Integrated vs. Separated Schooling for Refugee Children: Long-Term Effects on Education and Assimilation*

Caterina Pavese, Simon ter Meulen, and Ludger Woessmann†

Abstract

Does direct integration into native schools or temporary placement in separate schools better promote the education and cultural assimilation of refugee children? In the Netherlands, primary-school-aged refugee children are placed in either integrated or separated schools depending on their randomly assigned municipality. By contrast, all secondary-school-aged arrivals are initially separated. We exploit this institutional setting in two-way fixed-effects models with age-at-arrival and municipality fixed effects. Using rich administrative data, we find that integrated schooling reduces educational attainment at age 18, particularly for children of lower-educated parents. At the same time, integrated schooling increases cultural assimilation, measured by hormonal-contraceptive usage among teenage women, particularly for children from non-Muslim-majority countries. A distinctive feature of separated schooling that may account for its positive educational effect is that it allows employing teachers specialized in native-language teaching for non-native speakers.

Keywords: refugee children, schools, integration, educational attainment, contraceptives

JEL classification: I24, I21, J13, J15, O15

September 28, 2026

* We gratefully acknowledge comments and suggestions from Orazio Attanasio, Erich Battistin, Leah Boustan, Christian Dustmann, Francesco Fasani, Christina Felfe, Monique De Haan, Nadine Ketel, Theresa Kuchler, Björn Öckert, Hessel Oosterbeek, Umut Ozek, Erik Plug, Helmut Rainer, Michela Tincani, and seminar participants at the universities of Amsterdam, Bristol, Brno, Cagliari, Galway, Siena, SOFI Stockholm, TUM Munich, the ifo Center for the Economics of Education, the World Labor Conference in Toronto, the German Economic Association in Cologne, the CESifo Area Conference on Labor Economics in Munich, the IWAEE in Catanzaro, the CEP education conference at the LSE in London, the BSE Summer Forum in Barcelona, the DoPE Workshop in Oslo, the AIEL Conference in Milan, and the SIdE in Bertinoro. We are particularly grateful to Nienke Ruijs (Owinsp) and Boudien Bakker (Lowan) for help with understanding refugee education and data in the Netherlands. Data access was provided through Statistics Netherlands (CBS). The contribution of Woessmann is part of the German Science Foundation (DFG) project CRC TRR 190 and the DFG Cluster of Excellence TransforM (EXC-3137/1, Grant no. 533801653).

† Pavese: Università di Bologna, CESifo, caterina.pavese@unibo.it. Ter Meulen: ifo Institute, University of Munich, CESifo, termuelen@ifo.de. Woessmann: University of Munich, ifo Institute; Hoover Institution, Stanford University; CESifo, IZA@LISER, and RFBerlin, woessmann@ifo.de.

1. Introduction

The integration of refugee children remains one of the most pressing challenges facing many Western societies (Bahar, Brough, and Peri (2026)). For these children, schools constitute their first sustained exposure to the host country's language and cultural norms. A central policy question is therefore whether they should be placed directly into mainstream schools or first attend separate specialist schools. In this paper, we study whether this initial schooling decision has lasting effects on refugee children's education and cultural integration.

Conceptually, both integrated and separated schooling arrangements offer potential benefits and downsides. Separated schools for refugee children can provide targeted instruction and focus on language acquisition. They may also help refugee children recover from trauma and allow them to learn alongside peers with similar linguistic and cultural backgrounds. However, the lack of exposure to native students may slow integration by reinforcing social and linguistic separation. Integrated schools, by contrast, offer immediate exposure to native peers. This interaction can support faster language acquisition and help children adapt to the host culture by everyday interaction. However, directly having to follow the native curriculum and local culture may just be too hard and overwhelm newly arriving refugee children.

While most countries adopt a uniform approach to refugee education, the Dutch system provides unique variation. In the Netherlands, municipalities decide whether refugee children who arrive at primary-school age are placed directly into mainstream schools with native peers or first attend separate specialist schools for about one year. The Inspectorate of Education categorizes four school types. In our analysis, we classify refugee schools and migrant schools as separated schooling, whereas fully integrated schools and semi-integrated schools—where refugee children attend school with native peers but are partly taught in separate classrooms—are classified as integrated schooling. All refugee children who arrive at secondary-school age are initially placed in separated schools, regardless of the municipality they are assigned to.

We leverage this unique variation in two-way fixed-effects (TWFE) models. Our identification strategy exploits two features of the Dutch setting. First, upon arrival, refugees are randomly assigned to municipalities, ensuring that school placement is orthogonal to refugee characteristics. However, school regimes are determined by municipalities and may therefore correlate with other elements of local environments that independently affect outcomes of refugee children—a potential source of bias that often goes unaddressed in studies exploiting

random allotment of refugees or migrants to local communities. Thus, second, we leverage the fact that all refugee children who arrive at secondary-school age are placed in separated schools nationwide, generating within-municipality treatment variation by age at arrival. This setting allows us to estimate effects of school types in TWFE models with age-at-arrival and municipality fixed effects. In the spirit of a novel difference-in-differences design, we compare refugee children arriving at primary- and secondary-school age across municipalities that do or do not place primary-school-age arrivals in separated schools.

Several specification tests corroborate this identification design. We show that treatment is balanced with respect to pre-determined characteristics of refugee children such as gender, parental education, and country of origin. Likewise, observed municipality characteristics predict neither the treatment nor the outcome in the TWFE setting. Event-study estimates show no significant differences in trends across students who arrived at different ages during secondary school, equivalent to parallel "pre-trends" in conventional difference-in-differences applications. The design even allows for the inclusion of fixed effects for municipality by immigration year that account for time-varying municipality effects.

Our analysis utilizes rich administrative data from Statistics Netherlands on refugee children who arrived in the Netherlands between 2003 and 2010 when they were aged 6 to 17. The data is unique in that it contains information on students' refugee status, refugee parents' education, and—from the Inspectorate of Education—detailed information on school types and educational outcomes until age 21. To this, we merge data on the usage of various prescription medicines stemming from mandatory health insurance registries. Following Achard (2026), we employ the usage of hormonal contraceptives by immigrant teenage women as a measure of cultural assimilation. We also observe alternative medical outcomes such as usage of anti-inflammatory drugs, antihistamines, and antidepressants.

Our results show that refugee children placed directly into mainstream primary schools are significantly less likely to graduate from secondary school by age 18. The estimated effect is a 7.0 percentage-point decrease in the probability of obtaining a high-school diploma for students assigned to integrated schools compared to those who initially attend a separated school (viewed against a mean of 35.7 percent in our sample). Additional analysis suggests that this effect mainly emerges because early integration increases the likelihood of placement into the lowest secondary-school track, which does not lead to a secondary-school diploma. By contrast, there

are no significant differences in enrollment or dropout. When distinguishing the four underlying school types, outcomes are slightly better in migrant than in refugee schools and in fully than in semi-integrated schools. But the economically and statistically significant difference in the treatment effect on educational outcomes is between separated (refugee or migrant) schools and (semi- or fully) integrated schools, rather than within either group.

The negative impacts of directly integrated schooling on obtaining a diploma and attending the lowest ability track are concentrated among students with lower-educated parents and those from Muslim-majority countries.[1] By contrast, students with tertiary-educated parents and those from non-Muslim-majority countries are more likely to experience grade retention when directly integrated. This pattern reflects a fundamental trade-off in ability-based tracking systems: when students fall behind academically, they face a choice between being assigned to a lower track or repeating a grade. We do not find gender differences in the effects on educational outcomes.

Opposite to the *educational* effects, results suggest that integrated schooling improves the *cultural* assimilation of refugee students. Female students placed directly into integrated schools are more likely to use hormonal contraceptives by age 18, consistent with the interpretation that early exposure to native peers facilitates adaptation to host-country norms. This effect is entirely driven by fully integrated schools and does not apply to semi-integrated schools, where refugee children are still particularly exposed to refugee peers at least during some lessons. Furthermore, the effect is restricted to children from non-Muslim-majority countries. We do not find significant effects on the usage of other medications such as antihistamines and anti-inflammatory drugs, suggesting that results are specific to culturally salient consumption rather than medication use more generally. There is also no treatment effect on the usage of antidepressants.

A key feature of separated schools is that their specialization allows them to employ teachers specifically trained to teach non-native speakers. Municipalities with separated primary schools have substantially more teachers with a qualification to teach Dutch as a second language. The greater availability of these specialized teachers may explain why separated schooling has positive educational effects that do not transfer to cultural assimilation. Correspondingly, analysis of school inspectorate records suggests that separated primary schools are rated higher

[1] Our focus on Muslim-majority origin countries is motivated by the size of the inflow of Muslim migrants into Western Europe over the past two decades and the policy interest in fostering their successful integration into host societies, which is also reflected in the academic literature particularly in the school context (e.g., Bisin et al. (2008), Abdelgadir and Fouka (2020), Dahl et al. (2022), Maurin and Navarrete H. (2023)).

than integrated ones in terms of tailoring instruction to students requiring differentiated support and ensuring students' learning focus. By contrast, the separated and integrated schools that refugee children attend do not differ substantively in the overall quality of their school environments in terms of funding, student body, and future outcomes, both showing highly disadvantaged environments.

Our results contribute to the literature on the education of refugees. Many studies focus on their impact on native students (Gould, Lavy, and Paserman (2009), Geay, McNally, and Telhaj (2013), Ballatore, Fort, and Ichino (2018), Figlio and Özek (2019), Figlio et al. (2024), Facchetti et al. (2026)). Other studies examine how the share of migrant peers affects other migrants, typically using school fixed effects (Jensen and Rasmussen (2011), Boucher et al. (2021), Meschi and Pavese (2023)). Rather than relying on share variation, we exploit institutional variation in school integration practices. Most closely related to our study, Schilling and Höckel (2026) study the placement in preparatory language classes relative to regular classes (both in mainstream schools) of Syrian refugee students in Hamburg. They find negative short-term educational effects of separated classrooms, which is supported by our results on semi- vs. fully integrated schools. Beyond classroom placement within mainstream schools, our educational results suggest that a key distinction is between initial separation in specialist schools and any form of direct integration. Moreover, we examine long-run effects on educational attainment and cultural assimilation, providing a broad perspective on the implications of initial school arrangements. Our result on specialized teachers is consistent with the important role of teachers for the integration of refugee students shown by Tumen, Vlassopoulos, and Wahba (2026).

Our findings on hormonal-contraceptive usage highlight that school integration affects not only educational outcomes but also cultural assimilation. This aligns with a growing literature emphasizing the importance of cultural integration (Boucher et al. (2021), Alan et al. (2021), Tumen, Vlassopoulos, and Wahba (2025), Bailey et al. (2026)). Achard (2026) introduces hormonal-contraceptive usage as a proxy for cultural assimilation, finding limited effects of exposure to native neighbors on cultural outcomes of migrants in the Netherlands. We extend this line of research by examining how school placement at arrival influences assimilation outcomes. Moreover, we show that these effects vary by background, suggesting that the interaction between institutional exposure and family characteristics plays an important role in shaping cultural integration.

Our results also contribute to the literature on educational specialization. Instruction in homogeneous vs. heterogeneous groups is relevant for ability tracking in secondary school (Hanushek and Woessmann (2006), Matthewes (2021)), elite education (Abdulkadiroğlu, Angrist, and Pathak (2014), Ketel et al. (2023)), multi-grading (Leuven and Rønning (2016), Barbetta, Sorrenti, and Turati (2021)), special-needs education (Balestra, Eugster, and Liebert (2022)), and university tutorial groups (Booij, Leuven, and Oosterbeek (2017)). Our results show that the effects of specialization are not unimodal across subgroups and outcomes. Learning in heterogeneous groups may differentially affect educational and assimilation outcomes, implying that the optimal degree of educational specialization depends on the policy objective.

The remainder of the paper proceeds as follows. Section 2 describes the institutional setting, section 3 the data, and section 4 the identification strategy. Section 5 presents our results and section 6 discusses possible mechanisms. Section 7 concludes.

2. Institutional Setting

This section outlines the allocation procedure for refugees and the schooling arrangements for refugee children in the Netherlands, as well as the tracked nature of secondary education.

2.1 The Allocation of Refugees to Refugee Centers in the Netherlands

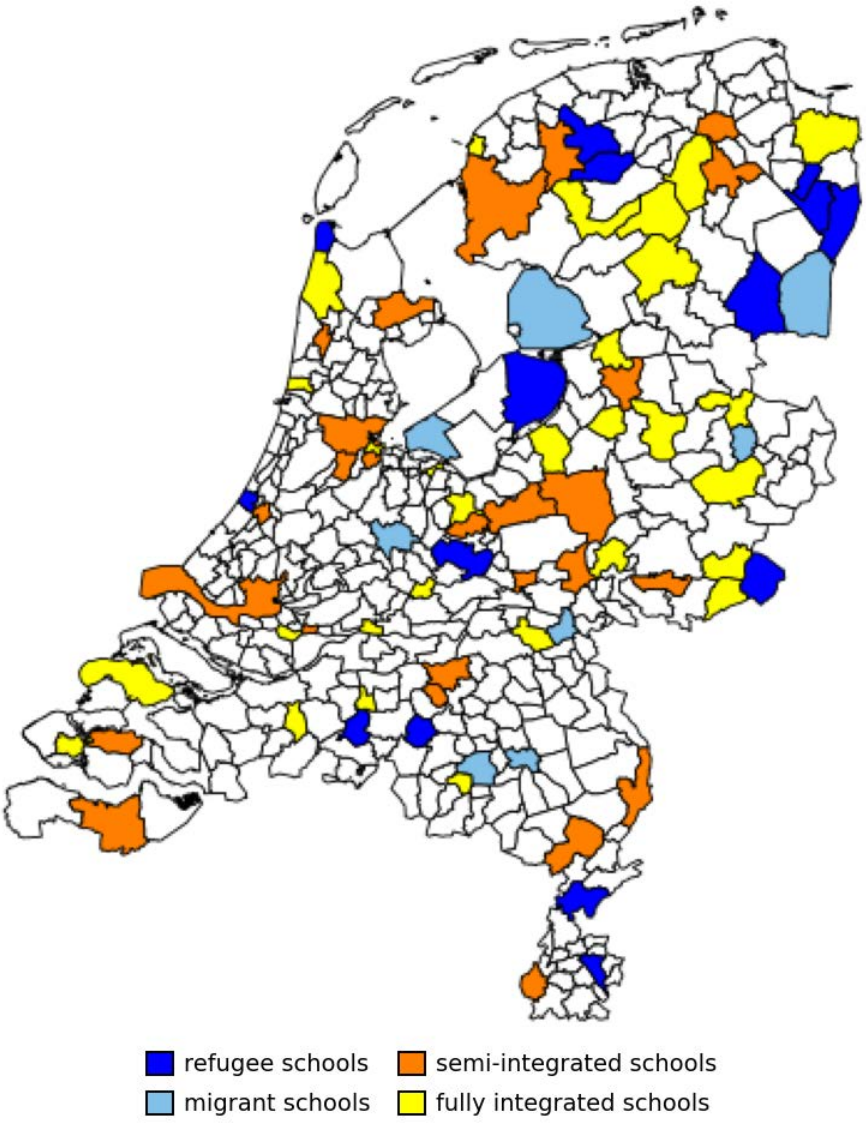
The allocation of asylum seekers in the Netherlands follows a standardized three-step administrative process (Kristiansen et al. (2021)). First, after reporting to local authorities, asylum seekers are sent to one of three national reception centers where they typically spend only a few days. Second, they transfer to refugee centers, determined based on availability, where they await the decision on their refugee status. Third, they transfer to social housing in a local municipality near the refugee center (Vereniging van Nederlandse Gemeenten (2015)), in ways that ensure that each municipality ultimately accepts a number of refugees proportional to its population size (Gerritsen, Kattenberg, and Vermeulen (2018)).

The second step of the allocation process is the focus of our analysis. This stage is centrally managed. Prior to 2016, allocation decisions were made without any consideration of refugees' background characteristics (Gerritsen, Kattenberg, and Vermeulen (2018)). In our analysis sample, refugee children spend on average 11 months (322 days) in these centers.[2] Thus,

[2] Appendix Figure A2 reports the days spent in the initially assigned refugee center by year of immigration.

children spend nearly a year in the educational arrangement determined by their initial random allocation to refugee centers, a substantial amount of time in the formative years. According to Kristiansen et al. (2021), this duration reflects the waiting time both for the asylum decision and for subsequent social housing in the assigned municipality. We study the schooling arrangement during this stay in the refugee center. The locations of these refugee centers during our study period (2003-2010) are indicated by the 125 colored municipalities in Figure 1.

Figure 1: Geographic distribution of refugee centers by school type



Notes: This figure shows the distribution of school types (varying at the municipality level) across the Netherlands. See Appendix A for details. Source: authors' depiction of *Central Orgaan opvang Asielzoekers* (COA) data available at https://www.coa.nl/en/locations.

2.2 The Schooling of Refugee Children

The municipality hosting the refugee center is responsible for the education of the refugee children in the center (Ministerie van Onderwijs (2006b)). Crucially, this responsibility differs substantially between children arriving at primary-school age (ages 6-11) and those arriving at secondary-school age (ages 12-18). For students arriving at secondary-school age, the Ministry of Education mandates enrollment in an International Transfer Class (*Internationale Schakelklas*) for approximately two years (Ministerie van Onderwijs (2017)). These students are separated from native peers and receive specialized instruction tailored to their linguistic and educational needs. In contrast, there are no national regulations or guidelines for the education of primary-school-aged refugees (Ministerie van Onderwijs (2017)), leaving it up to municipal councils to decide the type of educational arrangement.

In its inspection of primary schools, the Inspectorate of Education identifies four types of primary-school arrangements for refugee students (Inspectie van het Onderwijs (2013)). Appendix A lists the formal criteria for these school types. First, in fully integrated schools, refugee students attend classes alongside native students but may receive additional language support. Second, semi-integrated schools are a mixed arrangement in which refugee students might share some classes with Dutch students, receive Dutch language instruction for a few days per week, or are taught in separate classes. In our empirical analysis, we categorize fully and semi-integrated schools as *integrated schooling*, which defines our treatment group.

Third, refugee schools are typically located within the refugee centers and exclusively serve refugee children. Fourth, migrant schools cater to a broader audience, including children of non-native parents, and are generally situated near but outside the refugee centers. Both refugee and migrant schools differ from regular schools by offering tailored educational approaches with a primary focus on acquiring the Dutch language. Most students in these schools are expected to transition to a regular school within one to at most two years. We define these two educational arrangements as *separated schooling*, which constitutes our control group.

Inspectorate reports provide further insight into the functioning of refugee and migrant primary schools (Inspectie van het Onderwijs (2013)). These schools are generally small, typically comprising two to four classes, with student turnover occurring continuously throughout the academic year. They often serve students with highly diverse backgrounds, both in age and prior educational attainment. Some newly arrived students cannot be supported

immediately due to language barriers, while students from Afghan and Somali backgrounds often have not attended school at all and are unfamiliar with independent learning. The primary objective of the instructional approach of refugee and migrant schools is to facilitate Dutch language acquisition. As such, these schools adhere to the principles of Dutch as a Second Language (*Nederlands als Tweede Taal*, NT2), with learning goals tailored to newly arrived students. Moreover, 88 percent of the inspected refugee and migrant schools offer specific programs to teach Dutch social norms and customs.

Figure 1 plots the geographic distribution of the four primary-school types. The figure shows that all school types are widely dispersed across the country, with no segregation in a particular area. Still, it is not necessarily random which municipalities choose which primary-school type. For example, we show in section 4.2 below that larger municipalities and those with larger non-Western migrant shares are more likely to offer separated schooling. The school arrangement chosen by a municipality may thus correlate with other municipality characteristics.

Funding rules do not differ across school types. In the Dutch school system during our period of analysis, schools are funded on a per-student basis using a national formula that gives higher weight to disadvantaged students (Leuven et al. (2007)). In addition, the national government finances refugee education in municipalities through nationwide determined top-ups. Although municipalities and school boards have considerable discretion over how these funds are used (Ministerie van Onderwijs (2005)), eligibility does not depend on the type of school arrangement offered. Through the standard funding formula, school boards receive additional funding for refugee students whose parents have limited education (Ministerie van Onderwijs (2006a)), and municipalities receive temporary funding during refugee students' first year in the country (Ministerie van Onderwijs (2003)).

2.3 Tracking in the Dutch Secondary-Education System

In secondary education, which begins at age 12, there are six ability tracks that prepare students for higher education. These tracks differ in academic content: *Vmbo-basis* (basic track of pre-vocational secondary education) is the most vocational and *Vwo* (pre-university education) is the most academic track. Students who are unable to follow any of the six standard ability tracks may enter *Praktijkonderwijs* (practical education). According to the eligibility criteria to enter *Praktijkonderwijs* defined in national legislation, students must have low IQs and

severe learning deficits in several foundational domains such as reading comprehension and conceptual mathematics (Nederlandse overheid (2003)). The goal of this track is to prepare students for participation in society as decently as possible, mostly conveying basic practical skills such as filling out forms or cooking (Ministerie van Onderwijs (2026)). In what follows, we refer to this ability track for struggling students unable to follow regular education as the "lowest track". During our time of observation, this is the only secondary-education track that does not award a nationally recognized diploma upon completion.

The different tracks have enormous implications for outcomes later in life, and the lowest track is associated with particularly worse outcomes than the others. Table 1 shows how each ability track is associated with later outcomes in terms of educational degrees, earnings, and benefits for native-born students (in the 1993 birth cohort). On average, only 2.8 percent of native students attend the lowest track at age 15 (column 1)—much lower than the 17.1 percent of refugee children in our sample. By the end of compulsory education (age 18), only 12 percent of these lowest-track students have obtained a secondary-school diploma (after changing to a higher track), as opposed to the vast majority of students in all other tracks (column 2). The share subsequently increases but is still below half at age 21 (column 3). At age 30, average earnings of lowest-track students are far below all other tracks (column 4). More than a quarter of lowest-track students do not report any earnings at this age, compared to 12.5 percent for the next-highest track and below 8 percent for all other tracks (column 5). Nearly half of the lowest-track students receive public benefits at age 30, much higher than all other tracks (column 6).

3. Data: Administrative Records and Inspectorate of Education

This section describes our data sources and sample construction. In our primary data environment, Statistics Netherlands (*Centraal Bureau voor de Statistiek*, CBS), we merge data on the universe of refugee students arriving in the Netherlands between 2003 and 2010 from various administrative sources and track them longitudinally via unique anonymized identifiers. Data from the Inspectorate of Education allows us to observe school types and outcomes, as well as school inspection records on schools' pedagogical and classroom practices. Here we describe the main features of the data; Appendix B presents the different steps of the construction of the database in greater detail.

Table 1: Ability tracks and later-life outcomes: diplomas, earnings, and benefits

	Share	Diploma at 18	Diploma at 21	Earnings at 30	Zero earnings at 30	Benefits at 30
	(1)	(2)	(3)	(4)	(5)	(6)
Praktijkonderwijs	2.8	12.0	48.0	25,038	26.1	47.3
Vmbo-basis	12.6	90.7	91.7	37,610	12.5	17.9
Vmbo-kader	15.0	96.7	96.5	43,338	7.8	11.1
Vmbo-gemengd	5.5	96.3	96.6	48,165	4.7	8.2
Vmbo-theoretisch	20.3	94.1	94.3	49,160	5.4	9.0
Havo	22.2	83.2	94.5	56,279	4.1	8.2
Vwo	21.5	75.3	97.7	66,042	3.3	9.2

Notes: The distribution across the ability tracks indicated in the left column is measured at age 15. Numbers refer to the cohort born in the Netherlands between October 1992 and September 1993 (reflecting that *Praktijkonderwijs* enrollment is first recorded in 2007/08). Number of observations: 178,086. Columns 1-3 and 5-6: in percent. Column 4: earnings (in euro) before tax, including zero earnings.

3.1 Refugee Children

We identify refugee children using information on migration motives and residential addresses from administrative micro data provided by Statistics Netherlands. Migration motives are recorded at the individual level for all migrants, allowing us to distinguish refugees from other migrants. We define refugee centers as addresses where at least three refugee mothers are registered simultaneously. Our sample consists of refugee children whose first registered residence corresponds to one of these locations.[3]

We link each student in our sample to their parents using unique individual identifiers, which enables us to retrieve background characteristics such as parental education. This information is primarily recorded by the Dutch benefits office (*Uitvoeringsinstituut Werknemersverzekeringen*, UWV), as refugee parents often rely on social benefits upon arrival. As a result, we observe the educational attainment of at least one parent for 76.6 percent of

[3] Some children are initially registered at a transfer location. For these cases, we use the second registered address. We define transfer locations as addresses with an average refugee stay of less than a month.

children who arrived with a parent. For the 17.9 percent of children who arrived unaccompanied, no information on parental education is available.

3.2 Education Data

We use Inspectorate of Education data, linked to CBS registry data through a unique school identifier, to classify the types of educational arrangements refugee children are assigned to. The Inspectorate's data is filled out by professional inspectors to assess the quality of education provided to non-native students. The Inspectorate of Education began collecting data on refugee education in the 2006/07 school year (Inspectie van het Onderwijs (2013)).

The register data includes information on the attended primary and secondary school, the track and grade enrollment during secondary school, and attainment of a diploma. Our main educational outcome is the attainment of a secondary-school diploma by age 18, which marks the end of compulsory education.[4] We focus on this outcome because attaining a secondary-school degree plays a pivotal role in individuals' lives, serving as a gateway to higher education and improved labor-market outcomes (Angrist and Krueger (1991), Oreopoulos (2006)). Moreover, this outcome is observable for the broadest possible sample of students in our data. As explained below, we observe far fewer students who reach age 21 by the end of the available data range.

Additional educational outcome variables include the academic track attended in grade 9, the probability of lagging behind, and the probability of taking the exit exam from secondary education. We define students as lagging behind if, at age 15, they are enrolled in a grade below grade 9.[5]

For further analysis that investigates possible mechanisms, we make use of three additional data sources. First, we suggest a novel way of identifying teachers qualified to teach Dutch as a second language at the municipality by school-type level, combining social security payroll records with data on detailed degree titles. Second, we access a unique set of school inspection

[4] We do not find a treatment effect on whether refugee students emigrated by age 18, so treatment effects on age-18 outcomes—despite being observed many years after arrival for most refugee students—do not suffer from attrition bias.

[5] Throughout the paper, students' age refers to cohort age, defined according to the October-September school cohort. In the Dutch educational system, cohorts are determined by the October 1 school-entry cutoff (Oosterbeek, ter Meulen, and van der Klaauw (2021)). Rather than aligning with the calendar year, the age range within each grade thus spans from October of one year to September of the following year. Cohort age 15 thus refers to the nominal age at which students typically attend grade 9, which ranges from approximately 14 years and 10 months to 15 years and 9 months.

records in which the Inspectorate of Education scores the use of various pedagogical and classroom practices of primary schools according to common evaluation frameworks during its regular school visits and inspections. Third, to measure the ranking of the quality of the school environments of refugee students' primary schools within each municipality, we use enrollment records of all primary-school students in the Netherlands combined with data on students' funding scores and later university enrollment as well as parents' education and country of birth.

3.3 Medical Records

Using unique individual identifiers, we link our data to the national medical registers, which provide detailed annual information on prescription medication usage among children for the years 2006-2022. The data is not self-reported but derived from official health records, minimizing measurement error. It provides usage of different prescription medications in very fine-grained categorization on an annual basis.

We follow Achard (2026) in using the usage of hormonal contraceptives by immigrant teenage women, a culturally charged consumption, as a measure of cultural assimilation.[6] Widespread usage of hormonal contraception among young women has become a prime example of liberal views on female premarital sexuality in Western societies after the sexual revolution of the 1960s (Goldin and Katz (2002), Bailey (2006)),[7] whereas many non-Western societies and immigrant communities display more conservative gender norms (Algan et al. (2012), Adida, Laitin, and Valfort (2014)). Using data on young immigrant women in a survey on immigrant assimilation in France, Achard (2026) validates the construct by showing that hormonal contraceptive usage is strongly correlated with other assimilation measures commonly used in the literature such as having most of the friendship network from outside the country of origin and fully identifying as a national of the residence country.[8] According to Achard (2026), 86

[6] The literature has also used various other indicators to capture assimilation outcomes of refugees and migrants. A common approach is to use labor-market integration to measure cultural and economic assimilation (see Bahar, Brough, and Peri (2026) for a review). Other widely used methods examine intermarriages, naturalization rates, and names chosen by parents for their children as proxies for cultural assimilation (Abramitzky, Boustan, and Eriksson (2020), Fouka, Mazumder, and Tabellini (2021), Gagliarducci and Tabellini (2025)). Although we track refugee students over a long period, they are currently too young to make these outcomes informative in our context. We also lack target survey data that other studies have used to elicit assimilation indicators such as participation in extracurricular activities, feelings of loneliness, and friendship networks (Dahl et al. (2022)).

[7] Marie and Zwiers (2026) also show that access to contraceptives for minors in the Netherlands reduced teenage pregnancy and increased long-term human capital and work in high-paying jobs.

[8] See Achard (2026) for a more extensive discussion of the construct.

percent of natives, but only 34 percent of first-generation immigrants have ever used hormonal contraceptives by age 20. By contrast, the two groups do not differ in the usage of other medications such as anti-inflammatory drugs, making hormonal-contraceptive usage a good proxy for the adoption of Western views on sexuality and a revealed-preference measure of cultural assimilation. In the Netherlands, girls aged 16 and older do not require parental consent to obtain hormonal contraceptives, and parents cannot access the information. The available data refers to dispensed (not just prescribed) medicine. The medical category includes the pill, patches, injections, and implants (category G03A in the medical classification). To make sure that our measure does not capture arbitrary single usage but persistent usage, it captures whether hormonal contraceptives have been used in at least two years by the age of 18.[9]

We also study the effect of school arrangements on the usage of antidepressants. This may capture whether integrated vs. separated schooling affects the extent to which refugee children are able to overcome trauma from their refugee experiences in the long run. In the Netherlands, both hormonal contraceptives and antidepressants are not sold as over-the-counter medication; both require a prescription from a healthcare provider.

To test whether any effects of schooling arrangements are specific to these outcomes or relate to medical usage more generally, we additionally employ usage of alternative medications. These include anti-inflammatory drugs, antihistamines, corticosteroids, and beta-lactam antibiotics as the next most frequently used drugs among young women in the Netherlands (after hormonal contraceptives, which is the most frequent, see Achard (2026)).

3.4 Sample: Arrival Ages and Years

Our analysis sample refers to refugee children who arrived in the Netherlands at an age between 6 and 17 years. Arriving in this age range requires them to attend either primary or secondary school upon arrival.

For reasons of data availability, we focus on arrivals during the period 2003-2010. The latter year, the last arrival year usable in our analysis, is defined by whether a six-year-old can attain the age of the main outcome, diploma attainment at age 18. As the last data sheet (*Hoogstopgltab*) available to us refers to 2022, the last arriving six-year-olds who can be

[9] A measure of whether hormonal contraceptives have been used at least once yields qualitatively similar, albeit less statistically precise results (see section 5.2.1).

included in our analysis are those who arrived in 2010 (6 + 2022-2010 = 18). When studying diploma attainment at age 21, we lose three cohorts for our analysis.

Several of the data sheets available for our outcome measures, such as central exit exam participation and medical records, start in 2006. For 17-year-olds to be observed in these data, we would have to start our analysis in 2005. However, to increase statistical power, we also include arrival years 2003 and 2004, which allows us to observe children who arrived at a maximum age of 15. A child who arrived in 2003 at age 15 reached age 18 in the 2006 data.

3.5 Descriptive Statistics

Within the 2003-2010 time period, 7,801 refugee children between ages 6 and 17 arrived in the Netherlands, allocated across 125 municipalities.[10] Table 2 reports summary statistics for our analysis sample. The average age at arrival is 11.4 years. The age distribution is relatively balanced across ages, with each age group accounting for between 14.8 and 18.7 percent of the total sample (see Appendix Figure A3).

In terms of demographic characteristics, 45.3 percent of refugee students in our sample are female. One quarter have at least one parent who has completed tertiary education. 17.9 percent arrived without being accompanied by a parent. 70.3 percent arrived from a country of origin with a Muslim majority.[11] The sample is highly ethnically diverse, comprising children from 111 different nationalities. The five most represented countries of origin are Somalia, Iraq, Afghanistan, Iran, and Angola (see Appendix Figure A4). The composition of the refugee population by country of origin over time is characterized by distinctive waves in refugee inflows, reflecting shifting geopolitical conditions (see Appendix Figure A5).

Slightly more than half of the students in our sample (55.3 percent) live in municipalities that have integrated schooling at the primary-school level, the others have separated schooling. Among the four specific types of primary schools, 38.9 percent live in municipalities with refugee schools, 5.8 percent with migrant schools, 17.8 percent with semi-integrated schools, and 37.5 percent with fully integrated schools.

[10] This number excludes 28 singletons who drop out from our analysis when municipality by immigration-year fixed effects are included.

[11] Muslim-majority countries are defined as those where Islam is either the official state religion or the religion of the majority of the population. See Appendix Table A1 for a list of the countries in our sample with a Muslim majority.

Table 2: Descriptive statistics

	Mean	SD	N
	(1)	(2)	(3)
Demographic characteristics			
Age at arrival	11.425	3.587	7,801
Female	0.453	0.498	7,801
Higher-educated parent	0.254	0.435	4,906
Arrived accompanied by parent	0.821	0.383	7,801
Muslim-majority country of origin	0.689	0.463	7,801
Primary-school arrangement of municipality			
Integrated schooling	0.553	0.497	7,801
Refugee schools	0.389	0.488	7,801
Migrant schools	0.058	0.234	7,801
Semi-integrated schools	0.178	0.382	7,801
Fully integrated schools	0.375	0.484	7,801
Outcome measures			
High-school diploma by age 18	0.357	0.479	7,801
High-school diploma by age 21	0.543	0.498	6,365
Hormonal contraceptive usage	0.093	0.291	2,963

Notes: Descriptive statistics are based on our analysis sample, which refers to refugee students aged between 6 and 17 at arrival in the Netherlands during 2003-2010.

The refugee students show comparatively low values on the outcome measures. By age 18, only 35.7 percent of the refugee students have obtained a high-school diploma. This compares to 85.4 percent in the native student population (see Table 1). By age 21, the share increases to 54.3 percent among the refugee students (observed in a smaller sample) and to 93.9 among the native students, indicating some but by no means perfect catch-up. In terms of the cultural-assimilation measure, 9.3 percent of female refugee students have used hormonal contraceptives in more than one year by age 18, compared to 60.8 percent in the native student population.

4. Identification Strategy

This section describes our empirical model and provides specification tests.

4.1 Two-way Fixed-effects Model

Our identification strategy exploits two main features of the institutional setting that governs the schooling of refugee children in the Netherlands.

The first feature is that the allocation of refugees to municipalities is plausibly random (Achard (2026)). As in much of the literature that uses the random allocation of refugees across municipalities as a source of identification,[12] this ensures that treatment is balanced with respect to characteristics of the refugees such as their sociodemographic characteristics or countries of origin. However, random allocation does not ensure that treatment is balanced with respect to (other) characteristics of the municipalities (e.g., Schmandt, Tielkes, and Weinhardt (2025)), because municipalities themselves choose whether to offer direct integration or maintain separate schooling for refugee students in primary schools. As we show in the next section, this policy choice is related to municipality characteristics such as the size of the municipality or migrant population that may also affect refugee outcomes. Similarly, municipalities with higher shares of right-wing voters may be more likely to assign refugee children to separate schools, and prior work by Bracco et al. (2022) and Schilling and Stillman (2024) shows that right-wing political preferences can directly affect the social integration of refugees. Thus, relying just on random allocation can fail to disentangle the effect of local policies from the effect of location.

Therefore, we additionally exploit a second feature of the Dutch institutional system: the difference between arriving at primary- and secondary-school ages. While treatment status differs by municipality for students arriving at the primary level, students arriving at secondary-school age are initially taught in separated schools in all municipalities. This difference by age at arrival introduces treatment variation within municipalities, allowing us to separate the effect of treatment from the effect of place.

To identify the effect of direct integration, we estimate the following two-way fixed-effects (TWFE) specification:

[12] Several Western countries have implemented refugee dispersal policies that assign refugees to local authorities through quasi-random procedures; for examples, see Damm (2014) for Denmark, Couttenier et al. (2019) for Switzerland, and Campo, Giunti, and Mendola (2024) for Italy.

$$y_{imat} = \alpha_a + \beta_m + \sum_{\substack{b=-2 \\ b\neq 0}}^{3} \tau_b \cdot \mathbb{I}\{B_i = b\} \cdot D_{mt} + \gamma_t + \epsilon_{imat} \qquad (1)$$

where y_{imat} is the outcome of refugee student i, who was assigned to municipality m at arrival age a in year t. The treatment indicator D_{mt} equals one if municipality m provides integrated primary schooling for refugees in year t, and zero otherwise.

In the spirit of an event-study analysis, it is easiest to think of age at arrival a as running from the oldest to youngest age. This way, lower levels of the variable (ages 17-12) correspond to arrival ages at which no observation is treated, whereas higher levels correspond to arrival ages at which treatment status differs across municipalities (ages 11-6). To increase statistical power, arrival ages a are grouped into two-year bins: $b = -2$ corresponds to the oldest arrival ages 16-17, $b = -1$ to ages 14-15, and so on. The omitted baseline group consists of students who arrived at ages 12-13, the youngest group arriving at secondary-school age. Positive age bins refer to arrival at primary-school ages, with $b = 3$ corresponding to ages 6-7, the youngest age group in our analysis.

The main coefficients of interest are τ_b. Each τ_b captures the marginal effect of the treatment indicator D_{mt} interacted with arrival-age dummies, denoted by $\mathbb{I}\{B_i = b\}$. The regression allows τ_b to vary across these arrival-age bins to test whether the effects of direct integration are concentrated in the treatment ages, i.e., arrival at primary-school age.

The model includes fixed effects for the two dimensions that define treatment. Age-at-arrival fixed effects α_a capture any general ways in which outcomes may differ by the age at which a refugee child arrives in the country, independent of treatment.[13] Crucially, municipality fixed effects β_m absorb any municipality characteristics, ensuring that the estimated coefficients τ_b capture the causal effect of the educational arrangement, not other features of the local context. The model additionally controls for immigration-year fixed effects γ_t, which account for any differences in the integration of refugees over time. Finally, ϵ_{imat} is the error term. Throughout, standard errors are clustered at the municipality level, which is the level of treatment.

[13] As indicated above, age refers to cohort age throughout to reflect the Dutch school entry cutoff. For example, a student arriving at the age of 7 years and 9 months is assigned a cohort age of 7, whereas a student arriving at 7 years and 11 months is assigned a cohort age of 8.

Identification can be thought of in the spirit of a difference-in-differences model. But unlike a conventional difference-in-differences design that exploits variation over time, our specification leverages variation across arrival-age cohorts. Municipalities that directly integrate newly arrived primary-school students into native schools serve as the treatment group, while those maintaining separate schooling for these students serve as the control group. Within municipalities, the comparison is made across arrival-age cohorts: students who arrived at primary-school ages (6-11) are treated, while those who arrived at secondary-school ages (12-17) serve as the "pre-treatment" group. Each τ_b in equation (1) can be interpreted as a separate difference-in-differences coefficient. For example, for the calculation of τ_1, the treatment group includes students who arrived at ages 10-11, whereas the pre-treatment base group includes those who arrived at ages 12-13. τ_1 is the difference between municipalities with separating and integrating primary schools in the differences between these two age groups. As such, the construction of τ_1 corresponds to the canonical four-cell structure of a difference-in-differences design, and the same applies to the other τ_b coefficients.

The "parallel trends" assumption of our approach is that in the absence of directly integrating schooling, outcomes would evolve similarly across ages at arrival in directly integrating and initially separating municipalities. Although this assumption is inherently untestable, in our context we can test whether τ_{-1} and τ_{-2} differ from zero. This indicates whether cohort trends between integrating and separating municipalities differ across pre-treatment cohorts.

A possible concern with the indicated model is that integrating and separating municipalities may be on differential time trends for other reasons than differences in treatment status. Unlike standard event-study models, the identifying variation in our model is not over time, but across arrival ages. Therefore, treatment varies within municipalities and years, so that we can in fact extend the specification of equation (1) to include a full set of municipality by immigration-year fixed effects β_{mt}:

$$y_{imat} = \alpha_a + \beta_{mt} + \sum_{\substack{b=-2 \\ b\neq 0}}^{3} \tau_b \cdot \mathbb{I}\{B_i = b\} \cdot D_{mt} + \epsilon_{imat} \qquad (2)$$

In this specification, the identifying variation is only between primary- and secondary-school-aged arrivals within the same municipality in the same year.

To summarize our findings into one estimate and increase statistical power, we can combine the τ_b estimates of equation (2) into one coefficient μ:

$$y_{imat} = \alpha_a + \beta_{mt} + \mu \cdot \mathbb{I}[a_i < 12] \cdot D_{mt} + \epsilon_{imat} \qquad (3)$$

where $\mathbb{I}[a_i < 12]$ is an indicator variable for arriving at primary-school age.

It is worth summarizing how our empirical setup explicitly accounts for a wide range of endogeneity concerns. By exploiting additional variation within municipalities rather than relying solely on cross-municipality comparisons, our framework improves on existing cross-sectional refugee allocation designs. The performed TWFE estimations rule out that the estimated treatment effects are affected by the following confounding factors: (i) time-invariant municipality-specific differences (such as sentiments and other policies towards refugees), (ii) age-of-arrival-specific national differences (such as easier integration of children arriving at earlier ages), (iii) time-specific national differences (such as national changes in attitudes over time), and (iv) time-varying municipality-specific shocks or trends that affect children of different arrival ages similarly (such as natural disasters or municipality-level political or religious shocks or trends that do not differentially affect children of different arrival ages).

4.2 Specification Tests

We provide three specification tests to corroborate the validity of the empirical model. First, we show balancing of refugee characteristics in our TWFE setting. While any municipality differences are captured by the municipality fixed effects, refugee children arriving at primary- vs. secondary-school age might in principle differ in predetermined characteristics. However, with random assignment to municipalities, this should not be the case. We test this by regressing the realized treatment status (integrated schooling) on the gender, parental education, and indicators for the five most represented countries of origin of refugee students. The first column of Table 3 shows that (with one marginally significant exception) none of the characteristics predicts assignment to integrated schooling, and the joint *F*-statistic (1.36) fails to reject that the coefficients are jointly insignificant.[14] The same is true when in column 2, we add interactions of each characteristic with an indicator for age at arrival below 12, allowing for separate differences between primary- and secondary-school-aged arrivals.

[14] Similarly, Appendix Table A2 shows individual (rather than joint) balancing for each of the refugee characteristics when these variables are separately predicted by the integrated-schooling treatment.

Table 3: Balancing test

	(1)	(2)
Female	-0.001	0.001
	(0.003)	*(0.003)*
Female x $\mathbb{I}[a_i < 12]$		-0.003
		(0.003)
Higher-educated parent	-0.001	-0.002
	(0.010)	*(0.008)*
Higher-educated parent x $\mathbb{I}[a_i < 12]$		0.0005
		(0.004)
Parental education missing	-0.001	-0.005*
	(0.002)	*(0.003)*
Parental education missing x $\mathbb{I}[a_i < 12]$		0.006
		(0.004)
Afghanistan	-0.003	-0.001
	(0.006)	*(0.006)*
Afghanistan x $\mathbb{I}[a_i < 12]$		-0.002
		(0.002)
Angola	-0.003	0.007
	(0.005)	*(0.007)*
Angola x $\mathbb{I}[a_i < 12]$		-0.019*
		(0.011)
Iran	-0.002	0.003
	(0.001)	*(0.004)*
Iran x $\mathbb{I}[a_i < 12]$		-0.008
		(0.006)
Iraq	-0.0004	0.007
	(0.005)	*(0.006)*
Iraq x $\mathbb{I}[a_i < 12]$		-0.013
		(0.009)
Somalia	-0.014*	-0.011*
	(0.007)	*(0.006)*
Somalia x $\mathbb{I}[a_i < 12]$		-0.005
		(0.007)
Age-at-arrival fixed effects (FE)	Yes	Yes
Municipality x immigration year FE	Yes	Yes
F-test	1.36	1.01
Observations	7,801	7,801

Notes: This table shows estimation results from balancing tests. Dependent variable: indicator variable for realized treatment (integrated schooling), i.e., the interaction between integrating municipality and primary-school-age arrival ($D_{mt} \cdot \mathbb{I}[a_i < 12]$). Indicated countries refer to the five main countries of origin. Column 2 interacts each demographic characteristic with an indicator variable for age at arrival below 12. Standard errors adjusted for clustering at the municipality level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent.

Second, we test to what extent municipality characteristics are associated with the treatment and the main outcome. The educational arrangement chosen by a municipality—whether to integrate or separate in primary schools—is not necessarily orthogonal to other municipality characteristics. The first row of Table 4 shows that this is indeed not the case. In our data environment, we observe municipalities' size, average income, and two characteristics related to migration. The larger a municipality and the larger the share of non-Western migrants in the municipality population, the less likely the municipality is to choose integrated schooling. These associations may reflect scope effects, with larger cities and larger migrant populations allowing more easily for the operation of separated schools for refugees and migrants. Similarly, the number of inhabitants, the share of non-Western inhabitants, and the employment share among non-Western inhabitants are all significantly positively associated with our main outcome, obtaining a high-school diploma by age 18 (third row). Together, these associations would bias standard cross-sectional estimates of the policy treatment despite the random allocation of refugees to municipalities (Schmandt, Tielkes, and Weinhardt (2025)).

Importantly, all these associations of municipality characteristics with treatment and outcome become small and statistically insignificant in our TWFE specification (see second and fourth rows of Table 4). That is, when restricting the identifying variation to variation between primary- and secondary-school-aged arrivals within the same municipality, the municipality characteristics do not show a differential effect either on the treatment or on the outcome. These findings support the validity of the TWFE specification, indicating that the combined fixed effects for arrival ages and municipalities (by immigration year) reduce concerns that the estimates capture the effects of other municipality characteristics.

Third, we show the equivalent of parallel "pre-trends" in our setting. When estimating the event-study models of equation (2) for our main outcome measures (see the event-study graphs below), the τ_b coefficients are all insignificant for arrival ages 16-17 and 14-15 (relative to 12-13), indicating no differential trends for refugee students who arrived at different ages during secondary school. That is, outcomes evolve similarly across ages at arrival in separating and integrating municipalities within the window of secondary-school arrival ages where students in both types of municipalities are not treated.

Table 4: Association of municipality characteristics with treatment and outcome

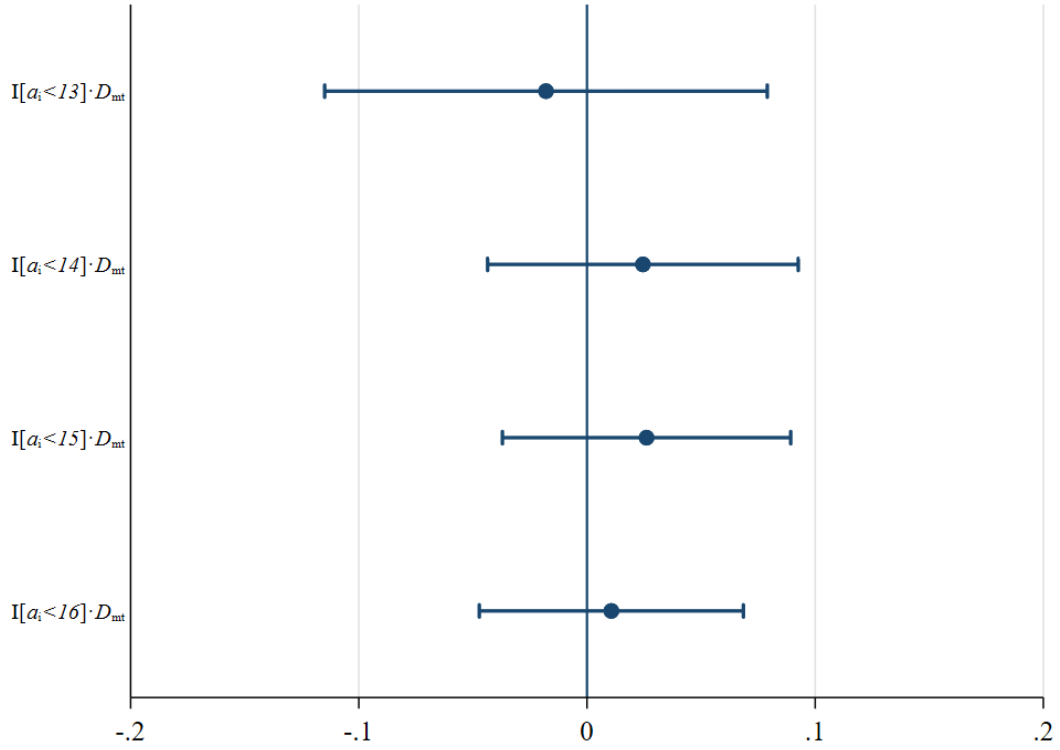
M =	Inhabitants	Average income	Non-Western	Employment non-Western
	(1)	(2)	(3)	(4)
Dep. var.: Integrated schooling				
OLS coefficient on M	-0.028***	0.026	-0.015**	0.001
	(0.010)	*(0.042)*	*(0.007)*	*(0.008)*
TWFE coefficient on M x $\mathbb{I}[a_i < 12]$	-0.0003	-0.001	-0.00005	-0.0003
	(0.0002)	*(0.001)*	*(0.00023)*	*(0.0002)*
Dep. var.: High-school diploma				
OLS coefficient on M	0.005***	0.006	0.003**	0.002**
	(0.002)	*(0.005)*	*(0.001)*	*(0.001)*
TWFE coefficient on M x $\mathbb{I}[a_i < 12]$	0.002	-0.008	0.002	-0.0005
	(0.002)	*(0.009)*	*(0.002)*	*(0.0015)*

Notes: This table shows coefficient estimates on the municipality characteristics indicated in the column headers. OLS coefficient on M: cross-sectional estimation of dependent variable on M (indicated in column header). TWFE coefficient on M x $\mathbb{I}[a_i < 12]$: coefficient on the interaction between M (indicated in column header) and primary-school-age arrival in a two-way fixed-effects model that includes age-at-arrival fixed effects and municipality by immigration-year fixed effects. Dependent variable: top part: indicator variable for realized treatment (integrated schooling), i.e., the interaction between integrating municipality and primary-school-age arrival ($D_{mt} \cdot \mathbb{I}[a_i < 12]$); bottom part: probability of obtaining a high-school diploma by age 18. Municipality characteristics (from Centraal Bureau voor de Statistiek (2006)): column 1: number of inhabitants in the municipality (in 100,000); column 2: average income in the municipality (in euro); column 3: share of non-Western non-natives in the municipality population; column 4: employment share among non-Western non-natives in the municipality. Standard errors adjusted for clustering at the municipality level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent.

In fact, we can run a series of placebo tests where we estimate "treatment effects" for different age cutoffs within the sample of secondary-school-aged arrivals. In Figure 2, we restrict the sample to refugee children who arrived at secondary-school age (12-17) and fictitiously delay the start of secondary schooling from age 12 to ages 13, 14, 15, and 16, respectively. In each case, we estimate the equivalent of equation (3) which has stronger statistical power than the separate pre-trend estimates of the event-study analysis. The period between age 12 and the respective fictitious start of secondary schooling is the new placebo treatment period. The presented coefficients show the interaction between these placebo treatment periods and the indicator for municipalities that provide integrated schooling for primary-school-age arrivals. Results are both statistically insignificant and precisely estimated around zero, indicating no

significant differences in the absence of actual treatment—i.e., in the arrival-age period that corresponds to secondary school, where educational arrangements are uniform across municipalities.

Figure 2: Placebo estimates of "effects" across differently aged secondary-school arrivals



Notes: This figure shows two-way fixed-effects estimates of the "effect" of integrated schooling by age at arrival on the probability of obtaining a high-school diploma by age 18. Each coefficient refers to a separate regression. The estimation sample is restricted to refugee students who arrived between 2003 and 2010 at secondary-school age (ages 12-17) throughout. The four estimates differ by the indicated age cutoff for "treatment" (the placebo intervention). Each specification includes age-at-arrival fixed effects and municipality by immigration-year fixed effects. Lines indicate 95 percent confidence intervals based on standard errors adjusted for clustering at the municipality level.

5. Results

This section presents our results on the effect of integrated schooling on the long-term educational outcomes and cultural assimilation of refugee students, in each case including further examination and heterogeneity analysis.

5.1 Educational Outcomes

A priori, it is unclear whether and how educational arrangements upon arrival will affect refugee children's educational outcomes. Immediately integrated schooling exposes refugee children to native students, which might facilitate educational integration, but following mainstream native curricula may also overstrain the new arrivals. By contrast, initially separated schooling allows for instruction to be targeted to the needs of the refugee students, but it limits exposure to the native language and behaviors. The net effect of these initial arrangements on ultimate educational outcomes is thus an empirical question.

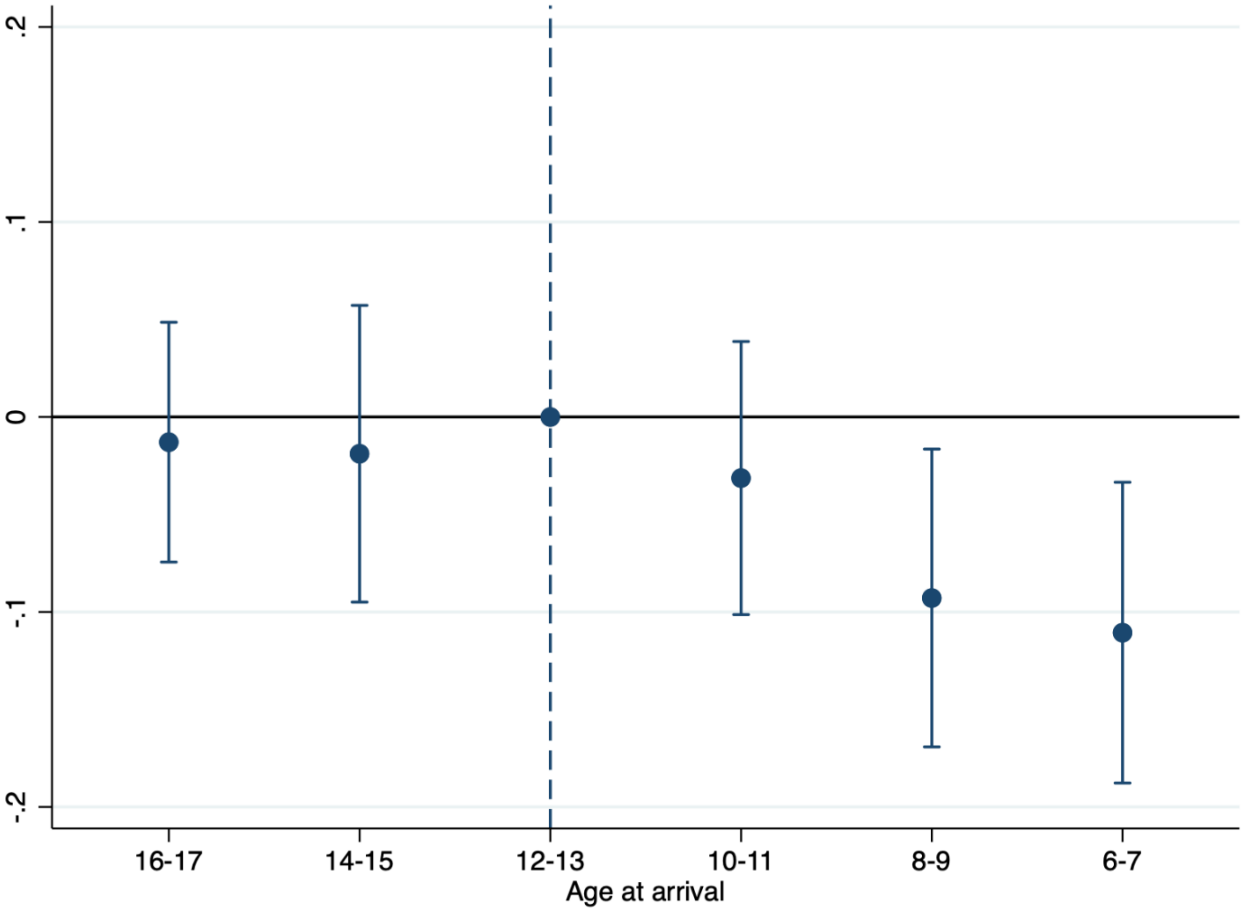
5.1.1 Obtaining a High-School Diploma

Our primary measure of long-term educational outcomes is the probability of obtaining a secondary-school diploma—i.e., graduating from high school—by age 18, the end of compulsory education. The event-study graph in Figure 3 plots the estimates of equation (2), the treatment effect of integrated schooling, for refugee students who arrived at different ages, going from older arrival ages at the left to younger ages at the right.

Results indicate that direct integration in primary schools reduces the probability of obtaining a high-school diploma. The event-study estimates of Figure 3 show that integrated schooling leads to lower educational attainment for refugee children who arrived between ages 6 and 11, compared to those initially enrolled in separated schools. The estimated treatment effect is relatively small and statistically insignificant for arrival ages 10-11 but becomes large and statistically significant for arrival ages 8-9 and 6-7. The smaller estimates for ages 10-11 may either reflect noncompliance with treatment assignment,[15] leading to attenuation bias from measurement error in the treatment variable, or genuinely larger effects of the integration treatment compared to separated schools at younger ages. The estimates for the two youngest age groups are similar and not significantly different from one another, indicating no difference in the integration treatment across this age range.

[15] As children who arrived at age 10 or 11 enter close to the start of secondary school, they are likely to spend little time in integrated schooling in municipalities that provide such educational arrangements. As a result, they may have received only partial or no exposure to integrated schooling, potentially leading to an underestimation of the true effect.

Figure 3: The effect of integrated schooling on obtaining a high-school diploma: Event-study estimates



Notes: This figure shows event-study estimates of the effect of integrated schooling by age at arrival (in two-year bins) on the probability of obtaining a high-school diploma by age 18. The econometric specification is a two-way fixed-effects model that includes age-at-arrival fixed effects and municipality by immigration-year fixed effects. All effects are computed relative to ages 12-13. The estimation sample includes refugee students who arrived between 2003 and 2010. Lines indicate 95 percent confidence intervals based on standard errors adjusted for clustering at the municipality level.

Results do not show significant differences in "pre-trends" across the different age-at-arrival groups at the secondary-school level. Compared to the baseline group (ages 12-13), there are no quantitatively or statistically significant differences for refugees who arrived at ages 16-17 or 14-15. As all age groups are assigned to the same treatment (separation) during the 12-17 age-at-arrival window, the lack of differences is consistent with the treatment effect at primary-school age groups capturing a genuine effect of educational arrangements, rather than any overall differential cohort trends between municipality assignments.

The estimated treatment effect on obtaining a high-school diploma persists at age 21. Descriptive statistics show some catch-up by age 21, with the share of refugee students in our sample who have a diploma increasing from 35.7 percent at age 18 to 54.3 percent at age 21 (see Table 2). While we do not yet observe age-21 outcomes for refugee children who arrived at ages 6-7 in our data, we can estimate the event-study model for arrival ages 8-17. Results in Appendix Figure A6 indicate that the treatment effect of integrated schooling on diplomas persists at the later age. Interestingly, by age 21 treatment effects for arrival ages 10-11 and 8-9 are both statistically significant and quantitatively indistinguishable, suggesting that the smaller point estimate for arrival ages 10-11 indicated in Figure 3 does not hold at older ages. Overall, results are mostly in line with a constant treatment effect of integrated schooling independent of the specific primary-school age at which treatment occurs.

Combined estimates suggest that integrated schooling leads to a seven percentage-point reduction in the likelihood of graduating from high school by age 18. Table 5 reports results of specifications equivalent to equation (3) which summarize the different age-at-arrival estimates of Figure 3 into one estimate. Column 1 shows the baseline TWFE model that includes fixed effects for age at arrival, municipality, and immigration year. Column 2 is our preferred model, which includes a full set of municipality by immigration-year fixed effects. Results are quite similar and statistically highly significant in the two specifications. The estimates suggest that directly integrated schooling decreases the probability of obtaining a diploma at age 18 by 6.7-7.0 percentage points. Compared to the mean of 35.7 percent in our sample, this constitutes a substantial effect. While our specification tests showed balancing of background characteristics and their inclusion as controls can change estimates even when they are irrelevant (e.g., Weinhardt (2026)), inclusion of covariates does not significantly change the estimate in our case, although the point estimate is slightly smaller at -5.9 (column 3). When we estimate a “donut” estimator that excludes arrival ages 11-12—the oldest primary- and the youngest secondary-school arrival age—where compliance may be a problem, the estimate is even larger at 9.8 percentage points (column 4).

Table 5: The effect of integrated schooling on obtaining a high-school diploma: Two-way fixed-effects estimates

	Baseline model	Preferred model	Including controls	Donut estimator	Types of integrated schooling	All underlying school types
	(1)	(2)	(3)	(4)	(5)	(6)
Integrated schooling	-0.067***	-0.070***	-0.059***	-0.098***		
	(0.024)	*(0.025)*	*(0.023)*	*(0.027)*		
Migrant school						0.037
						(0.039)
Semi-integration					-0.085*	-0.080*
					(0.044)	*(0.045)*
Full integration					-0.062**	-0.058**
					(0.018)	*(0.025)*
Age-at-arrival fixed effects (FE)	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Immigration year FE	Yes	Yes	Yes	Yes	Yes	Yes
Municipality x immigration year FE		Yes	Yes	Yes	Yes	Yes
Control variables			Yes			
Observations	7,801	7,801	7,801	6,612	7,801	7,801

Notes: This table presents two-way fixed-effects estimates of the effect of integrated schooling on the probability of obtaining a high-school diploma by age 18. Reported estimates refer to coefficients on the interaction between municipalities with the indicated school types and primary-school-age arrival. The estimation sample includes refugee students who arrived between 2003 and 2010. Control variables: student gender, parental education, family status, and indicators for each of the five main countries of origin (Afghanistan, Angola, Iran, Iraq, Somalia). The donut estimator in column 4 excludes students whose age at arrival is 11 or 12. The base group in column 6 is students allocated to municipalities with refugee primary schools, and in all other columns it is students allocated to municipalities with refugee or migrant primary schools. Standard errors adjusted for clustering at the municipality level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent.

Disaggregated analysis of the four underlying school types confirms the main pattern of separated vs. integrated schooling. Column 5 of Table 5 enters the two types of integrated schooling—semi- and fully integrated schools—separately. Both estimates show significant negative effects and are statistically indistinguishable from one another, warranting their combination into one integrated-schooling treatment. The fact that the point estimates suggest

worse outcomes in semi- than in fully integrated schools is consistent with the negative short-term effect of separated-classroom placement within integrated schools shown by Schilling and Höckel (2026). When also separating the two types of separated schooling in column 6, where the base category now becomes refugee schools (rather than all separated schools), migrant schools do not show a significant treatment effect compared to refugee schools, again warranting their combination into one separated-schooling treatment. While refugee schools are often situated within refugee centers, migrant schools (as well as all integrated schools) are generally not. In principle, not having to commute from refugee centers to schools might be part of the treatment effect in the case of refugee schools. However, the absence of significantly different treatment effects between refugee and migrant schools, as well as the existence of similar treatment effects between migrant and integrated schools, speak against it.[16]

Results also hold in a specification with family fixed effects. In our data, we can observe sibling connections, allowing us to include family fixed effects in a restricted sample of refugee children in families with at least one child arriving at primary- and one at secondary-school age. Such a specification allows us to restrict the identifying variation to children within the same family who arrive at different ages. At the same time, the family-fixed-effects specification may particularly suffer from attenuation from spillovers between siblings, both because of genuine influences among siblings and because of non-compliance with differential treatment between siblings (due to a desire not to separate vulnerable siblings). To reduce this concern, we exclude similarly aged siblings with different intended treatments (arrival ages 11-12) as in the donut specification. In the reduced sample of 1,949 refugee students, our preferred model yields an estimate of -8.4 (significant at the 5 percent level, see Appendix Table A3). Adding family fixed effects to this model hardly changes the estimate (-8.2, significant at the 10 percent level), corroborating the validity of our basic model.[17]

5.1.2 Enrollment and Track Attendance

To further explore why refugee students who are initially taught in integrated schools are less likely to obtain a high-school diploma, we investigate additional measures of educational

[16] When migrant schools are used as the base category, point estimates on semi-integration (-0.117) and full integration (-0.095) become even larger and are both statistically significant at the 5 percent level.

[17] In the hormonal-contraceptive analysis below, the requirement of restricting to female siblings leaves a sample that yields estimates that are too imprecise for meaningful interpretation.

outcomes. To receive a high-school diploma in the Netherlands, it is necessary to participate in the central exam at the end of secondary school. At a technical level, we find that directly integrated refugee students are in fact less likely to take part in the central exit exam (Appendix Figure A7).[18] This raises two options: integrated schooling may either lead to early dropout from high school, so that students are no longer in school at the time of the exams, or to attendance of the lowest academic track, where students do not take the central exam because the track does not lead to a diploma (see section 2.3).

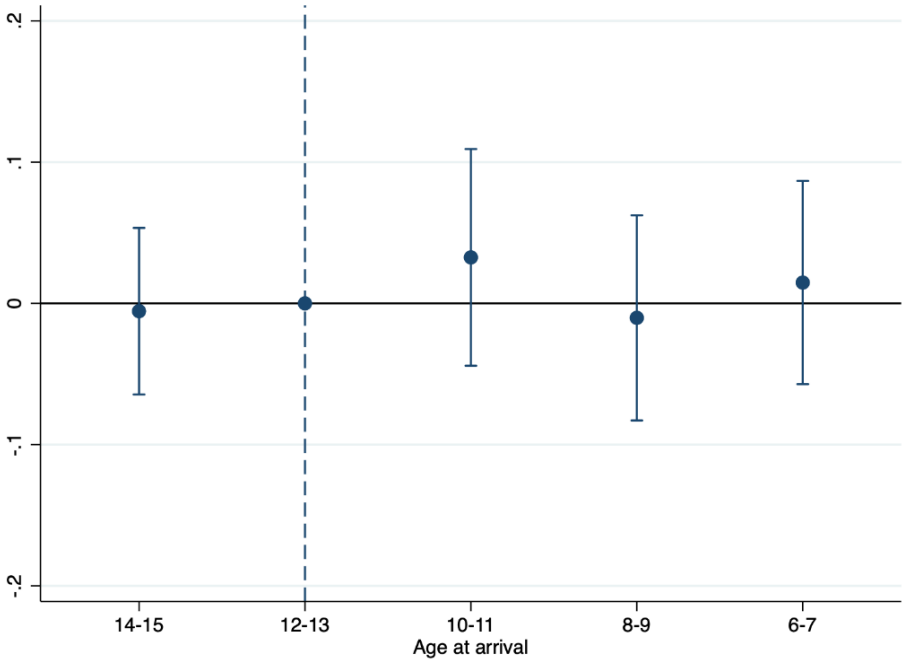
Results indicate that integrated schooling does not increase early school dropout. To study effects on dropout, we measure school enrollment at age 15. This is the nominal cohort age to attend grade 9, which is the grade just before it is possible for any students to take exit exams. The event-study analysis in Panel A of Figure 4 shows that there is no difference in enrollment at age 15 between directly integrated and initially separated refugee students. This is confirmed in column 1 of Table 6, which shows the summarizing TWFE specification of equation (3). These results suggest that directly integrated students are not more likely to drop out of school.

By contrast, integrated schooling significantly increases the likelihood that refugee students attend the lowest ability track. On average, 17.1 percent of the refugee students in our sample attend the lowest track (compared to only 2.8 percent in the native population, see Table 1). Results in Panel B of Figure 4 show that refugee students initially attending integrated schooling are disproportionately more likely to attend the lowest ability track for struggling students in grade 9. The estimate in column 2 of Table 6 suggests that integrated schooling increases lowest-track attendance by 8.4 percentage points. As such, lowest-track attendance could account for the lower rate of obtaining high-school diplomas of directly integrated refugee students. As the lowest track offers a dismal perspective for future options (see section 2.3), these results clearly signal strong academic challenges among many directly integrated refugee students.

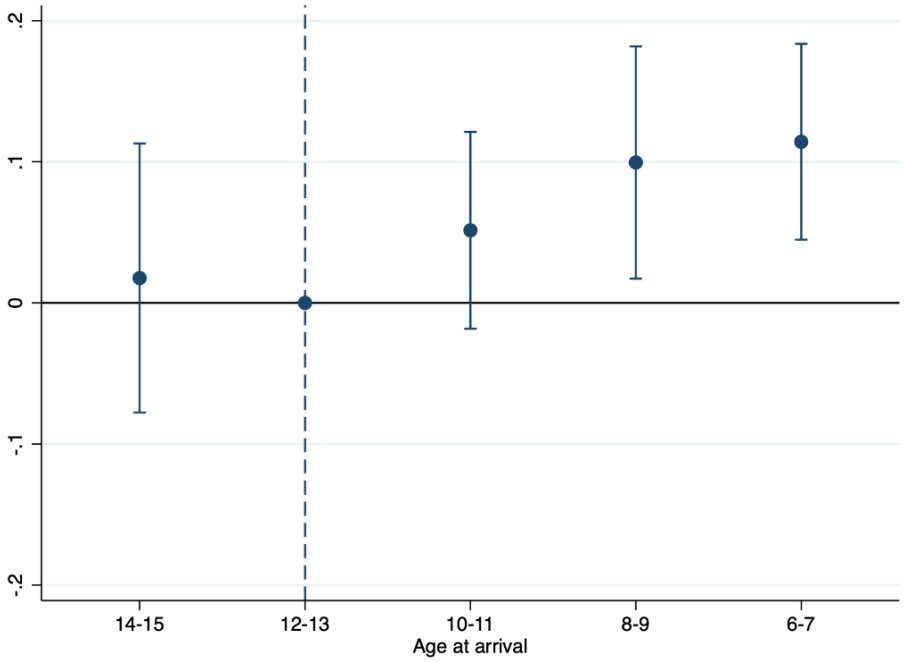
[18] In the summarizing specification equivalent to Tables 5 and 6, the treatment effect on taking the central exam is -0.055 (significant at the 5 percent level). Unfortunately, this treatment effect on central-exam taking, and the absence of large shares of refugee students in the central-exam data, prevents meaningful analysis of treatment effects on exam outcomes such as language test scores.

Figure 4: The effect of integrated schooling on enrollment and lowest-track attendance

A. Enrollment at age 15



B. Attending the lowest track



Notes: This figure shows event-study estimates of the effect of integrated schooling by age at arrival (in two-year bins) on the probability of being enrolled in school at age 15 (Panel A) and on the probability of attending the lowest track in grade 9 (Panel B). The econometric specification is a two-way fixed-effects model that includes age-at-arrival fixed effects and municipality by immigration-year fixed effects. All effects are computed relative to ages 12-13. The estimation sample includes refugee students who arrived between 2003 and 2010. Lines indicate 95 percent confidence intervals based on standard errors adjusted for clustering at the municipality level.

Table 6: The effect of integrated schooling on additional educational outcomes

			Higher track than				
	Enrolled at age 15	Lowest track	*Vmbo-basis*	*Vmbo-kader*	*Vmbo*	*Havo*	Lagging behind
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Integrated schooling	0.014	0.084**	-0.008	-0.007	-0.029	-0.018	0.074***
	(0.030)	*(0.033)*	*(0.033)*	*(0.041)*	*(0.028)*	*(0.011)*	*(0.026)*
Age-at-arrival fixed effects (FE)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality x immigration year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,330	4,204	4,204	4,204	4,204	4,204	6,330

Notes: This table presents two-way fixed-effects estimates of the effect of integrated schooling on different educational outcomes. Reported estimates refer to coefficients on the interaction between integrating municipality and primary-school-age arrival. The estimation sample includes refugee students who arrived between 2003 and 2010. Dependent variables: column 1: enrollment in school at age 15; column 2: attending the lowest track in grade 9; columns 3-6: attending a track higher than the indicated one in grade 9 (see section 2.3); column 7: enrollment below grade 9 at age 15. Standard errors adjusted for clustering at the municipality level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent.

The negative effect of integrated schooling on track attendance is concentrated at the lowest tracking threshold. The prior results suggest that the different educational arrangements have a strong effect on whether refugee students fall below *Vmbo-basis*, the lowest track leading to a high-school diploma. By contrast, results in columns 3-6 of Table 6 show that there are no significant treatment effects on obtaining any of the higher thresholds of the tracked Dutch school system that are above the *Vmbo-basis* track, including the most academic track of pre-university education.

Analysis of grade retention provides additional evidence that directly integrated students are struggling. We measure grade retention by whether students are lagging behind in the sense that at age 15, they are enrolled in a grade below grade 9. On average, 28.6 percent of the refugee students in our sample are lagging behind. Column 7 of Table 6 shows that integrated schooling increases the probability of lagging behind in secondary school by 7.4 percentage points. Grade retention indicates insufficient knowledge at a given grade level, but it can also reflect an extra year of investment; we will come back to this in the following heterogeneity analysis.

5.1.3 Heterogeneity of Effects on Educational Outcomes

In this section, we investigate heterogeneity of treatment effects by gender, parental education, and country of origin. We focus on obtaining a high-school diploma and attending the lowest track as the most consequential educational outcomes. As treatment effects on the probability of enrollment in secondary school at the age of 15 are statistically insignificant in all considered subgroups, we do not report these results. But there are interesting heterogeneities in terms of lagging behind, so we include this aspect in the analysis.

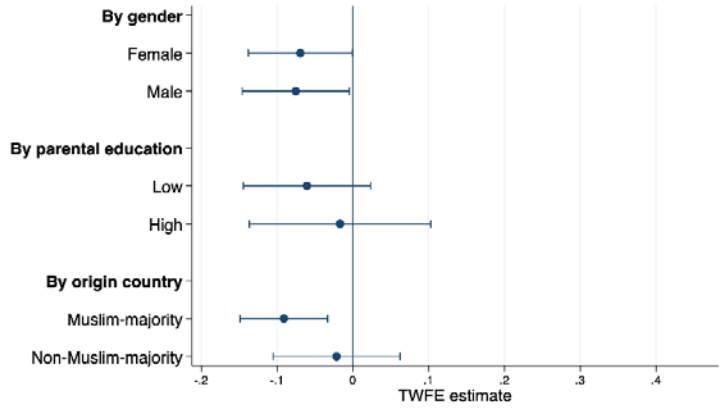
Results hardly differ by gender. Both for girls and for boys, integrated schooling has a significant negative effect on the probability of obtaining a high-school diploma, as seen in the top part of Panel A of Figure 5. Likewise, the effect on attending the lowest academic track is similar for both genders (Panel B). The effect on lagging behind is slightly larger (and significant only) for boys than for girls, but the difference is not statistically significant (Panel C).

By contrast, there are noteworthy differences in treatment effects by parental education. Highly educated parents may, for instance, be better equipped to communicate with teachers in English from the outset. Indeed, as shown in Figure 5, the negative effect of integrated schooling on obtaining a high-school diploma is larger (in absolute terms) for refugee students whose parents do not have higher education,[19] although the estimate is only marginally significant, indicating limitations of statistical power. By contrast, the effect is close to zero for students of higher-educated parents, although the difference between the two groups is not significant. More strikingly, there is a highly significant difference for lowest-track attendance: the treatment effect is strong and significantly positive for children of lower-educated parents, but negative (and not far from statistical significance at conventional levels) for children of higher-educated parents. Interestingly, a negative educational effect for the latter group surfaces when looking at grade retention: integrated schooling strongly increases the probability of lagging behind for children of higher-educated parents, but not of lower-educated parents.

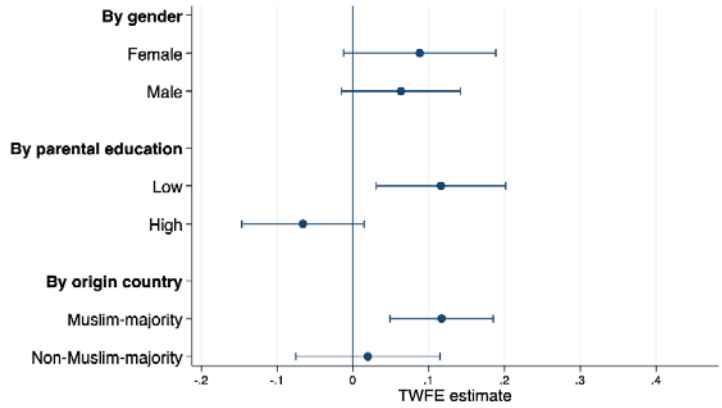
[19] Parental education is classified as high if at least one parent has completed tertiary education.

Figure 5: Heterogeneous effects of integrated schooling on educational outcomes

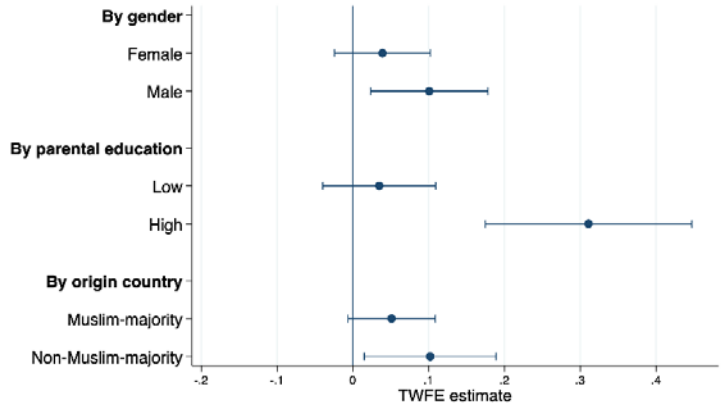
A. Obtaining a high-school diploma by age 18



B. Attending the lowest track



C. Lagging behind at age 15



Notes: This figure shows two-way fixed-effects estimates of the effect of integrated schooling by gender, parental education, and origin country on the probability of obtaining a high-school diploma by age 18 (Panel A), the probability of attending the lowest track in grade 9 (Panel B), and the probability of lagging behind at age 15 (Panel C). Each specification includes age-at-arrival fixed effects and municipality by immigration-year fixed effects. The estimation sample includes refugee students who arrived between 2003 and 2010. Lines indicate 95 percent confidence intervals based on standard errors adjusted for clustering at the municipality level. The subgroup estimates are significantly different from one another at the 1 percent level for parental education in Panels B and C and at the 5 percent level for origin country in Panel B.

This pattern of differences by parental education reflects a basic trade-off in tracked school systems. It aligns with the idea that more educated parents are more aware of the long-term implications of academic tracking in the Dutch system. These parents might prefer their children's placement in a higher-quality track, even if it means that their children fall behind. In educational systems that implement ability tracking, parents often face the decision between having their child repeat the grade or move to a lower academic track (Cockx, Picchio, and Baert (2019)). In this light, grade retention can be interpreted as an additional year of investment in a child's human capital, intended to improve their chances of accessing a more academically demanding track (ter Meulen (2026)).

A similar pattern of results emerges for refugee students coming from Muslim-majority and -minority origin countries. Prior research has shown differences in integration for Muslim migrants (e.g., Bisin et al. (2008), Abdelgadir and Fouka (2020), Dahl et al. (2022), Maurin and Navarrete H. (2023)). Our results show that the negative effect of integrated schooling on obtaining a high-school diploma and the positive effect on attending the lowest track are strong and statistically significant only for refugee students from Muslim-majority countries, whereas the estimates are small and insignificant for refugee students from non-Muslim-majority countries (Figure 5).[20] By contrast, the positive effect on lagging behind is larger and statistically significant only for refugee students from non-Muslim-majority countries. Overall, the results suggest that the negative educational effects of early school integration are most pronounced for children from lower-educated family backgrounds and from countries with Muslim majorities.

5.2 Cultural Assimilation

Educational settings may also influence refugees' cultural assimilation. Most obviously, immediately integrated schooling upon arrival will bring refugee children in touch with native peers early on, exposing them to the local language, behaviors, and norms right away. Direct integration allows refugee children to make native friends more easily, who may transmit local culture through everyday interactions. On this basis, integrated schooling may facilitate cultural assimilation more readily than separated schooling, which may reinforce cultural separation.

[20] The low correlation (0.06) between having lower-educated parents and coming from Muslim-majority countries warrants the separate heterogeneity analyses.

5.2.1 Hormonal Contraceptive Usage

To explore the effect of initial educational arrangements on refugee students' cultural assimilation, we follow Achard (2026) in leveraging the usage of hormonal contraceptives among females as a proxy for the adoption of Western views on sexuality and a revealed-preference measure of cultural assimilation (see section 3.3 for details). We test whether refugee teenage girls who were initially allocated to integrated schooling show systematic differences in the usage of hormonal contraceptives from those initially allocated to separated schooling.[21] To capture persistent usage, the measure refers to usage of hormonal contraceptives in more than one year by age 18.

Results suggest a positive effect of initial integrated schooling on cultural assimilation. The different columns of Table 7 show the TWFE specifications equivalent to those in the respective columns of Table 5. The estimates in the first four columns indicate that female refugee students in directly integrated schools have a 4.3-5.3 percentage points higher probability to use hormonal contraceptives by age 18 than those initially assigned to separated schools, with estimates being marginally significant in most specifications. Event-study estimates confirm consistent positive effects for the arrival-age groups 6-7, 8-9, and 10-11 (Appendix Figure A8).[22]

Specifications separating the underlying school types show that the effect on cultural assimilation is completely driven by fully integrated schools. As columns 5 and 6 of Table 7 indicate, initial placement in fully integrated schools increases hormonal-contraceptive usage by 7.7 percentage points compared to separated schools, whereas hormonal-contraceptive usage does not differ significantly between placement in semi-integrated and separated schools. The effect is large compared to the 9.3 percent average persistent usage in our sample of refugee students. It closes 15 percent of the overall gap in the contraceptive-usage measure between native and refugee students (0.077/(0.608–0.093)).[23]

[21] As the measure of cultural assimilation is available only for females, Appendix Table A4 confirms that the significant negative effect of integrated schooling on the main educational outcome, obtaining a high-school degree, is also given in the female sample available for the contraceptive analysis, even with slightly larger point estimates than in Table 5 for the full sample.

[22] The analysis of medical outcomes does not include arrival ages 16-17. Refugee children who arrived at ages 16-17 in 2003 would already be 19-20 in 2006, the first year for which medical-register data is available.

[23] Results are qualitatively similar for a measure of using hormonal contraceptives at least once (rather than persistently in at least two years), but point estimates are smaller and statistically more imprecise (Appendix Table A5). When estimating the effect of full integration relative to the other three categories, the effect on at-least-once usage is significant at the 6 percent level.

Table 7: The effect of integrated schooling on cultural assimilation

	Baseline model	Preferred model	Including controls	Donut estimator	Types of integrated schooling	All underlying school types
	(1)	(2)	(3)	(4)	(5)	(6)
Integrated schooling	0.048**	0.043*	0.043*	0.053*		
	(0.023)	*(0.024)*	*(0.023)*	*(0.029)*		
Migrant school						-0.035
						(0.031)
Semi-integration					-0.021	-0.025
					(0.036)	*(0.036)*
Full integration					0.077***	0.073***
					(0.022)	*(0.024)*
Age-at-arrival fixed effects (FE)	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Immigration year FE	Yes	Yes	Yes	Yes	Yes	Yes
Municipality x immigration year FE		Yes	Yes	Yes	Yes	Yes
Control variables			Yes			
Observations	2,963	2,963	2,963	2,391	2,963	2,963

Notes: This table presents two-way fixed-effects estimates of the effect of integrated schooling on females' probability of using hormonal contraceptives in at least two years by age 18. Reported estimates refer to coefficients on the interaction between municipalities with the indicated school types and primary-school-age arrival. The estimation sample includes female refugee students who arrived between 2003 and 2010. Control variables: parental education, family status, and indicators for each of the five main countries of origin (Afghanistan, Angola, Iran, Iraq, Somalia). The donut estimator in column 4 excludes students whose age at arrival is 11 or 12. The base group in column 6 is students allocated to municipalities with refugee primary schools, and in all other columns it is students allocated to municipalities with refugee or migrant primary schools. Standard errors adjusted for clustering at the municipality level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent.

The pattern of results suggests that the absence of migrant-group formation in school plays an important role in promoting cultural assimilation. While there are native students in semi-integrated schools, refugee children are separated out in migrant-only classrooms for part or all of their instruction. This may lead to the formation of all-migrant friendship networks at the detriment of migrant-native friendships. Only in fully integrated schools are refugee students attending all classes with native students, which may be important for friendship formation

outside the migrant community, and making native friends may be important for cultural assimilation. Consistent with this interpretation, Schilling and Höckel (2026) argue that one mechanism underlying the negative effects of preparatory language classes is that refugee students are more likely to remain with peers from their initial classroom after transitioning to regular classes. This pattern suggests that initial classroom assignment shapes subsequent friendship networks, thereby limiting sustained interaction with native peers. The interpretation is also consistent with results in Achard (2026), who finds that the cultural assimilation measure is only modestly affected by the share of native children living in the neighborhood, which again may not be enough to prevent migrant network formation. To affect cultural assimilation, fully integrated schooling at arrival may be key.

Together, the results thus show opposing effects of initial educational arrangements for refugee children. In the long term, integrated schooling reduces the educational attainment but increases the cultural assimilation of refugee students, in particular for fully integrated schooling.

5.2.2 Other Medical Outcomes

The availability of rich medical-record data allows us to pursue a number of additional analyses that provide further insight and corroborate the interpretation of our results.

First, we estimate the effect of integrated schooling on the usage of antidepressant medications. In principle, the specialized environments of separated schools might help refugee children to overcome the trauma associated with their refugee experiences. However, results in column 1 of Table 8 show no quantitatively or statistically significant effect of the initial educational arrangement on the usage of antidepressants. The absent net effect might reflect a combination of various and partly opposing effects of initial, educational, and cultural experiences between arrival and age 18.

Second, the absence of treatment effects on other frequently used medications indicates that the treatment effect on hormonal-contraceptive usage is specific to the culturally salient outcome and does not capture medical usage more generally. While hormonal contraceptives are by far the most commonly used medication by young women in the Netherlands, the next four most frequently used medications are anti-inflammatory drugs (for pain and fever reduction), antihistamines (against allergic reactions), corticosteroids (against skin itching and swelling), and beta-lactam antibiotics (for bacterial infections) (Achard (2026)). Results in columns 2-5 of

Table 8 show that usage of these medications is not significantly affected by the initial educational arrangements.

Table 8: The effect of integrated schooling on other medical outcomes

	Young women's most frequently used drugs					
	Anti-depressants	Anti-inflamma-tory drugs	Anti-histamines	Corti-costeroids	Beta-lactam antibiotics	Any medication
	(1)	(2)	(3)	(4)	(5)	(6)
Integrated schooling	0.003	-0.014	0.031*	0.019	0.010	0.038
	(0.005)	*(0.026)*	*(0.019)*	*(0.021)*	*(0.020)*	*(0.034)*
Age-at-arrival fixed effects (FE)	Yes	Yes	Yes	Yes	Yes	Yes
Municipality x immi-gration year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,963	2,963	2,963	2,963	2,963	2,963

Notes: This table presents two-way fixed-effects estimates of the effect of integrated schooling on females' probability of using the indicated medications in at least two years by age 18. Reported estimates refer to coefficients on the interaction between integrating municipality and primary-school-age arrival. The estimation sample includes female refugee students who arrived between 2003 and 2010. Dependent variables: column 1: antidepressant; column 2: anti-inflammatory drugs; column 3: antihistamines; column 4: corticosteroids; column 5: beta-lactam antibiotics; column 6: any entry in the medical records. Standard errors adjusted for clustering at the municipality level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent.

In fact, initially integrated and separated refugee students do not differ significantly in their usage of the medical system overall. Column 6 of Table 8 does not show a significant treatment effect of integrated schooling on the probability of having any entry in the medical records. The positive point estimate partly reflects the treatment effect on hormonal-contraceptive usage, but it is not statistically significant overall. Thus, while the treatment effect on hormonal-contraceptive usage could in principle reflect an impact either on cultural norms or on access to the medical system, these additional results corroborate the interpretation as an impact on cultural assimilation in particular.

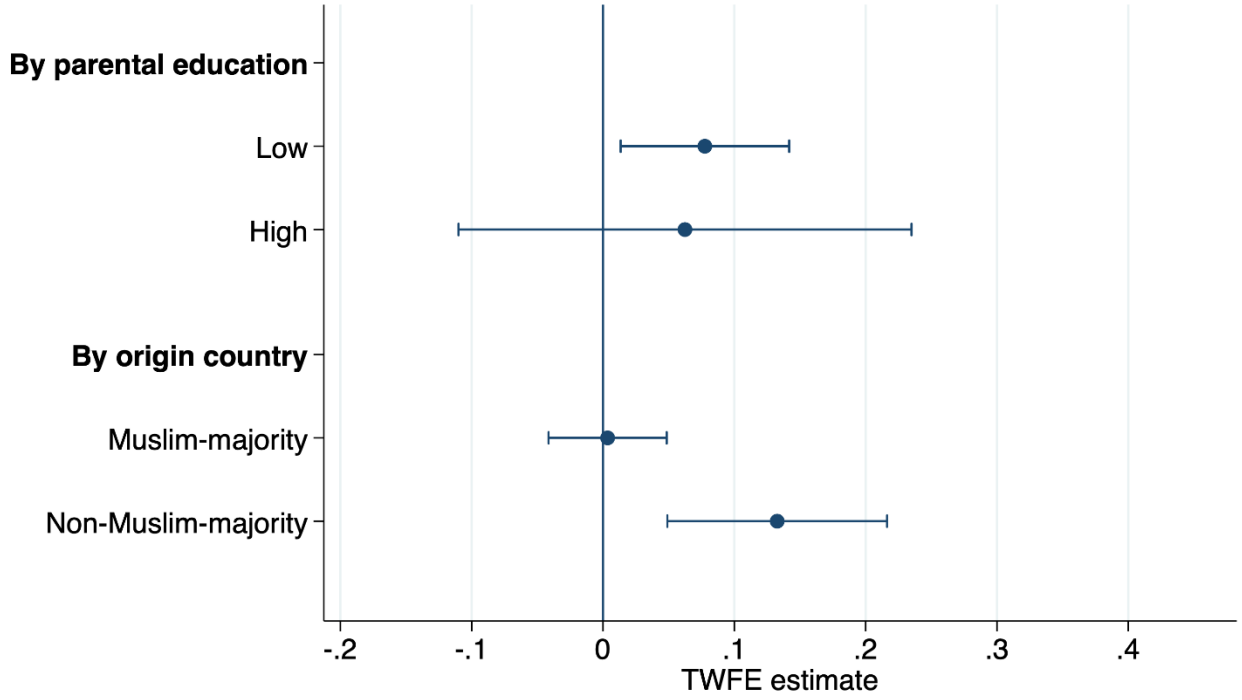
Finally, there is also no effect of integrated schooling on the likelihood of teenage pregnancy. On the one hand, hormonal-contraceptive usage may reduce pregnancies during the

teen years. On the other hand, if contraceptive usage indicates looser sexual norms and usage is not prudent, it may lead to increased teenage pregnancies. Results show that integrated schooling does not affect having a child by the age of 18 (Appendix Figure A9).

5.2.3 Heterogeneity of Effects on Assimilation Outcomes

While we observe our measure of cultural assimilation only for females, we can test for heterogeneity of treatment effects by parental education and country of origin. Results in Figure 6 show no significant differences in the effect of integrated schooling on using hormonal contraceptives between refugee students with or without higher-educated parents, although the coefficient for those with higher-educated parents is rather imprecisely estimated.

Figure 6: Heterogeneous effects of integrated schooling on using hormonal contraceptives



Notes: This figure shows two-way fixed-effects estimates of the effect of integrated schooling by parental education and origin country on females' probability of using hormonal contraceptives in at least two years by age 18. Each specification includes age-at-arrival fixed effects and municipality by immigration-year fixed effects. The estimation sample includes female refugee students who arrived between 2003 and 2010. Lines indicate 95 percent confidence intervals based on standard errors adjusted for clustering at the municipality level. The subgroup estimates by origin country are significantly different from one another at the 1 percent level.

By contrast, the effect of initial educational arrangements on contraceptive usage differs strongly between refugee students from Muslim-majority and -minority countries. The coefficient estimate is virtually zero, precise, and statistically insignificant for children from Muslim-majority countries, whereas it is large and statistically significant for children from non-Muslim-majority countries (Figure 6). Thus, the positive effect of integrated schooling on cultural assimilation is fully restricted to the latter group. The lack of an effect for children from Muslim-majority countries relates to findings in the literature that migrants with Muslim backgrounds tend to have more traditional cultural norms and show slower cultural assimilation (Bisin et al. (2008), Abdelgadir and Fouka (2020), Dahl et al. (2022), Maurin and Navarrete H. (2023)). Differences in initial schooling arrangements do not seem to affect the long-term cultural assimilation for this group of refugee students.

Overall, while *average* effects on cultural assimilation are only marginally significant in most specifications, there is strong evidence that placement in integrated schooling upon arrival increases cultural assimilation, measured by females' contraceptive usage, (a) for refugee children who come from origin countries without a Muslim majority and (b) when integration is fully implemented across all the classrooms that refugee students attend.[24]

6. Specialized Teachers as Distinctive Feature of Separated Schooling

Our results suggest that initially separated schooling is beneficial for the long-term educational (but not assimilation) outcomes of refugee children. In this section, we assemble novel data from a series of alternative sources to provide some insights into possible mechanisms.[25,26] We show that, first, municipalities with separated schools have substantially more teachers specialized in teaching Dutch as a foreign language at the primary level (but not at

[24] When separating semi- and fully integrated schools in the heterogeneity analysis, results are qualitatively similar: the estimate for fully integrated schools is 0.113 (significant at the 1 percent level) for students without higher-educated parents and 0.101 (imprecise) with higher-educated parents. It is 14.6 (significant at the 1 percent level) for children from non-Muslim-majority countries and 0.032 (insignificant) from Muslim-majority countries. The estimates for semi-integrated schools are statistically insignificant in all groups.

[25] While unique and novel in exploring possible mechanisms, the additional data does not allow for an elaborate mediation analysis because it refers to few years only, cannot be linked to individual schools, is not available for secondary schools, and/or is not linked to the full administrative data.

[26] On the cultural assimilation side, the difference between semi- and fully integrated schools suggests native friendship formation as a likely mechanism for the positive effect of integrated schooling on cultural assimilation. Unfortunately, we do not have access to data on friendship networks and cannot study intermarriages or related outcomes as refugee students in our sample are still too young for an informative analysis.

the secondary level) than municipalities with integrated schools. Having teachers who are particularly trained to teach refugee children could explain the better educational outcomes of separated schools. Second, and consistent with an important role of specially qualified teachers, inspectorate records tend to rate separated primary schools better than integrated primary schools in terms of various measures of focused support for students with special learning needs, but not in other dimensions. Third, differences in the overall quality of school environments are unlikely to contribute to the treatment effect, as the separated and integrated schools of refugee children hardly differ in terms of funding, student bodies, and overall future outcomes.

6.1 Teachers Qualified to Teach Dutch as a Second Language

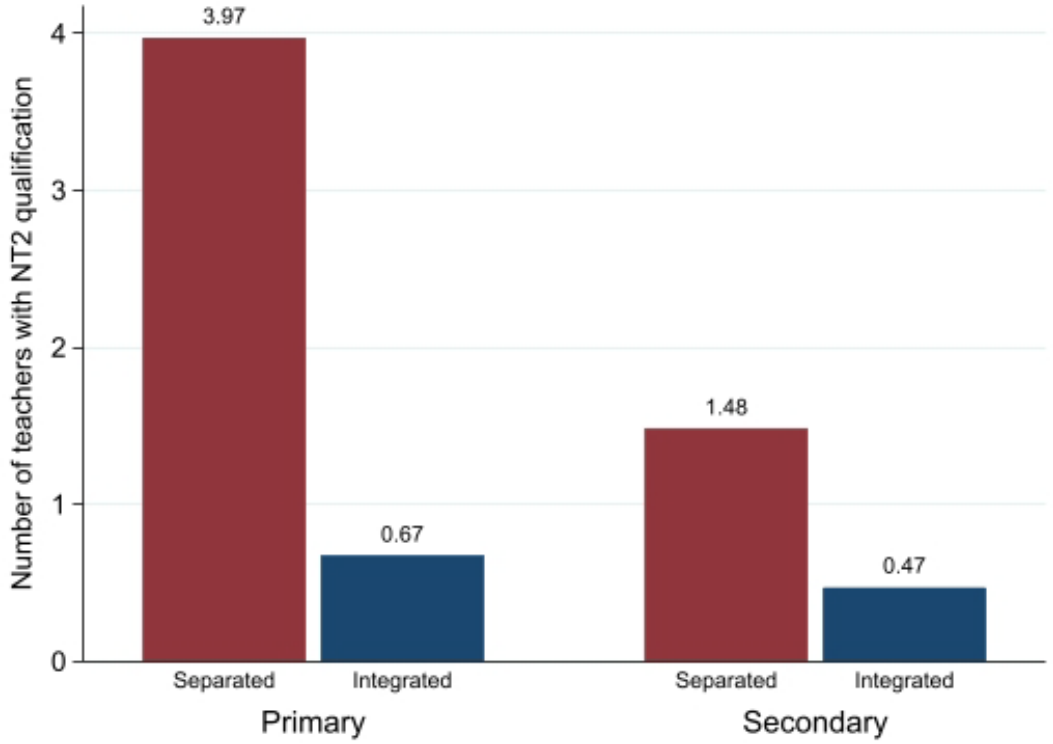
A key feature of separated schooling for refugee children is that it makes it much easier to employ teachers who are specifically trained to support refugee children. In particular, specialized schools are able to hire teachers who are qualified to teach the native language to foreign students. By contrast, in the case of integrated schools, hiring such specialized teachers would not be possible to the same extent. These schools usually do not have enough foreign-language learners to warrant the employment of specialized teachers, and teachers in integrated schools have to teach mainly native students. The effectiveness of separated schooling could thus reflect that their teachers are better trained to help refugee children.

To study the role of specialized teachers, we collected data on the availability of teachers with an NT2 (*Nederlands als tweede taal*—Dutch as a second language) qualification. These teachers have obtained a specialization that has trained them on methodology and best practices for teaching Dutch to non-native speakers. While we cannot study the availability of NT2 teachers at the school level, we combine social security payroll records with detailed degree title data in a novel way to count the number of NT2 teachers currently employed at primary and secondary schools within a municipality (see Appendix B.3 for details).

Results show that municipalities offering separated primary schooling indeed employ a disproportionately higher number of teachers trained to teach Dutch as a second language than municipalities offering integrated primary schooling. The difference is huge, amounting to an average of 4.0 teachers in separating municipalities compared to 0.7 teachers in integrating municipalities (Figure 7). Interestingly, the same does not apply at the secondary-school level, where only separated schooling is provided in both types of municipalities. Here, the difference

(and intensity) is much smaller (1.5 vs. 0.5 teachers), consistent with our institutional setting.[27] Thus, an important feature of separated primary schools is that their specialized nature allows them to hire teachers qualified in teaching the native language to non-native speakers.

Figure 7: Number of specialized teachers per municipality by school type



Notes: This figure shows the number of teachers with an NT2 (*Nederlands als tweede taal*—Dutch as a second language) qualification in municipalities with separated and integrated schooling by level of education. Data refers to 2003-2005. See Appendix B.3 for details.

6.2 Pedagogical Assessment of Schools

To test whether and how the greater availability of specialized teachers also translates into differences in classroom practices for refugee children, we access assessments of various qualitative features contained in school report cards coded by the Inspectorate of Education. As part of its regular school visits and inspections, the Inspectorate of Education conducts extensive

[27] The qualitative patterns are similar when the number of NT2 teachers is expressed relative to the municipalities' total number of teachers or to the number of refugee students in our sample.

assessments aimed at identifying strengths and weaknesses of each school. While questions and score scales differ somewhat across different assessment frameworks that have been used during our observation period, we focus on items that have been used consistently over time for both (separated and integrated) school types and harmonize the scaling of different frameworks and questions to scores ranging from 1 to 4 (see Appendix B.4 for details). To the best of our knowledge, this analysis is the first to exploit these inspection record data for quantitative research purposes. As this analysis is conducted outside the environment of the CBS registry data, the inspection records cannot be matched to our analysis sample and refer to any schools subjected to the inspectorate assessment.[28] As a result, we do not consider fully integrated schools (for which we cannot observe whether they have any refugee students) and thus restrict the analysis to refugee, migrant, and semi-integrated schools.[29] These data limitations—scattered information over different assessment frameworks and years for limited samples of schools—mean that the following analysis can only be considered explorative in nature.

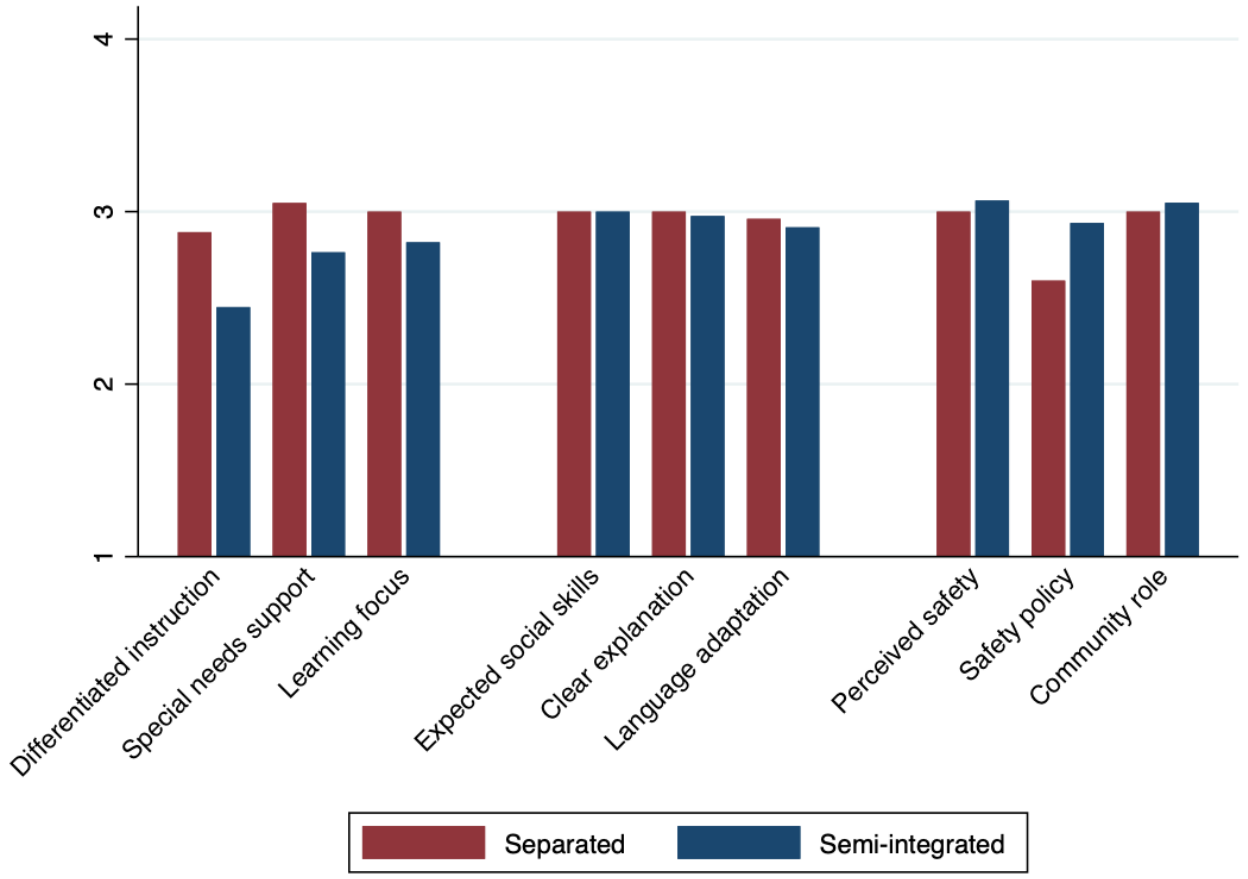
Consistent with an important role of specialized teachers, the inspectorate records suggest that separated schools are rated higher particularly on various measures of individualized pedagogical support. As indicated in Figure 8, separated schools are assessed more positively on average than semi-integrated schools on the dimensions of tailoring instruction to students' needs and providing support for students with special needs, indicating a focus on students requiring particular learning assistance. In addition, inspectorate records indicate that students in separated schools are more likely to be reported as being actively focused on learning. By contrast, while one might have expected otherwise, we do not find evidence that separated schools are rated higher than semi-integrated schools on dimensions such as expected social skills, clear explanation, or language adaptation. In fact, separated schools are rated worse on other dimensions such as perceived safety, school safety policies, and contributing to local communities. While these assessments likely reflect specific challenges of refugee-centered schools, they show that the superior assessment of separated schools does not reflect an

[28] Although the Inspectorate of Education supervises all major educational stages, data provision of inspection records was available only for primary schools.

[29] The negative treatment effect of integrated schooling on educational outcomes is separately statistically significant for both semi-integrated and fully integrated schools and does not differ significantly between the two forms (see Table 5). Focusing on semi-integrated schools likely provides a lower bound for the overall differences between separated and integrated schools. When we also include all fully integrated schools (likely including many that do not serve any refugee children) in the analysis, qualitative results are very similar (see Appendix B.4).

evaluation that is more positive overall but rather focused on pedagogical dimensions that reflect tailored support for students in need.

Figure 8: Pedagogical assessment of primary schools from inspection records by school type



Notes: This figure shows the average scores assigned by the Inspectorate of Education during school inspections for separated and semi-integrated primary schools, respectively. Scores are based on multiple survey items and range from 1 (lowest) to 4 (highest). Both on the raw data and when including fixed effects for assessment frameworks and years, differences in the left panel are statistically significant at the 1 percent level, differences in the middle panel are not statistically significant, and differences in the right panel are statistically significant at least at the 10 percent level. Data refers to 2003-2010. See Appendix B.4 for details.

6.3 Quality of the School Environment

Another potential mechanism for the treatment effect of integrated schooling is that the separated and integrated schools attended by refugee children might differ in the quality of their school environments, leading to different outcomes. To test this, we derive four measures of how

the environments of the primary schools to which refugee students are allocated rank within the distribution of primary schools of the respective municipality (see Appendix B.5 for details). The first measure refers to the input side: the average funding score as used in the school funding formula of the Dutch school system, which is based on various measures of disadvantage of the schools' student body.[30] The next two measures refer to the backgrounds of the students who attend the schools: the shares of students whose parents have a higher education and whose parents were not born in the Netherlands.[31] The final measure refers to the output side: the share of a school's students who later go on to university. For each of these measures, we compute the distribution of schools within a municipality and calculate the within-municipality ranking of the schools that refugee children attend (normalized to a 0-1 scale).[32]

Results indicate that the within-municipality ranking of the refugee children's primary school environments does not differ markedly between municipalities with separated or integrated schooling. As described in Section 2.2, schools receive the same amount of funding at the level of individual students based on a national funding formula (higher for refugee students than for higher-SES native students). Results in Figure 9 show that equal funding levels also apply at the school-board level, the administrative level that makes autonomous funding decisions in the Dutch school system. The administrative school units that refugee children are assigned to also do not differ markedly in the socioeconomic backgrounds of the overall student population that they attract nor in the outcome-based measure.[33] Based on all four measures, refugee children tend to attend schools in highly disadvantaged environments, but this does not differ between municipalities with separated or integrated schooling. While this may not be surprising for separated schools which cater for students with refugee and migrant backgrounds, the results for integrated schooling may indicate that widespread oversubscription of high-quality schools in the Dutch free-choice school system means that refugee children, upon their arrival in

[30] The data refers to the administrative unit that makes the funding decisions. When refugee or migrant schools do not have their own administrative entity, this will be their respective 'mother school' (related school with a joint school board). This will generally also be the school that refugee children attend after the initial separated phase.

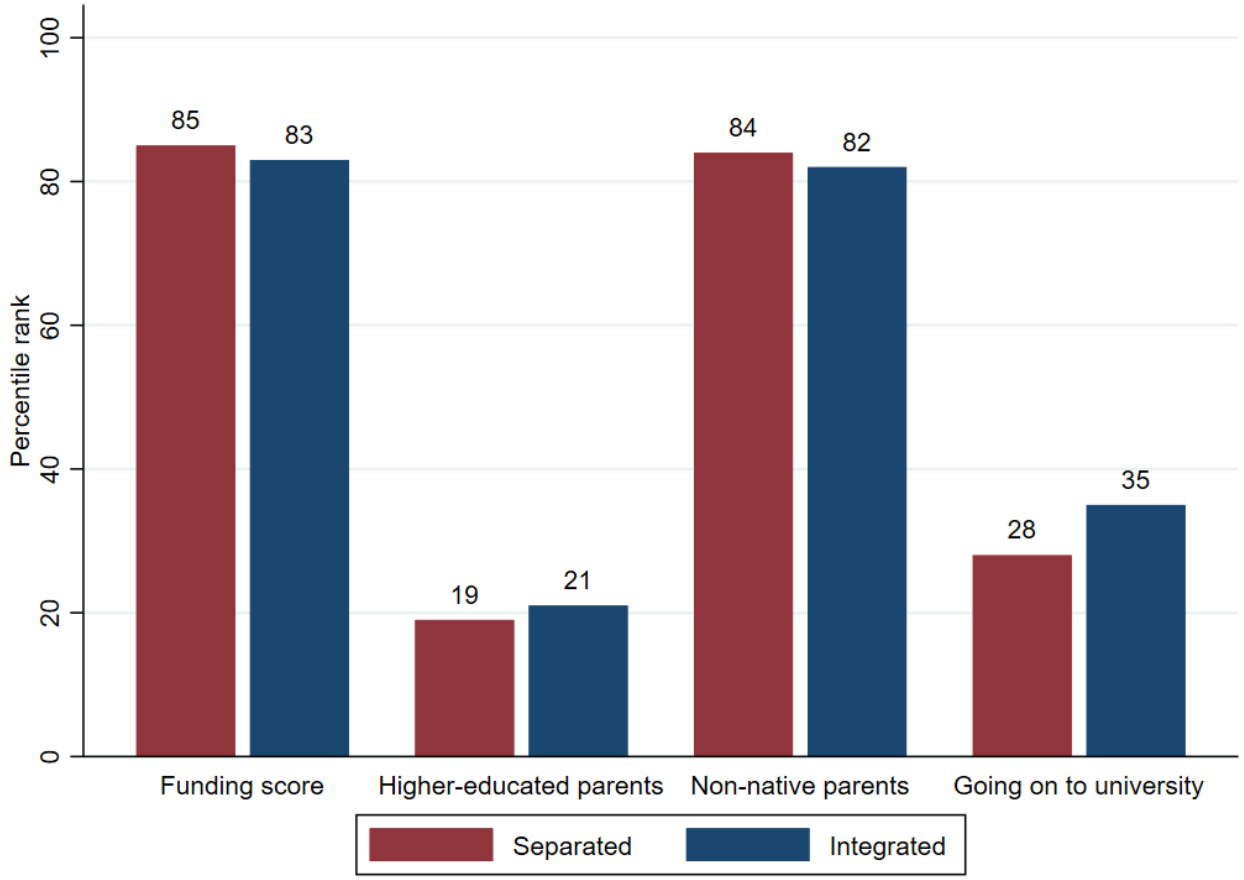
[31] We are not able to study peers at the classroom level as the data does not include classroom identifiers.

[32] As most municipalities generally have at most one secondary school, we cannot do a similar within-municipality ranking at the secondary-school level.

[33] The fact that separated schools are not ranked 100 on the share of non-native parents derives from the fact that for refugee and migrant schools, the data sometimes refers to their respective mother schools that also have students of native parents. Integrated schools even have a slightly larger share of students going on to university, which would work against the negative treatment effect on educational outcomes.

integrating municipalities, are allocated to vacant slots concentrated in relatively low-quality school environments. Overall, the results speak against the quality of school environments playing a major role in the overall treatment effect.

Figure 9: Characteristics of refugee children's primary schools by school type



Notes: This figure shows the percentile rank within the respective municipality (normalized to a 0-1 scale) of measures of the school environment of the primary schools that refugee children attend in municipalities with separated and integrated primary schooling, respectively. Measures: average funding score used in the school funding formula of the Dutch school system, share of students whose parents have a higher education, share of students whose parents were not born in the Netherlands, and share of students who later go on to university. See Appendix B.5 for details.

7. Conclusion

Many Western countries wonder about the best policies to support refugees and to effectively integrate them into their societies in the long run. A fundamental question arises right

upon arrival: Should refugee children be initially separated into specialized schools or directly integrated into mainstream schools? During the large refugee inflows over recent decades, different school systems have chosen different options, and uncertainty remains because a solid evidence base for this decision is missing. Our study sheds light on the long-term effects of different educational arrangements for refugee children—and the results suggest that the answer is not simple.

Our results highlight the complexity of school organization, suggesting a fundamental trade-off between integration and specialization in school systems. Immediately integrated schooling may be conducive to the cultural integration of refugees (measured by female hormonal-contraceptive usage), most importantly when not offering separate classrooms for refugee students. These positive effects are more pronounced for students from non-Muslim-majority countries, suggesting that integrated schools might promote social engagement and adherence to societal norms for refugee students from certain cultural backgrounds. By contrast, directly integrating refugee children into mainstream schools at a young age can lead to negative educational outcomes, particularly for students from low-educated parents and from Muslim-majority countries. These children are more likely to be placed in lower ability tracks, which results in lower chances of graduating from high school. An initial span of separated schooling—which allows for targeted support through specialized teachers—may thus be conducive to the long-term educational prospects of refugee children.

The trade-off between educational and cultural outcomes implies that there are no simple policy conclusions from the results. An overall assessment depends on the relative emphasis given to the different outcome dimensions. Whichever option is chosen, the detailed findings suggest possible considerations to reduce the trade-off by placing particular attention on alleviating the downsides of the chosen option. In integrated systems, the employment of specialized teachers and the provision of more focused support might help to reduce the negative educational implications. In separated systems, encouragement of native exposure in after-school activities and prevention of separation in subsequent schooling options might help to reduce the negative assimilation implications. Furthermore, an overall assessment will also consider potential implications beyond the refugee children themselves. While our analysis provides an in-depth investigation of impacts for refugee children, analysis of spillover effects on native students constitutes a fertile avenue for future research.

References

Abdelgadir, Aala, Vasiliki Fouka (2020). Political Secularism and Muslim Integration in the West: Assessing the Effects of the French Headscarf Ban. *American Political Science Review* 114 (3): 707-723.

Abdulkadiroğlu, Atila, Joshua Angrist, Parag Pathak (2014). The Elite Illusion: Achievement Effects at Boston and New York Exam Schools. *Econometrica* 82 (1): 137-196.

Abramitzky, Ran, Leah Boustan, Katherine Eriksson (2020). Do Immigrants Assimilate More Slowly Today Than in the Past? *American Economic Review: Insights* 2 (1): 125-141.

Achard, Pascal (2026). Exposure to Natives and Cultural Assimilation. *Journal of Human Resources*: forthcoming.

Adida, Claire L., David D. Laitin, Marie-Anne Valfort (2014). Women, Muslim Immigrants, and Economic Integration in France. *Economics & Politics* 26 (1): 79-95.

Alan, Sule, Ceren Baysan, Mert Gumren, Elif Kubilay (2021). Building Social Cohesion in Ethnically Mixed Schools: An Intervention on Perspective Taking. *Quarterly Journal of Economics* 136 (4): 2147-2194.

Algan, Yann, Alberto Bisin, Alan Manning, Thierry Verdier, eds. (2012). *Cultural Integration of Immigrants in Europe*. Oxford: Oxford University Press.

Angrist, Joshua D., Alan B. Krueger (1991). Does Compulsory School Attendance Affect Schooling and Earnings? *Quarterly Journal of Economics* 106 (4): 979-1014.

Bahar, Dany, Rebecca Brough, Giovanni Peri (2026). Refugees' Economic Integration. *Journal of Economic Literature* 64 (3): 950-983.

Bailey, Martha J. (2006). More Power to the Pill: The Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply. *Quarterly Journal of Economics* 121 (1): 289-320.

Bailey, Michael, Drew Johnston, Martin Koenen, Theresa Kuchler, Dominic Russel, Johannes Stroebel (2026). The Social Integration of International Migrants: Evidence from the Networks of Syrians in Germany. *Journal of Political Economy* 134 (4): 1159-1209.

Balestra, Simone, Beatrix Eugster, Helge Liebert (2022). Peers with Special Needs: Effects and Policies. *Review of Economics and Statistics* 104 (3): 602-618.

Ballatore, Rosario Maria, Margherita Fort, Andrea Ichino (2018). Tower of Babel in the Classroom: Immigrants and Natives in Italian Schools. *Journal of Labor Economics* 36 (4): 885-921.

Barbetta, Gian Paolo, Giuseppe Sorrenti, Gilberto Turati (2021). Multigrading and Child Achievement. *Journal of Human Resources* 56 (3): 940-968.

Bisin, Alberto, Eleonora Patacchini, Thierry Verdier, Yves Zenou (2008). Are Muslim Immigrants Different in Terms of Cultural Integration? *Journal of the European Economic Association* 6 (2-3): 445-456.

Booij, Adam S., Edwin Leuven, Hessel Oosterbeek (2017). Ability Peer Effects in University: Evidence from a Randomized Experiment. *Review of Economic Studies* 84 (2): 547-578.

Boucher, Vincent, Semih Tumen, Michael Vlassopoulos, Jackline Wahba, Yves Zenou (2021). Ethnic Mixing in Early Childhood: Evidence from a Randomized Field Experiment and a Structural Model. IZA Discussion Paper 14260. Bonn: IZA Institute of Labor Economics.

Bracco, Emanuele, Maria De Paola, Colin Green, Vincenzo Scoppa (2022). The Spillover of Anti-Immigration Politics to the Schoolyard. *Labour Economics* 75: 102141.

Campo, Francesco, Sara Giunti, Mariapia Mendola (2024). Refugee Crisis and Right-Wing Populism: Evidence from the Italian Dispersal Policy. *European Economic Review* 168: 104826.

Centraal Bureau voor de Statistiek (2006). *Kerncijfers Wijken En Buurten 2006*.

Cockx, Bart, Matteo Picchio, Stijn Baert (2019). Modeling the Effects of Grade Retention in High School. *Journal of Applied Econometrics* 34 (3): 403-424.

Couttenier, Mathieu, Veronica Petrencu, Dominic Rohner, Mathias Thoenig (2019). The Violent Legacy of Conflict: Evidence on Asylum Seekers, Crime, and Public Policy in Switzerland. *American Economic Review* 109 (12): 4378-4425.

Dahl, Gordon B., Christina Felfe, Paul Frijters, Helmut Rainer (2022). Caught between Cultures: Unintended Consequences of Improving Opportunity for Immigrant Girls. *Review of Economic Studies* 89 (5): 2491-2528.

Damm, Anna Piil (2014). Neighborhood Quality and Labor Market Outcomes: Evidence from Quasi-Random Neighborhood Assignment of Immigrants. *Journal of Urban Economics* 79: 139-166.

Facchetti, Elisa, Francesco Fasani, Elisabetta Pasini, Barbara Petrongolo (2026). When Non-Native Speakers Compete for Top Schools: Displacement and Peer Effects in Primary Education. Mimeo.

Figlio, David, Paola Giuliano, Riccardo Marchingiglio, Umut Ozek, Paola Sapienza (2024). Diversity in Schools: Immigrants and the Educational Performance of U.S.-Born Students. *Review of Economic Studies* 91 (2): 972-1006.

Figlio, David N., Umut Özek (2019). Unwelcome Guests? The Effects of Refugees on the Educational Outcomes of Incumbent Students. *Journal of Labor Economics* 37 (4): 1061-1096.

Fouka, Vasiliki, Soumyajit Mazumder, Marco Tabellini (2021). From Immigrants to Americans: Race and Assimilation During the Great Migration. *Review of Economic Studies* 89 (2): 811-842.

Gagliarducci, Stefano, Marco Tabellini (2025). Faith and Assimilation: Italian Immigrants in the US. *Economic Journal* 135 (670): 1814-1851.

Geay, Charlotte, Sandra McNally, Shqiponja Telhaj (2013). Non-Native Speakers of English in the Classroom: What Are the Effects on Pupil Performance? *Economic Journal* 123 (570): F281-F307.

Gerritsen, Sander, Mark Kattenberg, Wouter Vermeulen (2018). Regionale Plaatsing Vergunninghouders En Kans Op Werk. CPB Achtergronddocument 2 mei 2018. Den Haag: Centraal Planbureau.

Goldin, Claudia, Lawrence F. Katz (2002). The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions. *Journal of Political Economy* 110 (4): 730-770.

Gould, Eric D., Victor Lavy, M. Daniele Paserman (2009). Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence. *Economic Journal* 119 (540): 1243-1269.

Hackett, Conrad, Marcin Stonawski, Yunping Tong, Stephanie Kramer, Anne Fengyan Shi (2025). *Dataset of Global Religious Composition Estimates for 2010 and 2020*. Pew Research Center. https://doi.org/10.58094/vhrw-k516.

Hanushek, Eric A., Ludger Woessmann (2006). Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries. *Economic Journal* 116 (510): C63-C76.

Inspectie van het Onderwijs (2013). *Evaluatie Toezicht Op Voorzieningen Voor Nieuwkomers 2011/2012: De Kwaliteit Van Het Onderwijs Aan Nieuwkomers (Type 1 En 2). Rapport Inspectie Van Het Onderwijs*. Hoofdinspecteur primair onderwijs en expertisecentra: dr. A. Jonk.

Jensen, Peter, Astrid Würtz Rasmussen (2011). The Effect of Immigrant Concentration in Schools on Native and Immigrant Children's Reading and Math Skills. *Economics of Education Review* 30 (6): 1503-1515.

Keizer, Bé (2006). *Lumpsumbekostiging Basisschool*. Utrecht: Projectbureau Lumpsum Primair Onderwijs.

Ketel, Nadine, Hessel Oosterbeek, Sandor Sovago, Bas van der Klaauw (2023). The (Un)Importance of School Assignment. IZA Discussion Paper 16591. Bonn: IZA Institute of Labor Economics.

Kristiansen, Marcus H, Ineke Maas, Sanne Boschman, J Cok Vrooman (2021). Refugees' Transition from Welfare to Work: A Quasi-Experimental Approach of the Impact of the Neighbourhood Context. *European Sociological Review* 38 (2): 234-251.

Leuven, Edwin, Mikael Lindahl, Hessel Oosterbeek, Dinand Webbink (2007). The Effect of Extra Funding for Disadvantaged Pupils on Achievement. *Review of Economics and Statistics* 89 (4): 721-736.

Leuven, Edwin, Marte Rønning (2016). Classroom Grade Composition and Pupil Achievement. *Economic Journal* 126 (593): 1164-1192.

Marie, Olivier, Esmée Zwiers (2026). Religious Barriers to Birth Control Access. *Review of Economics and Statistics*: forthcoming.

Matthewes, Sönke H. (2021). Better Together? Heterogeneous Effects of Tracking on Student Achievement. *Economic Journal* 131 (635): 1269-1307.

Maurin, Éric, Nicolás Navarrete H. (2023). Behind the Veil: The Effect of Banning the Islamic Veil in Schools. *Economic Policy* 38 (113): 63-98.

Meschi, Elena, Caterina Pavese (2023). Ability Composition in the Class and the School Performance of Immigrant Students. *Labour Economics* 85: 102450.

Ministerie van Justitie (2002). *Rapportage Vreemdelingenketen September–December 2001, Tevens Jaarrapportage 2001*. Brief van de Staatssecretaris van Justitie, 26 februari 2002. Kamerstukken II 2001/02, 19 637 en 27 557, nr. 652.

Ministerie van Justitie (2008). *Naar Een Effectievere Asielprocedure En Een Effectiever Terugkeerbeleid*. Brief van de Staatssecretaris en de Minister van Justitie, 24 juni 2008. Kamerstukken II 2007/08, 29 344, nr. 67.

Ministerie van Onderwijs, Cultuur en Wetenschap (2003). Besluit Van 18 Juni 2003, Houdende Regels Met Betrekking Tot De Verstrekking Van Een Specifieke Uitkering Aan Gemeenten Ten Behoeve Van Het Onderwijs Aan Vreemdelingen (Besluit Onderwijs Aan Vreemdelingen). *Staatsblad van het Koninkrijk der Nederlanden* 2003: 307.

Ministerie van Onderwijs, Cultuur en Wetenschap (2005). Besluit Van 27 Oktober 2005 Tot Wijziging Van Onder Meer Het Bekostigingsbesluit Wpo En Het Formatiebesluit Wpo in Verband Met De Invoering Van Lumpsumbekostiging in Het Primair Onderwijs. *Staatsblad van het Koninkrijk der Nederlanden* 2005: 553.

Ministerie van Onderwijs, Cultuur en Wetenschap (2006a). Besluit Van 27 Mei 2006 Tot Wijziging Van Het Besluit Bekostiging Wpo En Enkele Andere Besluiten in Verband Met De Herziening Van De Gewichtenregeling Basisonderwijs. *Staatsblad van het Koninkrijk der Nederlanden* 2006: 283.

Ministerie van Onderwijs, Cultuur en Wetenschap (2006b). Wet Van 7 Juli 2006 Tot Wijziging Van De Wet Op Het Primair Onderwijs, De Wet Op De Expertisecentra En De Wet Op Het Voortgezet Onderwijs in Verband Met Wijzigingen in Het Onderwijsachterstandenbeleid. *Staatsblad van het Koninkrijk der Nederlanden* 2006: 340.

Ministerie van Onderwijs, Cultuur en Wetenschap (2017). *Informatiedocument Onderwijs Aan Asielzoekerskinderen*. https://www.poraad.nl/system/files/themas/school_kind_omgeving/informatiedocument_ocw_asielzoekers_en_nieuwkomers.pdf (accessed August 5, 2026).

Ministerie van Onderwijs, Cultuur en Wetenschap (2026). *Wat Is Praktijkonderwijs?* https://www.rijksoverheid.nl/onderwerpen/voortgezet-onderwijs/vraag-en-antwoord/hoe-zit-het-praktijkonderwijs-in-elkaar (accessed August 5, 2026).

Mulder, Lia, Wim Meijnen (2013). *Onderwijsachterstanden in De Bopo-Periode 2009-2012*. Nijmegen: ITS.

Nederlandse overheid (2003). Besluit Rvc's, Regionaal Zorgbudget En Praktijkscholen Met Declaratiebekostiging. *Staatsblad van het Koninkrijk der Nederlanden* 2003: 262.

Oosterbeek, Hessel, Simon ter Meulen, Bas van der Klaauw (2021). Long-Term Effects of School-Starting-Age Rules. *Economics of Education Review* 84: 102144.

Oreopoulos, Philip (2006). Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter. *American Economic Review* 96 (1): 152-175.

Schilling, Pia, Lisa Sofie Höckel (2026). Starting Off on the Right Foot—Language Learning Classes and the Educational Success of Refugee Children. *Economic Journal* 136 (676): 1217-1253.

Schilling, Pia, Steven Stillman (2024). The Impact of Natives' Attitudes on Refugee Integration. *Labour Economics* 87: 102465.

Schmandt, Marco, Constantin Tielkes, Felix Weinhardt (2025). Random Placement but Real Bias. RFBerlin Discussion Paper 161/25. Berlin: Rockwool Foundation Berlin.

ter Meulen, Simon (2026). Long-Term Effects of Grade Retention. *Economic Journal*: forthcoming.

Tumen, Semih, Michael Vlassopoulos, Jackline Wahba (2025). The Power of Language: Educational and Mental Health Impacts of Language Training for Refugee Children. IZA Discussion Paper 17593. Bonn: Institute of Labor Economics.

Tumen, Semih, Michael Vlassopoulos, Jackline Wahba (2026). Training Teachers for Diversity Awareness: Impact on School Outcomes of Refugee Children. *Journal of Human Resources* 61 (1): 84-122.

Vereniging van Nederlandse Gemeenten (2015). *Factsheets Asielzoekers En Vergunninghouders: Praktische Informatie Voor Gemeenten*. Den Haag.

Weinhardt, Felix (2026). Heterogeneous Effects and the Interpretation of OLS Coefficient Movements. Berlin School of Economics Discussion Paper 95. Berlin Berlin School of Economics.

Appendix A: School Types for Refugee Children

The Inspectorate of Education uses the following criteria to determine the school type arrangements for refugee children in primary schools (Inspectie van het Onderwijs (2013)). Figure A1 visualizes the four school types.

Type NK1: Refugee schools

Schools attached to asylum seekers' center.

Criteria:

- The school is attached to an asylum seekers' center (*asielzoekerscentrum, azc*).
- The school is located on the grounds of or close to the asylum seekers' center.

Type NK2: Migrant schools

Schools focused exclusively on initial reception and education of newcomers, possibly including children from asylum seekers' centers.

Criteria:

- The school provides education exclusively to newcomers.
- The school functions independently, but it may be linked to a 'mother school.'
- The school exists for a minimum of two years.
- The school has its own school plan and school guide, or—in the case of a link to a 'mother school'—occupies a clear place in the mother school's documents.
- Students can graduate to a (regular) primary school or to secondary education after a period of 1 to 1.5 years.

Type NK3: Semi-integrated schools

Classes aimed at initial reception/education of newcomers attached to a primary school; classes may have a function for the school concerned, for all schools of a board, for several boards, for the municipality, or for the region.

Criteria:

- The newcomer or first reception classes are linked to a (regular) school and are part of the primary school.
- They are separate classes containing only newcomers.
- The classes are aimed at learning the Dutch language.

- The classes provide education to newcomers for the duration of 1 to 1.5 years.
- The classes provide education for newcomers one or more days a week.
- The classes are the responsibility of the school management or the competent authority, who are responsible for the quality of reception and education in the newcomer classes.

Type NK4: Fully integrated schools

The newcomers are integrated in (regular) schools and are in regular primary-school classes.

Criteria:

- The newcomers are in regular primary-school classes.
- The newcomers may attend a separate class for newcomers for a few half-days.
- In the strict sense, there is no separate provision here, at most adapted education.

Appendix B: Data Appendix

This appendix presents the construction of our database. Section B.1 describes the sample of refugee children and the identification of municipalities' schooling arrangements. Section B.2 presents the construction of background variables and main outcomes. We proceed to discuss the data on Dutch-as-a-second-language teachers (section B.3), on school inspectorate assessments (section B.4), and on the quality of school environments (section B.5). Except for the inspection records, the work is carried out within the secured environment of Statistics Netherlands (CBS). CBS collects data from the different ministries and authorities and makes them available to researchers on an anonymized basis.

B.1 Sample Construction and Treatment Assignment

B.1.1 Refugee Student Sample

The basis of our analysis is the dataset from the Dutch Immigration and Naturalization Service (IND) on the migration motives of all non-Dutch nationals living in the Netherlands. This dataset records asylum seeking as a migration motive, which allows us to distinguish refugees from labor migrants.

From the municipality administration (*gemeentelijke basis administratie*), we obtain basic demographic information, including date of birth, ethnicity, and gender. The date of birth together with the arrival date from the IND data allows us to determine the age at arrival. As the treatment is split between primary and secondary school, we follow the school calendar rather than the calendar year. The school starting age cutoff in the Netherlands is October 1 (Oosterbeek, ter Meulen, and van der Klaauw (2021)), which means that school cohorts start with children born in October and end with those born in the following September. Using the school calendar prevents assigning students within the same age cohort partly to primary school and partly to secondary school but also helps with identifying educational delays.

Our sample consists of students with refugee status and an age at arrival between 6 and 17 whose initial residence in the Netherlands is in a refugee center. We identify refugee centers by identifying addresses where at least three refugee families live at the same time within a year. We use refugee mothers as a proxy for refugee families. The presence of three newly arrived refugee mothers at the same address (called a *rinobjectnummer* in the dataset) indicates that it is a larger facility capable of accommodating refugees. By restricting the sample to having three

mothers each year, we can account for opening and closing of refugee housing. We define these addresses as refugee centers.

B.1.2 School Types

We assign treatment based on the refugee schooling policy of the arrival municipality. A small number of locations serve as processing centers (*Aanmeldcentrum*) or as temporary reception facilities (*Tijdelijke Noodvoorziening Vreemdelingen*, see Ministerie van Justitie (2002, 2008)). Refugees remain at these locations only briefly before being assigned to regular refugee centers according to the assignment procedure described in section 2.1 in the main text. Accordingly, we assign treatment to these students based on their second location. We identify these processing and temporary locations as addresses where refugees on average stay for less than one month.

We identify municipalities' policy regarding refugee schooling on the basis of the presence of a refugee primary school, migrant primary school, or semi-integrated primary schools. The Inspectorate of Education's school register covers the universe of schools. It starts in 2002 and lists whether a school is a refugee school (NK1), migrant school (NK2), or semi-integrated school (NK3). We identify the municipality of each school by linking the encrypted school code (*brin*) to the address identifier (*rinobjectnummer*). The dataset includes start and end dates indicating the period during which the school is active. Combining refugees' arrival date with schools' operating period allows us to determine whether students reside in a treatment municipality. Fully integrated primary schools constitute the residual category after the other three school types are taken into account. The fully integrated schools are regular schools and are inspected with the standard inspection regime of the Inspectorate of Education. As such, they are not explicitly classified in the school universe database from the inspectorate of education, unlike the other school types.

We verify the Inspectorate's classifications using two additional sources. First, the web-scraped LOWAN databases enable us to verify the Inspectorate's classifications in more recent cohorts. The LOWAN databases are compiled from self-reported data provided by schools. Second, using special data on classroom composition from 2014 and 2015, we verify that fully integrated students are indeed placed in mixed classrooms with native students in those years. Additionally, the LOWAN secondary school database allows us to identify secondary schools

serving refugees.[34] Both verifications are consistent with the classification based on the Inspectorate's school registers.

B.2 Background and Outcome Variables

B.2.1 Parental Education and Country of Origin

Apart from refugee students' basic demographic information described above, we use records from the unemployment office (UWV) to identify parental education. The CBS register on the highest educational attainment (*hoogsteopltab)* combines information on educational qualifications from multiple sources, including information from the unemployment office. When benefit seekers apply to the UWV, they can let the UWV verify their diplomas, which can help with finding a new job. Using this information, we are able to identify whether or not a parent has any form of tertiary education.

We use data on the religious composition by country in 2010 in Hackett et al. (2025) to classify countries of origin as having a Muslim-majority or -minority population. Refugees' origin country is registered in the *gemeentelijke basis administratie* (GBA).

B.2.2 Educational and Assimilation Outcomes

The construction of the educational and assimilation outcome variables is relatively straightforward once the different data sources have been obtained. From the subsequent years of the registry of highest obtained education (*hoogsteopltab)*, we observe secondary school diplomas and graduation years. The graduation years, together with the students' birth dates, indicate whether students have obtained a diploma by the age of 18 or 21. Information on exam participation is obtained from the exam registry and indicates whether students have ever participated in the central exam (*examvotab*). The central exam is a requirement for participation in higher education. We measure enrollment at age 15 using the enrollment data (*onderwijsinschrtab*). Enrollment at age 15 serves as a proxy for secondary-school dropout. We use students' grade level at age 15 to determine whether they are at least one grade behind. The

[34] Due to the logic of the difference-in-differences design, we use only the primary school data for municipality treatment assignment. All secondary students, regardless of their assigned municipality, attend the same type of school.

enrollment data also provides the track students attend when they reach grade 9 (the third year of secondary school).

The medical data originates from the public health insurance system and covers prescription medications (*medicijntab)*. These medications are dispensed through licensed pharmacies and are covered by the compulsory basic health insurance. The data only indicates if a particular medication is used in a given year. Table A6 indicates the Anatomical Therapeutic Chemical (ATC) classification codes that we use in our study.

B.3 Dutch-as-a-Second-Language Teachers

The number of Dutch-as-a-second-language teachers per municipality is identified using degrees, educational program titles, and social security payroll records. The CBS register on the highest educational attainment (*hoogsteopltab)* includes degrees from about 80,000 educational programs. Approximately 50 of these programs are variants of *Dutch as Second Language* or *Dutch for Foreigners*. The source of these detailed program titles is the unemployment office. The unemployment office records educational program titles in greater detail than the higher education diploma registries. This implies that teachers need to have been unemployed in order to be in this dataset, and the analysis is only valid under the assumption that teachers in separating and integrating municipalities have equal chances of being unemployed.

The social security payroll records list all current contracts within the Netherlands, as participation is mandatory. While originally designed to determine eligibility and the level of social security benefits (such as unemployment or disability benefits), the data is also used for tax assessments by the tax authority. These records contain information on employers, their locations, employees, and their collective labor agreement. Until 2005, primary schools and secondary schools had different collective labor agreements. We use these labor agreements to identify which company identifiers are in fact primary and secondary schools. We merge the employees of these schools with the data on the highest educational attainment.

We collapse our main dataset into a list of municipalities and their refugee primary school policy. We merge the employee counts from the social security payroll records with these lists based on the listed location. As such, we get per year the number of teachers that are trained to teach Dutch as a second language at the municipality level.

B.4 Pedagogical Assessment of Schools

This section describes the school assessments obtained from inspections by the Inspectorate of Education. During on-site inspections, inspectors complete a standardized survey for each school. The survey items collected during these inspections measure each school's strengths and weaknesses along specific dimensions.

B.4.1 Data Source

The analysis exploits unique administrative inspection records from the Inspectorate of Education (*Inspectie van het Onderwijs*), which, to the best of our knowledge, have not previously been used for quantitative research purposes. The Dutch education system combines substantial school autonomy with a centralized system of quality assurance. The Inspectorate regularly evaluates schools to assess compliance with national educational standards and to verify that public resources are used appropriately.

Schools are assessed using formal supervisory frameworks (*toezichtkaders*). Each framework specifies the domains to be evaluated, the indicators used to measure those domains, and the criteria according to which schools are assessed. During our sample period (2003-2010), the Inspectorate revised its framework several times. Inspections combine information from multiple sources, including administrative records, school documentation, interviews with school staff, and classroom observations. Although the Inspectorate supervises all educational stages, data is available only for primary education (*Primair Onderwijs,* PO).

The main data file used in our analysis, the inspection indicators file, contains one observation for every indicator evaluated during each school inspection. Each record reports (i) the numerical score assigned by the inspector, (ii) the corresponding verbal assessment (e.g., *voldoende, onvoldoende,* "+" or "−"), (iii) the inspection date, (iv) the framework under which the inspection was conducted, and (v) the indicator code and its official Dutch description.

We merge this data with the school characteristics file. This file contains school-level information, including the special-characteristic codes (*SpecKenmerkCodes*) that identify the school types used throughout the paper (refugee, migrant, and semi-integrated schools). We construct and analyze the school assessments outside the CBS microdata environment. This prevents us from identifying fully integrated schools that are attended by refugee students, as they are not marked in the school characteristics file.

B.4.2 Crosswalk of Questions across Frameworks

The Inspectorate revised its supervisory framework several times during our sample period. We use three supervisory frameworks under which both separated and semi-integrated schools are observed: the 2003 supervision framework (PO_2003_01), the 2005 supervision framework (PO_2005_01), and the Quality Card (PO_KK_01). Thematic inspections (e.g., arithmetic or language inspections) and frameworks in which refugee, migrant, or semi-integrated schools are absent are excluded because they do not provide identifying variation for the comparison of interest. The framework revisions lead to changes in both the set of indicators collected and the interpretation of their score scales.

As a consequence of the revisions, schools inspected under different frameworks were evaluated using different indicator codes and, in some cases, slightly different wording, even when the underlying dimension remained unchanged. Directly comparing raw indicators across frameworks would therefore confound changes in measurement with genuine differences in pedagogical and classroom practices. We therefore construct a crosswalk that links framework-specific inspection indicators to a stable set of standardized dimensions capturing school pedagogical and classroom practices.

To obtain a consistent set of measures over time, we aggregate framework-specific indicators into a fixed set of standardized dimensions. The crosswalk is constructed through a systematic review of the official Dutch descriptions of all inspection indicators. Indicators that capture the same underlying educational concept are assigned to a common standardized dimension, irrespective of the supervisory framework under which they were collected. For example, the indicator “Teachers clearly explain” in the 2005 framework and “The explanation is clear” in the 2003 framework are both mapped to the standardized dimension “Clear explanation.” Each framework-specific indicator therefore inherits the standardized label associated with its underlying educational concept.

Only indicators that have a conceptual counterpart in at least one other supervisory framework are included in the crosswalk. Framework-specific items with no comparable equivalent are excluded because they cannot be meaningfully compared across frameworks. Additionally, we restrict attention to indicators that capture differences in pedagogical and classroom practices across school arrangements. We therefore exclude indicators that are

unrelated to the mechanisms investigated in this paper, such as measures of school management or the existence of formal self-evaluation systems.

Occasionally, several indicator codes within a given framework measure different aspects of the same standardized dimension. Retaining each indicator separately would give greater weight to frameworks that collect more detailed information. To avoid these inspections from receiving disproportionate weight in the analysis, we average all sub-indicators belonging to the same standardized dimension within each school inspection. As a result, every school contributes at most one observation for each standardized dimension in every inspection round.

B.4.3 Harmonization of Inspection Scores

Because the numerical coding of the evaluation scale differs across supervisory frameworks, raw inspection scores are not directly comparable over time. In particular, identical numerical values may correspond to different assessments across frameworks. For example, under the Quality Card framework, the numerical ordering is reversed relative to the earlier frameworks.

To ensure comparability, we harmonize all inspection scores onto a common four-point evaluation scale, where 1 denotes inadequate, 2 insufficient, 3 sufficient, and 4 good. Rather than relying on the raw numerical value reported by the Inspectorate, the harmonization is based on the accompanying verbal assessment whose interpretation remains substantially more stable across frameworks.

Inspection frameworks that employ only two evaluation categories (e.g., sufficient/ insufficient, positive/negative, or yes/no) are mapped onto the central categories of the common scale, assigning values of 2 (insufficient) and 3 (sufficient). Observations corresponding to administrative or non-evaluative codes, as well as CITO attainment classifications that do not measure pedagogical and assessment classroom practices, are treated as missing.

This harmonization produces a common scale that is comparable across successive frameworks while preserving the substantive interpretation of the Inspectorate's evaluations.

B.4.4 Sample Construction and Standardized Questions

The sample of schools for which there are consistent assessments derives from four data processing steps. First, we merge the inspection indicators with the school characteristics file using the unique school identifier. This merge is complete for all inspected schools. Schools that

did not receive an inspection during the sample period are not represented in the inspection database and therefore do not enter the analysis.

Second, we merge the inspection records using the described crosswalk. We retain only indicators that both have a conceptually equivalent standardized dimension across frameworks and capture aspects of the pedagogical and classroom practices relevant to our research question. Consequently, we exclude framework-specific indicators that are not comparable over time, as well as dimensions that are orthogonal to the educational experiences of refugee children.

Third, we restrict the sample to schools operating under one of the three refugee educational arrangements observed in the inspection data: refugee, migrant, and semi-integrated schools. Fully integrated schools are excluded because, outside the CBS environment, we cannot identify whether refugee children were enrolled in the school. Consequently, fully integrated schools cannot be distinguished from regular schools that do not serve refugee students. In a robustness analysis, we can include all regular schools.

Fourth, when an inspection framework contains multiple indicators measuring the same underlying dimension, we aggregate these sub-indicators into a single observation at the school-by-inspection-by-question level by averaging the corresponding scores.

Altogether, we retain 19 standardized questions in our analysis. For each of these questions, Table A7 shows the English translation of the wording of the framework-specific items that are aggregated into the respective question, as well as the number of separated and integrated schools for which the respective question is observed.

B.4.5 Differences between Separated and Integrated Schools

To study whether inspectorate assessments of the 19 standardized questions differ between separated and integrated primary schools, we separately regress the harmonized inspection score (ranging from 1 to 4) on an indicator for separated schools as well as fixed effects for assessment frameworks and inspection years. The inclusion of these fixed effects restricts comparisons to schools evaluated under the same supervisory framework and at the same point in time. This accounts for changes in the Inspectorate's framework over the sample period, whereas unconditional comparisons could partly reflect differences in inspection regimes rather than genuine differences between school arrangements.

The estimated coefficients on separated schooling for each standardized dimension are shown in Figure A10. Results for the nine dimensions discussed in the main text, shown on top of the figure, are consistent with the descriptive evidence presented in Figure 7. Separated schools receive significantly higher inspection scores on differentiated instruction, support for students with special educational needs, and students' engagement with learning. They also tend to perform better in fostering students' expected social skills, although this estimate is less precisely estimated. There are no statistically significant differences in the clarity of teachers' explanations or in language adaptation. The latter result is consistent with semi-integrated schools typically providing dedicated language classes or targeted language support for refugee children, thereby reducing differences between the two educational arrangements along this dimension. By contrast, separated schools receive significantly lower inspection scores on perceived safety, having a safety policy, and playing a functional role in the community. None of the remaining ten standardized dimensions, ordered by the absolute magnitude of the estimated coefficient, exhibits a statistically significant difference between the two types of school arrangements. Qualitative results are very similar in a robustness test that additionally includes fully integrated schools (which may include regular schools that do not serve refugee students, see Figure A11).

Overall, the analysis of inspection data shows that the only systematic differences in which separated schools outperform semi-integrated schools concern dimensions related to individualized pedagogical support. Independent inspectors consistently rate separated schools more highly in their ability to tailor instruction to students' needs, provide support for children requiring additional assistance, and actively engage students in the learning process. By contrast, there is little evidence of differences in broader aspects of the school environment such as parental involvement or organizational practices, and separated schools perform worse in terms of school safety and community engagement. Taken together, these findings suggest that the distinguishing feature of separated schools is not a generally different educational environment, but rather a stronger emphasis on adapting teaching and school support to the specific needs of children.

B.5 Quality of the School Environment

Indicators of the quality of the school environment are based on all students enrolled in schools in the Netherlands in 2010. These indicators characterize the context in which refugee students attend school. To calculate the measures, we first calculate the average of a particular indicator per school and then determine the percentile ranking within a municipality.[35] We subsequently match the school of the first enrollment of the refugee students with the school percentile ranking. If the refugee and migrant schools do not have a separate location identifier from their mother school, the merge is de facto at the level of the joined regular-refugee/migrant school environment.

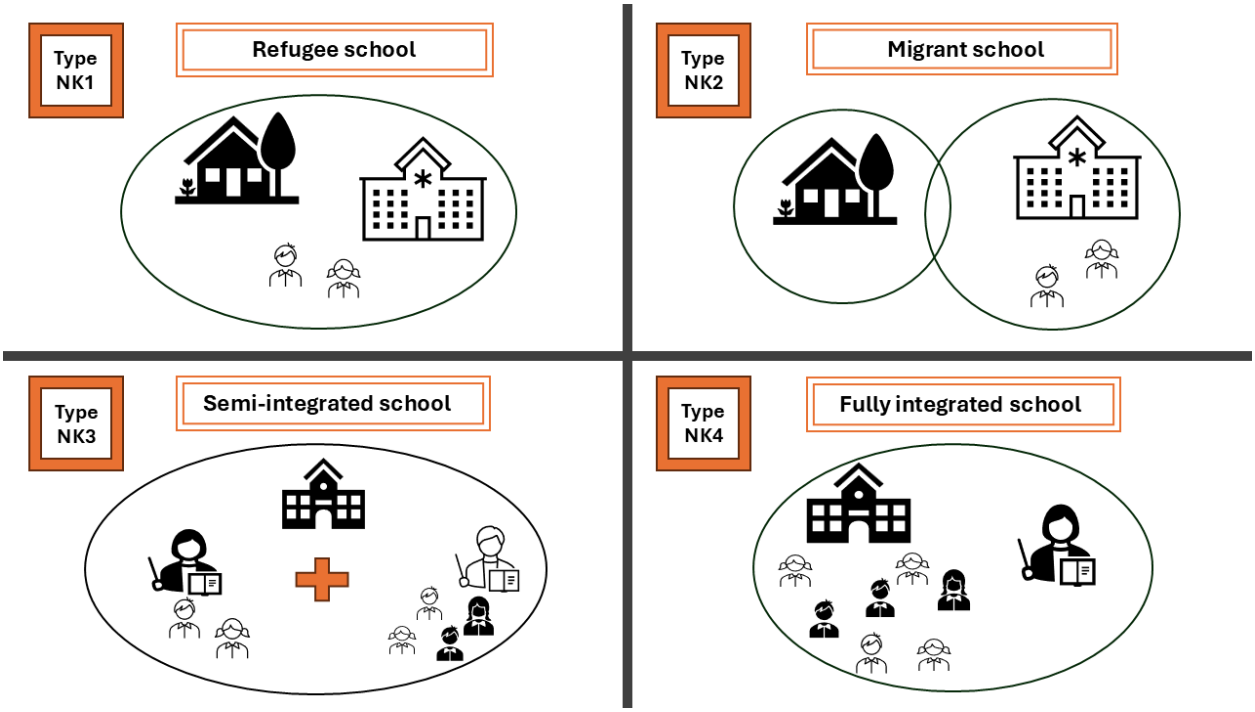
We calculate the percentile rank within the municipality for four measures of the school environment: the average funding weight of the students at the schools, the share of higher-educated parents, the share of non-native parents, and the share of students who ever end up going to university. The funding weight of the students is retrieved from the enrollment records (*onderwijsinschrtab*). These funding weights are part of the school funding formula that determines the funding of Dutch schools (Keizer (2006)). The weights are intended to reflect the chance of learning deficits and are based on students' background characteristics (Mulder and Meijnen (2013)). A higher weight indicates a greater likelihood of learning deficits.

The share of higher-educated parents reflects the percentage of students at each school who have at least one parent with a higher-education degree. We link parents and children through the child-parent link (*kindoudertab*). The share of students with non-native parents refers to the share of first- and second-generation migrants based on the birthplace of students' parents. We retrieve this information from the municipality records (*gbapersoonstab*). If parents are born outside of the Netherlands, we code the students as having non-native parents. Finally, we calculate the share of students at a school who eventually (by 2023) attend university.

[35] The percentile rank is calculated by $pr=(r-1)/(N-1)$, where r is the school's rank and N is the number of schools in the municipality. This ensures that percentile ranks range from 0 for the lowest school to 1 for the highest school.

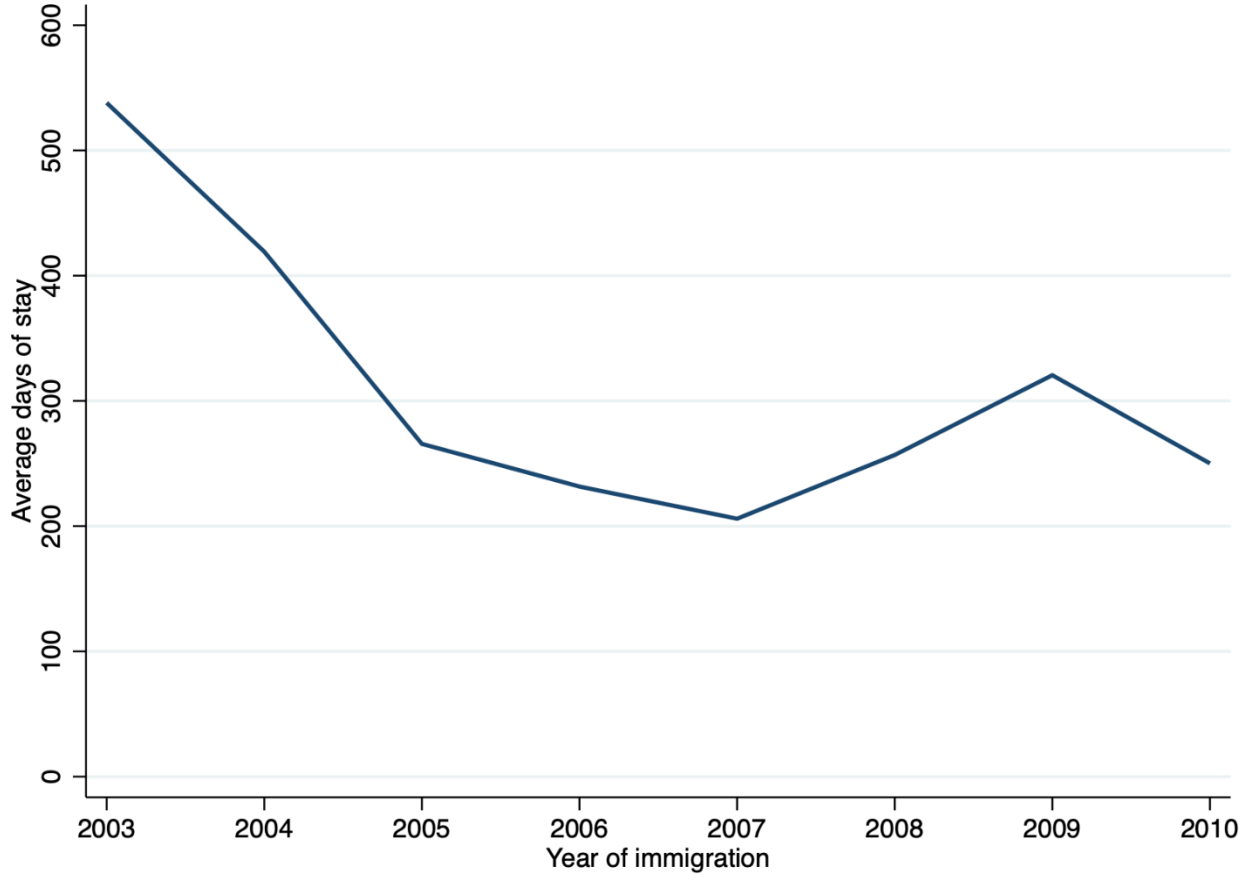
Appendix C: Additional Figures and Tables

Figure A1: Educational arrangements for refugee students upon arrival



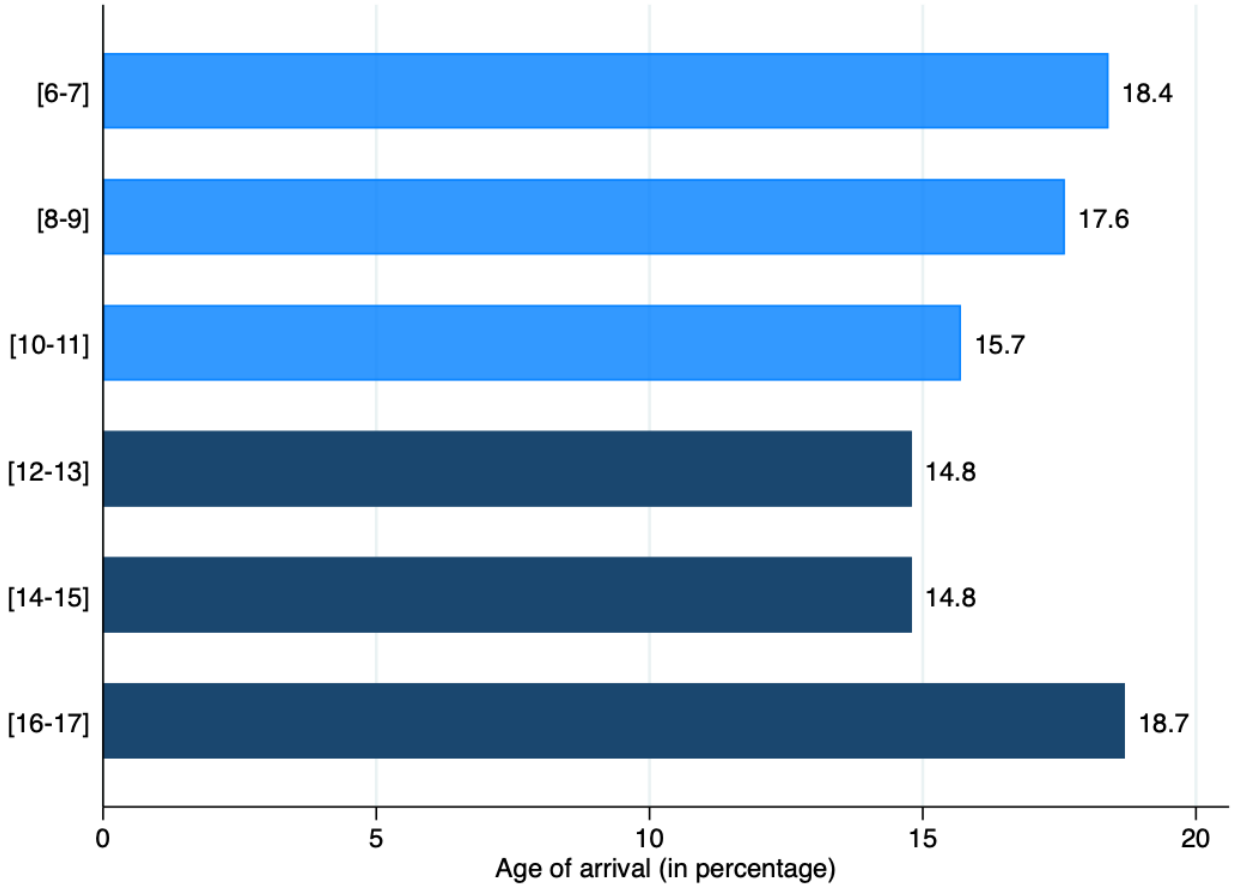
Notes: The figure exemplifies the content of the four types of educational arrangements. See section 2.2 and Appendix A for details.

Figure A2: Average length of stay in refugee centers by immigration year



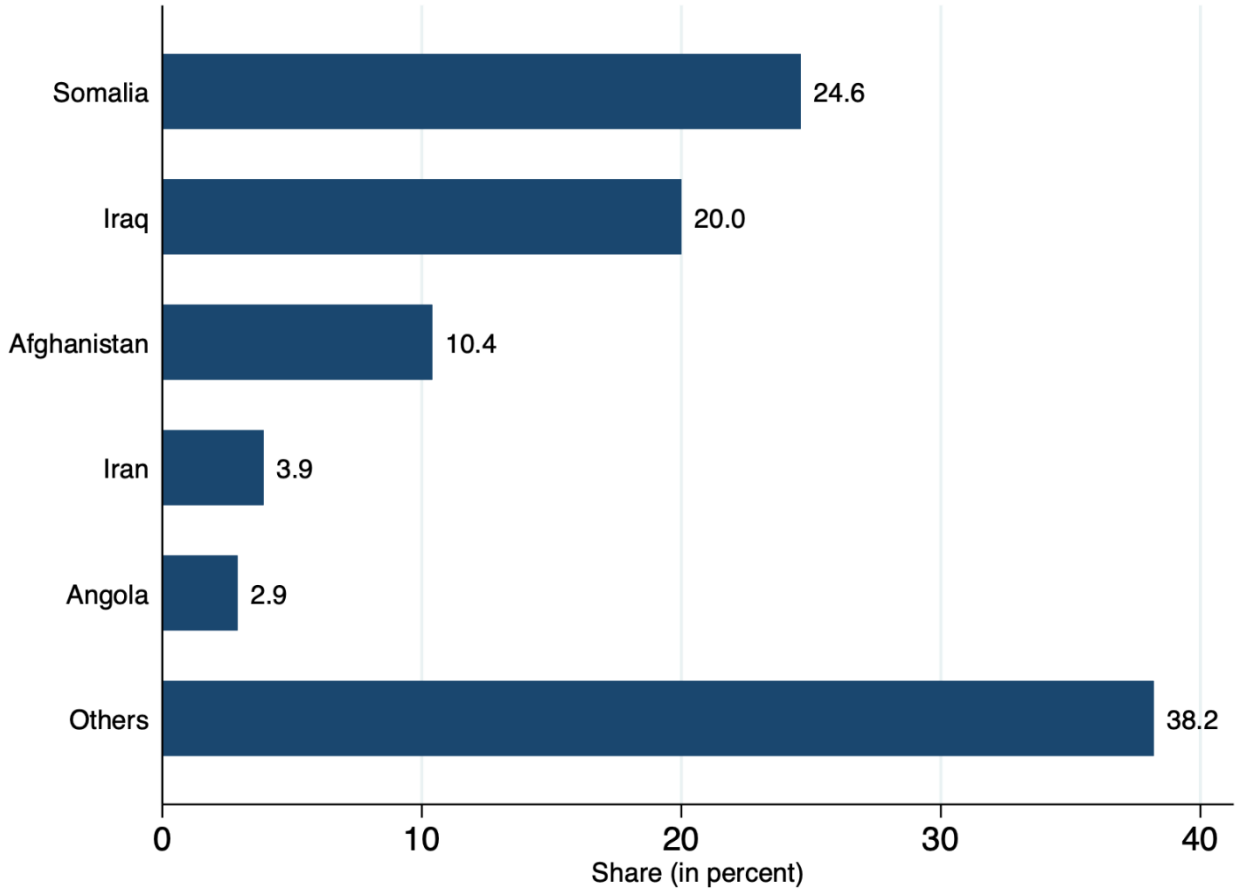
Notes: The figure reports the average number of days spent by refugee students in our sample in the refugee centers they are assigned to upon arrival for each immigration year.

Figure A3: Share of refugee students by age at arrival



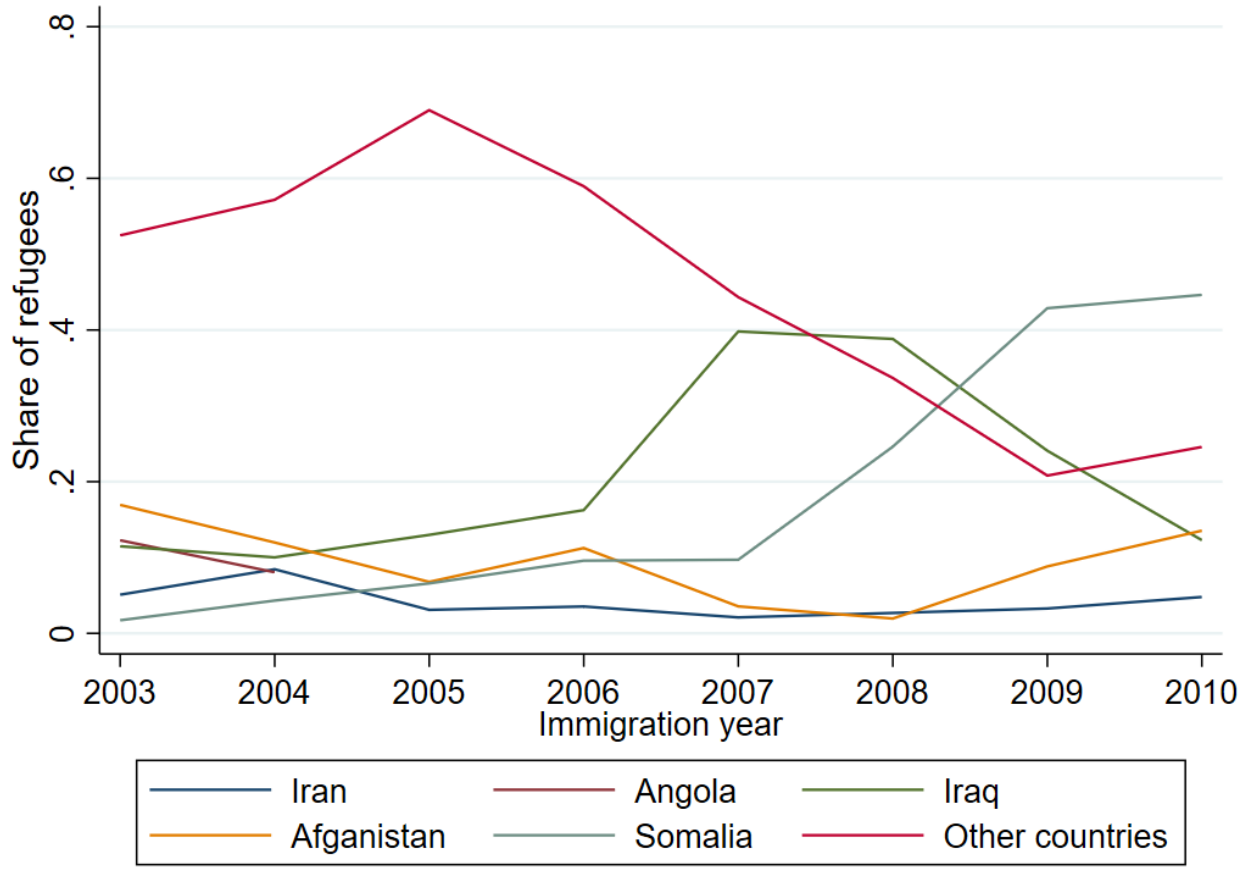
Notes: The figure displays the distribution of refugee students in our sample by age at arrival in the Netherlands between 2003 and 2010. Light bars indicate children who arrived at primary-school age, dark bars indicate children who arrived at secondary-school age.

Figure A4: Share of refugee students by country of origin



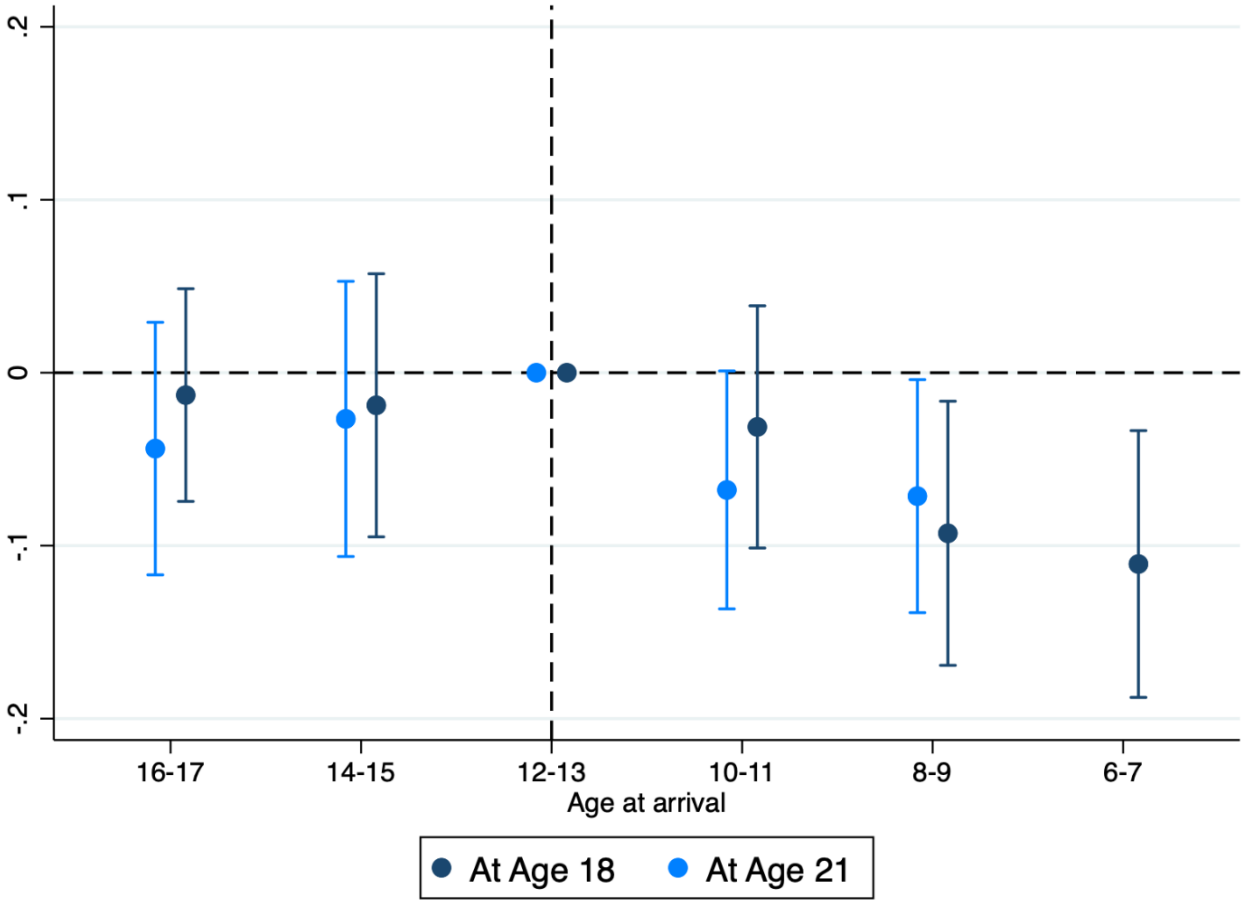
Notes: The figure shows the share of refugee students in our sample (who arrived in the Netherlands between 2003 and 2010) by country of origin for each of the five main countries of origin.

Figure A5: Share of refugee students over time for main countries of origin



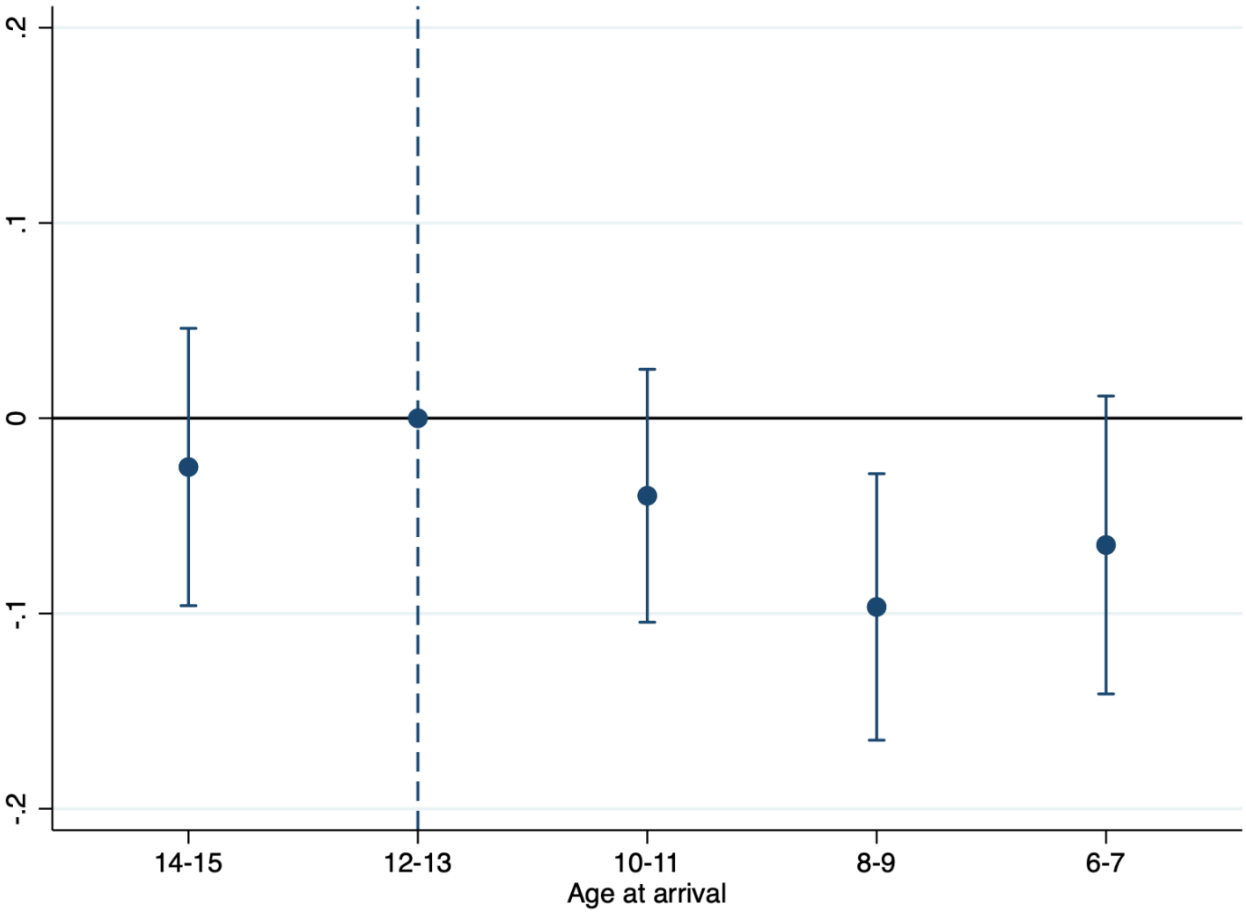
Notes: The figure shows the annual shares of refugee students in our sample (who arrived in the Netherlands between 2003 and 2010) for each of the five main countries of origin.

Figure A6: The effect of integrated schooling on obtaining a high-school diploma by ages 18 and 21



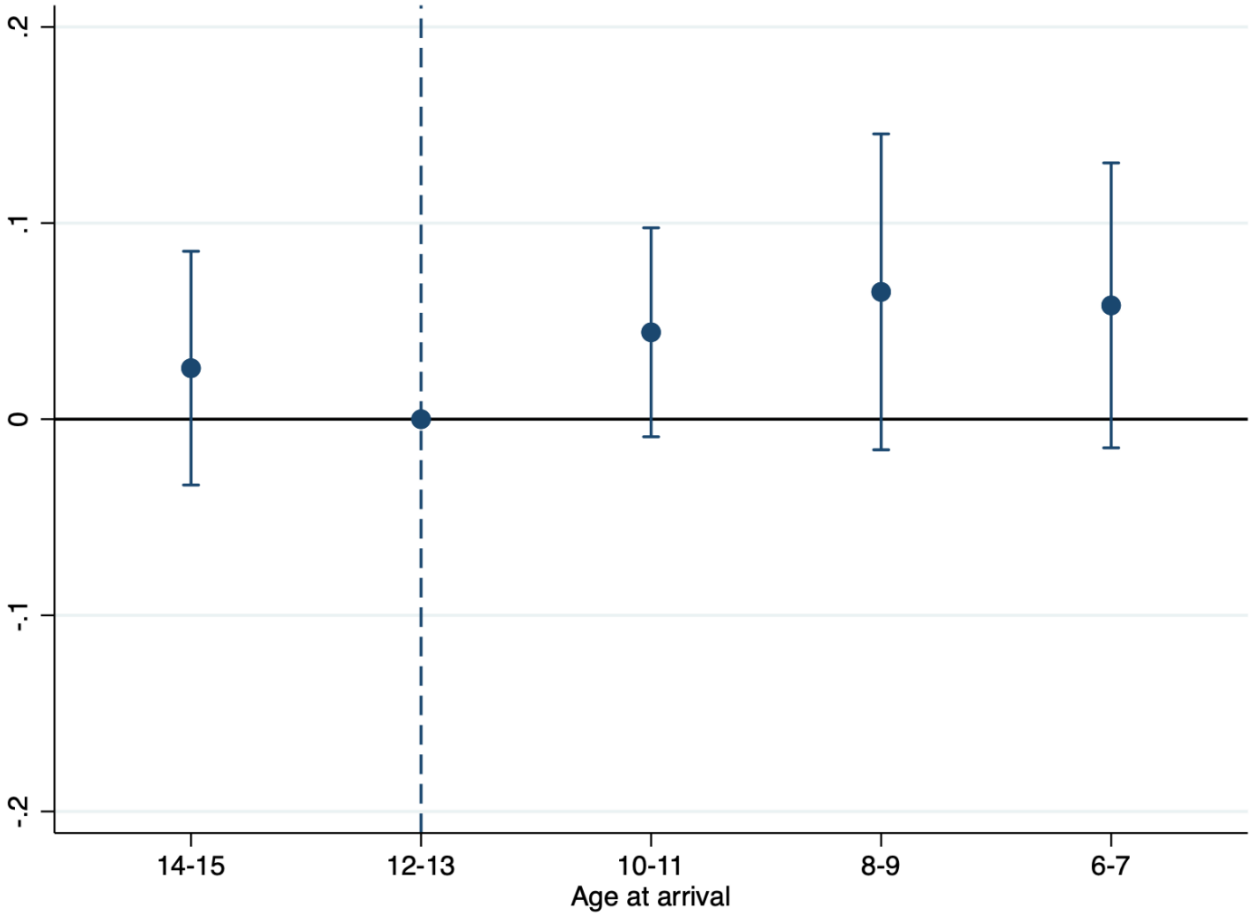
Notes: This figure shows event-study estimates of the effect of integrated schooling by age at arrival (in two-year bins) on the probability of obtaining a high-school diploma by ages 18 and 21. The econometric specification is a two-way fixed-effects model that includes age-at-arrival fixed effects and municipality by immigration-year fixed effects. All effects are computed relative to ages 12-13. The estimation sample includes refugee students who arrived between 2003 and 2010. Lines indicate 95 percent confidence intervals based on standard errors adjusted for clustering at the municipality level.

Figure A7: The effect of integrated schooling on the probability of taking the secondary-school exit exam



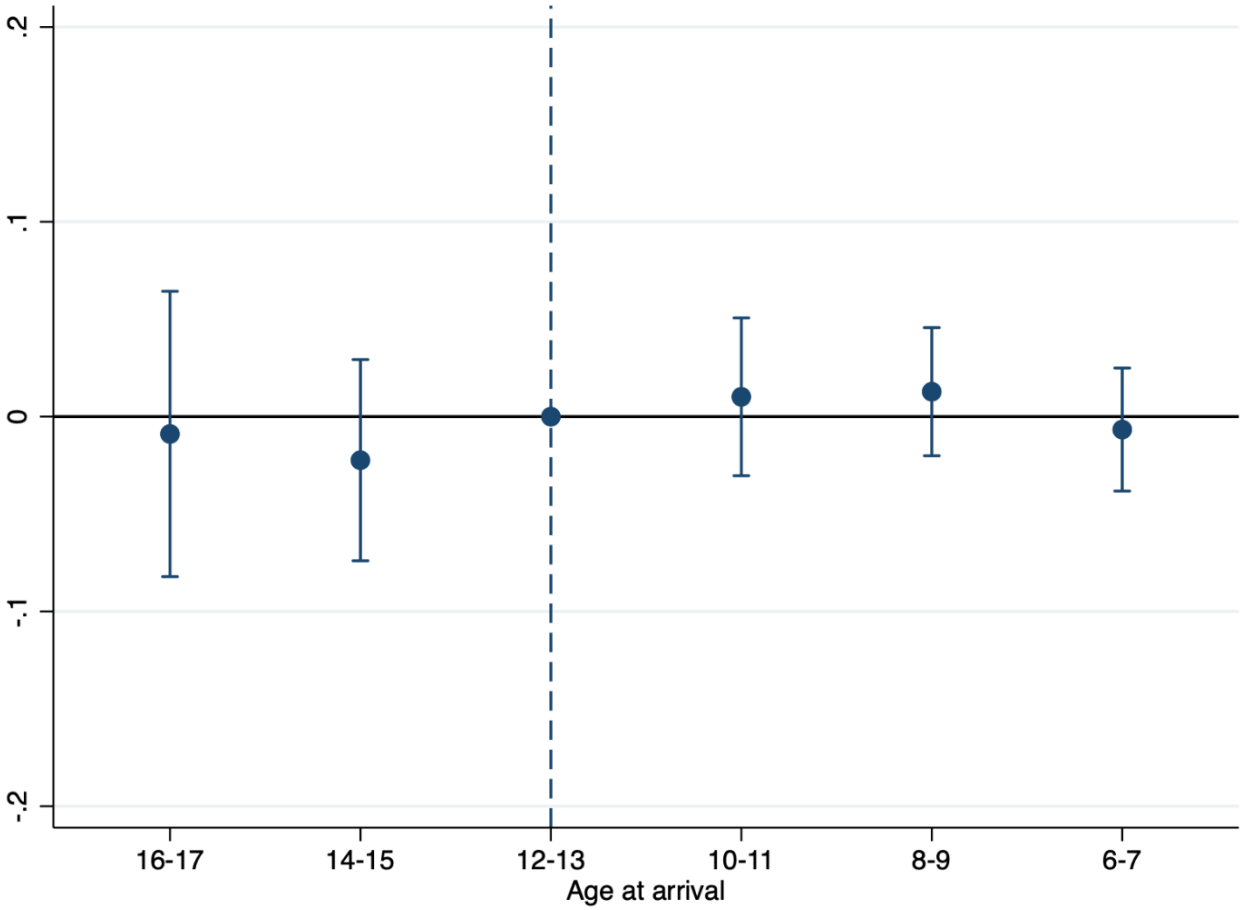
Notes: This figure shows event-study estimates of the effect of integrated schooling by age at arrival (in two-year bins) on the probability of taking the exit exam by age 18. The econometric specification is a two-way fixed-effects model that includes age-at-arrival fixed effects and municipality by immigration-year fixed effects. All effects are computed relative to ages 12-13. The estimation sample includes refugee students who arrived between 2003 and 2010. Lines indicate 95 percent confidence intervals based on standard errors adjusted for clustering at the municipality level.

Figure A8: The effect of integrated schooling on using hormonal contraceptives



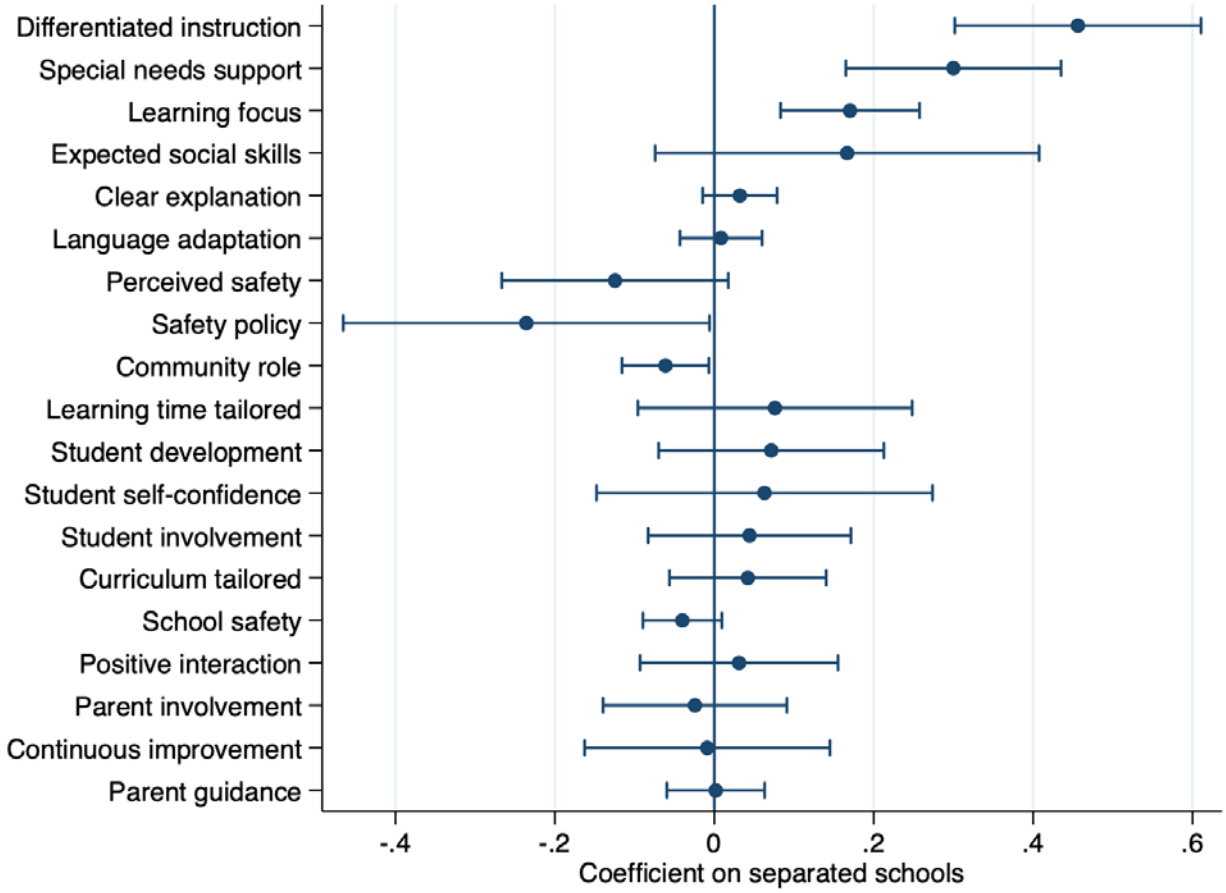
Notes: This figure shows event-study estimates of the effect of integrated schooling by age at arrival (in two-year bins) on females' probability of using hormonal contraceptives in at least two years by age 18. The econometric specification is a two-way fixed-effects model that includes age-at-arrival fixed effects and municipality by immigration-year fixed effects. All effects are computed relative to ages 12-13. The estimation sample includes female refugee students who arrived between 2003 and 2010. Lines indicate 95 percent confidence intervals based on standard errors adjusted for clustering at the municipality level.

Figure A9: The effect of integrated schooling on the probability of having a child by age 18



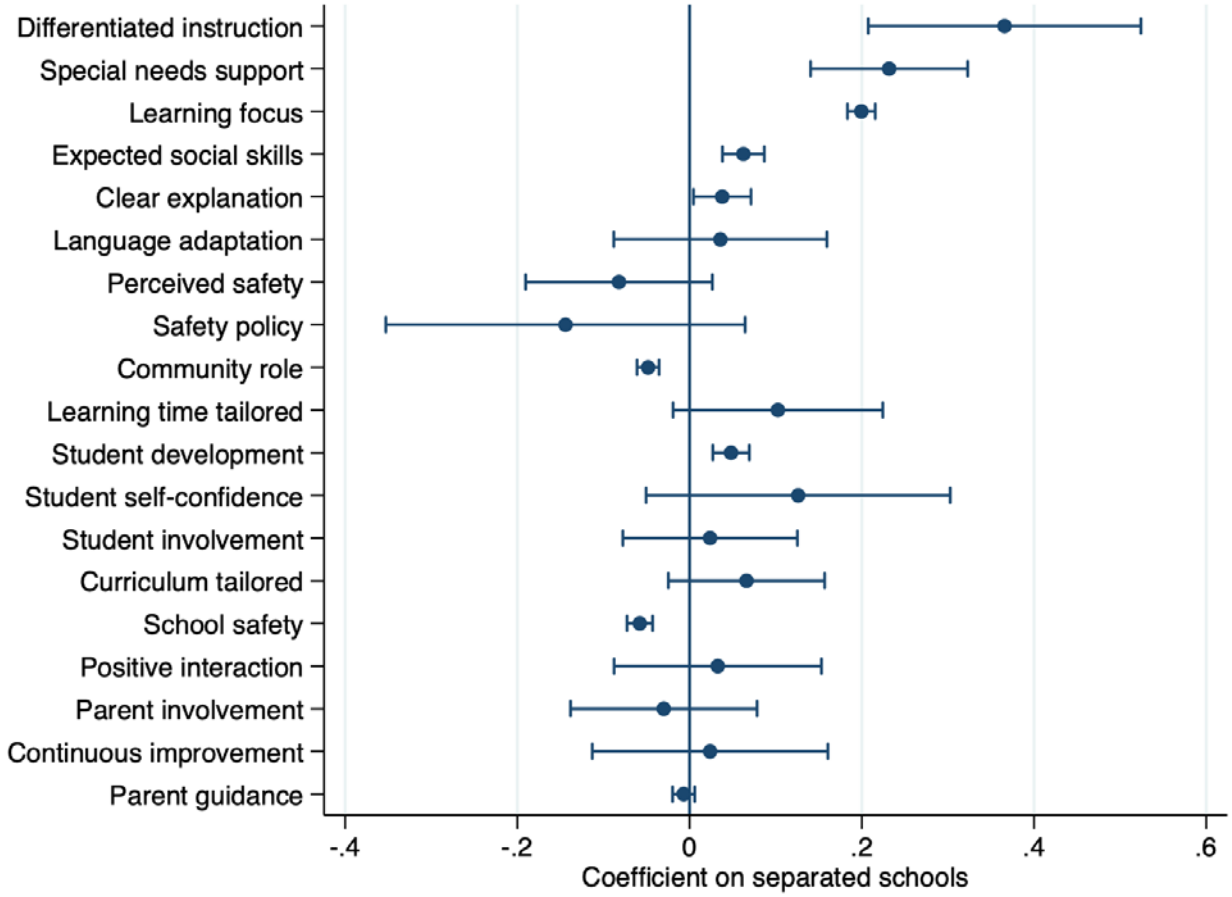
Notes: This figure shows event-study estimates of the effect of integrated schooling by age at arrival (in two-year bins) on the probability of having a child by age 18. Births are recorded by matching parents to their children in the data environment of Statistics Netherlands. Parents are required to declare births at the municipality, which records this information in the municipal population register (*Gemeentelijke Basisadministratie*) which is then made available by Statistics Netherlands. The econometric specification is a two-way fixed-effects model that includes age-at-arrival fixed effects and municipality by immigration-year fixed effects. All effects are computed relative to ages 12-13. The estimation sample includes refugee students who arrived between 2003 and 2010. Lines indicate 95 percent confidence intervals based on standard errors adjusted for clustering at the municipality level.

Figure A10: Differences in pedagogical assessments in inspection records between separated and integrated primary schools



Notes: This figure shows OLS estimates of the coefficient on separated schools (compared to semi-integrated schools) from separate regressions for each of the indicated assessments of pedagogical and classroom practices by the Inspectorate of Education, controlling for fixed effects for assessment frameworks and inspection years. Lines indicate 95 percent confidence intervals based on standard errors adjusted for clustering at the school level.

Figure A11: Differences in pedagogical assessments including fully integrated schools



Notes: This figure shows OLS estimates of the coefficient on separated schools (compared to integrated schools) from separate regressions for each of the indicated assessments of pedagogical and classroom practices by the Inspectorate of Education, controlling for fixed effects for assessment frameworks and inspection years. Lines indicate 95 percent confidence intervals based on standard errors adjusted for clustering at the school level.

Table A1: Countries of origin with a Muslim majority

Afghanistan	Iraq	Saudi Arabia
Albania	Ivory Coast	Senegal
Algeria	Jordan	Somalia
Azerbaijan	Kazakhstan	Sudan
Bangladesh	Kosovo	Syria
Bosnia and Herzegovina	Kuwait	Tajikistan
Burkina Faso	Kyrgyzstan	Tunisia
Djibouti	Lebanon	Turkey
Egypt	Libya	Turkmenistan
Gambia	Mauritania	United Arab Emirates
Gaza Strip and West Bank	Morocco	Uzbekistan
Guinea	Niger	Yemen
Iran	Pakistan	

Notes: This table lists all countries of origin with a Muslim majority, defined as those where Islam is either the official state religion or the religion of the majority of the population.

Table A2: Separate balancing tests

	Female	Higher-educated parent	Arrived with parent	Afgha-nistan	Angola	Iran	Iraq	Somalia
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Integrated schooling	-0.012	-0.043	-0.027	0.020	0.002	-0.004	0.013	0.006
	(0.025)	*(0.027)*	*(0.038)*	*(0.015)*	*(0.008)*	*(0.011)*	*(0.020)*	*(0.018)*
Age-at-arrival fixed effects (FE)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipality x immi-gration year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,801	4,881	7,801	7,801	7,801	7,801	7,801	7,801

Notes: This table shows estimation results from separate balancing tests. The respective dependent variable of each column is indicated in the column headers. Indicated countries refer to the five main countries of origin. Standard errors adjusted for clustering at the municipality level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent.

Table A3: The effect of integrated schooling on obtaining a high-school diploma: Family fixed effects

	Full donut sample	Family sample	
		Without family fixed effects	With family fixed effects
	(1)	(2)	(3)
Integrated schooling	-0.098***	-0.084**	-0.082*
	(0.027)	*(0.035)*	*(0.046)*
Age-at-arrival fixed effects (FE)	Yes	Yes	Yes
Municipality x immigration year FE	Yes	Yes	Yes
Family FE	No	No	Yes
Observations	6,612	1,949	1,949

Notes: This table presents two-way fixed-effects estimates of the effect of integrated schooling on the probability of obtaining a high-school diploma by age 18. Reported estimates refer to coefficients on the interaction between municipalities with the indicated school types and primary-school-age arrival. Column 1 replicates column 4 from Table 5. The family sample in columns 2 and 3 includes refugee students who arrived between 2003 and 2010 at ages below 11 or above 12 together with a sibling in the other (primary- or secondary-school) age type. Standard errors adjusted for clustering at the municipality level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent.

Table A4: The effect of integrated schooling on obtaining a high-school diploma: Females

	Baseline model	Preferred model	Including controls	Donut estimator	Types of integrated schooling	All underlying school types
	(1)	(2)	(3)	(4)	(5)	(6)
Integrated schooling	-0.083**	-0.085**	-0.085**	-0.136***		
	(0.036)	*(0.036)*	*(0.038)*	*(0.040)*		
Migrant school						0.021
						(0.083)
Semi-integration					-0.090	-0.087
					(0.055)	*(0.056)*
Full integration					-0.083**	-0.080*
					(0.041)	*(0.041)*
Age-at-arrival fixed effects (FE)	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Immigration year FE	Yes	Yes	Yes	Yes	Yes	Yes
Municipality x immigration year FE		Yes	Yes	Yes	Yes	Yes
Control variables			Yes			
Observations	2,963	2,963	2,963	2,391	2,963	2,963

Notes: This table presents two-way fixed-effects estimates of the effect of integrated schooling on females' probability of obtaining a high-school diploma by age 18. Reported estimates refer to coefficients on the interaction between municipalities with the indicated school types and primary-school-age arrival. The estimation sample includes females refugee students who arrived between 2003 and 2010, equivalent to the female medical-outcome sample in Tables 7 and 8. Control variables: student gender, parental education, family status, and indicators for each of the five main countries of origin (Afghanistan, Angola, Iran, Iraq, Somalia). The donut estimator in column 4 excludes students whose age at arrival is 11 or 12. The base group in column 6 is students allocated to municipalities with refugee primary schools, and in all other columns it is students allocated to municipalities with refugee or migrant primary schools. Standard errors adjusted for clustering at the municipality level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent.

Table A5: The effect of integrated schooling on using hormonal contraceptives at least once

	Baseline model	Preferred model	Including controls	Donut estimator	Types of integrated schooling	All underlying school types
	(1)	(2)	(3)	(4)	(5)	(6)
Integrated schooling	0.027	0.018	0.018	0.036		
	(0.033)	*(0.034)*	*(0.032)*	*(0.034)*		
Migrant school						-0.002
						(0.057)
Semi-integration					-0.053	-0.054
					(0.035)	*(0.035)*
Full integration					0.055	0.055
					(0.039)	*(0.040)*
Age-at-arrival fixed effects (FE)	Yes	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Immigration year FE	Yes	Yes	Yes	Yes	Yes	Yes
Municipality x immigration year FE		Yes	Yes	Yes	Yes	Yes
Control variables			Yes			
Observations	2,963	2,963	2,963	2,391	2,963	2,963

Notes: This table presents two-way fixed-effects estimates of the effect of integrated schooling on females' probability of using hormonal contraceptives at least once by age 18. Reported estimates refer to coefficients on the interaction between municipalities with the indicated school types and primary-school-age arrival. The estimation sample includes female refugee students who arrived between 2003 and 2010. Control variables: parental education, family status, and indicators for each of the five main countries of origin (Afghanistan, Angola, Iran, Iraq, Somalia). The donut estimator in column 4 excludes students whose age at arrival is 11 or 12. The base group in column 6 is students allocated to municipalities with refugee primary schools, and in all other columns it is students allocated to municipalities with refugee or migrant primary schools. Standard errors adjusted for clustering at the municipality level in parentheses. Significance level: *** 1 percent, ** 5 percent, * 10 percent.

Table A6: Anatomic therapeutic chemical (ATC) drug codes

Drug	ATC code
Hormonal contraceptives for systemic use	G03A
Antidepressants	N06A
Anti-inflammatory and antirheumatic products, non-steroids	M01A
Antihistamines for systemic use	R06A
Corticosteroids, plain (dermatological preparations)	D07A
Beta-lactam antibacterials, penicillins	J01C

Notes: This table lists the ATC drug codes of the medications used in our analysis.

Table A7: Standardized questions of school inspection data

Indicator	Item in inspectorate record	Separated schools	Integrated schools
Differentiated instruction	PO_2003_01 and PO_KK_01: "Teachers ensure that the educational learning process is tailored to the educational needs of the students." PO_2005_01: "The teachers coordinate the instruction and processing to the differences in development between the students."	23	155
Special needs support	PO_2003_01 and PO_KK_01: "The school provides guidance for students with specific educational needs."	6	99
Learning focus	PO_2003_01 and PO_KK_01: "Students are actively focused on learning."	6	96
Expected social skills	PO_2003_01: "At the end of the school period, the students control social skills at a level that can be expected." PO_2005_01: "The social skills of the students are at a level that can be expected."	2	22
Clear explanation	PO_2003_01: "The explanation is clear." PO_2005_01: "The teachers clearly explain."	22	133
Language adaptation	PO_2003_01: "In their language, teachers take into account the possible extra needs of students in the field of language." PO_2005_01: "The language use of the teachers fits in with the language needs of the students."	20	73
Perceived safety	PO_2003_01: "Students experience the school as safe." PO_2003_01: "Staff members experience the school as safe." PO_2005_01: "The students and the staff feel demonstrably safe at school."	22	133
Safety policy	PO_2003_01: "The school has a safety policy." PO_2005_01: "The school has a detailed safety policy aimed at the action after incidents."	10	80
Community role	PO_2003_01 and PO_KK_01: "The school plays a functional role within the local and regional community."	6	74
Learning time tailored	PO_2003_01: "The time for education and learning is tailored to the educational needs of students." PO_2003_01: "The time for education and learning is tailored to the characteristics of the student population." PO_2003_01: "The school varies the amount of time for education and learning, depending on the educational needs of students." PO_2005_01: "The school coordinates the amount of time for learning and teaching in Dutch language and arithmetic and mathematics to the educational needs of students."	23	134
Student development	PO_2003_01 and PO_KK_01: "Students are expected to develop."	2	31

(continued on next page)

Table A7 (continued)

Indicator	Item in inspectorate record	Separated schools	Integrated schools
Student self-confidence	PO_2003_01 and PO_2005_01: "The teachers promote the self-confidence of the students."	21	79
Student involvement	PO_2003_01: "Students are involved in the educational learning process." PO_2005_01: "The students are actively involved in the educational activities."	22	133
Curriculum tailored	PO_2003_01 and PO_KK_01: "The learning material range is tailored to the educational needs of the students."	6	97
School safety	PO_KK_01 and PO_2003_01: "The school provides safety."	6	87
Positive interaction	PO_2003_01: "Staff members and students interact with each other in a positive way." PO_KK_01: "Staff and students treat each other in a positive way."	6	97
Parent involvement	PO_2003_01 and PO_KK_01: "The school stimulates the involvement of parents and carers."	5	90
Continuous improvement	PO_2003_01: "The school is demonstrably working on improvements." PO_2005_01: "The school works systematically on improvement activities."	24	155
Parent guidance	PO_2003_01 and PO_KK_01: "The school involves parents/ guardians in the care and guidance of the students."	6	92

Notes: This table reports the standardized dimensions of pedagogical and classroom practices used in the analysis together with the corresponding framework-specific inspection indicators from which they are constructed. Indicator wording is translated from the official Dutch descriptions provided by the Inspectorate of Education. The final two columns report the number of unique schools contributing to each dimension by educational arrangement. Integrated schools refer to semi-integrated schools only.