence and find that while the gap increases over time, education mitigates the learning inequalities driven by socioeconomic status. Dämmrich and Triventi (2018) examined a similar question in 15 OECD countries, using a pseudo-longitudinal approach comparable to that used in this study. However, Skopek and Passaretta (2020) did not explore the performance of immigrant children over time and largely relied on the implicit assumption of cardinal comparability between different surveys.

Vonnahme (2021) is the closest European study to ours. It documents the immigrant achievement gap change in Germany at different stages of the school career by leveraging the longitudinal structure of NEPS for three student cohorts. Vonnahme (2021) reports the existence of a significant achievement gap of migrant children in both mathematics and reading in Grade 5, the cohort of students who enrolled in primary school in 2012. Consistent with our findings, Vonnahme (2021) shows that immigrant achievement gaps narrow more modestly between Grades 5 and 10 in mathematics than in reading, without ever closing in either subject. However, in contrast to our study, they did not differentiate between different generational groups of immigrant students, limiting their focus to students with at least one parent born abroad. Furthermore, our student cohort differs from those used in Vonnahme (2021).

In sum, (i) there is international evidence of significant achievement gaps of immigrant-background children (measured at different points of the educational curricula), and (ii) the potential reasons for such gaps are numerous and possibly heterogeneous across institutional environments. In the absence of systematic, cross-nationally comparable longitudinal studies – especially outside of the US – it remains however difficult to find robust evidence about the *progression* of achievement gaps over time. Against this background, we seek to answer the following empirical questions: How does the raw and adjusted immigrant-native achievement gap in mathematics and reading evolve between primary and secondary education in Europe and the TIC? Do we observe differences in immigrant children's academic progress over time between countries with and without early academic tracking? Is the change in the immigrant achievement gap between elementary and secondary education smaller among second-generation and mixed families in contrast to the first generation?

3 Data and Methods

Our study documents the academic trajectories of immigrant children between primary and secondary schools in 12 European countries and four TIC: Australia, Canada, New Zealand, and the United States. In line with our broad research questions, we categorize these countries into two groups: TIC and European countries. We further categorize European countries as those that track students before the age of twelve (tracking countries) and those that do not or track at a later age (non-tracking countries). In this study, tracking countries include Austria, Belgium, Germany, Luxembourg, and the Netherlands, whereas non-tracking countries include Denmark, Norway, Sweden, France, Italy, Spain, and Ireland.

3.1 Data sources

Educational progress over time is best measured with longitudinal data. The academic trajectories of ethnic minorities between kindergarten and elementary school is well-documented in the US owing to the availability of a rich collection of educational surveys. However, to the best of our knowledge, no comparable data are available to explore the academic trajectories of immigrant-background children between primary and secondary schools in either TIC or in Western Europe, with the exception of NEPS in Germany, as mentioned earlier. The Children of Immigrants Longitudinal Study (CILS) in the US and the Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU), its European equivalent, could potentially be exploited to explore the academic progress of immigrant children between 14 and 17 years of age across five countries. Unfortunately, no measure of scholastic achievement was collected before the age of 14 in these surveys.

We therefore follow Hanushek and Woessmann (2006) and Hanushek et al. (2022) and build synthetic cohort datasets by pooling independent cross-sections from PIRLS, PISA and TIMSS—three widely used international surveys in educational research—to circumvent the absence of longitudinal data.[2] Pooling surveys allows us to follow the academic trajectories in reading and mathematics of pupils drawn from three different cohorts in a cross-country setting. Earlier studies by Schubert and Becker (2010), Jakubowski and Pokropek (2015), Ruhose and Schwerdt (2016) have exploited a similar strategy. None of these contributions, however, have focused explicitly on documenting the change in the immigrant achievement gap between primary and secondary schools.

PIRLS targets pupils attending the grade corresponding to 4 years of formal elementary education (Grade 4) and provides a measure of scholastic achievement in reading. TIMSS

[2]PIRLS, TIMSS and PISA data and data documentation are available online at `http://timssandpirls.bc.edu/` and at `http://www.oecd.org/pisa/`.

targets pupils attending Grades 4 and 8 and provides a measure of scholastic achievement in mathematics and science. Both surveys cover nationally representative samples of pupils over a large number of countries. TIMSS started in 1995 and is run every four years. PIRLS started in 2001 and is run every five years. PISA targets older students. It provides a measure of academic achievement at age 15 (without targeting a specific grade) in reading, mathematics and science. PISA started in 2000 and is run every three years [3].

By design, the population of students in Grade 4 targeted in PIRLS in a given year is approximately the population that is targeted in PISA five years later. So, tracking the performance of children in PIRLS in year t and in PISA in year $t + 5$ informs us of the academic trajectory of one particular cohort of children. We take advantage of this design to study the academic progress between the ages of 10 and 15 for two separate cohorts of pupils in reading.[4]

Cohort 1 pools cross-sectional data from PIRLS 2001 and PISA 2006, which includes students born in 1990 or 1991. Cohort 2 pools cross-sectional data from PIRLS 2006 and PISA 2012. Note that this second cohort is not perfectly aligned since there is a 6 years period between PIRLS and PISA: PIRLS 2006 covers students born in 1995–96, while PISA 2012 covers students born in 1996–97. We assume here that the performance of the latter provides a good approximation of the performance of children born one year earlier.

Finally, Cohort 3 leverages cross-sectional data from TIMSS 2015 and TIMSS 2019. For this cohort, the population of Grade 4 students in TIMSS 2015 was approximately the population targeted for Grade 8 students in TIMSS 2019. The primary motivation for including Cohort 3 is to provide a measure of change in performance in mathematics. The secondary benefit is to provide evidence of changes in academic performance from the most recent data available to carry out this exercise. Note, as discussed earlier, that Cohort 3 results should not be compared to Cohort 1 and 2, given the differences in academic subjects (reading vs. mathematics) and compositional change in the immigrant population over time.

Table 1 presents the countries included in each of those three cohorts. Students from a particular country are included if the country took part in the study in both years used to construct the synthetic cohort datasets.

We focus on three distinct groups of immigrant-background students. Following the traditional definition used by PISA, first-generation immigrant students (1G) include all foreign-born students of foreign-born parents. Second-generation immigrants (2G) include all students

[3] With the exception of PISA 2022 that was delayed by one year due to the COVID-19 pandemic

[4] Note that, because TIMSS and PIRLS target Grade 4 (and not necessarily children aged 10), we restrict our elementary school samples to children born in the expected year to attend Grade 4. That excludes kids having already repeated a grade by age 10 or having started elementary school early. We only exclude a small number of cases and our estimation results are not sensitive to this sample selection criteria. Results on the unrestricted sample are available upon request.

of foreign-born parents born in the country of assessment (OECD, 2023). We further define students with one parent born in the host country and one foreign-born parent as a distinct group (Mixed). This last grouping is motivated by evidence that intermarried immigrants tend to integrate better into host society's labour market, making their children an interesting pivotal group between foreign- and native-born (Meng and Gregory, 2005, Meng and Meurs, 2009, Furtado and Song, 2015, Elwert and Tegunimataka, 2016, Justiniano Medina and Valentova, 2023).[5] With intermarriage becoming more common, mixed background children make a sizeable group in each country of this study. Note that the cohort population is fixed at Grade 4: we exclude from the PISA samples all migrant children who settled in the host country after the age of 10.

Tables 2 and 3 present sample sizes and the percentage of students with a migratory background included in our analytical samples. The shares of immigrant background students – 1G, 2G or Mixed – vary greatly across countries and cohorts. Reassuringly, these shares are generally similar in both PIRLS/TIMSS and in PISA samples in each cohort—as they should be given our definition of the samples. However, when more important differences prevail, the shares of immigrant children in PISA samples tend to be smaller than in PIRLS/TIMSS. These differences may be driven by non-random selection of students from different groups into secondary school – i.e. dropout rates may differ between native and immigrant background students. This potential source of bias is not likely to be a major concern here since the PISA target population includes all students between 15 and 16 years old and the age of compulsory schooling is 16 in most countries in our study. Caution is however warranted in the few countries where compulsory schooling ends at the of 15 and for the mixed background students whose share appears to differ more systematically.[6]

3.2 Measures of academic performance

An ideal measure of performance is of course difficult to design. The difficulty is compounded here by the need to obtain a measure that is comparable over time, across groups of children and across countries. PIRLS, TIMSS and PISA aim to assess pupils knowledge, abilities and problem-solving skills in basic domains (reading and/or mathematics). As in traditional student assessment exercises, this is done by asking participants to provide answers to a range of test questions of varying complexity and covering various skill types. It is assumed that

[5]Mixed background students form a fairly heterogeneous group as it does not only include students from "intermarriages" with a native parent but also families with first-generation and second generation parents.

[6]PISA targets all students aged from 15 years and 3 months to 16 years and 2 months at the beginning of the assessment period in April (OECD, 2012). During the time covered by our study, compulsory schooling ends at the age of 15 only in Australia, Austria, Italy and Luxembourg (OECD, 2010).

participants have some (unobserved or latent) ability that is revealed by the test responses. Under the additional assumption that these latent abilities are normally distributed among study participants, individual measures of student ability are provided in the form of a personal standardized score per domain (reading and/or mathematics). Their computation follows a series of steps involving Item Response Theory modeling (Mislevy and Sheehan, 1989) and Rubin (1987)'s multiple imputation method (Adams et al., 2007, Monseur and Adams, 2009, Wu, 2005). In each survey, standardized student-level test scores have a mean of 500 and a standard deviation of 100 over the whole set of participating children across all countries.

Standardized test scores conveniently summarize participants abilities and make it easy – at least mechanically – to compare performance among respondents across different, say, countries, schools, tracks, etc. These are however only proxies of students abilities and academic performance. Furthermore, because they capture a latent construct, direct cardinal comparisons of such scores are hazardous: whether a pupil with a score of, say, 500 is twice as able as a child with a score of 250, or whether there is the same gap between children of scores 200 and 250 and 500 and 550 fully hinge on the assumption of normality. Yet this assumption is arbitrary since the latent abilities are unobserved by definition. To make things more complicated, because the universe of students covered varies over time and across surveys (because different countries are taking part), re-normalization of test scores is often applied: the ability score of each participant is re-normalized to have mean 500 and standard deviation 100 within each country separately. This national standardization changes the numerical differences in ability scores across students within each country. Whereas the rank of students in the ability distribution remains unaffected by these transformations – or by the assumptions about the latent score distribution – the actual scores vary in ways that depend on multiple factors.

As with previous related studies, the reliability of our design also rests on the assumption that the cognitive tests of PIRLS, PISA, and TIMSS are substantially and quantitatively comparable; see Ammermüller (2013), Hanushek and Woessmann (2006), Ruhose and Schwerdt (2016), Schnepf (2007), among others. While PIRLS, PISA and TIMSS are all designed to provide reliable aggregate measures of students' abilities for major subpopulations, they were not explicitly designed to be compared to each other. These surveys share, however, many similar features and it appears legitimate to compare aggregate test scores across studies. Analysis of the equivalence of item difficulty report 80 per cent commonality of the total variance in item difficulties in PIRLS and PISA, and also a high correlation between national results (Grisay et al., 2007, 2009). Some differences are found in performance between TIMSS and PISA in mathematics. However, they are not large enough to cause significant concerns over the robustness of comparative results (Wu, 2010). Brown et al. (2007) reach a similar conclusion while outlining that these differences are also sensitive to the survey years considered.

3.3 Quantifying the achievement gap between natives and children with an immigration background

Our objective is to appraise differences in test scores within each country across students with different migration background—the achievement gap—and the evolution of these differences as students progress from primary to secondary education. To do so, we rely on two summary measures of the achievement gap.

The simplest measure is the difference in *average* test score between native-born children and children with a migration background (1G, 2G or mixed)

$$\text{Diff} = \mu^n - \mu^f$$

where μ^n and μ^f are the average of (nationally standardized) test scores of native-born and immigrant-background children respectively. This measure has two drawbacks. First, focusing on the average score hides heterogeneity in the performance of students. Second, the average test score is only meaningful if one accepts the 'cardinality' of test scores as a measure of academic achievement. Although this is common practice, this is a strong assumption – for reasons we discussed above. US studies have raised concerns about the inter-temporal comparability of test scores when the achievement measures are inherently ordinal (Bond and Lang, 2013, Ho and Reardon, 2012, Reardon, 2008). For example, Bond and Lang (2013) find that up to 13% of the widening of the Black-White test gap between school entry and grade three can be due to test score sensitivity to scale transformation, raising questions about the reliability of the information conveyed in test score gap changes between grades.

We address this cardinality issue by considering a second measure of achievement gap. Our second measure is a transposition to test scores of an index originally proposed in Gastwirth (1975) to measure earnings differentials. Gastwirth's PROB index provides an ordinal measure of the achievement gap by comparing the relative *ranking* of students from distinct groups in the distribution of tests scores. The index measures the probability that a randomly selected student with migration background has a higher score than a randomly selected native student. If there were no systematic deviation in the test scores of the two groups, this probability would be 0.5 (a random immigrant background student would be equally likely to do better or worse than a random native student). On the other hand, if *all* immigrant background pupils did worse than *all* native pupils, the probability would be zero. It would be one in the other extreme case of seeing all immigrant background students doing better than all native students.[7] The value of the PROB index is determined entirely by (differences in) the *ranks* of students

[7]Formally, the PROB index can be written

of the two groups in the overall (national) distribution of scores. Therefore, any monotonic transformation of the individual scores – by rescaling, translation, re-normalization, etc. – that leaves ranks unchanged in the national distribution has no impact on PROB. This is crucial in the context of test scores derived under an assumption of normality of a latent score distribution in cross-national survey instruments, such as those available in PIRLS, TIMSS and PISA. While deviations from normality within countries or any arbitrary re-standardization applied to each country's test score distribution may affect the value of the Diff index (and therefore assessments of its change over time), these do not change students' ranks in their own country distribution and therefore do not influence assessment based on the PROB index. This measure of achievement gap between native and immigrant background pupils is not impacted by the potential cardinal comparability bias of test scores. While such a measure may not be an 'ideal' measure of ability and ability differences – which arguably would allow cardinal comparisons in the quantification of the size of the gap – the ordinal approach avoids the pitfalls involved in using test scores at face value when scores are drawn from different countries and different survey instruments. Whatever index is considered (Diff or PROB), the change in achievement gap over time for a cohort of children is given by the change between measurements at age 10 (in primary school) and at age 15 (in secondary school)

$$\Delta_D = \text{Diff}^{15} - \text{Diff}^{10}$$

or

$$\Delta_P = \text{PROB}^{15} - \text{PROB}^{10}.$$

$\Delta_D > 0$ or $\Delta_P < 0$ indicates that the academic performance of immigrant background students is deteriorating over time compared to their native-born peers.

$$\text{PROB} = \int F^n(s) f^f(s) ds$$

where F^n is the cumulative probability distribution of scores among native pupils and f^f is the probability density function of scores among immigrant background pupils; see Gastwirth (1975) for methodological details, Le Breton et al. (2012) for a more recent discussion, Zhou (2012) for a re-expression of the index as a measure of stratification, and Ho (2009) for an application to test scores. With individual data on test scores for N students, PROB can be estimated as

$$\widehat{\text{PROB}} = \frac{1}{N^f} \sum_{i=1}^{N} I_i \, \hat{F}^n(s_i)$$

where s_i and I_i are student i's test score and migration background (I_i is 1 for immigrant background and 0 otherwise), $N^f = \sum_{i=1}^{N} I_i$ is the number of immigrant background students, and $\hat{F}^n(s_i) = \frac{1}{N^n} \sum_k^N (1 - I_k) \mathbb{1}(s_k \leq s_i)$ is the share of native students with a score lower than s_i.

3.4 Adjusting for differences in socioeconomic background

Differences in achievement between the group of immigrant background and native children reflect, at least in part, differences in socioeconomic background. For a discussion of the impact of migration background on academic achievement, there is therefore interest in examining measures that are adjusted for those differences, that is, measures that compare the achievement of pupils with migration background to the achievement of native pupils with similar socioeconomic background.

With access to individual data on test scores and socioeconomic characteristics, reweighting or 'direct standardization' methods offer an easy way to achieve this (DiNardo et al., 1996, Fortin et al., 2011). To calculate adjusted Diff or PROB measures, we need estimates of a counterfactual distribution of test scores among natives "if they had the characteristics of immigrant background students," which we denote F^c. Substituting F^n by F^c in the calculation of achievement gap measures provides 'adjusted' gap measures that reflect differences in achievement that are *not* explained by mere differences in observable SES characteristics. A direct standardization approach to constructing the latter involves reweighting native pupils in such a way that the socioeconomic characteristics of the reweighted sample of native pupils has the same characteristics as the sample of immigrant background pupils. The intuition is that characteristics x relatively more common among immigrant background children receive a weight larger than one while characteristics relatively rare receive a weight closer to zero. At one extreme, characteristics absent among immigrant background pupils receive a weight of zero and native students with those characteristics are effectively discarded in the reweighted sample of native children.[8]

[8]Formally, F^c is defined as follows:

$$F^c(s) = \int_{\Xi} F^n(s|X = x)h^f(x)dx$$

where $F^n(s|X = x)$ is the distribution of test scores among native children with characteristics x and $h^f(x)$ is the probability density of observing characteristics x among immigrant background children. The reweighting argument is that F^c can be expressed as

$$F^c(s) = \int_{\Xi} \Psi(x)F^n(s|X = x)h^n(x)dx$$

with the reweighting function $\Psi(x) = \frac{h^f(x)}{h^n(x)}$ defined as the ratio of densities of characteristics in the two groups. The counterfactual distribution is then easily estimated from the sample data as

$$\hat{F}^c(s) = \frac{1}{N^c}\sum_{i=1}^{N^n} \Psi(x_i)\mathbb{1}(s_i \leq s)$$

where $N^c = \sum_i^{N^n} \Psi(x_i)$. DiNardo et al. (1996) show that, by applying Bayes' rule, $\Psi(x)$ can be calculated as $\frac{\Pr(f|X=x)}{1-\Pr(f|X=x)}\frac{1-\Pr(f)}{\Pr(f)}$ with $\Pr(f|X = x)$ being the probability that a pupil with characteristics x has foreign background and $\Pr(f)$ the share of foreign background students in the sample. Each term can be estimated using,

Note that we 'reweight' native pupils towards immigrant background students and not the other way round. There are two reasons for this. The first is notional: we are interested in the progress of immigrant background students against native pupils with similar characteristics. Reweighting them to the native pupils characteristics would lose representativity of our group of interest. The second is technical: the reweighting technique requires that all possible patterns of characteristics found in the target group of interest are also found in the reference (reweighted) group—otherwise it is not possible to match the target population characteristics by reweighting. This condition is more often satisfied when the larger and more heterogeneous population—here the natives—is reweighted 'down' to the smaller population (Hildebrand et al., 2017).

We estimate the adjustment weights using a set of commonly used student and family background characteristics available in both PIRLS/TIMSS and PISA: student's gender, parental education, the number of books at home and whether students have a computer and a desk to study on. The choice of *books-at-home* as an additional proxy for family background characteristics is validated by the number of studies which consistently showed that it constitutes the most powerful predictor of children educational achievement, independent of parental social status; see Schütz et al. (2008), Woessmann (2008), Ammermüller and Pischke (2009), Evans et al. (2010), Pedro et al. (2013), Brunello et al. (2016) for further discussion. However, in view of its endogeneity and student reporting bias when quantifying the number of books (Engzell, 2021), we also account for the educational level of parents. Including both the number of books and parental education limited the number of survey years that could be exploited to build synthetic cohorts because of the lack of information on parental education for some survey years (for example, TIMSS 2003 and 2007 in Grade 4).

Tables 4 and 5 show the summary statistics of those variables in the pooled samples: 1G and 2G students are more likely to have fewer books at home and as likely to have their own computer in secondary school.

We denote the adjusted difference in average test score and the Gastwirth index by Diff^U and PROB^U where the U superscript indicates that they measure an achievement gap that is 'unexplained' by differences in socioeconomic background. Diff^U measured at age 15 ($\text{Diff}^{U,15}$) is comparable to the measure of educational *integration* of Schneeweis (2011): the distance separating the educational performance of immigrant children from that of native-born that is not explained by differences in student characteristics and family background. Our measures of progress adjusted for socioeconomic background are then

$$\Delta_D^U = \text{Diff}^{U,15} - \text{Diff}^{U,10}$$

e.g., logistic regression models. See Hildebrand et al. (2017) for an application to analysis of deprivation levels of immigrants.

and

$$\Delta_P^U = \text{PROB}^{U,15} - \text{PROB}^{U,10}.$$

and each measure captures the change in the immigrant achievement gap between 10 and 15 years old that cannot be explained by student and family characteristics. A narrowing of the unexplained part over time would suggest that the return to endowment of migrants are converging to that of natives. It can be viewed as a measure of educational assimilation.

4 Results

Given the concerns about the cardinal comparability of test scores obtained from surveys such as PIRLS, TIMSS and PISA, we focus our discussion on Gastwirth's PROB index results – summarized graphically in Figures 1 to 4 for reading, and in Figures 5 to 8 for mathematics – as these only require assuming ordinal comparability of individual scores.

Each figure displays the raw (observed) and adjusted gap point estimates with 95% confidence interval error bars. As a reminder, the adjusted gap measures the portion of the immigrant achievement gap unaccounted for by between-group differences in observable characteristics. Given the popularity of measures assuming cardinal comparability, we also obtained estimates from those measures.[9] Complete estimation results are reported in Appendix in Tables A.1 to A.4 for reading and Tables B.5 to B.8 for mathematics.[10] Reassuringly, the direction and magnitude of changes in the immigrant achievement gap implied by either measure of academic progress yield comparable results.

In each of Figures 1 to 8, the *x*-axis marks the achievement gap in primary school (grey circle symbol) and secondary school (black triangle symbol) with corresponding countries on the *y*-axis. Point estimates in countries where migrant-origin children would perform equally well than their native peers – either in primary or secondary school – would lay on the vertical axis at $x = 0.5$. Gastwirth's PROB index estimates displayed to the left (right) of the vertical axis at $x = 0.5$ indicate that the academic performance of immigrant-background children lag behind (outperform) their native peers. The change in the achievement gap between elementary and secondary school is visually observed for each country by the horizontal distance separating the black triangle symbol (immigrant achievement gap in secondary school) from the grey circle symbol (immigrant achievement gap in elementary school).

[9]Effect size statistics commonly used in educational research (e.g., Cohen's q) implicitly assume cardinal comparability. See Nielsen (2015) for a recent comprehensive discussion.

[10]Each table reports the observed (Raw) and adjusted (Adj) gaps in Grade 4 – primary school –, at age 15 (or age 14 for Cohort 3) – secondary school – and the resulting change in gaps over time (ΔGap) using our two measures of achievement – normalized standardized test scores and Gastwirth's PROB index with – with significance level at the at 1, 5 and 10 percent levels, respectively.

Overall, regardless of the academic subject considered, Figures 1, 2, 5 and 6 reveal that the academic performance of immigrant–background children lag behind their native-born peers in most European countries. This finding corroborates numerous studies that have consistently documented that 1G and 2G children lag behind their native peers (in both reading and mathematics); see, among many others, Schneeweis (2011), Riederer and Verwiebe (2015). However, most of these studies did not measure academic progress over time.[11] By contrast, in traditional immigration countries – Australia, Canada, New Zealand and the United States – migrant-origin children appear to perform at least as well as their native peers by the age of 15 (by the age of 14 for students in Cohort 3).

Accounting for differences in student characteristics narrows the achievement gap in almost all countries, as evidenced by a horizontal rightward shift of point estimates towards the vertical axis at x = 0.5 for most countries in the "Adjusted gap" plots. A sizeable portion, however, often remains unaccounted for, and the size of these unexplained differences, at either level of schooling, varies greatly across countries and cohorts within countries. This observation underscores the existence of a large diversity between immigrant populations across receiving countries.

Academic Progress in Reading

In traditional immigration countries (Canada, New Zealand, and the United States in our reading sample), 1G and 2G pupils generally perform at least as well as their native counterparts by age 15, depending on the cohort considered (see Figures 1 and 2). Note, however, the existence of a large gap among 1G fourth-graders in Canada (Cohort 1) and the United States (Cohort 2), which fully disappears by the time students reach secondary school. Graphically, the confidence intervals of the raw and adjusted point estimates in secondary school cross the vertical line at $x = 0.5$ suggesting the absence of a statistically significant gap between 1G students and their native peers by age 15 (see Figure 1). We find a similar pattern for 1G students in the United States, demonstrating significant academic progress between the ages of 10 and 15. In contrast, our figures reveal more modest academic progress in reading among both 1G and 2G pupils in most European countries.

With the exception of Cohort 1 in France and Norway and Cohort 2 in Austria and Luxembourg, 1G pupils in Europe generally exhibit large achievement gaps in Grade 4, which improve over time without never being fully absorbed by age 15. This translates graphically by a point estimate in secondary school to the left of the vertical bar at $x = 0.5$ without the 95%

[11] Jakubowski and Pokropek (2015) and Schnepf (2007) are a notable exceptions. However, no attempt is made to follow the educational progress of a particular cohort over time.

confidence interval crossing it (see Figure 1).[12]

The positive educational outcomes of 1G children in TIC is often attributed to selective migration policies, which explicitly target high skilled parents (Buchmann and Parrado, 2006, Entorf and Minoiu, 2005, Levels et al., 2008). Cattaneo and Wolter (2015) provide further evidence supporting this view in a study which reported a significant increase in test scores among immigrant children in Switzerland, following a policy change in the mid-1990s, which triggered a large exogenous increase in its share of high-skilled migrants. This finding supports the view that students' socioeconomic capital at the time of migration (the initial endowment) is an important determining factor of future academic success into the host country school system.

By the same token, our finding of more modest academic progress among 1G pupils in Europe compared to TIC is not fully surprising, given that immigrant children in Europe are usually more drawn heavily from families of unskilled labour or refugees (Sakellariou, 2018). Educational disparities with their native peers may, therefore, be a mere reflection of the additional difficulties faced by immigrant-background children from families with weak socioeconomic backgrounds to integrate smoothly into the host country's school system (Levels et al., 2008, Cattaneo and Wolter, 2015). Country-specific obstacles inherent to their school systems, which are also perhaps less prepared (or less willing) to meet the needs of increasing (and unplanned) flows of immigrant children, may also be at play. Levels et al. (2008), however, do not find robust evidence supporting this hypothesis.

Results from countries in which we observe two cohorts of children reveal that the size of the achievement gap is sensitive to the student cohort. For instance, while 1G children from Cohort 1 in Norway, and France no longer show significant differences in achievement with observationally comparable natives by the age of 15, their counterparts from Cohort 2 still experience a non-negligible statistically significant gap. While this finding is indicative of stagnant academic progress for the most recent cohort of 1G children in these countries, comparing results from Cohort 1 and 2, as pointed out in section 3.1, is difficult due to the confounding effects of compositional changes in the migrant populations over time and changes in migration policies.

Finally, Figure 1 does not show a systematic negative association between academic tracking and the academic progress of 1G children. For instance, the point estimate of the adjusted gap at age 15 in Germany is smaller than that in Italy and Sweden for Cohort 1 (and not statistically different), whereas for Cohort 2, Austria and Luxembourg are the only European countries where immigrant children appear to be closing the achievement gap by the age of 15.

[12]Cohort 2 from Denmark is a notable exception with a large and statistically significant increase in the achievement gap of 1G-pupils over time ($\Delta_P^U = -0.13$).

[Figure 1 about here]

The estimated academic progress of 2G children, displayed in Figure 2 provides interesting insights. 2G are children expected to better transition to primary education than their 1G peers as a result of having experienced from the onset the benefits of early education programs in their parent's destination country.[13] In line with expectations, 2G children appear to be closing the adjusted achievement gap in reading in more European countries than 1G children, regardless of the cohort considered, as evidenced by an adjusted gap not statistically different from the point of equal performance with native children at PROB = 0.5 in France, Italy, Sweden, and Norway (Cohort 1), and Luxembourg, Austria, Spain, Italy, Sweden, and Norway (Cohort 2). Interestingly, countries in which 2G students do not close the achievement gap in reading by the age of 15 are overwhelming those that track students at an early age: Netherlands, Germany (Cohort 1), and the Netherlands, Germany, and Belgium (Cohort 2). This finding is fully in line with Azzolini et al. (2012), van de Werfhorst and Heath (2019). By contrast, in line with prior expectations, students in TIC do equally well, or outperform their native peers.

[Figure 2 about here]

We further examine the academic progress of 2G pupils by contrasting their scholastic achievements with their 1G peers. Results in reading are displayed in Figure 3 and also reported in Table A.3). Consistently with prior research, 1G fourth graders from Cohort 1 generally perform more poorly than their 2G peers in almost all European countries after accounting differences in observable characteristics (except for the Netherlands where first-generation 4th graders outperform second-generation). We find broadly similar results for Cohort 2, but the adjusted gaps and changes in the adjusted gaps between primary and secondary schools are often statistically insignificant (Table A.3). This lack of statistical significance may be due to lack of power due the modest size of our immigrant samples. Despite this caveat, we view our results as providing little supporting evidence of significant educational progress of 2G children compared to their 1G peers in many European countries included in our sample, both within and across cohorts.

[Figure 3 about here]

We now turn to the results for the mixed-migrant children. Mixed migrant pupils outperform their 1G and 2G peers, as evidenced by a smaller (or insignificant) achievement gap in Grade 4, which fully disappears by the age of 15 in almost all countries (see Figure 4). These results are in line with evidence that children of intermarried parents have more social contact with

[13] Remember that our sample only considers 1G children who migrated before the age of 10.

natives, better language skills, and information about local institutions and customs than their 1G and 2G counterparts (Kalmijn, 2015). Information on the educational achievement of mixed-migrant children is scarce. Available evidence shows that their educational achievement lies somewhere between that of immigrant children (1G or 2G) and natives (Levels et al., 2008, Kalmijn, 2015). Interestingly, our results show that in most countries, the reading proficiency of mixed-migrant pupils at age 15 does not differ significantly from that of native-born pupils in all but three countries that track students early: Germany in both cohorts and Belgium and the Netherlands for Cohort 2. Our results for Germany and the Netherlands corroborate Kalmijn (2015). Finally, we do not observe significant differences between the educational achievements of mixed children in European countries that do not track students at an early age and TIC.

[Figure 4 about here]

Academic Progress in Mathematics

Mathematics results are drawn from the most recent cohort included in our study of children who entered Grade 4 in 2015, allowing us to document the academic progress of children who entered Grade 9 the year of the COVID-19 pandemic avoiding the period of online teaching imposed by stay-at-home orders in most countries covered by our sample. They reveal more striking differences between Europe and traditional immigration countries than in reading, with the exception of Ireland, where the experience of migrant-origin children appears more comparable to their peers in TIC (see Figures 5 to 8).

Neither 1G nor 2G immigrant children in TIC show any significant disparities in achievement with their native-born peers in elementary school, except for the United States, and outperform them significantly in all countries by Grade 8. 1G and 2G children in continental Europe exhibit more contrasting results. Both 1G and 2G children exhibit large and persistent achievement gaps in France and Sweden whereas immigrant-background children are closing the achievement gap in Italy and Norway (see Figure 6). Unfortunately, Cohort 3 does not include any European countries that tracks students early.

[Figure 5 about here]

[Figure 6 about here]

To further investigate these results, we again contrast the achievement of 2G children with that of their 1G peers (see Figure 7 and Table B.7). Consistent with prior research, 2G children are generally performing better at age 10 than their 1G peers in all European countries. However, as for reading, their (adjusted) educational advantage tends to narrow by age 14 in all

countries except Italy. Although the estimated changes are not always statistically significant, they provide additional evidence suggesting a slower rate of educational progress among 2G immigrant children. To some extent, these children appear to have hit some achievement glass ceilings. These lackluster results in mathematics may partly reflect the cumulative effect of poor reading skills in the early grades (Andon et al., 2014).

[Figure 7 about here]

As for reading, mixed-migrant children in Europe appear to be performing as well as their native counterparts except for France (where they lag behind native) and Ireland (where they outperform native) with no significant differences between Grade 4 and Grade 8. By contrast, mixed children in the TIC outperform their native peers by Grade 8 in both Australia and New Zealand.

[Figure 8 about here]

5 Discussion and conclusion

The goal of this paper is to provide empirical answers three main questions: Does the immigrant-native achievement gap, both raw and adjusted, in mathematics and reading change between primary and secondary education in Europe and the TIC? Do we observe discernible differences in immigrant children's progress over time between countries with vs. without early academic tracking in secondary education? Does the second-generation and mixed-origin group display a smaller immigrant achievement gap between elementary and secondary education in contrast to the first generation?

To answer these questions, the study uses independent cross-sections from PIRLS, TIMSS, and PISA and constructs three cohorts of students between the approximate age of 10 and 15 (14). It leverages a versatile method that exploits an ordinal measure of the achievement gap in reading and mathematics. This approach allows us to circumvent the untestable cardinality assumption of test scores from independent surveys to measure the change in academic performance over time in absence of longitudinal data. However, these data have certain limitations that need to be acknowledged before further discussing our results. Firstly, several studies have identified the importance of parental background in explaining differences in scholastic achievement between various ethnic groups within and between countries (Azzolini and Barone, 2013, Bygren and Szulkin, 2010, Jackson et al., 2012, Jonsson and Rudolphi, 2011, Levels et al., 2008). These factors include linguistic, cultural, and religious distance between immigrant and native groups (Fleischmann and Dronkers, 2010, Isphording, 2014, Isphording et al., 2016, Jonsson and

Rudolphi, 2011, Levels and Dronkers, 2008), the pre-migration context (measured through political stability and democratic tradition), economic prosperity in the country of origin (Levels et al., 2008), community concentration, and residential segregation in the host country (Bygren and Szulkin, 2010). Several other studies have identified the importance of ethnic origin and/or race in explaining heterogeneous education and labour market integration outcomes in TIC (Boyd, 2002, Gamoran et al., 2012), Europe (Levels et al., 2008, Fleischmann and Dronkers, 2010, Bygren and Szulkin, 2010, Cheung, 2014, Kalter and Kogan, 2006), and other destination countries (van de Werfhorst and Heath, 2019). This is an important dimension that could not be explored in this study due to the absence of data about the country of origin of parents and children in TIMSS and PIRLS.[14] Next, the data do not allow us to differentiate the academic trajectories of the children of economic migrants from those of refugees. This limitation may undermine our identification of the academic trajectories of immigrant children in countries with a large share of refugees in their immigrant population (Schipolowski et al., 2021), cognizant that both groups have been found to experience significantly different integration outcomes (Brell et al., 2020, Kanas and Steinmetz, 2020).

Lastly, as for any classification, the immigrant groups used in this study is not immune to criticism. In particular, what we refer to as mixed-origin children (Mixed) may include families with very different ethno-cultural backgrounds, such as intermarried couples, where the parent born in the destination country can hypothetically also be a 3rd generation immigrant that shares common ethnic/racial ancestry with the grandparents of its foreign-born spouse. We may hypothesize that their children may experience a different academic trajectory than children from intermarried couples between one native-born parent whose family resides in the destination country of its foreign-born partner for many generations. While these two groups of mixed children may experience a different upbringing and/or socioeconomic conditions, we do not see this as a major limitation in this study given our finding that the trajectory of mixed children is not significantly different from the academic trajectories of children from native-born families.

Overall, our analyses show a clear and consistent pattern. The academic achievement of 1G and 2G children in TIC converges to that of natives in reading and even surpasses them in mathematics. This finding contrasts with the more challenging experiences of immigrant children (both 1G and 2G) in European countries, where we find evidence of a more strenuous journey to close the achievement gap between elementary and secondary schools in many countries. Furthermore, in line with the existing literature, we find evidence of a negative association between early tracking and the academic progress of 2G migrants in reading (van de

[14]The sample size of our immigrant groups in most countries in our sample is probably also an issue had this information be available

Werfhorst and Heath, 2019).[15]

The limited academic progress of children with an immigration background relative to natives in European countries – not only of first but also of second-generation – is somewhat disappointing and a cause for concern. This points to a persistent disadvantage with potential long-lasting impacts since skills acquired in schools are key to adulthood outcomes. However, over the last decade, many European countries have engaged in major investments to address the educational needs of economically vulnerable children, including immigrant students. For instance, Germany improved access to quality Kindergarten programs, provided additional language support to immigrant children and their parents (Anderson et al., 2015, Klein and Becker, 2017), and began offering more flexible possibilities for moving between school tracks (Riederer and Verwiebe, 2015). Likewise, Austria, Belgium, Denmark, Ireland, the Netherlands, Norway, and Sweden introduced several measures to address the specific needs of immigrant children, including language skills screening among preschoolers, compulsory kindergarten enrollment, additional school teaching time, training on intercultural pedagogy, and language support throughout school, and provided extra funding to schools with a higher concentration of immigrant children (see, Nusche et al., 2009, 2010, Riederer and Verwiebe, 2015, Shewbridge et al., 2010, Taguma et al., 2009, 2010, for further details).

Evaluating the impact of such policies is beyond the scope of this paper, and we purposely avoid comparing the academic progress of immigrant children drawn from different cohorts due to the confounding effects of compositional changes in immigrant populations and migration policies over time. However, it is worth mentioning that Levels et al. (2008) did not find evidence supporting that the better educational achievement of migrant children in TIC could be explained by educational policies specifically targeting the needs of immigrant children using data from PISA 2003. Anderson et al. (2015) find some evidence that the reforms in Germany improved the equality of opportunities of the targeted population at the bottom of the achievement spectrum over the 2003–2009 period. However, they did not implicitly address the impact of these reforms on immigrant children's educational trajectories.

Our finding of diverging educational progress between immigrant children in TIC and our sample of European countries reinforces the importance of the initial socioeconomic endowment in shaping the academic trajectories of immigrant children. The selective migration policies of the TIC provide a valuable design for gauging the effect of parental background (children's initial educational endowment) at the time of migration on the achievement gap. With this in mind, our findings of slower academic progress in Europe may simply underscore the role of their migrant population's socioeconomic endowment at the time of migration (Ammermüller,

[15] Unfortunately, we do not have data to document the academic progress of migrant students in countries that track students early for Cohort 3.

2007, Levels et al., 2008, Cattaneo and Wolter, 2015), rather than being a by-product of European school systems' inability to accommodate their idiosyncratic needs.

Overall, these findings reinforce the importance of the cumulative effect of parental background at the time of migration (country of origin, parental socioeconomic background) in shaping the educational trajectories of immigrant-background children in destination countries, as between-group achievement differentials at age 15 (14) remain correlated with the performance gap in Grade 4.

The presence of a large achievement gap already in primary education suggests, in line with other studies, that inequalities in achievement are generated from younger ages (Passaretta et al., 2022). Equalizing learning opportunities for disadvantaged children before the start of obligatory education could be the most effective policy intervention. Quality early childhood education and care (ECEC) is considered the most promising measure for supporting educational achievement and attainment (Heckman, 2006). Among immigrant children attending the ECEC, the achievement gap to native students in PISA was reduced by 55 points compared to immigrant peers who did not attend pre-school education (OECD/EU, 2018)

Another potential policy that may be considered in an attempt to reverse the immigrant achievement gap is to reassess the ways in which ability grouping is implemented in secondary education. Our results again show, in line with other studies, that the achievement gap is smaller in countries with late tracking. In addition, other studies suggest that course-by-course ability grouping is less detrimental to socioeconomic inequality in achievement than academic/vocational ability grouping (Chmielewski, 2014).

We examined progression from primary education to what would be approximately the middle of secondary education in most countries. A natural extension of this research will be to relate the progression observed all the way through to final educational achievements of immigrant-background children – to post-secondary and tertiary education, study choices (e.g., in STEM and ICT) and early career – so as to fully assess how immigrant-background children's disadvantage evolves over the course of educational systems and shape long-term outcomes.

References

Adams, R. J., M. L. Wu, and C. H. Carstensen (2007). *Application of Multivariate Rasch Models in International Large-Scale Educational Assessments*, pp. 271–280. New York, NY: Springer New York.

Alba, R. and N. Foner (2015). Mixed unions and immigrant-group integration in North America and Western Europe. *The ANNALS of the American Academy of Political and Social Science 662*(1), 38–56.

Alba, R., J. Sloan, and J. Sperling (2011). The integration imperative: the children of low-status immigrants in the schools of wealthy societies. *Annual Review of Sociology 37*(1), 395–415.

Alesina, A., M. Carlana, E. L. Ferrara, and P. Pinotti (2018). Revealing stereotypes: evidence from immigrants in school. Technical report, Harvard University.

Algan, Y., C. Dustmann, A. Glitz, and A. Manning (2010). The economic situation of first- and second-generation immigrants in France, Germany, and the United Kingdom. *Economic Journal 120*(542), F4–F30.

Ammermüller, A. (2007). Poor background or low returns? Why immigrant students in Germany perform so poorly in the Programme for International Student Assessment. *Education Economics 15*(2), 215–230.

Ammermüller, A. (2013). Institutional features of schooling systems and educational inequality: Cross-country evidence from PIRLS and PISA. *German Economic Review 14*(2), 190–213.

Ammermüller, A. and J.-S. Pischke (2009). Peer effects in European primary schools: Evidence from the Progress in International Reading Literacy Study. *Journal of Labor Economics 27*(3), 315–348.

Anderson, G., T. Fruehauf, M. G. Pittau, and R. Zelli (2015). Evaluating progress toward an equal opportunity goal: assessing the German educational reforms of the first decade of the 21st century. Working Paper 552, University of Toronto, Department of Economics.

Andon, A., C. G. Thompson, and B. J. Becker (2014). A quantitative synthesis of the immigrant achievement gap across OECD countries. *Large-scale Assessments in Education 2*(1).

Antony-Newman, M. (2019). Parental involvement of immigrant parents: a meta-synthesis. *Educational Review 71*(3), 362–381.

Australian Bureau of Statistics (2024). Overseas migration (2022-23-financial-year). `https://www.abs.gov.au/statistics/people/population/overseas-migration/latest-release`.

Azzolini, D. and C. Barone (2013). Do they progress or do they lag behind? Educational attainment of immigrant's children in Italy: The role played by generational status, country of origin and social class. *Research in Social Stratification and Mobility 31*, 82 – 96.

Azzolini, D., P. Schnell, and J. R. B. Palmer (2012). Educational achievement gaps between immigrant and native students in two 'new' immigration countries: Italy and Spain in comparison. *The ANNALS of the American Academy of Political and Social Science 643*(1), 46–77. PMID: 23493944.

Becker, B., C. Gresch, and T. Zimmermann (2023). Are they still aiming high? The development of educational aspirations of lower secondary school students with immigrant backgrounds in Germany. *International Migration Review 3*(57), 1216–1248.

Bernardi, F. and G. Ballarino (2016). *Education, occupation and social origin: A comparative analysis of the transmission of socio-economic inequalities*. Edward Elgar Publishing.

Biroli, P., D. D. Boca, J. J. Heckman, L. P. Heckman, Y. K. Koh, S. Kuperman, S. Moktan, C. D. Pronzato, and A. Ziff (2017, May). Evaluation of the reggio approach to early education. Working Paper 23390, National Bureau of Economic Research.

Bol, T. and H. G. van de Werfhorst (2013). Educational systems and the trade-off between labor market allocation and equality of educational opportunity. *Comparative Education Review 57*(2), 285–308.

Bol, T., J. Witschge, H. G. van de Werfhorst, and J. Dronkers (2014). Curricular tracking and central examinations: Counterbalancing the impact of social background on student achievement in 36 countries. *Social Forces 92*(4), 1545–1572.

Bond, T. N. and K. Lang (2013). The evolution of the black-white test score gap in grades K-3: the fragility of results. *Review of Economics & Statistics 95*(5), 1468 – 1479.

Borgna, C. (2016). Multiple paths to inequality. how institutional contexts shape the educational opportunities of second-generation immigrants in Europe. *European Societies 18*(2), 180–199.

Boyd, M. (2002). Educational attainments of immigrant offspring: Success or segmented assimilation? *International Migration Review 36*(4), 1037–1060.

Breen, R. and J. O. Jonsson (2005). Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annu. Rev. Sociol. 31*, 223–243.

Breen, R., R. Luijkx, W. Müller, and R. Pollak (2009). Nonpersistent inequality in educational attainment: Evidence from eight european countries. *American journal of sociology 114*(5), 1475–1521.

Brell, C., C. Dustmann, and I. Preston (2020, February). The labor market integration of refugee migrants in high-income countries. *Journal of Economic Perspectives 34*(1), 94–121.

Brooks-Gunn, J., P. K. Klebanov, and G. J. Duncan (1996). Ethnic differences in children's intelligence test scores: Role of economic deprivation, home environment, and maternal characteristics. *Child Development 67*(2), 396–408.

Brown, G., J. Micklewright, S. V. Schnepf, and R. Waldmann (2007). International surveys of educational achievement: how robust are the findings? *Journal of the Royal Statistical Society: Series A (Statistics in Society) 170*(3), 623–646.

Brunello, G. and D. Checchi (2007). Does school tracking affect equality of opportunity? new international evidence. *Economic Policy 22*(52), 782–861.

Brunello, G., G. Weber, and C. T. Weiss (2016, 04). Books are Forever: Early Life Conditions, Education and Lifetime Earnings in Europe. *The Economic Journal 127*(600), 271–296.

Buchmann, C. and E. A. Parrado (2006). Educational achievement of immigrant-origin and native students: A comparative analysis informed by institutional theory. *International Perspectives on Education and Society 7*, 335–366.

Bukodi, E. and J. H. Goldthorpe (2013). Decomposing 'social origins': The effects of parents' class, status, and education on the educational attainment of their children. *European sociological review 29*(5), 1024–1039.

Burger, K. (2016). Intergenerational transmission of education in europe: Do more comprehensive education systems reduce social gradients in student achievement? *Research in Social Stratification and Mobility 44*, 54–67.

Bygren, M. and R. Szulkin (2010). Ethnic environment during childhood and the educational attainment of immigrant children in Sweden. *Social Forces 88*(3), 1305.

Cattaneo, M. A. and S. C. Wolter (2015). Better migrants, better PISA results: Findings from a natural experiment. *IZA Journal of Migration 4*(1), 1–19.

Cerna, L., O. Brussino, and C. Mezzanotte (2021, December). The resilience of students with an immigrant background. *OECD Education Working Papers* (261).

Cheung, S. Y. (2014). Ethno-religious minorities and labour market integration: generational advancement or decline? *Ethnic and Racial Studies 37*(1), 140–160.

Chmielewski, A. K. (2014). An international comparison of achievement inequality in within- and between-school tracking systems. *American Journal of Education 120*(3), 293–324.

Cobb-Clark, D. A., M. Sinning, and S. Stillman (2012). Migrant youths' educational achievement: The role of institutions. *The Annals of the American Academy of Political and Social Science 643*(1), 18–45.

Crul, M. (2013). Snakes and ladders in educational systems: Access to higher education for second-generation Turks in Europe. *Journal of Ethnic and Migration Studies 39*(9), 1383–1401.

Dämmrich, J. and M. Triventi (2018). The dynamics of social inequalities in cognitive-related competencies along the early life course–a comparative study. *International Journal of Educational Research 88*, 73–84.

DiNardo, J., N. M. Fortin, and T. Lemieux (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica 64*(5), 1001–1044.

DiPrete, T. A., T. Bol, C. C. Eller, and H. G. van de Werfhorst (2017, May). School-to-work linkages in the United States, Germany, and France. *American Journal of Sociology 122*(6), 1869–1938.

DiPrete, T. A. and G. M. Eirich (2006). Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. *Annual Review of Sociology 32*, 271–297.

Dollmann, J. (2021). Ethnic inequality in choice- and performance-driven education systems: A longitudinal study of educational choices in England, Germany, the Netherlands, and Sweden. *The British Journal of Sociology 72*(4), 974–991.

Dustmann, C., P. A. Puhani, and U. Schönberg (2017). The long-term effects of early track choice. *The Economic Journal 127*(603), 1348–1380.

Elwert, A. and A. Tegunimataka (2016). Cohabitation premiums in Denmark: Income effects in immigrant-native partnerships. *European Sociological Review 32*(3), 383–402.

Engzell, P. (2021). What do books in the home proxy for? a cautionary tale. *Sociological Methods & Research 50*(4), 1487–1514.

Engzell, P. and I. J. Raabe (2023). Within-school achievement sorting in comprehensive and tracked systems. *Sociology of Education 96*(4), 324–343.

Entorf, H. and N. Minoiu (2005). What a difference immigration policy makes: A comparison of PISA scores in Europe and traditional countries of immigration. *German Economic Review 6*(3), 355–376.

Erikson, R. and J. H. Goldthorpe (2002). Intergenerational inequality: A sociological perspective. *Journal of Economic Perspectives 16*(3), 31–44.

Erola, J., S. Jalonen, and H. Lehti (2016). Parental education, class and income over early life course and children's achievement. *Research in Social Stratification and Mobility 44*, 33–43.

Eurostat (2022, April). Children in migration - asylum applicants. Online.

Evans, M., J. Kelley, J. Sikora, and D. J. Treiman (2010). Family scholarly culture and educational success: Books and schooling in 27 nations. *Research in Social Stratification and Mobility 28*(2), 171 – 197.

Fiel, J. E. (2013). Decomposing school resegregation: social closure, racial imbalance, and racial isolation. *American Sociological Review 78*(5), 828–848.

Fleischmann, F. and J. Dronkers (2010). The socio-economic integration of immigrants in the EU – effects of characteristics of origin and destination countries on the first and second generation. In P. Attewell and K. S. Newman (Eds.), *Growing Gaps – Educational inequality around the World*, pp. 258–283. Oxford: Oxford University Press.

Forster, A. G. and H. G. van de Werfhorst (2019, 10). Navigating Institutions: Parents' Knowledge of the Educational System and Students' Success in Education. *European Sociological Review 36*(1), 48–64.

Fortin, N., T. Lemieux, and S. Firpo (2011). Decomposition methods. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 4A, pp. 1–102. Amsterdam: North-Holland.

Fryer, Roland G, J. and S. D. Levitt (2004). Understanding the black-white test score gap in the first two years of school. *The Review of Economics and Statistics 86*(2), 447–464.

Fryer, Roland G, J. and S. D. Levitt (2006). The Black-White test score gap through third grade. *American Law and Economics Review 8*(2), 249–281.

Furtado, D. and T. Song (2015). Intermarriage and socioeconomic integration trends in earnings premiums among U.S. immigrants who marry natives. *The ANNALS of the American Academy of Political and Social Science 662*(1), 207–222.

Gamoran, A. and R. D. Mare (1989). Secondary school tracking and educational inequality: Compensation, reinforcement, or neutrality? *American Journal of Sociology 94*(5), 1146–1183.

Gamoran, A., R. N. L. Turley, A. Turner, and R. Fish (2012). Differences between Hispanic and non-Hispanic families in social capital and child development: First-year findings from an experimental study. *Research in Social Stratification and Mobility 30*(1), 97 – 112. Inequality across the Globe.

Gastwirth, J. L. (1975). Statistical measures of earnings differentials. *The American Statistician 29*(1), 32–35.

Glock, S., S. Krolak-Schwerdt, F. Klapproth, and M. Böhmer (2013). Beyond judgment bias: How students' ethnicity and academic profile consistency influence teachers' tracking judgments. *Social Psychology of Education 16(4)*, 555–573.

Grisay, A., J. H. de Jong, E. Gebhardt, A. Berezner, and B. Halleux-Monseur (2007). Translation equivalence across PISA countries. *Journal of Applied Measurement 8(3)*, 249–266.

Grisay, A., E. J. Gonzalez, and C. Monseur (2009). Equivalence of item difficulties across national versions of the PIRLS and PISA reading assessments. *IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments 2*, 63–83.

Hanushek, E. A., J. D. Light, P. E. Peterson, L. M. Talpey, and L. Woessmann (2022). Long-run trends in the U.S. SES-achievement gap. *Education Finance and Policy 17*(4), 608–640.

Hanushek, E. A. and S. G. Rivkin (2006, October). School quality and the Black-White achievement gap. Working Paper 12651, National Bureau of Economic Research.

Hanushek, E. A. and S. G. Rivkin (2009). Harming the best: How schools affect the Black-White achievement gap. *Journal of Policy Analysis and Management 28*(3), 366–393.

Hanushek, E. A. and L. Woessmann (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *The Economic Journal 116(510)*, C63–C76.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science 312*(5782), 1900–1902.

Hildebrand, V. A., M. N. Pi Alperin, and P. Van Kerm (2017). Measuring and accounting for the deprivation gap of portuguese immigrants in luxembourg. *Review of Income and Wealth 63*(2), 288–309.

Hillmert, S. (2013). Links between immigration and social inequality in education: A comparison among five European countries. *Research in Social Stratification and Mobility 32*, 7–23.

Ho, A. D. (2009). A nonparametric framework for comparing trends and gaps across tests. *Journal of Educational and Behavioral Statistics 34*(2), 201–228.

Ho, A. D. and S. F. Reardon (2012). Estimating achievement gaps from test scores reported in ordinal "proficiency" categories. *Journal of Educational and Behavioral Statistics 37*(4), 489–517.

Hogrebe, N. and R. Strietholt (2016). Does non-participation in preschool affect children's reading achievement? international evidence from propensity score analyses. *Large-scale Assessments in Education 4*(1), 2.

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

Isphording, I. E. and S. Otten (2014). Linguistic barriers in the destination language acquisition of immigrants. *Journal of economic Behavior & organization 105*, 30–50.

Isphording, I. E., M. Piopiunik, and N. Rodriguez-Planas (2016). Speaking in numbers: The effect of reading performance on math performance among immigrants. *Economics Letters 139*, 52 – 56.

Jackson, M., J. O. Jonsson, and F. 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.

Jakubowski, M. and A. Pokropek (2015). Reading achievement progress across countries. *International Journal of Educational Development 45*(Supplement C), 77 – 88.

Janssen, L., D. Zwier, and H. van de Werfhorst (2023). Privatization of preschool education: The implications for educational inequalities in cross-national perspective. *Acta Sociologica 0*(0), 00016993231205201.

Jonsson, J. O. and F. Rudolphi (2011). Weak performance–strong determination: School achievement and educational choice among children of immigrants in Sweden. *European Sociological Review 27*(4), 487–508.

Justiniano Medina, A. C. and M. Valentova (2023). The effects of intermarriage on wages for immigrant women in Italy. *Journal of Family Issues 44*(1), 264–285.

Kalmijn, M. (2015). The children of intermarriage in four European countries: Implications for school achievement, social contacts, and cultural values. *The ANNALS of the American Academy of Political and Social Science 662*(1), 246–265.

Kalter, F. and I. Kogan (2006). Ethnic inequalities at the transition from school to work in Belgium and Spain: Discrimination or self-exclusion? *Research in Social Stratification and Mobility 24*(3), 259 – 274.

Kanas, A. and S. Steinmetz (2020, 11). Economic Outcomes of Immigrants with Different Migration Motives: The Role of Labour Market Policies. *European Sociological Review 37*(3), 449–464.

Kao, G. and M. Tienda (1998). Educational aspirations of minority youth. *American Journal of Education 106*(3), 349.

Klein, D. and M. Neugebauer (2023). A downside to high aspirations: Immigrants' (non-)success in tertiary education. *Acta Sociologica 0*(0), 00016993221148897.

Klein, O. and B. Becker (2017). Preschools as language learning environments for children of immigrants. Differential effects by familial language use across different preschool contexts. *Research in Social Stratification and Mobility 48*, 20–31.

Klein, O., N. Biedinger, and B. Becker (2014). The effect of reading aloud daily - differential effects of reading to native-born German and Turkish-origin immigrant children. *Research in Social Stratification and Mobility 38*, 43 – 56.

Kogan, I. (2016). Integration policies and immigrants' labor market outcomes in Europe. *Sociological Science 3*(16), 335–358.

Lavrijsen, J. and I. Nicaise (2016). Educational tracking, inequality and performance: New evidence from a differences-in-differences technique. *Research in Comparative and International Education 11*(3), 334–349.

Le Breton, M., A. Michelangeli, and E. Peluso (2012). A stochastic dominance approach to the measurement of discrimination. *Journal of Economic Theory 147*(4), 1342–1350.

Levels, M. and J. Dronkers (2008). Educational performance of native and immigrant children from various countries of origin. *Ethnic and Racial Studies 31*(8), 1404–1425.

Levels, M., J. Dronkers, and G. Kraaykamp (2008). Immigrant children's educational achievement in Western countries: Origin, destination, and community effects on mathematical performance. *American Sociological Review 73*(5), 835 – 853.

Luedemann, E. and G. Schwerdt (2013). Migration background and educational tracking. *Journal of Population Economics 26*(2), 455 – 481.

Melitz, J. (2018, April). English as a lingua franca: Facts, benefits and costs. *The World Economy 41*(7), 1750–1774.

Meng, X. and R. G. Gregory (2005, January). Intermarriage and the economic assimilation of immigrants. *Journal of Labor Economics 23*(1), 135–174.

Meng, X. and D. Meurs (2009). Intermarriage, language, and economic assimilation process: A case study of France. *International Journal of Manpower 30*(1/2), 127–144.

Merry, J. J. (2013). Tracing the U.S. deficit in PISA reading skills to early childhood: Evidence from the United States and Canada. *Sociology of Education 86(3)*, 234–252.

Mislevy, R. J. and K. M. Sheehan (1989). The role of collateral information about examinees in item parameter estimation. *Psychometrika 54*(4), 661–679.

Monseur, C. and R. J. Adams (2009). Plausible values: How to deal with their limitations. *Journal of Applied Measurement 10*(3), 320–34.

Mulhlenweg, A. M. (2008). Educational effects of alternative secondary school tracking regimes in Germany. *Schmollers Jahrbuch 128*, 351–379.

Muttarak, R. (2013). Is it (dis)advantageous to have mixed parentage? exploring education & work characteristics of children of interethnic unions in britain. Vienna Institute of Demography Working Papers 1/2013, Vienna.

Nielsen, E. R. (2015). Achievement gap estimates and deviations from cardinal comparability. Finance and Economics Discussion Series 2015-040, Washington: Board of Governors of the Federal Reserve System.

Noble, K. G., L. E. Engelhardt, N. H. Brito, L. J. Mack, E. J. Nail, J. Angal, R. Barr, W. P. Fifer, A. J. Elliott, and in collaboration with the PASS Network (2015). Socioeconomic disparities in neurocognitive development in the first two years of life. *Developmental Psychobiology 57*(5), 535–551.

Noble, K. G., M. F. Norman, and M. J. Farah (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental Science 8*(1), 74–87.

Nusche, D., C. Shewbridge, and C. L. Rasmussen (2009). OECD reviews of migrant education: Austria. Technical report, OECD.

Nusche, D., G. Wurzburg, and B. Naughton (2010). OECD reviews of migrant education: Denmark. Technical report, OECD.

OECD (2010). *Closing the Gap for Immigrant Students*.

OECD (2012). *PISA 2009 Technical Report*.

OECD (2019). *PISA 2018 Results (Volume II)*.

OECD (2023, September). *Education at a Glance 2023: OECD Indicators*. OECD.

OECD/EU (2015a). Income of immigrant households. In *Indicators of Immigrant Integration 2015: Settling In*. OECD Publishing, Paris.

OECD/EU (2015b). Young people with a migrant background. In *Indicators of Immigrant Integration 2015: Settling In*. OECD Publishing, Paris.

OECD/EU (2018). *Settling In 2018: Indicators of Immigrant Integration*. OECD Publishing, Paris/European Union, Brussels.

Park, H. and G. Sandefur (2010). Educational gaps between immigrant and native students in Europe: The role of grade. In J. Dronkers (Ed.), *Quality and Inequality of Education*, pp. 113–136. Springer Netherlands.

Passaretta, G., J. Skopek, and T. van Huizen (2022, February). Is social inequality in school-age achievement generated before or during schooling? a european perspective. *European Sociological Review 38*(6), 849–865.

Pedraja-Chaparro, F., D. S. González, and R. S. Rodríguez (2016). The impact of immigrant concentration in schools on grade retention in Spain: a difference-in-differences approach. *Applied Economics 48*(21), 1978–1990.

Pedro, C., M. Costas, and P. Matthias (2013). Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association 11*, 123–160.

Pfeffer, F. T. (2008, 05). Persistent Inequality in Educational Attainment and its Institutional Context. *European Sociological Review 24*(5), 543–565.

Reardon, S. F. (2008). Thirteen ways of looking at the Black-White test score gap. Working Paper 2008-08, Institute for Research on Education Policy and Practice.

Reardon, S. F. and C. Galindo (2009). The Hispanic-White achievement gap in math and reading in the elementary grades. *American Educational Research Journal 46*(3), 853–891.

Riederer, B. and R. Verwiebe (2015). Changes in the educational achievement of immigrant youth in Western societies: The contextual effects of national (educational) policies. *European Sociological Review 31*(5), 628–642.

Rubin, D. B. (1987). *Multiple imputation for non-response in surveys*. New York: John Wiley & Sons.

Ruhose, J. and G. Schwerdt (2016). Does early educational tracking increase migrant-native achievement gaps? Differences-in-differences evidence across countries. *Economics of Education Review 52*(Supplement C), 134 – 154.

Ruzek, E., M. Burchinal, G. Farkas, and G. J. Duncan (2014). The quality of toddler child care and cognitive skills at 24 months: Propensity score analysis results from the ECLS-B. *Early Childhood Research Quarterly 29*(1), 12 – 21.

Sakellariou, C. (2018, October). Selective immigration policies as screening devices and the immigrant–native educational achievement gap. *Educational Research and Evaluation 24*(6–7), 394–416.

Schipolowski, S., A. Edele, N. Mahler, and P. Stanat (2021, March). Mathematics and science proficiency of young refugees in secondary schools in germany. *Journal for Educational Research Online 2021*(1), 78–104.

Schneeweis, N. (2011). Educational institutions and the integration of migrants. *Journal of Population Economics 24*(4), 1281–1308.

Schnepf, S. V. (2007). Immigrants' educational disadvantage: an examination across ten countries and three surveys. *Journal of Population Economics 20*(3), 527–545.

Schubert, F. and R. Becker (2010). Social inequality of reading literacy: A longitudinal analysis with cross-sectional data of PIRLS 2001 and PISA 2000 utilizing the pair wise matching procedure. *Research in Social Stratification and Mobility 28*(1), 109–133.

Schütz, G., H. W. Ursprung, and L. Wößmann (2008). Education policy and equality of opportunity. *Kyklos 61*(2), 279–308.

Shewbridge, C., M. Kim, G. Wurzburg, and G. Hostens (2010). OECD reviews of migrant education: Netherlands. Technical report, OECD.

Skopek, J. and G. Passaretta (2020, 10). Socioeconomic inequality in children's achievement from infancy to adolescence: The case of Germany. *Social Forces 100*(1), 86–112.

Statistics Canada (2021). Focus on geography series, 2021 census of population. https://www12.statcan.gc.ca/census-recensement/2021/as-sa/fogs-spg/page.cfm?topic=9&lang=E&dguid=2021A000011124.

Taguma, M., M. K. andSatya Brink, and J. Teltemann (2010). OECD reviews of migrant education: Sweden. Technical report, OECD.

Taguma, M., M. Kim, G. Wurzburg, and F. Kelly (2009). OECD reviews of migrant education: Ireland. Technical report, OECD.

Triventi, M., J. Skopek, N. Kulic, S. Buchholz, and H.-P. Blossfeld (2020). Advantage 'finds its way': How privileged families exploit opportunities in different systems of secondary education. *Sociology 54*(2), 237–257.

van de Werfhorst, H. G. (2019). Early tracking and social inequality in educational attainment: Educational reforms in 21 european countries. *American Journal of Education 126*(1), 65–99.

van de Werfhorst, H. G. and A. Heath (2019, May). Selectivity of migration and the educational disadvantages of second-generation immigrants in ten host societies. *European Journal of Population 35*(2), 347–378.

Vonnahme, C. (2021). Do migrant-native achievement gaps narrow? evidence over the school career. techreport Ruhr Economic Papers, No. 932, RWI - Leibniz-Institut für Wirtschaftsforschung.

Woessmann, L. (2008). How equal are educational opportunities? family background and student achievement in Europe and the United States. *Zeitschrift für Betriebswirtschaftr Betriebswirtschaft 78*(1), 45–70.

Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation 31*(2), 114 – 128. Measurement, Evaluation, and Statistical Analysis.

Wu, M. (2010). Comparing the similarities and differences of PISA 2003 and TIMSS. OECD Education Working Papers 32, OECD.

Zhou, X. (2012). A nonparametric index of stratification. *Sociological Methodology 42*(1), 365–389.

Zwier, D., S. Geven, and H. G. van de Werfhorst (2020, December). Social inequality in shadow education: The role of high-stakes testing. *International Journal of Comparative Sociology 61*(6), 412–440.

6 Tables & Figures

Table 1: Age-Cohort–Data Summary

	Cohort 1	**Cohort 2**	**Cohort 3**
Primary school Secondary school	PIRLS 2001- PISA 2006	PIRLS 2006- PISA 2012	TIMSS 2015- TIMMS 2019
Tracking countries			
Austria		x	
Belgium		x	
Germany	x	x	
Luxembourg		x	
Netherlands	x	x	
Late- / Non-Tracking countries			
Denmark		x	
France	x	x	x
Ireland			x
Italy	x	x	x
Norway	x	x	x
Spain		x	
Sweden	x	x	x
Traditional immigration countries (TIC)			
Australia			x
Canada	x	x	
New Zealand	x	x	x
United States		x	x

Table 2: Percentage of Students with a Migratory Background (Reading Test Cohorts)

	N		1G		2G		Mixed	
Cohort 1 (2001-2006)	PIRLS	PISA	PIRLS	PISA	PIRLS	PISA	PIRLS	PISA
Tracking								
Germany	5350	4181	6.5	5.0	7.2	6.3	8.1	4.9
Netherlands	3599	4541	3.7	3.1	5.6	7.2	8.9	8.3
Late- / Non-Tracking countries								
France	2832	4337	3.0	2.4	10.9	8.9	12.1	12.5
Italy	3055	20279	2.5	1.8	1.1	0.5	6.5	5.1
Norway	3033	4335	2.6	2.3	2.7	2.8	11.1	8.0
Sweden	5200	4136	5.2	3.6	6.1	5.6	12.0	10.5
TIC								
Canada	5569	20569	7.8	6.7	12.1	11.2	14.5	11.5
New Zealand	1271	4126	11.5	9.5	9.6	7.0	16.7	16.4
Cohort 2 (2006-2012)	PIRLS	PISA	PIRLS	PISA	PIRLS	PISA	PIRLS	PISA
Tracking								
Austria	3973	4270	3.1	3.6	9.8	8.7	8.1	8.0
Belgium	7395	7602	4.5	4.6	6.4	7.1	13.4	13.2
Germany	5933	3661	3.2	2.0	10.2	9.4	9.8	7.3
Luxembourg	2395	4323	9.2	13.4	26.6	27.7	18.5	17.0
Netherlands	3712	4121	2.1	2.1	8.0	7.0	9.3	8.5
Late- / Non-Tracking countries								
Denmark	3018	6464	2.0	1.9	6.0	5.2	9.9	7.8
France	3513	4185	2.9	3.1	9.2	9.0	14.1	11.5
Italy	3281	28600	3.2	3.4	2.5	1.6	6.6	6.6
Norway	3216	4241	1.7	3.1	3.2	4.5	8.7	9.3
Spain	3650	23266	7.9	5.7	1.8	1.4	6.7	6.5
Sweden	3870	4121	2.5	3.1	9.3	7.6	11.8	11.7
TIC								
Canada	16248	19019	8.9	9.0	19.4	16.8	12.4	10.8
New Zealand	4837	3683	12.1	11.1	8.2	10.3	17.6	17.3
United States	3844	4523	6.3	5.3	10.9	14.9	9.0	8.6

Notes: All proportions are weighted using individual student weights. 1G=First-generation migrants, 2G=Second-generation migrants, Mixed=One native-born parents.

Table 3: Percentage of Students with a Migratory Background (Math Test Cohort)

	N		1G		2G		Mixed	
Cohort 3 (TIMMS 20015-2019)	G4	G8	G4	G8	G4	G8	G4	G8
Late- / Non-Tracking countries								
France	4163	3556	3.9	4.6	9.7	12.2	13.7	14.1
Ireland	3922	3787	5.9	9.5	6.2	6.6	14.9	16.9
Italy	3988	3471	3.1	3.6	7.9	6.8	9.8	9.1
Norway	3802	3962	5.3	6.5	7.8	8.5	13.2	11.9
Sweden	3771	3632	5.9	10.5	11.0	10.0	12.1	12.3
TIC								
Australia	5079	8243	10.7	13.5	12.0	12.9	17.5	17.4
New Zealand	5202	5398	12.5	14.6	13.0	12.1	20.1	18.0
United States	7214	7073	4.7	4.8	16.8	14.6	11.4	10.8

Notes: All proportions are weighted using individual student weights. 1G=First-generation migrants, 2G=Second-generation migrants, Mixed=One native-born parents. Own calculations using TIMSS 2015 (Grade 4) and TIMMS 2019 (Grade 8) data. G4 and G8 stand for Grade 4 and Grade 8 respectively. All proportions are weighted using individual student weights. 1G=First-generation migrants, 2G=Second-generation migrants, Mixed=One native-born parents.

Table 4: Descriptive Characteristics of Reading Sample by Cohorts

	Natives		1G		2G		Mixed	
Cohort 1 (2001-2006)	PIRLS	PISA	PIRLS	PISA	PIRLS	PISA	PIRLS	PISA
Student Demograhics								
Boys	0.50	0.50	0.54	0.50	0.47	0.50	0.52	0.50
Age	10.28	16.00	10.44	15.99	10.33	15.99	10.29	15.99
Grade	4.00	9.57	4.00	9.34	4.00	9.48	4.00	9.62
Birthyear	1991	1990	1991	1990	1991	1990	1991	1990
Education of Parents								
Primary or less	0.25	0.14	0.26	0.19	0.39	0.30	0.22	0.10
Secondary	0.51	0.42	0.43	0.26	0.42	0.29	0.46	0.36
Tertiary	0.24	0.45	0.30	0.55	0.19	0.41	0.32	0.54
Books at Home								
0-10	0.07	0.08	0.21	0.18	0.16	0.18	0.07	0.08
11-25	0.21	0.14	0.29	0.22	0.27	0.21	0.19	0.13
26-100	0.35	0.32	0.28	0.31	0.35	0.32	0.33	0.27
101+	0.37	0.47	0.22	0.29	0.23	0.29	0.41	0.53
Household Possessions								
Own computer	0.85	0.93	0.74	0.90	0.77	0.93	0.83	0.91
Own Desk	0.88	0.96	0.84	0.95	0.84	0.95	0.89	0.95
Cohort 2 (2006-2012)								
Student Demograhics								
Boys	0.51	0.50	0.50	0.51	0.49	0.49	0.51	0.49
Age	10.22	16.00	10.35	15.99	10.30	15.99	10.25	15.99
Grade	4.01	9.64	4.05	9.50	4.02	9.57	4.03	9.63
Birthyear	1996	1996	1996	1996	1996	1996	1996	1996
Education of Parents								
Primary or less	0.21	0.13	0.21	0.17	0.29	0.21	0.16	0.11
Secondary	0.43	0.33	0.35	0.27	0.42	0.33	0.42	0.32
Tertiary	0.36	0.54	0.44	0.57	0.29	0.46	0.41	0.58
Books at Home								
0-10	0.08	0.11	0.19	0.23	0.16	0.20	0.08	0.12
11-25	0.19	0.15	0.29	0.24	0.28	0.23	0.20	0.15
26-100	0.34	0.30	0.29	0.30	0.31	0.31	0.32	0.29
101+	0.39	0.45	0.23	0.24	0.25	0.26	0.40	0.44
Household Possessions								
Own computer	0.90	0.98	0.84	0.96	0.88	0.98	0.91	0.97
Own Desk	0.87	0.96	0.82	0.95	0.86	0.95	0.87	0.95

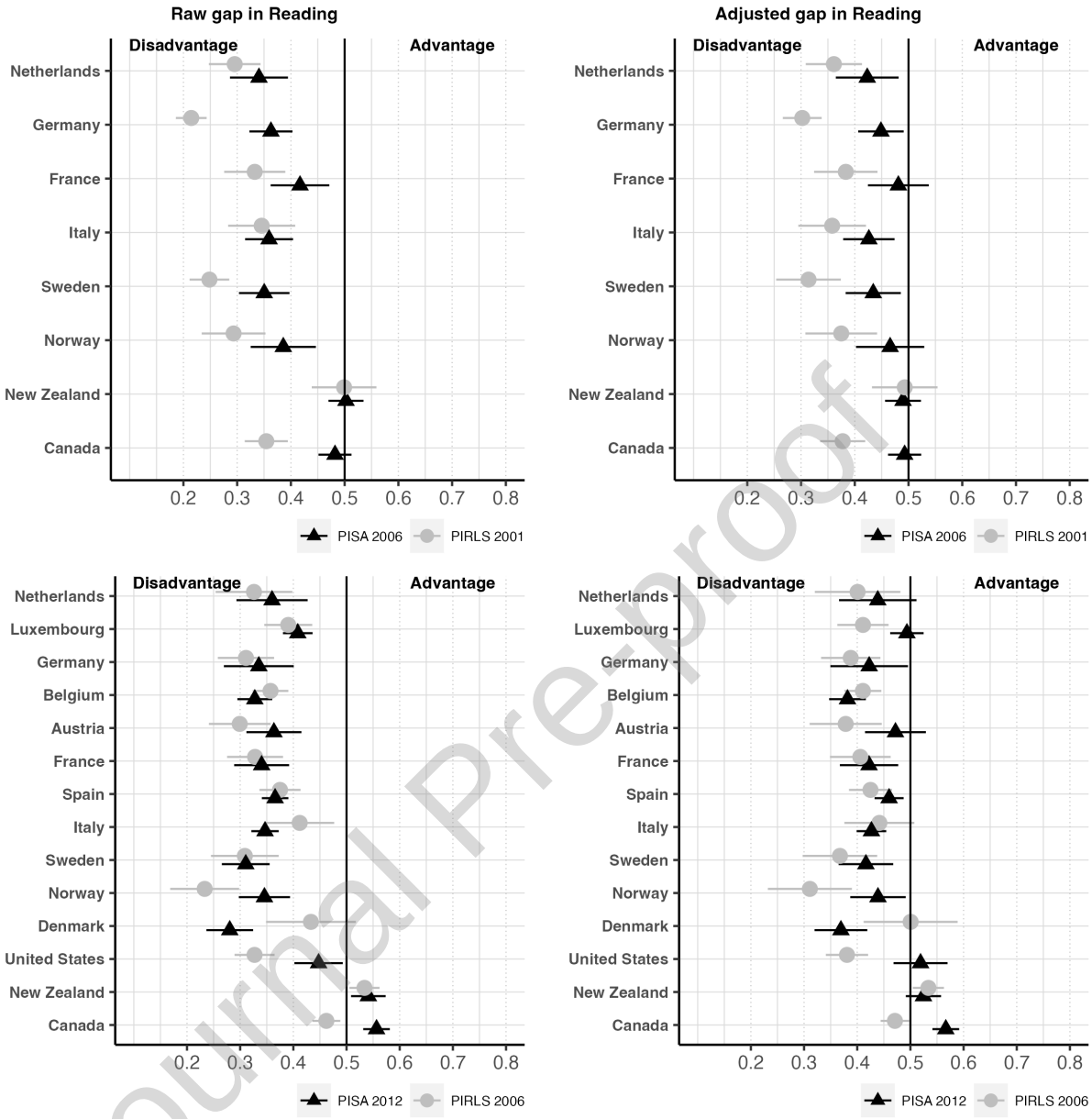
Notes: Own calculations on PIRLS waves 2001 and 2006 and PISA 2006 and 2012 data. 1G=First-generation migrants, 2G=Second-generation migrants, Mixed=One native-born parent. All proportions are weighted using individual student weights.

Table 5: Descriptive Characteristics of Mathematic Sample

	Natives		1G		2G		Mixed	
Cohort 3 (2015-2019)	G4	G8	G4	G8	G4	G8	G4	G8
Student Demograhics								
Boys	0.51	0.51	0.53	0.49	0.49	0.51	0.50	0.50
Age	9.91	13.90	10.07	14.16	9.94	13.97	9.93	13.93
Grade	4.06	8.05	4.10	8.10	4.07	8.06	4.08	8.07
Birthyear	2005.1	2005.1	2004.9	2004.8	2005.1	2005.0	2005.1	2005.1
Education of Parents								
Primary or less	0.10	0.13	0.17	0.17	0.23	0.22	0.11	0.12
Secondary	0.41	0.38	0.36	0.29	0.39	0.37	0.37	0.33
Tertiary	0.49	0.49	0.47	0.54	0.38	0.41	0.52	0.55
Books at Home								
0-10	0.10	0.15	0.22	0.32	0.19	0.28	0.10	0.15
11-25	0.25	0.22	0.31	0.31	0.33	0.30	0.25	0.25
26-100	0.35	0.28	0.28	0.21	0.31	0.24	0.35	0.25
101+	0.30	0.35	0.20	0.17	0.17	0.18	0.31	0.35
Household Possessions								
Own computer	0.63	0.98	0.69	0.96	0.67	0.97	0.68	0.98
Own Desk	0.77	0.92	0.75	0.88	0.75	0.86	0.77	0.91

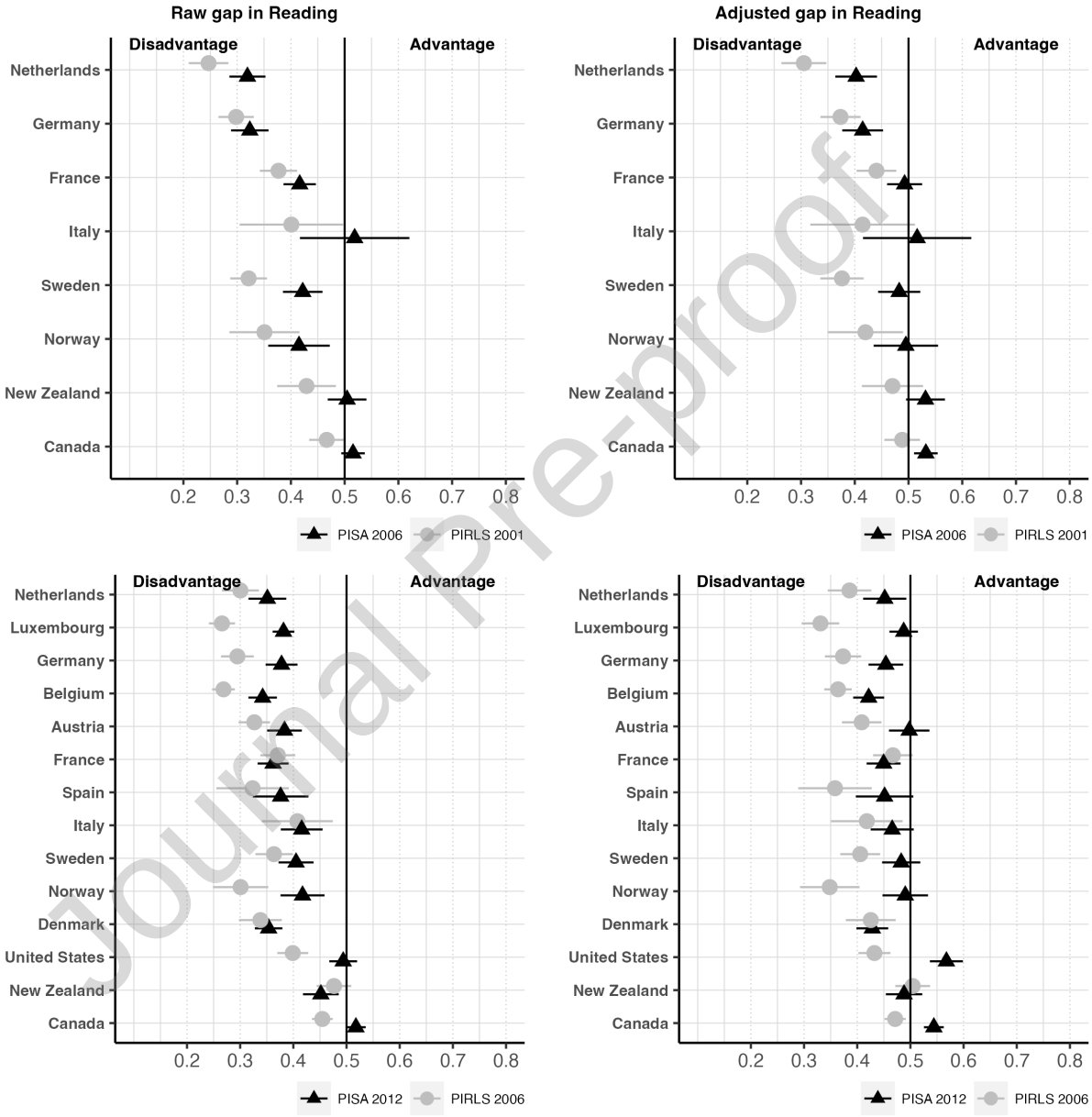
Notes: Own calculations using TIMSS 2015 (Grade 4) and 2019 (Grade 8) data. G4 and G8 stand for Grade 4 and Grade 8 respectively. 1G=First-generation migrants, 2G=Second-generation migrants, Mixed=One native-born parents. All proportions are weighted using individual student weights.

Figure 1: First-generation migrants vs. Natives; Gastwirth's PROB Index



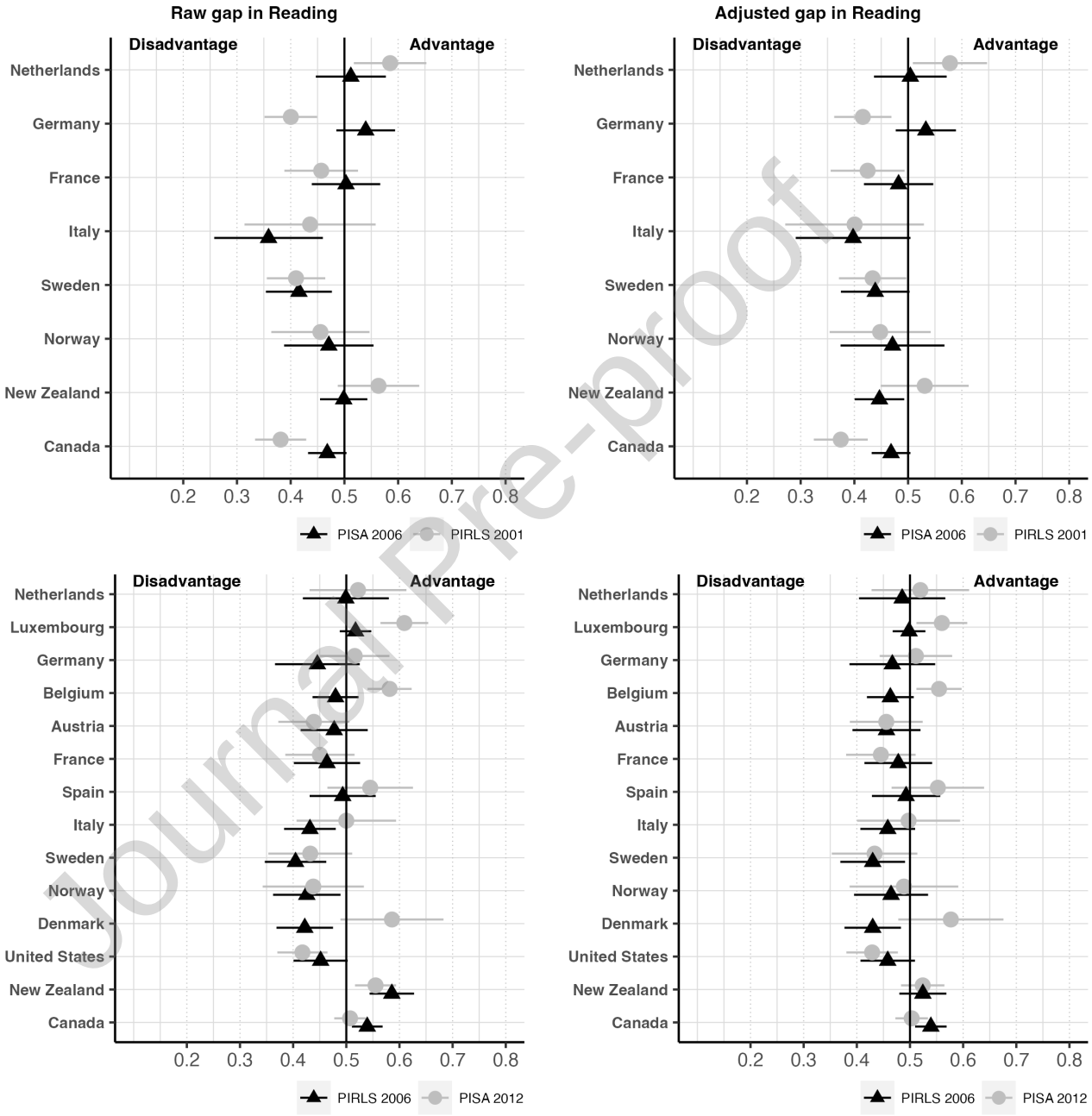
Note: Gastwirth index (PROB) with 95% confidence interval (CI) error bars. PROB < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on.

Figure 2: Second-generation migrants vs. Natives; Gastwirth's PROB Index



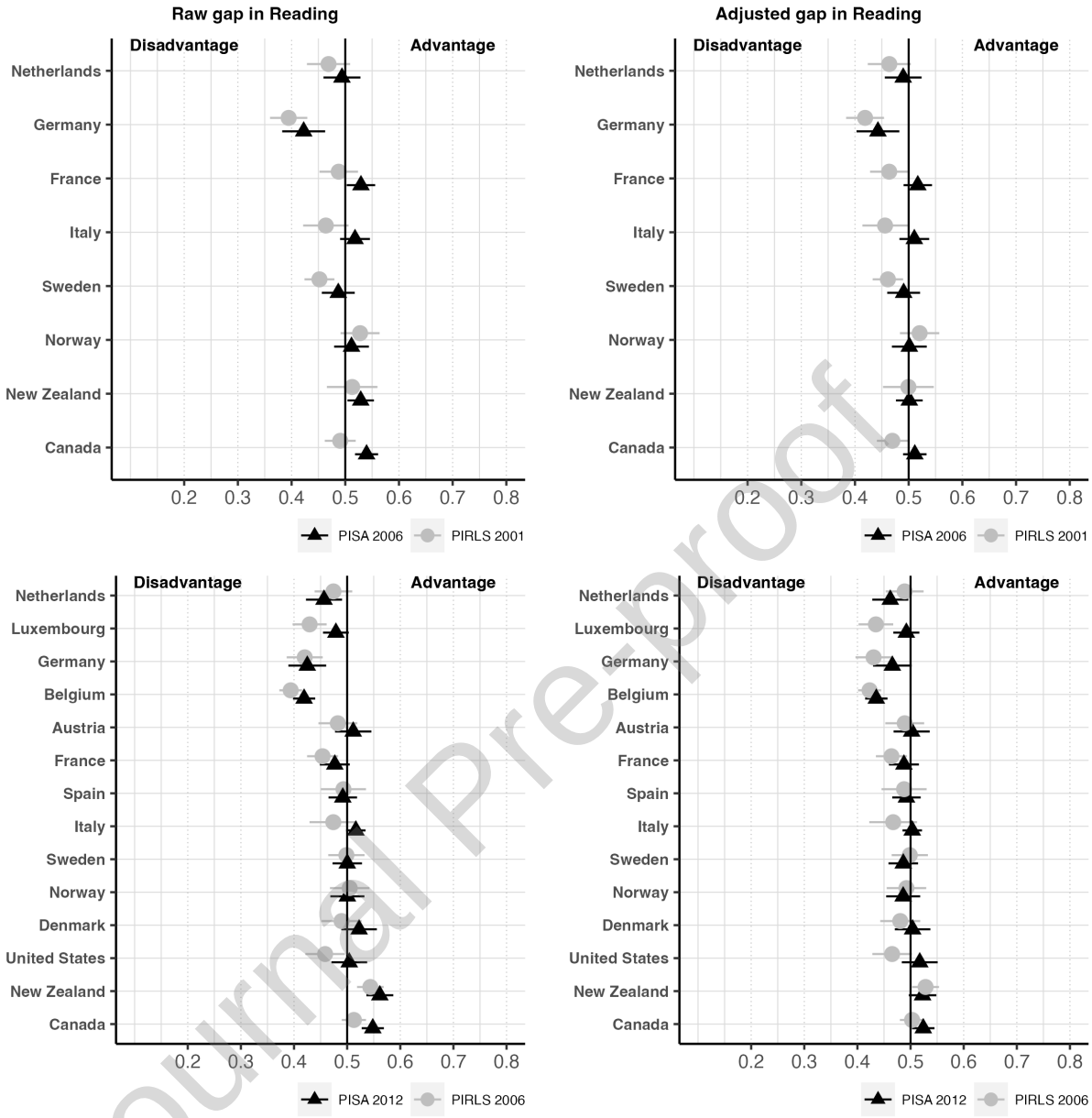
Note: Gastwirth index (PROB) with 95% confidence interval (CI) error bars. PROB < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on.

Figure 3: First-generation vs. Second-generation migrants; Gastwirth's PROB Index



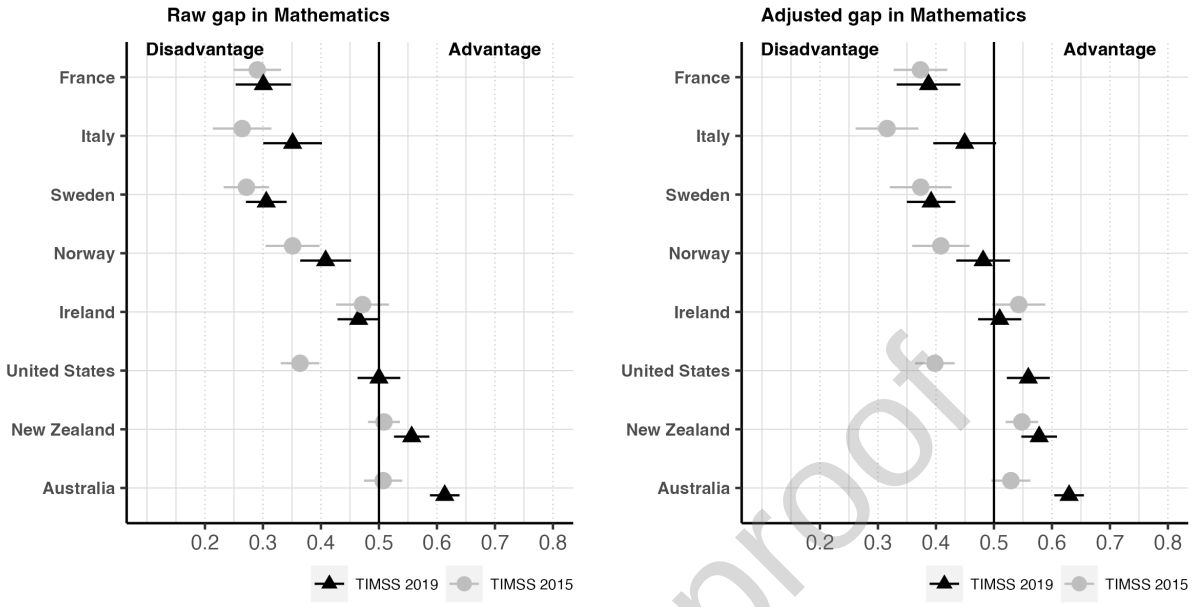
Note: Gastwirth index (PROB) with 95% confidence interval (CI) error bars. PROB < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on.

Figure 4: Mixed migrants vs. Natives; Gastwirth's PROB Index



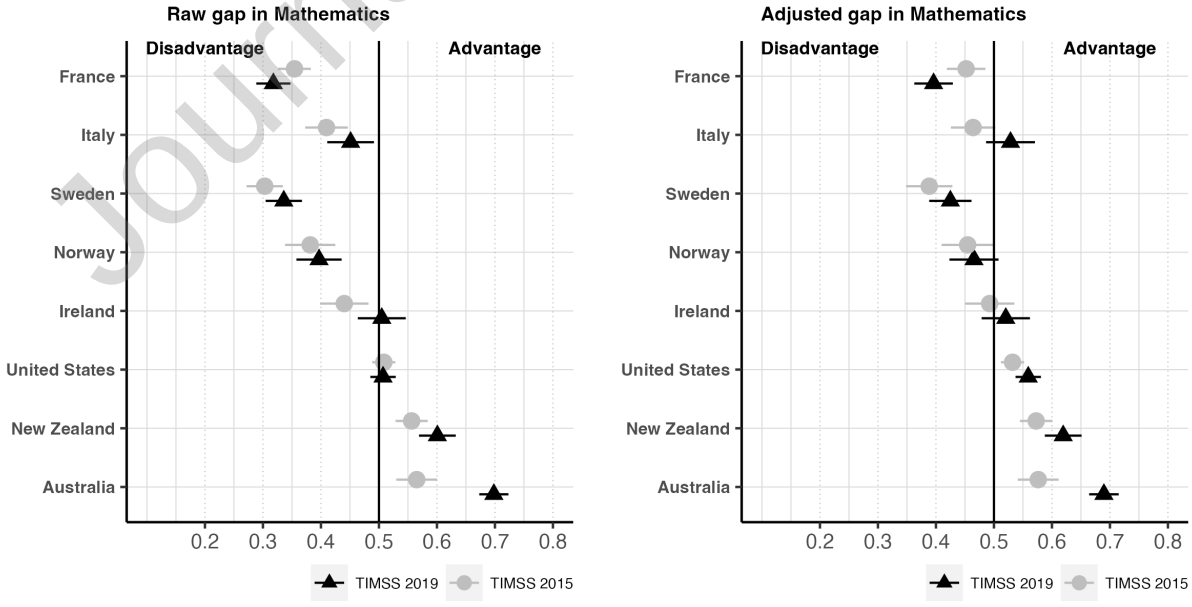
Note: Gastwirth index (PROB) with 95% confidence interval (CI) error bars. PROB < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on.

Figure 5: First-generation migrants vs. Natives; Gastwirth's PROB Index



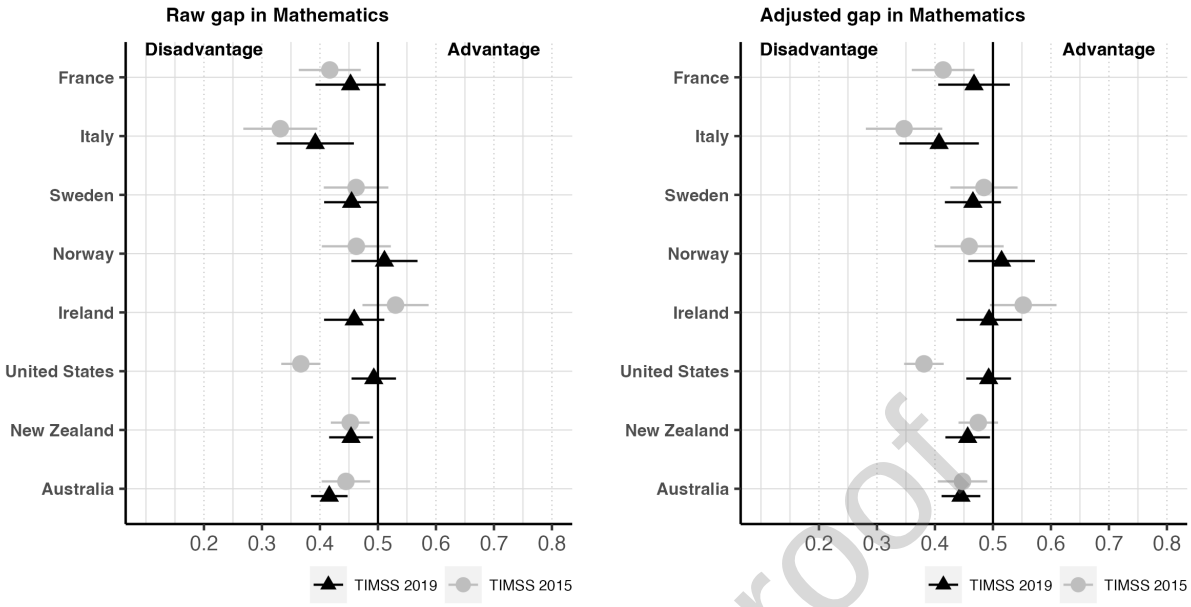
Note: Gastwirth index (PROB) with 95% confidence interval (CI) error bars. PROB < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on.

Figure 6: Second-generation migrants vs. Natives; Gastwirth's PROB Index



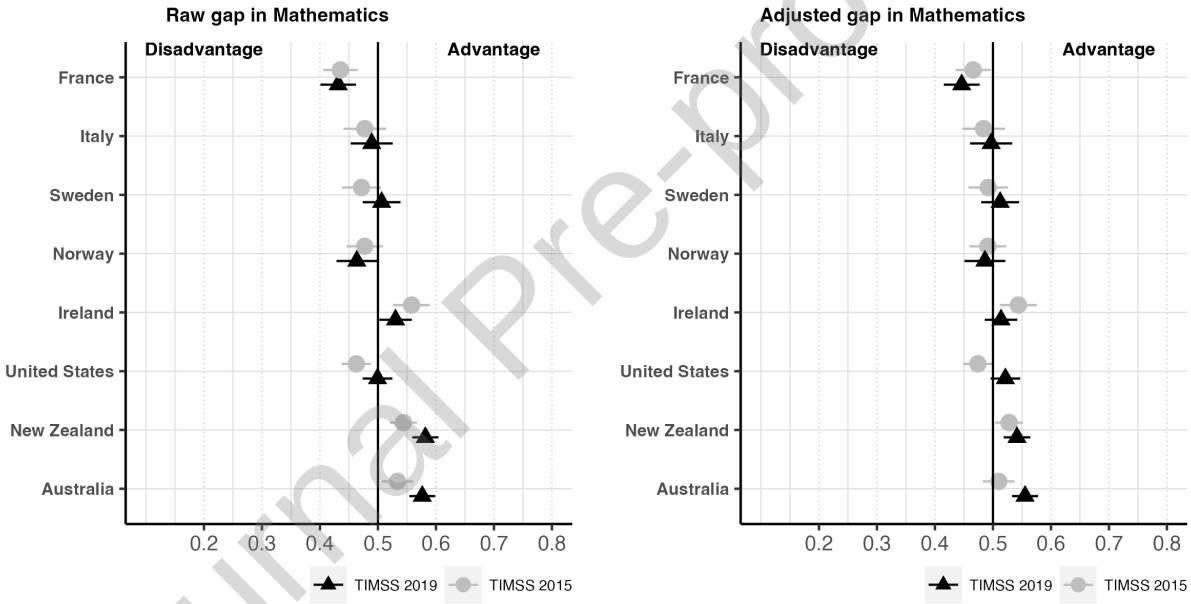
Note: Gastwirth index (PROB) with 95% confidence interval (CI) error bars. PROB < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on.

Figure 7: First-generation vs. Second-generation migrants; Gastwirth's PROB Index



Note: Gastwirth index (PROB) with 95% confidence interval (CI) error bars. PROB < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on.

Figure 8: Mixed migrants vs. Natives; Gastwirth's PROB Index



Note: Gastwirth index (PROB) with 95% confidence interval (CI) error bars. PROB < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on.

Appendices

Appendix A Results in Reading

Table A.1: Achievement Gap in Reading: Native-born vs First Generation Migrants

	Normalized Score						Gastwirth Index					
	PIRLS		PISA		Δ Gap		PIRLS		PISA		Δ Gap	
	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj
2001-2006												
Netherlands	64*	42*	61*	29*	−4	−13	0.30*	0.36*	0.34*	0.42†	0.05	0.06
Germany	109*	73*	53*	22†	−56*	−52*	0.21*	0.30*	0.36*	0.45†	0.15*	0.15*
France	58*	40*	29*	8	−28‡	−32†	0.33*	0.38*	0.42*	0.48	0.08†	0.10†
Italy	52*	48*	57*	32*	5	−16	0.35*	0.36*	0.36*	0.43*	0.01	0.07‡
Sweden	94*	66*	55*	25*	−39*	−41*	0.25*	0.31*	0.35*	0.43†	0.10*	0.12*
Norway	78*	49*	45*	15	−33‡	−34‡	0.29*	0.37*	0.39*	0.47	0.09†	0.09‡
New Zealand	3	5	0	5	−2	0	0.50	0.49	0.50	0.49	0.00	0.00
Canada	57*	48*	9	5	−48*	−43*	0.35*	0.38*	0.48	0.49	0.13*	0.12*
2006-2012												
Netherlands	67*	39†	60*	30†	−7	−9	0.33*	0.40†	0.36*	0.44‡	0.03	0.04
Luxembourg	39*	32*	32*	0	−7	−32*	0.39*	0.41*	0.41*	0.49	0.02	0.08*
Germany	69*	40*	69*	36†	0	−3	0.31*	0.39*	0.34*	0.42†	0.02	0.03
Belgium	53*	34*	65*	45*	12	11	0.36*	0.41*	0.33*	0.38*	−0.03	−0.03
Austria	81*	47*	48*	9	−33†	−39†	0.30*	0.38*	0.36*	0.47	0.06	0.09†
France	62*	34*	61*	30*	−0	−4	0.33*	0.41*	0.34*	0.42*	0.01	0.02
Spain	47*	29*	46*	13*	−1	−16‡	0.37*	0.42*	0.37*	0.46*	−0.01	0.03
Italy	33†	22‡	57*	27*	24‡	5	0.41*	0.44‡	0.35*	0.43*	−0.07‡	−0.01
Sweden	78*	56*	65*	28*	−13	−28	0.31*	0.37*	0.31*	0.42*	0.00	0.05
Norway	104*	71*	58*	24†	−46†	−47†	0.23*	0.31*	0.35*	0.44†	0.11*	0.13*
Denmark	26‡	1	81*	46*	55*	46*	0.43	0.50	0.28*	0.37*	−0.15*	−0.13†
United States	61*	42*	21†	−6	−41*	−48*	0.33*	0.38*	0.45†	0.52	0.12*	0.14*
New Zealand	−11†	−12†	−11‡	−5	−0	7	0.53†	0.53†	0.54†	0.52	0.01	−0.01
Canada	12*	9‡	−20*	−24*	−32*	−32*	0.46*	0.47†	0.56*	0.57*	0.09*	0.10*

Notes: Achievement gap measures based on normalized scores are rounded to the nearest integer. Gastwirth Index < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on. ∗, † and ‡ indicate achievement gap and achievement gap changes significance at 1, 5 and 10 percent levels, respectively.

Table A.2: Achievement Gap in Reading: Native-born vs Second Generation Migrants

	Normalized Score						Gastwirth Index					
	PIRLS		PISA		Δ Gap		PIRLS		PISA		Δ Gap	
	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj
2001-2006												
Netherlands	93*	69*	67*	35*	−27†	−34*	0.25*	0.31*	0.32*	0.40*	0.07*	0.10*
Germany	74*	47*	63*	30*	−11	−17‡	0.30*	0.37*	0.32*	0.41*	0.03	0.04
France	43*	21*	29*	3	−14	−17‡	0.38*	0.44*	0.42*	0.49	0.04‡	0.05†
Italy	34‡	29	−1	−1	−36	−30	0.40†	0.41‡	0.52	0.52	0.12‡	0.10
Sweden	63*	44*	24*	5	−39*	−39*	0.32*	0.38*	0.42*	0.48	0.10*	0.11*
Norway	59*	32†	31*	2	−28	−30‡	0.35*	0.42†	0.42*	0.49	0.06	0.08
New Zealand	21†	8	−1	−10	−22‡	−19	0.43†	0.47	0.50	0.53‡	0.08†	0.06‡
Canada	10‡	3	−5	−11*	−15†	−15†	0.47†	0.49	0.52	0.53*	0.05†	0.04†
2006-2012												
Netherlands	70*	38*	54*	19*	−16	−19‡	0.30*	0.39*	0.35*	0.45†	0.05†	0.07†
Luxembourg	81*	56*	40*	4	−41*	−52*	0.27*	0.33*	0.38*	0.49	0.12*	0.16*
Germany	72*	42*	43*	18*	−29*	−25*	0.29*	0.37*	0.38*	0.45*	0.08*	0.08*
Belgium	81*	47*	57*	28*	−24*	−19*	0.27*	0.36*	0.34*	0.42*	0.07*	0.06*
Austria	62*	33*	41*	0	−21†	−33*	0.33*	0.41*	0.38*	0.50	0.06†	0.09*
France	45*	11‡	46*	16*	2	5	0.37*	0.47‡	0.36*	0.45*	−0.01	−0.02
Spain	55*	42*	43*	16	−12	−26	0.32*	0.36*	0.38*	0.45‡	0.05	0.09†
Italy	29†	25†	32*	14‡	3	−11	0.41*	0.42†	0.42*	0.47‡	0.01	0.05
Sweden	45*	31*	35*	9	−10	−22†	0.36*	0.41*	0.41*	0.48	0.04‡	0.08*
Norway	79*	59*	27*	2	−52*	−58*	0.30*	0.35*	0.42*	0.49	0.12*	0.14*
Denmark	56*	26*	51*	23*	−5	−3	0.34*	0.43*	0.35*	0.43*	0.02	0.00
United States	33*	22*	−1	−25*	−34*	−47*	0.40*	0.43*	0.49	0.57*	0.09*	0.14*
New Zealand	6	−3	16*	4	10	7	0.48	0.50	0.45*	0.49	−0.02	−0.02
Canada	16*	11*	−6†	−16*	−23*	−26*	0.45*	0.47*	0.52‡	0.54*	0.06*	0.07*

Notes: Achievement gap measures based on normalized scores are rounded to the nearest integer. Gastwirth Index < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on. ∗, † and ‡ indicate achievement gap and achievement gap changes significance at 1, 5 and 10 percent levels, respectively.

Table A.3: Achievement Gap in Reading: Second vs First Generation Migrants

	Normalized Score						Gastwirth Index					
	PIRLS		PISA		Δ Gap		PIRLS		PISA		Δ Gap	
	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj
2001-2006												
Netherlands	$-29^{\dagger}$	$-27^{\dagger}$	-6	-4	23	24	$0.59^{\dagger}$	$0.58^{\dagger}$	0.51	0.50	-0.07	-0.07
Germany	35^{*}	30^{*}	-10	-8	-45^{*}	$-38^{\dagger}$	0.40^{*}	0.42^{*}	0.54	0.53	0.14^{*}	0.12^{*}
France	14	$24^{\ddagger}$	0	7	-14	-17	0.46	$0.42^{\dagger}$	0.50	0.48	0.05	0.06
Italy	18	30	58^{*}	$43^{\ddagger}$	40	13	0.44	0.40	0.36^{*}	$0.40^{\ddagger}$	-0.08	0.00
Sweden	30^{*}	$23^{\dagger}$	31^{*}	$22^{\ddagger}$	0	-0	0.41^{*}	$0.43^{\dagger}$	0.42^{*}	$0.44^{\ddagger}$	0.01	0.00
Norway	19	21	14	12	-5	-10	0.46	0.45	0.47	0.47	0.02	0.02
New Zealand	-18	-7	1	$21^{\dagger}$	19	28	0.56	0.53	0.50	$0.45^{\dagger}$	-0.06	$-0.08^{\ddagger}$
Canada	46^{*}	49^{*}	$14^{\dagger}$	$14^{\dagger}$	-32^{*}	-35^{*}	0.38^{*}	0.37^{*}	$0.47^{\ddagger}$	$0.47^{\ddagger}$	0.09^{*}	0.09^{*}
2006-2012												
Netherlands	-3	-1	6	11	9	12	0.52	0.52	0.50	0.49	-0.02	-0.03
Luxembourg	-41^{*}	$-22^{\dagger}$	-7	-0	34^{*}	$22^{\dagger}$	0.61^{*}	$0.56^{\dagger}$	0.52	0.50	-0.09^{*}	$-0.06^{\dagger}$
Germany	-3	-2	26	18	29	20	0.52	0.51	0.45	0.47	-0.07	-0.04
Belgium	-27^{*}	$-17^{\dagger}$	9	$15^{\ddagger}$	36^{*}	32^{*}	0.58^{*}	$0.55^{\dagger}$	0.48	0.46	-0.10^{*}	-0.09^{*}
Austria	19	13	7	15	-12	2	$0.44^{\ddagger}$	0.46	0.48	0.46	0.04	0.00
France	17	19	15	9	-2	-10	0.45	0.45	0.46	0.48	0.01	0.03
Spain	-8	-14	3	2	11	16	0.54	0.55	0.49	0.49	-0.05	-0.06
Italy	4	4	25^{*}	14	21	10	0.50	0.50	0.43^{*}	0.46	-0.07	-0.04
Sweden	$32^{\dagger}$	$32^{\dagger}$	29^{*}	$21^{\dagger}$	-3	-11	$0.43^{\ddagger}$	0.43	0.40^{*}	$0.43^{\dagger}$	-0.03	0.00
Norway	25	6	$31^{\dagger}$	16	6	10	0.44	0.49	$0.43^{\dagger}$	0.46	-0.01	-0.02
Denmark	$-30^{\ddagger}$	$-28^{\ddagger}$	29^{*}	27^{*}	60^{*}	55^{*}	$0.59^{\ddagger}$	0.58	0.42^{*}	0.43^{*}	-0.16^{*}	-0.15^{*}
United States	28^{*}	24^{*}	$21^{\dagger}$	$19^{\ddagger}$	-7	-6	0.42^{*}	0.43^{*}	$0.45^{\ddagger}$	0.46	0.03	0.03
New Zealand	-18^{*}	-7	-28^{*}	-5	-10	2	0.56^{*}	0.52	0.59^{*}	0.52	0.03	0.00
Canada	-4	-3	$-13^{\dagger}$	$-13^{\dagger}$	-9	-10	0.51	0.50	0.54^{*}	0.54^{*}	0.03	$0.04^{\ddagger}$

Notes: Achievement gap measures based on normalized scores are rounded to the nearest integer. Gastwirth Index < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on. $*$, † and ‡ indicate achievement gap and achievement gap changes significance at 1, 5 and 10 percent levels, respectively.

Table A.4: Achievement Gap in Reading: Native-born vs Mixed Migrants

	Normalized Score						Gastwirth Index					
	PIRLS		PISA		Δ Gap		PIRLS		PISA		Δ Gap	
	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj
2001-2006												
Netherlands	11	$12^{\ddagger}$	3	4	-7	-8	0.47	$0.46^{\ddagger}$	0.49	0.49	0.02	0.03
Germany	37^{*}	28^{*}	25^{*}	$18^{\dagger}$	-11	-11	0.39^{*}	0.42^{*}	0.42^{*}	0.44^{*}	0.03	0.02
France	6	$14^{\dagger}$	$-8^{\ddagger}$	-5	$-14^{\ddagger}$	$-19^{\dagger}$	0.49	$0.46^{\dagger}$	$0.53^{\dagger}$	0.52	$0.04^{\ddagger}$	$0.05^{\dagger}$
Italy	$15^{\ddagger}$	$18^{\dagger}$	-5	-3	$-20^{\dagger}$	$-21^{\dagger}$	$0.46^{\ddagger}$	$0.46^{\dagger}$	0.52	0.51	$0.05^{\dagger}$	$0.05^{\dagger}$
Sweden	16^{*}	$13^{\dagger}$	3	2	$-13^{\ddagger}$	-11	0.45^{*}	0.46^{*}	0.49	0.49	$0.04^{\ddagger}$	0.03
Norway	-10	-8	-4	-1	5	7	0.53	0.52	0.51	0.50	-0.02	-0.02
New Zealand	-6	-2	-11^{*}	-2	-4	0	0.51	0.50	$0.53^{\dagger}$	0.50	0.02	0.00
Canada	2	$9^{\ddagger}$	-15^{*}	-5	-18^{*}	$-15^{\dagger}$	0.49	$0.47^{\dagger}$	0.54^{*}	0.51	0.05^{*}	$0.04^{\dagger}$
2006-2012												
Netherlands	10	5	16^{*}	$14^{\dagger}$	7	10	0.47	0.49	$0.46^{\dagger}$	$0.46^{\dagger}$	-0.02	-0.03
Luxembourg	24^{*}	22^{*}	4	-0	-20^{*}	-23^{*}	0.43^{*}	0.43^{*}	$0.48^{\ddagger}$	0.49	$0.05^{\dagger}$	0.06^{*}
Germany	26^{*}	23^{*}	26^{*}	$12^{\ddagger}$	-1	-11	0.42^{*}	0.43^{*}	0.42^{*}	$0.47^{\ddagger}$	0.00	0.04
Belgium	39^{*}	29^{*}	28^{*}	22^{*}	$-12^{\dagger}$	-7	0.39^{*}	0.42^{*}	0.42^{*}	0.44^{*}	$0.03^{\ddagger}$	0.01
Austria	7	5	-4	-1	-11	-6	0.48	0.49	0.51	0.50	0.03	0.01
France	16^{*}	$12^{\dagger}$	7	3	-9	-9	0.45^{*}	$0.46^{\dagger}$	0.48	0.49	0.02	0.02
Spain	1	3	4	4	3	0	0.49	0.49	0.49	0.49	0.00	0.00
Italy	11	13	$-6^{\ddagger}$	-1	$-16^{\ddagger}$	-14	0.47	0.47	$0.52^{\ddagger}$	0.50	$0.04^{\ddagger}$	0.04
Sweden	1	1	-0	5	-1	4	0.50	0.50	0.50	0.49	0.00	-0.01
Norway	-2	3	1	6	2	3	0.50	0.49	0.50	0.49	0.00	-0.01
Denmark	4	6	-7	-0	-10	-6	0.49	0.48	0.52	0.50	0.03	0.02
United States	$13^{\ddagger}$	11	-3	-8	$-16^{\ddagger}$	$-18^{\dagger}$	$0.46^{\dagger}$	$0.47^{\ddagger}$	0.50	0.52	$0.05^{\ddagger}$	$0.05^{\dagger}$
New Zealand	-16^{*}	$-11^{\dagger}$	-22^{*}	$-8^{\ddagger}$	-5	3	0.54^{*}	$0.53^{\dagger}$	0.56^{*}	$0.52^{\ddagger}$	0.02	-0.01
Canada	-5	-2	-18^{*}	$-9^{\dagger}$	$-13^{\dagger}$	-8	0.51	0.50	0.55^{*}	$0.52^{\dagger}$	$0.04^{\dagger}$	0.02

Notes: Achievement gap measures based on normalized scores are rounded to the nearest integer. Gastwirth Index < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on. $*$, † and ‡ indicate achievement gap and achievement gap changes significance at 1, 5 and 10 percent levels, respectively.

Appendix B Results in Mathematics

Table B.5: Achievement Gap in Math: Native-born vs First Generation Migrants

	Normalized Score						Gastwirth Index					
	TIMSS		PISA		Δ Gap		TIMSS		PISA		Δ Gap	
	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj
2007-2012												
France	77*	46*	79*	42*	1	−4	0.29*	0.37*	0.30*	0.39*	0.01	0.01
Italy	89*	68*	49*	14	−40*	−54*	0.26*	0.32*	0.35*	0.45‡	0.09†	0.13*
Sweden	86*	47*	72*	37*	−14	−10	0.27*	0.37*	0.31*	0.39*	0.03	0.02
Norway	54*	31*	32*	6	−22‡	−25†	0.35*	0.41*	0.41*	0.48	0.06‡	0.07†
Ireland	10	−16‡	14†	−2	4	14	0.47	0.54‡	0.47‡	0.51	−0.01	−0.03
United States	49*	36*	1	−21*	−47*	−56*	0.36*	0.40*	0.50	0.56*	0.14*	0.16*
New Zealand	−2	−16*	−20*	−28*	−18†	−12	0.51	0.55*	0.56*	0.58*	0.05†	0.03
Australia	−5	−13†	−41*	−47*	−36*	−34*	0.51	0.53‡	0.61*	0.63*	0.11*	0.10*

Notes: Achievement gap measures based on normalized scores are rounded to the nearest integer. Gastwirth Index < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on. ∗, † and ‡ indicate achievement gap and achievement gap changes significance at 1, 5 and 10 percent levels, respectively.

Table B.6: Achievement Gap in Math: Native-born vs Second Generation Migrants

	Normalized Score						Gastwirth Index					
	TIMSS		PISA		Δ Gap		TIMSS		PISA		Δ Gap	
	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj
2007-2012												
France	49^{*}	16^{*}	62^{*}	35^{*}	$13^{\ddagger}$	$18^{\dagger}$	0.35^{*}	0.45^{*}	0.32^{*}	0.40^{*}	$-0.04^{\ddagger}$	$-0.06^{\dagger}$
Italy	29^{*}	$11^{\ddagger}$	$17^{\dagger}$	-10	-12	$-20^{\dagger}$	0.41^{*}	$0.46^{\ddagger}$	$0.45^{\dagger}$	0.53	0.04	$0.06^{\dagger}$
Sweden	72^{*}	40^{*}	57^{*}	25^{*}	$-15^{\ddagger}$	-15	0.30^{*}	0.39^{*}	0.34^{*}	0.42^{*}	0.03	0.04
Norway	37^{*}	10	38^{*}	13	0	3	0.38^{*}	$0.45^{\dagger}$	0.40^{*}	0.47	0.02	0.01
Ireland	20^{*}	1	-4	-10	$-25^{\dagger}$	-11	0.44^{*}	0.49	0.50	0.52	$0.06^{\dagger}$	0.03
United States	-4	-13^{*}	-4	-22^{*}	1	$-9^{\ddagger}$	0.51	0.53^{*}	0.51	0.56^{*}	0.00	$0.03^{\ddagger}$
New Zealand	-20^{*}	-25^{*}	-39^{*}	-46^{*}	$-19^{\dagger}$	-21^{*}	0.56^{*}	0.57^{*}	0.60^{*}	0.62^{*}	$0.04^{\dagger}$	$0.05^{\dagger}$
Australia	-26^{*}	-30^{*}	-73^{*}	-69^{*}	-47^{*}	-39^{*}	0.57^{*}	0.58^{*}	0.70^{*}	0.69^{*}	0.13^{*}	0.11^{*}

Notes: Achievement gap measures based on normalized scores are rounded to the nearest integer. Gastwirth Index < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on. $*$, † and ‡ indicate achievement gap and achievement gap changes significance at 1, 5 and 10 percent levels, respectively.

Table B.7: Achievement Gap in Math: Second vs First Generation Migrants

	Normalized Score						Gastwirth Index					
	TIMSS		PISA		Δ Gap		TIMSS		PISA		Δ Gap	
	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj
2007-2012												
France	29^{*}	30^{*}	16	11	-12	-19	0.42^{*}	0.41^{*}	0.45	0.47	0.04	0.05
Italy	60^{*}	54^{*}	32^{*}	$26^{\dagger}$	$-28^{\ddagger}$	$-28^{\ddagger}$	0.33^{*}	0.35^{*}	0.39^{*}	0.41^{*}	0.06	0.06
Sweden	14	6	$15^{\ddagger}$	11	0	5	0.46	0.48	$0.45^{\ddagger}$	0.47	-0.01	-0.02
Norway	16	18	-6	-7	-22	-25	0.46	0.46	0.51	0.51	0.05	0.06
Ireland	-10	$-18^{\ddagger}$	$18^{\ddagger}$	6	$29^{\dagger}$	25	0.53	$0.55^{\ddagger}$	0.46	0.49	$-0.07^{\ddagger}$	-0.06
United States	53^{*}	47^{*}	5	5	-48^{*}	-42^{*}	0.37^{*}	0.38^{*}	0.49	0.49	0.13^{*}	0.11^{*}
New Zealand	17^{*}	10	$18^{\dagger}$	$17^{\dagger}$	1	8	0.45^{*}	0.47	$0.45^{\dagger}$	$0.46^{\dagger}$	0.00	-0.02
Australia	21^{*}	$20^{\dagger}$	32^{*}	21^{*}	11	1	0.44^{*}	$0.45^{\dagger}$	0.42^{*}	0.44^{*}	-0.03	0.00

Notes: Achievement gap measures based on normalized scores are rounded to the nearest integer. Gastwirth Index < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on. $*$, $\dagger$ and $\ddagger$ indicate achievement gap and achievement gap changes significance at 1, 5 and 10 percent levels, respectively.

Table B.8: Achievement Gap in Math: Native-born vs Mixed Migrants

	Normalized Score						Gastwirth Index					
	TIMSS		PISA		Δ Gap		TIMSS		PISA		Δ Gap	
	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj	Raw	Adj
2007-2012												
France	22^{*}	$11^{\dagger}$	25^{*}	20^{*}	3	9	0.44^{*}	$0.47^{\dagger}$	0.43^{*}	0.45^{*}	0.00	−0.02
Italy	$12^{\ddagger}$	10	3	1	−8	−9	0.48	0.48	0.49	0.50	0.01	0.01
Sweden	$10^{\ddagger}$	3	−1	−3	−11	−6	0.47	0.49	0.51	0.51	0.03	0.02
Norway	8	3	$13^{\dagger}$	5	5	2	0.48	0.49	$0.46^{\dagger}$	0.49	−0.01	−0.01
Ireland	-21^{*}	-16^{*}	$-11^{\dagger}$	−5	11	11	0.56^{*}	0.54^{*}	$0.53^{\dagger}$	0.51	−0.03	−0.03
United States	14^{*}	$10^{\dagger}$	1	−7	$-13^{\dagger}$	-16^{*}	0.46^{*}	$0.47^{\dagger}$	0.50	$0.52^{\ddagger}$	$0.04^{\dagger}$	0.05^{*}
New Zealand	-14^{*}	$-8^{\dagger}$	-29^{*}	-16^{*}	-15^{*}	−8	0.54^{*}	$0.53^{\dagger}$	0.58^{*}	0.54^{*}	$0.04^{\dagger}$	0.01
Australia	-12^{*}	−4	-25^{*}	-18^{*}	$-12^{\dagger}$	$-13^{\dagger}$	$0.53^{\dagger}$	0.51	0.58^{*}	0.56^{*}	$0.04^{\dagger}$	$0.05^{\dagger}$

Notes: Achievement gap measures based on normalized scores are rounded to the nearest integer. Gastwirth Index < 0.5 (> 0.5) indicates that immigrant-background children lag behind (outperform) their native peers. Raw gap refers to the observed gap. Adjusted gap accounts for differences in students' characteristics including gender, parental education, number of books at home, and whether students have a computer and a desk to study on. $*$, † and ‡ indicate achievement gap and achievement gap changes significance at 1, 5 and 10 percent levels, respectively.