Peer Effects and Educational Inequality*

Horst Entorf†
Goethe University Frankfurt

Eirini Tatsi‡
Stockholm University, SOFI

Please do not cite or distribute.

Abstract

We show how social multipliers reinforce preexisting socioeconomic disadvantages of children due to the German early tracking system. We estimate social multipliers from classroom-ability-social-interactions effects for 9th-grade students through a Conditional Quasi-Maximum Likelihood approach: we identify the average effect of classmates' performance from the reduced form and treat a student's potential (self-)selection in a classroom as an omitted variables problem. We find that a 1-point decrease in peer average performance in science leads to a 7.2-points decrease in classroom performance for the vocational training (Hauptschule) students at the lower end of the ability distribution but only to a 2.4-points decrease for the university-path (Gymnasium) students standing at the top of the performance ladder. Moreover, we conclude that a native student's performance in the classroom is more important – either in a beneficial or detrimental way – than an immigrant's performance.

Keywords: educational inequality, social multiplier, early tracking, spatial econometrics
JEL Classification: C31, J15, J18, I20, O15, Z13

*We thank for useful comments participants of the Seminar at IFAU in Uppsala, the Seminar at Linnaeus University in Vaxjo, and the SOFI Brown Bag Seminar in Stockholm.

†Prof. Dr. for the Chair of Econometrics, Department of Applied Econometrics and International Economic Policy, Faculty of Economics and Business Administration, Goethe University Frankfurt, RuW, Postbox 48, Grueneburgplatz 1, 60323 Frankfurt am Main, Germany. Tel. +49(69) 798 34765. Email: entorf@wiwi.uni-frankfurt.de Website: http://www.wiwi.uni-frankfurt.de/abteilungen/ei/professoren/entorf/team/prof-dr-horst-entorf.html.

‡Corresponding author. Stockholm University, Swedish Institute for Social Research (SOFI), Universitetsvägen 10F, 10691, Stockholm, Sweden. Tel. +46(0) 816 1920. Email: eirini.tatsi@sofi.su.se. Website: sites.google.com/site/etatsi.

1 Introduction

We examine the role of peer effects as an enhancer of educational inequality through social multipliers. Prominent examples using the social multiplier to explain undesirable outcomes are Glaeser, Sacerdote, and Scheinkman (1996) on how social interactions magnify crime and Galbiati and Zanella (2012) on how scarcity of auditing services results in a social multiplier in tax evasion.[1] The economics of education view positive peer effects in ability as a useful behavioral outcome because of their potential in policy design. Instead, we place emphasis upon the negative side of social interactions: In case of positive social interactions in achievement, the existence of a social multiplier propagates preexisting inequalities among students in the German early tracking system so that the social multiplier triggers a *vicious* instead of virtuous circle of interdependencies.

First, we show that the German early tracking system tends to cluster students with immigration background and low socioeconomic status in the lower-ability school-type (Hauptschule). Then, we evaluate the extent of educational inequality stemming from peer effects in the classroom by estimating separately effects for each school track as well as from immigration background and native students. Our results show that endogenous social interactions, i.e. the effect of peers' scores on own scores, occur more vividly in the lower-performance "Hauptschule" type than in the top-performance "Gymnasium". More intense social interactions can be detrimental to the "Hauptschule" students, whose background relates to immigration for 31% of the subsample, through the social multiplier effect: A 1-point decrease in peer average performance in the subject of science results in a 7.2-points decrease for a "Hauptschule" student but only in a 2.4-points decrease for a "Gymnasium" student. *This is exactly how preexisting inequalities from early tracking propagate through the social multiplier and enhance educational inequality.* By assuming different ability levels of a newcoming student in the classroom, we conclude that *we minimize losses and maximize gains by increasing a classroom's variance in ability.* Furthermore, we estimate social interactions models that isolate effects stemming from native or immigrant students alone and target first, the whole classroom and second, students with either same or different immigration status. We conclude that *the average performance of native students is the predominant social effect for all school types*, implying policy should be more concerned about where to place a new native instead of a new immigration-background student.

The question of educational inequality is relevant as different academic ability students assume low-skilled and high-skilled occupations respectively later on in the German labor markets. Knowing if and how the classroom environment, and especially classroom

[1] Other applications include Glaeser, Sacerdote, and Scheinkman (2003) on the effects of education on wages, demographics on crime and membership among college roommates, Maurin and Moschion (2009) in mothers' labor force participation, and Bayer, Ferreira, and McMillan (2004) who document a social multiplier on the effect of school quality on neighborhood sorting.

composition, contributes towards an increasing inequality is a topic of policy interest. Quoting the OECD Better Life Index for Germany with respect to PISA scores, *"the average difference in results, between the students with the highest socioeconomic background and the students with the lowest socioeconomic background is 107 points, higher than the OECD average of 96 points. This suggests the school system in Germany does not provide equal access to high-quality education."*[2] Furthermore, the Open Society Justice Initiative filed a report on the discrimination of immigration-background students in several primary and secondary schools in Berlin (Open-Society-Justice-Initiative, 2012). Although in Germany wage inequality between educational groups has hardly changed until the end of the last century (Kahn, 2000; Acemoglu, 2003), wage gaps between educated and low-skilled groups do widen during the most recent years (Dustmann, Ludsteck, and Schoenberg, 2009; Card, Heining, and Kline, 2013). The latter extends to the distribution of employment and unemployment across educational groups as it is the low-educated workforce that suffers the most from economic shocks.[3]

It is only natural that questions of education increasingly grab the attention of educational scientists, labor economists as well as politicians. The debate immediately relates to the access and amount of education for disadvantaged groups while prompts research aiming at the analysis of the nature and causes of the persisting lack of social mobility as well as the examination of the avenues of improving the social mobility of less privileged children through schooling and higher education. Social mobility seems to be very low in Germany, as results in Entorf and Minoiu (2005) have shown: the so-called socioeconomic gradient, which reveals the intergenerational correlation between the socioeconomic status of parents and the educational performance of children, is particular high.[4]

Most of the existing contributions focus on the relatively poor socioeconomic background of immigrants (Gang and Zimmermann, 2000; Frick and Wagner, 2001; Bauer and Riphahn, 2007). More detailed international comparisons reveal that the performance of children with a parental migration background differs strongly across countries (Entorf and Minoiu, 2005). By using educational achievement data from different international sources (PISA, TIMMS and PIRLS), Hanushek and Woessmann (2006) as well as Schnepf (2007) confirm that besides highly variant socioeconomic backgrounds, language problems can also be considered a main source of internationally differing performances of immigrant students. After analyzing immigrants' disadvantage of ten high immigration countries, Schnepf (2007) concludes that natives are on average as much as one grade ahead in their maths skills compared to immigrants in Germany and Switzerland. The situation is different in traditional countries of immigration (New Zealand, Canada, Aus-

[2]http://www.oecdbetterlifeindex.org/topics/education.

[3]A recent example is the aftermath of the global financial crisis, that reportedly affected mainly younger males in export-oriented manufacturing firms (BfA, 2009).

[4]For studies on income mobility see Bjőrklund and Jäntti (1997) and Aaberge, Bjőrklund, Jäntti, Palme, Pedersen, Smith, and Wennemo (2002).

tralia) where the more privileged parental backgrounds of (selected) immigrants lead to less significant or even positive (Canada) differences between educational achievements scores of migrants and natives (Entorf and Minoiu, 2005). Cattaneo and Wolter (2012) recently verify the role of socioeconomic background by exploiting the change in immigration policy to highly-qualified immigrants in Switzerland. Interestingly, Krause, Rinne, and Schueller (2012) find that conditional on similar family backgrounds there are no differences between native and immigrant students with respect to recommendations necessary to enter a specific secondary school type in Germany.

However, all studies reveal that there remains a considerable educational disadvantage of immigrants not explained by observed individual heterogeneity. For instance, given that immigrants are concentrated in large cities and the suburban areas, equality of educational opportunities is limited by spatial and social segregation and the resulting emergence of "good" or "bad" neighborhoods, pointing to spatial selection. A comparatively neglected factor which still seems highly relevant for the composition of peers is the impact of schooling systems. Some authors discuss whether, in addition, the early tracking into different ability schools at the age of 10 as in Austria and Germany might have negative consequences on school performance for children who enter school with language and social deficits, a high proportion of whom come from families with a migration background (see Dustmann, 2004, for a critical assessment of the selective German school systems). Under such circumstances the question of peer effects becomes intertwined with the influence of prevailing national schooling systems and points to school-type selection. In this spirit, Murat, Ferrari, Frederic, and Pirani (2010) as well as Murat (2012) focus exactly on the role of educational systems with PISA data, and uncover that the immigration-native student gap is smaller when the underlying system enjoys greater flexibility, regardless of the comprehensive or tracking nature. Vardardottir (2015) estimates positive effects from students' socioeconomic status in math and problem solving using PISA data from the Swiss streaming system.

Entorf and Lauk (2008) investigated the role that peer effects and social integration of immigrants play for their schooling achievements in selected nations. Their approach is based on the idea that education might have additional positive (or negative) external effects due to social interactions and, more specifically, via the "social multiplier" (see Glaeser, Sacerdote, and Scheinkman, 2003). The novelty of the Entorf and Lauk paper is that by taking migrant-to-migrant, native-to-migrant and native-to-native peer relations into account, the authors are able to test and confirm the hypothesis that early tracking reinforces and even amplifies existing socioeconomic preschool disadvantages of children with an immigration background, because (high) social interactions mainly take place within the group of immigrants and within the group of natives (with detrimental educational effects to immigrants), but less so between both groups. Fruehwirth (2013) finds that peer effects in the classroom are stronger within same races and that low-achieving students benefit more than high achievers from an increase in average peer achievement. Finally, Figlio and Őzek (2017) find that refugees from Haiti have no effect on the educational outcomes of incumbent students in Florida's public schools.

On the other side, one might be concerned with the effect immigration background students have on the school performance of natives. Jensen and Rasmussen (2011) show with PISA data from Denmark that high immigrant school concentration has a negative impact on both natives and immigrants, the effect being larger for the former. In a similar context, Brunello and Rocco (2013) recently use aggregate PISA data for 19 countries and find that a higher share of immigrants can have a negative effect on the scores of natives depending on gender and family background characteristics.

We complement the literature by focusing on the behavioral channels in social interactions instead of the role of exogenous peer characteristics. We employ data from an early-tracking educational system and provide clean evidence on the negative side of social multipliers by using two separate outcomes for each subject (science, math and language) obtained at about the same period of time. Thus, we control for subject-specific unobserved student characteristics and take into account potential (self-) selection into the classroom. Our empirical results cast doubts on the equality of the German educational system.

The rest of the paper is organized as follows: Section 2 briefly explains the German educational environment. Section 3 discusses the data and implications for the German early tracking system. Section 4 provides insight on the estimated econometric models while Section 5 comments on the empirical findings. Finally, Section 6 concludes.

2 A Summary on the Compulsory and Upper Secondary German Educational System

Germany's educational system thrives on an intensive ability tracking route.[5] The Constitution (Grundgesetz) renders Federal States responsible for administration and laws on almost all educational levels, therefore, significant differences may be encountered among States. Preschool education (Kindergarten), relevant for children aged three to six years old, is not compulsory. Children who reach the age of six must attend the primary school (Grundschule) corresponding to grades one to four. At this stage, all students are taught the same curricula on the same subjects and there is no trait of official ability sorting.

After the fourth grade around the age of ten, students are split into a variety of secondary schools; the criteria for a student's transition in one of these paths is based on a combination of his/her academic ability, teachers' recommendations and the family's

[5]The 2011 edition of the German educational system in English can be found under http://eacea.ec.europa.eu/education/eurydice/documents/eurybase/national_summary_sheets/047_DE_EN.pdf.

wishes.[6] The different tracks mainly lead to academic or vocational qualifications with the completion of the compulsory education at the end of grade nine or ten. At the bottom of Germany's secondary education we find the general school (Hauptschule) completed at grade nine (sometimes ten), which leads to the respective general school leaving certificate (Hauptschulabschluss) and most probably to enrollment into further education in the form of vocational training and apprenticeship (Berufschule and Duales System) until the age of 18. A bit higher stands the intermediate school (Realschule), which offers a bit more extensive curriculum than the "Hauptschule" and leads to either full-time vocational training (Berufsfachschule or Fachoberschule) or higher education entrance qualifications. There also exists a school with several educational paths (Schularten mit Mehreren Bildungsgaengen-MBG) conjoining courses offered in the previous two types. Another type is the comprehensive school (Gesamtschule-GS), either cooperative or integrated (Kooperative/Integrierte Gesamtschule-KG/IG) at the end of which students can acquire different qualifications. The cooperative system separates and teaches students according to the available qualification types while the integrated one according to achievement levels for core subjects with no streaming for all other subjects. Obviously, the integrated comprehensive type of school applies further tracking within the school. At the top of the secondary educational system lies the grammar school (Gymnasium), attended by students who wish to obtain university education; its completion leads to the higher education entrance qualification (Abitur).[7] Additional to the aforementioned types of secondary education schools are private schools and, more specifically relevant to Germany's case, the Waldorf school (Waldorfschule) as well as schools for children with mental or physical disabilities (Förderschule or Sonderschule).

The grading scale in the German system ranges from one, being the best mark, to six being the worst. A student receiving five or six in several major courses might have to repeat the grade. It is also noticeable that boarding schools (Internate) do exist in Germany, although they are rather rare, but homeschooling is prohibited. Finally, financing public schools falls within the responsibilities of the State. The above can be deemed as the general rule; nevertheless, as mentioned before, one may find that some States depart substantially.

3 Data and Descriptive Statistics

3.1 The Data

According to the OECD:[8] *"The Programme for International Student Assessment (PISA) is a triennial international survey which aims to evaluate education systems worldwide by*

[6] See Dronkers and Skopek (2015) for details.
[7] See also Dustmann (2004) and Dustmann, Puhani, and Schoenberg (2012).
[8] http://www.oecd.org/pisa/aboutpisa.

testing the skills and knowledge of 15*-year-old students"*. PISA tests the literacy and competences of 15-years-old students in reading, mathematics, science, starting 2003 problem solving and, starting 2012, financial literacy. Each wave places emphasis on a specific subject; the 2006 survey focused on the subject of science. Within the PISA framework, first, a random sample of schools is selected for the international version (PISA-I); then, from each school at most 25 students participate in the survey. The German PISA Consortium enlarges the international PISA sample by testing 2 whole ninth grade classrooms from the same schools chosen for the international sample. We use student final weights to account for the probability of being selected in the sample and to make inferences about the whole population instead of the specific sample (Solon, Haider, and Wooldridge, 2015). We obtained the enlarged sample from the Institute for Educational Progress in Berlin, Germany.[9] For the 2006 survey, the test was based on 13 different booklets, i.e., set of questions; students faced equal probability of getting tested with booklet 1 to 13. Each booklet does not have the same number of questions on each subject but due to the fact that PISA 2006 focused on science, students encountered more questions in science than in math or reading. Thus, we expect the 2006 science PISA scores to be more representative of true ability in the subject of science than math or German language scores. Scores reported are random variables derived from an ability distribution through imputation given performance of students tested with different booklets; the dataset provides 5 plausible values for each subject; although with large sample sizes it is not necessary to use all plausible values (see OECD, 2009; Jerrim, Lopez-Agudo, Marcenaro-Gutierrez, and Shure, 2017), we choose to employ all 5 because we perform analyses on subsamples. Students were asked to complete a questionnaire concerning their background in around 30 minutes; the student dataset also provides information on grades obtained at school in German language, math, physics, chemistry and biology – an element unique to our dataset that we exploit for identification of social interactions parameters. Furthermore, a questionnaire was distributed and completed by the school's principal and teachers and another one by parents.

One concern is how PISA and school scores compare with regards to students' effort: grades obtained from tests at school can be characterized as "high-stakes" tests because failing in many subjects leads to repeating the ninth grade; on the other hand, not performing one's best at the PISA tests does not lead to consequences for the test-takers. Nonetheless, the PISA is considered a "high-stakes" test in Germany: the low performance of German students in the first PISA survey in 2000 was perceived as a national disaster coined as the "PISA shock". Citing Gruber (2006): *"Nowhere did PISA have a similar tsunami-like impact as in Germany. [...] Not just for weeks but for months*

[9]PISA 2006: Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E., & Pekrun, R. (2010): Programme for International Student Assessment 2006 (PISA 2006). Version: 1. IQB – Institut zur Qualitätsentwicklung im Bildungswesen. Datensatz. http://doi.org/10.5159/IQB_PISA_2006_v1.

all major German newspapers and television stations ran special issues and programmes [...]. The Bundestag, the German parliament, held specific "PISA sessions". Evidently, the PISA can be characterized as a "high-stakes" test in Germany – especially for waves $2003 - 2006$ – because performing above the OECD average became an issue of national pride.

We view that our choice of data, i.e., using performance measures form 2 different sources obtained during the same year, has certain advantages over panel data or exclusively PISA data as in Lavy (2015). Panel data are built on yearly observations on school performance alone. The first problem is measurement error that varies among years: apart from true ability, school grades might reflect a student's behavior – seen either as a reward or a punishment – or a teacher's generous/frugal attitude in grading. As this error is teacher-student specific, it varies over years and teachers; hence, the error cannot be captured by student fixed effects. In our case, the underlying measurement error does not vary with the PISA measure because the latter is constructed independently from prior academic performance. Second, as Ramsden, Richardson, Josse, Thomas, Caroline Ellis, Seghier, and Price (2011) discover, the teenage brain undergoes changes in both verbal and non-verbal intelligence, i.e., the Intelligence Quotient (IQ) is not stable between ages $12 - 16$ and $15 - 20$. Therefore, student fixed effects capturing unobserved ability vary from year-to-year, directing to the appropriateness of modeling interactive fixed effects, i.e., time-varying fixed effects. Unfortunately, under cross-sectional dependence – meaning dependence among students' observations as in the case of peer effects – estimation is possible only when the cross-sectional and time observations are both large (see Shi and fei Lee, 2017). We bypass the problem by using performance data roughly obtained at the same period of time, i.e., the ninth grade. Finally, we do not exploit within-student variation in science, math and language performance as in Lavy (2015) but within-variation in PISA and school performance for science, math and German language separately; therefore, first, our within-variation is not potentially driven by the imputation error in the construction of PISA plausible values (see Jerrim, Lopez-Agudo, Marcenaro-Gutierrez, and Shure, 2017), and, second, we allow for unobserved characteristics to vary among the 3 subjects.

3.2 Descriptive Statistics

The 2006 survey sampled $9,577$ students from 203 schools and 395 classrooms (there was only one classroom sampled from 11 schools). From the $9,577$ observations 1764 or 18.42% belong to the "Hauptschule", $3,330$ observations or 34.77% to the "Gymnasium", $2,777$ or 29% of the sample come from the "Realschule", 817 observations or 8.53% from the "MBG" type, and the rest, i.e., 889 or 9.28% of students, were sampled from the "IG" school type. In the analysis we omit the latter two school types as their organization and function are not as clear as the rest for they sort students according to ability and content. We have converted the PISA and school grades for science, math and German language from their original scale to the $0-100$ scale. Regarding PISA plausible values (PPV), there are no missing observations; for the school grades, though, there are missing values: Around 5% for science, 0.3% for math and 0.4% for German language. As discussed below, insofar the true classroom size is known, missing values do not induce measurement error in the peer variables. As there is no subject "science" in school, we created school-grade science as the average of the school grades obtained in physics, chemistry and biology. In Table 3.1 we present basic summary statistics, namely the mean and standard deviation and the number of observations for each variable. For all three subjects, the average grade obtained at school is higher than the average PISA plausible value. Students scored at the PISA test the highest on average in German language (or reading comprehension), followed by math and then science; at school, students obtained on average the highest grades in science, followed by German language and then math. We additionally calculated correlations between each PISA score and school grade; all three correlations are positive: 0.37 for science, 0.28 for math and 0.32 for German language.[10] Low and moderate correlations between PISA and school measures imply sufficient within variation to achieve identification – technically speaking. We also report summary statistics on students' characteristics deemed important in the educational context: gender, age, immigration background and highest (in the family or between the two parents) socioeconomic index of occupational status. 48.5% of the sample is females and the average age of the sampled students are above 15 years old with 15 being the usual age for attending the ninth grade in Germany. Around 17% has some immigration background, broadly defined as being a first or second generation immigrant (or at least one of the parents was born abroad or the student was born abroad if previous information is missing). On average, socioeconomic status corresponds to 49 index points (the $10th$ quantile amounts to 30 points and the $90th$ to 71). As educational inequality refers to both socioeconomic status and immigration background, we further look at the relationship of the two in Figure 3.1: On average, students with immigration background have lower socioeconomic status than their native fellow-students in every school type, while socioeconomic status increases as we move to schools with higher academic ability (lowest for immigrants in the "Hauptschule" and highest for natives in the "Gymnasium").

[10]Correlations between school grades and averages from the five PISA Plausible Values using student final weights.

Table 3.1: Summary Statistics

	Mean	Standard Deviation	Observations
PISA Plausible Value in Science$^{\alpha}$	49.028	0.167	7,871
School Grade in Science	59.132	16.268	7,448
PISA Plausible Value in Math$^{\alpha}$	50.034	0.172	7,871
School Grade in Math	55.659	20.302	7,846
PISA Plausible Value in German Language$^{\alpha}$	53.855	0.165	7,871
School Grade in German Language	58.110	17.111	7,841
Female	0.488	0.500	7,871
Age	15.666	0.630	7,871
Immigration Background	0.181	0.385	7,646
HISEI^{b}	49.885	16.484	7,412

$^{\alpha}$Using five PISA Plausible Values.

^{b}Highest socioeconomic index of occupational status.
Note: Summary statistics using student final weights.

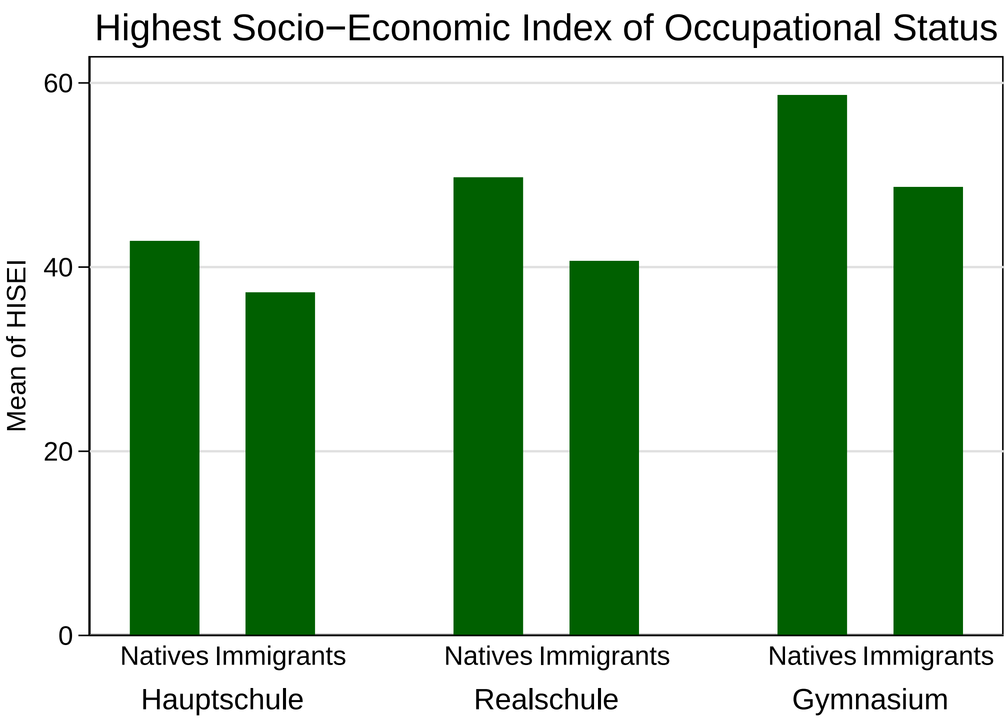

Figure 3.1: Mean of Highest Socioeconomic Index of Occupational Status (HISEI) over Immigration Background.

Germany's tracking secondary educational system unfolds in Figure 3.2 and Table 3.2. In Figure 3.2 we graph means of science, math and German language scores over the performance measure – PPV and school grades – and according to school type. Means of grades obtained at school are, first, higher than the PISA plausible values except for math and German language in the "Gymnasium", and second, more uniform across school types than the PISA plausible values. The latter is due to the fact that school teachers award grades within the same school-type population and syllabus, whereas PISA participants compete on the same test regardless of the school type they attend. Therefore, school grades cannot be directly comparable across school types whereas PISA plausible values can as they provide a more objective measure of competences: "Hauptschule" students score the lowest on average, while "Gymnasium" students stand at the top of the performance ladder. On the basis of performance at the PISA test, Figure 3.2 verifies that "Hauptschule" and "Gymnasium" students possess, respectively, the lower and upper part of the ability distribution. High performers attend the "Gymnasium" and can follow a University education if desired, whereas low performers attend the "Hauptschule" and receive at most vocational training.

In Table 3.2. we present estimation results from a Linear Probability Model (LPM). The dependent variable is the probability of attending a school type from "Hauptschule", "Gymnasium" and "Realschule". Females are less likely to join the "Hauptschule" and more likely to join the "Gymnasium": being female decreases the probability of attending the "Hauptschule" by 6.6% while increases the probability of attending the "Gymnasium" by 4.4%. Females are less likely to receive vocational training, which follows also from the positive sign for the "Realschule". Age serves as a natural proxy of ability – at this age range – as students younger than 15 years old in the ninth grade are grade skippers and those older are grade repeaters. Students one year older have a 12% higher probability of attending the "Hauptschule" and a 14.8% lower probability of attending the "Gymnasium". Older students are more likely to attend the "Realschule" but only by 2.8%. Using the parallel between age and ability in secondary compulsory education, we see that "Gymnasium" students are more likely to be grade skippers and "Hauptschule" students to be grade repeaters, reflecting the difference in academic ability. Immigration background has the largest effect in absolute value for the "Hauptschule": Being an immigrant increases the probability of getting at most vocational training by 11.1%. Immigrant students are less likely to attend the "Realschule" by 8.1% and the "Gymnasium" by 3%. The size of the effects follows the percentage of immigration background students in each school type: 31% for the "Hauptschule", 15.8% for the "Realschule" and 10.7% for the "Gymnasium". Thus, immigrants are more likely to attend the "Hauptschule" and, therefore, more likely to stand at the lower part of the ability distribution in Germany. Socioeconomic background captured by variable HISEI has a positive sign only for the "Gymnasium": A one-standard-deviation increase in the HISEI index increases the probability of attending the "Gymnasium" by about 16%. Between the other two school types, we find the largest negative effect for the "Hauptschule", which is consistent with the OECD's Better Life Index findings for Germany.[11] Table 3.2 verifies that Germany's early tracking system sorts students to different school types according to age (academic

[11]We do not include PPVs and especially school grades in the LPM as those might affect a student's decision to remain in the same school type or change it (reverse causality).

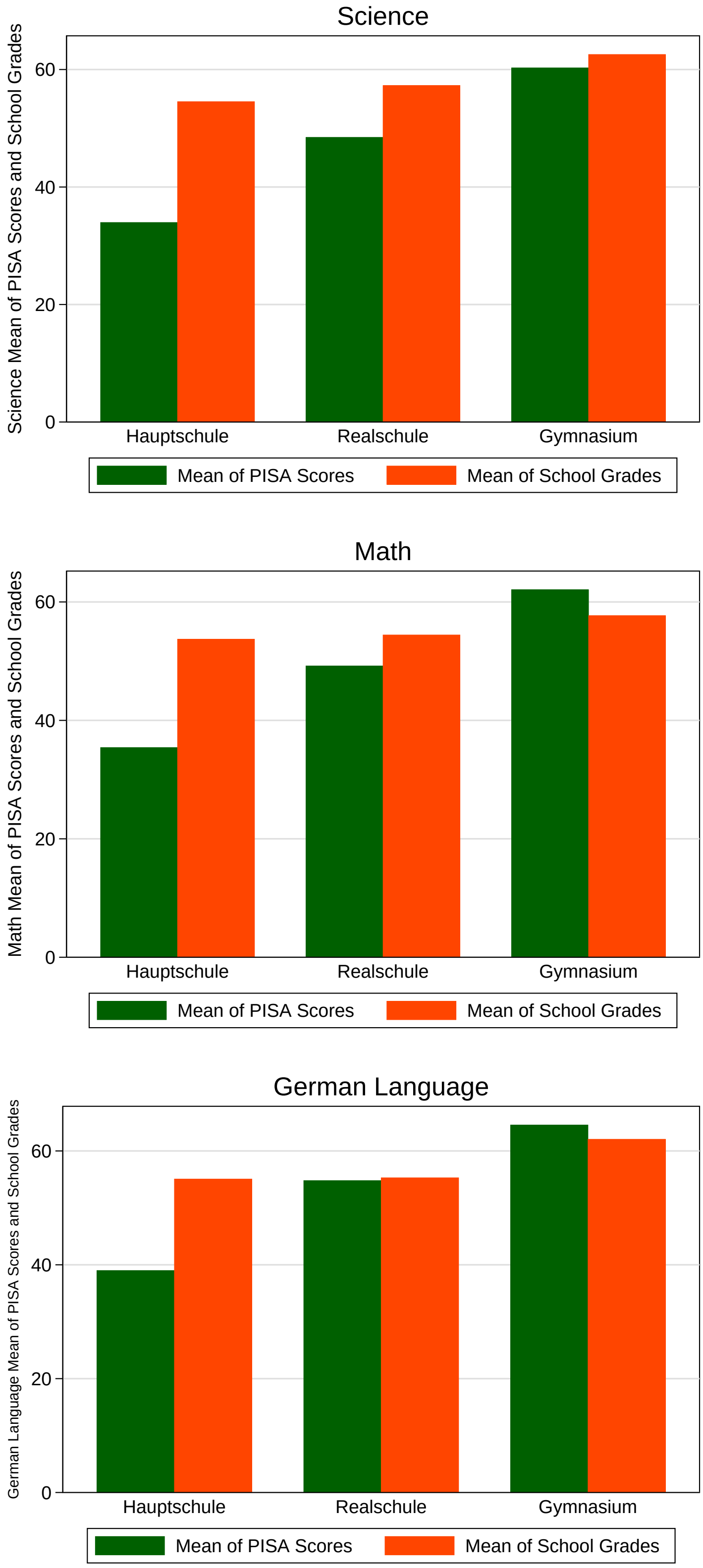

Figure 3.2: Mean of Science, Math and German Language PISA Scores (Average of five Plausible Values) and School Grades over School Type.

ability – grade skippers and repeaters), immigration and socioeconomic backgrounds.

Table 3.2: Determinants of Attending a School Type

Probability of Attending:	HS	GY	RS
	(1)	(2)	(3)
Female	-0.066***	0.044***	0.022**
	(0.011)	(0.011)	(0.011)
Age	0.120***	-0.148***	0.028***
	(0.010)	(0.008)	(0.009)
Immigration Background	0.111***	-0.030**	-0.081***
	(0.017)	(0.014)	(0.015)
HISEI	-0.008***	0.010***	-0.002***
	(0.000)	(0.000)	(0.000)
Number of Observations		7,336	

Note: Estimation of a Linear Probability Model (LPM) including a constant term and using student final weights. Robust to heteroscedasticity standard errors in parentheses. **, *** denote significance at 5% and 1% respectively. HS denotes "Hauptschule", GY "Gymnasium" and RS "Realschule".

4 Econometric Model

4.1 The Model

The peer effects models for student $i = 1, ..., n$ in classroom $c = 1, ..., C$ and performance measure $t = PPV$, *school grade* for subject $g = science$, $math$, $german$ is:

$$y_{itg} = \lambda \frac{1}{n_c - 1} \sum_{j=1, j \neq i}^{n_c} y_{jtg} + x_i \beta_1 + \frac{1}{n_c - 1} \sum_{j=1, j \neq i}^{n_c} x_j \beta_2 + \alpha_{ig} + \varepsilon_{itg} \tag{4.1}$$

in which y_{itg} denotes performance measured by school grades and PISA plausible values in science, math and German language. Average endogenous social interactions in the classroom are captured by term $\frac{1}{n_c-1} \sum_{j=1, j \neq i}^{n_c} y_{jtg}$ so that the scalar coefficient λ denotes the average effect of classmates' performance on own performance – the parameter of interest. The number of peers is simply the number of classmates, $n_c - 1$, and is the same for science, math and German language because all subjects are instructed in the same classroom. We model observed exogenous individual characteristics with the $1 \times k$ vector x_i. Such characteristics are a student's gender, age, socioeconomic status and parental background, immigration background or any trait relevant for an educational achievement equation. The effect of individual or own characteristics is measured by the $k \times 1$ vector β_1. The peer effects model also accommodates peer average characteristics through term $\frac{1}{n_c-1} \sum_{j=1, j \neq i}^{n_c} x_j$; such are the proportion of female or immigration-background classmates, the average age or socioeconomic status of fellow-students in the classroom, etc. The $k \times 1$ coefficient vector β_2 measures the effects from peer average characteristics, also known as contextual, compositional effects or as *exogenous* social interactions effects. The error term, $\boldsymbol{\varepsilon}_{itg}$, is independent and identically distributed across i and t, with mean zero and variance $\sigma^2_{\varepsilon_{tg}}$. We model unobserved individual heterogeneity with the inclusion of the subject-specific α_{ig}; thus, we model student unobserved effects that do not vary between school grades and PISA plausible values but vary with ability in science, math and German language. We allow for α_{ig} to be correlated with any of the explanatory variables, i.e., we treat α_{ig} as fixed and eliminate it with a within transformation. Notice that because $t = 2$, fixed effects and first difference estimators coincide.

The within-transformation sweeps out individual and contextual effects as well, $x_i\beta_1$ and $\frac{1}{n_c-1}\sum_{j=1,j\neq i}^{n_c} x_j\beta_2$ respectively, as they do not vary over the performance measure t (and subject g). Hence, by estimating the 3 econometric models embedded in equation (4.1) – one for each subject $g = science, math, german$ – we focus exclusively on peer behavioral effects keeping own and peer characteristics fixed. In contrast to the exogenous social effects, average peer performance is simultaneously determined with own performance in science, math and German language, therefore, called an *endogenous* social effect. From a policy perspective, estimates of λ are more interesting and important than estimates of β_1 or β_2, first, because although we can intervene and alter students' behavior, i.e., academic performance in this context, we cannot change a student's gender, age or parental background; second, because under average social interactions regrouping does not increase welfare for classrooms of equal size (see Hoxby and Weingarth, 2005).

In order to understand identification, estimation and interpretation of the model, it is useful to write equation (4.1) for classroom $c = 1, 2, ..., C$ and the whole sample $n = n_1 + n_2 + \cdots + n_C$. The socio-matrix for classroom c is:

$$\boldsymbol{W}_c = \frac{1}{n_c - 1}(\iota_{n_c}\iota_{n_c}^{'} - \boldsymbol{I}_{n_c}) = \begin{pmatrix} 0 & \frac{1}{n_c-1} & \cdots & \frac{1}{n_c-1} \\ \frac{1}{n_c-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{1}{n_c-1} \\ \frac{1}{n_c-1} & \cdots & \frac{1}{n_c-1} & 0 \end{pmatrix}, \tag{4.2}$$

in which ι_{n_c} denotes the $n_c \times 1$ vector of ones and $\boldsymbol{I}_{n_c}$ the $n_c \times n_c$ identity matrix. Elements on the main diagonal are zero in order to exclude self-influence. Socio-matrix $\boldsymbol{W}_c$ in equation (4.2) describes a complete network with uniform weights: All students are affected by all classmates in the same way. Dividing each influence with the number of classmates, $n_c - 1$, allows us to model *average* social interactions. For the full sample with C classrooms, the socio-matrix is:

$$\boldsymbol{W}_n = diag(\boldsymbol{W}_1, ..., \boldsymbol{W}_C) \tag{4.3}$$

so that interactions are permitted only within a classroom and not across classrooms even if classrooms belong to the same school. The reason is to avoid modeling non-existent interactions: although we can be certain that students spend many hours together in the classroom, we can only assume they interact with schoolmates outside their classroom. Socio-matrix $\boldsymbol{W}_n$ in equation (4.3) describes a non-complete network with varying weights; the latter depend on classroom sizes which vary from 7 to 34. We can rewrite model (4.1) in matrix notation as:

$$Y_{ntg} = \lambda \boldsymbol{W}_n Y_{ntg} + \boldsymbol{X}_n \beta_1 + \boldsymbol{W}_n \boldsymbol{X}_n \beta_2 + \iota_n \alpha_g + E_{ntg};\, t = PPV, school\, grade \tag{4.4}$$

As we focus on the role of social interactions in educational inequality stemming from the German early tracking system, we estimate model (4.1) or (4.4) for each school type, namely the "Hauptschule" (HS), "Realschule" (RS), and "Gymnasium" (GY) tracks. The most interesting comparison lies between the "Hauptschule" and the "Gymnasium" because they represent the bottom and the top of the academic ability and socioeconomic background distributions, respectively. The "Realschule" stands in the middle and along with the "Gymnasium" are attended by the majority of students.

Apart from socioeconomic background, educational equity refers to a student's immigration status. Table 3.2 reveals the tendency of the German early tracking system to cluster immigration background students more to the "Hauptschule" than the "Gymnasium" school type. Therefore, we estimate endogenous social interactions stemming from native or immigration background students in the classroom for each school type. Following Horrace, Liu, and Patacchini (2016) we define same and different immigration background social effects for classroom c. Let *nat* denote a native student and *im* an immigrant student in classroom c. The socio-matrices that model effects from same immigration-background students in the classroom are:

$$\boldsymbol{W}_{c,nat\,nat} = \begin{array}{c} \begin{matrix} nat & nat & nat & \cdots & im & im \end{matrix} \\ \begin{pmatrix} 0 & 1 & 1 & \cdots & 0 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \end{array}, \tag{4.5}$$

$$
\boldsymbol{W}_{c,im\,im} = \begin{array}{c} \begin{matrix} nat & nat & \cdots & im & im & im \end{matrix} \\ \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 1 \\ 0 & 0 & \cdots & 1 & 0 & 1 \\ 0 & 0 & \cdots & 1 & 1 & 0 \end{pmatrix} \end{array}, \tag{4.6}
$$

in which we have sorted native students first and immigrant students second. Socio-matrix (4.5) models the aggregate effects from native students on native students while socio-matrix (4.6) the effects from immigrant students on immigrant classmates. Native students affect only other native students and immigrant students only other immigrant students, which results in a socio-matrix with blocks of influences. Accordingly, we define the socio-matrices that capture effects from natives on immigrants and from immigrants on natives as:

$$
\boldsymbol{W}_{c,nat\,im} = \begin{array}{c} \begin{matrix} nat & nat & nat & \cdots & im & im \end{matrix} \\ \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & \ddots & 0 & 0 \\ 1 & 1 & 1 & \cdots & 0 & 0 \end{pmatrix} \end{array}, \tag{4.7}
$$

$$
\boldsymbol{W}_{c,im\,nat} = \begin{array}{c} \begin{matrix} nat & nat & nat & \cdots & im & im \end{matrix} \\ \begin{pmatrix} 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \ddots & 1 & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \end{array}. \tag{4.8}
$$

Finally, in order to estimate the effect from native and immigration background students on all other students in the classroom and not only on same/different immigration background classmates, we model the natives' and immigrants' socio-matrices:

$$\boldsymbol{W}_{c,nat} = \begin{array}{c} \begin{matrix} nat & nat & nat & \cdots & im & im \end{matrix} \\ \begin{pmatrix} 0 & 1 & 1 & \cdots & 0 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & \ddots & 0 & 0 \\ 1 & 1 & 1 & \cdots & 0 & 0 \end{pmatrix} \end{array}, \tag{4.9}$$

$$\boldsymbol{W}_{c,im} = \begin{array}{c} \begin{matrix} nat & nat & nat & \cdots & im & im \end{matrix} \\ \begin{pmatrix} 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & 0 & \ddots & 1 & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \end{array}. \tag{4.10}$$

Each of the socio-matrices above can be written as a linear combination of the other. The multicollinearity issue in using more than one socio-matrix simultaneously is more akin to social interactions than to spatial econometrics because in the latter weights-matrices are usually quite distinct and one is not a subset or a linear combination of the other. Therefore, in spatial econometrics it is quite common to estimate Higher Order Spatial Autoregressive (HSAR) models as in Lee and Liu (2010); Badinger and Egger (2011); Lee and Yu (2014). In social interactions, an example of a HSAR is Liu, Patacchini, and Zenou (2014), who estimate models with normalized and unnormalized socio-matrices, or Horrace, Liu, and Patacchini (2016), who include both same and cross type socio-matrices in a single equation. In both cases, effects from one of the two socio-matrices is not statistically significant because of multicollinearity. Thus, under such circumstances, the best strategy is to include only one socio-matrix – as in Horrace, Liu, and Patacchini (2016) – and then apply some criteria to decide on which socio-matrix fits the data best (see Footnote 5 in Lee, Liu, and Lin, 2010). Lee (2008) demonstrates through Monte Carlo simulations that the log-likelihood value from estimation with Maximum Likelihood (ML) increases as the chosen socio-matrix approaches the true socio-matrix. Following Lee (2008), we compare log-likelihood values for models corresponding to the same school type and number of observations to uncover which specification of the socio-matrix prevails.[12] Notice that estimation with any of the socio-matrices depicted in (4.5)-(4.10) is possible

[12] When estimating models with Two Stage Least Squares (2SLS) or Generalized Method of Moments (GMM), models with different socio-matrices can be compared with a J-test (see Kelejian, 2008; Kelejian and Piras, 2011).

when classrooms have at least two native and immigration background students so that we restrict the sample to those classrooms alone. Finally, for the sake of identification, we use the row-normalized versions of socio-matrices (4.5)-(4.10), meaning we divide each row with the number of peers (number of native or immigrant students) so that summing up elements of a row yields one and estimated coefficients measure average instead of aggregate effects.

4.2 Identification

When individuals affect and are affected by everyone in the peer group, identification relies on sufficient group size variation (Kelejian, Prucha, and Yuzefovich, 2006; Lee, 2007; Davezies, D'Haultfoeuille, and Fougére, 2009; Bramoullé, Djebbari, and Fortin, 2009; Lee, Liu, and Lin, 2010).[13] Classroom sizes vary from 7 to 34 for the full sample with 28 distinct group sizes. For the "Hauptschule" we encounter 18 different classroom sizes varying from 7 to 31 students, for the "Gymnasium" 22 sizes from 13 to 34, for the "Realschule" 17 sizes from 14 to 33. Thus, there is enough variation in classroom sizes to identify the endogenous social effect, λ, as classroom size assumes small, medium and large values for each school type.

As demonstrated in Lee (2007) for a model with only group fixed effects, identification is based on two sources: the mean regression function and the correlated distributions of the disturbances. Model (4.1) with student fixed effects uses only the correlated distributions of the disturbances as the within transformation eliminates own and peer characteristics; the within-transformed estimable model resembles a pure spatial autoregressive model (SAR), $Y_n = \lambda \boldsymbol{W}_n Y_n + E_n$, instead of a mixed SAR, $Y_n = \lambda \boldsymbol{W}_n Y_n + \boldsymbol{Z}_n \beta + E_n$, with additional covariates $\boldsymbol{Z}_n$ (see Ord, 1975; Lee, 2004; Bao and Ullah, 2007). Classroom size appears in the reduced form for the model in first differences written for student i:

$$\Delta y_{ig} = \frac{n_c - 1}{n_c - 1 + \lambda} \Delta \varepsilon_{ig} + \frac{\lambda}{(1-\lambda)(n_c - 1 + \lambda)} \sum_{i=1}^{n_c} \Delta \varepsilon_{ig}. \tag{4.11}$$

In practice, we estimate the within-transformed model instead of the first-differenced model but notice that the two are equivalent under $T = 2$ as in our case. We use an orthogonal transformation as proposed by Lee, Liu, and Lin (2010) and Lee and Yu (2010) in order to eliminate student fixed effects and exploit only the linearly independent disturbances. The within-transformed model uses the effective sample size, $n \times (T-1)$, or simply a cross-section:

$$Y^*_{ng} = \lambda \boldsymbol{W}_n Y^*_{ng} + E^*_{ng}, \tag{4.12}$$

[13] When group sizes are identical, it is possible to estimate *same* and *cross* type effects as proposed by Horrace, Liu, and Patacchini (2016).

in which Y^*_{ng} and E^*_{ng} are the within-transformed counterparts of Y_{ntg} and E_{ntg}. The socio-matrix, $\boldsymbol{W}_n$, is not affected by the transformation because it does not vary with performance measure t. The latter implies that model (4.12) preserves its autoregressive nature, meaning we can write its reduced form as:

$$Y^*_{ng} = (I_n - \lambda \boldsymbol{W}_n)^{-1} E^*_{ng}. \tag{4.13}$$

Obviously, $(I_n - \lambda \boldsymbol{W}_n)^{-1}$ exists if $\lambda \in (-1, 1)$ so that equation (4.13) represents a Nash equilibrium in a peer effects game (Calvó-Armengol, Patacchini, and Zenou, 2009). Since $\boldsymbol{W}_n$ is row-normalized, $(I_n - \lambda \boldsymbol{W}_n)^{-1}$ is non-singular if λ assumes values in interval $(-1, 1)$ (Kelejian and Prucha, 2010).

Are Peer Effects just Mechanical?

The endogenous social interactions parameter, λ, might be simply picking preferences or preexisting ability in science, math and German language. The behavioral peer effect does not necessarily reflect interactions or learning and becoming better in a subject from fellow students but merely the fact that students are more (or less) able in the specific subject even in the absence of true social interactions. Our definition of peer groups and our identification strategy ensure that peer effects estimated herein are not of spurious nature for two reasons: First, social interactions are real and not assumed. We choose the classroom – as opposed to the school or the neighborhood or any other vaguely defined peer group – to study peer effects: Students unavoidably spend many hours together with their classmates. Even if students do not want to interact with their classmates, their knowledge and performance is affected when a classmate asks or answers a question or when classmates are noisy and impede the learning process. Second, the peer effects model in equation (4.1) controls for unobserved student characteristics with the inclusion of α_{ig}. Therefore, our estimated peer effects do not reflect own preferences or ability for science, math and German language. Conditional on own preferences and subject-specific abilities, our peer effects pick the plausible-classroom social interactions.

4.2.1 Sources of Bias

Identification and estimation of social interactions parameters is a formidable task. Even though technically group size variation ensures the estimation of parameter λ, there exist multiple sources of bias that might be responsible for uncovering spurious peer effects. Below we discuss and tackle each of the problems in the current setting, namely correlated effects, endogenous group formation, measurement error and missing values.

Correlated effects

Correlated effects refer to common environmental factors that might be affecting students. For instance, schools can be private or public, operate in richer or poorer areas or offer better resources in terms of more experienced and able teachers, access to computer facilities, libraries, etc. Correlated effects also refer to common shocks, e.g., a one-time policy that improved performance but is erroneously perceived as the result of social interactions. Such effects can be captured by school or classroom fixed effects so that the within transformation at the student level removes correlated effects as well. In general, correlated effects refer to unobserved group instead of individual characteristics; therefore, group fixed effects suffice to capture them.

Endogenous Group Formation

Selectivity or endogenous group formation refers to an individual's decision to connect with another individual and form a network or a peer group according to own preferences and characteristics. There are two ways to view bias from endogenous group formation. First, the formation of a link, i.e., elements of the socio-matrix, depends on unobservable individual characteristics so that the socio-matrix is correlated with the error term (see, for instance, Lee, 2007; Bramoullé, Djebbari, and Fortin, 2009; Lee, Liu, and Lin, 2010; Bramoullé, Kranton, and D'Amours, 2014 for network and group fixed effects). Once the individual unobserved effects are controlled for – either with within-transforming or first-differencing the model – network or group formation is considered exogenous, i.e., conditionally on individual fixed effects, the socio-matrix is exogenous. Obviously, this solution requires at least two observations per cross-section. Second, we view the formation of a link and the outcome equation as a two-step procedure. Estimation of the link formation equation results in estimation of an individual level selectivity bias that enters the outcome equation and controls for possible endogeneity of the socio-matrix. Estimation of the network formation and the outcome equations can be simultaneous or sequential (see Goldsmith-Pinkham and Imbens, 2013; Hsieh and Lee, 2014; Qu and fei Lee, 2015; Qu, fei Lee, and Yu, 2017; Johnsson and Moon, 2017). This solution requires an exclusion criterion for the selectivity equation.

In our context, unobserved student characteristics, e.g., students' behavior, incidence of bullying or individual preferences, might be responsible for classroom formation in only a few cases: in reality, parents choose the neighborhood – and, therefore, the school – the school principal forms classrooms based on observed or unobserved-to-us students' characteristics and students cannot choose whether to attend school as the ninth grade is compulsory. Furthermore, students do not form their classrooms individually as we would end up with as many classrooms as students; thus, classroom formation is not akin to friendship formation. Nevertheless, under extraordinary circumstances, students might opt to change classrooms based on unobserved-to-us characteristics so that W_c is potentially endogenous due to its correlation with α_{ig}. We exploit unique information in our data that provides two distinct measures of science, math, and German language competences around the same period of time to control for student fixed effects and consider W_c as conditionally exogenous.

Measurement Error and Missing Values

In peer effects, measurement error refers to missing values, which create distortions in individual influences. Davezies, D'Haultfoeuille, and Fougére (2009) generalize Lee's (2007) identification results to accommodate for this incidence. If the true size of each group, m_r, is known, then, measurement error in the peer variables from missing observations does not pose a problem. To illustrate, let n_r denote the number of observed individuals in group r, and m_r the true number of members in the group. The effect of the missing is the same for all observed students in classroom c, $\frac{1}{m_c-1}\sum_{j=1,j\neq i}^{m_c-n_c} y_{jtg} = \theta_c$, and, therefore, gets eliminated by the within transformation. In practice, we divide with the known true classroom size, m_r, but sum across the observed number of students, n_r, so that the classroom socio-matrix in equation (4.2) becomes:

$$\boldsymbol{W}_c = \frac{1}{m_c - 1}(\iota_{n_c}\iota_{n_c}^{'} - \boldsymbol{I}_{n_c}). \tag{4.14}$$

Notice, first, that the within transformation also sweeps out any other peer influences, for instance ability peer effects from friends who did not take the PISA test or private tutoring at home. Second, the problem with missing values lingers when building socio-matrices on immigration background as in equations (4.5)-(4.10) because when information about immigration background is missing, we cannot infer the true number of natives and immigrants in the classroom. Nevertheless, the percent of missing regarding information on immigration background is only 2.86% in the whole sample.

4.3 Estimation

Term $\boldsymbol{W}_n Y_{ng}^*$ in equation (4.12) is endogenous because Y_{ng}^* and $\boldsymbol{W}_n Y_{ng}^*$ are simultaneously determined even if $\boldsymbol{W}_n$ is exogenous. In principle, estimation can be performed both with Instrumental Variables (IV) (Kelejian and Prucha, 1998; Lee, 2003) and with (Conditional) Quasi Maximum Likelihood (QML) on the reduced form (Lee, 2007; Lee, Liu, and Lin, 2010). Regarding the former, valid instruments emerge from the linearly independent columns of $\boldsymbol{W}_n^2 X_n$ or of higher powers. From the reduced form in equation (4.13) it becomes obvious there are no relevant instruments because exogenous students' characteristics are swept out by the within transformation. Therefore, model (4.12) can be estimated only with QML. The unconcentrated likelihood function derived from the reduced form 4.13 under normality of the error term for subject $g = science, math, german$ is:

$$lnL_n\left(\lambda, \sigma^2\right) = -\frac{n}{2}ln\left(2\pi\right) - \frac{n}{2}ln\sigma^2 + ln|\boldsymbol{I}_n - \lambda\boldsymbol{W}_n| - \frac{1}{2\sigma^2}E^{*\prime}_{ng}E^{*}_{ng}. \tag{4.15}$$

Estimation corresponds to both Conditional and Quasi Maximum Likelihood: the former because $\overline{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}$ provides a sufficient statistic for α_i – the student fixed effect – and the latter because innovations are i.i.d. but not necessarily normally distributed.

4.3.1 Interpretation

In theory, the endogenous social interactions parameter λ can take both negative and positive values. Negative values are quite uncommon in the peer effects literature in education as they signify anticonformist behavior (Sommer, 2016). According to Boucher and Fortin (2016) positive dependence implies conformity and/or complementarity in behavior. If $0 < \lambda < 1$, then the marginal effect from an increase (decrease) in peer average performance is not simply $\lambda\left(-\lambda\right)$ but

$$\left(I_n - \lambda W_n\right)^{-1} = I_n + \lambda W_n + \lambda^2 W_n^2 + \lambda^3 W_n^3 + \ldots\,. \tag{4.16}$$

Equation (4.16) reveals a dying out circle of interdependencies as each additional term approaches more and more to zero. When W_n is row-normalized as in our case, the average effect to a single observation denoted by $\frac{1}{n}\iota_n'\left(I_n - \lambda W_n\right)^{-1}\iota_n$ reduces to $\frac{1}{1-\lambda}$, known as the social multiplier. To illustrate why the marginal effect is $\frac{1}{1-\lambda}$ instead of λ in the current setting, consider the effect of a 1-point decrease in peer average scores on the focal student: a1-point decrease in peer average scores leads to a λ-point decrease in scores for the focal student. The contribution of the focal student to the classroom average is $\frac{1}{n}$, so that the λ-point decrease from the focal student induces a $\frac{\lambda}{n}$ points decrease in *classroom* average scores and a $\frac{\lambda}{n-1}(n-1) = \lambda$ points further decrease in *peer* average scores. The latter effect on the focal student is $\lambda \times \lambda = \lambda^2$ points. The λ^2 points decrease for the focal student decreases the *classroom* average by $\frac{\lambda^2}{n}$ points and the *peer* average by $\frac{\lambda^2}{n-1}(n-1) = \lambda^2$ points. Thus, for this circle of interdependencies the effect on the focal student is $\lambda^2 \times \lambda = \lambda^3$ points. Infinite repetition of the procedure yields the final impact of a 1-point decrease in *peer* average scores or simply the social multiplier:

$$1 + \lambda + \lambda^2 + \lambda^3 + \lambda^4 + \cdots = \frac{1}{1-\lambda}. \tag{4.17}$$

The closer the value of λ is to one, the higher the powers of term λW_n needed for the effect to fade away. As the social multiplier is a nonlinear function of λ, we calculate its standard error (se) with the delta method:

$$se\left(\frac{1}{1-\lambda}\right) = se\left(\lambda\right)\frac{1}{\left(1-\lambda\right)^2}. \tag{4.18}$$

The social multiplier follows from the autoregressive nature of the model – see the reduced form equation (4.13): Whenever we can write the reduced form as in equation (4.13), the marginal effect from a change in peer average performance is $\frac{1}{1-\lambda}$ (and not λ), incorporating both direct and indirect influences until they die out. The latter necessarily die out when the parameter space for λ is $(-1, 1)$ ensured by the row-normalization. When $\lambda = 1$ the model is said to have a unit root; in this case there is no equilibrium and the model is not estimable. Lee and Yu (2013) study the case when parameter λ is very close to 1, i.e., the near unit root model.

5 Estimation Results

In each of the Tables 5.1 to 5.4, we present first the estimated initial dependence among classmates' performances (endogenous effects, $\widehat{\lambda}$) and then the corresponding resulting social multiplier for each academic track ("Hauptschule", "Gymnasium" and "Realschule") and subject (science, math and German language). We base conclusions mainly on the subject of science – as per the 2006 PISA focus – but nevertheless report results for math and German language as well; as explained in Subsection 3.1, science PISA scores for 2006 are a more accurate measure than scores in math or German language because students encountered more questions on science.

Overall, first, all endogenous social interactions parameters, $\widehat{\lambda}$, are positive and highly statistically significant, which verifies that educational outcomes are positively interdependent; the result is conditional on own characteristics, classroom composition, and the possibility of (self-)selection or endogenous group formation. With only two exceptions, dependence is higher in the "Hauptschule" compared to the "Gymnasium": *students at the lower part of the ability distribution are affected more by their peers than students at the upper part of the ability distribution.* The magnitude of the endogenous effects parameter for the "Realschule" lies in most cases in the middle of the corresponding estimates for the "Hauptschule" and the "Gymnasium" for the subject of science and closer to that of the "Gymnasium" for the subjects of math and German language. Second, apart from effects of natives on immigrants for the "Gymnasium" and immigrants on natives for the "Hauptschule", there exists a significant social multiplier: *when the endogenous effects parameter is above* 0.5*, the final impact of an initial* 1*-point change in average peer scores is larger than* 2 *points.* Below we discuss more thoroughly our empirical findings.

5.1 Social Interactions According to School Type

Table 5.1 shows how the presence of positive social interactions propagates preexisting inequalities stemming from early tracking: through higher multiplier effects in the "Hauptshule", which is attended more likely by students of lower socioeconomic status and immigration background (see Table 3.2) than in the "Gymnasium". By looking at the first row of Table 5.1, we observe that students attending the lower-performing schools (HS) are more intensely influenced by classmates than students attending the top-performing schools (GY): compare 0.861 with 0.584 for the subject of science keeping in mind that parameter λ takes values between (0, 1). As the same is true for both math and German language, we infer that the "Hauptschule" students are affected more than the "Realschule" and, especially, the "Gymnasium" students from the same-amount change in average classroom performance in any of the three subjects. The exact magnitude of the impact depends on the social multiplier, presented below the endogenous effects estimate in Table 5.1: A 1-point change in peer average performance in science results in a 7.2-points change for a "Hauptschule" student but only in a 2.4-points change for a "Gymnasium" student. The change for the "Gymnasium" is only one third of the change in the "Hauptschule".

Notice that the estimates presented in Table 5.1 are – among other things – free of bias from missing values. We produced results using average social effects depicted in socio-matrix (4.14). Therefore, the difference in the number of observations between science and math or German language in the "Hauptschule" due to missing values is not alarming. As mentioned in Section (4.2), the within-transformation eliminates the effect of the missing as long as we construct social influences using the true number of peers – as we do following Davezies, D'Haultfoeuille, and Fougére (2009).

There are two ways to exploit the results of Table 5.1 from a policy perspective: In which school type to implement educational policies that increase performance and in which school-type classroom to place a new student. Regarding the former, any policy such as supplementary teaching or tutoring that results in an increase in classroom performance through any student, has a positive effect to all school types. Nevertheless, the greatest gains are found for the "Hauptschule" rather than the "Gymnasium" or even the "Realschule": *It is the students standing at the lower parts of the ability distribution that benefit the most from average increases in classroom performance.* To illustrate, assume two classrooms of the same size with the first belonging to the "Hauptschule" and the second to the "Gymnasium". Further assume that the same policy costing the same amount of euro is implemented in both classrooms: A 1-point increase in peer average performance in science from the policy translates into a 140% increase for the "Gymnasium" but in a 620% increase for the "Hauptschule" classroom attributable to the presence of the social multiplier. Clearly, the benefit per euro spent is much higher for the "Hauptschule".

Table 5.1: Social Interactions Models According to School Type

	Science			Math			German Language		
Average Social Effects	HS	GY	RS	HS	GY	RS	HS	GY	RS
Endogenous Effects, $\widehat{\lambda}$	0.861***	0.584***	0.749***	0.813***	0.601***	0.636***	0.820***	0.613***	0.610***
	(0.010)	(0.019)	(0.013)	(0.012)	(0.018)	(0.018)	(0.011)	(0.017)	(0.022)
Social Multiplier, $\frac{1}{1-\widehat{\lambda}}$	7.194***	2.403***	3.984***	5.348***	2.506***	2.747***	5.556***	2.584***	2.564***
	(0.518)	(0.110)	(0.206)	(0.343)	(0.113)	(0.136)	(0.339)	(0.113)	(0.145)
Log-likelihood	-8,513	-19,915	-16,748	-11,315	-20,815	-17,369	-11,047	-20,288	-16,767
Number of Observations	1,360	3,323	2,765	1,757	3,323	2,766	1,753	3,322	2,766

Note: Estimation with QML including individual fixed effects (CML) using student final weights and five PISA Plausible Values. Average social interactions in the classroom. Standard errors in parentheses. *** denotes significance at 1%. HS denotes “Hauptschule”, GY “Gymnasium” and RS “Realschule”.

Regarding the latter, i.e., in which classroom to place a new student, the answer depends on the student's performance relative to the classroom average. Figure 3.2 shows that "Hauptschule" mean performance either in PISA or at school is always lower from the respective "Gymnasium" mean performance for any of the three subjects (science, math and German language). Therefore, we discern three cases assuming again two classrooms of equal size, one in the "Hauptschule" and the other in the "Gymnasium":

1. The new student's performance is higher than both the "Hauptschule" and "Gymnasium"average classrooms' performance. Since adding the new student increases the classroom average in any case, the new student should be placed in the "Hauptschule" classroom because the benefit is much higher.

2. The new student's performance is lower than both the "Hauptschule" and "Gymnasium" average classrooms' performance so that adding the new student decreases the classroom average in any case. The new student should be placed in the "Gymnasium" classroom to minimize the negative effect of interdependencies.

3. The new student's performance lies between the mean performance in the "Hauptschule" and "Gymnasium" classrooms. Obviously, since the new student increases average performance for the "Hauptschule" classroom but decreases it for the "Gymnasium" classroom, the new student should be placed in the "Hauptschule" classroom.

The empirical exercise reveals that *we minimize losses and maximize gains by increasing the classroom's variance in ability.* Nevertheless, since the German early-tracking system relies on previous performance in order to recommend which track a student should follow at the age of 10 or afterwards, in reality it necessarily places low-performing students at the "Hauptschule". Furthermore, in Table 3.2 of Subsection 3.2 we established that students with lower socioeconomic status and immigration background are more likely to attend the "Hauptschule" instead of the "Gymnasium". Thus, lower-ability students with lower socioeconomic status get trapped in the vicious circle of the social multiplier from attending the same classroom with other lower-ability and lower socioeconomic status students: *This is exactly how preexisting inequalities from early tracking propagate through the social multiplier and enhance educational inequality that results in low social mobility.*

Comparing with policies such as Carrell, Sacerdote, and West (2013), notice that we do not create new or different peer groups in order to help certain students; we merely assume adding a new student to already existing peer groups of the same size in the environment of the "Hauptschule" or the "Gymnasium". Also, our policy exercise is conditional on the characteristics of both the peer group and the additional student. Thus, the result and policy implications are robust to any potentially new endogenous group formation from bringing a new student into the classroom.

5.2 Social Interactions According to School Type and Immigration Background

In Tables 5.2-5.4 we model average social interactions from native and immigration background students to, first, the whole classroom, second, to same immigration background classmates (native-to-native and immigrant-to-immigrant channels), and third, to different immigration background peers (native-to-immigrants and immigrants-to-native channels). As we omitted classrooms with none or only a single native or immigrant student as in that case identification of social interactions is not possible, the sample size for the subject of science decreases from $1,560$ to $1,201$ observations for the "Hauptschule", from $3,323$ to $2,026$ in the "Gymnasium" and from $2,765$ to $1,932$ in the "Realschule". The most striking reductions occurs for the "Gymnasium" schools: From 131 classrooms we use 80.

Results in Table 5.2 follow the pattern of classroom average social interactions presented in Table 5.1. For the "Hauptschule", effects stemming from native or immigrant students on the whole classroom are of similar magnitude. But in the higher-ability schools, namely the "Realschule" and the "Gymnasium", the natives' effect on the whole classroom are larger than the immigrants' effects on all classmates regardless of their immigration background. Overall, Table 5.2 shows that a native-student's performance is more important than an immigrant-student's performance for the whole classroom. Social multipliers are statistically significant for all cases with the largest impact found in the column of science and the "Hauptschule" native students: a 1-point increase in the average performance of natives increases classroom performance by about 6 points.

Although in Table 5.2 we can discern the origin of the performance effect (native or immigrant), we cannot distinguish between the targeted students—rather only a homogeneous classroom effect. Therefore, in Tables 5.3 and 5.4 we estimate same and cross-immigration background social interactions. Table 5.3 reveals that effects from native students to other native students in the classroom are always higher than respective effects from immigrant students to other immigrant students in the classroom: The native-to-native dependence in performance is higher than the immigrant-to-immigrant performance. The largest deviation is found for the "Gymnasium" and the subjects of science and math with a difference of about 0.4 points (compare 0.557 with 0.166 and 0.573 with 0.161). Although the native-to-native dependence is either moderate (meaning around 0.5) or high (meaning well above 0.5), the immigrant-to-immigrant dependence is rather moderate for the "Hauptschule" and low for the "Gymnasium" with the "Realschule" dependence in the middle. Low dependence implies social multipliers not very different than 1 point. Again, all specifications yield multiplier effects.

Finally, in Table 5.4 we present the cross-immigration-background effects to shed light on if and how the performance of native (immigrants) students affects the performance of immigrant (native) students in the classroom. The native-to-immigrant dependence is very high and similar between the "Hauptschule" and the "Gymnasium"; for the "Realschule" the parameter estimate has either a similar or lower magnitude. For the first time we encounter a lack of a multiplier effect for the "Gymnasium" schools. Thus, a 1-point increase in natives' average performance raises immigrants' performance only by 0.84 in the case of science instead of 6.4 points implied by the social multiplier. We conclude that the performance of native students on their immigrant peers is much more important in the lower-ability schools. Furthermore, in two out of three cases with absence of a multiplier effect, the estimated endogenous effects parameter, $\widehat{\lambda}$, is very close to 1 for math and German language resembling the near-unit-root case of Lee and Yu (2013). The latter holds for the immigrant-to-native channel in the "Hauptschule" and the subject of science but the social multiplier is present otherwise. The average effects from immigrants on natives are always larger in the "Hauptschule" than in the "Gymnasium" but the social multiplier effect is present only in math and German language for the "Hauptschule". Therefore, it is only for science that the immigrant-to-native effect is more important in the higher-ability schools (compare the social multiplier effect of 2.04 in the "Gymnasium" with the endogenous effects parameter 0.97 in the "Hauptshule"). Overall, although for the low and middle-ability schools the native-to-immigrant effects are more important than the immigrant-to-native effects, for the high-ability schools it is the immigrant-to-native effect that matters the most.

By comparing magnitudes in Tables 5.3 and 5.4 we observe that, first, for the low and middle-ability schools social interactions among classmates with and without immigration background are more intense than social interactions among students with the same immigration background; more specifically, the highest effect stems from natives' to immigrants' performance; second, for the high-ability schools the highest dependence in performance stems from the native-to-native channel. Therefore, as in Table 5.2: *The performance of native peers matters more* in terms of implied impacts.

In all three Tables 5.2-5.4, we see that social interactions emerge regardless of how we specify them merely because they exist. But comparison of the log-likelihood values reveals that the socio-matrix specification (see Subsection 4.1) that fits the data best is the average effects of native students on the whole classroom (Table 5.2). We conclude that *the average performance of native students is the predominant social effect for all school types and subjects*. From a policy perspective, the latter implies we should be more concerned about where to place a native instead of an immigration-background student. The issue is more relevant now than ever considering the influx of refugees—mainly from the war in Syria. Our results point to the same direction as the recent findings of Figlio and Őzek (2017) who show that refugees have zero effects on incumbent students.

Table 5.2: Social Interactions Models According to School Type: Effects from Natives and Immigrants on the Whole Classroom

	Science			Math			German Language		
	HS	GY	RS	HS	GY	RS	HS	GY	RS
Average Effects from Natives									
Endogenous Effects, $\widehat{\lambda}$	0.832***	0.587***	0.725***	0.792***	0.602***	0.607***	0.794***	0.610***	0.549***
	(0.013)	(0.024)	(0.017)	(0.014)	(0.022)	(0.023)	(0.013)	(0.022)	(0.033)
Social Multiplier, $\frac{1}{1-\widehat{\lambda}}$	5.952***	2.421***	3.636***	4.808***	2.513***	2.544***	4.854***	2.564***	2.217***
	(0.461)	(0.141)	(0.225)	(0.324)	(0.139)	(0.149)	(0.306)	(0.145)	(0.162)
Log-likelihood	-7,514	-12,146	-11,685	-9,716	-12,751	-12,131	-9,481	-12,374	-11,718
Number of Observations	1,201	2,026	1,932	1,510	2,026	1,932	1,506	2,025	1,932
Average Effects from Immigrants									
Endogenous Effects, $\widehat{\lambda}$	0.828***	0.392***	0.616***	0.759***	0.365***	0.397***	0.767***	0.416***	0.409***
	(0.012)	(0.027)	(0.021)	(0.016)	(0.027)	(0.028)	(0.014)	(0.027)	(0.031)
Social Multiplier, $\frac{1}{1-\widehat{\lambda}}$	5.814***	1.645***	2.604***	4.149***	1.575***	1.658***	4.292***	1.712***	1.692***
	(0.406)	(0.073)	(0.142)	(0.275)	(0.067)	(0.077)	(0.258)	(0.079)	(0.089)
Log-likelihood	-7,586	-12,202	-11,770	-9,774	-12,794	-12,196	-9,525	-12,419	-11,750
Number of Observations	1,201	2,026	1,932	1,510	2,026	1,932	1,506	2,025	1,932

Note: Estimation with QML including individual fixed effects (CML) using student final weights and five PISA Plausible Values. Average social interactions in the classroom. Standard errors in parentheses. *** denotes significance at 1%. HS denotes "Hauptschule", GY "Gymnasium" and RS "Realschule".

Table 5.3: Social Interactions Models According to School Type: Effects from Natives on Natives and Immigrants on Immigrants

	Science			Math			German Language		
	HS	GY	RS	HS	GY	RS	HS	GY	RS
Average Effects from Natives on Natives									
Endogenous Effects, $\widehat{\lambda}$	0.767***	0.557***	0.688***	0.733***	0.573***	0.574***	0.729***	0.564***	0.517***
	(0.013)	(0.025)	(0.018)	(0.015)	(0.024)	(0.025)	(0.014)	(0.024)	(0.033)
Social Multiplier, $\frac{1}{1-\widehat{\lambda}}$	4.292***	2.257***	3.205***	3.745***	2.342***	2.347***	3.690***	2.294***	2.070***
	(0.239)	(0.127)	(0.185)	(0.210)	(0.132)	(0.138)	(0.191)	(0.126)	(0.141)
Log-likelihood	-7,791	-12,178	-11,775	-9,936	-12,773	-12,180	-9,677	-12,409	-11,752
Number of Observations	1,201	2,026	1,932	1,510	2,026	1,932	1,506	2,025	1,932
Average Effects from Immigrants on Immigrants									
Endogenous Effects, $\widehat{\lambda}$	0.640***	0.166***	0.451***	0.575***	0.161***	0.319***	0.599***	0.211***	0.263***
	(0.020)	(0.039)	(0.033)	(0.023)	(0.043)	(0.040)	(0.021)	(0.045)	(0.047)
Social Multiplier, $\frac{1}{1-\widehat{\lambda}}$	2.778***	1.199***	1.821***	2.353***	1.192***	1.468***	2.494***	1.267***	1.357***
	(0.154)	(0.056)	(0.109)	(0.127)	(0.061)	(0.086)	(0.131)	(0.072)	(0.086)
Log-likelihood	-7,957	-12,287	-11,934	-10,124	-12,845	-12,266	-9,815	-12,496	-11,807
Number of Observations	1,201	2,026	1,932	1,510	2,026	1,932	1,506	2,025	1,932

Note: Estimation with QML including individual fixed effects (CML) using student final weights and five PISA Plausible Values. Average social interactions in the classroom. Standard errors in parentheses. *** denotes significance at 1%. HS denotes “Hauptschule”, GY “Gymnasium” and RS “Realschule”.

Table 5.4: Social Interactions Models According to School Type: Effects from Natives on Immigrants and Immigrants on Natives.

	Science			Math			German Language		
	HS	GY	RS	HS	GY	RS	HS	GY	RS
Average Effects from Natives on Immigrants									
Endogenous Effects, $\widehat{\lambda}$	0.884***	0.844***	0.853***	0.904***	0.944***	0.740***	0.903***	0.966***	0.638***
	(0.038)	(0.107)	(0.062)	(0.046)	(0.108)	(0.079)	(0.038)	(0.086)	(0.098)
Social Multiplier, $\frac{1}{1-\widehat{\lambda}}$	8.621***	6.410*	6.803***	10.417**	17.857	3.846***	10.309***	29.412	2.762***
	(2.824)	(4.397)	(2.869)	(4.991)	(34.439)	(1.169)	(4.039)	(74.394)	(0.748)
Log-likelihood	-7,910	-12,262	-11,898	-10,074	-12,824	-12,240	-9,767	-12,464	-11,782
Number of Observations	1,201	2,026	1,932	1,510	2,026	1,932	1,506	2,025	1,932
Average Effects from Immigrants on Natives									
Endogenous Effects, $\widehat{\lambda}$	0.973***	0.510***	0.730***	0.898***	0.448***	0.439***	0.893***	0.524***	0.500***
	(0.027)	(0.039)	(0.034)	(0.031)	(0.035)	(0.038)	(0.029)	(0.037)	(0.045)
Social Multiplier, $\frac{1}{1-\widehat{\lambda}}$	37.037	2.041***	3.704***	9.804***	1.816***	1.782***	9.346***	2.101***	2.000***
	(37.037)	(0.162)	(0.466)	(2.980)	(0.115)	(0.121)	(2.533)	(0.163)	(0.180)
Log-likelihood	-7,766	-12,196	-11,806	-9,906	-12,791	-12,215	-9,648	-12,415	-11,751
Number of Observations	1,201	2,026	1,932	1,510	2,026	1,932	1,506	2,025	1,932

Note: Estimation with QML including individual fixed effects (CML) using student final weights and five PISA Plausible Values. Average social interactions in the classroom. Standard errors in parentheses. *, ** and *** denote significance at 10%, 5% and 1% respectively. HS denotes "Hauptschule", GY "Gymnasium" and RS "Realschule".

6 Conclusion

This paper estimates spatial autoregressive type models for school types encountered in the German secondary educational system, meaning the lower-performing “Hauptschule” that leads at most to vocational training, the middle-performing “Realschule”, and the top-performing “Gymnasium” whose graduates predominantly attend the university later on. The empirical findings uncover that there is room for more inequality in the “Hauptschule” type rather than the “Realschule” or the “Gymnasium” schools, as endogenous social interactions are more intense in magnitude, revealing greater dependence among lower-achieving students. The higher magnitude translates into higher social multiplier effects, so that already preexisting inequalities due to early tracking tend to amplify to a larger degree. The implication is important as summary statistics reveal that around 31% of the “Hauptschule” students report having some immigration background (the latter defined in a broad sense). Moreover, our results suggest that students are more impervious to immigrant than native classmates’ performance. Our empirical exercise challenges the practice of early tracking in education as it compromises equity.

References

Aaberge, R., A. Bjőrklund, M. Jäntti, M. Palme, P. J. Pedersen, N. Smith, and T. Wennemo (2002): "Income Inequality and Income Mobility in the Scandinavian Countries Compared to the United States," *Review of Income and Wealth*, 48(4), 443–469.

Acemoglu, D. (2003): "Cross-country Inequality Trends," *The Economic Journal*, 113(485), F121–F149.

Badinger, H., and P. Egger (2011): "Estimation of Higher-order Spatial Autoregressive Cross-Section Models with Heteroscedastic Disturbances," *Papers in Regional Science*, 90(1), 213–235.

Bao, Y., and A. Ullah (2007): "Finite Sample Properties of Maximum Likelihood Estimator in Spatial Models," *Journal of Econometrics*, 137(2), 396–413.

Bauer, P., and R. Riphahn (2007): "Heterogeneity in the Intergenerational Transmission of Educational Attainment: Evidence from Switzerland on Natives and Second-generation Immigrants," *Journal of Population Economics*, 20(1), 121–148.

Bayer, P., F. Ferreira, and R. McMillan (2004): "Tiebout Sorting, Social Multipliers and the Demand for School Quality," Working Paper 10871, National Bureau of Economic Research.

BfA (2009): "Bundesagentur fuer Arbeit: Report on the Labour Market: the Labour Market in Germany - a Year of Crisis for the German Labour Market," Nuremberg 2009.

Bjőrklund, A., and M. Jäntti (1997): "Intergenerational Income Mobility in Sweden Compared to the United States," *The American Economic Review*, 87(5), 1009–1018.

Boucher, V., and B. Fortin (2016): *Some Challenges in the Empirics of the Effects of Networks*chap. 12, The Oxford Handbook of the Economics of Networks. Oxford University Press.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009): "Identification of Peer Effects Through Social Networks," *Journal of Econometrics*, 150(1), 41–55.

Bramoullé, Y., R. Kranton, and M. D'Amours (2014): "Strategic Interaction and Networks," *American Economic Review*, 104(3), 898–930.

Brunello, G., and L. Rocco (2013): "The Effect of Immigration on the School Performance of Natives: Cross Country Evidence Using PISA Test Scores," *Economics of Education Review*, 32(0), 234–246.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009): "Peer Effects and Social Networks in Education," *Review of Economic Studies*, 76(4), 1239–1267.

Card, D., J. Heining, and P. Kline (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality," *The Quarterly Journal of Economics*.

Carrell, S. E., B. I. Sacerdote, and J. E. West (2013): "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation," *Econometrica*, 81(3), 855–882.

Cattaneo, M. A., and S. C. Wolter (2012): "Migration Policy Can Boost PISA Results: Findings from a Natural Experiment," IZA Discussion Papers 6300, Institute for the Study of Labor (IZA).

Davezies, L., X. D'Haultfoeuille, and D. Fougére (2009): "Identification of Peer Effects Using Group Size Variation," *Econometrics Journal*, 12(3), 397–413.

Dronkers, J., and J. Skopek (2015): "Performance in Secondary School in German States - A Longitudinal Three-Level Approach," mimeo, European University Institute and Maastricht University, ROA.

Dustmann, C. (2004): "Parental Background, Secondary School Track Choice, and Wages," *Oxford Economic Papers*, 56(2), 209–230.

Dustmann, C., J. Ludsteck, and U. Schoenberg (2009): "Revisiting the German Wage Structure," *The Quarterly Journal of Economics*, 124(2), 843–881.

Dustmann, C., P. A. Puhani, and U. Schoenberg (2012): "The Long-term Effects of School Quality on Labor Market Outcomes and Educational Attainment," CReAM discussion paper 1208, Centre for Research and Analysis of Migration, Department of Economics, University College London.

Entorf, H., and M. Lauk (2008): "Peer Effects, Social Multipliers and Migrants at School: An International Comparison," *Journal of Ethnic and Migration Studies*, 34(4), 633–654.

Entorf, H., and N. Minoiu (2005): "What a Difference Immigration Policy Makes: A Comparison of PISA Scores in Europe and Traditional Countries of Immigration," *German Economic Review*, 6(3), 355–376.

Figlio, D. N., and U. Őzek (2017): "Unwelcome Guests? The Effects of Refugees on the Educational Outcomes of Incumbent Students," Working Paper 23661, National Bureau of Economic Research.

Frick, J. R., and G. G. Wagner (2001): "Economic and Social Perspectives of Immigrant Children in Germany," IZA Discussion Papers 301, Institute for the Study of Labor (IZA).

Fruehwirth, J. C. (2013): "Identifying Peer Achievement Spillovers: Implications for Desegregation and the Achievement Gap," *Quantitative Economics*, 4(1), 85–124.

Galbiati, R., and G. Zanella (2012): "The Tax Evasion Social Multiplier: Evidence from Italy," *Journal of Public Economics*, 96(5), 485 – 494.

Gang, I. N., and K. F. Zimmermann (2000): "Is Child like Parent? Educational Attainment and Ethnic Origin," *Journal of Human Resources*, 35(3), 550–569.

Glaeser, E. L., B. Sacerdote, and J. A. Scheinkman (1996): "Crime and Social Interactions," *The Quarterly Journal of Economics*, 111(2), 507–548.

Glaeser, E. L., B. I. Sacerdote, and J. A. Scheinkman (2003): "The Social Multiplier," *Journal of the European Economic Association*, 1(2-3), 345–353.

Goldsmith-Pinkham, P., and G. W. Imbens (2013): "Social Networks and the Identification of Peer Effects," *Journal of Business and Economic Statistics*, 31(3), 253–264.

Gruber, K. H. (2006): "The German PISA-shock: Some Aspects of the Extraordinary Impact of the OECD's PISA Study on the German Education System," pp. 195–208. Oxford, UK: Symposium Books.

Hanushek, E. A., and L. Woessmann (2006): "Does Educational Tracking Affect Performance and Inequality? Differences- in-Differences Evidence Across Countries," *Economic Journal*, 116(510), C63–C76.

Horrace, W. C., X. Liu, and E. Patacchini (2016): "Endogenous Network Production Functions with Selectivity," *Journal of Econometrics*, 190(2), 222–232, Endogeneity Problems in Econometrics.

Hoxby, C. M., and G. Weingarth (2005): "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects," Mimeo, Harvard University.

Hsieh, C.-S., and L. F. Lee (2014): "A Social Interactions Model with Endogenous Friendship Formation AND Selectivity," *Journal of Applied Econometrics*, pp. n/a–n/a.

Jensen, P., and A. W. 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.

Jerrim, J., L. A. Lopez-Agudo, O. D. Marcenaro-Gutierrez, and N. Shure (2017): "What Happens When Econometrics and Psychometrics Collide? An Example Using the PISA Data," IZA Discussion Papers 10847, Institute for the Study of Labor (IZA).

Johnsson, I., and H. R. Moon (2017): "Estimation Of Peer Effects In Endogenous Social Networks: Control Function Approach," Working Paper.

Kahn, L. M. (2000): "Wage Inequality, Collective Bargaining, and Relative Employment from 1985 to 1994: Evidence from Fifteen OECD Countries," *The Review of Economics and Statistics*, 82(4), 564–579.

Kelejian, H. H. (2008): "A Spatial J-Test for Model Specification Against a Single or a Set of Nonnested Alternatives," .

Kelejian, H. H., and G. Piras (2011): "An Extension of Kelejian's J-test for Non-Nested Spatial Models," *Regional Science and Urban Economics*, 41(3), 281–292.

Kelejian, H. H., and I. R. Prucha (1998): "A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances," *The Journal of Real Estate Finance and Economics*, 17(1), 99–121.

——— (2010): "Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances," *Journal of Econometrics*, 157(1), 53–67.

Kelejian, H. H., I. R. Prucha, and Y. Yuzefovich (2006): "Estimation Problems In Models With Spatial Weighting Matrices Which Have Blocks Of Equal Elements," *Journal of Regional Science*, 46(3), 507–515.

Krause, A., U. Rinne, and S. Schueller (2012): "Kick It Like Oezil? Decomposing the Native-Migrant Education Gap," IZA Discussion Papers 6696, Institute for the Study of Labor (IZA).

Lavy, V. (2015): "Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries," *The Economic Journal*, 125(588), F397–F424.

Lee, L.-F. (2003): "Best Spatial Two-Stage Least Squares Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances," *Econometric Reviews*, 22(4), 307–335.

——— (2004): "Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models," *Econometrica*, 72(6), 1899–1925.

——— (2007): "Identification and Estimation of Econometric Models with Group Interactions, Contextual Factors and Fixed Effects," *Journal of Econometrics*, 140(2), 333 – 374.

Lee, L.-F., and X. Liu (2010): "Efficient Gmm Estimation of High Order Spatial Autoregressive Models with Autoregressive Disturbances," *Econometric Theory*, 26(01), 187–230.

Lee, L.-F., X. Liu, and X. Lin (2010): "Specification and Estimation of Social Interaction Models with Network Structures," *Econometrics Journal*, 13(2), 145–176.

Lee, L.-F., and J. Yu (2010): "Estimation of Spatial Autoregressive Panel Data Models with Fixed Effects," *Journal of Econometrics*, 154(2), 165–185.

——— (2013): "Near Unit Root in the Spatial Autoregressive Model," *Spatial Economic Analysis*, 8(3), 314–351.

——— (2014): "Efficient GMM Estimation of Spatial Dynamic Panel Data Models with Fixed Effects," *Journal of Econometrics*, 180(2), 174–197.

Lee, S.-Y. (2008): "Three Essays on Spatial Econometrics and Empirical Industrial Organization," Ph.D. thesis, Department of Economics, Ohio State University.

Liu, X., E. Patacchini, and Y. Zenou (2014): "Endogenous Peer Effects: Local Aggregate or Local Average?," *Journal of Economic Behavior and Organization*, 103(0), 39–59.

Maurin, E., and J. Moschion (2009): "The Social Multiplier and Labor Market Participation of Mothers," *American Economic Journal: Applied Economics*, 1(1), 251–72.

Murat, M. (2012): "Do Immigrant Students Succeed? Evidence from Italy and France," *Global Economy Journal*, 12(3), 1–22.

Murat, M., D. Ferrari, P. Frederic, and G. Pirani (2010): "Immigrants, Schooling and Background. Cross-Country Evidence from PISA 2006," Center for Economic Research (RECent) 054, University of Modena and Reggio E., Dept. of Economics.

OECD (2009): "PISA 2006 Technical Report," OECD Publishing.

Open-Society-Justice-Initiative (2012): "Submission to the UN Human Rights Committee: Review of Germany," https://www.opensocietyfoundations.org/sites/default/files/hrc-germany-10172012.pdf.

Ord, K. (1975): "Estimation Methods for Models of Spatial Interaction," *Journal of the American Statistical Association*, 70(349), 120–126.

Qu, X., and L. fei Lee (2015): "Estimating a Spatial Autoregressive Model with an Endogenous Spatial Weight Matrix," *Journal of Econometrics*, 184(2), 209–232.

Qu, X., L. fei Lee, and J. Yu (2017): "QML Estimation of Spatial Dynamic Panel Data Models with Endogenous Time Varying Spatial Weights Matrices," *Journal of Econometrics*, 197(2), 173–201.

Ramsden, S., F. M. Richardson, G. Josse, M. S. C. Thomas, C. S. Caroline Ellis, M. L. Seghier, and C. J. Price (2011): "Verbal and Non-verbal Intelligence Changes in the Teenage Brain," *Nature*, (479), 113â116.

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

Shi, W., and L. fei Lee (2017): "Spatial Dynamic Panel Data Models with Interactive Fixed Effects," *Journal of Econometrics*, 197(2), 323–347.

Solon, G., S. J. Haider, and J. M. Wooldridge (2015): "What Are We Weighting For?," *Journal of Human Resources*, 50(2), 301–316.

Sommer, M. (2016): "A Non-Affine Social Interactions Model for Conformist and Anticonformist Behavior in Social Networks," Mimeo, Swedish Institute for Social Research (SOFI), Stockholm University.

Vardardottir, A. (2015): "The Impact of Classroom Peers in a Streaming System," *Economics of Education Review*, 49(Supplement C), 110–128.