Working Paper No. 201

Does Ethnic Diversity in Schools Affect Occupational Choices?

Damiano Pregaldini, Simone Balestra and Uschi Backes-Gellner

Universität Zürich
IBW – Institut für Betriebswirtschaftslehre

UNIVERSITÄT BERN

Working Paper No. 201

Does Ethnic Diversity in Schools Affect Occupational Choices?

Damiano Pregaldini, Simone Balestra and Uschi Backes-Gellner

December 2022

Please cite as:
"Does Ethnic Diversity in Schools Affect Occupational Choices?" Swiss Leading House "Economics of Education" Working Paper No. 201, 2022. By Damiano Pregaldini, Simone Balestra and Uschi Backes-Gellner.

The Swiss Leading House on Economics of Education, Firm Behavior and Training Policies is a Research Program of the Swiss State Secretariat for Education, Research, and Innovation (SERI).

www.economics-of-education.ch

Does Ethnic Diversity in Schools Affect Occupational Choices?*

Damiano Pregaldini† Simone Balestra‡ Uschi Backes-Gellner§

December 5, 2022

Abstract

We study how two distinct dimensions of peer ethnic diversity (ethnic fractionalization and ethnic polarization) affect occupational choice. Using longitudinal administrative data and leveraging variation in ethnic composition across cohorts within schools, we find evidence for two opposing effects. Ethnic fractionalization increases the likelihood of students sorting into people-oriented occupations while ethnic polarization reduces this likelihood. Using data on social and cognitive skills, we provide evidence that exposure to higher levels of ethnic fractionalization enhances the students' formation of social skills and increases the likelihood of students sorting into people-oriented occupations where the returns to these skills are higher.

JEL Classification: H75, I21, J18, J24

Keywords: ethnic diversity, fractionalization, polarization, school, occupational choice

*This study is funded by the Swiss State Secretariat for Education, Research, and Innovation (SERI) through its "Leading House VPET-ECON: A Research Center on the Economics of Education, Firm Behavior and Training Policies." We are grateful to Eric Bettinger, Aline Bütikofer, Thomas Dohmen, Beatrix Eugster, David Figlio, Roland Hodler, Dinand Webbink, and Stefan Wolter for their constructive comments. The paper benefited from valuable feedback at the following meetings: Causal Inference in Personality Research, Economics of Education Association, and Swiss Society of Economics and Statistics, and departmental seminars at the universities of Bamberg, St. Gallen, and Zurich. For the data, we thank the Swiss Federal Statistical Office (BFS) and the Stellwerk test service provider

†University of Zurich, Department of Business Administration, Plattenstrasse 14, CH-8032 Zurich. Email: damiano.pregaldini@business.uzh.ch.

‡University of St. Gallen and CESifo. Email: simone.balestra@gmail.com

§University of Zurich and IZA. Email: backes-gellner@business.uzh.ch.

I. Introduction

Globalization, international migration, and inclusive policies for better integration of under-represented minorities have contributed over the past 30 to 40 years to increasingly ethnically diverse school environments in most industrialized countries. For example, between 2000 and 2017, the fraction of Hispanic and Asian/Pacific Islander students in American public schools nearly doubled, from 20 to 38 percent. At the same time, the fraction of white students decreased from 61 to 48 percent, while the fraction of African Americans remained roughly unchanged (NCES, 2020). The National Center for Education and Statistics projects this trend to continue over the next decade, raising the question of the implications of school ethnic diversity for both educational production and student post-compulsory schooling outcomes.

Thus far, the economic literature on the effects of ethnic diversity has largely focused on short-term school performance such as test scores,[1] leaving evidence missing on the effects of ethnic diversity on a broader set of outcomes, particularly post-compulsory schooling outcomes. This lack of evidence is surprising, given that ethnic diversity in schools can potentially affect the development of skills (e.g., social skills) that test scores do not fully capture but that still have substantial implications for long-run outcomes (Jackson, 2018).

Most studies on peer ethnic composition use single measures of ethnic diversity. These measures include the fraction of immigrant peers (e.g., Ballatore, Fort, and Ichino, 2018; Brunello and Rocco, 2013; Figlio and Özek, 2017; Figlio et al., 2021; Jensen and Rasmussen, 2011; Ohinata and van Ours, 2013), the fraction of non-English speakers (e.g., Diette and Oyelere, 2017; Geay, McNally, and Telhaj, 2013), and ethnic fractionalization Indices(e.g., Bredtmann, Otten, and Vonnahme, 2021; Chevalier, Isphording, and Lisauskaite, 2020; Hall and Leeson, 2010). However, recent studies from the political economy literature have shown that not only ethnic fractionalization but also other dimensions of diversity can affect inter-group interactions and, in turn, individual outcomes and decisions. For example, Bazzi et al. (2019) find that while ethnic fractionalization increases integration in communities in Indonesia, ethnic polarization decreases it. Despite these findings, evidence on how different dimensions of diversity share individual student outcomes in educational settings remains missing.

This paper analyzes how peer ethnic diversity in compulsory schooling affects a major post-compulsory schooling outcome: occupational choice. Specifically, we focus on the choice of people-oriented occupations, that is occupations that have a clear component of providing help to individuals (i.e., occupations in the social and healthcare sectors) (Stoet and Geary, 2022) and that current demographic trends (e.g., aging populations) are making increasingly important.

Peer ethnic diversity can have an effect on the formation of social and cognitive skills in school. In so doing, peer ethnic diversity can affect occupational choice by changing the comparative advantage of students in different occupations. For example, peer ethnic diversity could

[1] Few studies analyze outcomes other than test scores. For example, Gould, Lavy, and Paserman (2009) analyze the probability of completing high school, and Anelli, Shih, and Williams (2017) analyze the probability of graduating with a STEM (science, technology, engineering, and mathematics) major.

enhance the formation of social skills in school, such as prosociality and empathy. In turn, the improved social skills could lead students to sort into people-oriented occupations, in which the returns to these skills are higher. Indeed, our data on occupational skill requirements shows that people-oriented occupations require above-average social skills compared to all other occupations.

However, peer ethnic diversity could also reduce the efficient provision of classroom teaching (e.g., Ballatore, Fort, and Ichino, 2018), thereby impairing the formation of cognitive skills, such as math and language skills. The reduced cognitive skills could lead students to avoid occupations in which the returns to these skills are high. Indeed, previous research shows that individuals tend to sort into occupations that more greatly reward their skills (Borghans, Weel, and Weinberg, 2014; Deming, 2017).

In our analysis, we differentiate between two dimensions of ethnic diversity that research has shown to shape intergroup relationships: ethnic fractionalization (Alesina and La Ferrara, 2005; Easterly and Levine, 1997) and ethnic polarization (Esteban and Ray, 1994; Montalvo and Reynal-Querol, 2005; Reynal-Querol, 2002). Considering both dimensions is crucial, because ethnic fractionalization and ethnic polarization may have opposing effects on individual outcomes and decisions (e.g., Bazzi et al., 2019). While these measures are widely applied to measuring diversity between countries and between regions, we are the first to use these well-established measures in the context of education.

Fractionalization, which captures the fragmentation of the peer group in ethnic subgroups, measures the probability that two randomly selected students belong to different ethnic groups. The higher this probability, the higher the chance of students interacting with students from other ethnic groups, thereby increasing the frequency of interethnic student interactions. Given that these interactions might enhance student development of social skills, we expect that ethnic fractionalization increases the likelihood of students sorting into people-oriented occupations.

In contrast, polarization captures the distribution of ethnic subgroups within the peer group. Polarization reaches its maximum when two equally sized ethnic groups exist within the peer group (Alesina et al., 2003) and decreases as one ethnic group becomes dominant in relative terms. Polarization is likely to affect occupational choice by affecting the frequency of the interactions between students belonging to different ethnic groups. For example, when a peer group has one dominant ethnic subgroup (low polarization), interactions between members of other minority groups and members of this majority group can be more frequent (Lazear, 1999), thereby increasing the frequency of interethnic interactions. In contrast, when the peer group contains a few large groups, interethnic antagonism might prevail, and interethnic interactions might become less frequent (Esteban and Ray, 1994). Therefore, we expect polarization to reduce the frequency of interethnic student interactions (and thus the development of social skills), in turn reducing the likelihood of students sorting into people-oriented occupations.

We investigate these hypotheses by using newly released longitudinal data on the universe of students in Switzerland and their occupational choices at the transition from compulsory

schooling to post-compulsory education, when students are 14 to 16 years old. Switzerland is an ideal setting for our analysis for two reasons. First, Switzerland is among the top five OECD ("Organisation for Economic Co-operation and Development") countries with both the largest fraction of foreign-born residents and the most ethnically diverse population (OECD, 2020a,b). This high level of ethnic diversity generates variation in school ethnic composition that we can exploit for identification. Second, the majority (two-thirds) of each student cohort in Switzerland enters a vocational education and training (VET) program in one of roughly 220 training occupations. These occupations are classified according to the International Standard Classification of Education (ISCED) and range from STEM-oriented occupations (e.g., electronics technician) to people-oriented occupations (e.g., social worker). Both the large numbers of students entering VET after compulsory schooling and the broad range of occupations provide substantial variation in the occupational choices we use in our analysis.

In the empirical analysis, we measure fractionalization and polarization (the two treatment variables) in the last two years of compulsory schooling (grades eight and nine) and the choice of a people-oriented occupation (the main outcome variable) one year later. To construct ethnic fractionalization and polarization Indices, we use the language spoken at home as a proxy for ethnicity, as this characteristic is one of the most critical components of ethnic identity (Stokes, 2017). To classify people-oriented occupations, we use the ISCED field of study of the VET occupation into which students sorted after compulsory schooling. Specifically, we consider all occupations in the field "health and welfare" as people-oriented.

To causally identify the effect of ethnic diversity on occupational choices, we exploit the natural variation in the ethnic composition of cohorts within schools. This strategy rests on the assumption that the variation in cohorts' ethnic composition is unrelated to other factors affecting occupational choices (similar approaches have been used in the peer effects literature, e.g., Angrist and Lang, 2004; Bifulco, Fletcher, and Ross, 2011; Brenøe and Zölitz, 2020; Carrell, Hoekstra, and Kuka, 2018). Testing this assumption with a number of balancing tests, we find that the observed within-school across-cohort variation in ethnic composition is consistent with variation generated from a random process.

Our results show that, holding polarization constant, ethnic fractionalization during compulsory school increases the likelihood of students choosing a people-oriented occupation. This result is consistent with our hypothesis that having many small ethnic subgroups in the peer group (high fractionalization) increases the frequency of interethnic student interactions, enhancing the development of social skills that are comparatively more rewarded in people-oriented occupations. Therefore, given the comparative advantage students have in these occupations, students are more likely to sort into them after compulsory schooling.

In contrast, we find that, holding fractionalization constant, ethnic polarization in grade nine decreases the likelihood of students choosing a people-oriented occupation. This result is consistent with Lazear's (1999) theory of cultural assimilation, suggesting that interethnic student interactions are less frequent when few ethnic subgroups dominate (high polarization).

In turn, fewer interethnic student interactions reduce the development of social skills, thereby leading students to develop a comparative disadvantage in people-oriented occupations.

When analyzing the effect of fractionalization and polarization on the choice of all other possible occupations, we find that the increased enrollment in people-oriented occupations comes at the cost of a reduced enrollment in business-related occupations but not in STEM. To uncover the labor market implications of these effects, we use data from the Swiss Labor Force Survey (SLFS) to analyze differences in the average wages and employment across different occupations. Although wages in Swiss people-oriented occupations are relatively high, they are up to 13% lower than wages in Swiss business-oriented occupations. However, we find that unemployment in business-oriented occupations is almost twice as high (9%) than in people-oriented occupations (5%). These insights, together with the results of our main analysis, suggest that, on one hand, ethnic fractionalization decreases the likelihood of sorting into high-paying jobs while, on the other hand, it decreases the likelihood of sorting into occupations with high unemployment risk.

Using additional rich data on students' social and cognitive skills collected in two large cantons in Switzerland we can investigate the mechanisms that may underly the observed effects. In these additional analysis we find that While ethnic fractionalization appears to enhance the student formation of social skills (particularly sociability), it has no effect on math skills and reduces the formation of language skills. These results, combined with the finding that people-oriented occupations require above-average social skills, suggest that students exposed to higher levels of ethnic fractionalization in school develop a comparative advantage in people-oriented occupations, making these students more likely to sort into these occupations.

Our paper contributes to the literature analyzing occupational choice (Pavan, 2011) and, specifically, the choice of people-oriented occupations. For example, Borghans, Weel, and Weinberg (2014) analyzes how workers sort into "people jobs" according to their skills and Kuhn and Wolter (2022) analyzes the differential sorting of male and female adolescents into people- and things-oriented occupations. Our analysis contributes to these studies by showing the peer ethnic composition in school is an additional factor determining the sorting of young adolescents into people-oriented occupations. Given current demographic trends (e.g., aging populations), labor demand in people-oriented occupations will likely increase over the next few decades. A better understanding of why and how young adolescents sort into these occupations is critical for designing policies that effectively ensure an adequate supply of qualified workers.

Our paper also adds to the peer effects literature (e.g., Brenøe and Zölitz, 2020) in general and to the literature on peer ethnic composition in particular. While previous research has mainly analyzed the effect of peer ethnic composition on short-run student performance (Ballatore, Fort, and Ichino, 2018; Figlio et al., 2021; Hanushek, Kain, and Rivkin, 2009), only a few studies have analyzed post-schooling outcomes (e.g. Chevalier, Isphording, and Lisauskaite, 2020; Chuard et al., 2022). Our paper analyzes a novel post-compulsory schooling outcome—occupational choices—and shows that peer ethnic composition in school can affect

these choices. A better understanding of how the school environment affects the occupational sorting of young adolescents' in their early careers is crucial not only because these choices have important implications for earnings, employment, and future career, but also because occupational sorting is a major driver of wage disparities across different groups (e.g., Hirsch and Macpherson, 2004; Pan, 2015; Plug, Webbink, and Martin, 2014; Sloane, Hurst, and Black, 2021).

Moreover, we contribute to the literature that analyzes the combined effect of different dimensions of ethnic diversity. For example, Bazzi et al. (2019) analyze the effect of ethnic fractionalization and polarization on nation building in Indonesia, finding that while ethnic fractionalization enhances nation building, ethnic polarization dampens this effect. Another example is Ager and Brückner (2013), who analyze the effect of cultural fractionalization and polarization on economic growth in the United States. Overall, these studies show that when analyzing diversity it is crucial to account for both fractionalization and polarization, and that these two measures of diversity can have opposing effects. While the political economy literature has long recognized the importance of analyzing the combined effect of fractionalization and polarization, this approach remains largely unexplored in educational settings. Our results show that even in educational settings, considering both fractionalization and polarization is crucial and that these two dimensions of diversity can lead to opposing effects on occupational choices.

Finally, we contribute to a better understanding of the formation of social skills in school, skills that are becoming increasingly important in the labor market (Deming, 2017; Kiener, Gnehm, and Backes-Gellner, 2020). While previous studies have analyzed the link between ethnic diversity in school and cognitive skills, finding, in some cases, negative effects (e.g. Ballatore, Fort, and Ichino, 2018; Geay, McNally, and Telhaj, 2013), we are the first to investigate the link between ethnic diversity and social skills and to show that ethnic fractionalization can enhance the formation of social skills.

II. Institutional background

The education system in Switzerland has a federal structure and gives the 26 cantons—regional administrative entities similar to U.S. states—some autonomy in educational policy decision-making. The degree of coordination among the cantons is relatively high. Moreover, since the Intercantonal Agreement on the Harmonization of Compulsory Education in the 1970s, the cantons have applied roughly the same common curriculum for compulsory schooling. This curriculum includes a two-year entry level (kindergarten) and nine years of compulsory schooling,[2] the first six years of which are at the primary level and the last three at the lower secondary

[2]In this paper, "grades" refers to the nine years of compulsory schooling and "school year" to the twelve years consisting of both kindergarten and compulsory schooling.

level.[3]

Almost all students in Switzerland complete compulsory education in public school; only 5% of students in a cohort go to private school. As inclusion is an important public school policy objective, children of different gender, ethnic, and socio-economic backgrounds are educated in an inclusive setting whenever possible. Importantly, students cannot freely choose their public school but are instead assigned to schools according to the ZIP code of their municipality of residence. This type of assignment reduces concerns of the potential sorting of students across schools.

After finishing compulsory schooling, students enter the upper secondary level by choosing among three types of education: VET, specialized professional schools, or Gymnasium.[4] Around two-thirds of each student cohort choose VET in one of the roughly 220 training occupations. VET combines part-time formal education with workplace training and experience. Students in a VET program study at a school for 1 to 1.5 days per week, and for 3.5 to 4 days a week they work as apprentices in the host companies with which they have an employment contract for their entire two- to four-year training period (Wolter and Ryan, 2011).

For students entering VET, the choice of a training occupation occurs between grades eight and nine, when students are 14-16 years old. This choice, while based primarily on the students' preferences, is also based on their academic performance in secondary school. The choice of a training occupation clearly has implications for future earnings, career trajectories, and occupational mobility. Students and their families therefore give this choice very careful consideration.

Given that the last two years of compulsory schooling are decisive for students' occupational choice, we measure the treatments of interest in these two years. Specifically, we calculate ethnic fractionalization and polarization separately in grades eight and nine, and use these variables to explain students' occupational choices after compulsory schooling. In our main analysis, we measure the treatment in grade nine. As we show in section VI., measuring the treatment in grade eight has no impact on the finding, because school cohorts remain largely unchanged during secondary school.[5]

[3] Switzerland has four language regions (German, French, Italian, and Romansh). Within each region, the compulsory school curricula are virtually identical, with some minor differences. For example, in a few cantons, primary school goes from grade one through five, while the lower secondary level goes from grade six through nine. Nonetheless, the total number of years of compulsory education is the same for all cantons.

[4] "Specialized professional schools" provide general and professional education in specific occupational fields. "Gymnasiums" are academic high schools, which grant direct access to university education. Approximately one-third of each student cohort chooses one or the other.

[5] In Switzerland, student mobility across schools is limited. In the subsample of students choosing VET after compulsory schooling, more than 85% of students stay in the same school during the last three years of compulsory schooling, and more than 92% remain at an institution within the same municipality. Therefore, student selection across schools in response to changes in school diversity is unlikely to affect our results. Studies on single cantons report even lower mobility rates among young adults (Balestra, Sallin, and Wolter, forthcoming).

III. Student data and measures of ethnic diversity

To analyze the effect of school ethnic diversity during compulsory schooling on the subsequent occupational choice, we draw on register data on the universe of students in Switzerland. This data is generated by the Swiss Federal Statistical Office's LABB ("Längsschnittanalysen im Bildungsbereich") program through linking different educational and labor registers and are therefore well-suited for analyzing the effect of school ethnic diversity on occupational choices.

This comprehensive dataset contains the universe of students completing compulsory schooling between 2011 and 2016.[6] For each student, we observe the grade, cohort, and a school identifier (the data contain both public and private schools). We use this information to identify peer groups at the school-cohort level.[7] Therefore, peers are students in the same school and cohort. The average size of these groups is 61 students.

Within peer groups, we measure student ethnicity with the student's primary language, i.e., the one they most frequently speak at home.[8] This approach is common in the literature on ethnic diversity (e.g., Alesina et al., 2003; Desmet, Ortuño-Ortín, and Weber, 2009). Indeed, language is one of the most critical components of national identity (Stokes, 2017) and is therefore correlated with ethnicity.[9] In total, our data contain 29 languages (see Table A.1 in the Appendix).

To measure ethnic diversity within peer groups, we use two Indices that are well-established in the literature: the fractionalization index (Alesina and La Ferrara, 2005; Easterly and Levine, 1997) and the polarization index (Esteban and Ray, 1994; Montalvo and Reynal-Querol, 2005; Reynal-Querol, 2002). Specifically, we construct the ethnic fractionalization index according to formula 1 and the ethnic polarization index according to formula 2.[10]

[6] Data are available from 2011, when linking different registers through the social security number became possible. The last cohort for which we observe both the last year of compulsory schooling and the first year of VET is 2016.

[7] Students in grade nine are typically divided into tracks according to their performance. Some schools divide students into tracks for only a few subjects (e.g., math and foreign languages). In these schools, students spend most of the instructional time in classes with other students from different tracks, so that they are likely to interact with students both in the same cohort and track and in the same cohort but different tracks. For these schools, we define the peer group at the school-cohort level. Other schools divide students into tracks for all subjects, so that students interact mainly with students in the same track and cohort. For these schools, we define the peer group at the school-track-cohort level. Around 20% of the schools have both mixed-track and single-track classes. For these schools, we define the peer group at the school-cohort level if more than 80% of the school classes are mixed-track. If fewer than 80% of the classes are single-track, we define the peer group at the school-cohort level.

[8] Each student in the data reports one primary language. If students have two or more primary languages and one is the language of instruction, the students report the language of instruction. If students have two or more primary languages and the language of instruction is not among them, the students report the language they used most frequently during early childhood.

[9] Clearly, there are exceptions in which language does not optimally capture ethnicity. For example, in the U.S., both Non-Hispanic White and African Americans speak English.

[10] We construct both Indices separately for grades eight and nine, using all students in the data, that is, the universe of students in grade eight or nine in a given year.

$$F_{ist} = 1 - \sum_{g=1}^{G} (s_g^{-i})^2 \tag{1}$$

$$P_{ist} = 1 - \sum_{g=1}^{G} (\frac{0.5 - s_g^{-i}}{0.5})^2 * s_g^{-i} \tag{2}$$

Where i indicates the student, s the school, and t the cohort. The subscript g indicates one of the G different languages in school s and cohort t. $s_g^{-i} = \frac{N_g - 1}{N-1}$ is the leave-out fraction of students with primary language g in school s and cohort c, i.e., the fraction of students with primary language g in school c and cohort t excluding student i. We use the leave-out fraction, rather than the full group fraction, to separate the student's own contribution to ethnic fractionalization and polarization from the peers' contribution (Angrist, 2014).

The ethnic fractionalization index equals the probability that two randomly selected (with replacement) students from the same school cohort speak a different (primary) language. Therefore, a higher value of the ethnic fractionalization index indicates a higher degree of ethnic heterogeneity of the peer group. The index varies between zero (all students in the school cohort speak the same primary language) and $\frac{G-1}{G}$ (all students in the school cohort speak a different primary language), where G is the number of language groups. In contrast, the ethnic polarization index indicates how far the distribution of language subgroups in the school cohort is from the bipolar distribution (Montalvo and Reynal-Querol, 2005; Reynal-Querol, 2002). The index is maximized when two groups are of the same size and decreases as one group becomes dominant in relative terms.

Our outcome of interest is occupational choice, which we measure one year after completion of compulsory school. Therefore, while we measure school ethnic diversity (the treatment) in grade nine (the last year of compulsory schooling), we measure occupational choice (the outcome) one year later (the first year of post-compulsory education). Figure A.6 in the Appendix summarizes at which stage of the educational system we measure the treatment and the outcome.

To construct the outcome variable, we consider all students who choose VET (and immediately start a VET program) after compulsory schooling as these students have to choose an occupation for their training. For students choosing VET, we observe the field of the occupation they choose. We classify these occupational choices according to the International Standard Classification of Education (ISCED) Field of Education and Training. Specifically, we define all ISCED "health and welfare" occupations as people-oriented.[11]

[11]The other ISCED fields of study that we observe are art and humanities; social sciences, journalism, and information; business, administration, and law; communication technologies; engineering, manufacturing, and construction; agriculture, forestry, fisheries, and veterinary; and services.

Our data contain the universe of students in Switzerland. To construct our estimation sample, we restrict the initial sample to the population of students entering a VET program right after grade nine and for whom we can observe the ISCED field of the occupation they choose (44% of all students completing compulsory schooling), so that we can classify their occupational choice. We exclude individuals taking a gap year after grade nine, because we cannot observe the level of ethnic diversity of their peer group during that year. Additionally, we exclude individuals in small school cohorts (fewer than ten students, because this cohort size is highly unusual) and schools that appear only once in the data (i.e., with no within-school variation in the treatment). Moreover, we restrict the sample to students aged between 14 and 17 in their last year of compulsory schooling, as this is the typical age at which students starting a VET program make their occupational choice. These additional restrictions further reduce the sample by 5.8%. The final estimation sample consists of 238,630 students. Table A.1 in the Appendix reports descriptive statistics for the estimation sample.

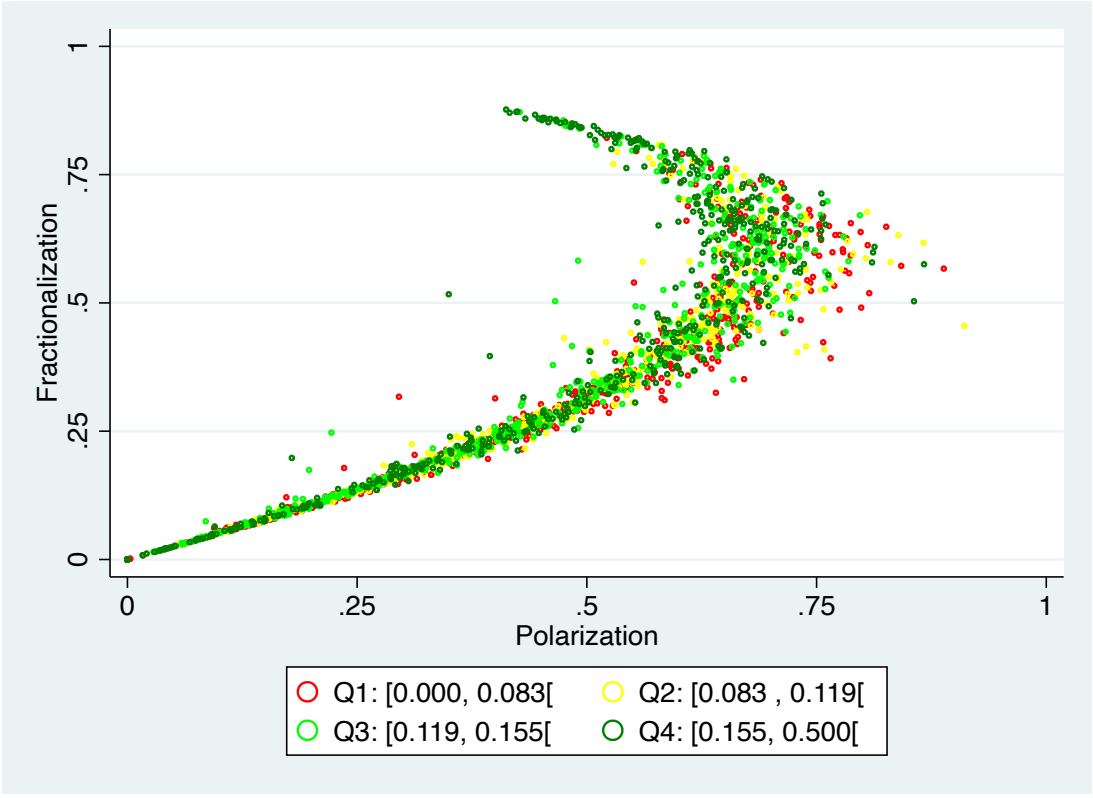


Figure 1: Ethnic fractionalization, polarization, and choice of a people-oriented occupation. *Note:* Data are collapsed at the school level: each dot is a school. Ethnic fractionalization and polarization are school averages. Each color indicates one quartile of the school-level fraction of students sorting into people-oriented occupations. The table's legend reports the quartiles boundaries.

Figure 1 shows the relationship of the two Indices with the outcome variable, that is, the choice of people-oriented occupations. For each school (one dot), we calculate the average levels of ethnic fractionalization and polarization across all cohorts. Moreover, we calculate the school-level fraction of students sorting into people-oriented occupations and rank schools

according to the quartiles of this variable. Dark green dots represent schools with fractions of students sorting into people-oriented occupations above the 75th percentile, while red dots represent schools with fractions of students sorting into people-oriented occupations below the 25th percentile.

Figure 1 is similar to the figures in Montalvo and Reynal-Querol (2005) and Bazzi et al. (2019). For low levels of fractionalization (< 0.25) and polarization (< 0.4), the two Indices are almost collinear, implying that separately identifying the effects of fractionalization and polarization at low levels of the Indices is difficult. In the mid-range, the correlation is roughly zero, whereas for high levels of fractionalization (> 0.6), the correlation is negative.

The figure shows that, for a given level of ethnic polarization, the density of schools with high fractions of students sorting into people-oriented occupations (dark green dots) increases with ethnic fractionalization. Conversely, for a given level of ethnic fractionalization, the density of schools with low fractions of students sorting into people-oriented occupations increases with ethnic polarization. Overall, Figure 1 provides descriptive evidence that the school-level fraction of student sorting into people-oriented occupations is positively associated with the level of ethnic fractionalization (keeping polarization constant) and negatively associated with the level of ethnic polarization (keeping fractionalization constant).

IV. Empirical strategy

1. Identification strategy

For identification, we follow an approach, first proposed by Hoxby (2000), that is now widely applied in the peer effects literature (e.g., Angrist and Lang, 2004; Bifulco, Fletcher, and Ross, 2011; Black, Devereaux, and Salvanes, 2013; Brenøe and Zölitz, 2020; Carrell and Hoekstra, 2010; Carrell, Hoekstra, and Kuka, 2018; Figlio and Özek, 2017; Gould, Lavy, and Paserman, 2009; Hanushek, Kain, and Rivkin, 2009; Lavy, Paserman, and Schlosser, 2012; Lavy and Schlosser, 2011). To overcome the potentially endogenous composition of the peer group, we leverage the variation in peer composition across cohorts within schools. In our setting, we exploit variation in the ethnic fractionalization and polarization Indices across cohorts within the same schools. Our approach consists of estimating the following model:

$$o_{ist} = \alpha + \beta F_{-ist} + \gamma P_{-ist} + \delta I_{ist} + \lambda C_{-ist} + \phi_s + \rho_t + \varepsilon_{ist} \tag{3}$$

Where o_{ist} is the binary indicator for student i in school s and year t choosing a people-oriented occupation after compulsory schooling. This indicator takes value one if the student chooses a people-oriented occupation and zero if he or she chooses any other occupation. F_{-ist} and P_{-ist} are the ethnic fractionalization and polarization Indices measured in grade nine at the school-

cohort level. For comparability, we standardize the Indices to have zero mean and unit variance. We also include a set of individual controls I_{-ist} (gender, age dummies, and nationality) and a set of school-cohort controls C_{-ist} (fraction of female peers, average age of the peers, and size of the school-grade cohort). Moreover, we include school fixed effects ϕ_s capturing heterogeneity at the school level and cohort fixed effects ρ_t capturing systematic differences across cohorts. In a more flexible specification, we additionally include linear and quadratic school-specific time trends to control for time-variant school-specific confounding characteristics. ε_{ist} is the idiosyncratic error term. Finally, we cluster the standard errors at the level of the treatment, the school-cohort level.

The identification of β and γ rests on the assumption that the cohort-to-cohort variation in student ethnic composition is random, conditional on school fixed effects, cohort fixed effects, and school trends. Put differently, this variation must be unrelated to differences in individual student characteristics that might simultaneously affect occupational choice. The next subsection provides empirical evidence supporting this assumption.

2. Validity of the identification strategy

To obtain precise estimates of the causal effect of diversity on occupational choices, we need sufficient variation in the ethnic fractionalization and polarization Indices within schools across cohorts. Moreover, our strategy rests on the assumption that variation in the ethnic fractionalization and polarization Indices across cohorts within the same school are as good as random, conditional on school fixed effects and cohort fixed effects. To assess the credibility of these assumptions, we run four tests.

First, we analyze how much of the variation in the two Indices remains after we remove school fixed effects, cohort fixed effects, and school trends. Table 1 reports these residual variations. After we account for school and cohort fixed effects, the standard deviation of the ethnic fractionalization index (panel A) drops by almost one third, from 0.25 to 0.09. Including school linear trends slightly reduces the standard deviation to 0.07. Panel B of Table 1 shows a similar pattern for the polarization index. These large reductions in variation indicate that our effect estimates are based on small changes in student ethnic composition. Therefore, our estimates are uninformative as to the impact of moving a single student across schools with very different ethnic compositions. Nonetheless, and most importantly, the data contains sufficient variation for us to estimate the effects of small changes in cohort ethnic composition with reasonable precision.

To visualize the full distribution of the Indices, in Figure 2 we plot the two Indices and their residuals. After we remove school and cohort fixed effects (red line) and school trends (green line), the distributions of the residuals appear normal for both Indices and are consistent with random variation in the Indices. Figures A.4 and A.5 compare the distribution of the residual variation of both Indices with the normal distribution and confirm that the two distributions are

Table 1: Variation in the ethnic fractionalization and polarization Indices

	N	Mean	SD	Min	Max
Panel A. Ethnic fractionalization					
Raw	238,630	0.371	0.243	0.000	0.915
Residuals after school FE and cohort FE	238,630	0.000	0.091	-0.588	0.561
Residuals after school FE, cohort FE, and school trends	238,630	0.000	0.075	-0.519	0.644
Panel B. Ethnic polarization index					
Raw	238,630	0.468	0.220	0.000	1.000
Residuals after school FE and cohort FE	238,630	0.000	0.109	-0.750	0.648
Residuals after school FE, cohort FE, and school trends	238,630	0.000	0.090	-0.752	0.553

Notes: Descriptive statistics of the ethnic fractionalization and polarization Indices after residualization.

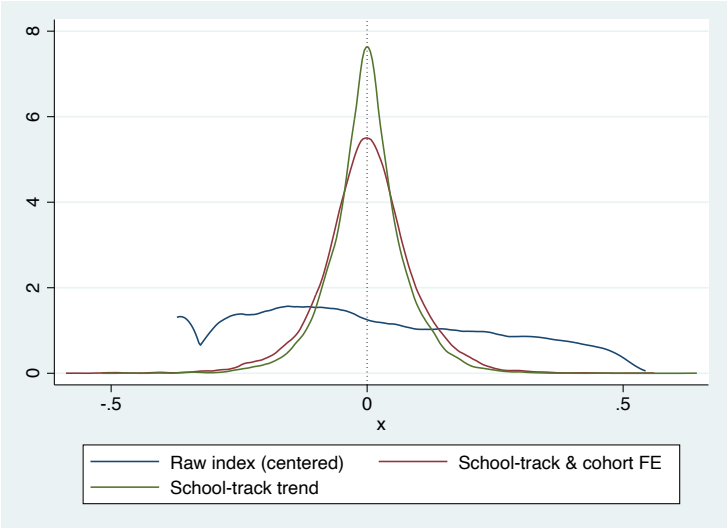
similar.

Second, we test whether the data shows students' systematic selection into schools and cohorts according to individual student background characteristics. In our case, students might select into school cohorts with higher or lower degrees of ethnic fractionalization and polarization according to individual (partly unobservable) characteristics related to occupational choice. While we cannot directly test whether students select into school cohorts according to their unobservable characteristics, we can test whether variation in observable student background characteristics within schools and across cohorts is correlated with variation in the ethnic fractionalization and polarization Indices.

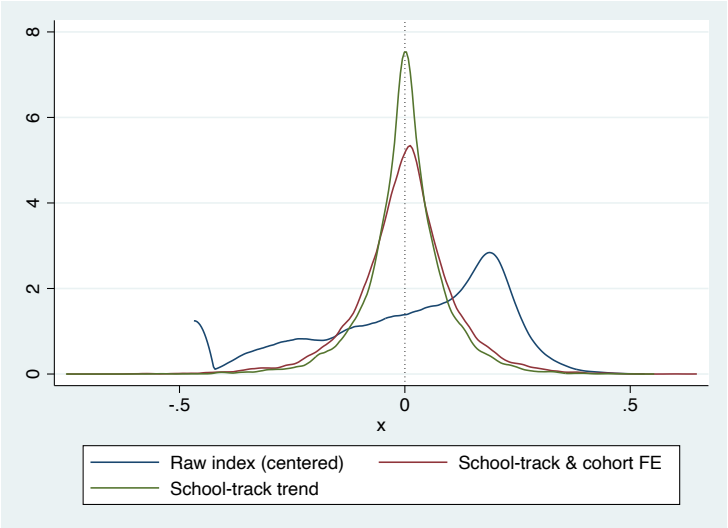
For this test, we analyze how many student characteristics are related to the two Indices. We run this test on the full sample of students and analyze 20 student background characteristics that we observe in the data: gender, four age dummies (one for each age between 14 and 17), nationality (Swiss national vs. non-Swiss national), three dummies for the type of residence area (urban, semi-rural, and rural), six dummies for parental education (no compulsory schooling, compulsory schooling, high school, VET, tertiary professional education, tertiary academic education), class size, and four dummies for the quarter of birth.

Table 2 summarizes the results. The number of characteristics for which the test cannot reject the null-hypothesis of no correlation between student background characteristics and school cohort ethnic fractionalization and polarization is close to the number that we would expect to be significant by chance, particularly in the models including controls and the school trends. Therefore, we exclude student selection according to individual observable characteristics into cohorts with different levels of ethnic diversity, conditional on school fixed effects, cohort fixed effects, and school trends.

Third, following Bifulco, Fletcher, and Ross (2011), we run simulations to analyze whether the variation in the Indices we observe in the data is consistent with random assignment of students to cohorts within schools. For this test, we simulate 250 random assignments of students to cohorts within schools. We then calculate school-cohort Indices of ethnic fractionalization



(a) Ethnic fractionalization



(b) Ethnic polarization

Figure 2: Distribution of residuals

and polarization according to these randomly assigned cohorts. Figure 3 shows that the actual and simulated distributions of the two Indices are nearly identical, supporting the assumption that variation in the Indices is as good as random.

Across 500 simulations, the average standard deviation of the residuals of the ethnic fractionalization index after we control for school and cohort fixed effects is 0.0724, as opposed to the actual 0.091 reported in Table 1. After we account for school trends, the average residual variation in the simulated ethnic fractionalization index is 0.0720, close to the residual variation of 0.075 we observe in the actual data. For the polarization index, the average standard deviation after we control for school and cohort fixed effects is 0.0896, as opposed to the actual 0.1062 reported in Table 1. After we remove school trends, the average residual variation in the simulated polarization index is 0.0891 and close to the residual variation observed in the data (0.0871). Overall, these results indicate that, after we control for school fixed effects, cohort

Table 2: Student background characteristics and Indices

	Ethnic fractionalization			Ethnic polarization			Expected
# performed tests	20	20	20	20	20	20	
# significant tests at the 1% level	1	2	1	2	2	0	0.2
# significant tests at the 5% level	5	2	2	3	3	1	1
# significant tests at the 10% level	5	3	2	5	5	2	2
Controls	No	Yes	Yes	No	Yes	Yes	
School FE	Yes	Yes	Yes	Yes	Yes	Yes	
School trend	No	No	Yes	No	No	Yes	

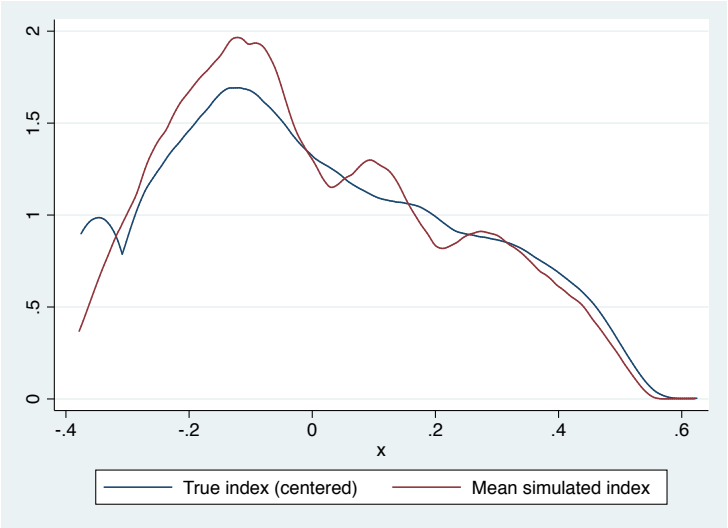
The table reports the number of student background characteristics for which the H_0 of no correlation with the ethnic fractionalization and polarization Indices could not be rejected. The column "expected" reports the number of characteristics we would expect to be significantly related to the index by chance at different significance levels. This number is the product of the significance level and the number of tests performed.

fixed effects, and school trends, variation in both Indices is consistent with random assignment of students to cohorts.

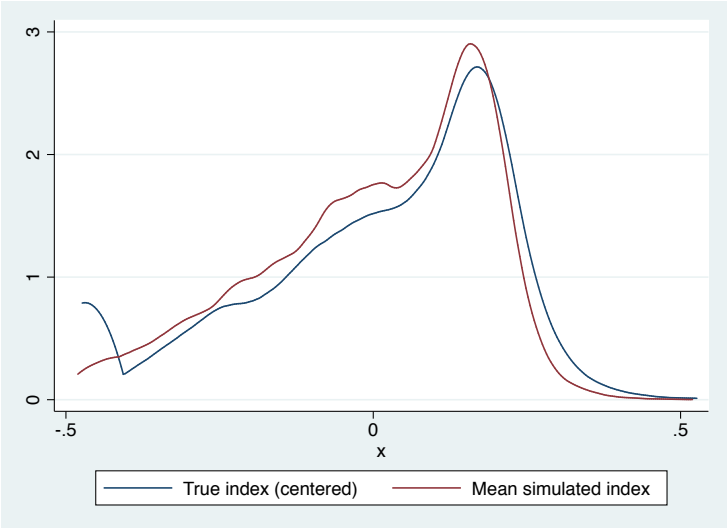
Fourth, we test the autocorrelation in the ethnic fractionalization and polarization Indices over time. Such autocorrelation might indicate the existence of time-varying unobservable confounders (Brenøe and Zölitz, 2020) and would support the inclusion of school-specific linear time trends for capturing linear autocorrelation within schools. In contrast, the absence of autocorrelation would indicate that the inclusion of school-specific trends in our regression model is unnecessary.

For this test, we run for each school and index a separate regression index in year t on the index in year $t-1$. Because this test requires a minimum of three years for each school, we can run the regression in 1,755 schools. For the ethnic fractionalization index, the percentage of schools for which we cannot reject the null hypothesis of no autocorrelation over time is 0.34 at the 1% significance level, 3.14 at the 5% significance level, and 6.52 at the 10 percent significance level. For the polarization index, the percentages are 0.86, 2.97, and 6.12, respectively. All these values are below the critical values that we would expect to observe by chance. We thus find no systematic autocorrelation of the diversity index over time. Therefore, controlling for school-specific time trends in our regression should not affect our estimates.

Despite having established that the residual variation in ethnic fractionalization and polarization is plausibly random, we still might be concerned that ethnic fractionalization and polarization might lead to (a) differential selection of students into VET compared to other upper-secondary types of education and therefore (b) systematic differences in the composition of students sorting into VET. Section II describes the three types of upper-secondary education



(a) Ethnic fractionalization



(b) Ethnic polarization

Figure 3: Distribution of actual and simulated Indices

that students can choose after compulsory schooling: VET, Gymnasium (the academic track), and specialized professional schools. Moreover, some students might opt for an "interim solution" after compulsory schooling.[12] To test whether ethnic fractionalization and polarization lead to systematic selection of students into VET, we consider the full sample of students at the end of compulsory schooling. We then analyze the effect of ethnic fractionalization and polarization on their choice of one of the three possible types of education at the upper secondary level or the choice of an interim solution.

Table A.12 in the Appendix reports the effect of ethnic fractionalization and polarization on the choice of the VET track (column1), Gymnasium (column 2), specialized schools (column 3), and interim solutions (column 4). The sample consists of students for whom we observe tracking choices after compulsory schooling. The coefficients do not show any significant effect

[12] In Switzerland, the choice postponement known as an "interim solution" usually entails an additional school year at the compulsory school level (for a detailed description, see Jaik and Wolter, 2016).

of ethnic fractionalization or polarization on the choice of any of these four options. We therefore find no evidence that ethnic fractionalization and polarization lead to systematic selection of students into VET.

Taken together, the results of the four tests support our assumption that—conditional on school fixed effects, cohort fixed effects, and school trends—the variation in the ethnic fractionalization and polarization Indices is as good as random. Moreover, the analysis of educational choices after compulsory schooling shows that our treatments do not lead to systematic changes in the composition of students selecting into VET, the type of education we consider when analyzing occupational choices in our main analysis.

V. Results

1. Main results

Table 3 reports the coefficients obtained by estimating Eq. 3, in which we subsequently add individual controls (age, gender, and nationality) (column 2); controls for cohort characteristics (fraction of female students, average age, and class size) (column 3); school-specific linear time trends (column 4); and school-specific quadratic time trends (column 5). Columns 2 and 3 show that the inclusion of control variables has little effect on the estimated coefficients, supporting the assumption that ethnic fractionalization and polarization are as good as random, conditional on school and year fixed effects (as discussed in Section 2.). Model 4, which includes school-specific linear time trends, is our preferred specification while model 5 includes school-specific quadratic time trends, allowing for a more flexible specification of the time trends. However, the coefficients from model 4 remain virtually unchanged.

The coefficients represent the effect (in percentage points) of a one standard deviation increase in each index. The estimated coefficients imply that—keeping polarization constant, that is, keeping the distribution of ethnic groups within school cohorts constant—a one standard deviation increase in ethnic fractionalization increases the likelihood of choosing a people-oriented occupation by roughly 1.12 to 1.14 percentage points. A one standard deviation increase in the ethnic fractionalization index is roughly 0.24, corresponding to a reduction of the probability that two randomly selected students from the same school cohort belong to the same ethnic group by 24 percentage points.[13] These effects suggest that going from a school cohort in the 25th percentile of the ethnic fractionalization distribution to one in the 75th percentile would increase, all else being equal, the likelihood of choosing a people-oriented occupation by roughly 2 percentage points. Given that about 12.5% of students in our sample choose a people-oriented occupation, this effect is economically significant.

[13] Consider a group of 20 students all belonging to the same ethnic group. If we were to replace three of these students with three students belonging to a different ethnic group, we would achieve an increase in the ethnic fractionalization index of 25.5 percentage points, corresponding to roughly one standard deviation in the index.

The coefficients also imply that—keeping ethnic fractionalization constant, that is, keeping the degree to which the school-cohort is split into distinct ethnic subgroups constant—a one standard deviation increase in the ethnic polarization index reduces the likelihood of choosing a people-oriented occupation by about 0.58 to 0.69 percentage points. A one standard deviation increase in the ethnic polarization index is roughly 0.22.[14] These effects suggest that going from a school cohort in the 25th percentile to one in the 75th percentile would decrease the likelihood of choosing a people-oriented occupation by roughly 1.6 percentage points. Again, given that the proportion of students choosing a people-oriented occupation is about 12.5%, this effect is economically significant.

Overall, these estimates confirm our expectation that ethnic fractionalization and polarization affects the choice of a people-oriented occupation. Students in school cohorts with many small groups (high fractionalization and low polarization) are more likely to sort into people-oriented occupations, possibly because interethnic interactions are more frequent and students develop social skills that are comparatively more rewarded in people-oriented occupations. In contrast, students in cohorts with few large ethnic groups (high polarization and low fractionalization) are less likely to sort into people-oriented occupations, possibly because ethnic polarization reduces the frequency of interethnic interactions by increasing interethnic antagonism (Esteban and Ray, 1994; Lazear, 1999).

Table 3: The effect of ethnic peer diversity on occupational choice

	(1)	(2)	(3)	(4)	(5)
Fractionalization	1.115***	1.134***	1.325***	1.119***	1.119***
	(0.300)	(0.273)	(0.294)	(0.337)	(0.338)
Polarization	-0.687***	-0.581***	-0.625***	-0.605***	-0.605***
	(0.221)	(0.203)	(0.204)	(0.231)	(0.232)
Year FE	yes	yes	yes	yes	yes
School FE	yes	yes	yes	yes	yes
Individual controls	no	yes	yes	yes	yes
Cohort controls	no	no	yes	yes	yes
School time trend	none	none	none	linear	quadratic
Obs.	238,630	238,630	238,630	238,630	238,630

The dependent variable is 1 if the occupation is people-oriented, and 0 otherwise. Ethnic fractionalization and polarization are standardized with zero mean and unit variance. The coefficients are multiplied by 100. Standard errors are in parentheses. Standard errors are clustered at the school cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[14] Consider a group of 20 students in which 17 students belong to the same ethnic group and three students belong to a different ethnic group. If we were to replace two students of the majority group with two students of the minority group, the ethnic polarization index would increase by 0.24.

2. Mechanisms and additional analyses: Skill formation

People tend to sort into those occupations in which they have a comparative advantage, that is, in which the returns to their skills are highest (Borghans, Weel, and Weinberg, 2014; Deming, 2017). Ethnic diversity can affect this comparative advantage through skill formation. On the one hand, ethnic fractionalization in school could enhance the formation of social skills by increasing the frequency of student inter-ethnic interactions. The interaction with students with a different ethnic background might make other students more prosocial and empathetic, increasing their likelihood to engage in people-oriented occupations in which these social skills are comparatively more rewarded (i.e., in which they have a comparative advantage).[15] However, empirical evidence is still scarce, and only a few studies analyze how ethnic diversity affects social skill formation (Burns, 2012; Rao, 2019).

On the other hand, ethnic fractionalization might reduce the formation of hard skills by making the efficient provision of classroom teaching more difficult (e.g., by increasing student disruption as in Lazear (1999)). In turn, the reduced provision of classroom teaching could adversely affect the students' formation of cognitive skills, such as math and reading skills, leading students to avoid occupations in which these skills are comparatively more rewarded (i.e., in which they have a comparative disadvantage). However, empirical evidence on the effects of ethnic diversity is mixed (e.g. Ballatore, Fort, and Ichino, 2018; Bredtmann, Otten, and Vonnahme, 2021; Figlio et al., 2021; Geay, McNally, and Telhaj, 2013; Jensen and Rasmussen, 2011). Therefore, it is a priori unclear how ethnic diversity might affect occupational choice through student cognitive skill formation.

In this section, we analyze how ethnic diversity affects skill formation in schools using data on cognitive and non-cognitive skills from the "Stellwerk 8" test. The Stellwerk 8 is a computer-based standardized test administered to all 8th graders by the cantonal ministries of education with the goal of providing a comparable measure of students' skills at the end of grade eight.[16] All children take the test in mandatory core subjects (e.g., math, German language, English language). This test is norm-referenced, self-scoring, and adaptive: the set of questions students face is not the same for all students. Instead, students face questions of varying difficulty depending on the number of correct answers given in the previous questions.

In addition to mandatory core subjects, a subset of students take an additional test on personal, social, and methodological competencies ("PSM test"). The PSM test comprises 162 items measuring different non-cognitive skills, from the ability and willingness to work in a team to empathy.[17]. Students rate each item on a scale from 1 (strongly disagree) to 6 (strongly agree). In the PSM test, all students face the same set of questions.

At the end of the test, students receive a certificate indicating the score achieved in the

[15] For ethnic polarization, we expect the opposite effect because ethnic polarization potentially reduces the frequency of student inter-ethnic interactions.

[16] The test is comparable to U.S. standardized tests such as the GRE.

[17] The 162 items are grouped in 14 skills: autonomy, resilience, empathy, emotional stability, flexibility and creativity, need for harmony, sociability, learning ability, consciousness, self-confidence, team play, tolerance, reliability

different subjects. Because the test provides a measure of students' competencies that is comparable across schools and regions, it has substantial signaling power. Indeed, students often have to provide their test results when applying to apprenticeship training positions. Therefore, Stellwerk 8 test scores can impact the set of occupations to which a student can and is applying for an apprenticeship training position.

For our analysis, we were able to acquire and use the Stellwerk 8 data of two large Swiss cantons. We aggregate these data at the school-cohort level, and we then match the aggregated data to our main data containing the Indices of ethnic fractionalization and polarization.[18] Our final sample comprises 504 school-by-cohort observations for the social skill data and 2.305 school-by-cohort observations for the cognitive skill data.[19]

To proxy social skills, we use the PSM test scores in items related to four skills that are related to people-oriented occupations: empathy, sociability, teamplay, and tolerance. To proxy cognitive skills, we use the test scores ranging between 0 and 1000 in math, German (i.e., the language of instruction), and English (i.e., the most common foreign language). For comparability, we standardize test scores in math, German, and English to have 0 mean and unitary standard deviation.

To analyze the effects of ethnic fractionalization and polarization on cognitive and social skill formation, we estimate the following model similar in spirit to our main model in equation 3:

$$z_{st} = \alpha + \beta F_{st} + \gamma P_{st} + \lambda C_{st} + \phi_s + \rho_t + \varepsilon_{st} \tag{4}$$

Where z_{it} is the average of skill z in school s and at time t. z is either one of the four social skills (empathy, sociability, team play, tolerance) or one of the three cognitive skills (math, German, English). F_{st} and P_{st} are the ethnic fractionalization and polarization Indices measured in grade eight (i.e., in the same grade in which students take the Stellwerk 8 test) at the school-cohort level. We additionally include a set of school controls C_{st} (fraction of female peers, average age of the peers, and size of the school-grade cohort). C_{st} also includes a linear school-specific time trend to capture school-specific confounding characteristics that linearly vary over time. Moreover, we include school fixed effects ϕ_s capturing heterogeneity at the school level and cohort fixed effects ρ_t capturing systematic differences across cohorts. ε_{st} is the idiosyncratic error term. Finally, we weigh each observation by the size of the school cohort, and we cluster the standard errors at the level of the treatment, that is, at the school-cohort level.

Table A.4 reports in columns (1) to (4) the estimated effects of ethnic fractionalization and polarization on social skills. The coefficients of the ethnic fractionalization index are all posi-

[18] For the matching, we use the municipality of the school and the size and year of the cohort. We do so, as we are not allowed to identify single students across the two data sets.

[19] While the test in core subjects is compulsory for all students of the two cantons, only a subsample of students (roughly 16% of all students in both cantons) take the PSM test. Therefore, the number of observations in the cognitive and social skill datasets is different.

tive, suggesting a positive association between ethnic fractionalization and social skills. However, only the estimated effect on sociability is statistically significant at the 10% level. For ethnic polarization, the estimated effects are all negative, but none of the coefficients is statistically significant. Overall, the results provide suggestive evidence of a positive link between ethnic fractionalization and social skill formation.

Columns (5) to (7) of Table A.4 report the estimated effects of ethnic fractionalization and polarization on cognitive skills. The coefficients associated with the ethnic fractionalization index are negative but only the estimated effects on German (the language of instruction) and English (the foreign language) are statistically significant. For math, the effect is non-significant at any conventional level. For ethnic polarization, the estimates are positive, but none of the coefficients is statistically significant. The estimates suggest that a one standard deviation increase in the ethnic fractionalization index is associated with a 0.12 standard deviation reduction in the German test scores and a 0.11 reduction in English test scores.

Given that we construct our measures of ethnic diversity using the language students speak at home, these results are easily explained. Indeed, schools with high ethnic fractionalization have many small groups of students speaking different languages. In these schools, local language proficiency varies greatly across students, and teaching languages becomes more difficult. The same does not hold for math class, in which differences in language proficiency might play less of a role. For this reason, we observe a negative effect of ethnic fractionalization on German and English but not on math.

Taken together, the results in Table A.4 provide suggestive evidence that while ethnic fractionalization in school could enhance the students' formation of social skills, it could also dampen the formation of language skills. Therefore, students exposed to higher levels of fractionalization are likely to develop a comparative advantage in occupations with high returns to social skills and low returns to language skills.

To reconcile these findings with the results of our main analysis, we consider the cognitive and social skill requirements of different occupational fields. Assuming that skill formation is driving the effects of ethnic fractionalization and polarization on the choice of people-oriented occupations, we would expect these occupations to require above-average social skills or, alternatively, below-average language skills. Indeed, Table A.2 shows that people-oriented occupations (i.e., care occupations) require above-average social skills, such as empathy and teamplay. However, these occupations also require above-average language skills. We interpret these findings as suggestive evidence that the effect of ethnic fractionalization and polarization on the choice of people-oriented occupations primarily operates through social skills formation.

3. Effect on other occupations

Given that ethnic diversity increases the likelihood of students choosing a people-oriented occupation, the question arises as to which occupations students are pulled away from. For example,

if ethnic diversity reduces the student acquisition of cognitive skills, students in more diverse schools could be less likely to choose an occupation that requires strong academic and cognitive skills, such as an occupation in STEM (e.g., Ballatore, Fort, and Ichino, 2018; Gould, Lavy, and Paserman, 2009; Hanushek, Kain, and Rivkin, 2009).

To answer this question, we estimate Eq. 3 for all ISCED fields that we observe in the data.[20] Table A.3 in the Appendix reports the estimated coefficient. In addition to the positive effect on the choice of a people-oriented occupation (column 6), we find no effect of the ethnic fractionalization index on the choice of an occupation in the field "engineering, manufacturing, and construction." Some occupations in this field are STEM occupations requiring mathematical skills, such as electronics engineering ("Elektroniker/in EFZ") and mechanical engineering ("Polymechaniker-in EFZ"). Therefore, diversity does not appear to pull students away from STEM occupations. In contrast, we find negative coefficients for choosing an occupation in the fields "business, administration, and law." These results suggest that the marginal students who move into people-oriented occupations when exposed in school to a higher degree of diversity are not those who would have otherwise chosen a people in STEM but rather those who would have chosen a career in business.

4. Wages and employment

To examine how labor market outcomes differ across occupational fields and for students who selected into people-oriented occupations in particular, we use the SLFS. We report in Table A.5 in the Appendix the 2017 median wages and the unemployment rate in these occupational fields.

Table A.3 shows that people-oriented occupations are not low-paying jobs. Indeed, full-time equivalent median wages in the people-oriented occupations are comparable to median wages in business-oriented occupations. According to Table A.3, these business-oriented occupations are those into which students sort less frequently when exposed to higher levels of ethnic fractionalization and lower levels of ethnic polarization. However, the raw wages of all workers and the wages of workers employed full-time in people-oriented occupations are around 12% below the wages in business-oriented occupations.

In contrast, the unemployment rate in business-oriented occupations is almost twice as high (9%) as the unemployment rate in people-oriented occupations (5%). Therefore, while students sorting into people-oriented occupations face lower wages, they also have a lower risk of unemployment.

[20] In the subsample of students choosing VET after compulsory schooling, we observe seven ISCED fields in total. For simplicity, we aggregate ISCED fields "arts and humanities" and "social sciences, journalism, and information" into one category.

5. Heterogeneity

i. Heterogeneity according to student characteristics

School ethnic diversity might have different effects on the occupational choices of students with different background characteristics, such as gender and nationality. Moreover, the effect of ethnic diversity could also differ according to whether students are in the ethnic minority or majority group within the peer group. To analyze effect heterogeneity, we estimate Eq. 3 in subsamples of male, female, Swiss, non-Swiss, minority, and majority students. We assign students to the majority group if their first language is the local language (e.g., German in the German-speaking regions, French in the French-speaking regions) and otherwise to the minority group. The estimated effects of fractionalization and polarization in the different subsamples appear in Figure 4.

Consistent with our main findings in Table 3, the estimates of the effect of fractionalization are positive, and the estimates of the effect of polarization are negative in all subsamples. However, both coefficients for the subsample of male students and the coefficient of the polarization index for the subsample of non-Swiss students are marginally significant. Figure 4 shows that the effect is different for male and female students: both ethnic fractionalization and polarization have stronger effects on the choices of female students.[21] For the subsamples of non-Swiss and minority students, Figure 4 shows little evidence of effect heterogeneity.

ii. Nonlinear effects

To analyze whether the effect of the ethnic fractionalization and polarization Indices is nonlinear, we construct a cubic spline with three knots. Figure A.1 plots the predicted outcome (Panel A) and estimated marginal effect (Panel B) of ethnic fractionalization. Figure A.2 plots the same results for ethnic polarization. The estimated effect of ethnic fractionalization appears nonlinear: the positive effect of the ethnic fractionalization index on the probability of choosing a people-oriented occupation decreases as the index increases. The effect is no longer statistically significant for high values of the index (index larger than 0.075), suggesting that an additional increase in ethnic fractionalization in school cohorts that already have high levels of ethnic fractionalization does not further increase the likelihood of sorting into people-oriented occupations. For the estimated effect of ethnic polarization, we find little evidence of nonlinearities.

6. Other outcomes

In Table A.6 we analyze the effects of ethnic fractionalization and polarization on other outcomes. In column 1 we analyze the choice of a people-oriented occupation. However, unlike

[21] The small number of male students choosing people-oriented occupations might partly explain why the coefficients are non-significant in this subsample. Indeed, only 8% of the students choosing people-oriented occupations are male.

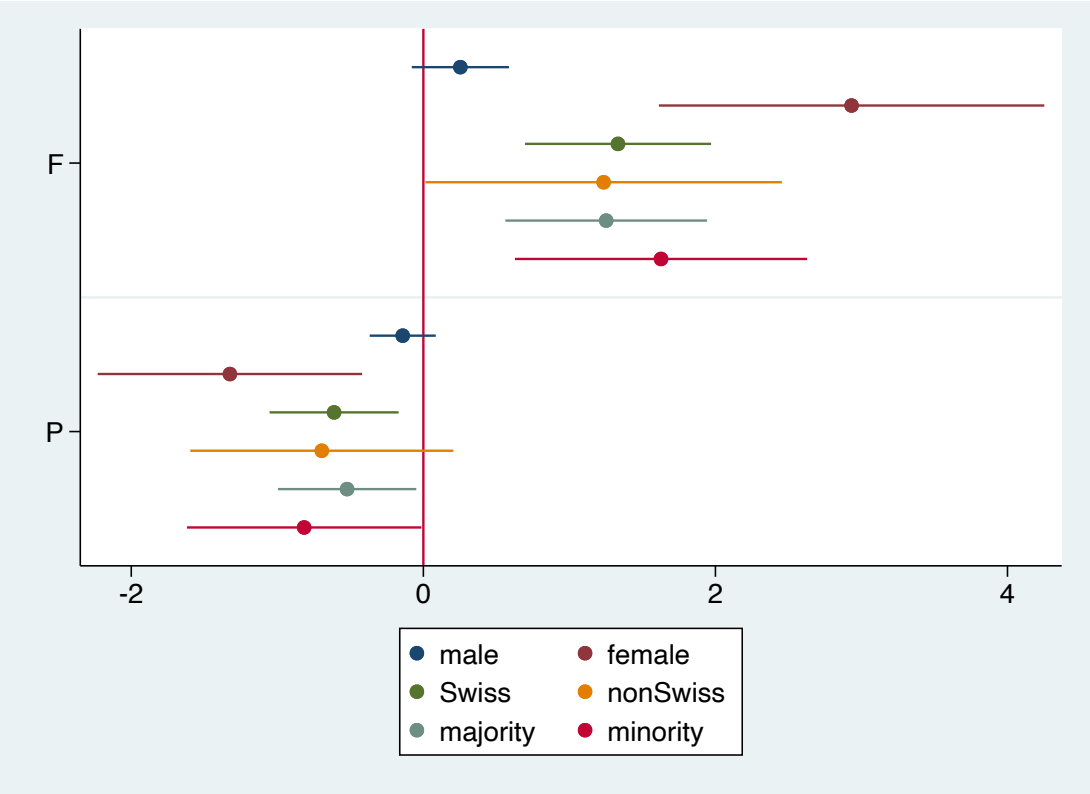


Figure 4: Estimates of the effect of fractionalization and polarization in different subsamples of female, male, Swiss, and non-Swiss students.

with our main specification, we consider the occupation two years after completion of compulsory schooling. By doing so, we analyze how ethnic fractionalization and polarization affect the likelihood not only of choosing a people-oriented occupation but also of staying in it up to two years after completing compulsory schooling. Consistent with our main analysis, we again find that ethnic fractionalization increases the probability of being in a people-oriented occupation two years after completion of compulsory schooling, while ethnic polarization decreases this probability.

Moreover, in column 2 we examine the probability of leaving the panel after compulsory schooling and therefore not having any outcome at the upper-secondary level. Students leave the panel if they do not start a formal education at the upper-secondary level, mostly because they enter the labor market right after compulsory schooling.[22] While we find no effect of ethnic fractionalization on the probability of leaving the panel, we find that ethnic polarization decreases the probability of leaving the panel.

Finally, in columns 3 and 4, we consider the probability (a) of leaving the country (out-migration) and (b) of retention (repeating the last year of compulsory schooling). None of the coefficients are significant, suggesting that ethnic fractionalization and polarization do not affect these probabilities.

[22] In a few cases, the student might start an education that is not recognized as formal or might start an education abroad. In such special cases, that education would not appear in our data.

VI. Sensitivity analysis

In this section, we analyze the robustness of our results for the timing of the treatment, the different specifications of the school time trend, and the exclusion of extreme values, single schools, cantons, and years.

1. Treatment in grade eight

Given that the choice of a training occupation typically occurs in grade eight or nine, we test whether measuring ethnic fractionalization and polarization in grade eight instead of grade nine produces different results. Table A.8 reports the results when the treatment is measured in grade eight, the second-to-last year of compulsory schooling. The results are in line with the results in Table 3, showing that, for a given level of polarization, ethnic fractionalization in grade eight has a positive effect on the likelihood of choosing a people-oriented occupation. Conversely, for a given level of ethnic fractionalization, polarization reduces the likelihood of choosing a people-oriented occupation.

Table A.9 in the Appendix reports the estimated effects of ethnic fractionalization and polarization in different occupations when the two Indices are measured in grade eight. Again, the results are largely consistent with those in Table A.3, when the two Indices are measured in grade nine. Therefore, measuring the treatments in grade eight instead of nine does not produce different results.

2. Time trends

The results of the test in subsection 2. suggest that linear autocorrelation of ethnic fractionalization and polarization within schools is not a concern in our setting. Indeed, the results in Table 3 show that including school-specific linear time trends has little effect on the estimates. However, school-specific time trends might be nonlinear. To deal with this concern, we follow Brenøe and Zölitz (2020), Hill (2017), and Lavy and Schlosser (2011) by including quadratic and cubic school-specific time trends in our regression model. We can thus identify the effect of ethnic fractionalization and polarization by exploiting the deviations of ethnic fractionalization and polarization from their school-specific time trends.

Table A.10 in the Appendix reports the results of these models. It shows that the inclusion of quadratic and cubic time trends has virtually no effect on the estimates of the effect of ethnic fractionalization and polarization. Therefore, we find that school-specific time trends do not drive our results.

3. Exclusion of single cantons, schools, years, and extreme values

To ensure that no single year, canton, or school drives our results, we estimate Eq. 3 by excluding one of the six years (2011 through 2016), one of the 26 cantons, and one of the 2,721 schools at a time. Figure A.7 in the Appendix reports the results, showing that the estimates of the effect of ethnic fractionalization and polarization are robust to the exclusion of single years, cantons, and schools.

To ensure that extreme values of ethnic fractionalization, polarization, and the outcome variable do not drive our results, we estimate Eq. 3 in subsamples excluding schools with average values of ethnic fractionalization, polarization, and the outcome variable above the 99th, 95th, and 90th percentiles and below the 1st, 5th and 10th percentiles.

Table A.11 in the Appendix reports the estimates of these models. The estimates are stable across all restricted subsamples, indicating that schools with extreme values of ethnic fractionalization, polarization, and the outcome variable (the choice of a people-oriented occupation) do not drive our results.

VII. Conclusion

This paper analyzes how school ethnic diversity affects occupational choice and, in particular, the choice of people-oriented occupations. One major novelty is that we distinguish between two dimensions of diversity: fractionalization and polarization. Our results show that an increase in ethnic fractionalization increases the likelihood of choosing people-oriented occupations, mainly at the cost of business-related occupations. Moreover, we find that polarization decreases the likelihood of students sorting into people-oriented occupations.

We show that the effect of ethnic diversity on students' occupational choice could operate through skill formation. Specifically, ethnic fractionalization appears to enhance the student formation of social skills (particularly sociability) but it reduces the formation of language skill. In turn, these effects on skill formation make students more likely to sort into people-oriented occupations in which social skills are comparatively more rewarded.

Overall, our analysis provides novel evidence on the effect of peer ethnic diversity on occupational choice. Our results add to the body of evidence on the effects of ethnic diversity in schools. While previous studies have mainly investigated short-run educational outcomes (e.g., test scores), we analyze a novel post-compulsory schooling outcome—occupational choice—and show that school ethnic diversity also affects post-compulsory schooling outcomes, with important implications for future earnings and career trajectories. Moreover, this paper shows that, even in educational settings, accounting for the combined effect of ethnic fractionalization and polarization is crucial, and that these two distinct dimensions of diversity can have opposing effects on the choice of people-oriented occupations. Given current demographic developments, these occupations are likely to become increasingly important over the next several decades.

Our results also show that teachers and school principals should factor in not only the presence of different ethnic groups (i.e., fractionalization) but also their distribution (i.e., polarization) when configuring classrooms. Indeed, the beneficial effects of ethnic diversity can be enhanced when polarization is low and interethnic interactions are potentially more frequent.

References

Ager, Philipp and Markus Brückner. 2013. "Cultural diversity and economic growth: Evidence from the US during the age of mass migration." *European Economic Review* 64 (C):76–97.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth* 8 (2):155–94.

Alesina, Alberto and Eliana La Ferrara. 2005. "Ethnic diversity and economic performance." *Journal of Economic Literature* 43 (3):762–800.

Anelli, Massimo, Kevin Shih, and Kevin Williams. 2017. "Foreign peer effects and STEM major choice." CESifo Working Paper No. 6466, CESifo.

Angrist, Joshua. 2014. "The perils of peer effects." *Labour Economics* 30 (C):98–108.

Angrist, Joshua and Kevin Lang. 2004. "Does school integration generate peer effects? Evidence from Boston's Metco Program." *American Economic Review* 94 (5):1613–1634.

Balestra, Simone, Aurélien Sallin, and Stefan Wolter. forthcoming. "High-ability influencers? The heterogeneous effect of gifted classmates." *Journal of Human Resources* .

Ballatore, Rosario, Margherita Fort, and Andrea Ichino. 2018. "Tower of babel in the classroom: immigrants and natives in Italian schools." *Journal of Labor Economics* 36 (4):885–921.

Bazzi, Samuel, Arya Gaduh, Alexander Rothenberg, and Maisy Wong. 2019. "Unity in diversity? How intergroup contact can foster nation building." *American Economic Review* 109 (11):3978–4025.

Bifulco, Robert, Jason Fletcher, and Stephen Ross. 2011. "The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health." *American Economic Journal: Economic Policy* 3 (1):25–53.

Black, Sandra, Paul Devereaux, and Kjell Salvanes. 2013. "Under pressure? The effect of peers on outcomes of young adults." *Journal of Labor Economics* 31 (1):119–153.

Borghans, Lex, Bas Ter Weel, and Bruce A. Weinberg. 2014. "People Skills and the Labor-Market Outcomes of Underrepresented Groups." *ILR Review* 67 (2):287–334. URL `https://doi.org/10.1177/001979391406700202`.

Bredtmann, Julia, Sebastian Otten, and Christina Vonnahme. 2021. "Linguistic diversity in the classroom, student achievement, and social integration." *Education Economics* 29 (2):121–142. URL `https://doi.org/10.1080/09645292.2020.1866499`.

Brenøe, Anne Ardila and Ulf Zölitz. 2020. "Exposure to more female peers widens the gender gap in STEM participation." *Journal of Labor Economics* 38 (4):1009–1054.

Brunello, Giorgio and Lorenzo Rocco. 2013. "The effect of immigration on the school performance of natives: Cross country evidence using PISA test scores." *Economics of Education Review* 32 (C):234–246.

Burns, Justine. 2012. "Race, diversity and pro-social behavior in a segmented society." *Journal of Economic Behavior & Organization* 81 (2):366–378.

Carrell, Scott and Mark Hoekstra. 2010. "Externalities in the classroom: how children exposed to domestic violence affect everyone's kids." *American Economic Journal: Applied Economics* 2 (1):211–228.

Carrell, Scott, Mark Hoekstra, and Elira Kuka. 2018. "The long-run effects of disruptive peers." *American Economic Review* 108 (11):3377–3415.

Chevalier, Arnaud, Ingo Isphording, and Elena Lisauskaite. 2020. "Peer diversity, college performance and educational choices." *Labour Economics* 64:101833.

Chuard, Caroline, Annatina Aerne, Simone Balestra, Beatrix Eugster, and Roland Hodler. 2022. "Ethnic Clustering in Schools and Early Career Outcomes." Cepr discussion paper no. 17505, CEPR.

Deming, David. 2017. "The growing importance of social skills in the labor market." *The Quarterly Journal of Economics* 132 (4):1593–1640.

Desmet, Klaus, Ignacio Ortuño-Ortín, and Shlomo Weber. 2009. "Linguistic diversity and redistribution." *Journal of the European Economic Association* 7 (6):1291–1318.

Diette, Timothy and Ruth Uwaifo Oyelere. 2017. "Do limited English students jeopardize the education of other students? Lessons from the North Carolina public school system." *Education Economics* 25 (5):446–461.

Easterly, William and Ross Levine. 1997. "Africa's growth tragedy: Policies and ethnic divisions." *The Quarterly Journal of Economics* 112 (4):1203–1250.

Esteban, Joan and Debraj Ray. 1994. "On the measurement of polarization." *Econometrica* 62 (4):819–851.

Figlio, David and Umut Özek. 2017. "Unwelcome guests? The effects of refugees on the educational outcomes of incumbent students." Nber working paper no. 23661, National Bureau of Economic Research.

Figlio, David N., Paola Giuliano, Riccardo Marchingiglio, Umut Oezek, and Paola Sapienza. 2021. "Diversity in Schools: Immigrants and the Educational Performance of U.S. Born Students." NBER Working Papers 28596, National Bureau of Economic Research, Inc. URL `https://ideas.repec.org/p/nbr/nberwo/28596.html`.

Geay, Charlotte, Sandra McNally, and Shqiponja Telhaj. 2013. "Non-native speakers of English in the classroom: What are the effects on pupil performance?" *The Economic Journal* 123 (570):F281–F307.

Gould, Eric, Victor Lavy, and Daniele Paserman. 2009. "Does immigration affect the long-term educational outcomes of natives? Quasi-experimental evidence." *Economic Journal* 119 (540):1243–1269.

Hall, Joshua C. and Peter T. Leeson. 2010. "Racial Fractionalization and School Performance." *American Journal of Economics and Sociology* 69 (2):736–758.

Hanushek, Eric, John Kain, and Steven Rivkin. 2009. "New evidence about Brown v. Board of Education: the complex effects of school racial composition on achievement." *Journal of Labor Economics* 27 (3):349–383.

Hill, Andrew. 2017. "The positive influence of female college students on their male peers." *Labour Economics* 44:151–160.

Hirsch, Barry T. and David A. Macpherson. 2004. "Wages, Sorting on Skill, and the Racial Composition of Jobs." *Journal of Labor Economics* 22 (1):189–210. URL `http://www.jstor.org/stable/10.1086/380407`.

Hoxby, Caroline. 2000. "The effects of class size on student achievement: New evidence from population variation." *The Quarterly Journal of Economics* 115 (4):1239–1285.

Jackson, Kirabo. 2018. "What do test scores miss? The importance of teacher effects on non-test score outcomes." *Journal of Political Economy* 126 (5):2072–2107.

Jaik, Katharina and Stefan Wolter. 2016. "Lost in transition: The influence of locus of control on delaying educational decisions." Leading House Working Paper No. 118, Swiss Leading House on Economics of Education.

Jensen, Peter and 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.

Kiener, Fabienne, Ann-Sophie Gnehm, and Uschi Backes-Gellner. 2020. "Non-Cognitive Skills in Training Curricula and Heterogeneous Wage Returns." Economics of Education Working Paper Series 0175, University of Zurich, Department of Business Administration (IBW).

Kuhn, Andreas and Stefan C. Wolter. 2022. "Things versus People: Gender Differences in Vocational Interests and in Occupational Preferences." *Journal of Economic Behavior & Organization* 203:210–234. URL `https://www.sciencedirect.com/science/article/pii/S0167268122003201`.

Lavy, Victor, Daniele Paserman, and Analia Schlosser. 2012. "Inside the black box of ability peer effects: evidence from variation in the proportion of low achievers in the classroom." *Economic Journal* 122 (559):208–237.

Lavy, Victor and Analia Schlosser. 2011. "Mechanisms and impacts of gender peer effects at school." *American Economic Journal: Applied Economics* 3:1–33.

Lazear, Edward. 1999. "Culture and language." *Journal of Political Economy* 107 (S6):S95–S126.

Montalvo, José and Marta Reynal-Querol. 2005. "Ethnic polarization, potential conflict, and civil wars." *American Economic Review* 95 (3):796–816.

NCES. 2020. "Racial/Ethnic Enrollment in Public Schools."

OECD. 2020a. *All Hands In? Making Diversity Work for All*. OECD Publishing.

———. 2020b. *International Migration Outlook 2020*. OECD Publishing.

Ohinata, Asako and Jan van Ours. 2013. "How immigrant children affect the academic achievement of native Dutch children." *The Economic Journal* 123 (570):F308–F331.

Pan, Jessica. 2015. "Gender Segregation in Occupations: The Role of Tipping and Social Interactions." *Journal of Labor Economics* 33 (2):365 – 408. URL https://EconPapers.repec.org/RePEc:ucp:jlabec:doi:10.1086/678518.

Pavan, Ronni. 2011. "Career Choice and Wage Growth." *Journal of Labor Economics* 29 (3):549–587. URL http://www.jstor.org/stable/10.1086/659346.

Plug, Erik, Dinand Webbink, and Nick Martin. 2014. "Sexual Orientation, Prejudice, and Segregation." *Journal of Labor Economics* 32 (1):123 – 159.

Rao, Gautam. 2019. "Familiarity does not breed contempt: generosity, discrimination, and diversity in Delhi schools." *American Economic Review* 109 (3):774–809.

Reynal-Querol, Marta. 2002. "Ethnicity, political systems, and civil wars." *Journal of Conflict Resolution* 46 (1):29–54.

Sloane, Carolyn M., Erik G. Hurst, and Dan A. Black. 2021. "College Majors, Occupations, and the Gender Wage Gap." *The Journal of Economic Perspectives* 35 (4):223–248. URL https://www.jstor.org/stable/27074132.

Stoet, Gijsbert and David C. Geary. 2022. "Sex differences in adolescents' occupational aspirations: Variations across time and place." *PLOS ONE* 17:1–18. URL https://doi.org/10.1371/journal.pone.0261438.

Stokes, Bruce. 2017. "What it takes to truly be one of us." Tech. rep., Pew Research Center.

Wolter, Stefan and Paul Ryan. 2011. "Apprenticeship." In *Handbook of the Economics of Education*, vol. 3, edited by Erik Hanushek, Stephen Machin, and Ludger Woessmann, chap. 11. Elsevier, 521–576.

Appendix

Table A.1: Descriptive statistics

	Obs.	Mean	SD	Min	Max
people occupation	238,630	0.119	0.32	0	1
Fractionalization	238,630	0.371	0.24	0	1
Polarization	238,630	0.468	0.22	0	1
School cohort size	238,630	58.343	45.53	10	410
Male	238,630	0.581	0.49	0	1
Age in grade nine	238,630	14.959	0.66	14	17
Swiss national	238,630	0.834	0.37	0	1
Languages					
German	238,630	0.619	0.49	0	1
French	238,630	0.132	0.34	0	1
Italian	238,630	0.052	0.22	0	1
Albanian	238,630	0.052	0.22	0	1
Portuguese	238,630	0.034	0.18	0	1
Serbian and Croatian	238,630	0.034	0.18	0	1
Turkish	238,630	0.015	0.12	0	1
Spanish	238,630	0.011	0.10	0	1
Indo-Aryan and Dravidian languages	238,538	0.009	0.09	0	1
East Asian languages	238,630	0.005	0.07	0	1
Other languages	238,630	0.005	0.07	0	1
English	238,630	0.004	0.06	0	1
Macedonian	238,630	0.004	0.06	0	1
African languages	238,630	0.003	0.05	0	1
Arabic	238,630	0.003	0.06	0	1
West Asian languages	238,630	0.003	0.05	0	1
Romansh	238,630	0.003	0.06	0	1
Other European languages	238,627	0.003	0.06	0	1
Other Slavic languages	238,630	0.003	0.06	0	1
Russian	238,630	0.002	0.04	0	1
Dutch	238,630	0.001	0.03	0	1
Romanian	238,630	<0.001	0.03	0	1
Czech and Slovak	238,630	<0.001	0.03	0	1
Finnish	238,630	<0.001	0.01	0	1
Greek	238,630	<0.001	0.03	0	1
Hungarian	238,630	<0.001	0.02	0	1
Polish	238,630	<0.001	0.03	0	1
Slovenian and Bulgarian	238,627	<0.001	0.02	0	1
Scandinavian languages	238,627	<0.001	0.02	0	1

Descriptive statistics of the estimation sample. Ethnic fractionalization and polarization are measured in grade nine.

Table A.2: Skill requirement in each occupational field

	Humanities Soc. sciences	Business	IT	Engineering	Agriculture	Care	Services
Cognitive skills							
Math skills	+	-	+	+	-	-	-
Language skills	+	+	-	-	-	+	-
Noncognitive skills							
Empathy	+	-	-	-	-	+	+
Team play	-	+	-	-	-	+	+

(1) Arts, humanities, social sciences, journalism, information (e.g., graphic designer); (2) Business, administration, law (e.g., commercial employee); (3) Communication technologies (e.g., information technologist); (4) Engineering, manufacturing, construction (e.g., electronics engineer); (5) Agriculture, forestry, fisheries, veterinary (e.g., farmer); (6) Health, welfare (e.g., social care worker); (7) Services (e.g., chefs). + (-) indicates that the average skill requirement of occupations in that field is above (below) the average skill requirement of all occupations. Data on cognitive skill requirement come from the website "Anforderungsprofile": for a subset of occupations in each field, experts rate the skill requirement of these occupations on a range from 0 to 100. Data on social skill requirement come from the website "Berufsberatung": for a subset of occupations in each field, experts list skills required in these occupations. For each occupation, we compute the proportion of skills that is either "Empathy" or "Teamplay" and average these proportions over all occupations in the field. Weighted averages by the number of apprentices each occupation.

Table A.3: All occupations - Treatment in grade nine

	Humanities Soc. sciences	Business	IT	Engineering	Agriculture	Care	Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fractionalization	0.006	-1.327***	-0.139	0.671	0.106	1.118***	-0.435
	(0.139)	(0.495)	(0.165)	(0.420)	(0.162)	(0.337)	(0.287)
Polarization	0.096	0.560*	0.006	-0.397	0.144	-0.603***	0.193
	(0.101)	(0.332)	(0.115)	(0.301)	(0.128)	(0.231)	(0.202)
Year FE	yes	yes	yes	yes	yes	yes	yes
School-track FE	yes	yes	yes	yes	yes	yes	yes
Individual controls	yes	yes	yes	yes	yes	yes	yes
Cohort controls	yes	yes	yes	yes	yes	yes	yes
School time trend	linear	linear	linear	linear	linear	linear	linear
Obs.	238,630	238,630	238,630	238,630	238,630	238,630	238,630

(1) Arts, humanities, social sciences, journalism, information (e.g., graphic designer); (2) Business, administration, law (e.g., commercial employee); (3) Communication technologies (e.g., information technologist); (4) Engineering, manufacturing, construction (e.g., electronics engineer); (5) Agriculture, forestry, fisheries, veterinary (e.g., farmer); (6) Health, welfare (e.g., social care worker); (7) Services (e.g., chefs). Fractionalization and Polarization are standardized with zero mean and unit variance. The coefficients are multiplied by 100. Standard errors in parentheses. Standard errors clustered at the school-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Effects of ethnic fractionalization and polarization on social skills

	social skills				Cognitive skills		
	(1) Empathy	(2) Sociability	(3) Team play	(4) Tolerance	(5) Math	(6) German	(7) English
Fractionalization	0.104	0.228*	0.138	0.002	-0.051	-0.115**	-0.110*
	(0.086)	(0.130)	(0.105)	(0.147)	(0.051)	(0.050)	(0.059)
Polarization	-0.061	-0.089	-0.094	-0.030	0.024	0.042	0.039
	(0.066)	(0.083)	(0.081)	(0.112)	(0.031)	(0.032)	(0.036)
Year FE	yes	yes	yes	yes	yes	yes	yes
School FE	yes	yes	yes	yes	yes	yes	yes
Cohort controls	yes	yes	yes	yes	yes	yes	yes
School time trend	linear	linear	linear	linear	linear	linear	linear
Obs.	504	504	504	504	2305	2305	2305
Adj.-R2	0.406	0.471	0.354	0.427	0.860	0.885	0.843

The dependent variable is the mean score in the social skills empathy, sociability, team play, and tolerance. The score ranges from 1 to 6, with 6 being the highest value. Each observation is a school cohort. Observations are weighted by the size of the cohort. Ethnic fractionalization and polarization are standardized with zero mean and unit variance. Standard errors are in parentheses. Standard errors are clustered at the school cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: Wages and employment in different ISCED fields

	Humanities Soc. sciences	Business	IT	Engineering	Agriculture	Care	Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Median annual wage 2017 (in CHF)	58,950	65,000	94,700	76,650	60,000	57,050	55,950
Median annual wage FTE, 2017 (in CHF)	76,850	81,550	98,333	80,000	66,300	79,450	69,500
Median annual wage fulltime only, 2017 (in CHF)	76,250	85,150	99,750	80,600	65,000	74,350	72,550
Unemployment	0.07	0.09	0.07	0.08	0.05	0.05	0.11

(1) Arts, humanities, social sciences, journalism, information (e.g., graphic designer); (2) Business, administration, law (e.g., commercial employee); (3) Communication technologies (e.g., information technologist); (4) Engineering, manufacturing, construction (e.g., electronics engineer); (5) Agriculture, forestry, fisheries, veterinary (e.g., farmer); (6) Health, welfare (e.g., social care worker); (7) Services (e.g., chefs). Annual wages and unemployment in each ISCED field in 2017 reported for the subsample of workers aged 22-59 with VET as first education after compulsory schooling (source: Swiss Labor Force Survey, we use the weights provided with the survey).

Table A.6: Other outcomes

	(1) people-oriented t+2	(2) No outcome	(3) Out-migration	(4) Retention
Fractionalization	0.915**	0.361	-0.024	-0.183
	(0.426)	(0.220)	(0.062)	(0.221)
Polarization	-0.747***	-0.513***	0.002	0.165
	(0.286)	(0.151)	(0.041)	(0.150)
Year FE	yes	yes	yes	yes
School FE	yes	yes	yes	yes
Individual controls	yes	yes	yes	yes
Cohort controls	yes	yes	yes	yes
School time trend	linear	linear	linear	linear
Obs.	236,355	456,736	456,736	456,736

The dependent variables are: 1) 1 if the occupation two years after completion of compulsory schooling is people-oriented and 0 otherwise; 2) 1 if we do not observe any post-compulsory schooling outcome and 0 otherwise; 3)1 if an individual out-migrated (i.e., left the country) and 0 otherwise; and 4) 1 if an individual repeated the last year of compulsory schooling and 0 otherwise. In column 1 the sample consists of individuals in the VET track for whom we observe the occupation two years after completion of compulsory schooling. In columns 2-4 the sample consists of all students in their last year of compulsory schooling. The coefficients are multiplied by 100. Standard errors in parentheses. Standard errors clustered at the school-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: Other outcomes

	Outcome					No outcome				
	Obs.	Mean	SD	Min	Max	Obs.	Mean	SD	Min	Max
Male	439,451	0.506	0.50	0	1	17,285	0.418	0.49	0	1
Age	439,451	14.877	0.68	14	17	17,285	15.011	0.76	14	17
Swiss	439,451	0.819	0.39	0	1	17,285	0.644	0.48	0	1
Area										
Urban	437,411	0.582	0.49	0	1	16,855	0.629	0.48	0	1
Suburban	437,411	0.240	0.43	0	1	16,855	0.216	0.41	0	1
Rural	437,411	0.178	0.38	0	1	16,855	0.155	0.36	0	1

Descriptive statistics of the samples of students with and without post-compulsory schooling outcome.

Table A.8: Treatment in grade eight

	(1)	(2)	(3)	(4)	(5)
Fractionalization	1.122***	0.990***	1.118***	1.043***	1.043***
	(0.343)	(0.314)	(0.346)	(0.392)	(0.394)
Polarization	-0.702***	-0.487**	-0.518**	-0.553**	-0.553**
	(0.251)	(0.230)	(0.232)	(0.261)	(0.262)
Year FE	yes	yes	yes	yes	yes
School-track FE	yes	yes	yes	yes	yes
Individual controls	no	yes	yes	yes	yes
Cohort controls	no	no	yes	yes	yes
School time trend	none	none	none	linear	quadratic
Obs.	196,983	196,983	196,983	196,983	196,983

Dependent variable is 1 if the occupation is people-oriented and 0 otherwise. Ethnic fractionalization and polarization are standardized with zero mean and unit variance. The coefficients are multiplied by 100. Standard errors in parentheses. Standard errors clustered at the school-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.9: All occupations - Treatment in grade eight

	Humanities Soc. sciences	Business	IT	Engineering	Agriculture	Care	Services
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fractionalization	0.293*	-1.892***	-0.053	0.642	-0.142	1.043***	0.108
	(0.151)	(0.546)	(0.174)	(0.472)	(0.183)	(0.392)	(0.327)
Polarization	-0.002	1.208***	-0.190	-0.287	0.157	-0.553**	-0.333
	(0.115)	(0.377)	(0.124)	(0.347)	(0.143)	(0.261)	(0.238)
Year FE	yes	yes	yes	yes	yes	yes	yes
School-track FE	yes	yes	yes	yes	yes	yes	yes
Individual controls	yes	yes	yes	yes	yes	yes	yes
Cohort controls	yes	yes	yes	yes	yes	yes	yes
School time trend	linear	linear	linear	linear	linear	linear	linear
Obs.	196983	196983	196983	196983	196983	196983	196983

(1) Arts, humanities, social sciences, journalism, information (e.g., graphic designer); (2) Business, administration, law (e.g., commercial employee); (3) Communication technologies (e.g., information technologist); (4) Engineering, manufacturing, construction (e.g., electronics engineer); (5) Agriculture, forestry, fisheries, veterinary (e.g., farmer); (6) Health, welfare (e.g., social care worker); (7) Services (e.g., chefs). F and P are standardized with zero mean and unit variance. The coefficients are multiplied by 100. Standard errors in parentheses. Standard errors clustered at the school-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.10: Time trends

	(1)	(2)	(3)	(4)
Fractionalization	1.325***	1.119***	1.119***	1.119***
	(0.294)	(0.337)	(0.338)	(0.340)
Polarization	-0.625***	-0.605***	-0.605***	-0.605***
	(0.204)	(0.231)	(0.232)	(0.233)
Year FE	yes	yes	yes	yes
School-track FE	yes	yes	yes	yes
Individual controls	yes	yes	yes	yes
Cohort controls	yes	yes	yes	yes
School time trend	none	linear	quadratic	cubed
Obs.	238,630	238,630	238,630	238,630

Dependent variable is 1 if the occupation is people-oriented and 0 otherwise. Ethnic fractionalization and polarization are standardized with zero mean and unit variance. The coefficients are multiplied by 100. Standard errors in parentheses. Standard errors clustered at the school-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.11: Exclusion of extreme values

	Choice of people-oriented occupation			Fractionalization			Polarization		
	p1/p99	p5/p95	p10/p90	p1/p99	p5/p95	p10/p90	p1/p99	p5/p95	p10/p90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Fractionalization	1.139***	1.050***	0.826**	1.060***	0.875***	0.978***	1.120***	1.129***	1.215***
	(0.332)	(0.348)	(0.377)	(0.331)	(0.317)	(0.308)	(0.339)	(0.346)	(0.359)
Polarization	-0.632***	-0.656***	-0.568**	-0.582**	-0.532**	-0.638***	-0.588***	-0.544**	-0.588***
	(0.232)	(0.247)	(0.272)	(0.229)	(0.230)	(0.238)	(0.228)	(0.220)	(0.218)
Year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
School FE	yes	yes	yes	yes	yes	yes	yes	yes	yes
Individual controls	yes	yes	yes	yes	yes	yes	yes	yes	yes
Cohort controls	yes	yes	yes	yes	yes	yes	yes	yes	yes
School time trend	linear	linear	linear	linear	linear	linear	linear	linear	linear
Obs.	232,802	214,736	190,928	233,811	214,768	190,818	233,712	214,761	190,934

Estimates in subsamples excluding schools with an average fraction of students sorting into people-oriented occupations above the 99th or below the 1st percentiles (1), above the 95th or below the 5th percentiles (2), and above the 90th or below the 10th percentiles (3). We do the same in columns (4)-(6) using the school average ethnic fractionalization and in columns (7)-(9) using the school average ethnic polarization. Ethnic fractionalization and polarization are standardized with zero mean and unit variance. The coefficients are multiplied by 100. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.12: Tracking

	(1) VET	(2) Academic	(3) Specialized schools	(4) AdInterim
Fractionalization	0.342	0.230	-0.169	-0.392
	(0.435)	(0.247)	(0.192)	(0.385)
Polarization	0.089	0.059	-0.079	-0.076
	(0.309)	(0.204)	(0.143)	(0.266)
Year FE	yes	yes	yes	yes
School-track FE	yes	yes	yes	yes
Individual controls	yes	yes	yes	yes
Cohort controls	yes	yes	yes	yes
School time trend	linear	linear	linear	linear
Obs.	439,451	439,451	439,451	439,451

Estimates of the effects of fractionalization and polarization on track choices after compulsory schooling: 1) Vocational education and training (VET); 2) Academic high schools; 3) Specialized schools; 4) Ad Interim (e.g., gap year). The sample consists of students for whom we observe the post-secondary tracking choice. Ethnic fractionalization and polarization are standardized with zero mean and unit variance. The coefficients are multiplied by 100. Standard errors in parentheses clustered at the school-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

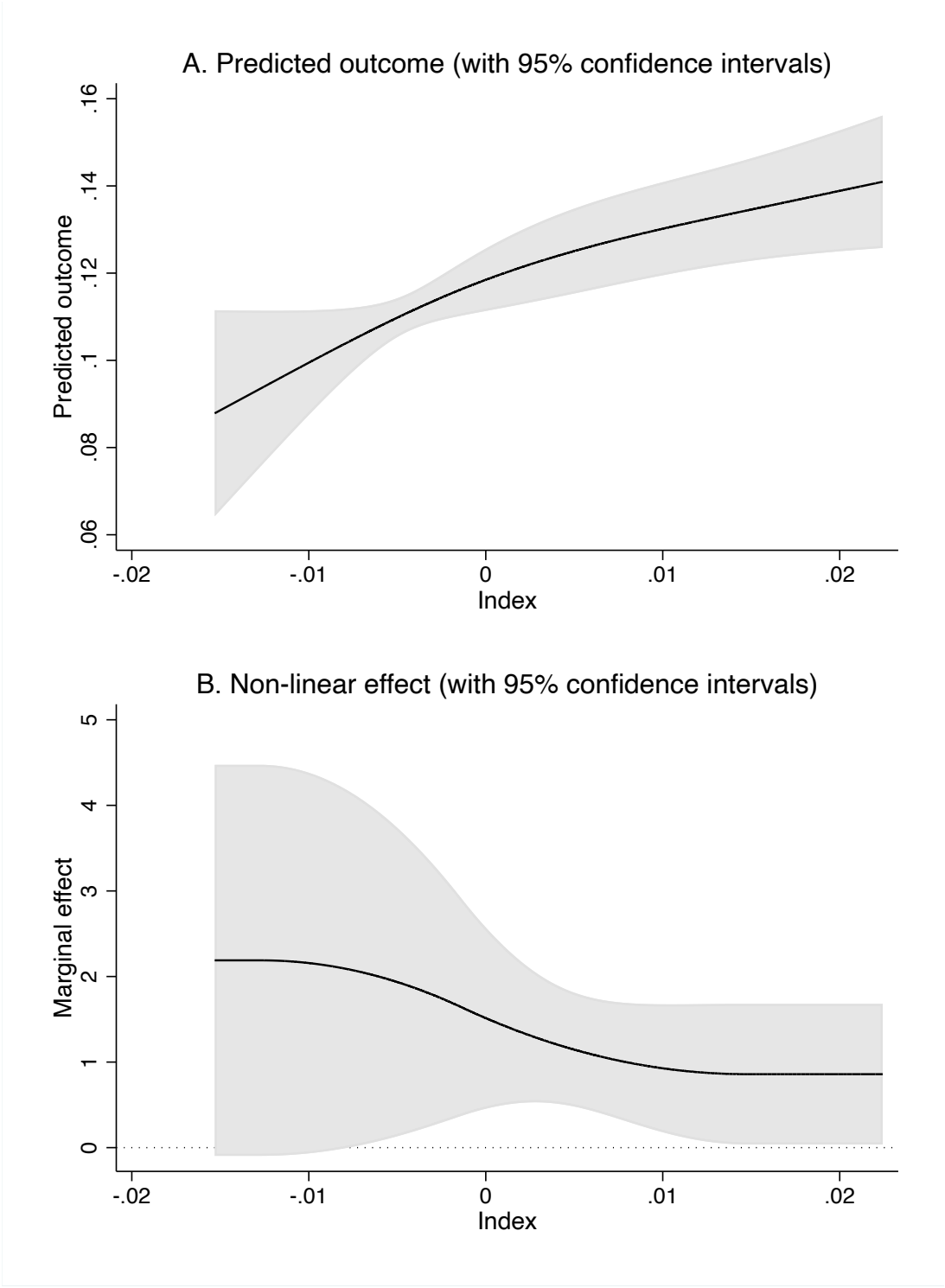


Figure A.1: Predicted outcome and marginal effects of ethnic fractionalization using restricted cubic splines. The estimates are obtained from a model including both ethnic fractionalization and polarization.

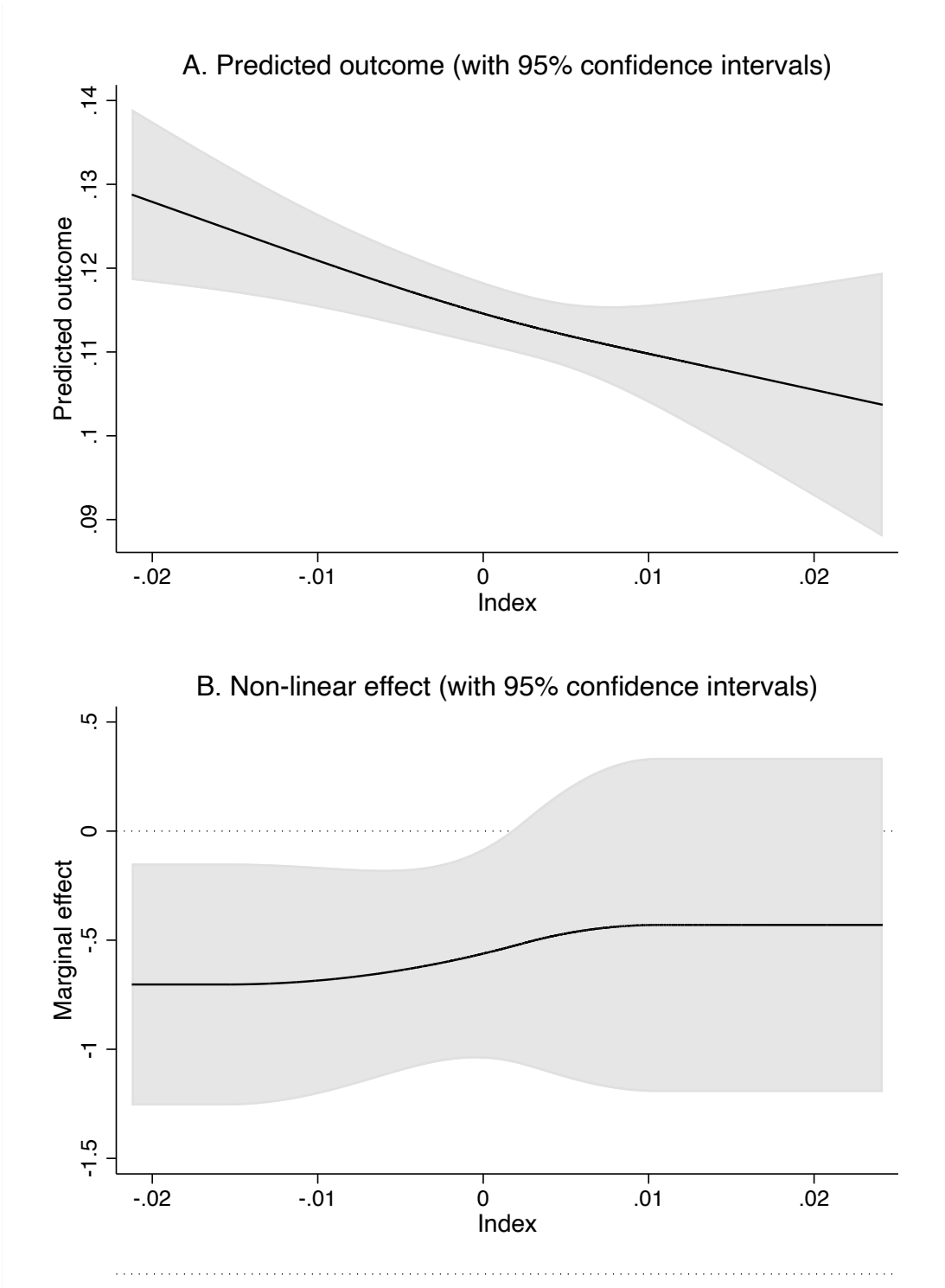
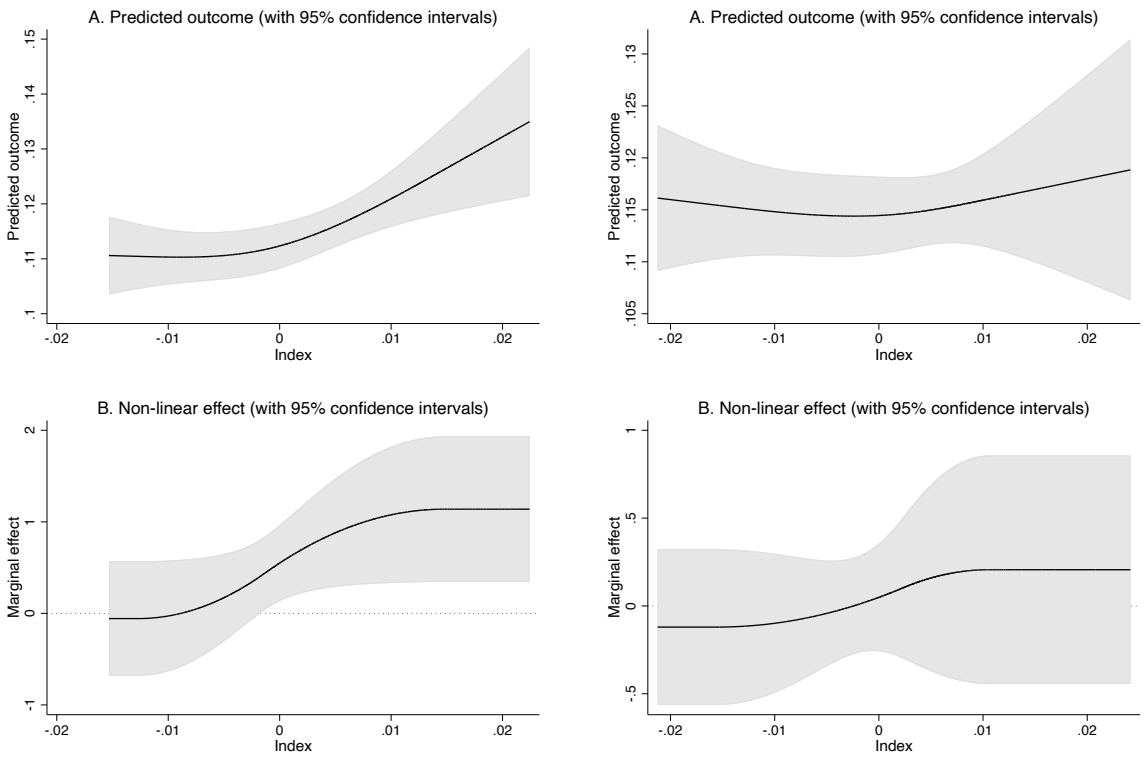


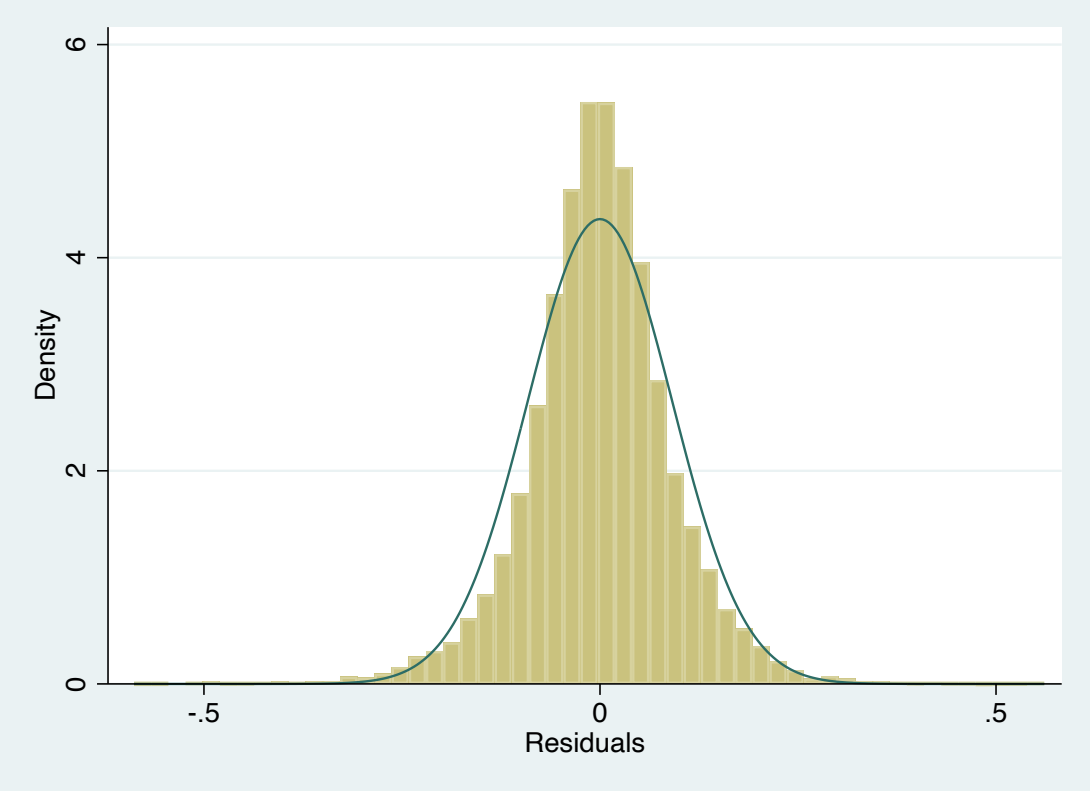
Figure A.2: Predicted outcome and marginal effects of ethnic polarization using restricted cubic splines. The estimates are obtained from a model including both ethnic fractionalization and polarization.



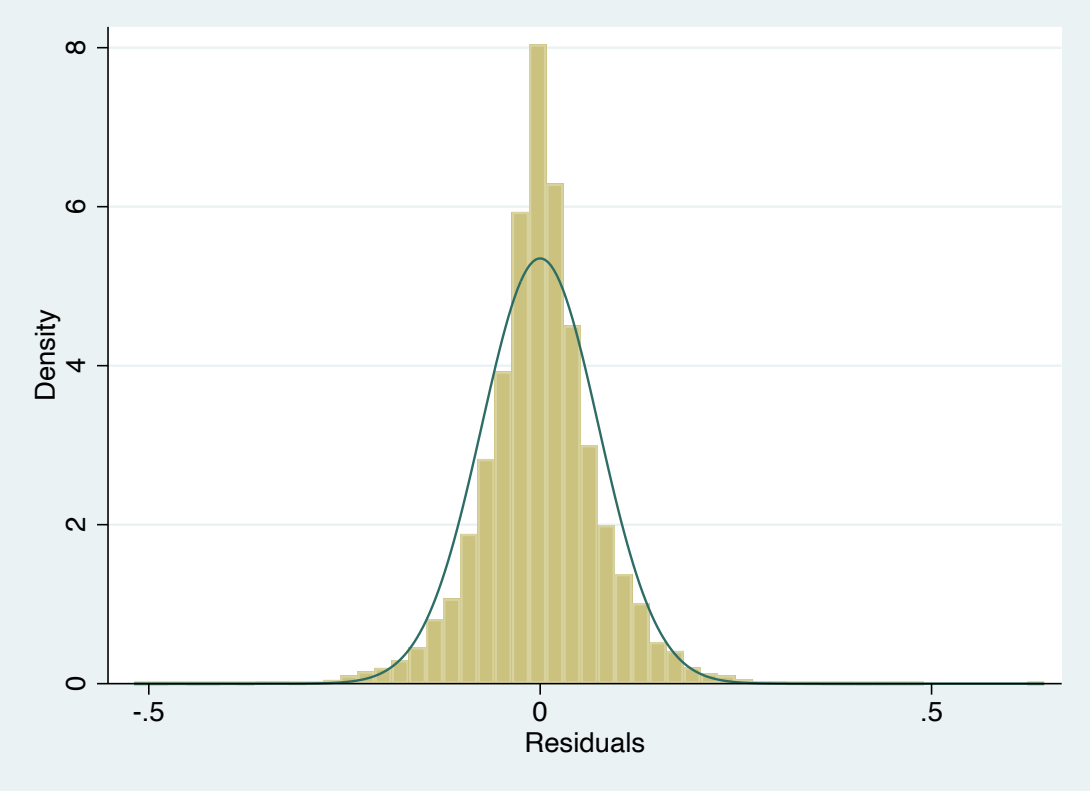
(a) Ethnic fractionalization only

(b) Polarization only

Figure A.3: Predicted outcome and marginal effects of ethnic fractionalization and polarization obtained using restricted cubic splines. The estimates in A.3a are obtained from a model including only ethnic fractionalization. The estimates in A.3b are obtained from a model including only ethnic fractionalization

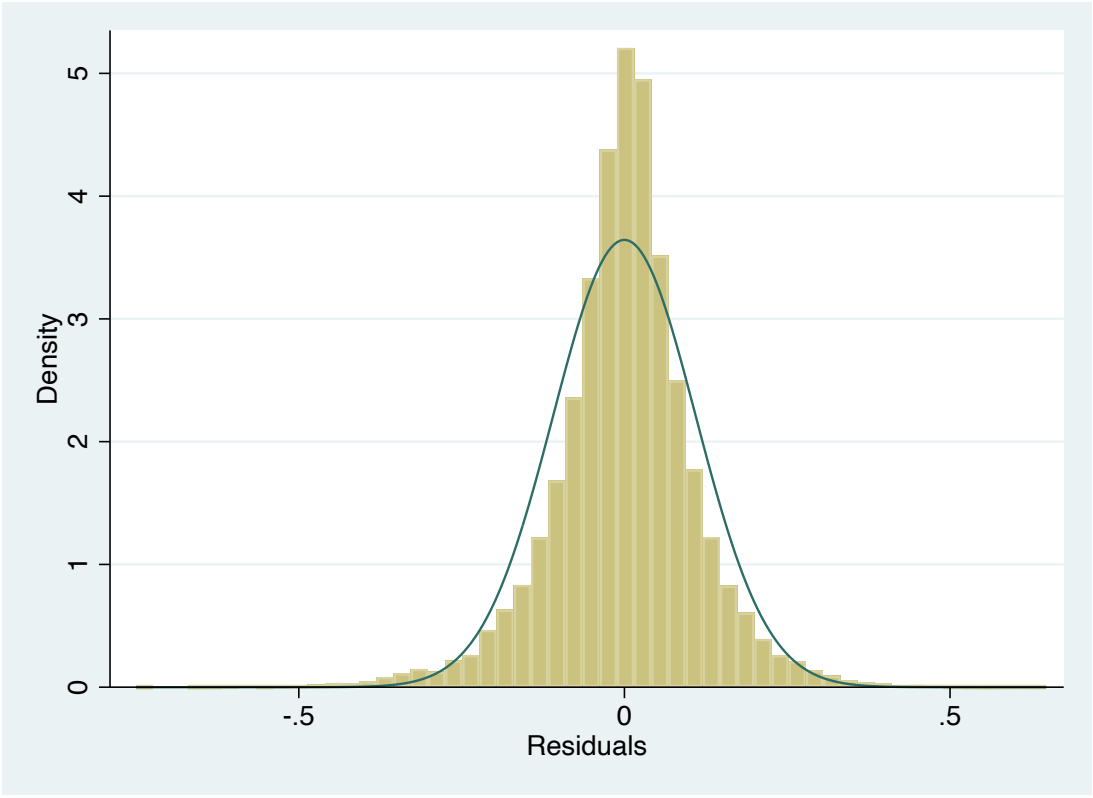


(a) Residuals after school and cohort FE

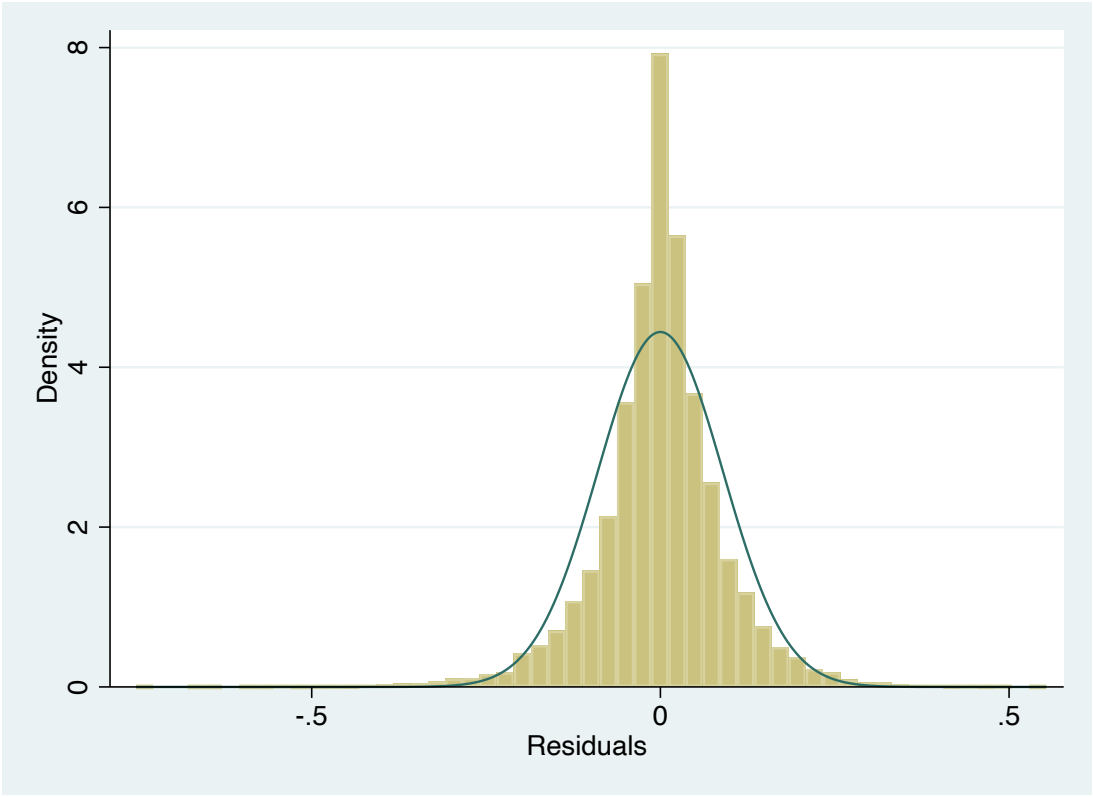


(b) Residuals after school and cohort FE and school trends

Figure A.4: Distribution of residuals and normal distribution (in green), ethnic fractionalization index



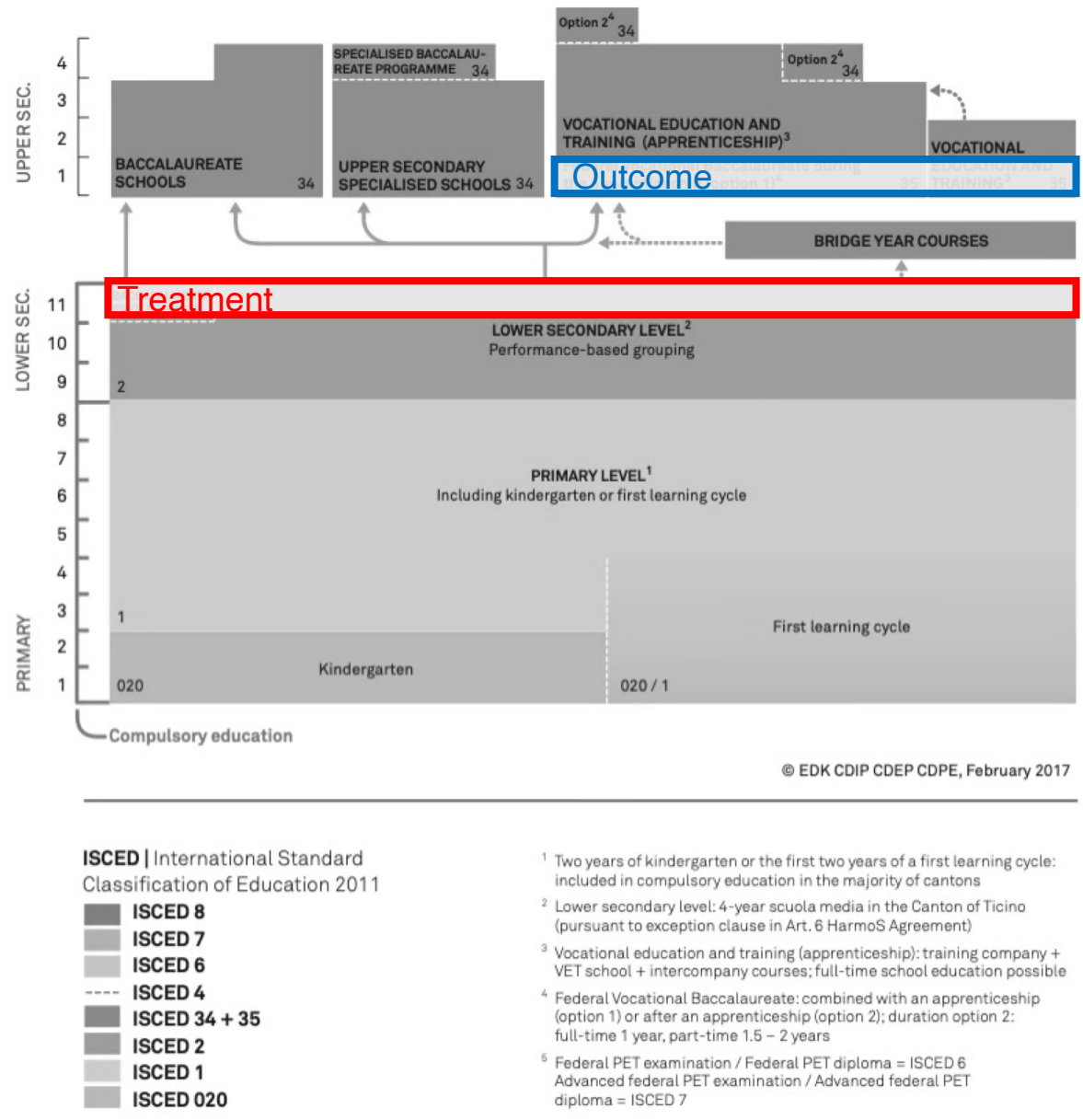
(a) Residuals after school and cohort FE



(b) Residuals after school and cohort FE and school trends

Figure A.5: Distribution of residuals and normal distribution (in green), polarization index

Figure A.6: Swiss educational system and measurement of treatment and outcome



ISCED | International Standard Classification of Education 2011

- **ISCED 8**
- **ISCED 7**
- **ISCED 6**
- **ISCED 4**
- **ISCED 34 + 35**
- **ISCED 2**
- **ISCED 1**
- **ISCED 020**

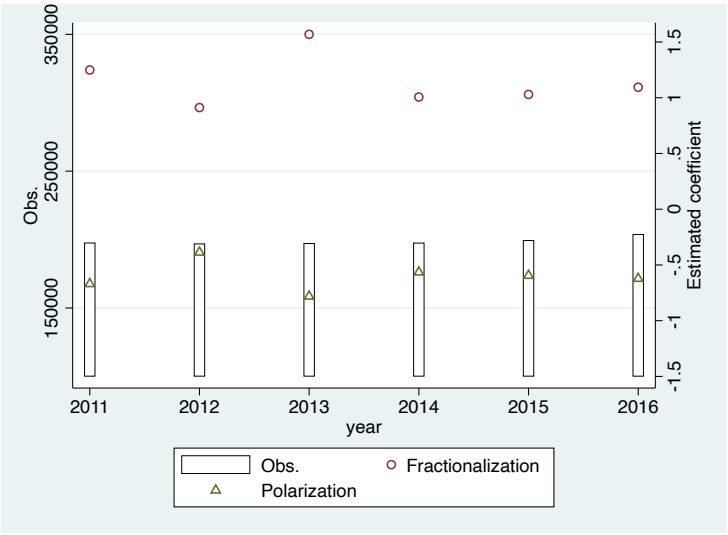
[1] Two years of kindergarten or the first two years of a first learning cycle: included in compulsory education in the majority of cantons

[2] Lower secondary level: 4-year scuola media in the Canton of Ticino (pursuant to exception clause in Art. 6 HarmoS Agreement)

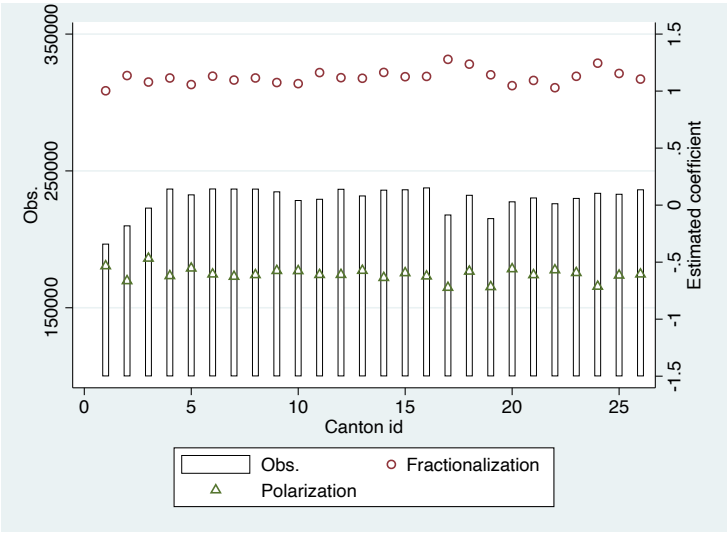
[3] Vocational education and training (apprenticeship): training company + VET school + intercompany courses; full-time school education possible

[4] Federal Vocational Baccalaureate: combined with an apprenticeship (option 1) or after an apprenticeship (option 2); duration option 2: full-time 1 year, part-time 1.5 – 2 years

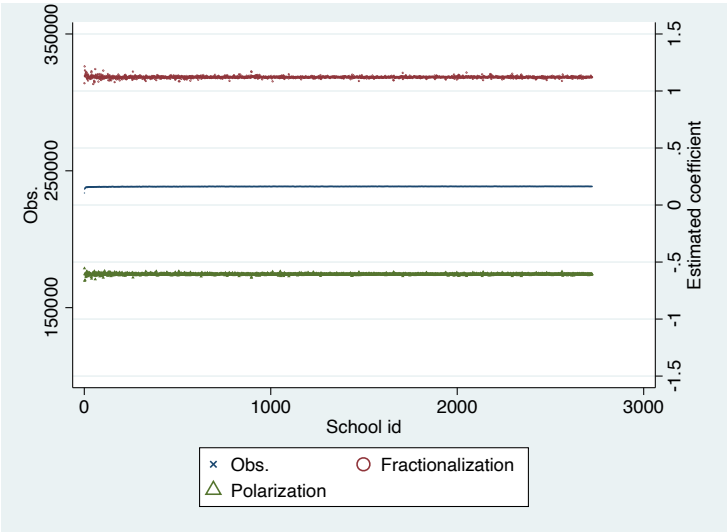
[5] Federal PET examination / Federal PET diploma = ISCED 6
Advanced federal PET examination / Advanced federal PET diploma = ISCED 7



(a) Estimated coefficients of ethnic fractionalization and polarization after excluding one year at a time.



(b) Estimated coefficients of ethnic fractionalization and polarization after excluding one canton at a time.



(c) Estimated coefficients of ethnic fractionalization and polarization after excluding one school at a time.

Figure A.7: Exclusion of years, cantons, and schools.

Figure A.8: Prevalence of ISCED fields of study